

# **Investigations into Intracellular Thiols of Biological Importance**

by

Christine Elizabeth Hand

A thesis  
presented to the University of Waterloo  
in fulfillment of the  
thesis requirement for the degree of  
Doctor of Philosophy  
in  
Chemistry

Waterloo, Ontario, Canada, 2007

© Christine Elizabeth Hand 2007

## AUTHOR'S DECLARATION

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

## Abstract

The presence of thiols in living systems is critical for the maintenance of cellular redox homeostasis, the maintenance of protein thiol-disulfide ratios and the protection of cells from reactive oxygen species. In addition to the well studied tripeptide glutathione ( $\gamma$ -Glu-Cys-Gly), a number of compounds have been identified that contribute to these essential cellular roles. Many of these molecules are of great clinical interest due to their essential role in the biochemistry of a number of deadly pathogens, as well as their possible role as therapeutic agents in the treatment of a number of diseases. A series of studies were undertaken using theoretical, chemical and biochemical approaches on a selection of thiols, ergothioneine, the ovolthiols and mycothiol, to further our understanding of these necessary biological components.

Ergothioneine is present at significant physiological levels in humans and other mammals; however, a definitive role for this thiol has yet to be determined. It has been implicated in radical scavenging *in vivo* and shows promise as a therapeutic agent against disease states caused by oxidative damage. Given the clinical importance of this intracellular thiol, further investigation into the behaviour of ergothioneine appeared warranted. A high level theoretical study was performed to determine the thermodynamic driving force behind the instability of the ergothioneine disulfide, as well as the thermodynamics of the reactions of ergothioneine with a selection of biologically relevant reactive oxygen species. These results were compared to those determined for a glutathione model compound, as well as the related ovolthiols. The latter are believed to act as hydrogen peroxide scavengers *in vivo* and are currently under review as possible therapeutics against oxidative damage. The structural differences between the ovolthiols and ergothioneine dramatically affect their reactivity and this study investigates the thermodynamic driving forces behind these differences.

Mycothiol is the major thiol found in the *Actinomycetales* bacteria, which include the causative agent of tuberculosis, and the enzymes which use mycothiol have been identified as important targets for the development of novel antimicrobials. To better understand the *in vivo* behaviour of mycothiol, a thorough conformational search was performed to determine what, if any, trends exist among the low energy conformers expected to be present in solution. Knowledge of the conformations preferred by mycothiol may aid in the design of substrate-based inhibitors targeted at mycothiol-dependent enzymes. In addition, the efforts towards the identification of a mycothiol-dependent glyoxalase system are described. The glyoxalase system is essential for the detoxification of methylglyoxal, a toxic by-product of glycolysis, and this system would serve as a target for the design of new therapeutics against tuberculosis and other pathogenic *Actinomycetales* bacteria.

In addition to the study of intracellular thiols, this work details a preliminary theoretical study of the thermodynamics of the phosphorylation of proteinaceous serine residues by inositol pyrophosphates in eukaryotic cell-free extracts. It has been postulated that this observed activity may represent a novel signalling pathway in eukaryotes. This study focused on the effect of inositol pyrophosphate structure and overall charge on the thermodynamics of these reactions. This information should contribute to our understanding of this novel cellular phosphorylation process.

## Acknowledgements

I am indebted to Dr. John F. Honek for his friendship, guidance and encouragement during the course of this work. This project presented a wide range of challenges and John's constant patience, faith and enthusiasm was always appreciated.

I would like to thank the members of the Honek lab who have been an excellent source of knowledge, laughter and support: Pei Chun Hang, Danish Khan, Meijun Lu, Kadia Mullings, Jennifer Steere, Nicole Sukdeo, Dr. Zhengding Su, Uthaiwan Suttisansanee, Dr. Mark Vaughan, and Paula Walasek. Dr. Elisabeth Daub was delightfully supportive and assisted in all aspects of the microbiology detailed herein.

Many thanks to the members of my committee for their valuable advice over the course of my doctorate: Drs. France-Isabelle Auzanneau, Gary Dmitrienko, and Michael Palmer. In addition, I would like to thank my external examiners, Drs. Barbara Moffatt and Jeffrey Atkinson for their participation in this process.

The technical assistance of Jan Venne, Robyn Landers, Val Goodfellow and Dr. Richard Smith was greatly appreciated. Dr. Nicholas J. Taylor performed the crystal structure determination of ergothioneine and was very helpful in the preparation of this manuscript. Sadly, Nick passed away before the completion of my degree and I am sure he will be missed by the department.

The arduous task of proof-reading this work fell to Emelyne Das, and Drs. Darren Anderson and Jennifer Belelie. I am very grateful for their quick reading, suggestions and support. Luna Guha's assistance was also appreciated.

I have been blessed in having two fantastic mentors: Drs. Janet Kumita and Jennifer Belelie. They have been patient sources of advice, encouragement and unfailing support for which I will always be grateful.

I am lucky to have been befriended by many wonderful people during my stay at UW, including Dr. Bryan Hill, Joe and Karen Meissner, Don and Jill Spratt, Jason Yaeck and Andrea Dupont, as well as Kristen Day and Scott Ruttgaizer, who are my Waterloo family.

Finally, I must thank Mom, Dad, Brandon, Gramma, Sean, and Heather, as well as Uncle Brian and his family, for all of their advice, encouragement and assistance. Bran.: Thank you for taking such good care of me; your calm perspective and unfailing support was a god-send. Mom and Dad: I can not even begin to express my gratitude; your thoughtfulness, generosity and love overwhelms me.

*For my parents, Bob and Linda Hand*

*“A happy family is but an earlier heaven”*

*~John Bowring*

# Table of Contents

<b>Abstract</b> .....	iii
<b>Acknowledgements</b> .....	iv
<b>Dedication</b> .....	v
<b>Table of Contents</b> .....	vi
<b>List of Tables</b> .....	xi
<b>List of Figures</b> .....	xv
<b>Table of Abbreviations</b> .....	xviii
<b>Chapter 1: Introduction to Intracellular Thiols</b> .....	1
<b>1.1. Ergothioneine</b> .....	2
1.1.1. Distribution.....	2
1.1.2. Biosynthesis.....	3
1.1.3. Chemical Properties.....	3
1.1.4. Biological Function.....	4
<b>1.2. Ovothiols</b> .....	5
1.2.1. Distribution.....	5
1.2.2. Biosynthesis.....	6
1.2.3. Chemical Properties.....	7
1.2.4. Biological Function.....	7
<b>1.3. Glutathione</b> .....	9
1.3.1. Distribution.....	9
1.3.2. Biosynthesis.....	10
1.3.3. Chemical Properties.....	10
1.3.4. Biological Function.....	10
<b>1.4. Mycothiol</b> .....	18
1.4.1. Distribution.....	19
1.4.2. Biosynthesis.....	19
1.4.3. Chemical Properties.....	21
1.4.4. Biological Function.....	22
<b>1.5. Objectives</b> .....	25
<b>1.6. Other Interesting myo-Inositol Derivatives</b> .....	25
<b>1.7. Publications</b> .....	25

<b>1.8. References</b> .....	26
<b>Chapter 2: Modelling of Ergothioneine and Ovothiols</b> .....	36
<b>2.1. Computational Chemistry</b> .....	36
2.1.1. Hartree-Fock Methods .....	37
2.1.2. Basis Functions and Sets.....	38
2.1.3. Density Functional Theory .....	39
2.1.4. Energy Minimization or Geometry Optimization.....	39
2.1.5. Gas Phase versus Solvation Modelling.....	41
2.1.6. Zero Point Energy Correction.....	42
2.1.7. Vibrational Frequencies.....	43
2.1.8. Plan of Action .....	43
<b>2.2. Materials and Methods</b> .....	43
2.2.1. Crystal Structure of Ergothioneine .....	44
2.2.2. Determination of Appropriate Basis Set.....	44
2.2.3. Detailed Electronic Structural Calculations.....	45
2.2.4. Free Energy of Reaction Calculations .....	46
<b>2.3. Results and Discussion</b> .....	46
2.3.1. The Use of a Crystal Structure as a Starting Point for Optimization and a Potential Benchmark for Accuracy.....	47
2.3.2. Crystal Structure Determination of Ergothioneine .....	47
2.3.3. Determination of Appropriate Basis Set.....	50
2.3.4. Use of Truncated Analogues.....	53
2.3.5. Modelling 2-Thiol-4-Methyl-Imidazole and its Disulfide.....	53
2.3.6. Modelling 4-Thiol-N <sup>1</sup> -Methyl-5-Methyl-Imidazole and the Stability of its Disulfide.....	57
2.3.7. Biologically Relevant Reactions.....	58
<b>2.4. Conclusions and Future Work</b> .....	63
<b>2.5. References</b> .....	63
<b>Chapter 3: Conformational Analyses of Mycothiol</b> .....	67
<b>3.1. Computational Chemistry</b> .....	67
3.1.1. Monte Carlo Multiple Minimization Methods.....	70
3.1.2. Molecular Mechanics Methods and Force Fields .....	71
3.1.3. Plan of Action .....	71
<b>3.2. Materials and Methods</b> .....	71
3.2.1. Geometry Optimizations of Mycothiol.....	71
3.2.2. Survey of Force Fields Available .....	71
3.2.3. Monte Carlo Multiple Minimization Searches of the Mycothiol Conformational Space.....	71

3.2.4. Resultant Conformer Analysis using XCluster.....	72
3.2.5. MOE2004 Calculations.....	73
<b>3.3. Results and Discussion</b> .....	<b>73</b>
3.3.1. Choice of Conformational Search Method.....	74
3.3.2. Force Field Selection.....	75
3.3.3. Ring Opening during the Conformational Search.....	76
3.3.4. Gas Phase versus Solvation Phase Modelling.....	76
3.3.5. Energy Differences Relevant to Biological Systems.....	77
3.3.6. Analysis of the OPLS-AA Data using Clustering Techniques.....	78
3.3.7. AMBER* Conformational Search and Comparison of Results to those found using OPLS-AA.....	80
3.3.8. Comparison of the Monte Carlo Multiple Minimization and MOE Stochastic Search Results.....	81
3.3.9. Comparison to the Structure of Bimane Derivative of Mycothiol, Obtained Using NMR Methods.....	82
3.3.10. Applications to Drug Design.....	84
<b>3.4. Conclusions and Future Work</b> .....	<b>85</b>
<b>3.5. References</b> .....	<b>85</b>
<b>Chapter 4: Efforts Towards the Identification of a Mycothiol-Utilizing Glyoxalase System</b> .....	<b>89</b>
<b>4.1. Introduction</b> .....	<b>90</b>
4.1.1. Parallel Functions of Glutathione and Mycothiol Utilizing Enzymes.....	90
4.1.2. A Brief Review of the Glutathione-Dependent Glyoxalase System.....	91
4.1.3. The Trypanothione-Dependent Glyoxalase System.....	92
4.1.4. Methods Used for the Detection of the Glyoxalase System.....	93
4.1.5. Introduction to <i>Streptomyces</i> Bacteria.....	93
4.1.6. The Synthesis of Mycothiol.....	95
4.1.7. The Use of Truncated Mycothiol as a Probe for Enzymatic Activity.....	99
4.1.8. Isolation of Mycothiol from Cell-Free Extracts.....	100
4.1.9. Plan of Action.....	101
<b>4.2. Materials and Methods</b> .....	<b>101</b>
4.2.1. Data Mining of the <i>Streptomyces coelicolor</i> Genome.....	103
4.2.2. Media and Buffers Used.....	103
4.2.3. Growth of <i>Streptomyces coelicolor</i> .....	105
4.2.4. Initial Assay for Glyoxalase Activity in <i>Streptomyces coelicolor</i> .....	105
4.2.5. Control Assays for Glyoxalase Activity in <i>Escherichia coli</i> .....	106
4.2.6. Synthesis of Des- <i>myo</i> -Inositol Mycothiol.....	106
4.2.7. Alternative Synthesis of Des- <i>myo</i> -Inositol.....	109
4.2.8. Synthesis of 2-S-(2'-thiopyridyl)-6-hydroxynaphthylsulfide.....	110
4.2.9. Growth of <i>Streptomyces jumonjinensis</i> , the Chosen Source of Mycothiol.....	111
4.2.10. Isolation of Mycothiol from <i>Streptomyces jumonjinensis</i> .....	111



4.2.11. Synthesis of a Mycothiol Analogue.....	112
<b>4.3. Results and Discussion .....</b>	<b>113</b>
4.3.1. Sequence Searching of the <i>Streptomyces coelicolor</i> Genome.....	113
4.3.2. Choice of Organism.....	127
4.3.3. Preliminary Assays for the Identification of a Mycothiol Dependent Glyoxalase System.....	127
4.3.4. Comparative Assays Performed in <i>Escherichia coli</i> .....	128
4.3.5. Synthesis of Des- <i>myo</i> -Inositol-Mycothiol .....	129
4.3.6. Alternative Synthesis of Des- <i>myo</i> -Inositol Mycothiol .....	129
4.3.7. Synthesis of 2- <i>S</i> -(2'-Thiopyridyl)-6-hydroxynaphthyl disulfide and Isolation of Mycothiol.....	130
4.3.8. Synthesis of Mycothiol Analogue.....	131
<b>4.4. Conclusions and Future Work .....</b>	<b>131</b>
<b>4.5. References .....</b>	<b>132</b>
<b>Chapter 5: Phosphorylation by Inositol Pyrophosphates .....</b>	<b>138</b>
<b>5.1. Introduction to Inositol Phosphates and Pyrophosphates .....</b>	<b>138</b>
5.1.1. Numbering .....	138
5.1.2. Distribution and Structure.....	139
5.1.3. Biosynthesis.....	140
5.1.4. Biological Role .....	141
<b>5.2. Objectives.....</b>	<b>141</b>
<b>5.3. Computational Studies of Phosphate Transfer.....</b>	<b>141</b>
5.3.1. The Modelling of Pyrophosphate.....	141
5.3.2. Modelling of the Inositol Phosphates .....	142
<b>5.4. Computational Chemistry.....</b>	<b>142</b>
5.4.1. Semi-Empirical Theory.....	142
5.4.2. Plan of Action .....	143
<b>5.5. Materials and Methods .....</b>	<b>143</b>
5.5.1. Initial Hartree-Fock, Density Functional Theory and Molecular Mechanics Geometry Optimizations.....	143
5.5.2. Semi-Empirical Geometry Optimizations.....	144
5.5.3. Geometry Optimizations of Model Phosphate Donors.....	144
5.5.4. Geometry Optimizations of the Inositol Pyrophosphates .....	145
5.5.5. Dihedral Angle Driving .....	148
5.5.6. Geometry Optimizations of Other Important Compounds .....	149
5.5.7. Calculation of the Thermodynamics of Phosphate Group Transfer .....	149
<b>5.6. Results and Discussion.....</b>	<b>149</b>
5.6.1. Selection of Modelling Technique.....	149

5.6.2. Exclusion of Counter-Ions .....	151
5.6.3. Use of Methanol as a Model for Serine .....	151
5.6.4. Geometry Optimizations of Inositol Pyrophosphates using Semi-Empirical Techniques .....	152
5.6.5. Phosphorylation of Methanol by Model Compounds .....	152
5.6.6. Phosphorylation of Methanol by the Inositol Pyrophosphates .....	153
5.6.7. Phosphorylation of Methanol by Adenosine Triphosphate and Other Model Compounds.....	157
<b>5.7. Conclusions and Future Work</b> .....	157
<b>5.8. References:</b> .....	158
<b>Appendix 1: Modelling of Ergothioneine and Ovothiol Supplementary Information</b> .....	161
<b>A1.1: Studies of Ergothioneine and 2-Thiol-4-Methyl-Imidazole</b> .....	161
A1.1.1: Crystal Structure Data.....	161
A1.1.2: Geometry Optimizations of the Ergothioneine Crystal Structure.....	162
A1.1.3: Geometry Optimizations of 2-Thiol-4-Methyl-Imidazole .....	171
<b>A1.2: Studies of Ovothiol A and 4-thiol-N<sup>1</sup>-methyl-5-methyl-Imidazole</b> .....	172
A1.2.1: Geometry Optimizations of 4-Thiol-N <sup>1</sup> -Methyl-5-Methyl-Imidazole.....	173
<b>A1.3: Studies of Methyl Mercaptan and Other Biologically Relevant Compounds</b> .....	173
A1.3.1: Geometry Optimizations of Methyl Mercaptan and its Derivatives.....	173
A1.3.2: Geometry Optimizations of the Ascorbate Model Compound .....	174
A1.3.3: Geometry Optimizations of Thiophenol and Diphenyldisulfide .....	174
A1.3.4: Geometry Optimizations of Other Biologically Relevant Molecules.....	175
<b>A1.4: Reaction Enthalpy Calculations</b> .....	175
<b>A1.5: Electron Density Determination</b> .....	180
<b>Appendix 2: Phosphorylation by Inositol Pyrophosphates</b> .....	181
<b>A2.1: Hartree-Fock and Density Functional Techniques</b> .....	181
<b>A2.2: Molecular Mechanics Methods</b> .....	182
<b>A2.3: Results from Semi-Empirical Methods</b> .....	185

## List of Tables

### Chapter 1: Introduction to Intracellular Thiols

<b>Table 1.1:</b> The pKa values of glutathione .....	10
<b>Table 1.2:</b> Thiols produced by <i>Actinomycetales</i> .....	20

### Chapter 2: Modelling of Ergothioneine and Ovothiol

<b>Table 2.1:</b> As a basis set becomes more complete, the computed energies approach the HF limit .....	39
<b>Table 2.2:</b> Bond lengths (in Å) observed in the crystal structures of ESH, part A. ....	47
<b>Table 2.3:</b> Bond lengths (in Å) observed in the crystal structures of ESH, part B. ....	47
<b>Table 2.4:</b> Bond angles (in degrees) observed in the crystal structures of ESH, part A. ....	48
<b>Table 2.5:</b> Bond angles (in degrees) observed in the crystal structures of ESH, part B. ....	48
<b>Table 2.6:</b> Torsional angles (in degrees) observed in the crystal structures of ESH part A ...	48
<b>Table 2.7:</b> Torsional angles (in degrees) observed in the crystal structures of ESH part B ...	48
<b>Table 2.8:</b> Bond lengths (in Å) returned after the geometry optimization of the ergothioneine crystal structure using DFT methods with implicit solvation .....	50
<b>Table 2.9:</b> Bond angles (in degrees) returned after the geometry optimization of the ergothioneine crystal structure using DFT methods with implicit water solvation .....	51
<b>Table 2.10:</b> Selected dihedral angles (in degrees) returned after the geometry optimization of the ergothioneine crystal structure using DFT methods with implicit water solvation .....	52
<b>Table 2.11:</b> The number of functions in each basis set for the geometry optimization of ergothioneine using the crystal structure as the starting conformation .....	53
<b>Table 2.12:</b> The number of functions in each basis set and time required for the geometry optimization of phenol .....	53
<b>Table 2.13:</b> Calculated energies, bond lengths and angles of TMI, * its tautomers and ergothioneine .....	54
<b>Table 2.14:</b> Energies and selected bond lengths and angles for the tautomers of (TMI-S) <sub>2</sub> .....	55
<b>Table 2.15:</b> (TMI-S) <sub>2</sub> + 2 CH <sub>3</sub> SH → 2 TMithione + H <sub>3</sub> CSSCH <sub>3</sub> .....	55
<b>Table 2.16:</b> (TMI-S) <sub>2</sub> + 2 CH <sub>3</sub> SH → 2 TMithiol + H <sub>3</sub> CSSCH <sub>3</sub> .....	56
<b>Table 2.17:</b> 2 TMI <sub>thiol</sub> → 2 TMI <sub>thione</sub> .....	56
<b>Table 2.18:</b> A comparison of C-S bond lengths the various ionic states of TMMI to thiophenol .....	57
<b>Table 2.19:</b> 2 RSH + H <sub>2</sub> O <sub>2</sub> → RSSR + 2 H <sub>2</sub> O .....	59
<b>Table 2.20:</b> RSH + HO• → RS• + H <sub>2</sub> O .....	59
<b>Table 2.21:</b> RS <sup>-</sup> + HO• + H <sub>3</sub> O <sup>+</sup> → RS• + 2 H <sub>2</sub> O .....	60
<b>Table 2.22:</b> 2 RSH + 2 HO• → RSSR + 2 H <sub>2</sub> O .....	60
<b>Table 2.23:</b> RS• + 2-23 → RSH + 2-24 .....	63

\* TMI = 2-thiol-4-methyl-imidazole; (TMI-S)<sub>2</sub> = the disulfide of TMI

### Chapter 3: Conformational Analyses of Mycothiol

<b>Table 3.1:</b> Quality ratings of the force field parameters available in MacroModel.....	75
<b>Table 3.2:</b> Hydrogen-bonding interactions for the global minimum conformations of MSH found <i>in vacuo</i> and with implicit water solvation using the OPLS-AA force field.....	77
<b>Table 3.3:</b> Comparison of the three clusters and global minima of the implicitly solvated MSH conformers generated in this study.....	80
<b>Table 3.4:</b> Comparison of selected cysteinyl torsional angles from representative structures of the implicitly solvated mycothiol conformers found with the OPLS-AA force field to those of the bimane derivative of mycothiol found using NMR techniques.....	83
<b>Table 3.5:</b> Comparison of key inter-hydrogen distances for representative structures of the implicitly solvated mycothiol conformers found with the OPLS-AA force field to those of the bimane derivative of mycothiol found using NMR techniques.....	84

### Chapter 4: Efforts Towards the Identification of a Mycothiol-Utilizing Glyoxalase System

<b>Table 4.1:</b> Comparison of known glutathione and mycothiol biochemistry.....	91
<b>Table 4.2:</b> Useful antibiotics produced by <i>Streptomyces</i> bacteria.....	94
<b>Table 4.3:</b> Thiols produced by <i>Streptomyces</i> bacteria.....	94
<b>Table 4.4:</b> Solvent gradient used to purify <i>N</i> -acetyl-L-cysteinyl-2-amino-2-deoxy- $\alpha,\beta$ -D-glucopyranoside.....	108
<b>Table 4.5:</b> Solvent gradient used to purify 6-hydroxynaphthyl-mycothiol disulfide.....	112
<b>Table 4.6:</b> Results for the sequence searches of the <i>S. coelicolor</i> genome for homologous sequences to known mycothiol dependent enzymes.....	114
<b>Table 4.7:</b> Results for the sequence searches of the <i>Streptomyces coelicolor</i> genome for homologous sequences to known glyoxalase enzymes.....	127
<b>Table 4.8:</b> A comparison of the amount of mycothiol produced by <i>Streptomyces</i> bacteria to that produced by <i>Mycobacterium smegmatis</i> .....	131

### Chapter 5: Phosphorylation by Inositol Pyrophosphates

<b>Table 5.1:</b> The pKa values of pyrophosphate and its cleavage products phosphate and methyl phosphate.....	145
<b>Table 5.2:</b> The pKa values known for InsP <sub>6</sub> .....	145
<b>Table 5.3:</b> Calculated $\Delta H_{rxn}$ values for phosphorylation of methanol by pyrophosphate.....	150
<b>Table 5.4:</b> Calculated $\Delta H_{rxn}$ values for the phosphorylation of methanol by pyrophosphate using molecular mechanics methods.....	150
<b>Table 5.5:</b> $\Delta H_{rxn}$ values determined for the phosphorylation of methanol to form methyl phosphate.....	153
<b>Table 5.6:</b> $\Delta H_{rxn}$ values calculated for the phosphorylation of methanol by 5-PP-InsP <sub>5</sub> ionization states expected at physiological pH.....	154
<b>Table 5.7:</b> $\Delta H_{rxn}$ values calculated for the phosphorylation of methanol by 5,6-bis-PP-InsP <sub>4</sub> ionization states expected at pH 5.7.....	155
<b>Table 5.8:</b> $\Delta H_{rxn}$ values calculated for the phosphorylation of methanol by 5,6-bis-PP-InsP <sub>4</sub> ionization states expected at pH 6.8 – 7.6.....	156

<b>Table 5.9:</b> $\Delta H_{\text{rxn}}$ values calculated for the phosphorylation of methanol by ATP, ADP and model compounds. ....	157
---	-----

## Appendix 1: Modelling of Ergothioneine and Ovothiol

<b>Table A1.1:</b> Bond lengths of ergothioneine determined using X-ray crystallography in this work. ....	162
<b>Table A1.2:</b> Bond angles of ergothioneine as determined by X-ray crystallography in this work. ....	162
<b>Table A1.3:</b> Bond distances (in Å) returned by geometry optimizations of the crystal structure of ergothioneine using density functional theory methods in the gas phase. ....	163
<b>Table A1.4:</b> Bond angles (in °) returned by a geometry optimization of the ergothioneine crystal structure using density function theory methods in the gas phase .....	164
<b>Table A1.5:</b> Bond lengths (in Å) returned after geometry optimizations of the ergothioneine crystal structure using density functional theory methods with implicit water solvation. ....	165
<b>Table A1.6:</b> Bond angles (in °) returned by geometry optimizations of the ergothioneine crystal structure using density functional theory methods with implicit water solvation. ....	166
<b>Table A1.7:</b> Bond lengths (in Å) returned by geometry optimizations of the ergothioneine crystal structure using Hartree-Fock methods in the gas phase .....	167
<b>Table A1.8:</b> Bond angles (in °) returned by the geometry optimizations of the ergothioneine crystal structure using Hartree-Fock methods in the gas phase .....	168
<b>Table A1.9:</b> Bond lengths (in Å) returned by geometry optimizations of the ergothioneine crystal structure using Hartree-Fock methods with implicit water solvation .....	169
<b>Table A1.10:</b> Bond angles (in °) returned after the geometry optimization of the ergothioneine crystal structure using Hartree-Fock methods with implicit water solvation. ....	170
<b>Table A1.11:</b> Energies, bond lengths and angles of 2-thiol-4-methylimidazole, its tautomers and ergothioneine determined by geometry optimization .....	171
<b>Table A1.12:</b> Energies, bond lengths and angles of the thiolate and thiyl radicals of 2-thiol-4-methylimidazole determined by geometry optimization .....	171
<b>Table A1.13:</b> Energies and selected bond lengths and angles of the various tautomers of the disulfide of 2-thiol-4-methylimidazole determined by geometry optimization .....	172
<b>Table A1.14:</b> Energies and selected bond lengths and angles for ovothiol A and 4-thiol- $N^1$ -methyl-5-methylimidazole, its thiolate, thiyl radical and zwitterion determined by geometry optimization .....	173
<b>Table A1.15:</b> Energies and C-S bond lengths for methylmercaptan, its thiolate, thiyl radical and disulfide determined by geometry optimization .....	173
<b>Table A1.16:</b> Energies of the model compounds of ascorbate and its radical determined by geometry optimization .....	174
<b>Table A1.17:</b> Energies and bond lengths of thiophenol and diphenyldisulfide determined by geometry optimization .....	174
<b>Table A1.18:</b> Energies of other biologically relevant molecules determined by geometry optimization. ....	175
<b>Table A1.19:</b> $\text{RSSR} + 2 \text{CH}_3\text{SH} \rightarrow 2 \text{RSH} + \text{H}_3\text{CSSCH}_3$ .....	175
<b>Table A1.20:</b> $2 \text{TMI}_{\text{thiol}} \rightarrow 2 \text{TMI}_{\text{thione}}$ .....	176

<b>Table A1.21:</b> $2 \text{TMMI}^\dagger \rightarrow 2 \text{TMMI}_{\text{zwitterion}}$ .....	176
<b>Table A1.22:</b> $\text{TMMI}_{\text{zwitterion}} + \text{OH}^- \rightarrow \text{TMMI}^- + \text{H}_2\text{O}$ .....	176
<b>Table A1.23:</b> $2 \text{RSH} + \text{H}_2\text{O}_2 \rightarrow \text{RSSR} + 2 \text{H}_2\text{O}$ .....	177
<b>Table A1.24:</b> $\text{RSH} + \text{HO}\cdot \rightarrow \text{RS}\cdot + \text{H}_2\text{O}$ .....	177
<b>Table A1.25:</b> $\text{RS}^- + \text{HO}\cdot + \text{H}_3\text{O}^+ \rightarrow \text{RS}\cdot + 2 \text{H}_2\text{O}$ .....	178
<b>Table A1.26:</b> $2 \text{RSH} + 2 \text{HO}\cdot \rightarrow \text{RSSR} + 2 \text{H}_2\text{O}$ .....	178
<b>Table A1.27:</b> $\text{RS}\cdot + 2\text{-}23 \rightarrow \text{RSH} + 2\text{-}24$ .....	179
<b>Table A1.28:</b> Mullikan charges and spin densities for 2-thiol-4-methyl-imidazole and 4-thiol-N1-methyl-5-methylimidazole.....	180

## Appendix 2: Phosphorylation by Inositol Pyrophosphates

<b>Table A2.1:</b> $\Delta G^\circ_f$ values calculated using HF and DFT techniques .....	181
<b>Table A2.2:</b> $\Delta G^\circ_{\text{rxn}}$ values for the phosphorylation of methanol by pyrophosphate as calculated using HF and DFT methods.....	182
<b>Table A2.3:</b> Calculated $\Delta G^\circ_{\text{rxn}}$ values determined for the phosphorylation of methanol by InsPPs using HF and DFT techniques .....	182
<b>Table A2.4:</b> The quality ratings of the parameters of the force fields available in MacroModel.....	183
<b>Table A2.5:</b> Calculated $\Delta H_f$ values determined using molecular mechanics methods.....	184
<b>Table A2.6:</b> Calculated $\Delta H_{\text{rxn}}$ values for the phosphorylation of methanol by pyrophosphate using molecular mechanics methods.....	185
<b>Table A2.7:</b> Calculated $\Delta H_{\text{rxn}}$ values for the phosphorylation of methanol by the InsPPs as determined by molecular mechanics methods.....	185
<b>Table A2.8:</b> Calculated $\Delta H_f$ values determined using semi-empirical techniques for the various $\text{InsP}_6$ charged states studied. ....	186
<b>Table A2.9:</b> $\Delta H_f$ values calculated for the various 5-PP- $\text{InsP}_5$ charged states studied.....	187
<b>Table A2.10:</b> $\Delta H_f$ values calculated for the various 5, 6-bis-PP- $\text{InsP}_4$ charged states studied. ....	188
<b>Table A2.11:</b> $\Delta H_{\text{rxn}}$ values determined for the phosphorylation of methanol by various pyrophosphates .....	189
<b>Table A2.12:</b> $\Delta H_{\text{rxn}}$ values calculated for the phosphorylation of methanol by 5-PP- $\text{InsP}_5$ .....	189
<b>Table A2.13:</b> $\Delta H_{\text{rxn}}$ values calculated for the phosphorylation of methanol by 5-PP- $\text{InsP}_5$ .....	190
<b>Table A2.14:</b> $\Delta H_{\text{rxn}}$ values calculated for the phosphorylation of methanol by 5, 6-bis-PP- $\text{InsP}_5$ .....	191
<b>Table A2.15:</b> $\Delta H_{\text{rxn}}$ values calculated for the phosphorylation of methanol by 5, 6-bis-PP- $\text{InsP}_5$ .....	192
<b>Table A2.16:</b> $\Delta H_{\text{rxn}}$ values calculated for the phosphorylation of methanol by 5, 6-bis-PP- $\text{InsP}_5$ .....	193
<b>Table A2.17:</b> $\Delta H_{\text{rxn}}$ values calculated for the phosphorylation of methanol by 5, 6-bis-PP- $\text{InsP}_5$ .....	194
<b>Table A2.18:</b> $\Delta H_{\text{rxn}}$ values calculated for the phosphorylation of methanol by ATP, ADP and model compounds .....	195

<sup>†</sup> TMMI = 4-thiol-*N*<sup>1</sup>-methyl-5-imidazole; (TMMI-S)<sub>2</sub> = the disulfide of TMMI.

## List of Figures

### Chapter 1: Introduction to Intracellular Thiols

<b>Figure 1.1:</b> Glutathione, ergothioneine, the ovothiols and mycothiol .....	1
<b>Figure 1.2:</b> $\beta$ -Hydroxyergothioneine .....	2
<b>Figure 1.3:</b> The biosynthetic pathway of ergothioneine. ....	3
<b>Figure 1.4:</b> The three different tautomeric states of ergothioneine.....	3
<b>Figure 1.5:</b> The adenochromines and imbricatine .....	6
<b>Figure 1.6:</b> The proposed biosynthetic pathway of the ovothiols.....	7
<b>Figure 1.7:</b> 4-Thiol- $N^1$ -methyl-5-methyl-imidazole: a model for the ovothiolss. ....	7
<b>Figure 1.8:</b> The detoxification by the ovothiols of $H_2O_2$ formed during the fertilization of sea urchin eggs.....	8
<b>Figure 1.9:</b> Glutathione.....	9
<b>Figure 1.10:</b> The biosynthetic pathway of glutathione .....	10
<b>Figure 1.11:</b> The gentisate pathway.....	11
<b>Figure 1.12:</b> The naturally occurring forms of vitamin E.....	12
<b>Figure 1.13:</b> Examples of known physiological substrates of the glutathione- <i>S</i> -transferases .....	13
<b>Figure 1.14:</b> Selected pathways for the metabolism of glutathione- <i>S</i> -conjugates.....	13
<b>Figure 1.15:</b> The reaction catalyzed by glutathione dependent formaldehyde dehydrogenase.....	14
<b>Figure 1.16:</b> Structural comparison of the two substrates of glutathione dependent formaldehyde dehydrogenase. ....	15
<b>Figure 1.17:</b> The isomerization catalyzed by triose phosphate isomerase.....	16
<b>Figure 1.18:</b> The glyoxalase system .....	17
<b>Figure 1.19:</b> Mycothiol. ....	19
<b>Figure 1.20:</b> The biosynthetic pathway of mycothiol.....	21
<b>Figure 1.21:</b> Selected substrates for mycothiol- <i>S</i> -conjugate amidase. ....	24

### Chapter 2: Modelling of Ergothioneine and Ovothiol

<b>Figure 2.1:</b> Ergothioneine and the ovothiols.....	36
<b>Figure 2.2:</b> A three dimensional potential energy surface .....	40
<b>Figure 2.3:</b> A schematic one-dimensional energy surface.....	41
<b>Figure 2.4:</b> The equilibrium between 4-hydroxypyridine and 4-pyridone. ....	41
<b>Figure 2.5:</b> Schematic of the treatment of a solute in a liquid by the implicit water solvation Polarized Continuum Model (PCM) .....	42
<b>Figure 2.6:</b> Vibrational modes of a system.....	43
<b>Figure 2.7:</b> Structures of ergothioneine, 2-thiol-4-methyl-imidazole and its disulfide. ....	45
<b>Figure 2.8:</b> Structures of ovothiol A and 4-thiol- $N^1$ -methyl-5-methyl-imidazole.....	46
<b>Figure 2.9:</b> Miscellaneous compounds used to model biologically relevant reactions. ....	46
<b>Figure 2.10:</b> Crystal packing diagram of ergothioneine crystals.....	49
<b>Figure 2.11:</b> Tautomers of 2-thiol-4-methyl-imidazole.....	54
<b>Figure 2.12:</b> The tautomers of the disulfide of 2-thiol-4-methyl-imidazole.....	54
<b>Figure 2.13:</b> The hydrolysis of phosphoenolpyruvate .....	56

<b>Figure 2.14:</b> 4-Thiol- <i>N</i> <sup>1</sup> -methyl-5-methyl-imidazole and its disulfide.....	57
<b>Figure 2.15:</b> Spin density localization for the 2-thiol-4-methyl-imidazole tautomer ( <b>2-9</b> ) in the gas phase and with implicit water solvation .....	61
<b>Figure 2.16:</b> Spin density localization for the 2-thiol-4-methyl-imidazole tautomer ( <b>2-10</b> ) in the gas phase and with implicit water solvation .....	61
<b>Figure 2.17:</b> Spin density localization for 4-thiol- <i>N</i> <sup>1</sup> -methyl-5-methyl-imidazole in the gas phase and with implicit water solvation.....	61
<b>Figure 2.18:</b> Model compounds used to study the reactions of ascorbic acid. ....	62
<b>Figure 2.19:</b> Tautomerization of the ascobate model. ....	62

### Chapter 3: Conformational Analyses of Mycothiol

<b>Figure 3.1:</b> Mycothiol .....	67
<b>Figure 3.2:</b> Variation in the energy of pentane with the variation of select angles .....	68
<b>Figure 3.3:</b> A one dimensional representation of the energy of all possible conformers of a simple molecule .....	68
<b>Figure 3.4:</b> Sampling the conformational space of a molecule using a Monte Carlo search method with a small number of iterations .....	70
<b>Figure 3.5:</b> Hydrolysis of modified mycothiol by mycothiol- <i>S</i> -conjugate amidase.....	74
<b>Figure 3.6:</b> Stereoview overlay of the global minimum conformations of mycothiol found for the gas and implicit water phases.....	77
<b>Figure 3.7:</b> Clustering of the $\varphi$ and $\psi$ angles of the low-energy mycothiol conformers. ....	78
<b>Figure 3.8:</b> Comparison of the $\varphi$ and $\psi$ angles and energies of the low-energy mycothiol conformers. ....	79
<b>Figure 3.9:</b> Stereoview of the superimposition of the lowest energy conformers of each cluster as determined by the OPLS-AA force field with implicit water solvation .....	79
<b>Figure 3.10:</b> The bimane derivative of mycothiol. ....	83
<b>Figure 3.11:</b> Trypanothione and select examples of substrate-based inhibitors of trypanothione reductase. ....	85

### Chapter 4: Efforts Towards the Identification of a Mycothiol-Utilizing Glyoxalase System

<b>Figure 4.1:</b> Glutathione and mycothiol.....	89
<b>Figure 4.2:</b> The two-enzyme glyoxalase system composed of GlxI and GlxII. ....	90
<b>Figure 4.3:</b> Trypanothione. ....	90
<b>Figure 4.4:</b> The isomerization catalyzed by triose phosphate isomerase.....	92
<b>Figure 4.5:</b> The biosynthesis of mycothiol. ....	95
<b>Figure 4.6:</b> The synthesis of mycothiol as reported by Jardine <i>et al.</i> ....	97
<b>Figure 4.7:</b> The synthesis of the bimane derivative of mycothiol as described by Nicholas <i>et al.</i> ....	98
<b>Figure 4.8:</b> The synthesis of mycothiol as reported by Lee and Rosazza.....	99
<b>Figure 4.9:</b> The synthesis of des- <i>myo</i> -inositol mycothiol as reported by Patel and Blanchard .....	100
<b>Figure 4.10:</b> The isolation of mycothiol using 2- <i>S</i> -(2'-thiopyridyl)-6- hydroxynaphthyldisulfide as described by Steenkamp and Vogt.....	101
<b>Figure 4.11:</b> The synthesis of Fmoc-Cys(Acm)GlcN.....	106



<b>Figure 4.12:</b> The synthesis of H <sub>2</sub> N-Cys(Acm)GlcN.....	107
<b>Figure 4.13:</b> The synthesis of AcCys(Acm)GlcN.....	108
<b>Figure 4.14:</b> The proposed synthesis of (AcCys-GlcN) <sub>2</sub> .....	109
<b>Figure 4.15:</b> The synthesis of AcCys(Trt)GlcN.....	109
<b>Figure 4.16:</b> The reduction of 6-hydroxynaphthyl disulfide with sodium borohydride.....	110
<b>Figure 4.17:</b> The synthesis of 2- <i>S</i> -(2'-thiopyridyl)-6-hydroxynaphthyl disulfide.....	110
<b>Figure 4.18:</b> The formation of 2- <i>S</i> -6-hydroxynaphthyl-mycothiols disulfide.....	111
<b>Figure 4.19:</b> The synthesis of a proposed mycothiol analogue.....	112
<b>Figure 4.20:</b> The mechanism for the racemization of <i>N</i> -acyl amino acids that can occur during the base-catalyzed coupling reaction.....	130

## Chapter 5: Phosphorylation by Inositol Pyrophosphates

<b>Figure 5.1:</b> Hexaphosphorylated <i>myo</i> -inositol, 5-diphospho- <i>myo</i> -inositol pentakisphosphate, and 5,6-bis-diphospho- <i>myo</i> -inositol tetrakisphosphate.....	138
<b>Figure 5.2:</b> Agranoff's turtle.....	139
<b>Figure 5.3:</b> The biosynthesis of <i>L</i> - <i>myo</i> -inositol-1-phosphate ( <i>D</i> - <i>myo</i> -inositol-3-phosphate) catalyzed by <i>L</i> - <i>myo</i> -inositol-1-phosphate synthase.....	140
<b>Figure 5.4:</b> The compounds used to model phosphate transfer.....	145
<b>Figure 5.5:</b> The InsP <sub>6</sub> <sup>‡</sup> ionization states modelled in this study.....	146
<b>Figure 5.6:</b> The 5-PP-InsP <sub>5</sub> <sup>‡</sup> ionization states modelled in this study.....	147
<b>Figure 5.7:</b> The 5,6-bis-PP-InsP <sub>4</sub> <sup>‡</sup> ionization states modelled in this study.....	148
<b>Figure 5.8:</b> The angles of 5-PP-InsP <sub>5</sub> <sup>‡</sup> and 5,6-bis-PP-InsP <sub>4</sub> <sup>‡</sup> altered by dihedral angle driving.....	149
<b>Figure 5.9:</b> The electrostatic potential for 5, 6-bis-PP-InsP <sub>4</sub> <sup>‡</sup> with a charge of -8.....	156

## Appendix 1: Modelling of Ergothioneine and Ovothiols Supplementary Information

<b>Figure A1.1:</b> Structures of 2-thiol-4-methyl-imidazole and ergothioneine.....	161
<b>Figure A1.2:</b> Ovothiol A and 4-thiol- <i>N</i> <sup>1</sup> -methyl-5-methyl-imidazole used to model OSH and its forms.....	172
<b>Figure A1.3:</b> Methylmercaptan structures used to model glutathione and its various forms.....	173
<b>Figure A1.4:</b> The compounds used to model ascorbic acid and its radical.....	174
<b>Figure A1.5:</b> Thiophenol and diphenyldisulfide are used as models for aromatic thiols.....	174

<sup>‡</sup> InsP<sub>6</sub>: Hexaphosphorylated *myo*-inositol; 5-PP-InsP<sub>5</sub>: 5-diphospho-*myo*-inositol pentakisphosphate; 5,6-bis-PP-InsP<sub>4</sub>: 5,6-bis-diphospho-*myo*-inositol tetrakisphosphate

## Table of Abbreviations

$\alpha_1$ -AP	$\alpha_1$ -antiproteinase
<i>A. methanolica</i>	<i>Amycolaptopsis methanolica</i>
<i>A. thaliana</i>	<i>Arabidopsis thaliana</i>
Ac	Acetyl
Acetyl-CoA	Acetyl-coenzyme A
Ac <sub>2</sub> O	Acetic anhydride
AcOH	Acetic acid
Acm	Acetamidomethyl
ACV	$\delta$ -(L- $\alpha$ -Amino-adipyl)-L-cysteinyl-D-valine
AdoMet	S-Adenosyl-L-methione
ADP	Adenosine diphosphate
AgOTf	Silver triflate
AMP	Adenosine monophosphate
ATCC	American Type Culture Collection
ATP	Adenosine triphosphate
B3	Becke 3 hybrid scheme
B3LYP	Becke 3 hybrid scheme combined with the Lee-Yang-Parr correction functional
B3PW91	Becke 3 hybrid scheme combined with the Perdew-Wang exchange and correlation functional
1,4-bis-PP-InsP <sub>4</sub>	1,4-Diphospho-D- <i>myo</i> -inositol-tetrakisphosphate
1,5-bis-PP-InsP <sub>4</sub>	1,5-Diphospho-D- <i>myo</i> -inositol-tetrakisphosphate
3,6-bis-PP-InsP <sub>4</sub>	3,6-Diphospho-D- <i>myo</i> -inositol-tetrakisphosphate
5,6-bis-PP-InsP <sub>4</sub>	5,6-Diphospho-D- <i>myo</i> -inositol-tetrakisphosphate
bis-PP-InsP <sub>4</sub>	Diphospho-D- <i>myo</i> -inositol-tetrakisphosphate
BnBr	Benzyl bromide
<i>C. efficiens</i>	<i>Corynebacterium efficiens</i>
<i>C. glutamicum</i>	<i>Corynebacterium glutamicum</i>
<i>C. fasciculata</i>	<i>Crithidia fasciculata</i>
<sup>13</sup> C-NMR	<sup>13</sup> Carbon nuclear magnetic resonance
CH <sub>3</sub> CN	Acetonitrile
CoA	Coenzyme A
CPU	Central processing unit
Cyclohex-Pi	Cyclohexyl-phosphate
Cyclohex-PPi	Cyclohexyl-pyrophosphate
Cys	L-cysteine
Cys-GlcN-Ins	1-D- <i>myo</i> -inosityl-2-(L-cysteinyl)-amido-2-dexoy- $\alpha$ -D-glucopyranoside
$\Delta G^\circ$	Standard Gibbs free energy
$\Delta H_f$	Heat of formation
$\Delta H_{rxn}$	Change in enthalpy of a reaction
D	Number of degrees of freedom
<i>D. imbricate</i>	<i>Dermasterias imbricate</i>
<i>D. discoideum</i>	<i>Dictyostelium discoideum</i>
DEPC	Diethylphosphoryl cyanide

DFT	Density functional theory
DGAP	D-glyceraldehyde-3-phosphate
DHAP	Dihydroxyacetone phosphate
DMAP	4- <i>N,N</i> -dimethylaminopyridine
DMF	Dimethylformamide
DMSO	Dimethyl sulfoxide
DMTI	1, 5-Dimethyl-5-thioimidazole
DNA	Deoxyribonucleic acid
DTNB	5,5'-dithio-bis(2-nitrobenzoic acid)
DPPH	2, 2-Diphenyl-1-picrylhydrazyl
DZ	Double zeta
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	ethylenediaminetetraacetic acid
EI	Electron impact
Equiv.	Equivalents
ES•	Ergothioneine radical
ESH	Ergothioneine
ESH <sub>thiol</sub>	Thiol form of ergothioneine
ESH <sub>thione</sub>	Thione form of ergothioneine
ESI	Electrospray
ESSE	Disulfide of ergothioneine
EtOAc	Ethyl acetate
EtOH	Ethanol
FAD	Flavin adenine dinucleotide
FD	Formaldehyde dehydrogenase
ferMb	Ferrylmyoglobin
Fmoc	9-Fluorenylmethoxycarbonyl
g	Grams
GB/SA	Generalize Born / Surface Area
GD	Glutathione dependent
GD-FD	Glutathione dependent formaldehyde dehydrogenase
GlcN	D-Glucosamine
GlcN-HCl	D-Glucosamine hydrochloride
GlcNAc-Ins(3)P <sub>1</sub>	3-phospho-1-D- <i>myo</i> -inosityl-2-acetamido-2-deoxy- $\alpha$ -D-glucopyranoside
GlcNAc-Ins	1-D- <i>myo</i> -inosityl-2-acetamido-2-deoxy- $\alpha$ -D-glucopyranoside
GlcN-Ins	1-D- <i>myo</i> -inosityl-2-amino-2-deoxy- $\alpha$ -D-glucopyranoside
Glx	Glyoxalase
GlxI	Glyoxalase I
GlxII	Glyoxalase II
GlxIII	Glyoxalase III
GR	Glutathione disulfide reductase
GSH	Glutathione
GS•	Glutathione radical
GSNO	<i>S</i> -Nitrosoglutathione
GSSG	Disulfide of glutathione

GSSG• <sup>-</sup>	Glutathione disulfide radical anion
GST	Glutathione-S-transferase
GTP	Guanosine triphosphate
H <sub>2</sub> O	Water
H <sub>2</sub> O <sub>2</sub>	Hydrogen peroxide
H <sub>3</sub> O <sup>+</sup>	Hydronium ion
H <sub>2</sub> S	Hydrogen sulfide
Hb	Hemoglobin
HATU	O-(7-Azabenzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate
HBTU	O-(Benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate
HCl	Hydrochloric acid
HF	Hartree-Fock
HO•	Hydroxide radical
HOAt	1-Hydroxy-7-azabenzotriazole
HOBt	1-Hydroxybenzotriazole
HOCl	Hypochlorous acid
HPLC	High performance liquid chromatography
Hr	Hour
IEF-PCM	Integral equation formalism polarized continuum model
Ins	D- <i>myo</i> -Inositol
Ins(3)P <sub>1</sub>	D- <i>myo</i> -Inositol-3-phosphate
Ins(5)P <sub>1</sub>	D- <i>myo</i> -Inositol-5-phosphate
Ins(3,4)P <sub>2</sub>	D- <i>myo</i> -inositol-3,4-diphosphate
Ins(3,6)P <sub>2</sub>	D- <i>myo</i> -inositol-3,6-diphosphate
Ins(1,3,4)P <sub>3</sub>	D- <i>myo</i> -inositol-1,3,4-triphosphate
Ins(1,4,5)P <sub>3</sub>	D- <i>myo</i> -Inositol-1,4,5-triphosphate
Ins(3,4,6)P <sub>3</sub>	D- <i>myo</i> -Inositol-3,4,6-triphosphate
Ins(1,3,4,5)P <sub>4</sub>	D- <i>myo</i> -Inositol-1,3,4,6-tetraphosphate
Ins(1,3,4,6)P <sub>4</sub>	D- <i>myo</i> -Inositol-1,3,4,6-tetraphosphate
Ins(1,4,5,6)P <sub>4</sub>	D- <i>myo</i> -Inositol-1,4,5,6-tetraphosphate
Ins(1,3,4,5,6)P <sub>5</sub>	D- <i>myo</i> -Inositol-1,3,4,5,6-pentaphosphate
InsP <sub>6</sub>	D- <i>myo</i> -inositol hexaphosphate, also known as phytic acid
InsP <sub>6</sub> K	D- <i>myo</i> -inositol hexaphosphate kinase
InsPPs	Inositol pyrophosphates
InsPs	Inositol phosphates
ISP	International <i>Streptomyces</i> Project
kcal/mol	Kilocalories per mole
kJ/mol	Kilojoules per mole
<i>K. pneumoniae</i>	<i>Klebsiella pneumoniae</i>
L	Litre
<i>L. infantum</i>	<i>Leishmania infantum</i>
<i>L. major</i>	<i>Leishmania major</i>
LDH	Lactate dehydrogenase
LMP2	Local second-order Møller-Plesset

LYP	Lee-Yang-Parr correction functional
<i>M. avium</i>	<i>Mycobacterium avium</i>
<i>M. bovis</i>	<i>Mycobacterium bovis</i>
<i>M. leprae</i>	<i>Mycobacterium leprae</i>
<i>M. luteus</i>	<i>Mycobacterium luteus</i>
<i>M. smegmatis</i>	<i>Mycobacterium smegmatis</i>
<i>M. tuberculosis</i>	<i>Mycobacterium tuberculosis</i>
M	Molar
Mb	Myoglobin
MC	Monte Carlo
MCA	Mycothiols-S-conjugate amidase
MCMM	Monte Carlo multiple minimization
MD	Mycothiols dependent
MD-FD	Mycothiols dependent formaldehyde dehydrogenase
MDyn	Molecular dynamics
MeOH	Methanol
MeOPi	Methyl phosphate
MeOPPi	Methyl pyrophosphate
MeOPPPi	Methyl triphosphate
metHb	Methemoglobin
metMb	Metmyoglobin
MG	Methylglyoxal
min	Minutes
mL	Millilitre
mM	Millimolar
mmol	Millimole
MM	Molecular mechanics
MOE	Molecular Operative Environment
MP	Møller-Plesset perturbation theory
MR	Mycothiols disulfide reductase
MS	Mass spectrometry
MSH	Mycothiols
MshA	UDP-N-acetylglucosamine: 3-D- <i>myo</i> -inositol-1-phosphate-1- $\alpha$ -D-N-acetylglucosaminyltransferase
MshA2	An as yet unidentified kinase responsible for the removal of the phosphate group from 3-phospho-1-D- <i>myo</i> -inosityl-2-acetamido-2-deoxy- $\alpha$ -D-glucopyranoside
MshB	N-Acetyl-1-D- <i>myo</i> -inosityl-2-amino-2-deoxy- $\alpha$ -D-glucopyranoside deacetylase
MshC	L-Cysteine: 1-D- <i>myo</i> -inosityl-2-amino-2-deoxy- $\alpha$ -D-glucopyranoside ligase
MshD	Mycothiols synthase
<i>msha</i>	The gene which codes for MshA
<i>mshb</i>	The gene which codes for MshB
<i>mshc</i>	The gene which codes for MshC
<i>mshd</i>	The gene which codes for MshD

MsmB	Monobromobimane derivative of MSH
MSNO	S-Nitrosomycothiols
MSNH <sub>2</sub> OH	N-Hydroxysulfenamide of mycothiol
MSSM	Disulfide of mycothiol
NAD <sup>+</sup>	Nicotinamide adenine dinucleotide
NADH	Reduced nicotinamide adenine dinucleotide
NADP <sup>+</sup>	Nicotinamide adenine dinucleotide phosphate
NADPH	Reduced nicotinamide adenine dinucleotide phosphate
NaOH	Sodium hydroxide
N-formyl-Cys- GlcN-Ins	1-D- <i>myo</i> -inosityl-2-(N-formyl-L-cysteinyl)-amido-2-deoxy- $\alpha$ -D-glucopyranoside
NH <sub>3</sub>	Ammonia
nm	Nanometer
NMR	Nuclear Magnetic Resonance
NO	Nitric oxide
NOE	Nuclear Overhauser Effect
<sup>1</sup> O <sub>2</sub>	Singlet oxygen
O <sub>2</sub> <sup>•-</sup>	Superoxide
ONOO <sup>-</sup>	Peroxynitrite
OPfp	Pentafluorophenyl
OS <sup>-</sup>	Ovothiol anion
OS <sup>•</sup>	Ovothiol radical
OSHs	Ovothiols
OSH <sub>A</sub>	Ovothiol A
OSH <sub>B</sub>	Ovothiol B
OSH <sub>C</sub>	Ovothiol C
OSSO	Disulfide of the ovothiols
OSSO <sub>A</sub>	Disulfide of ovothiol A
OSSO <sub>C</sub>	Disulfide of ovothiol C
<i>O. vulgaris</i>	<i>Octopus vulgaris</i>
oxyHb	Oxyhemoglobin
<i>P. dumerilii</i>	<i>Platynereis dumerilii</i>
<i>P. pallidum</i>	<i>Polysphondylium pallidum</i>
PBS	Phosphate Buffered Saline
PCM	Polarized continuum model
P <sub>E</sub>	Potential energy
PhMe	Toluene
Pi	Phosphate
5-PP-Ins	D- <i>myo</i> -Inositol-5-pyrophosphate
5-PP-InsP <sub>5</sub>	5-Diphospho-D- <i>myo</i> -inositol-pentakisphosphate
6-PP-InsP <sub>5</sub>	6-Diphospho-D- <i>myo</i> -inositol-pentakisphosphate
PP-InsP <sub>5</sub>	Diphospho-D- <i>myo</i> -inositol-pentakisphosphate
PPi	Pyrophosphate
PPPi	Triphosphate
PRCG	Polak-Ribiere Conjugate Gradient

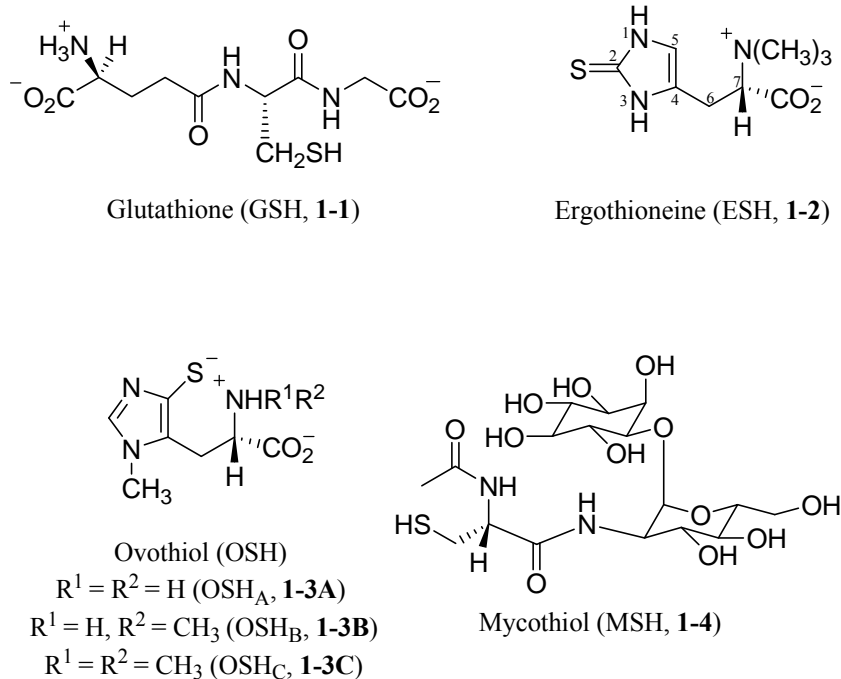
iPr <sub>2</sub> EtN	Diisopropylethylamine
RBCs	Red blood cells
Redox	Reduction-oxidation
RHF	Restricted Hartree-Fock
RIPS	Random incremental pulse search
RMS	Root-mean-square
RNA	Ribonucleic acid
RNS	Reactive nitrogen species
Rpm	Rotations per minute
RO•	Alkoxy radical
ROS	Reactive oxygen species
RS <sup>-</sup>	Alkyl thiolate
RS•	Alkyl thiyl radical
<i>S. ambofaciens</i>	<i>Streptomyces ambofaciens</i>
<i>S. antibioticus</i>	<i>Streptomyces antibioticus</i>
<i>S. argillaceus</i>	<i>Streptomyces argillaceus</i>
<i>S. aureofaciens</i>	<i>Streptomyces aureofaciens</i>
<i>S. avermitilis</i>	<i>Streptomyces avermitilis</i>
<i>S. caespitosus</i>	<i>Streptomyces caespitosus</i>
<i>S. cattleya</i>	<i>Streptomcyes cattleya</i>
<i>S. cerevisiae</i>	<i>Saccharomyces cerevisiae</i>
<i>S. coelicolor</i>	<i>Streptomyces coelicolor</i>
<i>S. fradiae</i>	<i>Streptomyces. fradiae</i>
<i>S. graminofaciens</i>	<i>Streptomyces graminofaciens</i>
<i>S. griseus</i>	<i>Streptomyces griseus</i>
<i>S. hygrosopicus</i>	<i>Streptomyces hygrosopicus</i>
<i>S. jumonjinensis</i>	<i>Streptomyces jumonjinensis</i>
<i>S. kanamyceticus</i>	<i>Streptomyces kanamyceticus</i>
<i>S. lincolnensis</i>	<i>Streptomyces lincolnensis</i>
<i>S. niveus</i>	<i>Streptomyces niveus</i>
<i>S. orchidaceus</i>	<i>Streptomyces orchidaceus</i>
<i>S. pristinaespiralis</i>	<i>Streptomyces pristinaespiralis</i>
<i>S. purpuratus</i>	<i>Strongylocentrotus purpuratus</i>
<i>S. rimosus</i>	<i>Streptomyces rimosus</i>
<i>S. roseosporus</i>	<i>Streptomyces roseosporus</i>
<i>S. spp.</i>	<i>Streptomyces spp.</i>
<i>S. spectabilis</i>	<i>Streptomyces spectabilis</i>
<i>S. tenebrarius</i>	<i>Streptomyces tenebrarius</i>
<i>S. venezuelae</i>	<i>Streptomyces venezuelae</i>
<i>S. verticillus</i>	<i>Streptomyces verticillus</i>
SASA	Solvent accessible surface area
SCF	Self-consistent field
SCRF	Self-consistent reaction field
Sec	Seconds
SGI	Silicon Graphics Inc.
S-HM-GSH	S-Hydroxymethyl glutathione

<i>T. brucei brucei</i>	<i>Trypanosoma brucei brucei</i>
TB	Tuberculosis
TEA	Triethylamine
TES	<i>N</i> -(Tris(hydroxymethyl)methyl)-2-aminoethanesulfonic acid
TFA	Trifluoroacetic acid
TIM	Triose phosphate isomerase
TLC	Thin layer chromatography
TMI	2-Thiol-4-methylimidazole
(TMI-S) <sub>2</sub>	Disulfide of 2-thiol-4-methyl-imidazole
TMI <sub>thiol</sub>	Thiol form of 2-thiol-4-methyl-imidazole
TMI <sub>thione</sub>	Thione form of 2-thiol-4-methyl-imidazole
TMMI	4-Thiol- <i>N</i> <sup>1</sup> -methyl-5-methyl-imidazole
TMMI <sup>-</sup>	4-Thiol- <i>N</i> <sup>1</sup> -methyl-5-methyl-imidazole anion
TMMI•	4-Thiol- <i>N</i> <sup>1</sup> -methyl-5-methyl-imidazole radical
(TMMI-S) <sub>2</sub>	Disulfide of 4-thiol- <i>N</i> <sup>1</sup> -methyl-5-methyl-imidazole
TMSOTf	Trimethylsilyl triflate
T <sub>opt</sub>	Time required for a geometry optimization
TSH <sub>2</sub>	Trypanothione ( <i>N</i> <sup>1</sup> , <i>N</i> <sup>8</sup> -bis(glutathione)-spermidine)
<i>p</i> -TsOH	<i>p</i> -Toluenesulfonic acid
T <sub>single</sub>	Time required for a single point energy calculation
TZ	Triple zeta
μm	Micromolar
μL	Microlitre
UDP	Uridine diphosphate
UDP-GlcNAc	Uridine diphosphate- <i>N</i> -acetylglucosamine
UCSD	University of California San Diego
WHO	World Health Organization
YEME	Yeast extract malt extract
ZPE	Zero point energy



## Chapter 1: Introduction to Intracellular Thiols

The presence of thiols in living systems is critical for the maintenance of cellular redox potentials and protein-disulfide ratios, as well as for the protection against reactive oxygen species (ROS). In addition to the well-studied tripeptide  $\gamma$ -L-glutamyl-L-cysteinylglycine (glutathione, GSH **1-1**), a number of compounds have been identified that contribute to these essential cellular roles. Our research focused on the properties of several critically important and naturally occurring thiols (Figure 1.1): ergothioneine (ESH, **1-2**), the ovothiols (OSHS, **1-3A-C**) and mycothiol (MSH, **1-4**). ESH, although only biosynthesized by fungi and the *Actinomycetales* bacteria, is present at significant levels in humans and may contribute to single electron redox reactions in human cells. The OSHs appear to function as important modulators of reactive oxygen toxicity and act as small molecule mimics of GSH peroxidase, a key enzyme in the detoxification of ROS. MSH is essential to the biochemistry of a number of human pathogens, including *Mycobacterium tuberculosis*, and the enzymes utilizing this thiol have been recognized as important novel drug targets.



**Figure 1.1:** Glutathione, ergothioneine, the ovothiols and mycothiol

Despite the involvement of these intracellular thiols in essential biological processes and their possible therapeutic applications, very little is known regarding their fundamental behaviour. To gain insight into the detoxification reactions of ESH and the OSHs, we performed a series of calculations to determine the thermodynamics of these reactions. We also applied theoretical techniques to determine the preferred conformations of MSH, which may aid in the design of substrate based inhibitors. In addition, we sought to identify a new MSH-utilizing enzyme system in *Actinomycetales* bacteria, which could serve as a novel drug target against the pathogenic members of this class. Due to their critical involvement in cellular biochemistry, any information gained regarding these thiols, should increase our understanding of intracellular thiol behaviour.

## 1.1. Ergothioneine

ESH has been identified in a diverse group of organisms including humans and the pathogenic *Mycobacteria*, and may be of great clinical importance, not only due to its role in pathogenic species, but also due to its possible therapeutic applications. ESH may assist in the treatment of chronic lung inflammations such as asthma<sup>1</sup> and may have a role in the protection against Alzheimer's disease.<sup>2</sup> While many possible roles have been suggested for ESH, much remains to be determined regarding the definitive role that ESH plays in the cell.

### 1.1.1. Distribution

ESH was first isolated from ergot (*Claviceps purpurea*), the fungal infection of rye grain, by Tanret in 1909<sup>3</sup> and its structure was determined to be a novel betaine of 2-thiol-L-histidine.<sup>4</sup> ESH has since been detected in plants,<sup>5-8</sup> fungi,<sup>9-11</sup> bacteria,<sup>11-13</sup> humans<sup>14,15</sup> and other animals<sup>16-19</sup> at millimolar concentrations. To date, evidence for the biosynthesis of ESH has only been noted in certain fungi<sup>9,11,20</sup> and the *Actinomycetales* bacteria.<sup>12,13</sup> It has been shown that *Bacillus*, *Clostridium*, *Corynebacterium*, *Escherichia*, *Lactobacillus*, *Propionibacterium*, *Proteus*, *Pseudomonas*, *Staphylococcus*, *Streptococcus* and *Vibrio* bacteria do not biosynthesize ESH.<sup>12</sup> The presence of an ESH derivative,  $\beta$ -hydroxyergothioneine, has recently been identified in the mushroom *Lyophyllum connatum* (Figure 1.2, 1-5).<sup>21</sup>

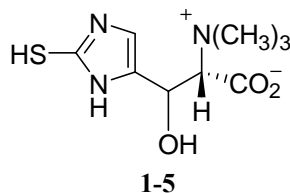
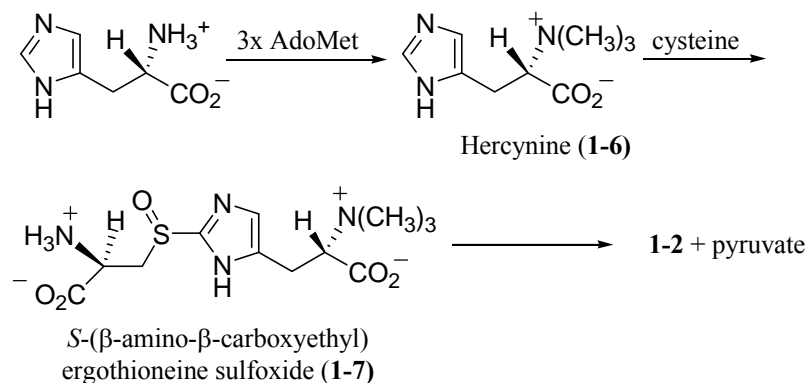


Figure 1.2:  $\beta$ -Hydroxyergothioneine

It is believed that organisms that do not biosynthesize ESH obtain it from their environment. Plants are thought to obtain ESH from the soil,<sup>6</sup> as oat seedlings have been shown to incorporate ESH via their root system.<sup>22</sup> Animals are believed to obtain ESH from dietary sources,<sup>18</sup> since it has been observed that diet can affect an organism's ESH content.<sup>5,8</sup> ESH levels are found to vary between species<sup>16</sup> and tissue type,<sup>23</sup> but it is mainly found in tissues exposed to oxidative stress including red blood cells (RBCs),<sup>17,24-26</sup> seminal fluid,<sup>27,28</sup> kidneys,<sup>29</sup> liver,<sup>29,30</sup> the ocular lens<sup>15,16</sup> and the brain.<sup>19</sup> RBCs sequester ESH at concentrations two to nine times greater than that of blood plasma. It was originally believed that ESH was incorporated into RBCs during cell formation,<sup>31</sup> but recent work indicates that RBCs gradually take up ESH<sup>32</sup> and maintain ESH levels against a concentration gradient. The uptake of ESH is believed to be transporter dependent; however, a specific ESH transporter in RBCs has not yet been identified.<sup>32</sup> An ESH transporter has recently been identified in human kidney cells and cells expressing this transporter were found to accumulate ESH to high levels and avidly retain it. Cells without this transporter were found to lack ESH as ESH cannot pass through the plasma membrane.<sup>33</sup>

### 1.1.2. Biosynthesis

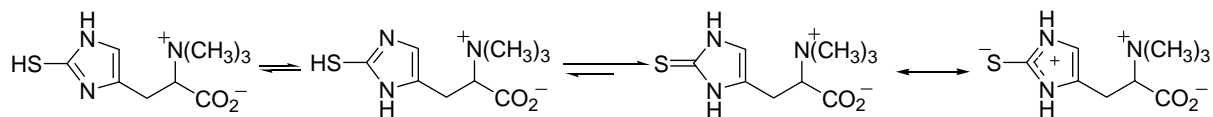
Initial studies of the biosynthesis of ESH (Figure 1.3) in fungi have shown that the pathway begins with the enzymatic methylation of the  $\alpha$ -amino group of L-histidine three times by *S*-adenosyl-L-methionine to yield hercynine (**1-6**).<sup>20,22,34</sup> The sulfur group is then added by the coupling of cysteine to hercynine to form an *S*-( $\beta$ -amino- $\beta$ -carboxyethyl)-ergothioneine sulfoxide intermediate (**1-7**) through an undetermined pathway requiring oxygen and iron (II).<sup>20,35</sup> This intermediate is then converted to ESH and pyruvate via an unknown pyridoxal-requiring enzyme.<sup>35</sup> Detailed studies of the biosynthesis of ESH in bacteria have not yet been performed; however, the use of radioactive precursors has indicated that the pathway parallels that of fungi.<sup>13</sup>



**Figure 1.3:** The biosynthetic pathway of ergothioneine. AdoMet = *S*-adenosyl-L-methionine.

### 1.1.3. Chemical Properties

ESH can exist as three different tautomers (Figure 1.4) with the thione being favoured under physiological conditions.<sup>36</sup> The C-S bond length has been determined by X-ray crystallography to be intermediate to the C-S single and double bond lengths, confirming the contribution of the thione.<sup>37</sup> The thione/thiol tautomerism of ESH causes the reactivity of ESH to differ from that of other thiols. ESH will react with certain thiol reactive agents including iodoacetamide,<sup>38</sup> the bromobimanes,<sup>10</sup> and 2,2-diphenyl-1-picrylhydrazyl (DPPH),<sup>39</sup> a stable free radical known to react with sulfhydryl groups to form thiyl radicals.<sup>40</sup> Unlike standard aliphatic thiols, ESH will not react with *N*-methyl-*N'*-nitro-*N*-nitroguanidine,<sup>41</sup> metmyoglobin (metMb),<sup>42</sup> nor carbon centered or aliphatic peroxy radicals.<sup>43</sup> At a pH greater than 7, ESH will ionize to the thiolate, which is a better electron donor; however, it is still resistant to carbon centered radicals.<sup>43</sup>



**Figure 1.4:** The three different tautomeric states of ergothioneine.

The most significant difference between ESH and other thiols is the instability of the ESH disulfide (ESSE) under physiological conditions. Most low molecular weight thiols are known to form disulfides upon exposure to aerated aqueous solutions; their

cognant organisms produce disulfide reductases to maintain these thiols in their biologically useful free thiol form.<sup>44</sup> ESSE does not form under conditions known to oxidize GSH and has only been shown to be stable under strongly acidic conditions, where the free and oxidized forms exist simultaneously.<sup>45</sup> Even if ESSE was formed *in vivo*, the large concentrations of GSH or MSH in these organisms would likely maintain ESH in its reduced state.

#### 1.1.4. Biological Function

Despite the large amount of literature related to ESH, its definitive role in the cell has yet to be determined. ESH has been implicated as a possible antioxidant and recent studies have shown that ESH may have superior scavenging ability for oxygen centered radicals as compared to the well-known antioxidant GSH. ESH may therefore be involved in, or be a useful treatment for, diseases known to involve oxidative stress. Susceptibility to the chronic inflammatory diseases Crohn's disease and rheumatoid arthritis have been associated with variants in the gene coding for the ESH transporter.<sup>33,46,47</sup> In addition, ESH has been shown to inhibit the release of certain mediators of inflammation in the lung, which may indicate a possible therapy to inhibit the chronic immune response known to occur in airway inflammations such as asthma and adult respiratory distress syndrome.<sup>48</sup>

ESH may also be a useful therapeutic agent for disease states caused by oxidative stress such as Alzheimer's disease. Alzheimer's patients show neuronal degradation and increased levels of nitrosotyrosine caused by peroxynitrite ( $\text{ONOO}^-$ ).<sup>49-51</sup> ESH has been shown to attenuate intracellular formation of  $\text{ONOO}^-$  through  $\text{ONOO}^-$  scavenging,<sup>2</sup> which is consistent with the observation that ESH can inhibit the nitration of tyrosine.<sup>52</sup> ESH may therefore decrease the neuronal apoptosis caused by Alzheimer's disease.<sup>2</sup>

In addition to its possible therapeutic applications, ESH may be important for the cellular response to oxidative stress caused by ROS and reactive nitrogen species (RNS). The hydroxide radical ( $\text{HO}\cdot$ ), for example, is one of the most reactive biologically relevant free radicals, attacking a number of essential cellular components.<sup>53</sup> Studies have shown that ESH can scavenge  $\text{HO}\cdot$  in a biologically relevant manner,<sup>54</sup> through the formation of a putative radical cation intermediate,<sup>55</sup> which is believed to decay into ESSE.<sup>54</sup> ESH has also been shown to inhibit the copper and iron catalyzed production of  $\text{HO}\cdot$ <sup>54</sup> and 2-thiol-imidazoles, such as ESH, have been shown to protect ascorbic acid from copper (II) catalyzed oxidation.<sup>56</sup> It has been suggested that this protection indicates that a cooperative interaction may exist between ESH and ascorbic acid in the protection of biological systems from ROS.<sup>55</sup>

Most thiols, including GSH, are oxidized to form disulfides and peroxides in the presence of heavy metals, such as copper or iron.<sup>57,58</sup> ESH however, has been shown to chelate divalent metals including cadmium (II), cobalt (II), copper (II), mercury (II), nickel (II) and zinc (II), and is not auto-oxidized in the presence of these metals to ESSE.<sup>59</sup> At physiological pH, ESH-copper (II) complexes are stable and, unlike GSH, do not decompose to give radicals.<sup>60,61</sup> The chelation of copper into redox inactive forms may be a major function of ESH.<sup>54</sup>

Hydrogen peroxide ( $\text{H}_2\text{O}_2$ ) is not as reactive as  $\text{HO}\cdot$ ; however it is membrane permeable and once in the cell, it can react with thiols, inactivating thiol-containing enzymes.<sup>62</sup>  $\text{H}_2\text{O}_2$  can interact with iron (II) and copper (I) to form  $\text{HO}\cdot$ , which is highly toxic to the cell.<sup>62</sup>  $\text{H}_2\text{O}_2$  can also oxidize metMb very quickly to ferrylmyoglobin

(ferMb),<sup>54</sup> which has peroxidase activity against many cellular components such as membrane lipids and fatty acids.<sup>63</sup> The addition of ESH to ferMb has been shown to quickly recover myoglobin (Mb) and its addition to mixtures of Mb or hemoglobin (Hb) and H<sub>2</sub>O<sub>2</sub> has been shown to inhibit the peroxidation of fatty acids through the transient formation of the ergothioneine radical (ES•) and ESSE.<sup>54</sup> The reactivity of ESH with ferMb is greater than that of GSH and is closer to that of ascorbate,<sup>63</sup> a known radical scavenger.<sup>62,64</sup> In addition, the formation of methemoglobin (metHb) from oxyhemoglobin (oxyHb) is also inhibited by 2-thiol-imidazoles such as ESH,<sup>56,65</sup> which would aid in maintaining the useful Hb concentration in biological systems.

Hypochlorous acid (HOCl) is a potent oxidizing and chlorinating agent that can react with primary amines to form cytotoxic *N*-chloroamines.<sup>66</sup> HOCl can induce cell lysis at low concentrations and can oxidize plasma membrane thiol groups causing the inactivation of glucose and amino acid transporters and a loss in potassium ion pumping activity.<sup>67</sup> An important biological target of HOCl is  $\alpha_1$ -antiproteinase ( $\alpha_1$ -AP), the major inhibitor of serine proteinases in bodily fluids. Few compounds can protect  $\alpha_1$ -AP against HOCl inactivation; however, ESH was shown to scavenge HOCl in a biologically relevant manner at concentrations as low as 500  $\mu$ M.<sup>54</sup>

To date, the role of ESH appears to be as an antioxidant, active in the prevention of damage caused by ROS and RNS. Further investigation into the antioxidant behaviour of ESH may elucidate the true role of ESH in biological systems and any information gained may be applicable to the development of therapeutics against disease states caused by oxidative damage. The role of ESH in the pathogenic *Mycobacteria* has yet to be explored, and studies in this area may provide novel targets for the development of antimycobacterial agents.

## 1.2. Ovothiols

The OSHs are not as widely found as ESH; however, they too may be of medical significance. In sea urchin eggs, the OSHs have been shown to be scavengers of H<sub>2</sub>O<sub>2</sub>,<sup>68</sup> which has led to the development of medicinally useful antioxidants based upon the OSHs.<sup>69-71</sup> The OSHs have also been identified in the pathogenic *Trypanosoma*,<sup>72</sup> and therefore may serve as useful drug targets against diseases such as African sleeping sickness,<sup>73</sup> Chagas disease<sup>74</sup> and visceral leishmaniasis.<sup>75</sup>

### 1.2.1. Distribution

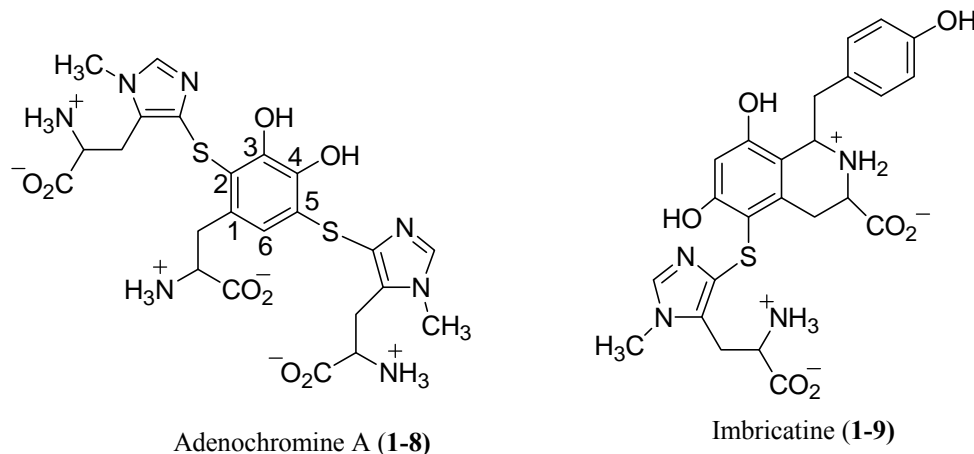
The disulfide of OSH<sub>A</sub>\* (OSSO<sub>A</sub>) has been identified in a number of echinoderms, including the sea urchins *Paracentrotus lividus*, *Sphaerechinus granularis* and *Arabacia lixula*, the holothuroid *Holothuria tubulosa*, and the asteroids *Astropecten auratiacus* and *Marthasterias glacialis*,<sup>76</sup> as well as in the marine invertebrate *Platynereis dumerili*,<sup>77</sup> and the mottle sea star *Evasterias troschelli*.<sup>78</sup> OSH<sub>A</sub> (1-3A) and 4-thiol-L-histidine have been identified in the common octopus *Octopus vulgaris*, the European squid *Loligo vulgaris*<sup>79</sup> and in all insect stages of the pathogenic Kinetoplastida,<sup>80</sup> such as *Crithidia fasciculata*<sup>81</sup> and *Leishmania donovani*.<sup>72</sup> OSH<sub>B</sub> (1-3B) and OSH<sub>C</sub> (1-3C) have been identified in the spiny scallop *Chlamys hastata* and the purple sea urchin *Strongylocentrotus purpuratus*

---

\* Initial work by Palumbo *et al.* (1984) identified the OSHs as 1-methyl-5-thiol-L-histidine due to confusion surrounding the nomenclature and numbering of substituted histidine rings. The corrected numbering system assigns OSHs as 1-methyl-4-thiol-L-histidine (Turner *et al.*, 1987).

respectively.<sup>78</sup> OSHs have also been identified in the eggs of rainbow trout and Coho salmon<sup>68</sup> and in the halo-tolerant green alga *Dunaliella salina*.<sup>82</sup>

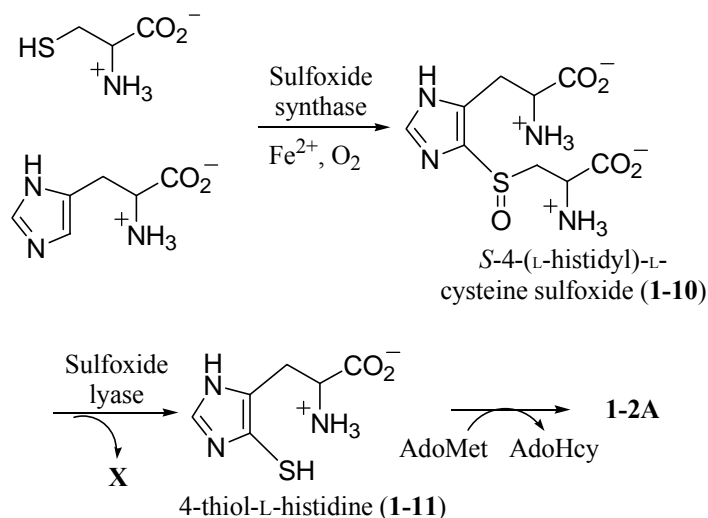
The OSHs are also components of the adenochromines and imbricatinine. The adenochromines are found in *O. vulgaris* and consist of two OSH<sub>A</sub> moieties attached at various locations to a substituted phenyl ring: in adenochromine A (Figure 1.5, **1-8**), the attachment points are positions 2 and 6 while in B and C the attachment points are positions 5 and 6 and positions 2 and 6 respectively.<sup>83</sup> Imbricatinine is produced by the starfish *Dermasterias imbricata* and, like the adenochromines, has OSH<sub>A</sub> attached to a substituted phenyl ring (Figure 1.5, **1-9**).<sup>84</sup>



**Figure 1.5:** The adenochromines are found in adenochrome, the purple iron (III) sequestering pigment of *Octopus vulgaris*, and differ in the position of the othiols on the phenyl ring. Imbricatinine (**1-8**) is found in *Dermasteria imbricata*, a starfish.

### 1.2.2. Biosynthesis

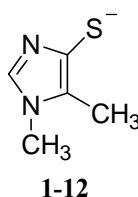
The exact biosynthetic pathway of the OSHs is unknown. Preliminary studies in *Crithidia fasciculata* (Figure 1.6) indicate that cysteine is first coupled to intact histidine to form a *S*-4-(L-histidyl)-L-cysteine sulfoxide intermediate (**1-10**) via a sulfoxide synthase enzyme which requires iron (II) and oxygen. The cysteine is then cleaved by an enzyme requiring a pyridoxal phosphate cofactor to yield 4-thiol-L-histidine (**1-11**) and a side product which can be reduced with dithiothreitol to pyruvate and cysteine. The methylation of 4-thiol-L-histidine by *S*-adenosyl-L-methionine is catalyzed by a methyltransferase to form OSH<sub>A</sub>.<sup>85</sup>



**Figure 1.6:** The proposed biosynthetic pathway of OSH in *Crithidia fasciculata*, where **X** is an unknown product that can be reduced with dithiothreitol to yield pyruvate and cysteine.

### 1.2.3. Chemical Properties

Unlike ESH, the OSHs are unable to form a thione due to the location of the sulfur atom on the imidazole ring. The thiol group is instead expected to exist as a thiolate at physiological pH: the pKa of OSH<sub>A</sub> model compound 4-thiol-*N*<sup>1</sup>-methyl-5-methyl-imidazole (TMMI, Figure 1.7, **1-12**) is only 2.3 for the thiol moiety.<sup>86</sup> In contrast, GSH has a pKa between 8.7 and 9.2<sup>87,88</sup> for the thiol moiety and exists as a thiol under physiological conditions.



**Figure 1.7:** 4-Thiol-*N*<sup>1</sup>-methyl-5-methyl-imidazole is often used as a model to study the properties of the OSHs.

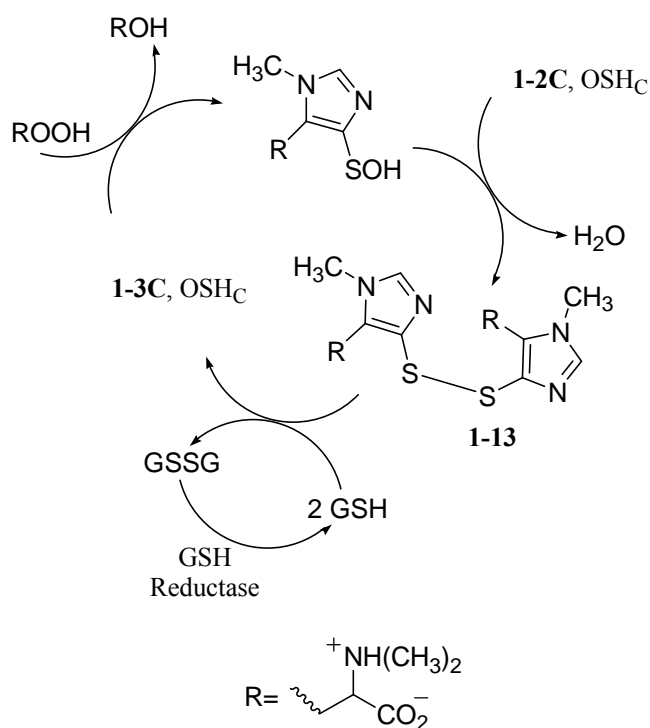
The presence of the thiolate may explain the superior nucleophilicity of TMMI as compared to GSH: iodoacetamide has been shown to react faster with TMMI than with GSH.<sup>86</sup> In addition, TMMI is a superior radical scavenger to GSH as it reacts faster than GSH with several radicals,<sup>86,89</sup> including tyrosyl radicals.<sup>89</sup> The increased reactivity of TMMI with radicals is likely related to the aromatic nature of TMMI, which allows for delocalization of the thiyl radical onto the imidazole ring, increasing the stability of the TMMI radical.<sup>89</sup> Semi-empirical calculations of the TMMI radical indicate that the free radical density is not localized on the sulfur atom, but is shared between the sulfur atom and carbon atoms 4 and 5 (see Figure 1.1 for numbering).<sup>71</sup>

### 1.2.4. Biological Function

The OSHs have been linked to a number of diverse activities in marine invertebrates. They can act as a pheromone in *P. dumerilii*,<sup>77</sup> as a component of the

adenochromines which sequester iron in adenochrome of *O. vulgaris*,<sup>79</sup> and as a component of imbricatine in *D. imbricata* which elicits a swimming response from *Stomphia coccinea*.<sup>84</sup> OSH<sub>C</sub> has been shown to act as an H<sub>2</sub>O<sub>2</sub> scavenger in sea urchin eggs.<sup>68</sup>

During fertilization, the sea urchin *S. purpuratus* uses H<sub>2</sub>O<sub>2</sub> to crosslink tyrosine residues in its egg's fertilization envelope.<sup>90,91</sup> The concentration of H<sub>2</sub>O<sub>2</sub> must be tightly controlled to protect the embryo from oxidative damage and this goal is accomplished through the non-enzymatic reaction of OSH<sub>C</sub> with H<sub>2</sub>O<sub>2</sub><sup>68</sup> to form OSSO<sub>C</sub> and water (Figure 1.8, **1-13**).<sup>68,92</sup> OSSO<sub>C</sub> is believed to be reduced back to its free thiol form by GSH, as there have been no reports of an OSSO reductase in any OSH containing organism. The OSH<sub>C</sub> system accounts for 73% of the metabolism of H<sub>2</sub>O<sub>2</sub> in unfertilized and fertilized sea urchin eggs, while catalase, an enzyme used to convert H<sub>2</sub>O<sub>2</sub> to water, accounts for only 27%.<sup>68</sup> OSH<sub>C</sub> may therefore play a critical role in the prevention of oxidative damage during fertilization.



**Figure 1.8:** The redox cycle for the detoxification by ovothiol C of peroxides formed during the fertilization of sea urchin eggs. The disulfide of ovothiol C (**1-13**) is formed, which is believed to be reduced by glutathione back to the free thiol form.

The role of OSH in the *Trypanosoma* parasites is not clear. These organisms utilize a trypanothione-based peroxidase system to scavenge H<sub>2</sub>O<sub>2</sub>, relegating the non-enzymatic scavenging of H<sub>2</sub>O<sub>2</sub> by OSH to a relatively minor role.<sup>80</sup> OSH<sub>A</sub> has been shown to decompose the *S*-nitroso groups of *S*-nitrosogluthathione and di-*S*-nitrosotrypanothione,<sup>93</sup> the intermediate products of nitric oxide detoxification. Reactive nitrogen species (RNS) such as nitric oxide are produced by the parasite's host organisms as a defence<sup>94</sup> and the protection of the parasite from RNS may prove to be an essential role for OSHs in these organisms.

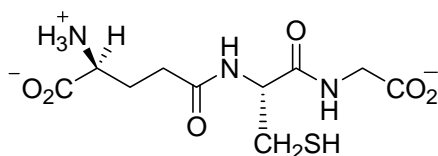
OSHs and their derivatives may be of therapeutic importance as GSH peroxidase mimics. GSH peroxidase catalyzes the reduction of peroxides by GSH, thereby protecting



the cell from oxidative damage.<sup>95</sup> Peroxides have been implicated in a number of disease states, including inflammation and cancer.<sup>53,96,97</sup> Recently, there has been intense interest in the development of small-molecule mimetics of GSH peroxidase, as these compounds could aid in the control of peroxide levels in the cell and possibly act as a treatment in prevention of oxidative damage.<sup>98-100</sup> The 4-thiol-imidazoles are of interest as therapeutic antioxidants, as they can reduce oxygen-centered radicals,<sup>89</sup> many of which have been implicated in a variety of disease processes.<sup>53,96</sup> Analogues of OSH have been successfully synthesized in hopes of finding pharmaceutically relevant antioxidants.<sup>69</sup> Some of these derivatives are potent scavengers of HOCl, HO•, and H<sub>2</sub>O<sub>2</sub> and are able to inhibit copper induced low density lipoprotein peroxidation.<sup>70</sup> Due to their apparent non-toxic properties, these compounds might contribute to the medicinal control of ROS.<sup>68</sup> Further research in this area may result in the development of useful therapeutic agents to defend against oxidative damage.

### 1.3. Glutathione

GSH (Figure 1.9, **1-1**) is the best characterized of the intracellular thiols. Much is known about GSH and its relationship to cellular reduction-oxidation (redox) processes, its activity as an enzyme cofactor, and its biosynthesis and degradation. Once considered ubiquitous, it has been determined that not all organisms produce GSH; the pathogenic *Mycobacteria* for example produce MSH in the place of GSH. The functions of MSH are believed to parallel that of GSH; therefore, GSH and its enzymes may serve as a template for the discovery of new MSH functionalities. Consequently, before discussing MSH, one of our thiols of interest, GSH will be briefly discussed to give context to the studies of MSH, and to highlight the similarity between these two thiols.



Glutathione (GSH, **1-1**)

**Figure 1.9:** Glutathione is composed of  $\gamma$ -L-glutamyl-L-cysteinyl-glycine.

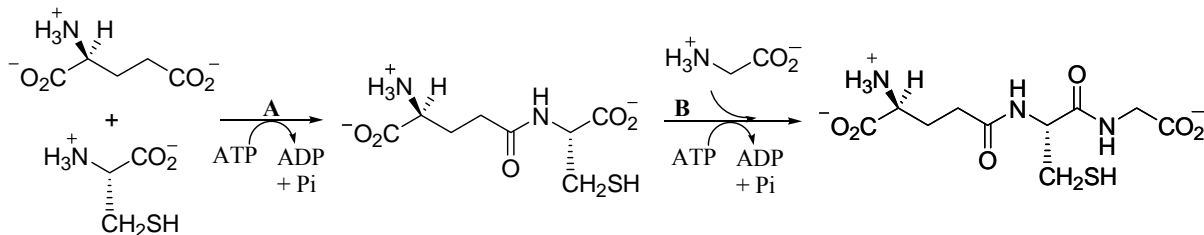
#### 1.3.1. Distribution

GSH is the predominant thiol in eukaryotic organisms including humans, animals, plants, and fungi.<sup>101</sup> All mammalian cells synthesize GSH<sup>102</sup> and it is maintained at an intracellular concentration between 0.5 to 10 mM.<sup>103</sup> Approximately 98% of GSH is found in its free thiol form. The remaining is present as the GSH disulfide (GSSG), mixed disulfides or as thioethers.<sup>102</sup>

Most Gram-negative bacteria,<sup>104</sup> including *Escherichia coli*, the cyanobacteria and the purple bacteria use GSH as their major thiol,<sup>101</sup> as do certain Gram-positive bacteria including *Streptococcus agalactiae* and *Streptococcus lactis*.<sup>104</sup> The pathogenic *Trypanosoma* and the *Leishmania* also produce GSH; however, it is mostly bound as trypanothione (*N*<sup>1</sup>-*N*<sup>8</sup>-bis(glutathionyl)spermidine), the major thiol in these organisms.<sup>105,106</sup>

### 1.3.2. Biosynthesis

The well characterized biosynthesis of GSH (Figure 1.10) begins with the coupling of L-glutamate with L-cysteine catalyzed by glutamate-cysteine ligase followed by the addition of glycine catalyzed by GSH synthase.<sup>107</sup>



**Figure 1.10:** The biosynthetic pathway of glutathione. A:  $\gamma$ -glutamyl-cysteine synthase; B: glutathione synthase

### 1.3.3. Chemical Properties

Unlike ESH, GSH has no structural features that lead to unusual bond lengths, angles or energies.<sup>87</sup> The pKa values for GSH have been determined (Table 1.1) and at physiological pH it is found as a thiol rather than a thiolate, with an overall charge of -1.<sup>87</sup> GSH is highly soluble in water as is its disulfide. Also, unlike ESH, GSSG is stable and found under physiological conditions.

**Table 1.1:** The pKa values of glutathione.<sup>87</sup>

pKa	Ionizable Group			
	SH	NH <sub>3</sub> <sup>+</sup>	CO <sub>2</sub> H	CO <sub>2</sub> H
	9.2	8.9-9.6	2.6	3.4

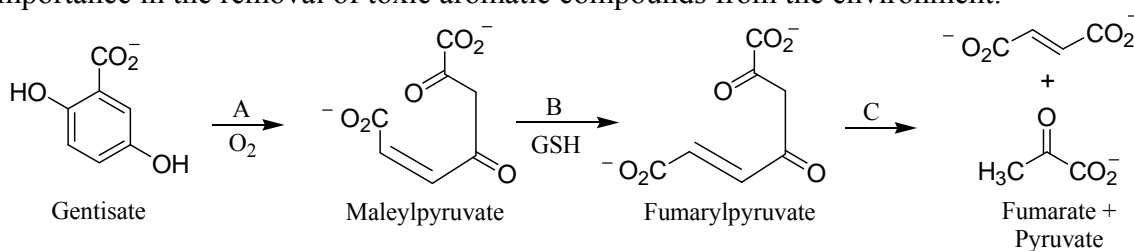
GSH can act as a nucleophile and a scavenger of free radicals. Once oxidized to the thiyl radical, GSH is not further oxidized to the thiol cation and often dimerizes to form the disulfide.<sup>87</sup> GSH is not as easily oxidized as cysteine, which is attributed to the blocking of the carboxyl and amino groups of cysteine by peptide bonds, which slows autooxidation in the presence of trace metals such as copper (II) and iron (II).<sup>108,109</sup>

### 1.3.4. Biological Function

Since its discovery in 1888,<sup>110</sup> GSH has been implicated in a wide variety of processes, including the metabolism of aromatic carbon sources, the maintenance of cellular redox potentials, and the removal of toxic substances from the cell. It is a co-factor for a number of vital enzymes including many enzymes essential in detoxification pathways. An encompassing description of the role that GSH plays in the cell is not feasible within the confines of this introduction<sup>†</sup> and therefore this discussion will be limited to the details which pertain to the work detailed herein. The reader is referred to the extensive review literature available for GSH.<sup>102,103,111-115</sup>

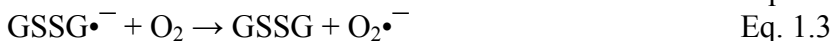
<sup>†</sup> A search of the PubMed database, a service of the National Library of Medicine and the National Institute of Health, for glutathione returned 73 334 references, 4 022 of which are review articles (accessed November 13, 2006).

The metabolism of aromatic compounds by microorganisms is important for the removal of toxic compounds from the environment.<sup>116</sup> Gentisate and its derivatives are key intermediates in the metabolism of a large number of aromatic compounds, including 3-hydroxybenzoate,<sup>117</sup> substituted phenols,<sup>118-120</sup> salicylate,<sup>116</sup> and naphthalene.<sup>121,122</sup> The cleavage of the aromatic ring of gentisate is catalyzed by gentisate 1,2-dioxygenase to form maleylpyruvate. In *Salmonella typhimurim*,<sup>117,123</sup> *Pseudomonas acidovorans*,<sup>124</sup> *Klebsiella pneumoniae*,<sup>125</sup> *Ralstonia* sp. Strain U2<sup>126</sup> and *Moraxella* OA3,<sup>127,128</sup> maleylpyruvate is converted to fumarylpyruvate by maleylpyruvate isomerase, using GSH as an essential cofactor (Figure 1.11),<sup>129</sup> followed by the hydrolysis of fumarylpyruvate to pyruvate and fumarate. Maleylpyruvate isomerase activity was first identified by Lack in 1959;<sup>129</sup> however, very little research has been reported regarding the further characterization of this enzyme despite the confirmation of this pathway in a number of organisms and its importance in the removal of toxic aromatic compounds from the environment.

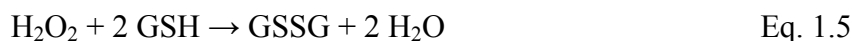


**Figure 1.11:** The gentisate pathway. A: gentisate dioxygenase; B: maleylpyruvate isomerase; C: fumarylpyruvate hydrolase.

The removal of ROS is essential to the viability of the cell as they can react with carbohydrates, DNA, lipids, and proteins, inactivating cellular components.<sup>62</sup> GSH can react with many ROS, detoxifying them, and in the process forming its disulfide, GSSG. The reaction of GSH with free radicals goes through a glutathionyl radical (GS•) which is less reactive than most ROS species, before ultimately forming GSSG and superoxide ( $\text{O}_2\cdot^-$ ) (Eqs. 1.1 to 1.4). The latter is removed by superoxide dismutase.<sup>62</sup>

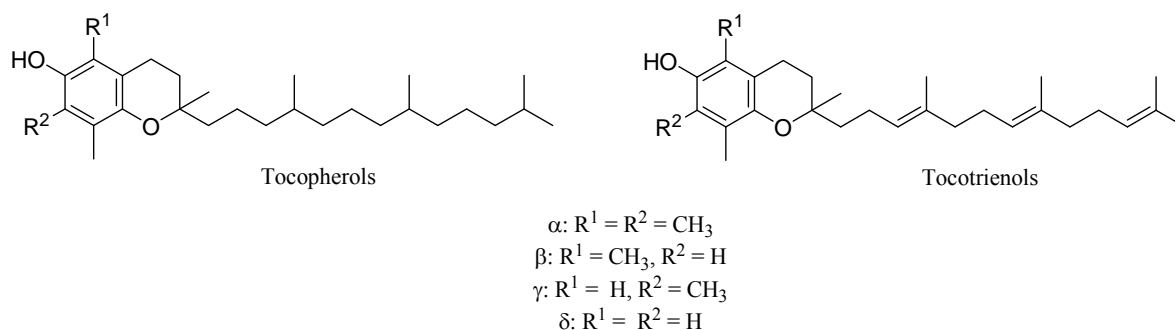


GSH will also react with  $\text{H}_2\text{O}_2$  to form GSSG.  $\text{H}_2\text{O}_2$  is less reactive than free radicals such as  $\text{HO}\cdot$ ; however, it is known to inactivate thiol-containing enzymes and can generate  $\text{HO}\cdot$  upon interaction with iron (II) and copper (I).<sup>62</sup> The intracellular concentration of  $\text{H}_2\text{O}_2$  is therefore tightly controlled by several peroxidases,<sup>62</sup> including GSH peroxidase (Eq. 1.5)<sup>130,131</sup> and catalase (Eq. 1.6).<sup>62</sup>

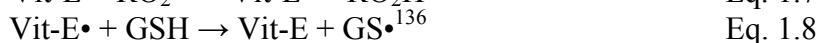
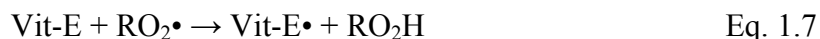


In addition, GSH has been implicated in the repair of vitamin E (Figure 1.12),<sup>62</sup> an essential lipid soluble antioxidant responsible for the protection of the cellular membrane against ROS.<sup>132-134</sup> Cell membranes contain polyunsaturated fatty acids, which are readily

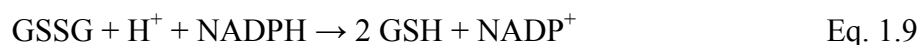
auto-oxidized, and if left unchecked, can cause a loss in the structural integrity of the membrane.<sup>134</sup> Vitamin E reacts with lipid peroxy radicals, preventing lipid peroxidation<sup>62</sup> by breaking the free radical chain reaction (Eq. 1.7).<sup>134</sup> The resulting fatty acyl hydroperoxide is presumed to be reduced to a hydroxyl acid by GSH peroxidase<sup>135</sup> and the vitamin E radical must then be reduced back to its active form for continued protection of the membrane. The membrane-bound vitamin E radical can react with cytosolic antioxidants, such as GSH, to regenerate vitamin E (Eq. 1.8).<sup>62</sup> As detailed above, generation of GS• ultimately results in GSSG (Eq. 1.2 and 1.4).



**Figure 1.12:** The naturally occurring forms of vitamin E.



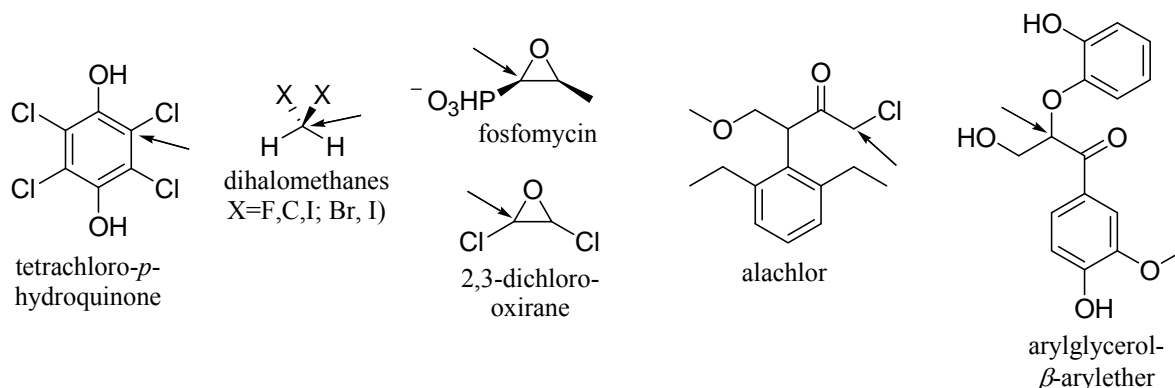
Once formed by these reactions, GSSG must be reduced back to its biologically useful free thiol form.<sup>102,137</sup> GSSG reductase (GR), which reduces GSSG to GSH, has been identified in bacteria, fungi, plants, protozoa and animals.<sup>138</sup> GR is a dimer of identical subunits<sup>138</sup> and belongs to the family of flavin adenine dinucleotide (FAD)-containing pyridine nucleotide:disulfide oxidoreductases.<sup>137</sup> GR reduces GSSG to GSH using one equivalent of nicotinamide adenine dinucleotide phosphate (NADPH) (Eq. 1.9).



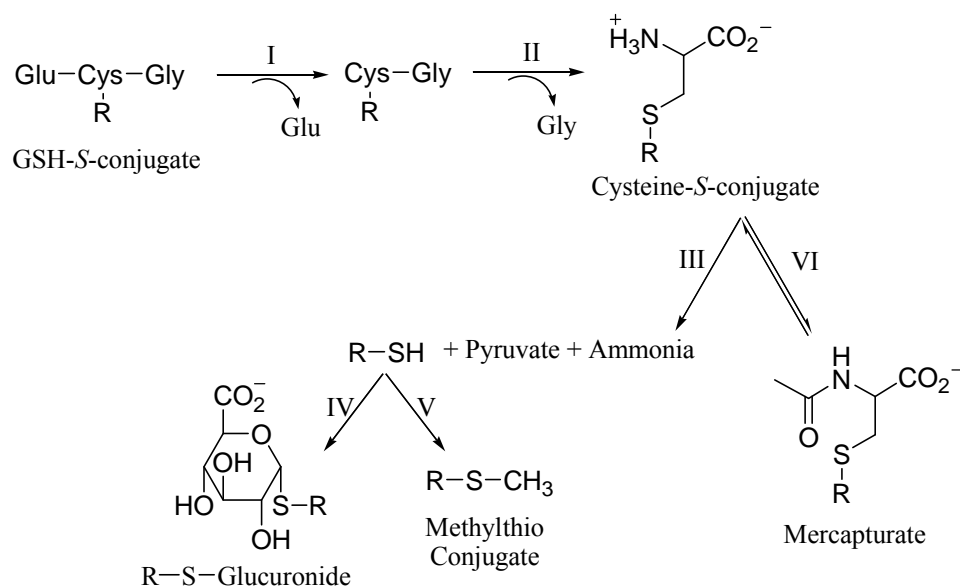
The enzyme binds NADPH, which reduces the FAD and NADP<sup>+</sup> is released. The active site contains a redox active disulfide, which is reduced by the reduced FAD and GSSG binds. The newly formed active site sulfur attacks GSSG, forming a mixed disulfide and one GSH molecule is released. The other active site sulfur attacks the mixed disulfide, reforming the active site disulfide and releasing the second GSH molecule.<sup>139</sup> Without this enzyme, the concentration of biologically active GSH would be depleted, hampering the organism's ability to maintain redox homeostasis and protect the cell from ROS.

Cells must also be protected from electrophilic species which can modify nucleophilic cellular components such as DNA and proteins. These compounds can often be linked to GSH to form GSH-S-conjugates which are further metabolized by the cell for excretion; these conjugation reactions are catalyzed by GSH-S-transferases (GSTs),<sup>140-144</sup> whose active site environment lowers the pK<sub>a</sub> of GSH to ~6.2, which increases the thiol group nucleophilicity and facilitates the reaction.<sup>142</sup> The most common GST reactions involve attack on an electrophilic carbon atom,<sup>102</sup> such as those shown in Figure 1.13. Once a GSH-S-conjugate is formed, it is further metabolized with the exact pathway and

ultimate mode of excretion varying with species and xenobiotic (an example pathway is shown in Figure 1.14).<sup>103</sup>



**Figure 1.13:** Examples of known physiological substrates of glutathione-*S*-transferases. The arrows indicate the electrophilic carbon atoms where GSH addition occurs. Adapted from Vuilleumier (1997).<sup>145</sup>

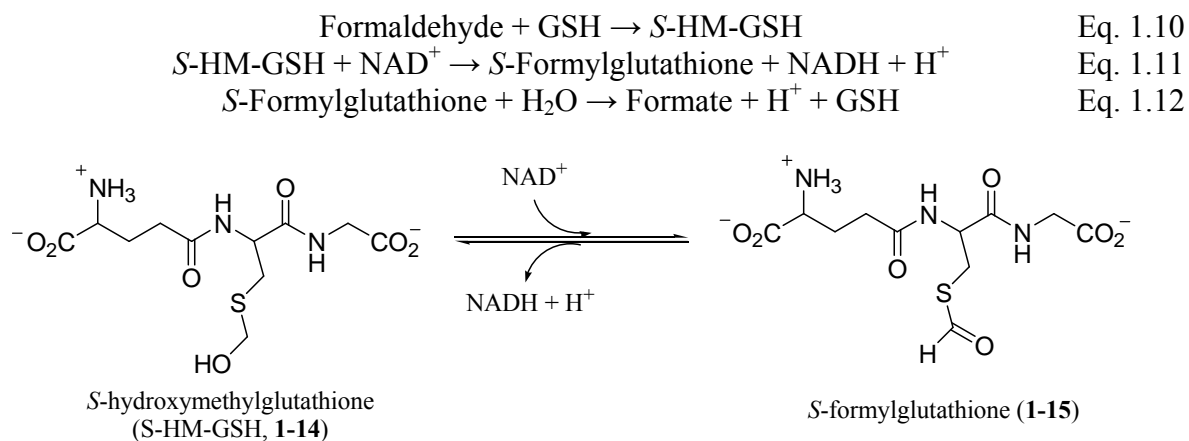


**Figure 1.14:** Selected pathways for the metabolism of glutathione-*S*-conjugates. I:  $\gamma$ -glutamyl transpeptidase; II: cysteinyl glycine dipeptidase; III: cysteine conjugate  $\beta$ -lyase; IV: UDP-glucuronyl transferase; V: *S*-methyltransferase; VI: *N*-acetyltransferase catalyzes the forward reaction while a deacetylase catalyzes the reverse reaction. Adapted from Pickett and Lu (1989).<sup>141</sup>

GSTs have been noted in mammals, certain insects, plants, squid and bacteria.<sup>142</sup> These enzymes can be divided into multiple classes<sup>140</sup> and each organism can have more than one isoform from each class. All known GSTs are dimeric with one active site per monomer.<sup>142</sup> The dimers need not be homodimeric but the monomers must be of the same class, as the features of subunit recognition differ between classes.<sup>144</sup> In addition to the removal of xenobiotics, GSTs are also involved in the formation of biological mediators and a number of physiological substrates have been identified.<sup>146,147</sup>

Cells must also be protected from toxic electrophilic substances produced as by-products of cellular processes. Formaldehyde for example, is produced during a number of

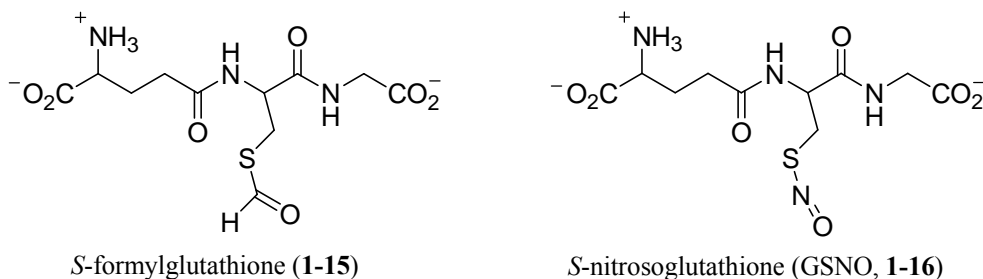
metabolic and detoxification processes,<sup>148-151</sup> but it must be removed due to its non-specific reactivity with proteins and nucleic acids. One of the most widespread methods of formaldehyde detoxification involves the oxidation of formaldehyde to formate via a two enzyme system. The first enzyme in this pathway, formaldehyde dehydrogenase (FD), has been identified in all animal tissues, pea and corn seeds, Baker's yeast, *E. coli*, all methanol-utilizing yeasts, and many methanol-utilizing methylotrophic bacteria. FD has been shown to be GSH dependent (GD) in these organisms<sup>152</sup> with the true substrate for GD-FD being the non-enzymatically formed hemithioacetal adduct of formaldehyde and GSH, *S*-hydroxymethylglutathione (*S*-HM-GSH Eq. 1.10, Figure 1.15, **1-14**). *S*-HM-GSH is reversibly converted to *S*-formylglutathione (Eq. 1.11, Figure 1.15, **1-15**), whose hydrolysis is catalyzed by *S*-formylglutathione hydrolase to formate and GSH (Eq. 1.12).



**Figure 1.15:** The reaction catalyzed by glutathione dependent formaldehyde dehydrogenase.

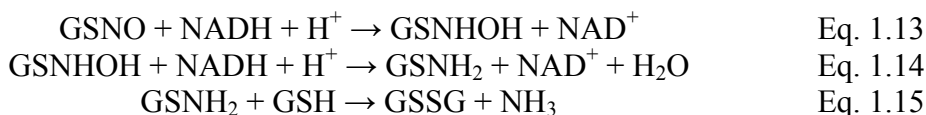
GD-FD is highly specific for GSH and will not accept oxidized GSH, or other thiols such as cysteinylglycine, ESH, cysteine, or *N*-acetylcysteine.<sup>153</sup> The specificity for the cofactor is less rigorous and GD-FD can accept both NAD<sup>+</sup> and NADP<sup>+</sup> as cofactors.<sup>153</sup> Based upon structural studies, GD-FD belongs to the family of alcohol dehydrogenases and is homologous to the class III alcohol dehydrogenases.<sup>154</sup> In agreement with this observation, GD-FD has been shown to catalyze the oxidation of long chain aliphatic alcohols independent of GSH; however, this activity is low at physiological pH.<sup>154</sup>

Unexpectedly, GD-FD has also been implicated in the metabolism of *S*-nitrosoglutathione (GSNO, Figure 1.16, **1-16**).<sup>155,156</sup> GSNO is formed by the nitrosation of GSH by nitric oxide (NO) in the presence of oxygen. NO performs a number of essential roles in the cell, including the control of blood pressure and the destruction of pathogens.<sup>62</sup> However, it can also damage metalloproteins and DNA,<sup>62</sup> and can react with O<sub>2</sub><sup>•-</sup> to form ONOO<sup>-</sup>, which can attack many biologically important molecules.<sup>62</sup> The formation of GSNO has been suggested to be a mechanism for the storage and transport of NO, as well as a protective method against NO toxicity. The structural similarity between GSNO and *S*-formylglutathione is visibly apparent (Figure 1.16), which may account for this activity of GD-FD.



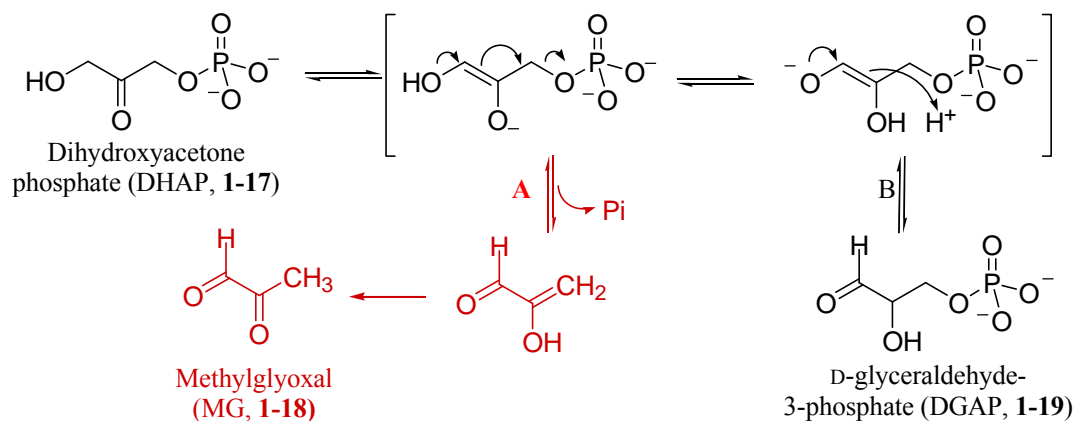
**Figure 1.16:** Structural comparison of the two substrates of glutathione dependent formaldehyde dehydrogenase.

The main products of the GSNO metabolism by *E. coli* GD-FD, ammonia (NH<sub>3</sub>) and GSSG, are consistent with Eqs. 1.13-1.15.<sup>156</sup>



GD-FD has been identified in *E. coli*,<sup>156</sup> mammals,<sup>155-158</sup> *Arabidopsis thaliana*,<sup>157</sup> and yeasts.<sup>156,157</sup> It has been determined that GD-FD is essential for cellular protection against nitrosative stress in yeast and it regulates and protects against the *S*-nitrosylation of proteins.<sup>156,158</sup>

Formaldehyde is not the only toxic by-product produced during metabolism. Methylglyoxal (MG, Figure 1.17, **1-18**) is produced by a number of pathways and is known to modify proteins and nucleic acids, resulting in cell death if its concentration is not controlled. A major source of MG is its spontaneous production by triose phosphate isomerase (TIM) during the conversion of dihydroxyacetone phosphate (DHAP, Figure 1.17, **1-17**) to D-glyceraldehyde-3-phosphate (DGAP, Figure 1.17, **1-19**) during glycolysis.<sup>159</sup> MG is formed when phosphate elimination occurs rather than reprotonation.<sup>160</sup> While reprotonation occurs 10<sup>6</sup> times more frequently than phosphate release, the high concentration of TIM in the cells allows for a significant accumulation of MG in the cells: the rate of MG accumulation due to TIM has been postulated to be as high as 0.4 mM/day.<sup>159,160</sup>

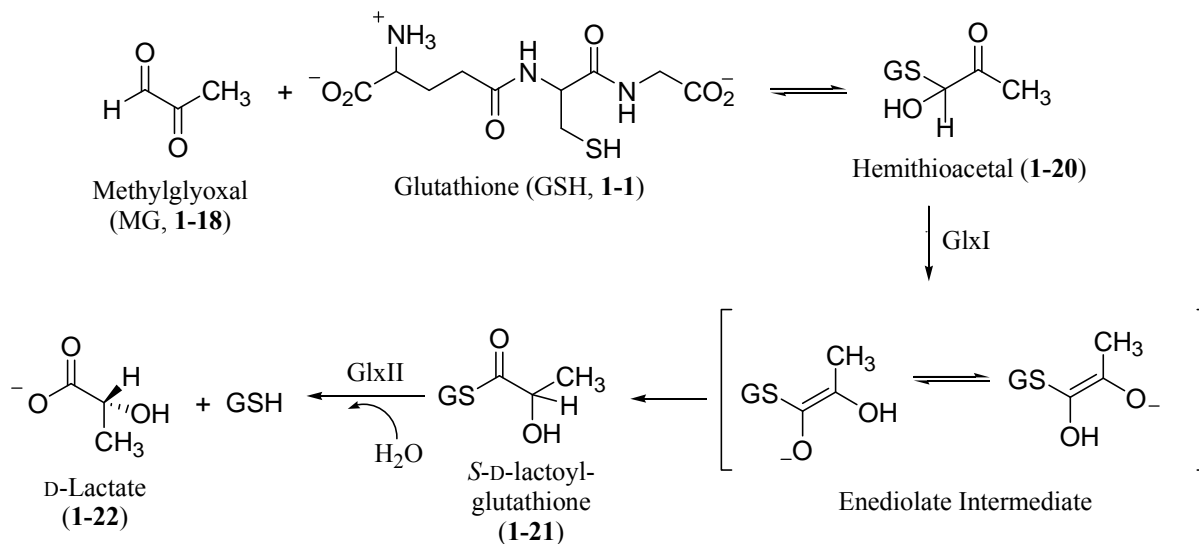


**Figure 1.17:** The isomerization catalyzed by triose phosphate isomerase: dihydroxyacetone phosphate (DHAP, **1-17**) is converted to D-glyceraldehyde-3-phosphate (DGAP, **1-19**) via a deprotonation and reprotonation mechanism. When inorganic phosphate (Pi) is eliminated (pathway A) methylglyoxal (MG, **1-18**) is formed.<sup>160</sup>

Microorganisms can also produce MG directly from DHAP by MG synthase in a postulated bypass to phosphorylating glycolysis.<sup>161,162</sup> DHAP can accumulate when cells have limited sources of phosphate and MG synthase, combined with other enzymes in the bypass, this may serve as a pathway for the formation of pyruvate, with the release of phosphate. This enzyme is tightly controlled through feedback inhibition by phosphate and activation by DHAP.<sup>163</sup> The concentration of MG must be tightly controlled by the cell as it, and other  $\alpha$ -ketoaldehydes, can form adducts with RNA and proteins and can inhibit protein biosynthesis.<sup>164</sup>

Many organisms control the intracellular concentration of MG by converting it to D-lactate (Figure 1.18, **1-22**) using the two-enzyme glyoxalase (Glx) system composed of GlxI and GlxII (Figure 1.18). The Glx system requires GSH and is believed to be present in all GSH-utilizing organisms.<sup>165</sup> GlxI sequences have been identified in humans,<sup>166</sup> mice,<sup>167</sup> yeasts,<sup>168,169</sup> plants,<sup>170-177</sup> insects,<sup>165</sup> protozoa,<sup>178</sup> fungi<sup>179</sup> and at least 42 species of bacteria,<sup>165,180</sup> including *E. coli*.<sup>181</sup> The true substrate for GlxI is the non-enzymatically formed hemithioacetal (Figure 1.18, **1-20**) of MG and GSH. The product of the GlxI catalyzed reaction, *S*-D-lactoylglutathione (Figure 1.18, **1-21**), is the substrate for GlxII, which hydrolyzes the thioester to GSH to form D-lactate (**1-22**). The mechanism is believed to proceed via the deprotonation of **1-20** to form the endiolate intermediate, followed by reprotonation to yield *S*-D-lactoylglutathione (**1-21**).<sup>182</sup> GlxI will accept either of the two possible stereoisomers of the hemithioacetal as substrates,<sup>183</sup> and will accept hemithioacetals formed from the condensation of GSH with other  $\alpha$ -ketoaldehydes, thereby protecting the cell against modification by  $\alpha$ -ketoaldehydes other than MG.





**Figure 1.18:** The glyoxalase system: a two enzyme system which converts cytotoxic methylglyoxal (MG, 1-17) to D-lactate (1-19) using glutathione (GSH, 1-1) as an essential cofactor.

The human, bacterial, and plant GlxI enzymes are known to be dimeric, while the yeast enzyme has been shown to be a monomer<sup>165</sup> containing two copies of a segment which is equivalent to two copies of the human enzyme.<sup>165</sup> Crystal structures have been solved for the human<sup>184</sup> and *E. coli*<sup>181</sup> GlxI enzymes and both structures indicate that there are two active sites, each of which is located at the interface of the dimers and contains a metal ion which interacts with amino acid residues from both subunits. The human enzyme has been shown to be most active with zinc (II)<sup>185</sup> which is coordinated by glutamine-33A,<sup>‡</sup> glutamate-99A, histidine-126B, glutamate-172B and two water molecules.<sup>184</sup> In contrast, the *E. coli* enzyme is maximally activated by nickel (II)<sup>186</sup> which is bound by histidine-5A,<sup>‡</sup> glutamate-56A, histidine-74B, glutamate-122B and two water molecules.<sup>181</sup> Both the human<sup>166</sup> and *E. coli*<sup>181</sup> enzymes have an octahedral coordination state for their preferred metal ion. When *E. coli* binds zinc (II), the coordination sphere is trigonal bipyramidal, which may account for the difference in activity.<sup>181</sup>

While GlxI enzymes have been thoroughly characterized from a number of organisms, much less is known regarding GlxII. GlxII has been identified in the tissues of humans,<sup>187,188</sup> rat liver,<sup>189,190</sup> yeasts (*Saccharomyces cerevisiae*,<sup>191</sup> and *Hansenula markii*<sup>192</sup>), a limited number of plants,<sup>193-196</sup> and *Candida albicans*.<sup>197</sup> Also, the cloning of GlxII from humans,<sup>198</sup> *A. thaliana*<sup>199</sup>, *S. cerevisiae*<sup>200</sup>, and *Neisseria meningitidis*<sup>201</sup> has been reported. While S-D-lactoylglutathione is the major substrate for this enzyme, it will accept other S-D-lactoylthiols, as well as non-GSH thioesters, though it has a high specificity for GSH.<sup>202</sup> Interestingly, while GlxI has only been identified in the cytosol, the presence of GlxII in both the cytosol and mitochondria<sup>189,190,195,196,203</sup> of various organisms has been reported. The definitive role of mitochondrial GlxII has yet to be identified; however, S-D-lactoylglutathione of unknown origin has been isolated from mitochondria. Multiple isozymes of GlxII have been reported in *A. thaliana*,<sup>203</sup> *S. cerevisiae*<sup>200,204</sup> and humans,<sup>205</sup> which differ in localization within the cell: mitochondrial or cytosolic.

<sup>‡</sup> Where A and B indicate different subunits.

In all cases, GlxII is a monomeric enzyme with no sequence similarity to GlxI. GlxII requires metal ions for activity<sup>206</sup> and the crystal structure of the human enzyme indicates that the enzyme has two metal binding sites.<sup>207</sup> Initial evidence indicated that GlxII is a zinc metalloenzyme; however, further studies have indicated that GlxII from *A.thaliana* demonstrates flexibility in its metal binding and can incorporate iron, manganese and zinc *in vivo*.<sup>208</sup> A variety of metal ion combinations were detected observed in the active site including iron (III)<sup>§</sup>-iron (II), iron (III)-zinc (II) and manganese (II)-manganese (II) centers, and similar catalytic efficiencies ( $k_{cat}/K_M$ ) were observed with the different ions.<sup>208</sup> The crystal structure indicates that the two metals are bridged by a water or hydroxide molecule and an oxygen from an aspartate residue.<sup>207</sup> The first ion is also coordinated by the nitrogen atoms of histidine-54, histidine-56, and histidine-110, while the second ion is coordinated by an oxygen atom from aspartate-58 and nitrogen atoms from histidine-59 and histidine-173.<sup>207</sup> All of the coordinating residues are stabilized by direct or indirect hydrogen bonding interactions with other residues in a second coordination sphere.<sup>207</sup> The bound water/hydroxide molecule is suspected to attack the carbonyl of the substrate to form a tetrahedral intermediate that collapses to release GSH and the corresponding carboxylic acid.<sup>207</sup>

A GSH-independent Glx system has been reported in *E. coli* which accepts MG as a substrate and yields D-lactate as a product. This monomeric enzyme, termed glyoxalase III (GlxIII), does not catalyze the formation or breakdown of the GlxI product *S*-D-lactoylglutathione. GlxIII was found to have a greater substrate specificity than the Glx system as it will only accept MG and phenylglyoxal as substrates, activity with phenylglyoxal being only 15 % of that seen with MG.<sup>209</sup> To date, this enzyme has not been further characterized.

As mentioned, GSH is the best characterized of the intracellular thiols. It is involved in a diverse range of cellular activities, only a small sampling of which was presented here. Not all organisms use GSH, and instead use other low molecular weight thiols in its place. It may be true that what is known about GSH can act as a guide for the discovery of enzymes that use these other thiols.

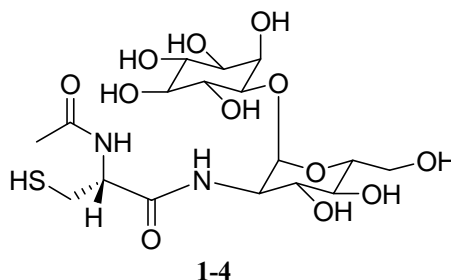
#### **1.4. Mycothiol**

The *Actinomycetales* bacteria, such as the pathogenic *Mycobacteria*, do not produce GSH and instead use a unique thiol termed MSH (1-D-*myo*-inosityl-2-(*N*-acetyl-L-cysteinyl)amido-2-deoxy- $\alpha$ -D-glucopyranoside, Figure 1.19, **1-4**). The most well known disease associated with *Mycobacteria* is tuberculosis (TB), caused by *M. tuberculosis* in humans<sup>210</sup> and *Mycobacterium bovis* in bovines.<sup>211</sup> The emergence of antibiotic resistant strains of *M. tuberculosis* has brought this disease back to the forefront of clinical research: strains resistant to single drugs have been reported in every country surveyed and strains that are resistant to all of the major TB drugs have been documented.<sup>210</sup> According to the World Health Organization (WHO), a third of the world's population is currently infected with TB.<sup>210</sup> *Mycobacteria* are linked to a number of other clinically and agriculturally important diseases including human leprosy and Johne's disease which affects cattle and causes losses to the dairy industry in excess of \$1.5 billion (US) per year.<sup>212</sup> Despite the medical importance of the organisms that utilize MSH, very little is known concerning its

---

<sup>§</sup> The presence of a ferric ion is consistent with the oxidative experimental condition and likely does not reflect the *in vivo* oxidation state (Schilling *et al.* 2003).

behaviour and reactions. The role of MSH appears to parallel that of GSH, indicating that the function of GSH and its enzymes may serve as a template for the discovery of new MSH functionalities. MSH and its enzymes act as a novel target for the development of novel therapeutics for use against the pathogenic *Mycobacteria*.



**Figure 1.19:** Mycothiol.

#### 1.4.1. Distribution

MSH was first identified as a disulfide (MSSM) in *Streptomyces* sp. AJ9463<sup>213</sup> and as a free thiol<sup>214</sup> in *M. bovis*<sup>72</sup> and *Streptomyces clavuligerus*.<sup>215</sup> Further studies identified MSH in species of *Micrococcus*, *Nocardia*, *Rhodococcus*, *Mycobacterium*, *Micromonospora*, *Actinomadura* and *Nocardiosis*,<sup>216</sup> but not in representatives of the *Arthrobacter*, *Agromyces* or *Actinomyces* genera (Table 1.2).<sup>216</sup> MSH was not detected in a selection of Gram-positive bacteria including representative species of *Listeria*, *Staphylococcus*, *Enterococcus*, *Bacillus* and *Clostridium*.<sup>216</sup> MSH was also absent from *E. coli* and *Pseudomonas aeruginosa*.<sup>216</sup> Tests of rat liver, bonito muscle, *Candida albicans*, *Neurospora crassa*, and spinach leaves were also negative.<sup>216</sup> Based upon these results it was assumed that MSH production is restricted to the *Actinomycetales* bacteria and that it is the major thiol in most *Actinomycetales* bacteria.<sup>101,216</sup>

#### 1.4.2. Biosynthesis

The steps in the biosynthetic pathway of MSH in *Mycobacteria* have recently been fully elucidated (Figure 1.20). The pathway begins with the coupling of D-*myo*-inositol-3-phosphate (Ins(3)P<sub>1</sub>, (1-23)) and UDP-*N*-acetylglucosamine (UDP-GlcNAc, (1-24))<sup>217,218</sup> to form 3-phospho-1-D-*myo*-inosityl-2-acetamido-2-deoxy- $\alpha$ -D-glucopyranoside (GlcNAc-Ins(3)P<sub>1</sub> (1-25)) which is catalyzed by a glycosyltransferase termed MshA.<sup>217</sup> GlcNAc-Ins(3)P<sub>1</sub> is then dephosphorylated by an inositol kinase MshA2 to yield GlcNAc-Ins (1-26).<sup>217</sup> GlcNAc-Ins is deacetylated by MshB to yield GlcN-Ins (1-27).<sup>219</sup> The coupling of GlcN-Ins to cysteine is catalyzed by the ligase MshC which activates cysteine using ATP to form a cysteine-AMP derivative. This derivative is attacked by the amino group of GlcN-Ins to form Cys-GlcN-Ins (1-28); AMP and pyrophosphate are also released.<sup>220</sup> The final step of MSH biosynthesis is the acylation of the cysteinyl amino group of Cys-GlcN-Ins by acetylCoA, catalyzed by MshD.<sup>221</sup>

**Table 1.2:** Thiols produced by *Actinomycetales*.<sup>a</sup>

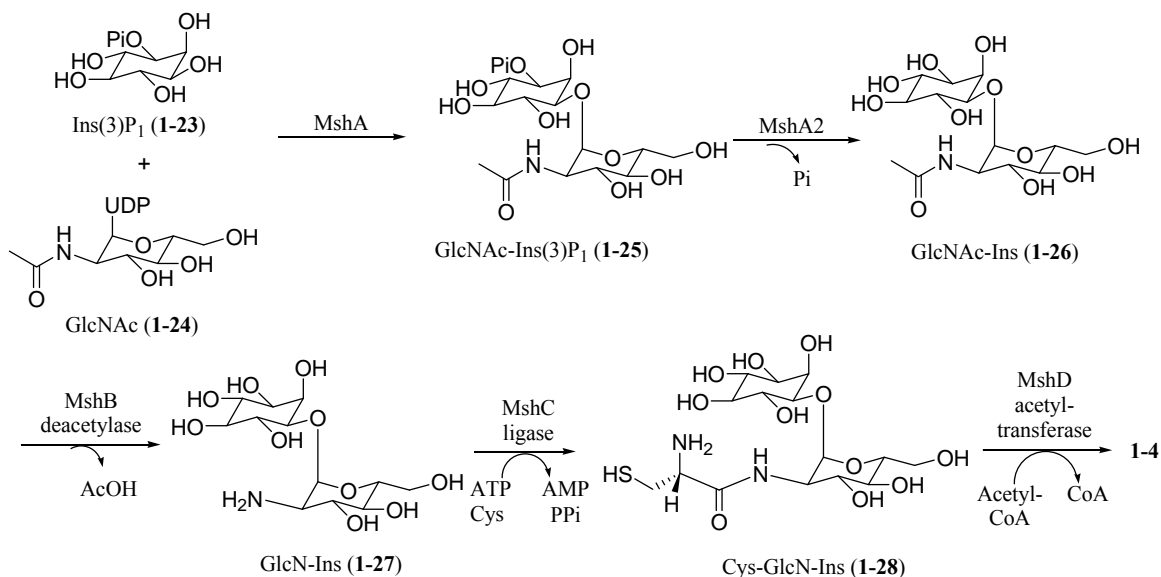
Strains (sorted by subclass)	Thiol( $\mu\text{mol/g}$ [residual dry weight])					
	MSH	Cys	GSH	H <sub>2</sub> S	CoA	Other
<b>Actinomycineae</b>						
<i>Actinomyces israelii</i>	<0.02	0.22	<0.02	0.04	0.17	
<b>Corynebacterineae</b>						
<i>Corynebacterium diphtheriae</i>	0.18	0.015	<0.01	0.04	0.19	0.14 <sup>b</sup>
<i>Mycobacterium avium</i> <sup>c</sup>	2.70	<0.02	<0.02	1.8	0.7	
	4.60	<0.03	<0.02	1.8	1.0	
<i>Mycobacterium chelonae</i>	6.60	<0.05	<0.02	0.8	5	
<i>Mycobacterium fortuitum</i>	9.40	<0.03	<0.02	0.9	3	
<i>Mycobacterium smegmatis</i>	19.00	0.09	<0.02	0.9	3	
<i>Mycobacterium tuberculosis</i> <sup>d</sup>	12.00	<0.02	<0.02	1.1	3	
	9.00	<0.02	<0.02	0.66	1.0	
<i>Nocardia asteroides</i>	3.50	<0.06	<0.05	0.35	0.66	0.01 <sup>b</sup>
<i>Nocardia brasiliensis</i>	2.80	<0.03	<0.03	0.47	1.0	0.03 <sup>c</sup>
<i>Rhodococcus roseus</i>	3.00	0.2	<0.03	0.3	0.15	
<b>Micrococcineae</b>						
<i>Arthrobacter globiformis</i>	<0.01	0.11	<0.01	0.4	0.34	0.28 <sup>b</sup>
<i>Micrococcus agilis</i>	0.90	<0.01	<0.01	0.4	0.47	
<i>Micrococcus luteus</i> <sup>f</sup>	3.30	0.11	<0.01	0.5	1.1	
	3.60	<0.02	<0.02	0.3	1.1	0.48 <sup>g</sup>
<i>Micrococcus kristinae</i>	0.60	<0.01	<0.01	0.10	0.50	
<i>Agromyces ramosus</i>	<0.03	<0.04	<0.03	0.15	0.38	
<b>Micromonosporineae</b>						
<i>Micromonospora carbonacea</i>	1.50	0.20	<0.01	3.8	0.3	2.4 <sup>h</sup>
<i>Micromonospora floridensis</i>	14.00	0.90	<0.01	2.5	0.91	4.2 <sup>h</sup>
<i>Micromonospora fulvopurpurea</i>	2.40	0.60	<0.01	2.3	0.58	1.9 <sup>h</sup>
<b>Propionibacterineae</b>						
<i>Propionibacterium acnes</i>	<0.03	0.08	<0.02	0.28	0.4	
<b>Pseudonocardineae</b>						
<i>Nocardiopsis flava</i>	2.70	1.10	0.22	0.74	0.39	0.69 <sup>i</sup>
<i>Nocardiopsis mutabilis</i>	0.44	0.06	<0.02	0.29	0.9	0.85 <sup>j</sup>
<b>Streptomycineae</b>						
<i>Streptomyces clavuligerus</i>	5.40	0.78	<0.01	1.3	0.8	1.55 <sup>k</sup>
<i>Streptomyces coelicolor A3(2)</i>	2.80	0.56	<0.02	0.75	0.5	<0.02 <sup>l</sup>
<i>Streptomyces jumonjinensis</i>	6.70	1.30	<0.09	0.83	1.9	1.1 <sup>k</sup>
<i>Streptomyces lactamdurans</i>	5.90	1.30	<0.08	1.3	1.1	23.05 <sup>m</sup>
<i>Streptomyces lividans</i>	6.60	0.58	<0.06	1.3	1.0	0.51 <sup>k</sup>
<b>Streptosporangineae</b>						
<i>Actinomadura hibisca</i>	1.50	0.43	<0.01	1.0	0.17	2.41 <sup>n</sup>

<sup>a</sup>MSH: mycothiol, Cys: cysteine, GSH: glutathione, H<sub>2</sub>S: hydrogen sulfide, CoA: coenzyme A, Adapted from Newton *et al.* (1996).<sup>216</sup> <sup>b</sup>Unknown thiol. <sup>c</sup>Two strains of *M. avium* were tested NJH9151 and NJH1854-4.

<sup>d</sup>Two strains of *M. tuberculosis* were tested: ATCC 25618 and PZ<sup>A</sup> UCSD 100. <sup>e</sup>4'-phosphopantetheine.

<sup>f</sup>Two strains of *M. luteus* were tested: UCSD 22 and ATCC 4698. <sup>g</sup>*N*-acetylcysteine; <sup>h</sup>thiosulfate and *N*-acetylthiosulfate. <sup>i</sup>*N*-acetylthiosulfate, thiosulfate and an unknown thiol. <sup>j</sup>unknown thiols. <sup>k</sup> $\delta$ -(L- $\alpha$ -amino-adipyl)-L-cysteinyl-D-valine (ACV) and an unknown thiol. <sup>l</sup>ACV. <sup>m</sup>ACV and ergothioneine.

<sup>n</sup>*N*-acetylcysteine and an unknown thiol.



**Figure 1.20:** The biosynthetic pathway of mycothiol.

Studies of *M. tuberculosis* and *Mycobacterium smegmatis* mutants deficient in the various biosynthetic enzymes have been performed. Inactivation of the gene for MshB does not fully block the biosynthesis of MSH as an alternative deacetylase activity exists in these organisms which produces a modest amount of GlcN-Ins and ultimately a limited amount of MSH.<sup>222,223</sup> Mutants of *M. tuberculosis* and *M. smegmatis* deficient in MshD were able to produce low levels of MSH as well as another thiol, *N*-formyl-Cys-GlcN-Ins, which can act as a poor substitute for MSH in some biochemical reactions.<sup>224</sup> The removal of MshA<sup>218</sup> or MshC<sup>225</sup> results in a complete halt in the biosynthesis of MSH in *M. smegmatis*, and in the case of *M. tuberculosis*, no cellular growth.<sup>217</sup> The biosynthetic enzymes involved in the production of MSH are therefore excellent targets for the development of antituberculin and the development of inhibitors to these enzymes is ongoing.<sup>226</sup>

Homologues of the *M. tuberculosis* MshC sequence have been identified in *Streptomyces coelicolor* and *Corynebacterium striatum* and orthologues of MshC were found in *M. bovis*, *Mycobacterium leprae*, the causative agent of leprosy, *Corynebacterium diphtheriae*,<sup>220</sup> the causative agent of diphtheria, and in *Mycobacterium avium*.<sup>220</sup> Homology studies of the *S. coelicolor* genome have identified the *msha*, *b*, *c* and *d* genes and the identity of their gene products were proven experimentally.<sup>227</sup> These genes have also been identified in *Corynebacterium glutamicum*, *Corynebacterium efficiens*, and *Corynebacterium jeikeium*.<sup>228</sup> It is therefore believed that the biosynthetic pathway of MSH is conserved throughout all MSH biosynthesizing organisms.

#### 1.4.3. Chemical Properties

Despite the intense interest in MSH as a possible target for antimycobacterial agents, there is little data available regarding the chemical properties of MSH. Most of the research published is related to the elucidation of the biosynthetic pathway of MSH and the identification of MSH-utilizing enzymes as possible targets for the development of novel pharmaceuticals. It is known that MSH will react with monobromobimanes,<sup>215,216</sup>

electrophiles often used to tag intracellular thiols<sup>229</sup> in cell extracts for purification and identification.<sup>10</sup> The pKa of MSH has been estimated to be close to that of cysteine (pKa = 8.33) and GSH (pKa = 9.2), which would suggest that the nucleophilicity of MSH would decrease at physiological pH<sup>72</sup> and would be similar to that of GSH. GSH does not react with most electrophiles under normal physiological conditions and conjugation reactions proceed with the assistance of GSTs.<sup>140-144</sup> It is therefore likely that MSH does have some nucleophilic character; however, at physiological pH, it is not enough to react unassisted with most electrophiles.

Thiols are known to auto-oxidize in the presence of oxygen and heavy metals such as copper. GSH is known to auto-oxidize much slower than cysteine which is attributed to the blockage of the amino and carboxyl groups of cysteine by amide bonds with glutamate and glycine. MSH, in turn has been shown to auto-oxidize at a rate seven-fold slower than that of GSH and 30-fold slower than that of cysteine.<sup>215</sup> This decreased rate of oxidation cannot be accounted for by the blockage of the amino and carboxyl groups of cysteine alone. It is suspected that the GlcN-Ins scaffold may be an important factor in the resistance of MSH to autooxidation.<sup>215</sup>

#### 1.4.4. Biological Function

Current research indicates that the function of MSH in *Actinomycetales* bacteria parallels that of GSH in other organisms. MSH is produced in a similar pattern to GSH during cellular growth and the MSH-utilizing enzymes identified to date, parallel GSH-utilizing enzymes in function. These enzymes are involved in the metabolism of the aromatic carbon sources and the protection of the cell from ROS and RNS, xenobiotics and endogenous electrophiles, and may serve as novel drug targets for the development of novel antimycobacterials.

Recently, MSH producing *C. glutamicum* has been noted to robustly metabolize aromatic compounds, and mutants deficient in the production of MSH were unable to use gentisate as a carbon source. An MSH-dependent (MD) gentisate pathway was subsequently identified which, similar to the GSH pathway, involves 1,2-dioxygenase, maleyl pyruvate isomerase and fumaryl pyruvate hydrolase, with maleyl pyruvate isomerase requiring MSH as an essential cofactor. The gene encoding MD maleyl pyruvate isomerase was not homologous to GSH-dependent enzyme found in *K. pneumoniae* or *Ralstonia* U2. Genomic mining of other *Actinomycetales* bacteria found sequences homologous to the MD maleyl pyruvate isomerase in *C. efficiens*, *Streptomyces* strain WA46, *S. coelicolor*, *Streptomyces avermitilis*, and *Nocardia farcinica*.<sup>228</sup> While this enzyme may not be a viable drug target against tuberculosis, it does illustrate the parallel evolution of GSH and MSH dependent enzymes.

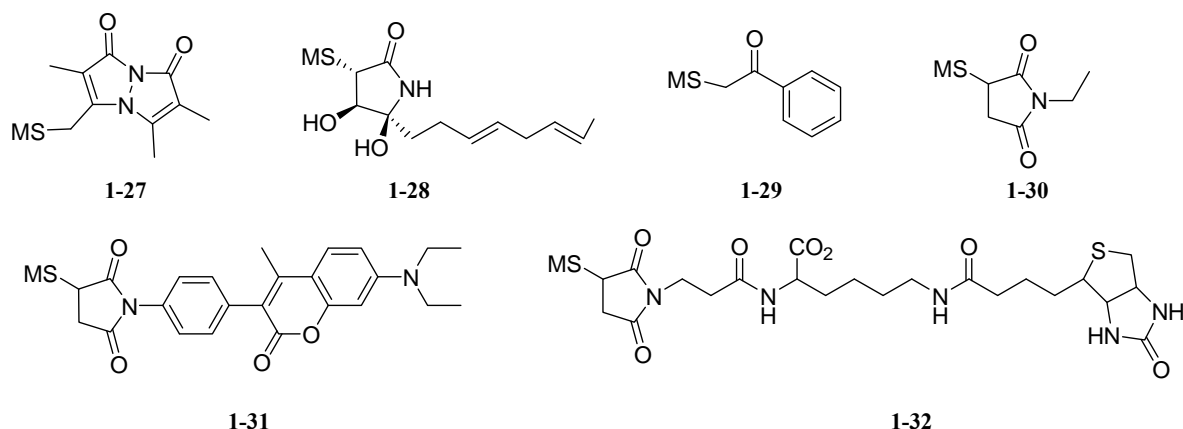
The removal of ROS is essential for cellular viability. While there have been no reported studies of the direct reaction of MSH with ROS, studies of mutants of *Mycobacteria* deficient in the biosynthesis of MSH have shown that MSH is important for the modulation of oxidative stress in these bacteria.<sup>230,231</sup> Mutants of *M. tuberculosis* deficient in MSH biosynthesis are not viable, as MSH contributes to the protection of this organism from oxidative stress.<sup>223</sup> MSH is also important for the growth of *M. smegmatis*, as mutants deficient in MSH have a much slower growth rate.<sup>232</sup> The difference in viability is believed to be related to the longer doubling time of *M. tuberculosis*, which is six times that of *M. smegmatis*. This longer growth time may allow for the accumulation of lethal

levels of oxidative damage to *M. tuberculosis*, while the faster doubling time of *M. smegmatis* results in less damage and therefore some growth is observed.<sup>225</sup> Mutants of *M. smegmatis* deficient in MSH production have increased sensitivity to H<sub>2</sub>O<sub>2</sub>,<sup>232</sup> as well as redox-cycling agents which increase the concentration of the O<sub>2</sub>•<sup>-</sup> in the cell.<sup>230</sup> The reactions of the thiols, such as GSH with peroxides and free radicals, usually lead to the formation of disulfides which must be reduced back to their biologically useful free thiol form. In keeping with this observation, an MSH disulfide (MSSM) reductase (MR)\*\* has been identified which converts MSSM to MSH<sup>233</sup> using NADPH as a reducing agent.<sup>234,235</sup> Similar to GR, MR is homodimeric, binds FAD, and has a redox active disulfide in the active site. MR is 31% homologous to the GR from *E. coli* and belongs to the same family of FAD-containing pyridine nucleotide:disulfide oxidoreductases. The chemical mechanism of MR parallels that of GR, using NADPH to reduce FAD followed by reduction of the active site disulfide. One of the active site thiols can then attack bound MSSM to form a mixed disulfide, releasing one molecule of MSH. The mixed disulfide is attacked by the other active site thiol, reforming the disulfide and releasing the second molecule of MSH. MR will not accept GSSG as a substrate. The reduction of MSSM to free MSH is essential for the protection of the cell from oxidative stress and the unique structure of MSSM makes MR an attractive drug for the development of novel antituberculin.<sup>235</sup>

The emergence of antibiotic resistant strains of TB represents a major challenge to the medical community. The presence of MSH has been shown to influence the sensitivity of *M. tuberculosis* to rifampin, an important electrophilic antituberculin.<sup>223</sup> In GSH-utilizing organisms, GSTs catalyze the coupling of electrophilic xenobiotics to GSH and these *S*-conjugates are then metabolized for excretion from the cell, often as mercapturic acids. Similarly, MSH is believed to form *S*-conjugates with rifampin<sup>223</sup> and other electrophilic xenobiotics, which are then cleaved to release mercapturic acids and GlcN-Ins.<sup>236</sup> An MSH-*S*-transferase has yet to be identified in these organisms but the resultant MSH-*S*-conjugates are known to be cleaved by MSH-*S*-conjugate amidase (MCA).<sup>236</sup> MCA is a monomeric<sup>236</sup> zinc metalloenzyme,<sup>237</sup> with high specificity for *S*-conjugates of MSH and little selectivity for the structure of the conjugate.<sup>237</sup> A sampling of known MCA substrates are shown in Figure 1.21. MCA is believed to be common to most MSH-utilizing organisms, especially those known to produce antibiotics. Mercapturic acids have been found in the fermentation broths of the *Streptomyces* bacteria which are known to produce a large number of antibiotics.<sup>238</sup> In addition, genomic searches have returned homologous sequences in *M. tuberculosis*, *M. avium*, *M. leprae*, *M. bovis*,<sup>236</sup> *C. glutamicum*, *Thermobifida fusca*,<sup>237</sup> *S. coelicolor*,<sup>237</sup> <sup>238</sup> *Amycolatopsis mediterranei*, *Saccharopolyspora erythraea*, *Streptomyces lincolnensis*, and *Streptomyces rochei*.<sup>238</sup> Mutants of *M. smegmatis* deficient in MCA have been shown to be more susceptible to electrophilic toxins including the antibiotic streptomycin.<sup>239</sup> Because MCA are not found in GSH-producing mammals, drugs inhibiting MCA may be less toxic than those for other drug targets.<sup>237</sup> MCA may therefore be an important target for the development of novel drugs to combat drug resistant tuberculosis<sup>237</sup> and the search of inhibitors of MCA is ongoing.<sup>240-245</sup>

---

\*\* Also known as mycothione reductase



**Figure 1.21:** Selected substrates for mycothiol-*S*-conjugate amidase. Structures **1-27** through **1-32** are mycothiol-*S*-conjugates of monobromobimane (**1-27**), cerulenin (**1-28**), acetophenone (**1-29**), *N*-ethylmaleimide (**1-30**), 7-(diethyl-amino)-3-(4'-maleimidylphenyl)-4-methylcoumarin (**1-31**), and 3-(*N*-maleimidopropionyl)biocytin (**1-32**).

In addition to the removal of xenobiotics, the detoxification of toxic electrophiles produced as by-products of cellular processes is essential to cellular viability. A number of metabolic and detoxification processes produce formaldehyde as a by-product, which must be removed due to its reactivity with proteins and nucleic acids. Many GSH-utilizing organisms oxidize formaldehyde to formic acid using GD-FD and *S*-formylglutathione hydrolase. In keeping with the parallel functions of MSH to GSH, an MD-formaldehyde dehydrogenase (MD-FD) has been identified in *Amycolaptopsis methanolica*, *Rhodococcus erythropolis*<sup>246</sup> and *M. smegmatis*,<sup>247</sup> which is the first step in the oxidation of formaldehyde to formic acid. MD-FD is similar to GD-FD in that it requires NADH and appears to belong to the class III alcohol dehydrogenases as it has a similar subunit molecular mass, is zinc dependent, and can oxidize higher aliphatic alcohols;<sup>246</sup> MD-FD differs from GD-FD in that it appears to be a trimer.<sup>246</sup> The residues typical of GSH binding are not present, which is consistent with the differences in structure between GSH and MSH,<sup>248</sup> and MD-FD does not bind GSH, cysteine, or their derivatives.<sup>249</sup> The conjugate of MSH and formaldehyde, *S*-hydroxymethylmycothiol, is likely the true substrate of MD-FD, forming *S*-formylmycothiol as the product.<sup>246</sup> In GSH-utilizing organisms, the conversion of *S*-formylglutathione to formic acid is catalyzed by *S*-formylmycothiol; however an *S*-formylmycothiol hydrolase has not been identified to date.

Further studies have shown that similar to GD-FD, MD-FD may be involved in the defence of *Actinomycetales* bacteria against nitrosative stress.<sup>247</sup> RNS are an essential part of the defense of a host organism against pathogenic species.<sup>250</sup> Therefore, the control of nitrosative stress is vital for the viability of pathogens like *Mycobacteria*. Upon exposure to NO, or NO donors, long lived *S*-nitrosothiols form which can nitrosate proteinaceous thiol residues, disturbing their cellular function.<sup>247,251,252</sup> *Mycobacteria* must therefore have a method for handling RNS and preventing nitrosative damage in order to survive the stresses placed upon it by the host organism's defence response. Studies of *M. smegmatis* have demonstrated that GSNO is metabolized to form GSSG and nitrate, with the involvement of MSH and MD-FD.<sup>247</sup> MSH is believed to react with GSNO to form *S*-nitrosomycothiol (MSNO), a substrate for MD-FD. MSNO is then reduced to an *N*-hydroxysulfenamide (MSNHOH) by MD-FD using NADH as a cofactor. The remaining



steps in this pathway have not yet been fully elucidated; however, this new MD-FD activity presents a possible route for the detoxification of RNS in *Mycobacteria*.<sup>247</sup> All of the MSH-utilizing enzymes thus far identified directly parallel GSH-utilizing enzymes in function. It is therefore expected that many known GSH-utilizing enzymes have MSH counterparts in the *Actinomycetales* bacteria. Because of the large differences in the structure of MSH and GSH, it is likely that these enzymes have very different criteria for substrate and cofactor binding and therefore are orthogonal in terms of selectivity. This relationship may allow for the design and development of novel antimicrobials against the pathogenic *Mycobacteria*.

### **1.5. Objectives**

Intracellular thiols are essential components of biological systems as they are responsible for the control of redox homeostasis and can act as cofactors for a number of vital enzymes. ESH and the OSHs have been shown to scavenge ROS in a biologically relevant manner and this behaviour has led to investigations into their use as novel therapeutics against disease states caused by oxidative stress. We performed a detailed theoretical study, as described in *Chapter 2*, of the thermodynamics of several important reactions that ESH and the OSHs may undergo in the cell and determined the thermodynamic driving force of the instability of the ESSE disulfide, which, unlike other known thiols, is unstable under physiological conditions. Our results should contribute to the further understanding of the fundamental chemistry and biochemistry of these important intracellular thiols

In addition to ESH and the OSHs, we investigated the behaviour of MSH, the major intracellular thiol of the *Actinomycetales* bacteria such as pathogenic *M. tuberculosis*. Our investigations commenced with a theoretical study of the conformational preferences of MSH both *in vacuo* and with implicit water solvation as described in *Chapter 3*. This information may aid in the design of substrate-based inhibitors aimed at MSH-utilizing enzymes. The functions of known MSH-utilizing enzymes parallel those of the GSH-utilizing enzymes in function. With this in mind, we sought to determine if a MSH requiring glyoxalase system exists in the *Actinomycetales* bacteria (*Chapter 4*). The identification of this enzyme system would increase the fundamental knowledge of MSH and would provide another target for the development of novel targets for the development of antimicrobials aimed at the pathogenic *Mycobacteria*.

### **1.6. Other Interesting myo-Inositol Derivatives**

In addition to our studies of ESH, the OSHs, and MSH, we investigated the thermodynamics of phosphorylation reactions involving inositol pyrophosphates. These *myo*-inositol-based species are believed to phosphorylate serine residues in eukaryotes without enzymatic assistance. Because these compounds differ greatly from the main focus of this work, we have restricted all discussion of the inositol pyrophosphates to *Chapter 5*.

### **1.7. Publications**

The research detailed in this thesis has resulted in four publications and one crystal structure deposition to the Cambridge Crystallographic Database:

Hand, C. E.; Taylor, N. J.; Honek, J. F. *Ab initio* studies of the properties of intracellular thiols ergothioneine and ovoidiol. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 1357-1360

Hand, C. E.; Honek, J. F. Biological chemistry of naturally occurring thiols of microbial and marine origin. *J. Nat. Prod.* **2005**, *68*, 293-308

Hand, C. E.; Auzanneau, F.-I.; Honek, J. F. Conformational analyses of mycothiol, a critical intracellular glycothiol in Mycobacteria. *Carbohydr. Res.* **2006**, *341*, 1164-1173

Hand, C. E.; Honek, J. F. Phosphate transfer from inositol pyrophosphates InsP<sub>5</sub>PP and InsP<sub>4</sub>PP<sub>2</sub>: A semi-empirical investigation, *Bioorg. Med. Chem. Lett.* **2007**, *7*, 183-188

Hand, C. E.; Taylor, N. J.; Honek, J. F. L-Ergothioneine, Cambridge Crystallographic Database. **2004**, Deposition No. FOLRII.

### ***1.8. References***

1. Guijaro, M. V.; Indart, A.; Aruoma, O. I.; Viana, M.; Bonet, B. *Food Chem. Tox.* **2002**, *40*, 1751-1755.
2. Jang, J.-H.; Aruoma, O. I.; Jen, L.-S.; Chung, H. Y.; Surh, Y.-J. *Free Radic. Biol. Med.* **2004**, *36*, 288-299.
3. Tanret, C. *Compt. Rend.* **1909**, *149*, 222-224.
4. Barger, G.; Ewins, A. J. *J. Chem. Soc.* **1911**, 2336-2341.
5. Baldridge, R. C.; Lewis, H. B. *J. Biol. Chem.* **1953**, *202*, 169-176.
6. Audley, B. G.; Tan, C. H. *Phytochemistry* **1968**, *7*, 1999-2000.
7. Melville, D. B.; Eich, S. *J. Biol. Chem.* **1956**, *218*, 647-651.
8. Eagles, B. A.; Vars, H. M. *J. Biol. Chem.* **1928**, *80*, 615-622.
9. Ban, R. W.; Stowell, E. C. *Arch. Biochem. Biophys.* **1956**, *63*, 259-260.
10. Fahey, R. C.; Newton, G. L.; Dorian, R.; Kosower, E. M. *Anal. Biochem.* **1981**, *111*, 357-365.
11. Melville, D. B.; Genghof, D. S.; Inamine, E.; Kovalenko, V. *J. Biol. Chem.* **1956**, *223*, 9-17.
12. Genghof, D. S. *J. Bacteriol.* **1970**, *103*, 475-478.
13. Genghof, D. S.; van Damme, O. *J. Bacteriol.* **1964**, *87*, 852-862.
14. Epand, R. M.; Epand, R. F. *J. Clin. Chem. Clin. Biochem.* **1988**, *26*, 623-626.
15. Shukla, Y.; Kulshrestha, O. P.; Khuteta, K. P. *Indian J. Med. Res.* **1981**, *73*, 472-473.
16. Shires, T. K.; Brummel, M. C.; Pulido, J. S.; Stegiink, L. D. *Comp. Biochem. Physiol.* **1997**, *117C*, 117-120.
17. Kawano, H.; Higuchi, F.; Mayumi, T.; Hama, T. *Chem. Pharm. Bull.* **1982**, *30*, 2611-2613.
18. Kawano, H.; Otani, M.; Takeyama, K.; Kawai, Y.; Mayumi, T.; Hama, T. *Chem. Pharm. Bull.* **1982**, *30*, 1760-1765.
19. Briggs, I. *J. Neurochem.* **1972**, *19*, 27-35.
20. Melville, D. B.; Eich, S.; Ludwig, M. L. *J. Biol. Chem.* **1957**, *224*, 871-877.
21. Kimura, C.; Nukina, M.; Igarashi, K.; Saugawara, Y. *Biosci. Biotechnol. Biochem.* **2005**, *69*, 357-363.
22. Melville, D. B. *Vitam. Horm.* **1959**, *17*, 155-204.

23. Melville, D. B.; Horner, W. H.; Lubschez, R. *J. Biol. Chem.* **1954**, *206*, 221-228.
24. Muda, M.; Pelizzoni, F.; Sello, G.; Mussini, E. *J. Chromatogr.* **1988**, *434*, 191-195.
25. Rae, C. D.; Sweeney, K. J. E.; Kuchel, P. W. *Magn. Reson. Med.* **1993**, *29*, 826-829.
26. Kumosani, T. A. *Exp Mol Med* **2001**, *33*, 20-22.
27. Leone, E.; Mann, T. *Nature* **1951**, *168*, 205-206.
28. Mann, T.; Leone, E. *Biochem. J.* **1953**, *53*, 140-148.
29. Deiana, M.; Rosa, A.; Casu, V.; Piga, R.; Dessi, M. A.; Aruoma, O. I. *Clin. Nutr.* **2004**, *23*, 183-193.
30. Mayumi, T.; Kawano, H.; Sakamoto, Y.; Suehisa, E.; Kawai, Y.; Hama, T. *Chem. Pharm. Bull.* **1978**, *26*, 3772-3778.
31. Hartman, P. E. *Meth. Enzymol.* **1990**, *186*, 310-317.
32. Mitsuyama, H.; May, J. M. *Clin. Sci.* **1999**, *97*, 407-411.
33. Gründemann, D.; Harlfinger, S.; Golz, S.; Geerts, A.; Lazar, A.; Berkels, R.; Jung, N.; Rubbert, A.; Schömig, E. *Proc. Natl. Acad. Sci.* **2005**, *102*, 5256-5261.
34. Askari, A.; Melville, D. B. *J. Biol. Chem.* **1962**, *237*, 1615-1618.
35. Ishikawa, Y.; Isreal, S. E.; Melville, D. B. *J. Biol. Chem.* **1974**, *249*, 4420-4427.
36. Motohashi, N.; Mori, I.; Sugiura, Y. *Chem. Pharm. Bull.* **1976**, *24*, 1737-1741.
37. Sugihara, A.; Uemura, K.; Matsuura, Y.; Tanaka, N.; Ashida, T.; Kakudo, M. *Acta Crystallogr.* **1976**, *B32*, 181-185.
38. Carlsson, J.; Kierstan, M. P.; Brocklehurst, K. *Biochem. J.* **1974**, *139*, 237-242.
39. Smith, R. C.; Gore, J. Z. *Phosphorus, Sulfur, Silicon* **1991**, *62*, 105-110.
40. Russell, K. E. *J. Phys. Chem.* **1954**, *58*, 437-439.
41. Hartman, P. E.; Hartman, Z.; Citardi, M. J. *Rad. Res.* **1988**, *113*, 319-330.
42. Motohashi, N.; Mori, I. *Chem. Pharm. Bull.* **1983**, *31*, 1702-1707.
43. Jovanovic, S. V.; Simic, M. G. In *Anticarcinogenesis and Radiation Protection 2*; O. F. Nygaard, A. C. Upton, Eds.; Plenum Press: New York, 1991, pp 229-235.
44. Hand, C. E.; Honek, J. F. *J. Nat. Prod.* **2005**, *68*, 293-308.
45. Heath, H.; Toennies, G. *Biochem. J.* **1958**, *68*, 204-210.
46. Peltekova, V. D.; Wintle, R. F.; Rubin, L. A.; Amos, C. I.; Huang, Q.; Gu, X.; Newman, B.; van Oene, M.; Cescon, D.; Greenberg, G.; Griffiths, A. M.; St. George-Hyslop, P. H.; Siminovitch, K. A. *Nat. Genet.* **2004**, *36*, 471-475.
47. Tokuhira, S.; Yamada, R.; Chang, X.; Suzuki, A.; Kochi, Y.; Sawada, T.; Suzuki, M.; Nagasaki, M.; Ohtsuki, M.; Ono, M.; Furukawa, H.; Nagashima, M.; Yoshino, S.; Mabuchi, A.; Sekine, A.; Saito, S.; Takahashi, A.; Tsunoda, T.; Nakamura, Y.; Yamamoto, K. *Nat. Genet.* **2003**, *35*, 341-348.
48. Rahman, I.; Gilmour, P. S.; Jimenez, L. A.; Biswas, S. K.; Antonicelli, F.; Aruoma, O. I. *Biochem. Biophys. Res. Comm.* **2003**, *302*, 860-864.
49. Smith, M. A.; Harris, P. L. R.; Sayre, L. M.; Beckman, J. S.; Perry, G. *J. Neurosci.* **1997**, *17*, 2653-2657.
50. Law, A.; Gautier, S.; Quirion, R. *Brain Res. Rev.* **2001**, *35*, 73-96.
51. Butterfield, D. A.; Drake, J.; Pocernich, C.; Castegna, A. *Trends Mol. Med.* **2001**, *7*, 548-554.
52. Aruoma, O. I.; Whiteman, M.; England, T. G.; Halliwell, B. *Biochem. Biophys. Res. Comm.* **1997**, *231*, 389-391.
53. Baskin, S. I.; Salem, H., Eds. *Oxidants, Antioxidants, and Free Radicals*; Taylor & Francis: Washington, D.C., 1997.

54. Akanmu, D.; Cecchini, R.; Aruoma, O. I.; Halliwell, B. *Arch. Biochem. Biophys.* **1991**, *288*, 10-16.
55. Asmus, K.-D.; Bensasson, R. V.; Bernier, J.-L.; Houssin, R.; Land, E. J. *Biochem. J.* **1996**, *315*, 625-629.
56. Smith, R. C.; Reeves, J. C.; Dage, R. C.; Schnettler, R. A. *Biochem. Pharmacol.* **1987**, *36*, 1457-1460.
57. Sundquist, A. R.; Fahey, R. C. *J. Biol. Chem.* **1989**, *264*, 719-725.
58. Tsen, C. C.; Tappel, A. L. *J. Biol. Chem.* **1958**, *233*, 1230-1232.
59. Motohashi, N.; Mori, I.; Sugiura, Y.; Tanaka, N. *Chem. Pharm. Bull.* **1974**, *22*, 654-657.
60. Rowley, D. A.; Halliwell, B. *FEBS Lett.* **1982**, *142*, 39-41.
61. Rowley, D. A.; Halliwell, B. *FEBS Lett.* **1982**, *138*, 33-36.
62. Acworth, I. N.; McCabe, D. R.; Maher, T. J. In *Oxidants, Antioxidants and Free Radicals*; S. I. Baskin, H. Salem, Eds.; Taylor and Francis: Washington, D. C., 1997, pp 22-77.
63. Arduini, A.; Eddy, L.; Hochstein, P. *Arch. Biochem. Biophys.* **1990**, *281*, 41-43.
64. Ternay, A. L., Jr.; Sorokin, V. In *Oxidants, Antioxidants, and Free Radicals*; S. I. Baskin, H. Salem, Eds.; Taylor and Francis: Washington, D. C., 1997.
65. Arduini, A.; Mancinelli, G.; Radatti, G. L.; Hochsetein, P.; Cadenas, E. *Arch. Biochem. Biophys.* **1992**, *294*, 398-402.
66. Korthuis, R. J.; Carden, D. L.; Granger, D. N. In *Biological Consequences of Oxidative Stress: Implications for Cardiovascular Disease and Carcinogenesis*; A. D. Bloom, Ed.; Oxford University Press: New York, 1992, pp 50-77.
67. Cochrane, C. G. *Am. J. Med.* **1991**, *91*, 3C:23S-30S.
68. Turner, E.; Hager, L. J.; Shapiro, B. M. *Science* **1988**, *242*, 939-941.
69. Zoete, V.; Bailly, F.; Catteau, J.-P.; Bernier, J.-L. *J. Chem. Soc. Perkin Trans. 1* **1997**, 2983-2988.
70. Zoete, V.; Bailly, F.; Vezin, H.; Teissier, E.; Duriez, P.; Fruchart, J.-C.; Catteau, J.-P.; Bernier, J.-L. *Free Radic. Res.* **2000**, *32*, 515-524.
71. Zoete, V.; Vezin, H.; Bailly, F.; Vergoten, G.; Catteau, J.-P.; Bernier, J.-L. *Free Radic. Res.* **2000**, *32*, 525-533.
72. Spies, H. S.; Steenkamp, D. J. *Eur. J. Biochem.* **1994**, *224*, 203-213.
73. World Health Organization, *Fact sheet No. 259: African Trypanosomiasis (Sleeping Sickness)*; <http://www.who.int/mediacentre/factsheets/fs259/en/>
74. World Health Organization, *Chagas Disease*; <http://www.who.int/tdr/diseases/chagas/files/chagas-poster.pdf>
75. World Health Organization, *Fact sheet No. 116: The leishmaniases and Leishmania/HIV co-infections*; <http://www.who.int/mediacentre/factsheets/fs116/en/>
76. Palumbo, A.; Misuraca, G.; D'ischia, M.; Donaudy, F.; Prota, G. *Comp. Biochem. Physiol.* **1984**, *78B*, 81-83.
77. Rohl, I.; Scheider, B.; Schmidt, B.; Zeeck, E. *Z. Naturforsch.* **1999**, *54C*, 1145-1147.
78. Turner, E.; Klevit, R.; Hager, L. J.; Shapiro, B. M. *Biochemistry* **1987**, *26*, 4028-4036.
79. Rossi, F.; Nardi, G.; Palumbo, A.; Prota, G. *Comp. Biochem. Physiol.* **1985**, *80B*, 843-845.
80. Ariyanayagam, M. R.; Fairlamb, A. H. *Mol. Biochem. Parasitol.* **2001**, *115*, 189-198.
81. Steenkamp, D. J.; Spies, H. S. *Eur. J. Biochem.* **1994**, *223*, 43-50.

82. Selman-Reimer, S.; Duhe, R. J.; Stockman, B. J.; Selman, B. R. *J. Biol. Chem.* **1991**, *266*, 182-188.
83. Ito, S.; Nardi, G.; Palumbo, A.; Prota, G. *J. Chem. Soc. Perkin Trans. 1* **1979**, 2617-2623.
84. Pathirana, C.; Andersen, R. J. *J. Am. Chem. Soc.* **1986**, *108*, 8288-8289.
85. Vogt, R. N.; Spies, H. S.; Steenkamp, D. J. *Eur. J. Biochem.* **2001**, *268*, 5229-5241.
86. Holler, T. P.; Hopkins, P. B. *J. Am. Chem. Soc.* **1988**, *110*, 4837-4838.
87. Kosower, E. M. In *Glutathione: Chemical, Biochemical and Medical Aspects*; D. Dolphin, O. Avramovic, R. Poulson, Eds.; John Wiley and Sons: New York, 1989; Vol. A, pp 103-146.
88. Reuben, D. M. E.; Bruice, T. C. *J. Am. Chem. Soc.* **1976**, *98*, 114-121.
89. Holler, T. P.; Hopkins, P. B. *Biochemistry* **1990**, *29*, 1953-1961.
90. Foerder, C. A.; Shapiro, B. M. *Proc. Natl. Acad. Sci. USA* **1977**, *74*, 4214-4218.
91. Hall, H. G. *Cell* **1978**, *15*, 343-355.
92. Shapiro, B. M. *Science* **1991**, *252*, 533-536.
93. Vogt, R. N.; Steenkamp, D. J. *Biochem. J.* **2003**, *371*, 49-59.
94. James, S. L. *Microbiol. Rev.* **1995**, *59*, 533-547.
95. Flohe, L. In *Glutathione: Chemical, Biochemical, and Medical Aspects*; D. Dolphin, O. Avramovic, R. Poulson, Eds.; John Wiley and Sons: New York, 1989; Vol. A, pp 643-731.
96. Sies, H., Ed. *Oxidative Stress*; Academic Press Inc.: London, 1985.
97. Maytin, M.; Leopold, J.; Loscalzo, J. *Curr. Atheroscler. Rep.* **1999**, *1*, 156-164.
98. Muges, G.; du Mont, W.-W.; Sies, H. *Chem. Rev.* **2001**, *101*, 2125-2179.
99. Bailly, F.; Azaroual, N.; Bernier, J.-L. *Bioorg. Med. Chem.* **2003**, *11*, 4623-4630.
100. Bailly, F.; Zoete, V.; Catteau, J.-P.; Bernier, J.-L. *FEBS Lett.* **2000**, *486*, 19-22.
101. Newton, G. L.; Fahey, R. C. *Methods Enzymol.* **1995**, *251*, 148-166.
102. Wang, W.; Ballartori, N. *Pharmacological Reviews* **1998**, *50*, 335-355.
103. Meister, A.; Anderson, M. E. *Ann. Rev. Biochem.* **1983**, *52*, 711-760.
104. Fahey, R. C.; Brown, W. C.; Adams, W. B.; Worsham, M. B. *J. Bacteriol.* **1978**, *133*, 1126-1129.
105. Fairlamb, A. H.; Blackburn, P.; Ulrich, P.; Chait, B. T.; Cerami, A. *Science* **1985**, *227*, 1485-1487.
106. Fairlamb, A. H.; Cerami, A. *Mol. Biochem. Parasitol.* **1985**, *14*, 187-198.
107. Meister, A. *Meth. Enzymol.* **1995**, *252*, 26-30.
108. Mannervick, B.; Carlberg, I.; Larson, K. In *Glutathione: Chemical, Biochemical and Medical Aspects*; D. Dolphin, R. Poulson, O. Avramovic, Eds.; John Wiley and Sons: New York, 1989; Vol. Part A, pp 476-516.
109. Rabenstein, D. L. In *Glutathione: Chemical, Biochemical, and Medical Aspects*; D. Dolphin, R. Poulson, O. Avramovic, Eds.; John Wiley and Sons: New York, 1989; Vol. A, pp 147-186.
110. Meister, A. *Trends Biochem. Sci.* **1988**, *13*, 185-188.
111. Dickinson, D. A.; Forman, H. J. *Biochem. Pharmacol.* **2002**, *64*, 1019-1026.
112. Pastore, A.; Federici, G.; Bertini, E.; Piemonte, F. *Clin. Chim. Acta* **2003**, *333*, 19-39.
113. Pompella, A.; Visvikis, A.; Paolicchi, A.; De Tata, V.; Casini, A. F. *Biochem. Pharmacol.* **2003**, *66*, 1499-1503.
114. Dolphin, D.; Avramovic, O.; Poulson, R., Eds. *Glutathione : Chemical, Biochemical, and Medical Aspects*; Wiley-Interscience: New York, 1989; Vol. A.

115. Dolphin, D.; Avramovic, O.; Poulson, R., Eds. *Glutathione : Chemical, Biochemical, and Medical Aspects*; Wiley-Interscience: New York, 1989; Vol. B.
116. Rani, M.; Prakash, D.; Sobti, R. C.; Jain, R. K. *Biochem. Biophys. Res. Comm.* **1996**, *220*, 377-381.
117. Goetz, F. E.; Harmuth, L. J. *FEMS Microbiol. Lett.* **1992**, *97*, 45-49.
118. Crawford, R. L.; Frick, T. D. *Appl. Environ. Microbiol.* **1977**, *34*, 170-174.
119. Jain, R. K. *Applied Microbiol. Biotechnol.* **1996**, *45*, 502-508.
120. Poh, C. L.; Bayly, R. C. *J. Bacteriol.* **1980**, *143*, 59-69.
121. Grund, E.; Denecke, B.; Eichenlaub, R. *Appl. Environ. Microbiol.* **1992**, *58*, 1874-1877.
122. Fuenmayor, S. L.; Wild, M.; Boyes, A. L.; Williams, P. A. *J. Bacteriol.* **1998**, *180*, 2522-2530.
123. Goetz, F. E.; Harmuth, L. J. *FEMS Microbiol.* **1992**, *97*, 45-50.
124. Hareland, W. A.; Crawford, R. L.; Chapman, P. J.; Dagley, S. *J. Bacteriol.* **1975**, *121*, 272-285.
125. Robson, N. D.; Parrott, S.; Cooper, R. A. *Microbiology* **1996**, *142*, 2115-2120.
126. Zhou, N. Y.; Fuenmayor, S. L.; Williams, P. A. *J. Bacteriol.* **2001**, *183*, 700-708.
127. Crawford, R. L.; Hutton, S. W.; Chapman, P. J. *J. Bacteriol.* **1975**, *12*, 794-799.
128. Crawford, R. L.; Frisk, T. D. *Appl. Environ. Microbiol.* **1977**, *34*, 170-174.
129. Lack, L. *J. Biol. Chem.* **1961**, *236*, 2835-2840.
130. Flohe, L. In *Glutathione: chemical, biochemical and medical aspects*; O. Avramovic, Ed.; John Wiley and Sons: New York, 1989; Vol. III, Part A, pp 643-731.
131. Epp, O.; Ladenstein, R.; Wendel, A. *Eur. J. Biochem.* **1983**, *133*, 51-69.
132. Qian, J.; Atkinson, J.; Manor, D. *Biochemistry* **2006**, *45*, 8236-8242.
133. Brigelius-Flohe, R.; Traber, M. G. *FASEB J.* **1999**, *13*, 1145-1115.
134. Burton, G. W.; Ingold, K. U. *Acc. Chem. Res.* **1986**, *19*, 194-201.
135. Banks, M. A. In *Antioxidants and Free Radicals*; S. I. Baskin, H. Salem, Eds.; Taylor and Francis: Washington, 1997, pp 95-111.
136. Sciuto, A. M. In *Oxidants, Antioxidants and Free Radicals*; S. I. Baskin, H. Salem, Eds.; Taylor and Francis: Washington, 1997, pp 171-191.
137. López-Barea, J.; Bárcena, J. A.; Bocanegra, J. A.; Florindo, J.; García-Alfonso; López-Ruiz, A.; Martínez-Galisteo; Peinado, J. In *Glutathione: Metabolism and Physiological Function*; J. Viña, Ed.; CRC Press: Boca Raton, 1990, pp 105-116.
138. Schirmer, H.; Krauth-Siegel, R. L.; Schulz, G. E. In *Glutathione: Chemical, Biochemical and Medical Aspects*; D. Dolphin, R. Poulson, O. Avramovic, Eds.; John Wiley and Sons: New York, 1989; Vol. Part A, pp 553-596.
139. Karplus, P. A.; Schulz, G. E. *J. Mol. Biol.* **1989**, *210*, 163-180.
140. Eaton, D. L.; Bammler, T. K. *Tox. Sci.* **1999**, *49*, 156-164.
141. Pickett, C. B.; Lu, A. Y. *Ann. Rev. Biochem.* **1989**, *58*, 743-764.
142. Salinas, A. E.; Wong, M. G. *Curr. Med. Chem.* **1999**, *6*, 279-309.
143. Strange, R. C.; Spiteri, M. A.; Ramachandran, S.; Fryer, A. A. *Mut. Res.* **2001**, *482*, 21-26.
144. Armstrong, R. N. *Chem. Res. Toxicol.* **1997**, *10*, 2-18.
145. Vuilleumier, S. *J. Bacteriol.* **1997**, *179*, 1431-1441.
146. Marrs, K.; Alfenito, M. R.; Lloyd, A. M.; Walbot, V. *Nature* **1995**, *375*, 397-400.

147. Kolm, R. H.; Danielson, U. H.; Zhang, Y.; Talalay, P.; Mannervick, B. *Biochem. J.* **1995**, *311*, 453-459.
148. Waydhas, C.; Weigl, K.; Sies, H. *Eur. J. Biochem.* **1978**, *89*, 143-150.
149. Jones, D. P.; Thor, H.; Andersson, B.; Orrenius, S. *J. Biol. Chem.* **1978**, *253*, 6031-6037.
150. Yurimoto, H.; Kato, N.; Sakai, Y. *Chem. Rec.* **2005**, *5*, 367-375.
151. Vorholt, J. A. *Arch. Microbiol.* **2002**, *178*, 239-249.
152. Uotila, L.; Koivusalo, M. In *Functions of Glutathione: Biochemical, Physiological, Toxicological and Clinical Aspects*; A. Larsson, S. Orrenius, A. Homgren, B. Mannervick, Eds.; Raven Press: New York, NY, USA, 1983, pp 175-186.
153. Uotila, L.; Koivusalo, M. In *Glutathione: Chemical, Biochemical and Medical Aspects*; D. Dolphin, R. Poulson, O. Avramovic, Eds.; John Wiley and Sons: New York, 1989; Vol. Part A, pp 517-552.
154. Koivusalo, M.; Baumann, M.; Uotila, L. *FEBS Lett.* **1989**, *257*, 105-109.
155. Jensen, D. E.; Belka, G. K.; Du Bois, G. C. *Biochem. J.* **1998**, *331*, 659-668.
156. Liu, L.; Hausladen, A.; Zeng, M.; Que, L.; Heitman, J.; Stamler, J. S. *Nature* **2001**, *410*, 490-494.
157. Fernández, M. R.; Biosca, J. A.; Parés, X. *Cell. Mol. Life Sci.* **2003**, *60*, 1013-1018.
158. Haqqani, A. S.; Do, S. K.; Birnboim, H. C. *Nitric Oxide* **2003**, *9*, 172-181.
159. Richard, J. P. *Biochemistry* **1991**, *30*, 4581-4585.
160. Richard, J. P. *Biochem. Soc. Trans.* **1993**, *21*, 549-553.
161. Inoue, Y.; Kimura, A. *Adv. Microb. Physiol.* **1995**, *37*, 177-227.
162. Hopper, D. J.; Cooper, R. A. *Biochem. J.* **1972**, *128*, 321-329.
163. Ferguson, G. P.; Totemeyer, S.; MacLean, M. J.; Booth, I. R. *Arch. Microbiol.* **1998**, *170*, 209-218.
164. Fraval, H. N.; McBrien, D. C. *J. Gen. Microbiol.* **1980**, *117*, 127-134.
165. Thornalley, P. J. *Biochem. Soc. Trans.* **2003**, *31*, 1343-1348.
166. Cameron, A. D.; Ridderstrom, M.; Olin, B.; Kavarana, M. J.; Creighton, D. J.; Mannervik, B. *Biochemistry* **1999**, *38*, 13480-13490.
167. Hovatta, I.; Tennant, R. S.; Helton, R.; Marr, R. A.; Singer, O.; Redwine, J. M.; Ellison, J. A.; Schadt, E. E.; Verma, I. M.; Lockhart, D. J.; Barlow, C. *Nature* **2005**, *438*, 662-666.
168. Inoue, Y.; Kimura, A. *J. Biol. Chem.* **1996**, *271*, 25958-25965.
169. Wood, V.; Gwilliam, R.; Rajandream, M. A.; Lyne, M.; Lyne, R.; Stewart, A.; Sgouros, J.; Peat, N.; Hayles, J.; Baker, S.; Basham, D.; Bowman, S.; Brooks, K.; Brown, D.; Brown, S.; Chillingworth, T.; Churcher, C.; Collins, M.; Connor, R.; Cronin, A.; Davis, P.; Feltwell, T.; Fraser, A.; Gentles, S.; Goble, A.; Hamlin, N.; Harris, D.; Hidalgo, J.; Hodgson, G.; Holroyd, S.; Hornsby, T.; Howarth, S.; Huckle, E. J.; Hunt, S.; Jagels, K.; James, K.; Jones, L.; Jones, M.; Leather, S.; McDonald, S.; McLean, J.; Mooney, P.; Moule, S.; Mungall, K.; Murphy, L.; Niblett, D.; Odell, C.; Oliver, K.; O'Neil, S.; Pearson, D.; Quail, M. A.; Rabbinowitsch, E.; Rutherford, K.; Rutter, S.; Saunders, D.; Seeger, K.; Sharp, S.; Skelton, J.; Simmonds, M.; Squares, R.; Squares, S.; Stevens, K.; Taylor, K.; Taylor, R. G.; Tivey, A.; Walsh, S.; Warren, T.; Whitehead, S.; Woodward, J.; Volckaert, G.; Aert, R.; Robben, J.; Grymonprez, B.; Weltjens, I.; Vanstreels, E.; Rieger, M.; Schafer, M.; Muller-Auer, S.; Gabel, C.; Fuchs, M.; Dusterhoft, A.; Fritzc, C.; Holzer, E.; Moestl, D.; Hilbert, H.; Borzym, K.; Langer, I.; Beck, A.; Lehrach, H.; Reinhardt, R.; Pohl, T. M.;

- Eger, P.; Zimmermann, W.; Wedler, H.; Wambutt, R.; Purnelle, B.; Goffeau, A.; Cadieu, E.; Dreano, S.; Gloux, S.; Lelaure, V.; Mottier, S.; Galibert, F.; Aves, S. J.; Xiang, Z.; Hunt, C.; Moore, K.; Hurst, S. M.; Lucas, M.; Rochet, M.; Gaillardin, C.; Tallada, V. A.; Garzon, A.; Thode, G.; Daga, R. R.; Cruzado, L.; Jimenez, J.; Sanchez, M.; del Rey, F.; Benito, J.; Dominguez, A.; Revuelta, J. L.; Moreno, S.; Armstrong, J.; Forsburg, S. L.; Cerutti, L.; Lowe, T.; McCombie, W. R.; Paulsen, I.; Potashkin, J.; Shpakovski, G. V.; Ussery, D.; Barrell, B. G.; Nurse, P. *Nature* **2002**, *415*, 871-880.
170. Theologis, A.; Ecker, J. R.; Palm, C. J.; Federspiel, N. A.; Kaul, S.; White, O.; Alonso, J.; Altafi, H.; Araujo, R.; Bowman, C. L.; Brooks, S. Y.; Buehler, E.; Chan, A.; Chao, Q.; Chen, H.; Cheuk, R. F.; Chin, C. W.; Chung, M. K.; Conn, L.; Conway, A. B.; Conway, A. R.; Creasy, T. H.; Dewar, K.; Dunn, P.; Etgu, P.; Feldblyum, T. V.; Feng, J.; Fong, B.; Fujii, C. Y.; Gill, J. E.; Goldsmith, A. D.; Haas, B.; Hansen, N. F.; Hughes, B.; Huizar, L.; Hunter, J. L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C. J.; Koo, H. L.; Kremenetskaia, I.; Kurtz, D. B.; Kwan, A.; Lam, B.; Langin-Hooper, S.; Lee, A.; Lee, J. M.; Lenz, C. A.; Li, J. H.; Li, Y.; Lin, X.; Liu, S. X.; Liu, Z. A.; Luros, J. S.; Maiti, R.; Marziali, A.; Militscher, J.; Miranda, M.; Nguyen, M.; Nierman, W. C.; Osborne, B. I.; Pai, G.; Peterson, J.; Pham, P. K.; Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.; Salzberg, S. L.; Schwartz, J. R.; Shinn, P.; Southwick, A. M.; Sun, H.; Tallon, L. J.; Tambunga, G.; Toriumi, M. J.; Town, C. D.; Utterback, T.; Van Aken, S.; Vaysberg, M.; Vysotskaia, V. S.; Walker, M.; Wu, D.; Yu, G.; Fraser, C. M.; Venter, J. C.; Davis, R. W. *Nature* **2000**, *408*, 816-820.
171. Parani, M.; Parida, A. *Unpublished data*, Accession number AAK06838.
172. Booker, J. P. *Unpublished data*, Accession number Q39366.
173. Veena; Reddy, V. S.; Sopory, S. K. *Plant J.* **1999**, *17*, 385-395.
174. Skipsey, M.; Andrews, C. J.; Townson, J. K.; Jepson, I.; Edwards, R. *Arch. Biochem. Biophys.* **2000**, *374*, 261-268.
175. Usui, Y.; Nakase, M.; Hotta, H.; Urisu, A.; Aoki, N.; Kitajima, K.; Matsuda, T. *J. Biol. Chem.* **2001**, *276*, 11376-11381.
176. Espartero, J.; Sanchez-Aguayo, I.; Pardo, J. M. *Plant. Mol. Biol.* **1995**, *29*, 1223-1233.
177. Johansen, K. S.; Svendsen, I. I.; Rasmussen, S. K. *Plant Science* **2000**, *155*, 11-20.
178. Iozef, R.; Rahlfs, S.; Chang, T.; Schirmer, H.; Becker, K. *FEBS Lett.* **2003**, *554*, 284-288.
179. Castro, N. S.; Bastos, K. P.; Jesuino, R. S. A.; Ffleipe, M. S. S.; Pereira, M.; Soares, C. M. A. *Unpublished data*, Accession number AAP03992.
180. Clugston, S. L.; Daub, E.; Honek, J. F. *J. Mol. Evol.* **1998**, *47*, 230-234.
181. He, M. M.; Clugston, S. L.; Honek, J. F.; Matthews, B. W. *Biochemistry* **2000**, *39*, 8719-8727.
182. Sellin, S.; Eriksson, L. E.; Mannervik, B. *Biochemistry* **1982**, *21*, 4850-4857.
183. Landro, J. A.; Brush, E. J.; Kozarich, J. W. *Biochemistry* **1992**, *31*, 6069-6077.
184. Cameron, A. D.; Olin, B.; Ridderstrom, M.; Mannervik, B.; Jones, T. A. *Embo J* **1997**, *16*, 3386-3395.
185. Ridderstrom, M.; Mannervik, B. *Biochem J* **1996**, *314*, 463-467.
186. Clugston, S. L.; Barnard, J. F.; Kinach, R.; Miedema, D.; Ruman, R.; Daub, E.; Honek, J. F. *Biochemistry* **1998**, *37*, 8754-8763.
187. Uotila, L. *Biochemistry* **1973**, *12*, 3944-3951.
188. Allen, R. E.; Lo, T. W.; Thornalley, P. J. *Eur. J. Biochem.* **1993**, *213*, 1261-1267.



189. Talesa, V.; Uotila, L.; Koivusalo, M.; Principato, G.; Giovannini, E.; Rosi, G. *Biochim. Biophys. Acta* **1988**, *955*, 103-110.
190. Talesa, V.; Uotila, L.; Koivusalo, M.; Principato, G.; Giovannini, E.; Rosi, G. *Biochim. Biophys. Acta* **1989**, *993*, 7-11.
191. Murata, K.; Saikusa, T.; Watanabe, K.; Fukuda, Y.; Makoto, S.; Kimura, S. *J. Ferment. Technol.* **1986**, *64*, 1-4.
192. Inoue, T.; Kimura, A. *J. Ferment. Bioeng* **1992**, *73*, 271-276.
193. Norton, S. J.; Principato, G. B.; Talesa, V.; Lupattelli, M.; Rosi, G. *Enzyme* **1989**, *42*, 189-196.
194. Norton, S. J.; Talesa, V.; Yuan, W. J.; Principato, G. B. *Biochem Int* **1990**, *22*, 411-418.
195. Talesa, V.; Rosi, G.; Contenti, S.; Mangiabene, C.; Lupattelli, M.; Norton, S. J.; Giovannini, E.; Principato, G. B. *Biochem. Int.* **1990**, *22*, 1115-1120.
196. Saxena, M.; Bisht, R.; Roy, S. D.; Sopory, S. K.; Bhalla-Sarin, N. *Biochem Biophys Res Commun* **2005**, *336*, 813-819.
197. Talesa, V.; Rosi, G.; Bistoni, F.; Marconi, P.; Norton, S. J.; Principato, G. B. *Biochem Int* **1990**, *21*, 397-403.
198. Ridderstrom, M.; Saccucci, F.; Hellman, U.; Bergman, T.; Principato, G.; Mannervik, B. *J Biol Chem* **1996**, *271*, 319-323.
199. Ridderstrom, M.; Mannervik, B. *Biochem J* **1997**, *322*, 449-454.
200. Bito, A.; Haider, M.; Hadler, I.; Breitenbach, M. *J. Biol. Chem.* **1997**, *272*, 21509-21519.
201. Kizil, G.; Wilks, K.; Wells, D.; Ala'Aldeen, D. A. *J Med Microbiol* **2000**, *49*, 669-673.
202. Uotila, L. *Biochemistry* **1973**, *12*, 3944-3951.
203. Maiti, M. K.; Krishnasamy, S.; Owen, H. A.; Makaroff, C. A. *Plant Mol Biol* **1997**, *35*, 471-481.
204. Bito, A.; Haider, M.; Briza, P.; Strasser, P.; Breitenback, M. *Protein Expr. Purif.* **1999**, *17*, 456-464.
205. Cordell, P. A.; Futers, S. T.; Grant, P. J.; Pease, R. J. *J. Biol. Chem.* **2004**, *279*, 28653-28661.
206. Crowder, M. W.; Maiti, M. K.; Banovic, L.; Makaroff, C. A. *FEBS Lett.* **1997**, *418*, 351-354.
207. Cameron, A. D.; Ridderstrom, M.; Olin, B.; Mannervick, B. *Structure* **1999**, *7*, 1067-1078.
208. Schilling, O.; Wenzel, N.; Naylor, M.; Vogel, A.; Crowder, M. W.; Makaroff, C. A.; Meyer-Klaucke, W. *Biochemistry* **2003**, *42*, 11777-11786.
209. Misra, K.; Bannerjee, A. B.; Ray, S.; Ray, M. *Biochem. J.* **1995**, *305*, 999-1003.
210. World Health Organization, *Tuberculosis*;  
[www.who.int/mediacentre/factsheets/fs104/en/print.html](http://www.who.int/mediacentre/factsheets/fs104/en/print.html)
211. Food and Agriculture Organization of the United Nations; *Tuberculosis*;  
<http://www.fao.org/ag/AGAInfo/subjects/en/health/diseases-cards/cards/tuberculosis.html>.
212. Rice, D. N.; Rogers, D. G.; *Johne's Disease (Paratuberculosis)*;  
<http://ianrpubs.unl.edu/animaldisease/g977.htm>, University of Nebraska at Lincoln Institute of Agriculture and Natural Resources.
213. Sakuda, S.; Zhou, Z. Y.; Yamada, Y. *Biosci. Biotech. Biochem.* **1994**, *58*, 1347-1348.

214. Newton, G. L.; Fahey, R. C.; Cohen, G.; Aharonowitz, Y. *J. Bacteriol.* **1993**, *175*, 2734-2742.
215. Newton, G. L.; Bewley, C. A.; Dwyer, T. J.; Horn, R.; Aharonowitz, Y.; Cohen, G.; Davies, J.; Faulkner, D. J.; Fahey, R. C. *Eur. J. Biochem.* **1995**, *230*, 821-825.
216. Newton, G. L.; Arnold, K.; Price, M. S.; Sherrill, C.; Delcardayre, S. B.; Aharonowitz, Y.; Cohen, G.; Davies, J.; Fahey, R. C.; Davis, C. *J. Bacteriol.* **1996**, *178*, 1990-1995.
217. Newton, G. L.; Ta, P.; Bzymek, K. P.; Fahey, R. C. *J. Biol. Chem.* **2006**, *In press*.
218. Newton, G. L.; Koledin, T.; Gorovitz, B.; Rawat, M.; Fahey, R. C.; Av-Gay, Y. *J. Bacteriol.* **2003**, *185*, 3476-3479.
219. Newton, G. L.; Av-Gay, Y.; Fahey, R. C. *J. Bacteriol.* **2000**, *182*, 6958-6963.
220. Sareen, D.; Steffek, M.; Newton, G. L.; Fahey, R. C. *Biochemistry* **2002**, *41*, 6885-6890.
221. Koledin, T.; Newton, G. L.; Fahey, R. C. *Arch. Microbiol.* **2002**, *178*, 331-337.
222. Rawat, M.; Kovacevic, S.; Billman-Jacobe, H.; Av-Gay, Y. *Microbiology* **2003**, *149*, 1341-1349.
223. Buchmeier, N. A.; Newton, G. L.; Koledin, T.; Fahey, R. C. *Mol. Microbiol.* **2003**, *47*, 1723-1732.
224. Rengarajan, J.; Bloom, B. R.; Rubin, E. J. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 8327-8332.
225. Sareen, D.; Newton, G. L.; Fahey, R. C.; Buchmeier, N. A. *J. Bacteriol.* **2003**, *185*, 6736-6740.
226. Gammon, D. W.; Hunter, R.; Steenkamp, D. J.; Mudzunga, T. T. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 2045-2049.
227. Park, J. H.; Cha, C. J.; Roe, J. H. *J. Microbiol.* **2006**, *44*, 121-125.
228. Feng, J.; Che, Y.; Milse, J.; Yin, Y.-J.; Liu, L.; Rückert, C.; Shen, X.-H.; Qi, S.-W.; Kalinowski, J.; Liu, S.-J. *J. Biol. Chem.* **2006**, *271*, 10778-10785.
229. Kosower, E. M.; Kosower, N. S. *Meth. Enzymol.* **1995**, *251*, 133-148.
230. Rawat, M.; Newton, G. L.; Ko, M.; Martinez, G. J.; Fahey, R. C.; Av-Gay, Y. *Antimicrob. Agents Chemother.* **2002**, *46*, 3348-3355.
231. Hayward, D.; Wiid, I.; van Helden, P. *IUBMB Life* **2004**, *56*, 131-138.
232. Newton, G. L.; Unson, M. D.; Anderberg, S. J.; Aguilera, J. A.; Oh, N. N.; delCardayre, S. B.; Av-Gay, Y.; Fahey, R. C. *Biochem. Biophys. Res. Commun.* **1999**, *255*, 239-244.
233. Patel, M. P.; Blanchard, J. S. *J. Am. Chem. Soc.* **1998**, *120*, 11538-11539.
234. Patel, M. P.; Blanchard, J. S. *Biochemistry* **1999**, *38*, 11827-11833.
235. Patel, M. P.; Blanchard, J. S. *Biochemistry* **2001**, *40*, 5119-5126.
236. Newton, G. L.; Av-Gay, Y.; Fahey, R. C. *Biochemistry* **2000**, *39*, 10739-10749.
237. Steffek, M.; Newton, G. L.; Av-Gay, Y.; Fahey, R. C. *Biochemistry* **2003**, *42*, 12067-12076.
238. Newton, G. L. *Arch. Microbiol.* **2002**, *178*, 388-394.
239. Rawat, M.; Uppal, M.; Newton, G. L.; Steffek, M.; Fahey, R. C.; Av-Gay, Y. *J. Bacteriol.* **2004**, *186*, 6050-6058.
240. Nicholas, G. M.; Newton, G. L.; Fahey, R. C.; Bewley, C. A. *Org. Lett.* **2001**, *3*, 1543-1545.
241. Fetterolf, B.; Bewley, C. A. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 3785-3788.

242. Knapp, S.; Gonzalez, S.; Myers, D. S.; Eckman, L. L.; Bewley, C. A. *Org. Lett.* **2002**, *4*, 4337-4339.
243. Nicholas, G. M.; Eckman, L. L.; Ray, S.; Hughes, R. O.; Pfefferkorn, J. A.; Barluenga, S.; Nicolaou, K. C.; Bewley, C. A. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 2487-2490.
244. Knapp, S.; Amorelli, B.; Darout, D.; Ventocilla, C. C.; Goldman, L. M.; Huhn, R. A.; Minnihan, E. C. *J. Carbohydr. Chem.* **2005**, *24*, 103-130.
245. Mahadevan, J.; Nicholas, G. M.; Bewley, C. A. *J. Org. Chem.* **2003**, *68*, 3380-3386.
246. van Ophem, P. W.; Van Beeumen, J.; Dune, J. A. *Eur. J. Biochem.* **1992**, *206*, 511-518.
247. Vogt, R. N.; Steenkamp, D. J.; Zheng, R.; Blanchard, J. S. *Biochem. J.* **2003**, *374*, 657-666.
248. Norin, A.; van Ophem, P. W.; Piersma, S. R.; Persson, B.; Duine, J. A.; Jörnvall, H. *Eur. J. Biochem.* **1997**, *248*, 282-289.
249. Misset-Smits, M.; van Ophem, P. W.; Sakuda, S.; Duine, J. A. *FEBS Lett.* **1997**, *409*, 221-222.
250. Nathan, C.; Shiloh, M. U. *Proc. Natl. Acad. Sci.* **2000**, *97*, 8841-8848.
251. Clancy, R. M.; Levartovsky, D.; Leszczynska-Piziak, J.; Yegudin, J.; Abramson, S. B. *Proc. Natl. Acad. Sci.* **1994**, *91*, 3680-3684.
252. Kharitonov, V. G.; Sundquist, A. R.; Sharma, V. S. *J. Biol. Chem.* **1995**, *270*, 28158-28164.

## Chapter 2: Modelling of Ergothioneine and Ovothiol

Unlike other biological thiols, the ergothioneine disulfide (ESSE) is not stable under physiological conditions and it has not been detected under conditions that are known to oxidize glutathione (GSH) to its disulfide (GSSG).<sup>1</sup> It is likely that even if formed *in vivo*, ESSE would be quickly reduced to its free thiol form (ESH, Figure 2.1, **2-1**) by the large concentrations of GSH and mycothiol (MSH) found in ESH-containing organisms. ESH has been implicated in radical scavenging activities *in vivo*, and shows promise as a therapeutic agent against disease states caused by radical damage such as is found in Alzheimer's disease.<sup>2</sup> Studies have shown that during the scavenging of hydroxide radicals (HO•), ESSE may be formed transiently before being reduced to ESH.<sup>3</sup> Despite the possible importance of the formation of ESSE, very little thermodynamic information exists regarding its instability. In addition, very little data is available regarding the thermodynamics of the known reactions of ESH with HO•<sup>3</sup> and hydrogen peroxide (H<sub>2</sub>O<sub>2</sub>),<sup>3,4</sup> as well as the possible cooperative relationship believed to exist between ESH and ascorbic acid in the protection of the cell from reactive oxygen species (ROS).<sup>5</sup> With this in mind, we performed a detailed computational study of ESH and its possible reactions to determine the driving force behind the ESSE instability and the thermodynamic nature of some important detoxification events.

In addition, we performed a detailed study of the related ovothiols (OSHS, Figure 2.1, **2-2A-C**) to determine if there is a thermodynamic basis for the differences in the reactivity between ESH and the OSHS. These thiols differ in the placement of the sulfur atom on the imidazole ring and this difference alters the reactivity of OSH as compared to ESH, such that the OSH disulfide (OSSO) is stable under physiological conditions.<sup>6,7</sup> The OSHS are currently under review as possible therapeutic agents in the prevention of oxidative damage<sup>6,8</sup> as their proposed role in sea urchin eggs is to scavenge the H<sub>2</sub>O<sub>2</sub> produced during fertilization.<sup>6</sup>

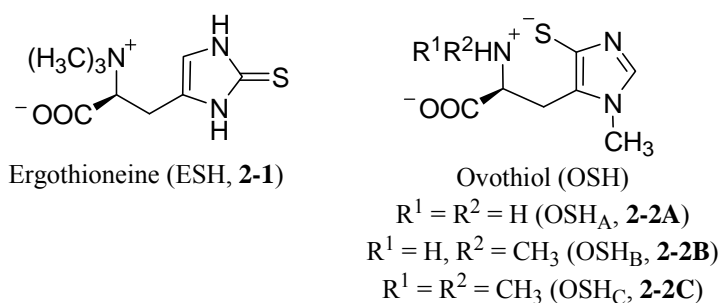


Figure 2.1: Ergothioneine and the ovothiols.

### 2.1. Computational Chemistry

Computational chemistry uses the results of theoretical chemistry to determine the structures and properties of molecules and molecular systems.<sup>9</sup> A wide variety of properties can be described using these techniques including minimum energy and transition state structures, enthalpies, vibrational frequencies and spectroscopic properties.<sup>10</sup> The results obtained from computational studies cannot be defined as exact values as even the use of quantum mechanics cannot give an exact result for a system more complicated than a

hydrogen atom.<sup>11</sup> As such, all computational chemistry methods use approximations and the number and sophistication of the approximations determines the accuracy of the results. While these techniques do not give exact values, they can give useful information regarding trends and insight into systems which are difficult to study experimentally.

There are a variety of molecular modelling techniques available. The *ab initio* molecular orbital calculations for example, are derived directly from quantum mechanical theory<sup>11</sup> and include approximations that are usually mathematical in nature and do not incorporate experimental data.<sup>11</sup> One of the most well-known *ab initio* techniques is the Hartree-Fock (HF) method, which builds the many-electron Schrödinger equation from many, simpler, one-electron equations. Though computationally expensive, *ab initio* techniques give very good qualitative results.<sup>11</sup> An alternative to the HF method is density functional theory (DFT) which uses the electron density, as opposed to the Schrödinger equation, to determine the energetics of the molecule.<sup>11</sup> This method can give results for a wide variety of properties that are comparable or superior to *ab initio* calculations.<sup>9</sup>

While the HF and DFT methods are known to give very good results, they are computationally demanding with respect memory and time. Other less rigorous methods have been developed, including the semi-empirical (see *Chapter 5*) and molecular mechanics (see *Chapter 3*) methods. Semi-empirical methods take the same form as an HF calculation; however, they simplify the calculation by approximating or eliminating certain integrals. They do so by incorporating experimental data or values obtained from *ab initio* calculations as parameters to fit their data to model compounds. While these approximations hasten the calculation, they reduce the accuracy as these results are dependent upon the system used to develop the parameter set. Compared to the *ab initio* techniques, fewer properties can be predicted reliably using semi-empirical methods.<sup>11</sup> The molecular mechanics methods use classical mechanics calculations to determine the properties of a system. These calculations are parameterized and are more sensitive than semi-empirical methods to the system used to develop the parameter set. While the semi-empirical and molecular mechanics methods are much less time consuming and are computationally less expensive than HF and DFT techniques, they are often less accurate. These methods are very useful in the modelling of large or difficult systems where time and resource constraints are issues.<sup>11</sup>

Based upon the above considerations, we chose HF and DFT techniques for our study of ESH and OSH. These molecules are relatively small and therefore the time and computational resource demands of the geometry optimizations would not be prohibitive and we desired the increased accuracy that these techniques would provide.

### 2.1.1. Hartree-Fock Methods

The HF method is the most commonly used, purely *ab initio* technique and is based on quantum mechanical theory. One of the fundamental postulates of quantum mechanics describes electrons as waves whose behaviour can be defined by a wave function. This wave function can describe the probability of the electrons' locations but not their exact position. The HF method builds the wave function of a system from many simpler one-electron equations which can be solved to give one-electron wave functions.<sup>11</sup> These one-electron equations are commonly called *basis functions* and a complete set of basis functions is called a *basis set*. The accuracy of the basis set scales with the number of functions used to describe the system (see 2.1.2. *Basis Functions and Sets*).

The HF method is an approximate quantum mechanical method with the primary approximation being the omission of an explicit definition of electron-electron repulsion, also known as electron correlation. This approximation is called the central field approximation and describes the behaviour of the electron in a net field of electrons giving an average effect of the electron-electron repulsion. As a result, all energies calculated using the HF method will be greater than the actual energy of the system; the limit to how close the calculated energies can approach the actual energy of the system is termed the HF limit<sup>11</sup>. There are a number of methods available to correct for the lack of electron correlation, as in general, its inclusion improves the accuracy of the calculated energies and molecular geometries. One popular method is the Møller-Plesset perturbation theory (MP $n$ , where  $n$  indicates the level of theory), which uses perturbation theory to improve the results. The MP method is attractive in that it is size-independent; however, it can give energies lower than the “true” energy of the system.<sup>9</sup> The inclusion of electron correlation is not always necessary; in the case of organic molecules correlation is not generally required to obtain quantitative results.<sup>11</sup>

Despite the HF limit, the energies calculated using HF methods can be very useful in calculating overall reaction thermodynamics. If the bond types in the products and reactants of a reaction are essentially identical, it is expected that the bond-by-bond electron correlation energy errors would mostly cancel in the computed heat of reaction resulting in relatively accurate results.<sup>10</sup> If comparing structures with similar bond types, the errors caused by a lack of electron correlation energy would also likely be reduced and the trends noted would therefore be valid.

### 2.1.2. Basis Functions and Sets

The choice of basis set is one of the largest determinants in the results obtained by HF and DFT calculations. A minimal basis set uses only the number of functions required to account for all of the atomic orbitals in the system. For example, in the case of hydrogen and helium, the basis set would include a single function to describe the 1s orbital, while the second row of the periodic table, lithium to neon would require five functions (the 1s, 2s, and three 2p orbitals) and third row, sodium to argon, would require nine functions.<sup>12</sup> Minimal basis sets unfortunately have several failings, especially when dealing with atoms at the end of a period. Atoms such as oxygen and fluorine are described using the same number of functions as atoms at the beginning of the period, even though they have more electrons. One solution to address the deficiencies of the minimal basis set is to increase the number of functions used to describe each orbital.<sup>9</sup> The double-zeta (DZ) basis sets replace each basis function in the minimal basis set with two, while in the triple-zeta set (TZ) three basis functions are used to describe each of the orbitals. Unfortunately, doubling or tripling the number of basis functions increases the computational demands of the calculations. As an alternative, the split valence method doubles the number of functions used to describe the valence electrons but uses a single function to describe the inner orbitals. These core orbitals do not have a large effect on chemical properties and vary only slightly between molecules.<sup>9</sup> Split valence methods include the 6-31G and 6-311G basis sets.

The split valence method does not solve all of the problems related to basis set accuracy. The basis sets often use functions that are centered on the nucleus and therefore do not take into account the change in the shape of the electron cloud when an atom is part of a molecule.<sup>9</sup> The performance of a basis set can be improved by the addition of polarization functions, which give more flexibility to the shape of the orbitals, often yielding more

accurate geometries and vibrational frequencies.<sup>11</sup> The application of polarization functions to heavy atoms is indicated by “\*”, and the further addition to helium and hydrogen atoms is indicated by “\*\*”. Even with the addition of polarization functions, these basis sets have difficulty describing atoms with lone pairs and anions, because both have a large amount of electron density distal from the nucleus. The addition of diffuse functions aids in the description of the orbital far from the nucleus.<sup>9</sup> These functions describe van der Waals and other long distance interactions,<sup>11</sup> as well as account for the effects of lone pairs and anions. The addition of diffuse functions is denoted by “+” for heavy atoms and “++” for hydrogen.

The accuracy of the calculation is dependent upon the basis set in use. The better a basis set can describe the system, the closer the calculated energy will approach the HF limit for that system. As can be seen in Table 2.1, as the basis sets becomes more complete, the calculated energy approaches the HF limit.<sup>12</sup>

**Table 2.1:** As a basis set becomes more complete, the computed energies approach the HF limit<sup>a,b</sup>

<i>Basis Set</i>	<i>H<sub>2</sub></i>	<i>N<sub>2</sub></i>	<i>CH<sub>4</sub></i>	<i>NH<sub>3</sub></i>	<i>H<sub>2</sub>O</i>
4-31G	-1.127	-108.754	-40.140	-56.102	-75.907
6-31G*	-1.127	-108.942	-40.195	-56.184	-76.011
6-31G**	-1.131	-108.942	-40.202	-56.195	-76.023
HF limit	-1.134	-108.997	-40.225	-56.225	-76.065

<sup>a</sup>The energies are expressed and were calculated using HF methods as multiples of the hartree,  $E_h=4.360 \times 10^{-18}\text{J}$ ; <sup>b</sup>Adapted from Atkins and Friedman (2005).<sup>12</sup>

### 2.1.3. Density Functional Theory

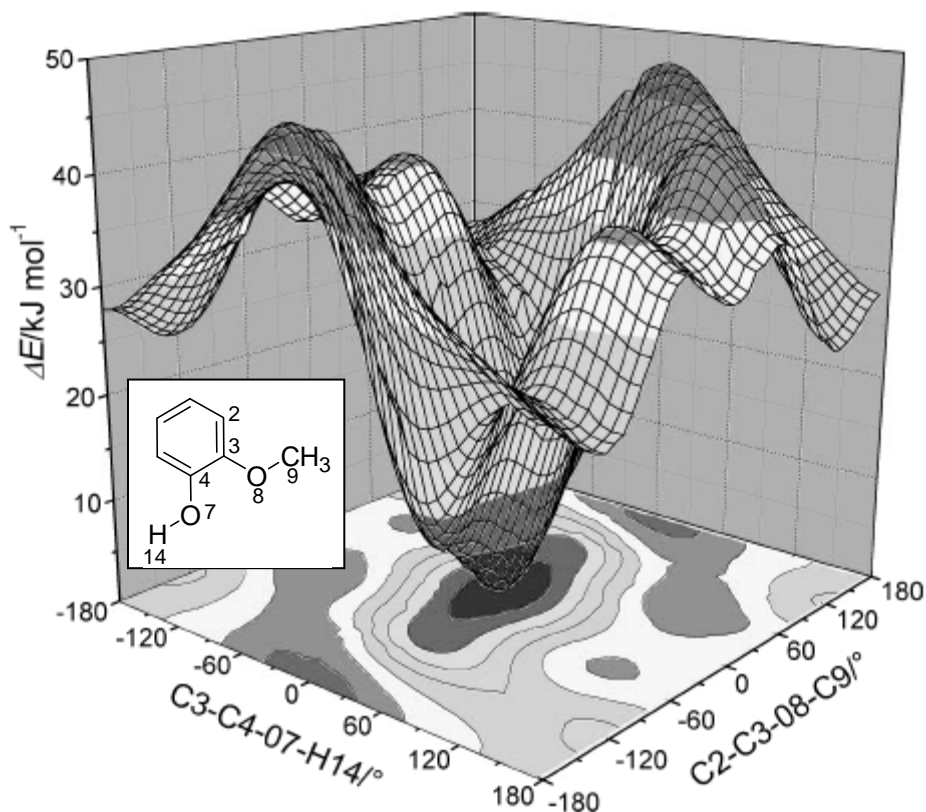
A popular alternative to HF methods is DFT which uses the electron density of a molecule to determine its energy. The electron density is expressed as a linear combination of basis functions, similar to the HF calculations.<sup>11</sup> A *functional* is a mathematical function applied to another function.<sup>12</sup> Density functionals can be broken down into several classes and the functional used can affect the accuracy of the calculation. In general, the two most accurate methods are the gradient-corrected functionals, which use the electron density and its gradient, and the hybrid method which combines functionals from other methods with aspects of the HF method to determine the energy of a system.<sup>11</sup> A commonly used functional is the B3LYP<sup>13-15</sup> hybrid functional which has been shown to yield accurate results for a wide range of molecules, particularly organic molecules.<sup>11</sup>

The accuracy of the DFT methods has been shown to depend on the basis set and density functional chosen. The results obtained from DFT calculations can match and even surpass the accuracy of HF calculations in some cases. In addition, DFT techniques are often less time consuming and require less computational resources.<sup>11</sup> These factors, combined with the inclusion of electron correlation, make DFT techniques an attractive alternative to HF methods for determining the energy of a system.

### 2.1.4. Energy Minimization or Geometry Optimization

The potential energy of a molecule is a function of the coordinates of its atoms. The way in which this energy changes with respect to changes in the molecular geometry defines a potential energy surface. These surfaces are multi-dimensional for all but the simplest of molecules; therefore, they are impossible to visualize except in cases where the changes in energy are a result of modifications to only one or two coordinates, such as a rotation around one or two torsional angles, as seen in Figure 2.2. The peaks of a potential energy surface

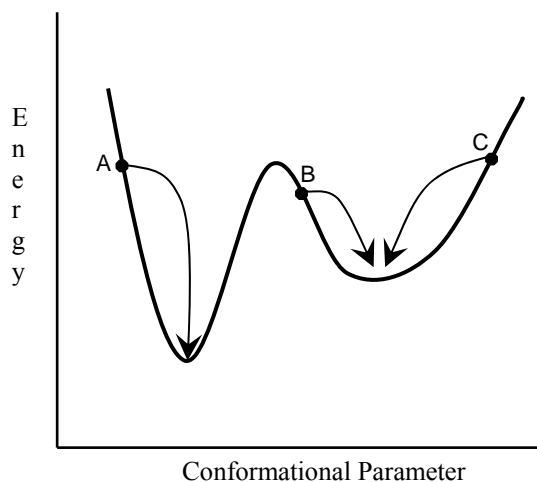
indicate transition structures while the valleys correspond to minimum energy structures. Multiple minima can exist on a surface with the very lowest energy point being the global energy minimum. All other minimum energy points are termed local energy minima. Any movement away from a minimum energy structure results in an increase in energy.



**Figure 2.2:** Potential energy surface for *o*-methoxy-phenol observed for the simultaneous rotation of the torsional angles of the hydroxyl and methoxy groups.<sup>16</sup>

To identify minimum energy structures, and therefore optimize a molecule's geometry, the coordinates of the atoms are gradually changed to produce configurations with lower and lower energies until a minimum is reached. Most minimization methods move down the potential energy surface and therefore find minima that are nearest to the starting structure (Figure 2.3).<sup>9</sup> To find multiple local minima, or to ensure that the global minimum energy structure was found, multiple starting structures must be generated and minimized. There are a variety of methods available to generate these starting structures and they are more thoroughly discussed in *Chapter 3*. These methods include molecular dynamics, which follows the time dependent behaviour of a system, dihedral angle driving, which rotates one or more torsional angles through a defined rotation, and the Monte Carlo methods, which randomly alter torsional angles within a molecule to generate new structures.



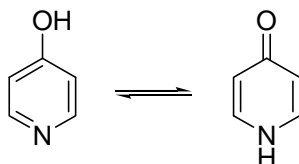


**Figure 2.3:** A schematic one-dimensional energy surface. Most minimization algorithms move downhill in energy towards the lowest energy structure. “A” would move towards the global energy minimum while “B” and “C” move towards a local energy minimum. Adapted from Leach.<sup>9</sup>

### 2.1.5. Gas Phase versus Solvation Modelling

The geometries and energies obtained during minimization will be more accurate if the structures are optimized in the phase in which they are normally found and studied. For example, to determine the behaviour of a molecule under physiological conditions, it is important to include water solvation. Calculations performed *in vacuo* treat the molecule as an isolated system and may not yield an accurate representation of the molecule’s behaviour in solution. This discrepancy is especially great when the molecule of interest contains highly polar or charged functional groups. The omission of solvation can exaggerate the importance of intramolecular interactions, resulting in a closed, folded structure that may not exist in solution. Water and other polar solvents can provide electrostatic shielding and competing hydrogen bonding interactions, resulting in a very different conformation than that observed in the gas phase.<sup>17</sup>

Solvation can also have a dramatic effect on reaction thermodynamics as it can change the stability of the reactants and products, altering equilibria. For example, the equilibrium of the tautomers of 4-hydroxypyridine and 4-pyridone (Figure 2.4) is dramatically affected by solvation as the pyridone is stabilized by high polarity solvents. Experimental data indicates that the equilibrium is dramatically shifted towards the pyridone in water as compared to the gas phase ( $K_{vapour} = <1/10$  vs.  $K_{aq.} = 2000/1$ ).<sup>10,18</sup> As demonstrated by this example, it is essential to perform thermodynamic calculations in the phase in which the reactions would occur to account for increased stabilization that the solvent may impart.

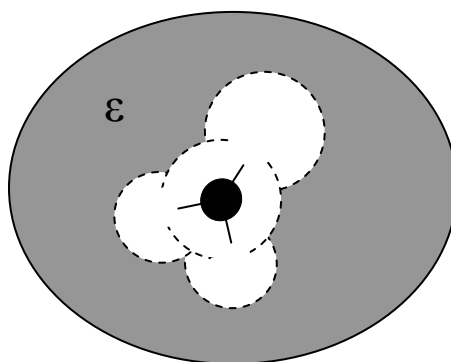


**Figure 2.4:** The equilibrium between 4-hydroxypyridine and 4-pyridone.

Multiple methods are available for the inclusion of solvation in molecular modelling. The most accurate method, explicit solvation, defines all of the solvent molecules involved in each solvation sphere individually to account for all possible hydrogen bonding and

intermolecular interactions which may affect the solute. The number of solvent molecules that must be defined increases exponentially with each solvation sphere included. The biggest disadvantage to this method is the large amount of computational resources and time required to complete such a calculation. Few researchers have access to the resources required to model such a system in a reasonable amount of time.<sup>10</sup>

To reduce the computational demands of solvation modelling *implicit solvation* methods were developed which treat the solvent as a continuum which surrounds the solute with a given dielectric (Figure 2.5). The main assumption of implicit solvation is that it is reasonable to remove the individual water molecules from the model as long as the space around the solute is modified to have properties consistent with those of the solvent.<sup>10</sup> While implicit solvation cannot account for specific hydrogen bonding interactions, it can provide electrostatic shielding to minimize the over-emphasis on certain intramolecular interactions. Due to the savings in computational resources and time involved, as well as the complexity of predicting the location of any explicitly defined solvent molecules, implicit methods are the most common choice for modelling solvation.



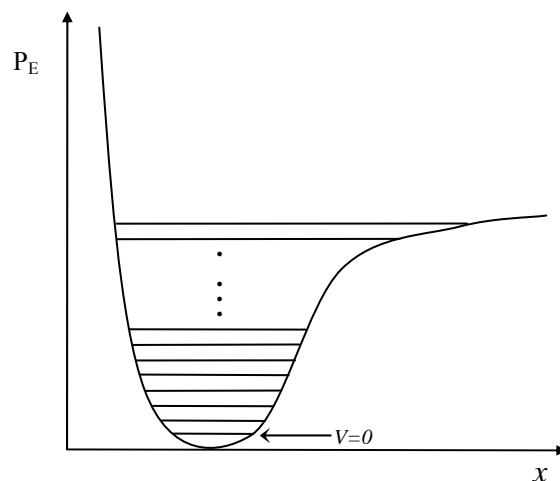
**Figure 2.5:** Schematic of the treatment of a solute in a liquid by the implicit water solvation Polarized Continuum Model (PCM). The solute cavity is defined as the union of a series of interlocking atomic spheres within a continuum of a uniform dielectric constant ( $\epsilon$ ). Adapted from Foresman and Frisch.<sup>19</sup>

#### 2.1.6. Zero Point Energy Correction

The energy determined during a geometry optimization is a minimum on the potential energy curve; however, it does not include the inherent vibrational energy of the system. Even at absolute 0 K, a molecule will have some vibrational energy, otherwise both its kinetic and potential energies would equal zero and the particle would always be at the origin, violating the Heisenberg uncertainty principle.\* The lowest vibrational mode a system can have,  $V = 0$ , is a non-zero value (Figure 2.6) termed the zero-point energy (ZPE). Many modelling programs calculate a ZPE correction which is then added to the total energy of the optimized structure. The ZPE for a single vibrational mode is quite small; however medium or large sized molecules can have multiple vibrational modes resulting in a substantial overall ZPE. In the case of 1,3-butadiene, there are 24 vibrational modes yielding a ZPE of  $\sim 50$  kcal/mol, a significant contribution to the overall energy of the molecule. This

\*  $\Delta x \Delta p_x \geq \frac{1}{2} (h/2\pi)$ , where  $x$  is a particle's  $x$  coordinate,  $\Delta p_x$  is the measure of the uncertainty of  $x$  and  $h$  is Planck's constant ( $= 6.62608 \times 10^{-34}$  J s).

correction is therefore essential for the accurate comparison of the energetics of conformations and isomers and should give results closer to experimental data.<sup>11,20</sup>



**Figure 2.6:** Vibrational modes of a system, where PE is the potential energy and x is the average distance between the nuclei involved in the vibration. The difference between the minimum energy and the  $V = 0$  level is the zero-point energy correction.

### 2.1.7. Vibrational Frequencies

To determine the legitimacy of a structure obtained through geometry optimization, the various vibrational frequencies of the resultant conformation are often calculated. If all of the vibrational frequencies are positive, the structure is likely a minimum energy conformation. This conformation may not be the global minimum structure, but rather a local minimum between two or more low energy structures. If there are negative or imaginary frequencies returned, the structure is possibly a transition structure and is not a minimum structure.<sup>10</sup>

### 2.1.8. Plan of Action

With the above considerations in mind, we undertook a detailed study of the thermodynamics of select biologically relevant reactions involving ESH and the OSHs. We wished to determine which combination of modelling method (HF or DFT) and basis set was best suited for studying ESH; therefore, we obtained a crystal structure of ESH to use both as a starting structure and as a benchmark for the accuracy of the modelling methods. We then performed geometry optimizations on ESH using either HF or DFT methods with a wide variety of basis sets either *in vacuo* or with implicit water solvation. The resultant minimum energy structures were compared to the crystal structure to determine which method and basis set combination gave the most accurate results. This combination was then used for all other calculations performed in this study. The calculated heats of formation were then used to determine the thermodynamics of a selection of the possible detoxification reactions of ESH and the OSHs.

## 2.2. Materials and Methods

ESH dihydrate was used as obtained from Sigma-Aldrich. For crystallization, Milli-Q quality water and HPLC grade ethanol (Fisher Scientific, Ottawa, ON) were used.

Preliminary calculations were performed using Jaguar 5.0 R.22 with structures drawn and visualized using Maestro 5.1, both of Schrödinger Inc. (Portland, OR). Detailed electronic structure calculations were performed using Gaussian 03, Revision B.05<sup>21</sup> (Gaussian Inc. Wallingford, CT) with structures drawn and visualized in GaussView 3.09 (Gaussian Inc.). All calculations were run on either a Silicon Graphics Inc. (SGI, Mountain View, CA) O<sub>2</sub> workstation (Violin) or Flexor, the University of Waterloo's multi-CPU SGI Origin 3800 system. Violin uses a MIPS R10000 processor chip, a MIPS R10010 floating point chip and a 195 MHz IP32 processor and has 750 MB of RAM. Violin runs the IRIX 3.5X operating system. Flexor consists of forty 400 MHz MIPS R12000 CPUs and twelve 500 MHz MIPS R14000 CPUs with 52 GB of RAM. The home and scratch directories reside on nine 36 GB fibre-channel disks in a RAID-5 configuration which is controlled by a TP9100 disk array with a 1 Gb hardware RAID controller. Third party software resides on a 36 GB fibre-channel disk outside of the RAID-5. Network connections include a 10/100 MB interface and a fibre-optic Gigabit Ethernet interface. Flexor runs the 64-bit IRIX 6.5.27f operating system.

### 2.2.1. Crystal Structure of Ergothioneine

ESH dihydrate was dissolved in a minimal amount of water and ethanol was added drop-wise until the solution became cloudy. This was allowed to sit at room temperature until crystals formed. The crystal was coated with a perfluoropolyether oil and frozen onto a glass fibre. The X-ray diffraction data was collected, and the structure solved by Dr. Nicholas J. Taylor of the University of Waterloo Chemistry Department. Data was collected on a Bruker Apex diffractometer at 150(1) K from a crystal with a size of 0.30{100} x 0.15{011} mm; from 9536 reflections with  $2.19 < \theta < 35.48$  collected and 4649 unmerged Friedels were used for structure analysis. The radiation source was a fine focus sealed tube,  $\lambda = 0.71073 \text{ \AA}$  and the radiation type was MoK $\alpha$ . Data was collected using SMART, the integration was performed using SAINT and the structure refinement was performed by and the graphics were prepared using SHELXTL, all of Bruker AXS (Karlsruhe, Germany) and all run on an NT4 system. Final R indices for all data were  $R_1 = 0.0438$  and  $wR_2 = 0.0731$ , the refinement was based on the  $wR(F^2)$  and the goodness of fit was 1.792. The absorption correction used face-indexed analytical integration with  $0.928 < T < 0.962$ . The protons were found and freely refined isotropically. The absolute configuration was determined by refinement of the Flack parameter.<sup>22</sup> The coordinates were deposited in the Cambridge Crystallographic Database, deposition FOLRII.

### 2.2.2. Determination of Appropriate Basis Set

A series of geometry optimizations of ESH were performed using the crystal structure as a starting conformation. The coordinates were imported into Maestro as a pdb file and the geometry was optimized by Jaguar using HF and DFT methods with the 3-21G,<sup>23-25</sup> 6-31G,<sup>26-31</sup> 6-311G,<sup>32-35</sup> cc-pVDZ<sup>36-38</sup> and cc-pVTZ(-f)<sup>36-38</sup> basis sets with the presence and absence of additional diffuse and polarization functions.<sup>†</sup> The DFT calculations used the B3LYP functional,<sup>13-15,39-42</sup> These calculations were performed both *in vacuo* and with the implicit water solvation model developed by Schrödinger Inc. for use with Jaguar. Vibrational

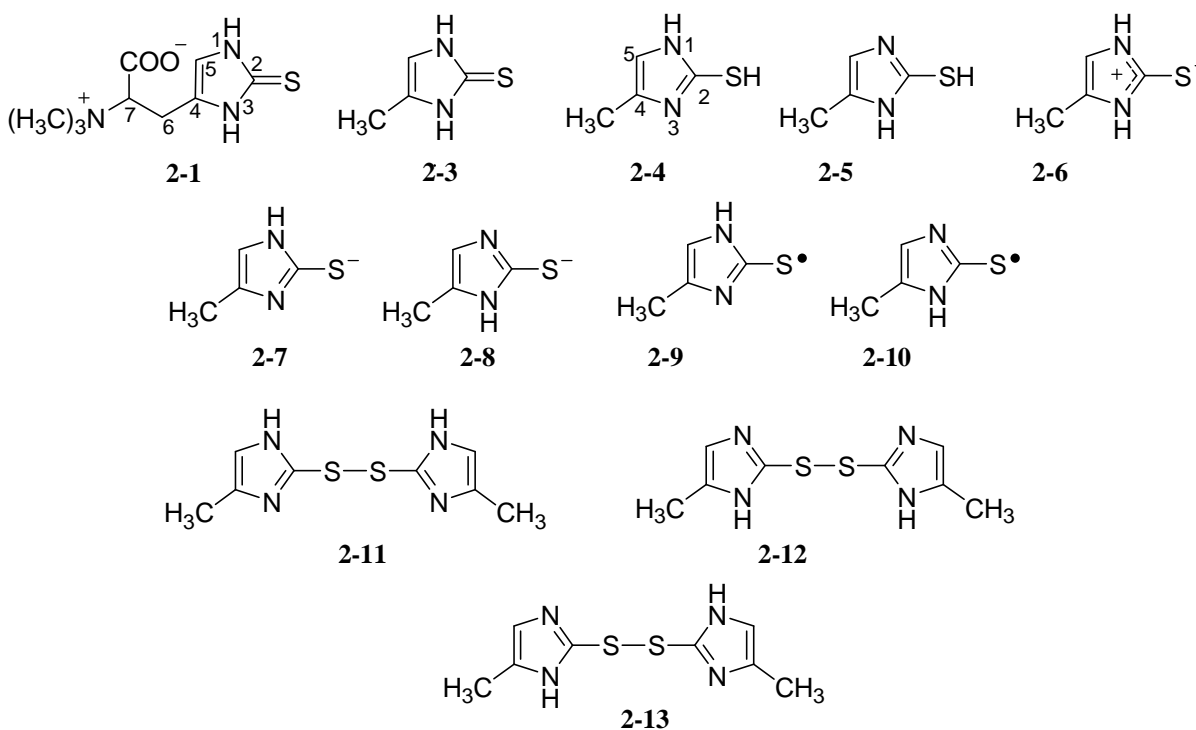
---

<sup>†</sup> In the case of 3-21G, polarization is only applied to the sulfur atom, in all other cases it is applied to all atoms of ESH. The cc-pVDZ and cc-pVTZ(-f) basis sets incorporate polarization and therefore calculations cannot be performed in the absence of polarization.

frequencies were calculated for all optimized structures; those that returned imaginary frequencies were discarded. In some cases, a minimum energy structure could not be obtained for a particular method and basis set combination. An optimized structure obtained using the same method but a simpler basis set was therefore used as a starting structure for the calculation. An optimization was also performed using DFT with the B3PW91<sup>13,39,43,44</sup> functional at the cc-PVTZ(-f) level both *in vacuo* and with implicit solvation to compare to the results of B3LYP. The bond distances and internal angles of the imidazole ring were used to compare the geometry optimized structures to the ESH crystal structure to determine the minimum basis set required for accurate geometry prediction.

### 2.2.3. Detailed Electronic Structural Calculations

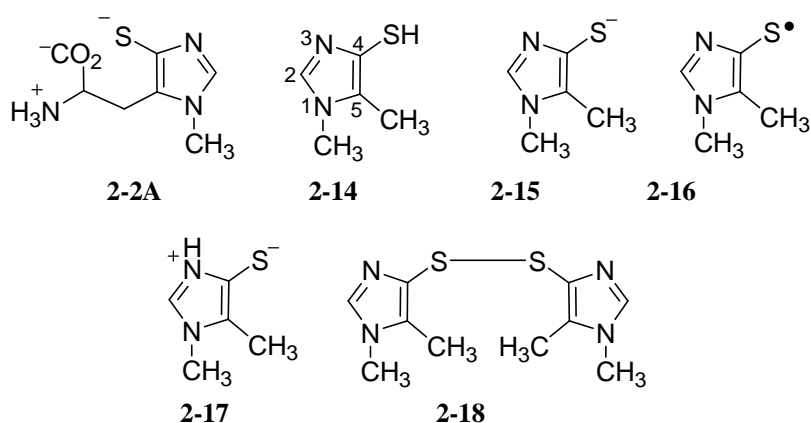
All structures were drawn and visualized in GaussView and all optimizations were performed *in vacuo* and with the IEF-PCM implicit solvation model<sup>45-47</sup> at the B3LYP<sup>14,15,40,42</sup>/6-311G<sup>34,35,48-54</sup> level with the addition of polarization and diffuse functions<sup>‡</sup> as implemented by Gaussian 03 (Gaussian Inc.). ESH was modelled using 2-thiol-4-methyl-imidazole (TMI, see *Results and Discussion*) in the various protonation states available to its thione, thiol, thiolate, thiyl radical and disulfide (TMI-S)<sub>2</sub> (Figure 2.7, (2-3) through (2-13)). Radicals were modelled using an unrestricted wavefunction.



**Figure 2.7:** Structures of ergothioneine and 2-thiol-4-methyl-imidazole and its disulfide.

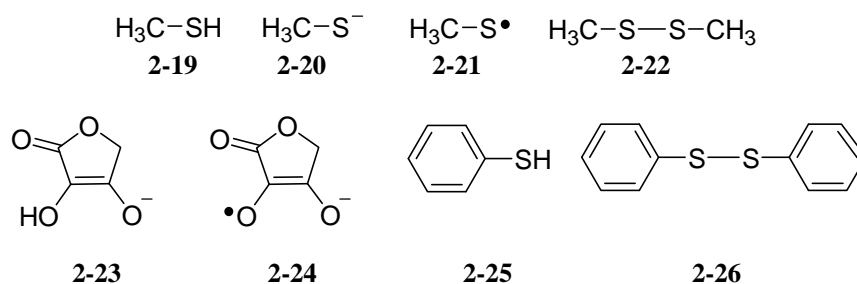
The OSHs were modelled by 4-thiol-*N*<sup>l</sup>-methyl-5-methyl-imidazole (TMMI, see *Results and Discussion*) as a thiol, thiolate, zwitterion, thiyl radical and disulfide (TMMI-S)<sub>2</sub> (Figure 2.8, (2-14) through (2-18)). Calculations were also performed on OSH<sub>A</sub> as a comparison (Figure 2.1, (2-2A)).

<sup>‡</sup> Denoted by 6-311++G(d,p) in Gaussian 03 notation (Frisch *et al.* (2003))



**Figure 2.8:** Structures of OSH<sub>A</sub> and 4-thiol-*N*<sup>1</sup>-methyl-5-methyl-imidazole used for modelling the OSH.

Certain biologically relevant molecules, such as H<sub>2</sub>O<sub>2</sub> and HO•, were studied to examine the thermodynamics of possible detoxification reactions involving ESH and the OSHs. GSH and its disulfide (GSSG) were modelled as methyl mercaptan and dimethyl disulfide (Figure 2.9, (2-19) through (2-22)), and ascorbic acid was modelled as 2-23 and 2-24 (Figure 2.9). To compare TMI and TMMI to standard aromatic thiols, thiophenol and diphenyl disulfide (Figure 2.9, (2-25) and (2-26)) were also modelled.



**Figure 2.9:** Miscellaneous compounds used to model biologically relevant reactions.

#### 2.2.4. Free Energy of Reaction Calculations

To determine the thermodynamics of a number of biologically relevant oxidation and reduction reactions, the sum of the thermal and electronic energies (ZPE corrected) were taken as  $\Delta G^\circ$  for each compound.

### 2.3. Results and Discussion

ESH is an important molecule in biological systems; however, there is very little known regarding the thermodynamics of the instability of its disulfide and the detoxification actions it is believed to perform. The placement of the sulfur atom on the imidazole ring of ESH differs from that of the OSHs, which is believed to lead to the difference in the stability of the two disulfides under physiological conditions. Both ESH and the OSHs have been suggested as therapeutic agents against the damage caused by oxidative stress. Our theoretical study into the thermodynamics of these compounds gives insight into the driving forces behind the differences in reactivity and the ability of these thiols to scavenge reactive oxygen species.

### 2.3.1. The Use of a Crystal Structure as a Starting Point for Optimization and a Potential Benchmark for Accuracy

Geometry optimizations usually move down the potential energy surface towards a minimum; therefore, the minimum energy structure determined somewhat resembles the starting structure.<sup>9</sup> The choice of the starting structure determines which local minimum energy structure is found by the minimization algorithm. As a starting point for our optimizations, we felt that the use of a crystal structure of ESH would be most appropriate as it would act as a pre-minimization, and might give a better orientation for the less rigid amino acid portion of ESH. In addition, we felt that using the crystal structure as a starting point might result in more accurate final geometries.

To evaluate the quality of the calculated geometries we used the crystal structure as a comparison point for the bond lengths, bond angles and torsional angles determined. This is a commonly used practice in the literature for determining the accuracy of the minimization method.<sup>55-59</sup> While a crystal structure of ESH had been previously determined by Sugihara *et al.*,<sup>60</sup> it included water molecules which, because of our use of implicit solvation, could not be included in our calculations. In addition, this crystal structure did not determine the absolute configuration of ESH. We therefore obtained a crystal structure which excluded water molecules and used this as our starting geometry and as a comparison for the accuracy of our calculated geometries.

### 2.3.2. Crystal Structure Determination of Ergothioneine

The crystals were grown optimally in water/ethanol and were orthorhombic and colourless and belong to the space group P2<sub>1</sub>2<sub>1</sub>2<sub>1</sub>. The unit cell had the dimensions of a = 6.0885(2) Å, b = 12.4212(5) Å and c = 13.9867(6) Å, a volume of 1057.76(7) Å<sup>3</sup> and Z = 4. The C<sup>2</sup>-S bond length was determined to be 1.69 Å, intermediate to the C<sup>2</sup>-S single and double bond lengths of 1.82 and 1.56 Å respectively, which is consistent with a substantial contribution of the thione form to the tautomeric equilibrium (Table 2.2). This result is in agreement with the Raman and <sup>13</sup>C-NMR spectral data which indicate that ESH exists predominantly as the thione under physiological conditions.<sup>61</sup> In addition, our crystal structure is in very good agreement with a previously published structure obtained for ESH crystals grown from water<sup>60</sup> (Table 2.2 through Table 2.7). All other crystallographic data can be found in *Appendix 1*

**Table 2.2:** Bond lengths (in Å) observed in the crystal structures of ESH, part A.

<i>Conditions</i>	C <sup>2</sup> -S	N <sup>1</sup> -C <sup>2</sup>	C <sup>2</sup> -N <sup>3</sup>	N <sup>1</sup> -C <sup>5</sup>	N <sup>3</sup> -C <sup>4</sup>	C <sup>4</sup> -C <sup>5</sup>	C <sup>4</sup> -C <sup>6</sup>	C <sup>6</sup> -C <sup>7</sup>
Water/Ethanol <sup>a</sup>	1.69	1.35	1.35	1.38	1.39	1.35	1.50	1.52
Water only <sup>b</sup>	1.69	1.35	1.35	1.38	1.38	1.35	1.49	1.53

<sup>a</sup>this work; <sup>b</sup>Sugihara *et al.*<sup>60</sup>

**Table 2.3:** Bond lengths (in Å) observed in the crystal structures of ESH, part B.

<i>Conditions</i>	C <sup>7</sup> -N <sup>11</sup>	C <sup>7</sup> -C <sup>8</sup>	C <sup>8</sup> -C <sup>10</sup>	C <sup>8</sup> -O <sup>9</sup>	N <sup>11</sup> -C <sup>12</sup>	N <sup>11</sup> -C <sup>13</sup>	N <sup>11</sup> -C <sup>14</sup>
Water/Ethanol <sup>a</sup>	1.54	1.55	1.25	1.25	1.50	1.50	1.51
Water only <sup>b</sup>	1.54	1.56	1.23	1.26	1.51	1.49	1.50

<sup>a</sup>this work; <sup>b</sup>Sugihara *et al.*<sup>60</sup>

**Table 2.4:** Bond angles (in degrees) observed in the crystal structures of ESH, part A.

<i>Conditions</i>	$S-C^2-N^1$	$S-C^2-N^3$	$C^2-N^3-C^4$	$C^2-N^1-C^5$	$N^1-C^2-N^3$	$N^3-C^4-C^5$	$N^3-C^4-C^6$	$N^1-C^5-C^4$	$N^3-C^4-C^6$	$C^5-C^4-C^6$	$C^4-C^6-C^7$
Water/Ethanol <sup>a</sup>	126.6	127.9	110.4	110.3	105.6	106.5	123.7	107.2	123.7	129.8	112.4
Water only <sup>b</sup>	128.0	126.6	110.3	110.8	105.4	107.1	121.9	106.4	121.9		109.7

<sup>a</sup>this work; <sup>b</sup>Sugihara *et al.*<sup>60</sup>**Table 2.5:** Bond angles (in degrees) observed in the crystal structures of ESH, part B.

<i>Conditions</i>	$C^6-C^7-C^8$	$C^6-C^7-N^{11}$	$C^7-C^8-O^9$	$C^7-C^8-O^{10}$	$C^7-N^{11}-C^{12}$	$C^7-N^{11}-C^{13}$	$C^7-N^{11}-C^{14}$	$C^{14}-N^{11}-C^{13}$	$C^{14}-N^{11}-C^{12}$	$C^{13}-N^{11}-C^{12}$
Water/Ethanol <sup>a</sup>	111.4	110.8	115.5	116.3	109.7	108.6	113.4	108.0	108.4	108.7
Water only <sup>b</sup>	109.9	112.2	115.9	118.0	108.9	109.5	111.5	109.6	108.8	108.4

<sup>a</sup>this work; <sup>b</sup>Sugihara *et al.*<sup>60</sup>**Table 2.6:** Torsional angles (in degrees) observed in the crystal structures of ESH, part A.

<i>Conditions</i>	$S-C^2-N^3-C^4$	$S-C^2-N^1-C^5$	$C^2-N^3-C^4-C^5$	$N^3-C^4-C^5-N^1$	$C^4-C^5-N^1-C^2$	$N^3-C^4-C^6-C^7$	$C^2-N^3-C^4-C^6$	$N^1-C^5-C^4-C^6$	$C^5-C^4-N^1-C^2$	$C^4-C^6-C^7-C^8$	$C^4-C^6-C^7-N^{11}$
Water/Ethanol <sup>a</sup>	-177.9	178.3	-0.9	0.5	0.1	84.6	176.5	-176.7	-98.7	55.9	179.9
Water only <sup>b</sup>	-179.7	-179.7	-1.0	0.8	0.1	79.7	-179.0	178.5	-97.7	67.4	-168.6

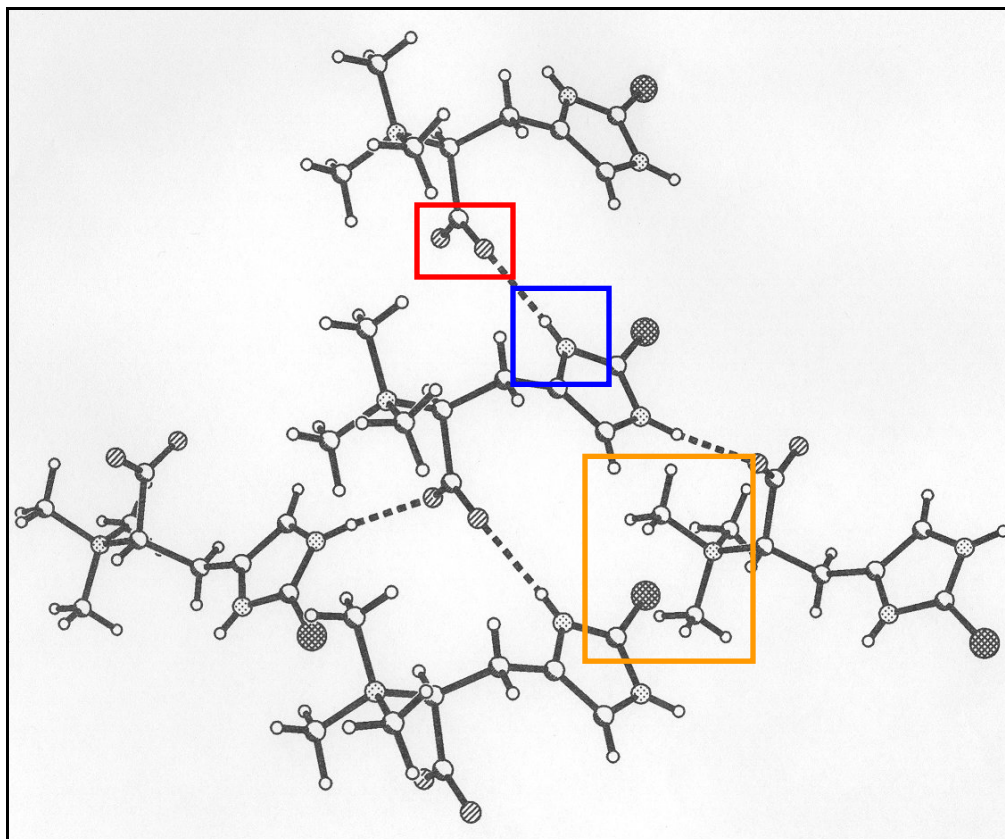
<sup>a</sup>this work; <sup>b</sup>Sugihara *et al.*<sup>60</sup>**Table 2.7:** Torsional angles (in degrees) observed in the crystal structures of ESH, part B.

<i>Conditions</i>	$C^6-C^7-C^8-O^{10}$	$C^6-C^7-C^8-O^9$	$C^6-C^7-N^{11}-C^{12}$	$C^6-C^7-N^{11}-C^{13}$	$C^6-C^7-N^{11}-C^{14}$	$C^8-C^7-N^{11}-C^{12}$	$C^8-C^7-N^{11}-C^{13}$	$C^8-C^7-N^{11}-C^{14}$	$O^9-C^8-C^7-N^{11}$	$O^{10}-C^8-C^7-N^{11}$
Water/Ethanol <sup>a</sup>	50.1	-128.6	70.4	-171.0	-50.9	-165.3	-46.7	73.4	107.5	-73.8
Water only <sup>b</sup>	46.4	-132.4	69.3	-172.3	-50.8	-167.3	-48.9	72.6	102.9	-78.3

<sup>a</sup>this work; <sup>b</sup>Sugihara *et al.*<sup>60</sup>



A packing diagram of the ESH crystal (Figure 2.10) indicates that each oxygen atom of the carboxylate moiety forms a hydrogen bond to the N-H of the imidazole ring of an adjacent ESH. Each ESH molecule is therefore involved in four hydrogen bonds: one for each N-H and one for each carboxylate oxygen. The sulfur atom is found packed in close proximity to the methyl groups of the quaternary ammonium group and does not appear to interact in any hydrogen bonding arrangement.



**Figure 2.10:** Crystal packing diagram of ergothioneine crystals grown from water/ethanol. The carboxyl oxygens (red) hydrogen bond to an adjacent N-H (blue) from the imidazole ring while the sulfur atom is packed in close proximity to the methyl groups of the quaternary ammonium group (orange).

### 2.3.3. Determination of Appropriate Basis Set

As mentioned, the geometries produced by minimization methods are often compared to crystal structure data to confirm the accuracy of the computed structure.<sup>55-57,59</sup> To determine which method and basis set combination was best suited for the modelling of ESH and the OSHs, the optimized structures of ESH obtained from the HF and DFT calculations were compared to the crystal structure.<sup>§</sup> A comparison of the DFT and HF methods using implicit water solvation indicated that the DFT calculations yielded structures in better agreement to the crystal structure (see *Appendix I*). A comparison of the implicitly solvated structures found using the B3LYP and B3PW91 functionals were very similar, with the B3LYP functional yielding results that were slightly closer to the crystal structure (Table 2.8,

<sup>§</sup> All comparisons were performed using the crystal structure determined in this work.

Table 2.9, and *Appendix I*). The B3LYP functional was therefore used for all further calculations. The addition of polarization functions to the basis set was important to obtain the correct C-S bond length as the absence of polarization functions resulted in a longer bond. The inclusion of diffuse functions did not have a large effect on the computed structures (see Table 2.8, Table 2.9 and *Appendix I*). Diffuse functions will be important for later modelling the thiolate moieties of ESH and the OSHs as they improve the description of the wavefunction of anionic species (see 2.1.2. *Basis Sets and Functions* for more details).

**Table 2.8:** Bond lengths (in Å) returned after the geometry optimization of the ergothioneine crystal structure using DFT methods with implicit solvation.<sup>a</sup>

<i>Basis set</i>	**	++	<i>C</i> <sup>2</sup> - <i>S</i>	<i>N</i> <sup>1</sup> - <i>C</i> <sup>2</sup>	<i>C</i> <sup>2</sup> - <i>N</i> <sup>3</sup>	<i>N</i> <sup>1</sup> - <i>C</i> <sup>5</sup>	<i>N</i> <sup>3</sup> - <i>C</i> <sup>4</sup>	<i>C</i> <sup>4</sup> - <i>C</i> <sup>5</sup>	<i>C</i> <sup>4</sup> - <i>C</i> <sup>6</sup>
<b>B3LYP</b>									
6-31G	**	++	1.71	1.36	1.36	1.39	1.39	1.36	1.50
	**	None	1.71	1.36	1.36	1.39	1.39	1.36	1.50
6-311G	None	None	1.75	1.36	1.36	1.40	1.40	1.37	1.50
	**	++ <sup>b</sup>	1.70	1.35	1.35	1.38	1.39	1.36	1.50
	**	None <sup>c</sup>	1.71	1.35	1.35	1.40	1.40	1.37	1.50
cc-pVDZ	None	None	1.75	1.36	1.36	1.39	1.40	1.36	1.50
	Incl.	++	ND <sup>d</sup>	ND	ND	ND	ND	ND	ND
cc-pVTZ(-f)	Incl.	None	1.70	1.35	1.36	1.38	1.39	1.36	1.49
	Incl.	++	1.70	1.35	1.35	1.38	1.39	1.35	1.49
	Incl.	None	1.71	1.35	1.35	1.38	1.40	1.36	1.49
<b>B3PW91</b>									
cc-pVTZ(-f)	Incl.	++	1.70	1.35	1.35	1.38	1.39	1.35	1.49
<b>Crystal<sup>a</sup></b>	--	--	1.69	1.35	1.35	1.38	1.39	1.35	1.50

<sup>a</sup>The crystal structure obtained in this work was used as the starting structure and the implicit solvation was applied as implemented by Jaguar; <sup>b</sup>The optimized structure obtained using HF/6-311++G\*\* methods and implicit water solvation was used as the starting structure; <sup>c</sup>The optimized structure obtained using HF/6-311G\*\* methods and implicit water solvation was used as the starting structure; <sup>d</sup>Not determined due to difficulties in convergence.

**Table 2.9:** Bond angles (in degrees) returned after the geometry optimization of the ergothioneine crystal structure using DFT methods with implicit water solvation.<sup>a</sup>

<i>Basis set</i>	**	++	<i>S-C<sup>2</sup>-N<sup>1</sup></i>	<i>S-C<sup>2</sup>-N<sup>3</sup></i>	<i>C<sup>2</sup>-N<sup>3</sup>-C<sup>4</sup></i>	<i>C<sup>2</sup>-N<sup>1</sup>-C<sup>5</sup></i>	<i>N<sup>1</sup>-C<sup>2</sup>-N<sup>3</sup></i>	<i>N<sup>3</sup>-C<sup>4</sup>-C<sup>5</sup></i>	<i>N<sup>1</sup>-C<sup>5</sup>-C<sup>4</sup></i>	<i>N<sup>3</sup>-C<sup>4</sup>-C<sup>6</sup></i>
<b>B3LYP</b>										
6-31G	**	++	127.57	127.18	111.06	110.61	105.24	105.88	107.20	123.14
	**	None	127.68	127.30	111.24	110.77	105.02	105.80	107.18	122.50
6-311G	None	None	127.42	127.33	110.98	110.68	105.25	105.88	107.21	122.81
	**	++ <sup>b</sup>	127.15	127.46	110.96	110.57	105.36	105.86	107.23	123.04
	**	None <sup>c</sup>	127.22	127.38	110.95	110.58	105.39	105.93	107.14	122.76
cc-pVDZ	None	None	127.22	127.29	110.77	110.55	105.49	106.02	107.17	123.13
	Incl.	++	ND <sup>d</sup>	ND	ND	ND	ND	ND	ND	ND
cc-pVTZ(-f)	Incl.	None	127.62	127.39	111.21	110.76	104.99	105.74	107.30	122.87
	Incl.	++	127.67	126.98	110.89	110.48	105.36	106.00	107.27	123.12
	Incl.	None	127.68	127.15	111.02	110.67	105.17	105.93	107.22	122.56
<b>B3PW91</b>										
cc-pVTZ(-f)	Incl.	None	128.81	127.97	111.92	111.71	103.21	105.94	107.21	123.25
<b>Crystal<sup>a</sup></b>	--	--	126.55	127.89	110.41	110.34	105.55	106.48	107.22	123.71

<sup>a</sup>The crystal structure obtained in this work was used as the starting structure and the implicit solvation was applied as implemented by Jaguar; <sup>b</sup>The optimized structure obtained using HF/6-311++G\*\* methods and implicit water solvation was used as the starting structure; <sup>c</sup>The optimized structure obtained using HF/6-311G\*\* methods and implicit water solvation was used as the starting structure.; <sup>d</sup>Optimization using this basis set was not successful.

Once polarization and diffuse functions were included, all four basis sets produced results in fairly good agreement with the crystal structure for the bond lengths and angles of ESH. When the examination was expanded to include the dihedral angles of ESH found with implicit solvation, it was noted that the cc-pVTZ(-f) and 6-311++G\*\* basis sets returned geometries which were in better agreement with the ESH crystal structure (Table 2.10).

**Table 2.10:** Selected dihedral angles (in degrees) returned after the geometry optimization of the ergothioneine crystal structure using DFT methods with implicit water solvation.<sup>a</sup>

<i>Basis set</i>	**	++	$C^5-C^4-C^6-C^7$	$C^4-C^6-C^7-N^{11}$
<b>B3LYP</b>				
6-31G	**	++	-101.0	178.0
	**	None	-80.3	-174.9
	None	None	-83.7	-176.6
6-311G	**	++ <sup>b</sup>	-85.0	179.8
	**	None <sup>c</sup>	-79.5	-177.0
	None	None	-101.5	175.4
cc-pVDZ	Incl.	++	ND <sup>d</sup>	ND
	Incl.	None	-93.8	-179.6
cc-pVTZ(-f)	Incl.	++	-101.0	176.2
	Incl.	None	-92.6	-173.3
<b>Crystal<sup>a</sup></b>	--	--	-98.7	179.9

<sup>a</sup>The crystal structure obtained in this work was used as the starting structure and implicit solvation was applied as implemented by Jaguar; <sup>b</sup>The optimized structure obtained using HF/6-311++G\*\* methods and implicit water solvation was used as the starting structure; <sup>c</sup>The optimized structure obtained using HF/6-311G\*\* methods and implicit water solvation was used as the starting structure; <sup>d</sup>Optimization using this basis set was not successful.

While the cc-pVTZ(-f) basis set gave results which were closer to the crystal structure than the 6-311++G\*\* basis set, it uses more basis functions in its calculations (Table 2.11). An increase in the number of basis functions exponentially increases the time and computational resources required for the calculations.<sup>11</sup> These effects are clearly demonstrated by the time required to complete a geometry optimization of phenol (Table 2.12). It was felt that the slight increase in accuracy obtained by the use of the cc-pVTZ(-f) basis set did not warrant the increased time and computational demands caused by the larger number of basis functions. It was concluded that given the large number of calculations to be performed, the B3LYP/6-311++G\*\* combination gave adequate accuracy with an acceptable use of time and computational resources and hence was used for all subsequent calculations.

The reactions in which we are interested occur under physiological conditions and are therefore in an aqueous environment. As demonstrated by the equilibrium of 4-hydroxypyridine and 4-pyridone (see *Section 2.1.5. Gas Phase versus Solvation Modelling*), the effect of solvation on polar compounds can alter the stability of reactants and products, affecting the overall thermodynamics of reactions and equilibria. To exclude solvation in our study would likely invalidate our results due to the polarity of the compounds involved. All further discussion therefore focuses on the results obtained using implicit solvation, as implemented by Jaguar. Gas phase results can be found in *Appendix 1*.

**Table 2.11:** The number of functions in each basis set for the geometry optimization of ergothioneine using the crystal structure as the starting conformation.

<i>Basis Set</i>	<i>Number of Functions</i>		
	<i>Alone<sup>a</sup></i>	<i>**<sup>b</sup></i>	<i>**++<sup>b,c</sup></i>
3-21G <sup>d</sup>	169	175	250
6-31G	169	304	379
6-311G	248	368	443
cc-pVDZ	-- <sup>e</sup>	289	484
cc-pVTZ(-f)	-- <sup>e</sup>	484	679

<sup>a</sup>No polarization or diffuse functions; <sup>b</sup>with polarization; <sup>c</sup>with diffuse functions; <sup>d</sup>polarization is only applied to the sulfur atom; <sup>e</sup>polarization is included in these basis sets.

**Table 2.12:** The number of functions in each basis set and time required for the geometry optimization of phenol.<sup>a</sup>

<i>Basis Set</i>	<i>Time Required</i>
6-311++G**	1790 s
cc-pVTZ(-f)++	6196 s

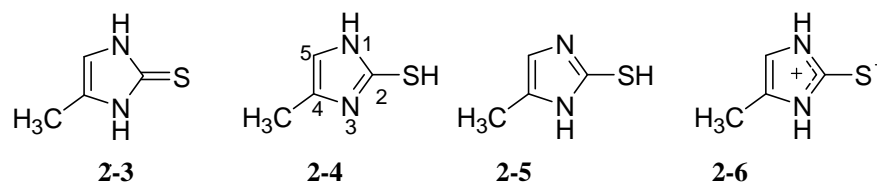
<sup>a</sup>Performed using DFT methods and the B3LYP functional with implicit solvation as implemented by Jaguar.

#### 2.3.4. Use of Truncated Analogues

The structures of ESH and the OSHs can be divided into the amino acid backbone and the thioimidazole side chain. Studies have shown that the amino acid portion of ESH does not affect its biological activity as 2-thiol-imidazole behaves in the same manner.<sup>5</sup> Similarly, the OSHs are often studied using truncated analogues.<sup>8,62-64</sup> Based upon this precedence, we used truncated versions of ESH and the OSHs in our studies. ESH was represented by TMI and the OSHs by TMMI. The use of these simpler structures decreases the number of degrees of freedom available to these compounds, simplifying the calculation, and ensures that the energies determined are related to the reactive portion of the molecule and not to the rather floppy backbone. In addition, the use of TMMI eliminates any confusion caused by the variation in methylation of the amino group of the OSHs (see Figure 2.1).

#### 2.3.5. Modelling 2-Thiol-4-Methyl-Imidazole and its Disulfide

ESH and by extension, TMI, can exist in multiple protonation states, with the ESH<sub>thione</sub> being observed under physiological conditions. To obtain a complete picture of the stability of ESH, four structures of TMI (Figure 2.11) were modelled. TMI<sub>thione</sub> (**2-3**) and its charge separated resonance structure (**2-6**) were found to be the most thermodynamically stable forms of TMI, and were closer in structure to the crystal and geometry optimized structures of ESH (Table 2.13). The results found for **2-3** and **2-6** were essentially identical, which is expected as *ab initio* calculations would treat them as the same structure. During an *ab initio* calculation, the overall charge is not localized to a specific atom but is applied to the whole molecule. The tautomeric forms **2-4** and **2-5** do not yield identical results due to the difference in protonation which slightly alters the calculated bond angles and bond lengths.



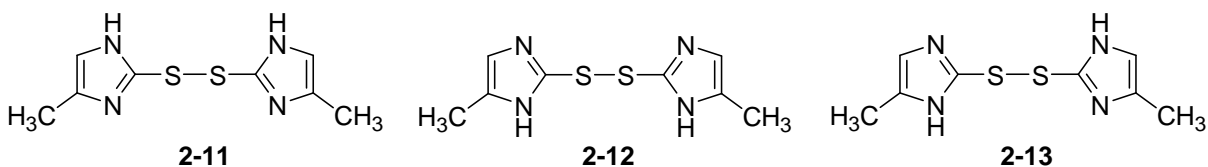
**Figure 2.11:** To simplify the calculations of ergothioneine, the amino acid portion was omitted leaving the ESH analogue 2-thiol-4-methyl-imidazole. Structures **2-3** and **2-6** are resonance structures, which would be in equilibrium with their tautomers **2-4** and **2-5**.

**Table 2.13:** Calculated energies, bond lengths and angles of 2-thiol-4-methyl-imidazole, its tautomers and ergothioneine.

<i>Tautomer</i>	<i>Energy (kcal/mol)</i>	$\Delta E$ (kcal/mol) <sup>a</sup>	$N^1-C^2-N^3$ <sup>b</sup>	$N^1-C^2$	$C^2-N^3$	$C^2-S$
ESH crystal <sup>c</sup>	--	--	105.5	1.38	1.39	1.69
ESH <sup>d</sup>	-668261.72	--	104.7	1.36	1.40	1.68
<b>2-3</b>	-416545.39	+ 00.02	104.7	1.35	1.36	1.71
<b>2-4</b>	-416532.12	+ 13.27	111.8	1.36	1.32	1.77
<b>2-5</b>	-416531.86	+ 13.27	110.0	1.32	1.37	1.77
<b>2-6</b>	-416545.41	0	104.7	1.35	1.36	1.71

<sup>a</sup>Energy difference between the lowest energy structure (2-6) and the other form studied; <sup>b</sup>the angles in °, bond lengths in Å; <sup>c</sup>Crystal structure determined in this work; <sup>d</sup>Geometry optimized structure.

The disulfide of ESH, and by extension TMI, can exist in multiple protonation states (Figure 2.12). Regardless of the protonation state, ESSE is unstable under physiological conditions<sup>1</sup> and ESH is not oxidized by conditions that readily oxidize GSH.<sup>1</sup> ESSE has been observed as a transient intermediate in redox reactions *in vitro* and may occur *in vivo* but is decomposed too quickly to be seen.<sup>3</sup> This decomposition may be caused by endogenous levels of GSH or mycothiol, which are found in ESH containing organisms. GSH is more easily oxidized than ESH and may act as a reducing agent for ESSE. To determine the thermodynamic driving force behind this instability the thermodynamics of the three (TMI-S)<sub>2</sub> tautomers (Figure 2.12, (**2-11** through **2-13**)) and their reduction by methylmercaptan, a model for GSH, were determined.



**Figure 2.12:** The tautomers of the disulfide of 2-thiol-4-methyl-imidazole

The optimized structures of TMI disulfides **2-11** and **2-12** differed in free energy by only 0.05 kcal/mol (Table 2.14). The bond lengths and angles of these disulfides were closer to those determined for the thiol form of TMI rather than that of the thione (see *Appendix 1*). Unfortunately, the optimization of the geometry of the TMI disulfide **2-13** was not successful and this structure is omitted from any further discussion.

**Table 2.14:** Energies and selected bond lengths and angles for the tautomers of the disulfide of 2-thiol-4-methyl-imidazole.

<i>Form</i>	<i>Energy</i> <sup>a</sup>	<i>Ring A/Ring B</i>				<i>S-S</i>	$\overset{A2}{C}-\overset{A}{S}-\overset{B}{C}-\overset{B2}{S}$
		$N^1-C^2-N^3$ <sup>b</sup>	$N^1-C^2$ <sup>c</sup>	$C^2-N^3$	$C^2-S$		
<b>2-11</b>	-832299.34	111.8/110.9	1.36/1.36	1.32/1.76	1.77/1.76	2.15	83.8
<b>2-12</b>	ND <sup>d</sup>	ND	ND	ND	ND	ND	ND
<b>2-13</b>	-832299.29	111.7/110.9	1.36/1.33	1.32/1.36	1.77/1.76	2.15	-84.0

<sup>a</sup>Energy in kcal/mol; <sup>b</sup>Angles in °; <sup>c</sup>Bond lengths in Å; <sup>d</sup>ND = not determined.

The reduction of ESSE by GSH is predicted to be favourable based upon the oxidation potentials of both compounds: ESH:  $E_o' = -0.06V$ <sup>65</sup> and GSH:  $E_o' = -0.25V$ .<sup>66</sup> The thermodynamics of the reduction to ESH<sub>thione</sub> were estimated using (TMI-S)<sub>2</sub> and methyl mercaptan, as shown in Eq. 2.1.



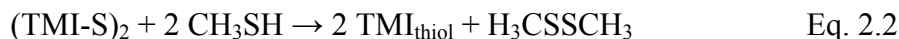
This reaction was favourable (Table 2.15), indicating that the reduction of ESSE by GSH is likely thermodynamically favourable. Interestingly, the free energy released by the reduction of (TMI-S)<sub>2</sub> to TMI<sub>thione</sub> is much greater than the energy released by the reduction of diphenyl disulfide (Table 2.15), indicating that TMI<sub>thione</sub> differs from a regular sulfhydryl and perhaps the thione character of the C-S bond plays a role in the thermodynamics.

**Table 2.15:** (TMI-S)<sub>2</sub> + 2 CH<sub>3</sub>SH → 2 TMI<sub>thione</sub> + H<sub>3</sub>CSSCH<sub>3</sub> (Eq. 2.1).

<i>Reactant</i>	<i>Product</i>	$\Delta G^\circ$ (kcal/mol)
<b>2-11</b>	2 <b>2-3</b>	-30.32
<b>2-12</b>	2 <b>2-3</b>	ND <sup>a</sup>
<b>2-13</b>	2 <b>2-3</b>	-30.37
Diphenyldisulfide	Thiophenol	-3.60

<sup>a</sup>Not determined

To determine the influence of the tautomerization of TMI<sub>thiol</sub> to TMI<sub>thione</sub> on free energy, Eq. 2.1 was broken down into its component steps: the reduction of the disulfide to the thiol followed by tautomerization (Eq. 2.2 and 2.3).



The free energy values for the reduction of (TMI-S)<sub>2</sub> to TMI<sub>thiol</sub> were found to be comparable to the reduction of diphenyldisulfide, which was used as a model aromatic thiol (Table 2.16) The tautomerization to TMI<sub>thione</sub> must therefore be a large contributor to the overall thermodynamics of the reaction.

**Table 2.16:**  $(\text{TMI-S})_2 + 2 \text{CH}_3\text{SH} \rightarrow 2 \text{TMI}_{\text{thiol}} + \text{H}_3\text{CSSCH}_3$  (Eq. 2.2)

Reactant	Product	$\Delta G^\circ$ (kcal/mol)
2-11	2 2-4	-3.78
2-12	2 2-5	ND <sup>a</sup>
2-13	2-4 + 2-5	3.47
Diphenyldisulfide	Thiophenol	-3.60

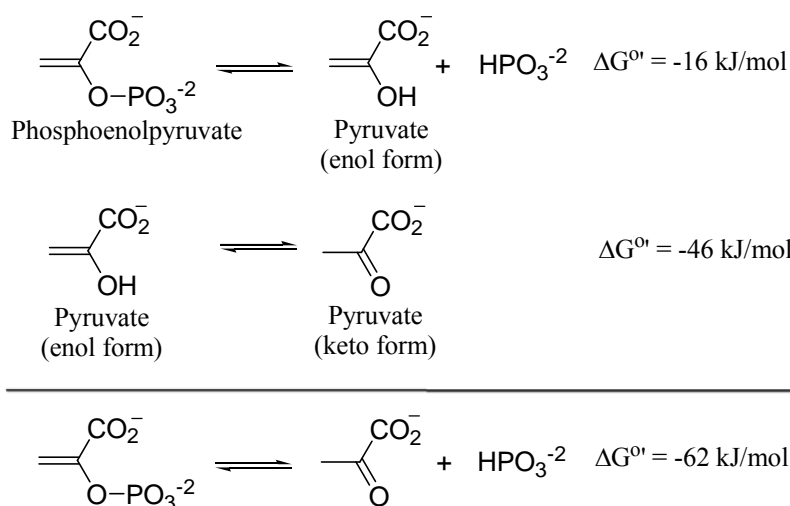
<sup>a</sup>Not determined

The conversion of the  $\text{TMI}_{\text{thiol}}$  to  $\text{TMI}_{\text{thione}}$  was found to be extremely favourable (Table 2.17) and accounts for over 85% of the energy released in Eq. 2.1 (Table 2.15). The free energy release found with implicit solvation was almost double that released in the gas phase (see *Appendix I*), indicating that solvation may play a role in the stabilization of the thione.

**Table 2.17:**  $2 \text{TMI}_{\text{thiol}} \rightarrow 2 \text{TMI}_{\text{thione}}$  (Eq. 2.3).

Reactant	Product	$\Delta G^\circ$ (kcal/mol)
2 2-4	2 2-3	-26.53
2 2-5	2 2-3	-27.06
2-4 + 2-5	2 2-3	-26.80

The use of tautomerization as a driving force for a reaction in biological systems is not isolated to the reduction of ESSE. In the hydrolysis of phosphoenolpyruvate to the keto form of pyruvate, the tautomerization of enol pyruvate contributes almost 75 % of the overall free energy of hydrolysis (Figure 2.13).<sup>67</sup>



**Figure 2.13:** The hydrolysis of phosphoenolpyruvate. The reaction is broken down into two steps, hydrolysis and tautomerization. The tautomerization of enol pyruvate to pyruvate is the major contributor to the overall  $\Delta G^\circ$  of hydrolysis.<sup>67</sup>

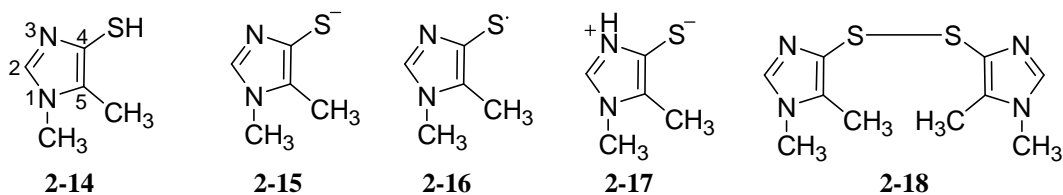
The thermodynamics of the reduction of  $(\text{TMI-S})_2$  indicate that the thermodynamic instability of ESSE may be due to the free energy released by the tautomerization of  $\text{ESH}_{\text{thiol}}$  to  $\text{ESH}_{\text{thione}}$ . Should ESSE be formed by redox reactions in the cell, it would likely be reduced by endogenous GSH or mycothiol. This reduction could account for the inability to



detect ESSE under physiological conditions and the absence of an ESSE reductase in the ESSE containing organisms.

### 2.3.6. Modelling 4-Thiol-*N*<sup>1</sup>-Methyl-5-Methyl-Imidazole and its Disulfide

The OSHs differ from ESH by the placement of a sulfur atom. This variation dramatically alters the properties of the OSHs in comparison to ESH. The OSHs exist mainly in the zwitterionic state, as  $pK_{aNH} = 10.3$  and  $pK_{aSH} = 2.6$ , with no reported presence of a thione moiety.<sup>62</sup> Geometry optimizations were performed on the model compound TMMI in the uncharged, anionic and zwitterionic forms (Figure 2.14). The fully protonated cationic form would only be present at very acidic pH and therefore was not included in this study. The zwitterionic form was found to be the most stable thermodynamically, in agreement with the  $pK_a$  values of the OSHs. The calculated C-S bond length was consistent with the calculated C-S bond length of thiophenol, indicating the presence of a thiol rather than a thione functionality (Table 2.18).



**Figure 2.14:** 4-Thiol-*N*<sup>1</sup>-methyl-5-methyl-imidazole and its various forms and disulfide.

**Table 2.18:** A comparison of the C-S bond lengths of the various ionic states of 4-thiol-*N*<sup>1</sup>-methyl-5-methyl-imidazole to thiophenol determined with implicit water solvation using the B3LYP/6-311++G\*\* combination

Species	C-S (Å)
2-14	1.78
2-15	1.77
2-16	1.74
Thiophenol	1.80

Also unlike ESH, the OSH disulfide has been detected *in vivo*,<sup>68-70</sup> suggesting that it may be more stable than ESSE. In sea urchin eggs OSSO is formed during the detoxification of H<sub>2</sub>O<sub>2</sub>,<sup>6,71</sup> however, there have been no reports of an OSSO reductase to reduce OSSO to its biologically useful free thiol form. In lieu of an OSSO reductase, it has been proposed that GSH may act as an endogenous reducing agent. To assess the thermodynamics of OSSO reduction by GSH, the disulfide of TMMI, (TMMI-S)<sub>2</sub> was modelled and the thermodynamics of its reaction with methyl mercaptan determined (Eq. 2.4). There are multiple protonation states possible for (TMMI-S)<sub>2</sub>; however, the methylation of TMMI to *S*-methyl-TMMI decreases the  $pK_a$  of N<sup>3</sup>H to ~6, similar to that of imidazole.<sup>62</sup> It is likely that the formation of (TMMI-S)<sub>2</sub> would have the same effect on N<sup>3</sup>H and therefore (TMMI-S)<sub>2</sub> was modelled as a neutral molecule (Figure 2.14).



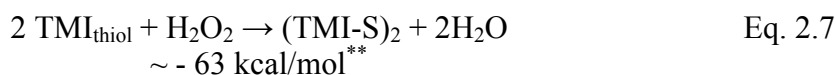
The reduction of (TMMI-S)<sub>2</sub> by methyl mercaptan to the zwitterion was determined to be favourable ( $\Delta G^\circ = -6.40$  kcal/mol). These results predict that GSH is able to act as a reducing agent for OSSO under physiological conditions. It is clear that under identical

conditions, the reduction of ESSE is much more favourable than the reduction of OSSO. This result is consistent with the detection of OSSO *in vivo*.

### 2.3.7. Biologically Relevant Reactions

ESH is known to react with a number of toxic ROS<sup>3</sup> including HO• and H<sub>2</sub>O<sub>2</sub>, and it has been suggested that ESH may work in concert with ascorbic acid to protect the cell against oxidative damage.<sup>5</sup> Similarly, OSHs act as peroxidase mimics, protecting sea urchin eggs from cytotoxic H<sub>2</sub>O<sub>2</sub>.<sup>6,71</sup> To further our understanding of these important reactions, the thermodynamics of the interactions of TMI and TMMI with H<sub>2</sub>O<sub>2</sub> and HO•, as well as the interaction of TMI• with an ascorbic acid analogue, were determined. These results were compared to the reactions of methyl mercaptan as a model for known ROS scavenger GSH.

The reduction of H<sub>2</sub>O<sub>2</sub> (Eq. 2.5) to water was favourable for all three thiols (Table 2.19). However, the protonation state of TMI and TMMI modulates the thermodynamics of these reactions. In the case of TMI, the thiol form is a thermodynamically superior reducing agent compared to the thione form. The difference between the two tautomers corresponds to the energy required to convert the thione to the thiol form. It is therefore likely that ESH is thermodynamically a good scavenger of H<sub>2</sub>O<sub>2</sub>.



In the case of TMMI, the effect of protonation state on the thermodynamics of Eq. 2.5 is even more striking. Reduction by the anion is thermodynamically 2.5 times more likely than by the zwitterion. This result is consistent with experimental evidence which suggests that the OSH anion is the primary reducing agent,<sup>62</sup> even though it accounts for only 0.002% of the OSH population at physiological pH.<sup>††</sup> The use of the anionic form of TMMI as a reducing agent for H<sub>2</sub>O<sub>2</sub> was more than twice as thermodynamically favoured over reduction by methyl mercaptan, in agreement with the role that OSHs act as H<sub>2</sub>O<sub>2</sub> scavengers in sea urchin eggs, in the presence of GSH.<sup>6,71</sup>

---

\*\* See Tables 2.13 and 2.14 for exact values.

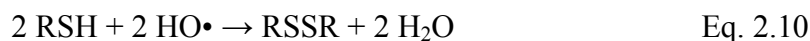
†† As calculated from pKa data.

**Table 2.19:** 2 RSH + H<sub>2</sub>O<sub>2</sub> → RSSR + 2 H<sub>2</sub>O (Eq. 2.5)

<i>Reactant</i>	<i>Product</i>	<i>ΔG° (kcal/mol)</i>
<b>TMI</b>		
2 2-3	2-11	-36.63
	2-12	ND <sup>a</sup>
	2-13	-36.58
2 2-4	2-11	-63.17
2 2-5	2-12	ND
2-4 + 2-5	2-13	-63.38
<b>TMMI</b>		
2 2-14	2-18	-69.13
2 2-15 <sup>b</sup>	2-18	-152.41
2 2-17	2-18	-60.55
<b>Methyl Mercaptan</b>		
2 2-19	2-22	-75.97

<sup>a</sup>Not determined; <sup>b</sup>Balanced equation becomes 2 TMMI<sup>-</sup> + 2 H<sub>3</sub>O<sup>+</sup> + H<sub>2</sub>O<sub>2</sub> → (TMMI-S)<sub>2</sub> + 4 H<sub>2</sub>O

The scavenging of HO• was studied using three different possible reaction schemes (Equations 2.8 to 2.10). All pathways were favourable for the three thiol analogues (Table 2.20, Table 2.21, and Table 2.22). Similar to the H<sub>2</sub>O<sub>2</sub> results, the anionic form of TMMI was the most thermodynamically favourable reducing agent. These results are consistent with experimental observations that ESH, the OSHs and GSH are potent scavengers of HO•.<sup>3,8,62,64,72,73</sup>

**Table 2.20:** RSH + HO• → RS• + H<sub>2</sub>O (Eq. 2.8)

<i>Reactant</i>	<i>Product</i>	<i>ΔG° (kcal/mol)</i>
<b>TMI</b>		
2-3	2-9	-33.83
2-3	2-10	-34.05
2-4	2-9	-47.10
2-5	2-10	-47.58
<b>TMMI</b>		
2-14	2-16	-45.88
<b>Methyl Mercaptan</b>		
2-19	2-21	-31.04

**Table 2.21:**  $\text{RS}^- + \text{HO}^\bullet + \text{H}_3\text{O}^+ \rightarrow \text{RS}^\bullet + 2 \text{H}_2\text{O}$  (Eq. 2.9).

<i>Reactant</i>	<i>Product</i>	$\Delta G^\circ$ (kcal/mol)
<b>TMI</b>		
<b>2-7</b>	<b>2-9</b>	-83.12
<b>2-8</b>	<b>2-10</b>	-83.57
<b>TMMI</b>		
<b>2-15</b>	<b>2-16</b>	-87.52
<b>Methyl Mercaptan</b>		
<b>2-20</b>	<b>2-21</b>	-78.97

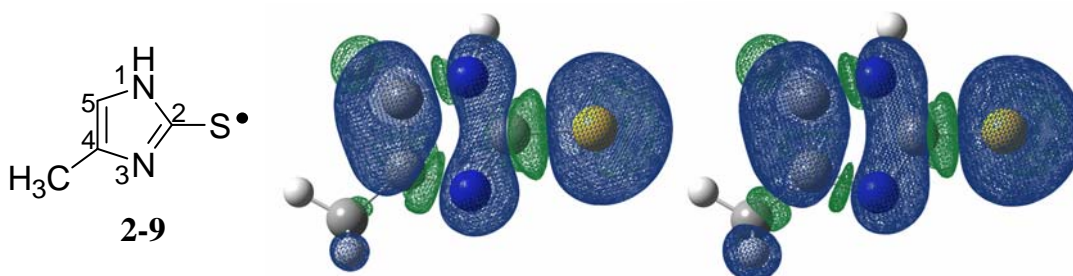
**Table 2.22:**  $2 \text{RSH} + 2 \text{HO}^\bullet \rightarrow \text{RSSR} + 2 \text{H}_2\text{O}$  (Eq. 2.10).

<i>Reactant</i>	<i>Product</i>	$\Delta G^\circ$ (kcal/mol)
<b>TMI</b>		
<b>2 2-3</b>	<b>2-11</b>	-69.27
	<b>2-12</b>	ND
	<b>2-13</b>	-69.21
<b>2 2-4</b>	<b>2-11</b>	-95.80
<b>2 2-5</b>	<b>2-12</b>	ND
<b>2-4 + 2-5</b>	<b>2-13</b>	-72.85
<b>TMMI</b>		
<b>2-14</b>	<b>2-18</b>	-101.77
<b>2-15<sup>a</sup></b>	<b>2-18</b>	-185.04
<b>2-17</b>	<b>2-18</b>	-93.19
<b>Methyl Mercaptan</b>		
<b>2-19</b>	<b>2-22</b>	-99.58

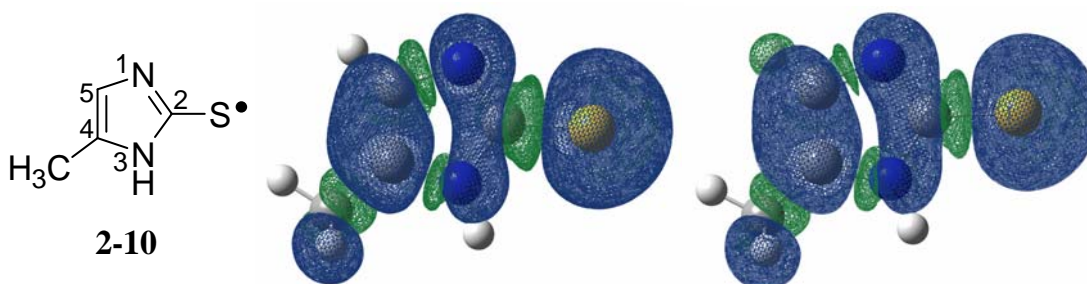
<sup>a</sup>Balanced equation becomes  $2 \text{OS}^- + 2 \text{HO}^\bullet + 2 \text{H}_3\text{O}^+ \rightarrow \text{OSSO} + 4 \text{H}_2\text{O}$

To further understand the behaviour of the ESH and OSH radicals the spin density of the TMI and TMMI radicals were modeled *in vacuo* and with implicit water solvation to determine the localization of the unpaired electron. In TMI, the spin density was found on multiple atoms: the S, C<sup>5</sup> and N<sup>3</sup> atoms of **2-9** and the S, C<sup>4</sup> and N<sup>1</sup> atoms of **2-10** *in vacuo* (see Figure 2.15 and Figure 2.16). When modeled with implicit solvation, the spin density on the nitrogen atom of each tautomer decreased, leaving most of the density on the S and C<sup>4</sup> or C<sup>5</sup> atoms (Figure 2.15 and Figure 2.16, see *Appendix 1* for a detailed listing of the spin densities of each atom). In contrast, the spin density of TMMI is localized on the S and C<sup>4</sup> atoms regardless of solvation (Figure 2.17 and *Appendix 1*), which is in quantitative agreement with published AM1 calculations for OSH derivatives.<sup>64</sup> The delocalization of the radical into the imidazole ring may aid in the stabilization of the ESH and OSH radicals. This increased stability may explain why the formation of the TMI and TMMI radicals was thermodynamically favoured over the formation of the methyl mercaptan radical. The

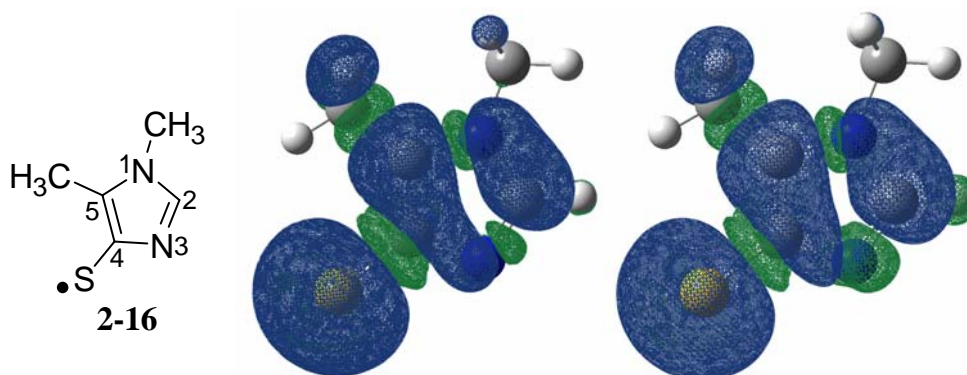
increased stability may also account for the superior ability of the OSHs to scavenge radicals as compared to GSH.<sup>8,62</sup>



**Figure 2.15:** Spin density localization (Isovalue for isosurface: 0.0004 electrons/au<sup>3</sup>) for 2-thiol-4-methylimidazole tautomer **2-9** in the gas phase (left) and with implicit water solvation (right).



**Figure 2.16:** Spin density localization (Isovalue for isosurface: 0.0004 electrons/au<sup>3</sup>) for 2-thiol-4-methylimidazole tautomer **2-10** in the gas phase (left) and with implicit water solvation (right).

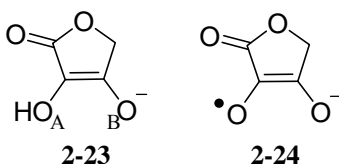


**Figure 2.17:** Spin density localization (Isovalue for isosurface: 0.0004 electrons/au<sup>3</sup>) for 4-thiol-*N*<sup>1</sup>-methyl-5-methylimidazole (**2-16**) in gas (left) and water (right)

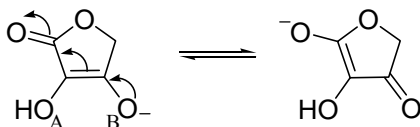
The scavenging of radicals by thiols often creates a thiyl radical which, though less reactive than ROS such as HO•, can still cause oxidative cellular damage. Ascorbic acid is an effective scavenger of a number of ROS and of thiyl radicals. The ascorbate radical formed by these reactions is much less reactive than most of the radicals scavenged, making it a powerful antioxidant. It has been suggested that a cooperative interaction may exist between ESH and ascorbic acid as the exposure of ESH/ascorbic acid solutions to azide radicals results in the formation of the ESH radical, which decays while the ascorbate radical forms, indicating the ascorbate repairs ESH.<sup>5</sup> Ascorbate has also been implicated in the

repair of vitamin E, an essential lipid soluble antioxidant that protects the cell membrane against ROS.<sup>72</sup> Vitamin E reduces lipid peroxy radicals, breaking the free radical propagation chain reaction, and in the process, vitamin E is oxidized to its radical form.<sup>74,75</sup> Continued membrane protection requires the reduction of the vitamin E radical; this is accomplished by a number of cytosolic antioxidants, including ascorbate.<sup>72</sup> In a similar fashion, ascorbate may be involved in the reduction of ES• to ESH.<sup>5</sup>

To study the thermodynamics of this interaction, we modeled the reactions of TMI, TMMI and methylmercaptan radicals with an ascorbic acid analogue (Figure 2.18, **2-23**, Eq. 2.11) to form **2-24**. These truncated analogues of ascorbic acid were used to reduce the number of degrees of freedom involved in the geometry optimization. Based upon the pKa values of ascorbic acid (pKa<sub>1</sub> = 4.2, pKa<sub>2</sub> = 11.6), **2-23** is expected to exist in the anionic form, as shown in Figure 2.18). It is expected that hydroxyl group B is more acidic than A, as the resulting anion is resonance stabilized as seen in Figure 2.19. Hydroxyl A remains protonated at physiological pH, and therefore can act as a hydrogen donor to harmful radical species such as HO•.



**Figure 2.18:** Model compounds used to study the reactions of ascorbic acid.



**Figure 2.19:** Tautomerization of the ascorbate model **2-23**.

The reduction of the thyl radicals of all three thiols by **2-23** was favourable indicating that ascorbate could play a role in returning ESH, the OSHs and GSH to their biologically useful thiol form (see *Appendix I*).

**Table 2.23:** RS• + 2-23 → RSH + 2-24 (Eq. 2.11)

<i>Reactant</i>	<i>Product</i>	$\Delta G^\circ$ (kcal/mol)
<b>TMI</b>		
<b>2-9</b>	<b>2-3</b>	-16.39
	<b>2-4</b>	-3.12
<b>2-10</b>	<b>2-3</b>	-16.17
	<b>2-5</b>	-2.64
<b>TMMI</b>		
<b>2-16</b>	<b>2-14</b>	-4.34
	<b>2-15<sup>a</sup></b>	37.3
	<b>2-17</b>	-8.63
<b>Methyl Mercaptan</b>		
<b>2-21</b>	<b>2-20</b>	-19.18

<sup>a</sup>Balanced equation modified to TMMI• + 2-23 + H<sub>2</sub>O → TMMI<sup>-</sup> + 2-24 + H<sub>3</sub>O<sup>+</sup>

## 2.4. Conclusions and Future Work

This computational approach explored the theoretical aspects of the thermodynamics of several of the important reactions that ESH and the OSHs may undergo in the cell. A key result from this work is the indication that the instability of ESSE may be driven by the energy released by the tautomerization of the thiol form of ESH to the thione.

Further computational studies of ESH and OSH could involve the determination of the thermodynamics of the reactions of TMI and TMMI with reactive nitrogen species (RNS) such as peroxyxynitrate (ONOO<sup>-</sup>), and with *S*-nitroso-glutathione and di-*S*-nitroso-trypanothione. ESH is believed to scavenge cytotoxic ONOO<sup>-</sup> in a biologically relevant manner<sup>2,76,77</sup> and both ESH and OSH have been shown to decompose *S*-nitrosothiols, intermediates in the detoxification of RNS.<sup>78,79</sup>

The calculations and results presented herein should contribute to our further understanding of the fundamental chemistry and biochemistry of these important intracellular thiols.

## 2.5. References

1. Heath, H.; Toennies, G. *Biochem. J.* **1958**, *68*, 204-210.
2. Jang, J.-H.; Aruoma, O. I.; Jen, L.-S.; Chung, H. Y.; Surh, Y.-J. *Free Radic. Biol. Med.* **2004**, *36*, 288-299.
3. Akanmu, D.; Cecchini, R.; Aruoma, O. I.; Halliwell, B. *Arch. Biochem. Biophys.* **1991**, *288*, 10-16.
4. Arduini, A.; Eddy, L.; Hochstein, P. *Arch. Biochem. Biophys.* **1990**, *281*, 41-43.
5. Asmus, K.-D.; Bensasson, R. V.; Bernier, J.-L.; Houssin, R.; Land, E. J. *Biochem. J.* **1996**, *315*, 625-629.
6. Turner, E.; Hager, L. J.; Shapiro, B. M. *Science* **1988**, *242*, 939-941.
7. Pathirana, C.; Andersen, R. J. *J. Am. Chem. Soc.* **1986**, *108*, 8288-8289.
8. Holler, T. P.; Hopkins, P. B. *Biochemistry* **1990**, *29*, 1953-1961.

9. Leach, A. R. *Molecular Modelling: Principles and Applications*; Second ed.; Prentice Hall: Harlow, Essex, England, 2001.
10. Cramer, C. J. *Essentials of Computational Chemistry*; Second ed.; John Wiley and Sons Ltd.: West Sussex, 2004.
11. Young, D. C. *Computational Chemistry: A Practical Guide for Applying Techniques for Real World Problems*; Wiley Interscience: New York, 2001.
12. Atkins, P.; Friedman, R. *Molecular Quantum Mechanics*; Fourth ed.; Oxford University Press: New York, 2005.
13. Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098-3100.
14. Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 1372-1377.
15. Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785-789.
16. Agache, C.; Popa, V. I. *Monatshefte für Chemie* **2006**, *137*, 55-68.
17. Boström, J.; Norrby, P.-O.; Liljefors, T. *J. Comput. Aided Des.* **1998**, *12*, 383-396.
18. Beak, P. *Acc. Chem. Res.* **1977**, *10*, 186-192.
19. Foresman, J. B.; Frisch, A. E. *Exploring Chemistry with Electronic Structure Methods*; Second ed.; Gaussian Inc.: Pittsburgh, PA, USA, 1996.
20. Levine, I. N. *Quantum Chemistry*; Fifth ed.; Prentice Hall: Upper Saddle River, NJ, USA, 2000.
21. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A.; B.05 ed.: Pittsburgh, PA, 2003.
22. Flack, H. D. *Acta Crystallogr.* **1983**, *A39*, 876-881.
23. Binkley, J. S.; Pople, J. A.; Hehre, W. J. *J. Am. Chem. Soc.* **1980**, *102*, 939-947.
24. Gordon, M. S.; Binkley, J. S.; Pople, J. A.; Pietro, W. J.; Hehre, W. J. *J. Am. Chem. Soc.* **1982**, *104*, 2797-2803.
25. Pietro, W. J.; Francl, M. M.; Hehre, W. J.; DeFrees, D. J.; Pople, J. A.; Binkley, J. S. *J. Am. Chem. Soc.* **1982**, *104*, 5039-5048.
26. Ditchfield, R.; Hehre, W. J.; Pople, J. A. *J. Chem. Phys.* **1971**, *54*, 724-728.
27. Hehre, W. J.; Pople, J. A. *J. Chem. Phys.* **1972**, *56*, 4233-4234.
28. Binkley, J. S.; Pople, J. A. *J. Chem. Phys.* **1977**, *66*, 879-880.
29. Hariharan, P. C.; Pople, J. A. *Theor. Chim. Acta* **1973**, *28*, 213-222.
30. Hehre, W. J.; Ditchfield, R.; Pople, J. A. *J. Chem. Phys.* **1972**, *56*, 2257-2261.
31. Francl, M. M.; Pietro, W. J.; Hehre, W. J.; Binkley, J. S.; Gordon, M. S.; DeFrees, D. J.; Pople, J. A. *J. Chem. Phys.* **1982**, *77*, 3654-3665.



32. Clark, T.; Chandrasekhar, J.; Spitznagel, G. W.; Schleyer, P. V. *J. Comput. Chem.* **1983**, *4*, 294-301.
33. Frisch, M. J.; Pople, J. A.; Binkley, J. S. *J. Chem. Phys.* **1984**, *80*, 3265-3269.
34. Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. *J. Chem. Phys.* **1980**, *72*, 650-654.
35. McLean, A. D.; Chandler, G. S. *J. Chem. Phys.* **1980**, *1980*, 5639-5648.
36. Kendall, R. A.; Dunning, T. H., Jr.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6796-6806.
37. Woon, D. E.; Dunning, T. H., Jr. *J. Chem. Phys.* **1993**, *98*, 1358-1371.
38. Woon, D. E.; Dunning, T. H., Jr. *J. Chem. Phys.* **1994**, *100*, 2975-2988.
39. Slater, J. C. *Quantum Theory of Molecules and Solids, Vol. 4: The Self-Consistent Field for Molecules and Solids*; McGraw-Hill: New York, 1974.
40. Vosko, S. H.; Wilk, L.; Nusair, M. *Can. J. Phys.* **1980**, *58*, 1200-1211.
41. Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623-11627.
42. Miehlich, B.; Savin, A.; Stoll, H.; Preuss, H. *Chem. Phys. Lett.* **1989**, *157*, 200-206.
43. Perdew, J. P. In *Electronic Structure Theory of Solids*; P. Ziesche, H. Eschrig, Eds.; Akademie Verlag: Berlin, 1991.
44. Perdew, J. P.; Chevary, J. A.; Vosko, S. H.; Jackson, K.; Perderson, M. R.; Singh, D., J.; Fiolhais, C. *Phys. Rev. B* **1992**, *46*, 6671-6687.
45. Cancès, M. T.; Mennucci, B.; Tomasi, J. *J. Chem. Phys.* **1997**, *107*, 3032-3041.
46. Mennucci, B.; Tomasi, J. *J. Chem. Phys.* **1997**, *106*, 5151-5158.
47. Tomasi, J.; Mennucci, B.; Cancès, M. T. *J. Mol. Struct. (THEOCHEM)* **1999**, *464*, 211-226.
48. Blaudeau, J.-P.; McGrath, M. P.; Curtiss, L. A.; Radom, L. *J. Chem. Phys.* **1997**, *107*, 5016-5021.
49. Watchers, A. J. H. *J. Chem. Phys.* **1980**, *52*, 1033-1036.
50. Hay, P. J. *J. Chem. Phys.* **1977**, *66*, 4377-4384.
51. Raghavachari, K.; Trucks, G. W. *J. Chem. Phys.* **1989**, *91*, 1062-1065.
52. Binning, R. C., Jr.; Curtiss, L. A. *J. Comput. Chem.* **1990**, *11*, 1206-1216.
53. Curtiss, L. A.; McGrath, M. P.; Blaudeau, J.-P.; Davis, N. E.; Binning, R. C., Jr. *J. Chem. Phys.* **1995**, *103*, 6104-6113.
54. McGrath, M. P.; Radom, L. *J. Chem. Phys.* **1991**, *94*, 511-516.
55. Huang, M.-J. *Int. J. Quantum Chem.* **2004**, *96*, 374-379.
56. Virovets, A. V.; Lider, E. V.; Elokhina, V. N.; Bushuev, M. B.; Lavrenova, L. G. *J. Struct. Chem.* **2005**, *46*, 358-362.
57. Jadeja, R. N.; Shirsat, R. N.; Suresh, E. *Struct. Chem.* **2005**, *16*, 515-520.
58. Bastos, B.; Mena, A.; Alcântara, A. F. d. C.; Beraldo, H. *Tetrahedron* **2005**, *61*, 7045-7053.
59. Papanikolaou, P. A.; Christidis, P. C.; Chaviara, A. T.; Bolos, C. A.; Tsipis, A. C. *Eur. J. Inorg. Chem.* **2006**, *2006*, 2083-2095.
60. Sugihara, A.; Uemura, K.; Matsuura, Y.; Tanaka, N.; Ashida, T.; Kakudo, M. *Acta Crystallogr.* **1976**, *B32*, 181-185.
61. Motohashi, N.; Mori, I.; Sugiura, Y. *Chem. Pharm. Bull.* **1976**, *27*, 1737-1741.
62. Holler, T. P.; Hopkins, P. B. *J. Am. Chem. Soc.* **1988**, *110*, 4837-4838.
63. Zoete, V.; Bailly, F.; Vezin, H.; Teissier, E.; Duriez, P.; Fruchart, J.-C.; Catteau, J.-P.; Bernier, J.-L. *Free Radic. Res.* **2000**, *32*, 515-524.

64. Zoete, V.; Vezin, H.; Bailly, F.; Vergoten, G.; Catteau, J.-P.; Bernier, J.-L. *Free Radic. Res.* **2000**, *32*, 525-533.
65. Jocelyn, P. C. *The Biochemistry of the SH Group*; Academic Press: New York, 1972.
66. Scott, E. M.; Duncan, I. W.; Ekstrand, V. *J. Biol. Chem.* **1963**, *238*, 3928-3933.
67. Voet, D.; Voet, J. G. *Biochemistry*; Second ed.; John Wiley and Sons: New York, 1995.
68. Palumbo, A.; Misuraca, G.; D'ischia, M.; Donaudy, F.; Prota, G. *Comp. Biochem. Physiol.* **1984**, *78B*, 81-83.
69. Rohl, I.; Scheider, B.; Schmidt, B.; Zeeck, E. *Z. Naturforsch.* **1999**, *54C*, 1145-1147.
70. Turner, E.; Klevit, R.; Hager, L. J.; Shapiro, B. M. *Biochemistry* **1987**, *26*, 4028-4036.
71. Shapiro, B. M. *Science* **1991**, *252*, 533-536.
72. Acworth, I. N.; McCabe, D. R.; Maher, T. J. In *Oxidants, Antioxidants and Free Radicals*; H. Salem, Ed.; Taylor and Francis: Washington, D. C., 1997, pp 22-77.
73. Sciuto, A. M. In *Oxidants, Antioxidants, and Free Radicals*; S. I. Baskin, H. Salem, Eds.; Taylor and Francis: Washington, D. C., 1997, pp 171-191.
74. Brigelius-Flohe, R.; Traber, M. G. *FASEB J.* **1999**, *13*, 1145-1115.
75. Burton, G. W.; Ingold, K. U. *Acc. Chem. Res.* **1986**, *19*, 194-201.
76. Aruoma, O. I.; Spencer, J. P. E.; Mahmood, N. *Food Chem. Toxicol.* **1999**, *37*, 1043-1053.
77. Aruoma, O. I.; Whiteman, M.; England, T. G.; Halliwell, B. *Biochem. Biophys. Res. Comm.* **1997**, *231*, 389-391.
78. Vogt, R. N.; Steenkamp, D. J.; Zheng, R.; Blanchard, J. S. *Biochem. J.* **2003**, *374*, 657-666.
79. Misiti, F.; Castagnola, M.; Zuppi, C.; Giardina, B.; Messana, I. *Biochem. J.* **2001**, *356*, 799-804.

## Chapter 3: Conformational Analyses of Mycothiol

The novel thiol mycothiol (MSH, Figure 3.1, **3-1**) is currently of intense interest due to its essential role in the cellular physiology of *Mycobacteria*,<sup>1-5</sup> such as *Mycobacterium tuberculosis*, and its possible role in antimycobacterial drug resistance.<sup>6-8</sup> Despite the medical importance of the organisms that use MSH as their major intracellular thiol, little is known concerning MSH biochemistry.<sup>†</sup> In addition, there is very little information available regarding the chemical reactivity and conformational preferences of MSH, information that would be useful in the design of novel inhibitors of MSH-utilizing enzymes. Based upon these considerations, a detailed conformational study of MSH was undertaken to determine if any conformational preferences exist in solution. The findings of this study should not only contribute to our fundamental understanding of MSH, but should also aid in the design of MSH analogues, such as conformationally restricted derivatives, which may serve as important leads in the development of novel antimycobacterial agents.

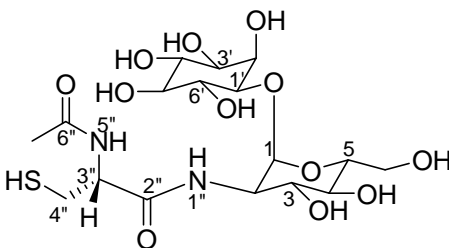
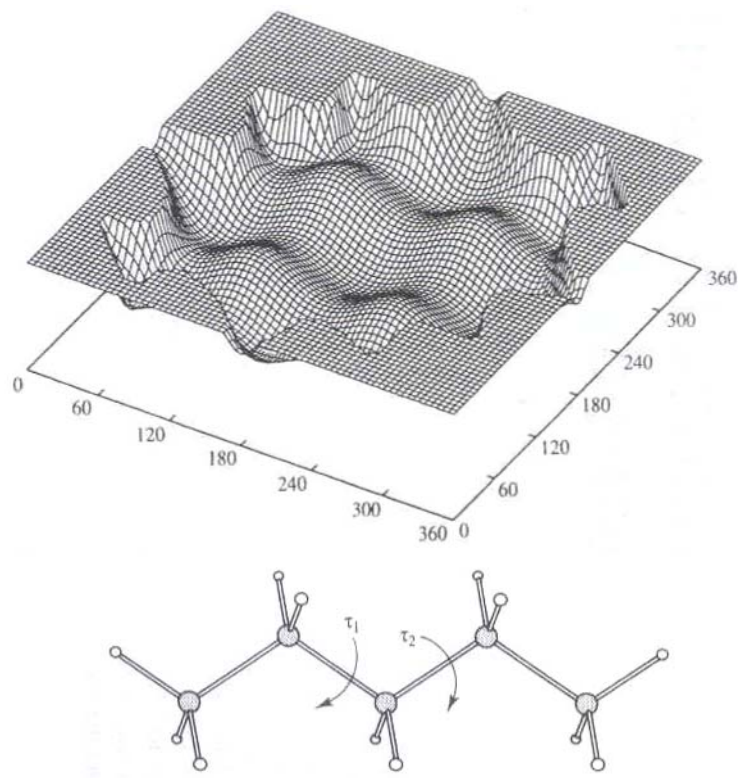


Figure 3.1: Mycothiol (MSH, 3-1)

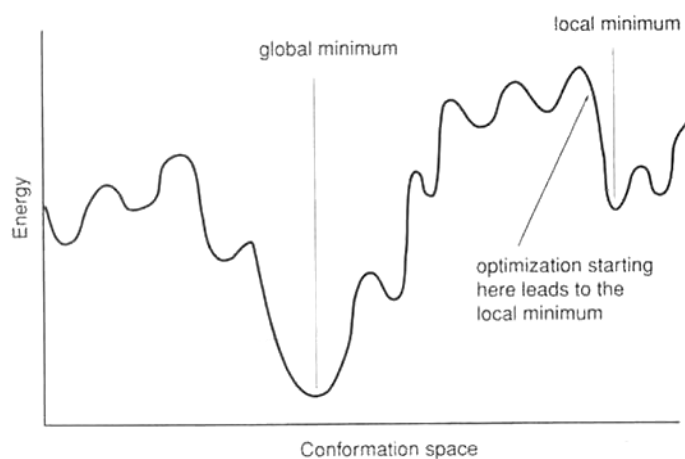
### 3.1. Computational Chemistry

As discussed in *Chapter 2*, the potential energy of a molecule is a function of the coordinates of its atoms. The way in which this energy changes with respect to geometry defines a potential energy surface, like the one seen in Figure 3.2. The valleys of a potential energy surface correspond to minimum energy structures with the lowest overall energy point being the global minimum. Most geometry optimization algorithms move down the potential energy surface and therefore can only locate the minimum energy which is closest to the starting structure (Figure 3.3).<sup>9</sup> Consequently, the optimized structure obtained by a geometry optimization may not be a global minimum but rather a local minimum energy structure (Figure 3.3). To ensure that a global minimum structure has been found, or to generate multiple local minimum energy structures and explore the conformational space of a molecule, multiple starting structures must be generated and their geometries optimized.

<sup>†</sup> A search of the PubMed database, a service of the National Library of Medicine and the National Institute of Health, for MSH returned 77 references, 8 of which are review articles while a search for glutathione, the main intracellular thiol of eukaryotes and most gram-negative bacteria, returned 73 334 references, 4 022 of which are review articles (accessed November 13, 2006).



**Figure 3.2:** Variation in the energy of pentane with the variation of the angles indicated. Only the lowest energy regions are shown.<sup>9</sup>



**Figure 3.3:** A one dimensional representation of the energy of all possible conformers of a simple molecule<sup>10</sup>

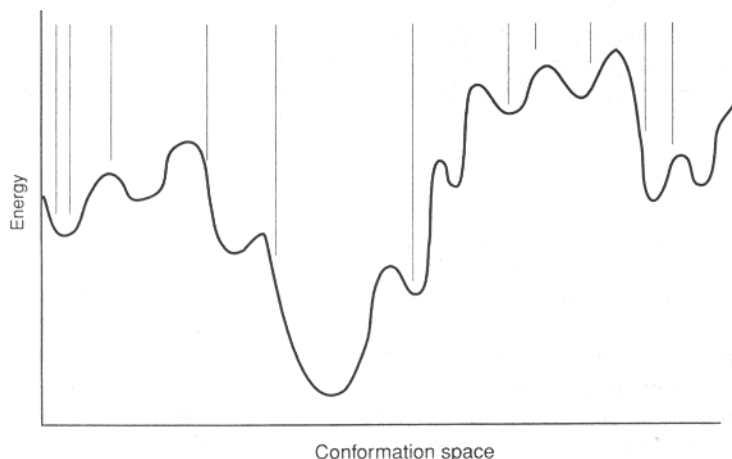
There are a variety of methods available for the generation of multiple starting structures and the exploration of a compound's conformational space. For example, the stochastic search algorithms, such as the Monte Carlo multiple minimization (MCMM) method,<sup>11</sup> make random changes to a molecule's structure to generate new geometries for optimization.<sup>10</sup> Another popular method is the use of molecular dynamics (MDyn) simulations which apply Newton's laws of motion and temperature to generate new starting geometries for optimization. MDyn methods are not suitable for the study of all systems as

the forces applied to the individual atoms of the molecule may not be sufficient to overcome energy barriers between energy minima, resulting in inadequate sampling of the conformational space of the molecule.<sup>9</sup> In addition, MDyn simulations can be very time consuming.<sup>10</sup> The most thorough method available to explore the conformations of a molecule is dihedral angle driving, or adiabatic mapping, which rotates the torsional angles of a molecule by a predefined value to thoroughly cover the conformational space of the compound. This method is not suited for larger molecules as the number of dihedral angle combinations sampled increases exponentially with the size of the molecule.<sup>9</sup> For example, if a molecule had 10 dihedral angles and each one was rotated in 20° increments for a total of 360°, one would generate 18<sup>10</sup>, or over 3.5 × 10<sup>12</sup>, structures.<sup>12</sup>

The optimization of the numerous structures generated by conformational search algorithms is very demanding with respect to time and computational resources. These demands often dictate the method chosen for the optimization of the resulting starting structures. The Hartree-Fock (HF) and Density Functional Theory (DFT) methods are the most accurate; however, they are also the most demanding due to the large number of calculations involved. Even the semi-empirical methods, which are less intensive than the HF and DFT techniques, can result in a prohibitively long study. As an alternative, molecular mechanics methods are often used in conformational searches due to their speed and relatively low demand on computational resources. With these factors in mind, we chose to use the MCMM method in combination with molecular mechanics methods to explore the conformational space of MSH.

### *3.1.1. Monte Carlo Multiple Minimization Methods*

Monte Carlo (MC) simulations generate new conformations by applying random changes to the positions of the constituent atoms of a molecule.<sup>9</sup> If enough random changes (or steps) are allowed, this method should give a statistically valid exploration of the conformational space of a molecule (Figure 3.4).<sup>10</sup> An MCMM conformational search begins with the choice of a stable starting structure which then undergoes random variations in its dihedral angles, followed by geometry optimization and a comparison of the new conformer to any previously determined minima. The optimized structure is either stored as new and unique or rejected as identical and the cycle is repeated. If the molecule under study contains a ring, the ring is opened prior to dihedral angle variation and then closed before optimization.<sup>11</sup> These methods are attractive as they require less computer time than MDyn simulations<sup>10</sup> and can overcome energy barriers that may impede other methods such as MDyn simulations.<sup>9</sup>



**Figure 3.4:** Sampling the conformational space of a molecule using a Monte Carlo search with a small number of iterations<sup>10</sup>

### 3.1.2. Molecular Mechanics Methods and Force Fields

Molecular mechanics calculations are very useful for the study of systems which are too large for quantum mechanical techniques. Quantum mechanical calculations focus on the electrons in a system, and even if some of these electrons are disregarded, as in semi-empirical calculations, there are still a large number of particles to be considered and the calculations are very time consuming.<sup>9</sup> Molecular mechanics methods are more feasible for large systems as they use the positions of the nuclei of a molecule to determine the overall energy of the system, with key energetic contributions from bond stretching, bond angle changes, rotations around single bonds and non-bonded interactions, such as electrostatic and van der Waals interactions.<sup>9</sup>

The energy of the system is calculated using a collection of functions and parameters called a *force field*.<sup>13</sup> These parameters are defined through the fitting of these functions to spectroscopic data or to the results of *ab initio* calculations for a given system.<sup>10</sup> It is desirable to have a force field that can accurately describe a system other than the one used to develop its parameters.<sup>9</sup> One method used to broaden the applicability of a force field is the use of atom types to describe the component atoms of a system. For example, the MM2, MM3 and MM4 force fields distinguish between  $sp^3$ ,  $sp^2$ ,  $sp$ , carbonyl, cyclopropane, and cyclopropene carbon atoms, as well as carbon radicals, and carbonium ions.<sup>9,14-23</sup> These atom types describe not only hybridization, but can also be used to describe functional groups, such as carbonyl versus carboxylic  $sp^2$  carbon atoms, as well as the local environment of the atom in question.<sup>9</sup> Force fields designed for modeling specific classes of molecules, such as proteins and carbohydrates, often use more specific atom types than those force fields designed for general use.<sup>9,10</sup>

The accuracy of molecular mechanics calculations is related to how well the parameters of the force field in use are suited to the system in question. A force field parameterized for the study of hydrocarbons may not yield accurate results for the study of biological molecules. For example, the rings in sugar molecules are known to pose a problem to general force fields, likely due to the unique properties of the anomeric position; therefore, these molecules should be modelled using force fields designed specifically for the study of carbohydrates.<sup>10</sup> With the use of a well chosen force field, molecular mechanics

calculations can, in some cases, provide results with similar accuracy to higher level quantum mechanical calculations.

### 3.1.3. Plan of Action

Based upon the above factors, we performed a detailed search of the conformational space of MSH using MCMM and molecular mechanics methods. We first evaluated the ability of the force fields available to us to model MSH, then performed a MCMM conformational search both *in vacuo* and with implicit solvation. The resultant structures were subjected to cluster analysis to determine if any patterns existed with respect to specific dihedral angles within MSH. The resultant global minimum and other low energy minima structures were compared to a recently published structure of an MSH-derivative found using NMR techniques.

## 3.2. Materials and Methods

All structures were drawn and visualized using Maestro (Schrödinger Inc., Portland). A conformational study of MSH was undertaken using MacroModel (versions 8.1 and 9.0, Schrödinger Inc., Portland). The Generalized Born/Surface Area (GB/SA) continuum model was used to create the implicit water solvation model as implemented by MacroModel. All calculations were performed on Violin or Flexor (see *Chapter 2* for details of these systems).

### 3.2.1. Geometry Optimizations of Mycothiol

All geometry optimizations were performed using the same protocol. The electrostatic treatment was dependent upon the force field in use with a dielectric constant of 1.0 and the charges were defined by the force field library. Hydrogen-bonding interactions were cut off at 4 Å. Van der Waals and charge-charge interactions were cut off at 7 and 12 Å respectively, in the gas phase, and at 8 and 20 Å respectively under implicit solvation. A Polak-Ribiere Conjugate Gradient (PRCG) was used and each structure was allowed up to 5000 iterations. The minima converged on a gradient with a threshold of 0.05.

### 3.2.2. Survey of Force Fields Available

An initial survey of the force fields available in MacroModel version 8.1 was undertaken to determine which force field was best suited to model MSH. Minimizations of MSH were performed using each of the following force fields:<sup>‡</sup> AMBER\*,<sup>24</sup> AMBER94,<sup>25</sup> MM2\*,<sup>15</sup> MM3\*,<sup>16</sup> MMFF,<sup>26</sup> OPLS<sup>27</sup> and OPLS-AA.<sup>28</sup> Each returned information regarding the number and quality of the bend, stretch, and torsional parameters available in each force field for the atom types of MSH. Of the force fields studied, OPLS-AA appeared best suited to model MSH (see *Results and Discussion*); therefore, OPLS-AA was used as the main force field in our conformational study of MSH.

### 3.2.3. Monte Carlo Multiple Minimization Searches of the Mycothiol Conformational Space

An X-ray structure of MSH is not currently available; therefore, the structure of MSH was drawn into Maestro and its geometry optimized both in the gas phase and with implicit

---

<sup>‡</sup> As implemented by MacroModel v. 8.1.

water solvation (GB/SA)<sup>29</sup> using the OPLS-AA force field.<sup>§</sup> The minimized structures were then used as the starting structures for MCMM conformational searches in their respective phases (gas or implicit water). Each subsequent step used the “least used” structure from the conformations thus far generated, based upon the number of times each conformer had been used as a starting conformer. In this study the structures were allowed 500 000 MC steps and structures within 50 kJ/mol of the global minimum were retained. For each starting structure,  $x$  number of degrees of freedom, or in this case torsional angles, were altered, where  $x$  is a randomly generated number between 2 and 5. The bonds were rotated plus or minus a randomly generated number between 0 and 180°. To adjust the torsional angles within the D-glucosamine and *myo*-inositol rings, the bonds between C<sup>3</sup> and C<sup>4</sup> and C<sup>2'</sup> and C<sup>3'</sup> (see Figure 3.1 for numbering) were temporarily broken to open each ring. The dihedral angles were modified and the rings closed. These newly formed bonds were constrained to between 0.5 and 2.5 Å in length. The absolute values of the torsion angles of the amide bonds were monitored and the structure was discarded if these values were outside of 0° - 90° for C<sup>2</sup>-N<sup>1''</sup>-C<sup>2''</sup>-O (D-glucosamine) and 90° - 180° for H-N<sup>5''</sup>-C<sup>6''</sup>-O (*N*-acetyl-cysteine, see Figure 3.1 for numbering).

The resultant conformation was then minimized using the OPLS-AA force field. After 1666 iterations, a preliminary energy test was performed and the minimization was aborted if the structure's energy was more than 100 kJ/mol higher than that of the current minimum energy structure, speeding the overall conformational search. The chiralities of the stereogenic atoms, the carbon atoms of the D-glucosamine and *myo*-inositol rings and the  $\alpha$ -C of the *N*-acetyl-cysteine moiety, were monitored and any structure whose chirality was altered was rejected. The structure was finally compared to the previously saved conformers for redundancy. Structures whose least squares superimposition was less than 0.25 Å were considered identical and the higher energy conformer was then discarded. Initially all heavy atoms and polar hydrogen atoms were used for comparison resulting in 13 074 and 137 764 unique structures in the gas phase and with implicit water solvation, respectively. These conformers were then subjected to the redundant conformation elimination module of MacroModel to eliminate conformers whose heavy atom placement differed by less than 0.25 Å without consideration of the polar hydrogen atoms.

A second search of the conformational space of MSH was performed using AMBER<sup>\*\*</sup> to account for any inadequacies of the OPLS-AA force field (see *Results and Discussion*). MSH was minimized using AMBER\* with implicit solvation (GB/SA) and this optimized structure was used as the starting point of the MCMM conformational search. The MCMM conformational search was performed in a similar fashion to the OPLS-AA search except that the torsional angles within the *myo*-inositol and D-glucosamine rings were not modified nor were the rings opened. The torsional angles between the rings were subjected to rotation (see *Results and Discussion*). After the initial conformational search, 33 124 conformers were found, which were then subjected to the redundant conformer elimination module.

#### 3.2.4. Resultant Conformer Analysis using XCluster

The conformers modelled with implicit water solvation which were calculated to have energies within 3 kcal/mol of the global minimum (OPLS-AA: 198 conformers, AMBER\*:

---

<sup>§</sup> As implemented by MacroModel v. 8.1

<sup>\*\*</sup> As implemented by MacroModel v. 9.0



116 conformers) were further analyzed using the XCluster module of MacroModel v. 9.0.<sup>30</sup> The conformers were grouped into families based upon the  $\phi$  (C<sup>1</sup>-O-C<sup>1'</sup>-H) and  $\psi$  (C<sup>1</sup>-O-C<sup>1'</sup>-H) angles (see Figure 3.1 for numbering) of the disaccharide portion of MSH. Conformers containing five axial substituents on the *myo*-inositol ring were excluded from further analysis (OPLS-AA: 50 conformers, see *Results and Discussion*).

### 3.2.5. MOE2004 Calculations

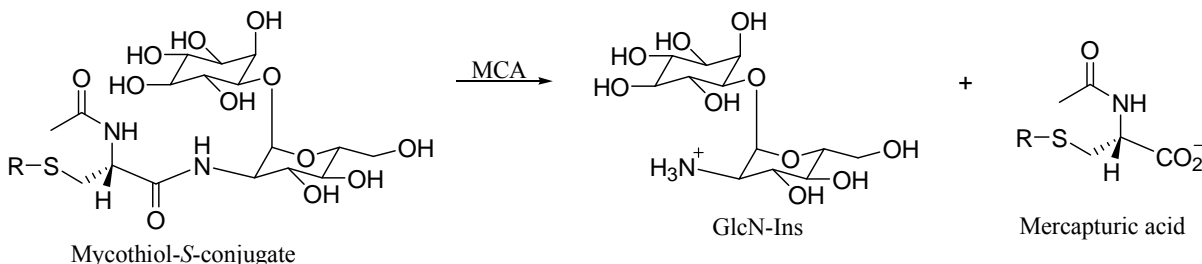
For comparison, a conformational search of MSH was also performed by Professor France-Isabelle Auzanneau of the University of Guelph Chemistry Department, using the Molecular Operating Environment program suite (MOE2004). All MOE2004 calculations were performed on a PC Pentium 4, 2.40 GHz CPU. Similar to the random incremental pulse search (RIPS)<sup>31</sup> method, new MSH conformations were generated by the random rotation of all of the bonds in MSH, including the ring bonds, in addition to Cartesian coordinate perturbation. Unfortunately, this method led to changes in the absolute configuration of the carbon atoms of the *myo*-inositol ring. MSH was therefore minimized with the PEF95SAC force field<sup>32</sup> and the heavy atoms of the *myo*-inositol ring were held fixed for all subsequent conformation searches. All other dihedral angles were rotated randomly with a sum-of-Gaussians distribution with peaks at multiples of 30°. To improve the efficiency of the search a simultaneous Cartesian perturbation of 0.4 Å was also applied. These random perturbations were repeated 250 000 times per conformational search and the resulting conformations were optimized using the PEF95SAC force field invoking the GB/SA implicit solvation model, both as implemented by MOE2004. Initially each new conformation was checked against the previously determined conformations using a root-mean-square (RMS) tolerance of 0.1 Å on all atoms, including the hydrogen atoms to account for possible hydrogen bonding. A second search was performed using a more stringent RMS cut off of 0.01 Å, which was applied to the heavy atoms only. Both searches were allowed 250 000 iterations with the limitation that the search would terminate if no new conformations were found after 10 000 iterations in a row. This limitation was never reached in either search; therefore, both searches terminated after 250 000 iterations. Structures that did not have the <sup>4</sup>C<sub>1</sub> conformation for D-glucosamine were eliminated (see *Results and Discussion*) and the results of the two searches were combined. Conformations with energies of 10 kcal/mol above the global minimum were rejected yielding a total of 25 852 structures.

The combined database was re-minimized using the AMBER94 force field<sup>25</sup> using the GB/SA implicit solvation model, both as implemented by MOE2004, yielding 10 656 structures within 10 kcal/mol of the AMBER94 defined global minimum. The structures within 3 kcal/mol of the global minimum for each force field (PEF95SAC: 1860 conformations, AMBER94: 1067 conformations) were then grouped based upon their  $\phi$  and  $\psi$  angles.

## 3.3. Results and Discussion

MSH is a clinically important glycothiol believed to be essential for the survival of *Mycobacterium tuberculosis*,<sup>33</sup> however, there is very little known regarding its biochemistry and conformational preferences.<sup>4,12</sup> Information regarding the conformational preferences of MSH may be useful for the design of novel inhibitors aimed at MSH-utilizing enzymes. MSH-S-conjugate amidase (MCA, Figure 3.5) for example, is an attractive target for the development of inhibitors as it is believed to be involved in the removal of electrophilic

species from *Mycobacteria*, and may be involved in the detoxification of thiol reactive antibiotics, such as rifamycin S and cerulenin.<sup>7</sup> The design and synthesis of MSH-utilizing enzyme inhibitors is ongoing, with some of these being synthetic MSH analogues.<sup>34-37</sup> The results of our modelling study may aid in the design of inhibitors by identifying key dihedral angles in MSH, which are highly populated and found in solution.



**Figure 3.5:** Hydrolysis of modified mycothiol by mycothiol-S-conjugate amidase (MCA).

### 3.3.1. Choice of Conformational Search Method

To accurately model the conformations of a molecule, multiple conformations must be generated that explore all, or at least a substantial portion of the conformational space of the molecule. The MCMM methods sample a wide variety of conformations by making random changes to the torsional angles of a molecule, generating new geometries that do not necessarily resemble the starting structure. If a sufficient number of conformers are generated, the MCMM methods should sample a large portion of the conformational space. An indication of a sufficient number of steps would be the repeated rejection of newly generated structures based upon their similarity to previously found conformations. The MCMM methods are also not hindered by energy barriers between conformations as they can make “non-physical moves” which significantly enhance the ability of the MCMM method to explore the potential energy surface of the molecule.<sup>9</sup> MCMM methods have been useful in many studies including the study of the coil-to-helix transition in polyalanine<sup>38</sup> and the determination of the low energy conformations of synthesized carbohydrates.<sup>39</sup> The ability of the MCMM method to overcome energy barriers to rotation makes it well suited for the study of MSH as it will not be hindered by the low mobility of the torsional angles between *myo*-inositol and D-glucosamine. This method was selected for our study of the conformations of MSH.

Other known conformational search methods include torsional angle driving, also known as adiabatic mapping, and MDyn simulations (see *Section 3.2.3*). Torsional angle driving, while one of the most thorough techniques, was rejected as a useful method for this study as the number of angles to be sampled in MSH would be prohibitively large. If one ignored the dihedral angles in the disaccharide rings, there would still remain nine dihedral angles to evaluate, assuming only heavy atoms were considered. If these dihedral angles were modified in 20° increments, 18<sup>9</sup> or over 198 billion possible starting conformations would be generated to examine; a major computational undertaking. Dihedral angle driving is therefore not a feasible option for the study of MSH. MDyn, another popular method, was also rejected for this study as these simulations can miss minimum energy structures if a large energy barrier exists between the starting structure and other minima on the potential energy surface.<sup>9</sup> While the amino acid portion of MSH is relatively flexible, the torsional angles between the disaccharide rings have less mobility due to the steric bulk of *myo*-inositol and D-glucosamine. The energy barriers to rotation around these dihedral angles is

likely quite high and therefore an MDyn simulation would be unlikely to be able to fully sample the conformational space of MSH. In addition, MDyn simulations are limited by the computational resources available and simulations of a few nanoseconds are standard; it is possible that carbohydrate molecules undergo dynamic events on longer time scales which could not be investigated by these methods.<sup>40</sup>

### 3.3.2. Force Field Selection

MSH is an unusual thiol in that it contains a disaccharide moiety coupled to cysteine. This results in a wide variety of atom and bond types within MSH, complicating force field selection; many force fields are not parameterized for all of the different MSH bond and atom types. MacroModel returns information regarding the quality (high, medium, or low) of the parameters in use for a particular structure (Table 3.1). The lower the quality of the parameter, the less experimental or *ab initio* data that has been used in fitting the force field equation. A low quality rating indicates a general parameter whose use can lead to inaccurate conformational energy differences and geometries. The use of low quality stretch parameters in particular, can give crude partial charges, which in turn, can cause computed charges and solvation energies to be inaccurate.<sup>41</sup>

**Table 3.1:** Number of each quality of the various force field parameters available in MacroModel for bond stretching and angle bending and torsional angle rotation in MSH.

Force Field <sup>a</sup>	Stretch			Bend			Torsion		
	High	Med	Low	High	Med	Low	High	Med	Low
AMBER*	37	0	0	111	1	0	152	11	0
AMBER94 <sup>b</sup>	--	--	--	--	--	--	--	--	--
MM2	75	8	0	140	23	0	143	108	0
MM3	35	14	14	58	73	137	82	20	1
MMFF	63	0	0	113	0	0	173	0	6
OPLS	63	0	0	112	0	1	170	9	0
<b>OPLS-AA</b>	<b>63</b>	<b>0</b>	<b>0</b>	<b>113</b>	<b>0</b>	<b>0</b>	<b>179</b>	<b>0</b>	<b>0</b>

<sup>a</sup>As implemented by MacroModel v. 8.1; <sup>b</sup>Insufficient force field constants were available.

Based upon the quality rating returned by MacroModel for the parameters of the various force fields, the OPLS-AA force field appeared best suited for modelling MSH and therefore, was used as the main force field in the MCMM conformational search. The OPLS-AA force field has traditionally been used to model amino acids, but has also been parameterized,<sup>42</sup> and successfully used to model carbohydrates including disaccharides<sup>43-45</sup> and cyclodextrin.<sup>46</sup> In addition, as implemented by MacroModel, OPLS-AA has been parameterized for the sulfur atom of thiols.<sup>28</sup>

MSH was also modelled using MCMM methods with the AMBER\* force field. While this force field returned a medium bend parameter for the bend of the O<sub>ring</sub>-C<sup>1</sup>-O<sub>linkage</sub> angle and medium torsional angle parameters for the cysteine portion of MSH (Table 3.1), its use in the study of carbohydrates is well documented.<sup>47</sup> This second conformational search was performed to detect any inadequacies in the OPLS-AA results.

### 3.3.3. Ring Opening during the Conformational Search

The initial MCMM conformational search using the OPLS-AA force field opened the *myo*-inositol and D-glucosamine rings for torsional angle variation followed by ring closure.

This ring opening allowed for an unbiased conformational search. All of the conformers found to be within 3 kcal/mol of the implicitly solvated global minimum structure had a  ${}^4C_1$  orientation for D-glucosamine. This orientation is consistent with the known crystal<sup>48</sup> and solution phase structures<sup>49</sup> of *N*-acetyl-D-glucosamine (as seen in Figure 3.1). The *myo*-inositol ring, however, was found in two orientations with either one equatorial or five equatorial substituents. It is known from solution phase NMR<sup>50</sup> and X-ray crystallographic<sup>51</sup> data that *myo*-inositol prefers the five equatorial/one axial (5 eq/1ax) substituent conformation (as seen in Figure 3.1); therefore, all structures with the 1 eq/5 ax orientation were eliminated from further consideration (50 structures within 3 kcal of the global minimum).

### 3.3.4. Gas Phase versus Solvation Phase Modelling

MSH was modelled in the gas phase, as well as with implicit water solvation using the GB/SA implicit water solvation model. The GB/SA model is a combination of the generalized Born and the solvent accessible surface area methods (SASA); the Born model is based on electrostatic interactions, dielectric permittivity and orbital overlaps while the SASA model assumes that the greatest interaction of the solvent with the solute is in the area very close to the solute molecule.<sup>10</sup> This method has been shown to be effective in modelling carbohydrates<sup>47</sup> and has been successfully applied to the conformational analyses of ganglioside head groups,<sup>52</sup> and mannobiosides and triosides.<sup>53</sup> The overlay of the global minima of MSH in the gas phase and with implicit water solvation demonstrates the effect of solvent shielding on the conformation of a polar compound. In the gas phase the importance of hydrogen bonding and electrostatics can be exaggerated, and folded or collapsed structures become more probable.<sup>54</sup> In a true aqueous environment, water molecules can provide competing hydrogen bonding interactions, which would also affect the structure of a molecule. While implicit water solvation cannot mimic the effect of solute-solvent hydrogen bonding, it can provide electrostatic shielding, and lessen the effects of electrostatics to the overall structure of the molecule. As can be seen in Table 3.2, the gas phase and implicitly solvated global minimum structures differ in hydrogen bond placement with the gas phase global minimum having a more closed structure with the amino acid portion of MSH hydrogen bonding to the disaccharide (Figure 3.6).

**Table 3.2:** Hydrogen-bonding interactions for the global minimum conformations of MSH found *in vacuo* and with implicit water solvation using the OPLS-AA force field.

<i>Phase</i>	<i>Number of Intramolecular Hydrogen Bonds</i>	<i>Atoms Involved<sup>a</sup></i>
<i>in vacuo</i>	3	$C^{6''}=O\cdots H-O-C^{2''}$ $C^{6''}=O\cdots H-N^{1''}$ $C^{2''}=O\cdots H-O-C^6$
water	2	$C^{2''}=O\cdots H-S$ $C^{6''}=O\cdots H-N^{1''}$

<sup>a</sup>See Figure 3.1 for numbering



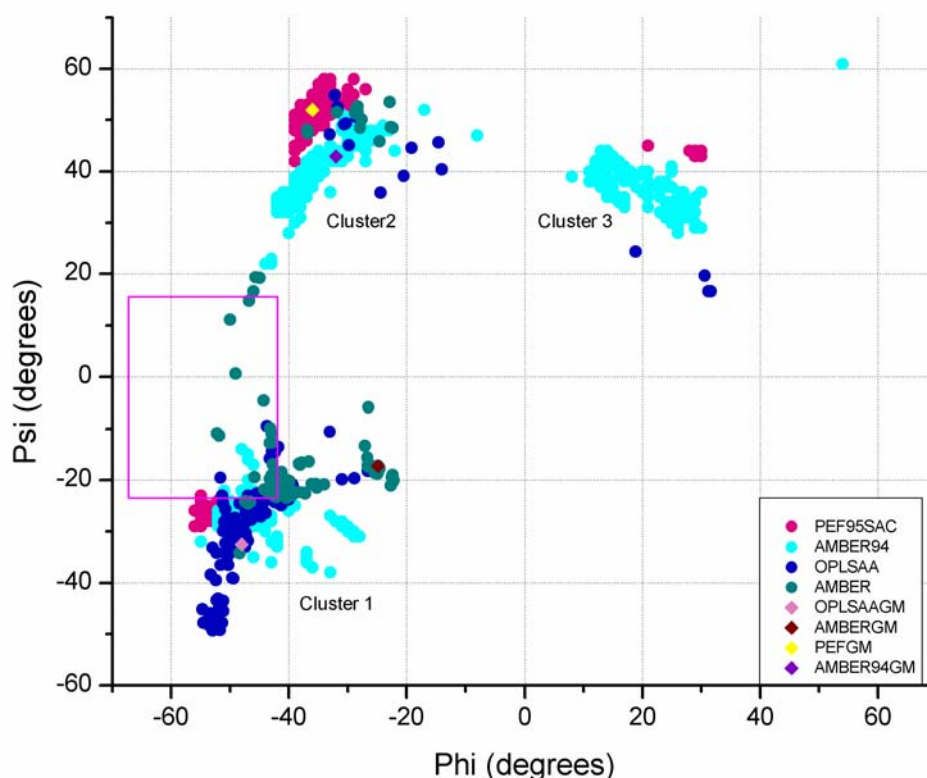
**Figure 3.6:** Stereoview overlay of the global minimum conformations of mycothiol found for the gas (coral) and implicit water (cyan) phases using the OPLS-AA force field.

### 3.3.5. Energy Differences Relevant to Biological Systems

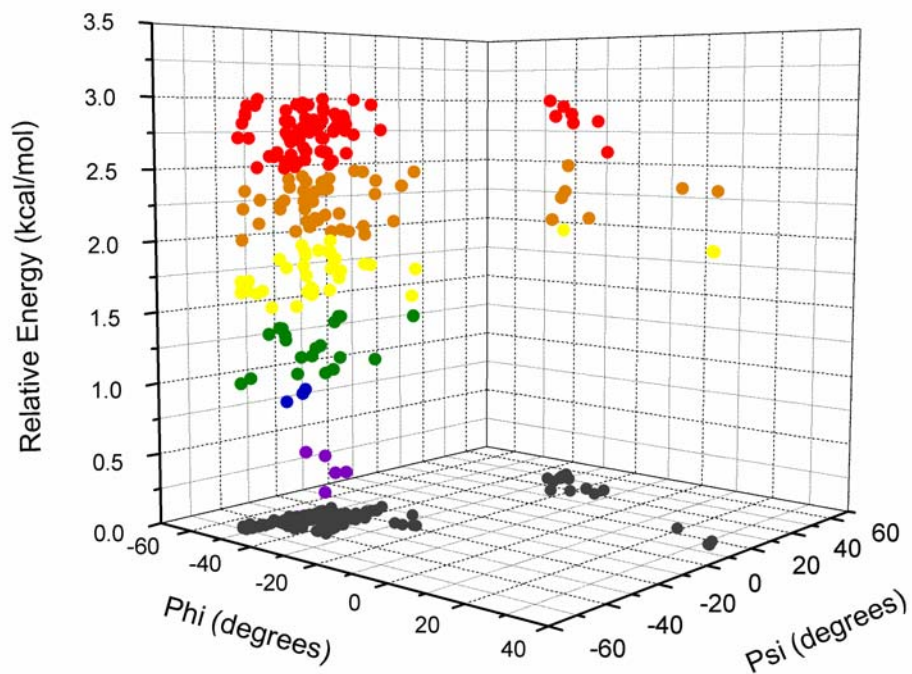
The global minimum conformation is the most energetically favourable form of the molecule; however, it may not be the conformation required for protein binding.<sup>55</sup> A study of 33 ligand-protein complexes found that for majority of the complexes, the protein bound ligand conformation was not the global minimum, but rather a conformation with an energy within 3 kcal/mol of the global minimum.<sup>54</sup> In the case of highly flexible ligands, binding conformations with energies of 5 kcal/mol greater than the global minimum have been identified.<sup>56</sup> It has been suggested that when performing a conformational study, a cut off of 3 kcal/mol should be applied,<sup>55</sup> except in the case of highly flexible ligands.<sup>56</sup> While MSH has a high degree of flexibility in its cysteine component, there is much less flexibility around the disaccharide linkage of *myo*-inositol and D-glucosamine; hence a 3 kcal/mol cut off may be sufficient for the study of MSH. Only those conformers within 3 kcal/mol of the implicitly solvated global minimum structure were studied further.

### 3.3.6. Analysis of the OPLS-AA Data using Clustering Techniques

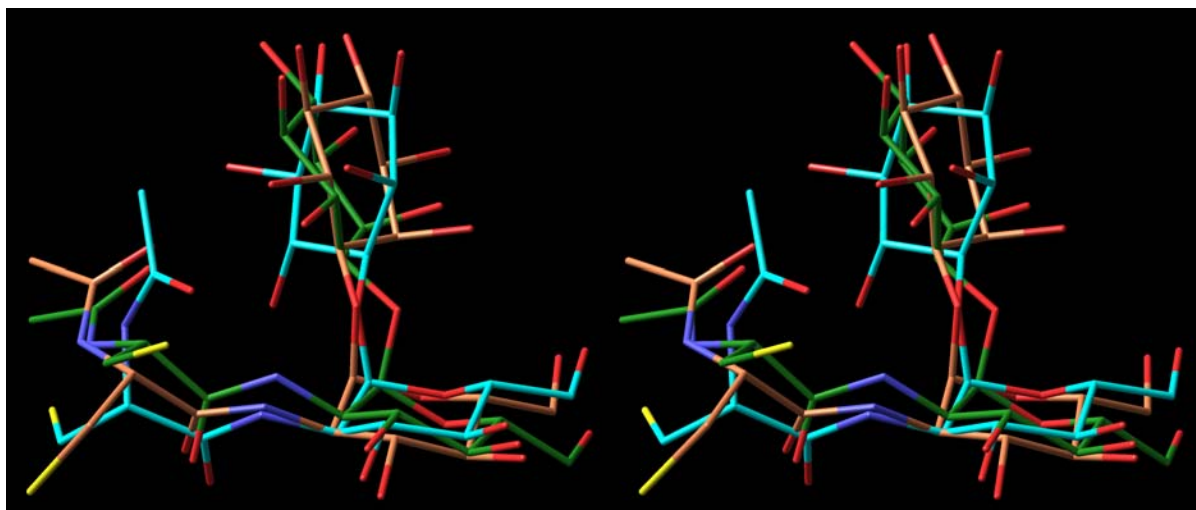
Clustering was performed using XCluster, which has been utilized to study a wide range of systems including the antibiotic mutacin 1140,<sup>57</sup> thrombin inhibitors,<sup>58</sup> agouti-related peptides<sup>59</sup> and peptide-tannin systems.<sup>60</sup> Clustering of the implicit water OPLS-AA data, using the  $\phi$  and  $\psi$  angles of interest, indicated the presence of three distinct clusters (Figure 3.7). The conformation of disaccharides is largely dependent upon the orientation of around these two angles as the flexibility of a pyranose ring is rather limited and the substituents have little influence on the conformational space of the disaccharide.<sup>40</sup> Cluster 1 is the most highly populated group, with 131 members and contains the OPLS-AA implicitly solvated global minimum structure as well as the majority of the structures within 2 kcal/mol of the global minimum (Figure 3.8). Clusters 2 and 3 contain 13 and 4 members, respectively (Figure 3.7 and Table 3.3). An overlay of the local minimum energy conformers from each cluster highlights the differences between these angles (Figure 3.9).



**Figure 3.7:** Clustering of the  $\phi$  and  $\psi$  angles of mycothiol conformers found to be within 3 kcal/mol of their respective global minima. These conformers were generated using the OPLS-AA, AMBER\*, AMBER94 and PEF95SAC force fields. The purple box represents the  $\phi$  and  $\psi$  angle ranges of MSH-S-bimane as determined by NMR data.<sup>12</sup> There are two conformers found using AMBER\*, which are not shown in this figure with  $(\phi, \psi) = (-27^\circ, -155^\circ)$  and  $(-28^\circ, -155^\circ)$ .



**Figure 3.8:** Comparison of the  $\phi$  and  $\psi$  angles and energy (kcal/mol) determined for MSH using the OPLS-AA force field and implicit water solvation.



**Figure 3.9:** Stereoview of the superimposition of the lowest energy conformers (C1-C3) of each of the cluster as determined by the OPLS-AA force field with implicit water solvation: cyan = Cluster 1 (C1), also the global minimum; coral = Cluster 2 (C2); green = Cluster 3 (C3).

**Table 3.3:** Comparison of the three clusters and global minima of the implicitly solvated MSH conformers generated in this study.

	<i>Cluster 1</i>	<i>Cluster 2</i>	<i>Cluster 3</i>	<i>GM<sup>a</sup></i>
<b>MacroModel</b>				
<b>OPLS-AA<sup>b</sup></b>				
$\phi$ Dihedral	-55° to -26°	-33° to -14°	19° to 32°	-48°
$\psi$ Dihedral	-49° to -10°	36° to 55°	17° to 24°	-33°
No. of members	131	13	4	--
<b>AMBER*<sup>c,d</sup></b>				
$\phi$ Dihedral	-52° to -22°	-37° to -23°	--	-25°
$\psi$ Dihedral	-34° to 19°	46° to 54°	--	-17°
No. of members	101	13	--	--
<b>MOE2004</b>				
<b>AMBER94<sup>e</sup></b>				
$\phi$ Dihedral	-55° to -28°	-44° to -8° <sup>f</sup>	8° to 30° <sup>g</sup>	-31°
$\psi$ Dihedral	-38° to -13°	22° to 51°	28° to 43° <sup>g</sup>	43°
No. of members	391	390	285	--
<b>PEF95SAC<sup>e</sup></b>				
$\phi$ Dihedral	-56° to -47°	-39° to 27°	20° to 30°	-36°
$\psi$ Dihedral	-29° to -23°	42° to 58°	43° to 45°	52°
No. of members	66	1782	12	
<b>Mycothioliol-S-bimane</b> <b>(NMR data)</b>				
$\phi$ Dihedral	-66° to -42°			
$\psi$ Dihedral	-22° to 18°			

<sup>a</sup>Implicitly solvated global minimum; <sup>b</sup>As implemented by MacroModel v. 8.1; <sup>c</sup>As implemented by MacroModel v. 9.0; <sup>d</sup>Two outlying angle combinations were found with  $(\phi, \psi) = (-27^\circ, -155^\circ)$  and  $(-28^\circ, 155^\circ)$ , no angle combinations corresponding to Cluster 3 were found; <sup>e</sup>As implemented by MOE2004; <sup>f</sup>Only two conformers were found to have  $\phi$  angles of  $-8^\circ$ , all others in this cluster were below  $-17^\circ$ ; <sup>g</sup>One outlying conformer was determined with  $(\phi, \psi) = (54^\circ, 61^\circ)$ ; <sup>h</sup>Angle ranges suggested by NOE data reported for mycothiol-S-bimane.

### 3.3.7. AMBER\* Conformational Search and Comparison of Results to OPLS-AA

The OPLS-AA force field was determined to be best suited to model the diverse atom types of MSH. To ensure accurate modelling of the carbohydrate portion of MSH, AMBER\*, a force field well-known for modelling carbohydrates,<sup>47</sup> was used to detect any modelling inadequacies in the OPLS-AA results. Within the 148 structures found by the OPLS-AA conformational search to be within 3 kcal/mol of the global minimum conformation with the correct orientation of *myo*-inositol and D-glucosamine, there was very little variation in the ring torsional angles. Therefore, the AMBER\* search froze the internal torsion angles of the two carbohydrate rings after an initial minimization with AMBER\*.



This modification shortened the overall time of the conformational search by reducing the number of torsional angles available for random changes.

Similar to the OPLS-AA results, the majority of the structures found to be within 3 kcal/mol determined using AMBER\*, including the global minimum, were found in Cluster 1 (101 conformers, Table 3.3). There were some outlying structures found with the AMBER\* force field that belonged to Cluster 1 with a  $\phi$  angle of  $\sim 50^\circ$  and  $\psi$  angles ranging from  $-20^\circ$  to  $20^\circ$ . These structures are not seen with any other force field used (OPLS-AA, PEF95SAC or AMBER94). A medium quality parameter is present for the bond of the  $O_{\text{ring}}-C^1-O_{\text{linkage}}$  bond angle. Lower quality parameters can affect the relative energetics of optimized geometries. It is possible that this medium quality parameter altered the calculated energetics, allowing for  $\phi$  and  $\psi$  angle combinations within 3 kcal/mol of the global minimum which were found to be of higher energy with the other force fields (Table 3.3). This low quality parameter may also be responsible for the absence of any  $\phi$  and  $\psi$  angle combinations consistent with Cluster 3 among the conformations found to be within 3 kcal/mol of the global minimum (Figure 3.7) and the identification of two outlying low-energy conformers with  $(\phi, \psi) = (-27^\circ, -155^\circ)$  and  $(-28^\circ, -155^\circ)$ .

### 3.3.8. Comparison of the Monte Carlo Multiple Minimization and MOE Stochastic Search Results

Initially the PEF95SAC conformational search performed by Professor France-Isabelle Auzanneau did not hold the *myo*-inositol or D-glucosamine rings rigid which allowed for ring flipping. In the case of *myo*-inositol, this flipping led to changes in the absolute configuration of the carbon atoms of the *myo*-inositol ring. The *myo*-inositol ring was therefore held fixed after initial optimization for all further studies. In addition, any resultant conformations of MSH found not to have the  ${}^4C_1$  orientation for D-glucosamine were eliminated from the database as *N*-acetyl-D-glucosamine has been shown to prefer  ${}^4C_1$  orientation.<sup>48,49</sup>

The MOE stochastic search used random torsion angle changes combined with a Cartesian perturbation, as opposed to the MCOMM method. A comparison of the MOE results with those of the OPLS-AA force field were in good agreement with respect to the  $\phi$  and  $\psi$  angle ranges found (Figure 3.7 and Table 3.3). This agreement implies that both methods adequately covered the conformational space of MSH and it is unlikely that low energy conformations<sup>††</sup> exist with  $\phi$  and  $\psi$  angles outside of these ranges.

There was some variation in the cluster populations between the various force fields. The conformations found using the AMBER94 force field were almost evenly distributed between the clusters, with Clusters 1 and 2 being slightly favoured. In contrast, over 95 % of the conformations found using the PEF95SAC force field belonged to Cluster 2 (1782 conformations) while only 66 and 12 conformations belonged to Clusters 1 and 3. Unlike the OPLS-AA and AMBER\* results, the global minimum structures of the AMBER94 and PEF95SAC searches were found in Cluster 2. While the relative energetics of the clusters differed between the various force fields, the overall  $\phi$  and  $\psi$  angle ranges were in agreement, indicating that the conformations obtained are independent of force field (Figure 3.7 and Table 3.3).

---

<sup>††</sup>Those with energies within 3 kcal/mol of the global minimum.

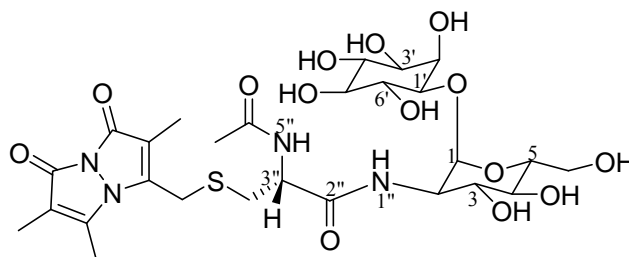
The determination of which basis set is best suited for the study of carbohydrates is a difficult process. There is a lack of suitable experimental data available to act as benchmarks for the thorough testing of the force fields available for carbohydrate modelling.<sup>47</sup> A comparison of 20 force fields commonly used for the modelling of carbohydrates, including PEF95SAC, AMBER94 and AMBER\*, indicated that it is impossible to give an objective determination of the suitability and quality of the force fields available.<sup>47</sup> A number of criteria were used to evaluate the force fields including their ability to model the geometry and energy of methyl glucopyranoside. Only the AMBER\* force field was able to successfully predict energy differences between the  $\alpha$  and  $\beta$  anomers. AMBER94 gave a small preference for the  $\beta$  anomer while PEF95SAC returned unreasonably high preference for the  $\alpha$ -anomer (~7 kcal/mol compared to the experimental value of 0.4 kcal/mol). These energy discrepancies could account for the location of the global minimum energy structures determined by PEF95SAC and AMBER94 for MSH in Cluster 2, as opposed to in Cluster 1 similar to OPLS-AA and AMBER\*. In addition, as implemented by MacroModel, AMBER94 was missing parameters for the atom types of MSH. MOE does not return information regarding the quality or number of parameters available for the study of a particular molecule; but it is likely that the implementations of the force fields are similar between the two software packages. Personal communications with MOE indicated that in the event of a missing parameter for a particular atom type, a generalized one is used in its place. This generalized parameter could also contribute to the location of the global minimum in Cluster 2. AMBER\* was found to have some medium quality parameters for the atom types of MSH. These parameters may account for the outlying  $\phi$  and  $\psi$  angle combinations found by this force field to be within 3 kcal/mol of the global minimum.

Based upon the above considerations, and the solely high quality rating for the parameters of the OPLS-AA force field, it was concluded that this force field was probably best suited for the study of MSH and the results obtained by this force field are the most reliable. The dihedral angle ranges determined using OPLS-AA were in fairly good agreement with the other three force fields, indicating that likely these ranges are valid. The energetic differences between the force fields may be related to variations in parameterization. Based upon the energetics of the OPLS-AA results (Figure 3.8) MSH would most likely be found with the  $\phi$  and  $\psi$  angle combinations consistent with Cluster 1, with some contributions from Clusters 2 and 3.

### *3.3.9. Comparison to the Structure of a Bimane Derivative of Mycothiol Obtained Using NMR Methods.*

A crystal structure of MSH is not yet available for comparison; however a solution phase structure of a bimane derivative of MSH (MSmB, **3-2**, Figure 3.10) has been recently reported.<sup>12</sup> This structure was determined by NMR using key inter-hydrogen atom distances obtained from NOE data. MSmB is a substrate for MCA, the enzyme responsible for the cleavage of MSH-S-conjugates formed by the reaction of MSH with electrophiles.<sup>6-8</sup> The structure of MSmB was therefore used as a comparison against the implicitly solvated MSH conformers found in this study. The local minimum energy structures (**C1-C3**, Figure 3.9) of the three clusters, as found by the OPLS-AA force field, were used as representative structures of Clusters 1 to 3. In the case of Cluster 1, the local minimum energy structure was also the global minimum conformation. The structures were compared using important

torsional angles, as well as the key interatom distances used to determine the structure of MSmB.<sup>12</sup>



**Figure 3.10:** The bimane derivative of mycothiol (MSmB, **3-2**), a substrate for mycothiol-*S*-conjugate amidase.

The MSmB structure is in good agreement with **C1** with respect to the select cysteinyl angles shown in Table 3.4. A comparison of the key interatom distances used to determine the NMR structure of MSmB is shown in Table 3.5. Among the three representative structures, **C2** and **C3** are more similar to each other than to **C1**, as can be seen in Figure 3.9. This overlay demonstrates that the *myo*-inositol ring of **C1** is oriented away from the *myo*-inositol rings of **C2** and **C3**. All three representative structures satisfy the interatom distance ranges determined from NOE data for H<sup>1</sup>-H<sup>1'</sup>, H<sup>1''</sup>-H<sup>3''</sup>, H<sup>5''</sup>-H<sup>3''</sup>, H<sup>5</sup>-H<sup>1'</sup>, and CH<sub>3</sub>-H<sup>5''</sup> (Table 3.5, see Figure 3.10 for numbering). Only **C1** satisfies the distance ranges reported for H<sup>1</sup>-H<sup>6'</sup>, H<sup>5</sup>-H<sup>2'</sup> and CH<sub>3</sub>-H<sup>6'</sup> (Table 3.5). **C2** and **C3** satisfy the range reported for H<sup>1</sup>-H<sup>2'</sup>, while **C1** does not. The orientation of the *myo*-inositol ring in **C1** brings H<sup>6'</sup> close to H<sup>1</sup> and places H<sup>2'</sup> away from H<sup>1</sup>, such that it cannot satisfy the H<sup>1</sup>-H<sup>2'</sup> NOE predicted distance range. The reverse is true for **C2** and **C3**. The inability of one local minima conformer to satisfy the criteria of the MSmB structure indicates that MSH does not exist as a static structure in solution, but is likely conformationally flexible. This hypothesis is supported by a comparison of the  $\phi$  and  $\psi$  angles calculated in this study and those determined for MSmB. There is some overlap between the NOE predicted angle ranges and those found in Cluster 1, as seen in Table 3.3 and Figure 3.7. MSmB is likely able to sample conformations of all three clusters on the time scale of the NMR experiment, leading to averaged values that do not correspond to one cluster alone. Carbohydrates are fairly flexible molecules with respect to their glycosidic linkage and multiple conformations can exist in solution; therefore, any NMR data obtained will likely represent a time-averaged conformation.<sup>61</sup> Care must be taken when using NOE data for structural determination as the measured value is an average which can lead to a single nonexistent conformation.<sup>61</sup> It is therefore likely that the solution phase structure predicted for MSmB by NMR is a weighted average of the conformations found within each cluster.

**Table 3.4:** Comparison of selected cysteinyl torsional angles from representative structures of the implicitly solvated MSH conformers found with the OPLS-AA force field to those of the bimane derivative of mycothiol found using NMR techniques.

<i>Torsion Angles</i>	<i>NMR data</i>	<i>C1<sup>b</sup></i>
H-C <sup>3''</sup> -C <sup>2''</sup> -O	Trans	125.3°
O-C <sup>2''</sup> -N <sup>1''</sup> -H	Trans	174.0°
H-N <sup>5''</sup> -C <sup>3''</sup> -H	-150°	-140.1°

<sup>a</sup>Mahadevan *et al.* (2003);<sup>12</sup> <sup>b</sup>**C1** is also the global minimum.

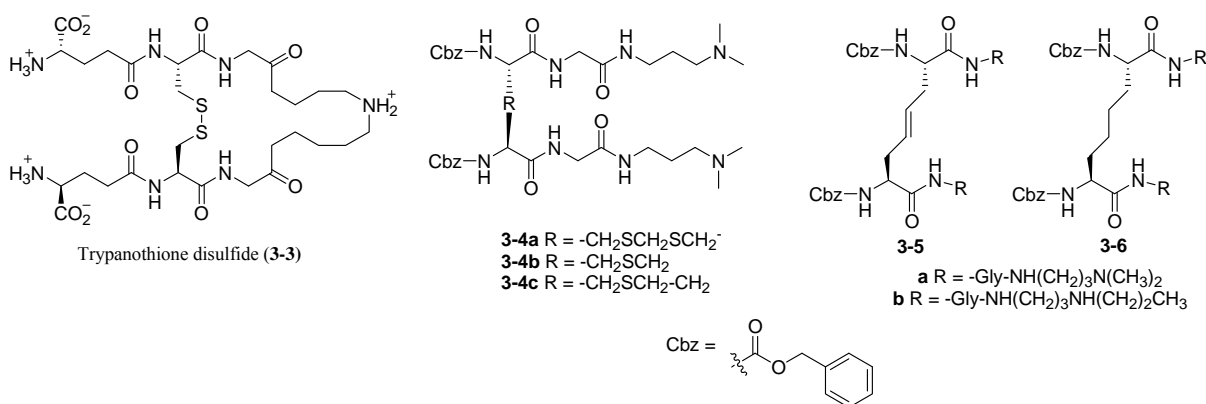
**Table 3.5:** Comparison of key inter-hydrogen distances for representative structures of the implicitly solvated mycothiol conformers found with the OPLS-AA force field to those of the bimane derivative of mycothiol found using NMR techniques.

Structure	Atoms <sup>a</sup>								
	H <sup>1</sup> -H <sup>1'</sup>	H <sup>1</sup> -H <sup>2'</sup>	H <sup>1</sup> -H <sup>6'</sup>	H <sup>5</sup> -H <sup>1'</sup>	H <sup>5</sup> -H <sup>2'</sup>	H <sup>1''</sup> - H <sup>3''</sup>	H <sup>5''</sup> - H <sup>3''</sup>	CH <sub>3</sub> - H <sup>6''b</sup>	CH <sub>3</sub> - H <sup>5''</sup>
NMR <sup>c</sup>	S <sup>d</sup>	M <sup>d</sup>	M	W <sup>d</sup>	S	S <sup>d</sup>	W'	W'' <sup>d</sup>	S''
C1 <sup>e</sup>	<b>2.6<sup>f</sup></b>	4.4	<b>3.4</b>	<b>3.2</b>	<b>2.7</b>	<b>2.5</b>	<b>2.9</b>	<b>3.1</b>	<b>3.0</b>
C2	<b>2.3</b>	<b>2.7</b>	4.5	<b>4.3</b>	3.6	<b>2.2</b>	<b>2.9</b>	7.1	<b>2.9</b>
C3	<b>2.3</b>	<b>2.7</b>	4.5	<b>4.6</b>	4.7	<b>2.2</b>	<b>2.9</b>	7.8	<b>2.9</b>

<sup>a</sup>See Figure 3.10 for numbering; <sup>b</sup>The CH<sub>3</sub> distance was obtained by averaging the distance from each H of the CH<sub>3</sub> group to either H<sup>6'</sup> or H<sup>5'</sup>; <sup>c</sup>Mahadevan *et al.*,<sup>12</sup> <sup>d</sup>S: strong (1.8 – 3.0 Å), M: medium (1.8 – 4.0 Å), W: weak (1.8 – 5.0 Å), S': strong (1.8 – 3.2 Å), W': (1.8 – 5.2 Å), S'': (1.8 – 3.5 Å), W'': (1.8 – 5.5 Å); <sup>e</sup>Representative structures from Clusters 1 (C1), 2 (C2) and 3 (C3). These values are the implicitly solvated local minima from each cluster determined using the OPLS-AA force field; <sup>f</sup>Values in bold are in agreement with the NMR data.

### 3.3.10. Applications to Drug Design

Knowledge of the important angles or conformations of MSH could be of great assistance for the design of antimicrobials against *Mycobacteria*. MSH is not found in eukaryotes and mutants of *Mycobacteria* deficient in the production of MSH have shown increased sensitivity to electrophiles, free radicals and antibiotics.<sup>62</sup> The enzymes known to require MSH as an essential cofactor are therefore potential targets for the development of new therapeutics against *Mycobacteria*. For example, a number of natural products and similar synthetic inhibitors have been identified for MCA,<sup>34,35,63</sup> and there is intense interest in the development of substrate based inhibitors targeted at MCA and other MSH-utilizing enzymes.<sup>36,37,64</sup> Substrate or transition state based inhibitors show some structural similarity to the substrate or mimic the properties of the transition state of the reaction. For example, enzymes that use trypanothione, the major thiol of the *Leishmania* and *Trypanosoma*<sup>65,66</sup> are targets for development for antipathogenic therapeutics.<sup>67</sup> A number of substrate based analogues have been developed which have inhibitory activity<sup>68,69</sup> against trypanothione reductase, the enzyme responsible for maintaining trypanothione in its biologically useful reduced form.<sup>70-75</sup> Select examples of these analogues can be seen in Figure 3.11. The development of substrate-based inhibitors of MSH is ongoing<sup>64</sup> and any information gained in this study regarding the preferred  $\phi$  and  $\psi$  angles of MSH may aid in the design of inhibitors aimed at MSH-utilizing enzymes.



**Figure 3.11:** Trypanothione and select examples of substrate based inhibitors of trypanothione reductase.

### 3.4. Conclusions and Future Work

A detailed conformational investigation was performed on the novel intracellular glycothiol MSH, which is found solely in the *Actinomycetales* bacteria, including the medically important *M. tuberculosis*. The OPLS-AA force field was used for MCMM calculations performed on MSH *in vacuo* and with implicit water solvation. The global minimum energy structures of MSH in the two phases were found to be similar; however, intramolecular hydrogen-bonding played a greater role in the conformation of MSH *in vacuo* than when modelled with implicit water solvation. This result is attributed to the lack of electrostatic shielding present *in vacuo*, as compared to solvated MSH.

To detect any inadequacies in the OPLS-AA results, MSH was also modelled using AMBER\* and the MCMM method, and using AMBER94 and PEF95SAC with a stochastic search method. Conformers within 3 kcal/mol of the respective global minimum structure were compared and it was found that these conformations were independent of the force fields used. All conformations clustered into three distinct families based upon the  $\phi$  and  $\psi$  angles of the disaccharide. Based upon the population of these clusters it is expected that MSH will be found predominantly with  $\phi$  and  $\psi$  angles ranges of  $-55^\circ$  to  $-26^\circ$  and  $-49^\circ$  to  $-10^\circ$  or  $-44^\circ$  to  $-14^\circ$  and  $22^\circ$  to  $58^\circ$ . This data compares favourably to a previously determined NMR structure of the bimane derivative of MSH. Knowledge of the preferred conformations of MSH may aid in the design of substrate based inhibitors aimed at MSH-utilizing enzymes.

Future conformational studies of MSH could include the use of a higher level of theory to minimize the geometries generated by the MCMM algorithm. In addition, the determination of a solution phase NMR or an X-ray crystal structure of MSH would provide greater insight into the behaviour and interactions of this critical glycothiol. The determination of the structure of MSH bound to an enzyme would greatly enhance our understanding of the interactions which are essential to MSH binding and would aid in the design of MSH-based enzyme inhibitors.

### 3.5. References

1. Newton, G. L.; Arnold, K.; Price, M. S.; Sherrill, C.; Delcardayre, S. B.; Aharonowitz, Y.; Cohen, G.; Davies, J.; Fahey, R. C.; Davis, C. *J. Bacteriol.* **1996**, *178*, 1990-1995.

2. Buchmeier, N. A.; Newton, G. L.; Koledin, T.; Fahey, R. C. *Mol. Microbiol.* **2003**, *47*, 1723-1732.
3. Vogt, R. N.; Steenkamp, D. J.; Zheng, R.; Blanchard, J. S. *Biochem. J.* **2003**, *374*, 657-666.
4. Hand, C. E.; Honek, J. F. *J. Nat. Prod.* **2005**, *62*, 293-308, and references therein.
5. Newton, G. L.; Fahey, R. C. *Arch. Microbiol.* **2002**, *178*, 388-394.
6. Newton, G. L.; Av-Gay, Y.; Fahey, R. C. *Biochemistry* **2000**, *39*, 10739-10746.
7. Steffek, M.; Newton, G. L.; Av-Gay, Y.; Fahey, R. C. *Biochemistry* **2003**, *42*, 12067-12076.
8. Rawat, M.; Uppal, M.; Newton, G.; Steffek, M.; Fahey, R. C.; Av-Gay, Y. *J. Bacteriol.* **2004**, *186*, 6050-6058.
9. Leach, A. R. *Molecular Modelling: Principles and Applications*; Second ed.; Prentice Hall: Harlow, Essex, England, 2001.
10. Young, D. C. *Computational Chemistry: A Practical Guide for Applying Techniques for Real World Problems*; Wiley Interscience: New York, 2001.
11. Chang, G.; Guida, W. C.; Still, W. C. *J. Am. Chem. Soc.* **1989**, *111*, 4379-4386.
12. Mahadevan, J.; Nicholas, G. M.; Bewley, C. A. *J. Org. Chem.* **2003**, *68*, 3380-3386.
13. Cramer, C. J. *Essentials of Computational Chemistry*; Second ed.; John Wiley and Sons Ltd.: West Sussex, 2004.
14. Allinger, N. L.; Chen, K.; Katzenelbogen, J. A.; Wilson, S. R.; Anstead, G. M. *J. Comput. Chem.* **1996**, *17*, 747-755.
15. Allinger, N. L. *J. Am. Chem. Soc.* **1977**, *99*, 8127-8134.
16. Allinger, N. L.; Yuh, Y. H.; Lii, J. H. *J. Am. Chem. Soc.* **1989**, *111*, 8551-8566.
17. Allinger, N. L.; Li, F.; Yan, L. *J. Comput. Chem.* **1990**, *11*, 848-867.
18. Allinger, N. L.; Li, F.; Yan, L.; Tai, J. C. *J. Comput. Chem.* **1990**, *11*, 868-895.
19. Allinger, N. L.; Chen, K.; Lii, J. H. *J. Comput. Chem.* **1996**, *17*, 642-668.
20. Lii, J. H.; Allinger, N. L. *J. Am. Chem. Soc.* **1989**, *111*, 8566-8582.
21. Nevins, N.; Chen, K.; Allinger, N. L. *J. Comput. Chem.* **1996**, *17*, 730-746.
22. Nevins, N.; Chen, K.; Allinger, N. L. *J. Comput. Chem.* **1996**, *17*, 669-694.
23. Nevins, N.; Chen, K.; Allinger, N. L. *J. Comput. Chem.* **1996**, *17*, 695-729.
24. McDonald, D. Q.; Still, W. C. *Tetrahedron Lett.* **1992**, *33*, 7743-7746.
25. Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179-5197.
26. Halgren, T. A. *J. Comput. Chem.* **1999**, *20*, 730-748.
27. Jorgensen, W. L.; Tiradorives, J. *J. Am. Chem. Soc.* **1988**, *110*, 1657-1666.
28. Kaminski, G. A.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L. *J. Phys. Chem. B.* **2001**, *105*, 6474-6487.
29. Qiu, D.; Shenkin, P. S.; Hollinger, F. P.; Still, W. C. *J. Phys. Chem. A* **1997**, *101*, 3005-3014.
30. Schrödinger Inc. *MacroModel 9.0 XCluster Manual*, 2005.
31. Ferguson, D. M.; Raber, D. J. *J. Am. Chem. Soc.* **1989**, *111*, 4371-4378.
32. Fabricius, J.; Engelsen, S. B.; Rasmussen, K. *J. Carbohydr. Chem.* **1997**, *16*, 751-772.
33. Newton, G. L.; Ta, P.; Bzymek, K. P.; Fahey, R. C. *J. Biol. Chem.* **2006**, *281*, 33910-33920.
34. Fetterolf, B.; Bewley, C. A. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 3785-3788.

35. Nicholas, G. M.; Eckman, L. L.; Newton, G. L.; Fahey, R. C.; Ray, S.; Bewley, C. A. *Bioorg. Med. Chem.* **2003**, *11*, 601-608.
36. Knapp, S.; Amorelli, B.; Darout, E.; Ventocilla, C. C.; Goldman, L. M.; Huhn, R. A.; Minnihan, E. C. *J. Carbohydr. Chem.* **2005**, *24*, 103-130.
37. Knapp, S.; Gonzalez, S.; Myers, D. S.; Eckman, L. L.; Bewley, C. A. *Org. Lett.* **2002**, *4*, 4337-4339.
38. van Giessen, A. E.; Straub, J. E. *J. Chem. Phys.* **2005**, *122*, Article No. 024904.
39. DeMatteo, M. P.; Snyder, N. L.; Morton, M.; Baldisseri, D. M.; Hadad, C. M.; Peczuha, M. W. *J. Org. Chem.* **2005**, *70*, 24-38.
40. Perez, S.; Kouwijzer, C. E.; Mazeau, K.; Engelsens, S. B. *J. Mol. Graph.* **1996**, *14*, 307-321.
41. Schrodinger Inc. *MacroModel 9.0 User Manual*, 2005.
42. Damm, W.; Frontera, S.; Tirado-Rives, J.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1997**, *119*, 1955-1970.
43. Vishnyakov, A.; Widmalm, G.; Kowalewski, J.; Laaksonen, A. *J. Am. Chem. Soc.* **1999**, *121*, 5403-5412.
44. Conrad, P. B.; de Pablo, J. J. *J. Phys. Chem. A* **1999**, *103*, 4049-4055.
45. Umemura, M.; Hayashi, S.; Nakagawa, T.; Urakawa, H.; Kajiwara, K. *J. Mol. Struct. (Theochem)* **2003**, *636*, 215-228.
46. Starikov, E. B.; Bräsicke, K.; Knapp, E. W.; Saenger, W. *Chem. Phys. Lett.* **2001**, *336*, 504-510.
47. Perez, S.; Imberty, A.; Engelsens, S. B.; Gruza, J.; Mazeau, K.; Jimenez-Barbero, J.; Poveda, A.; Espinosa, J. F.; van Eyck, B. P.; Johnson, G.; French, A. D.; Louise, M.; Kouwijzer, C. E.; Grootenuis, P. D. J.; Bernardi, A.; Raimondi, L.; Senderowitz, H.; Durier, V.; Vergoten, G.; Rasmussen, K. *Carbohydr. Res.* **1998**, *314*, 141-155.
48. Johnson, L. N. *Acta Crystallogr.* **1966**, *21*, 885-891.
49. Boyd, J.; Porteous, R.; Soffe, N.; Delepierre *Carbohydr. Res.* **1985**, *139*, 35-46.
50. Abraham, R. J.; Byrne, J. J.; Griffiths, L.; Koniotou, R. *Magn. Reson. Chem* **2005**, *43*, 611-624.
51. Rabinowitz, H. N.; Kraut, J. *Acta Crystallogr.* **1964**, *17*, 159-168.
52. Brocca, P.; Bernardi, A.; Raimondi, L.; Sonnino, S. *Glycoconjugate J.* **2000**, *17*, 283-299.
53. Bernardi, A.; Colombo, A.; Sánchez-Medina, I. *Carbohydr. Res.* **2004**, *339*, 967-973.
54. Boström, J.; Norrby, P.-O.; Liljefors, T. *J. Comput. Aided Mol. Des.* **1998**, *12*, 383-396.
55. Nicklaus, M. C.; Wang, S.; Driscoll, J. S.; Milne, G. W. A. *Bioorg. Med. Chem.* **1995**, *3*, 411-428.
56. Perola, E.; Charifson, P. S. *J. Med. Chem.* **2004**, *47*, 2499-2510.
57. Smith, L.; Zachariah, C.; Thirumoorthy, R.; Rocca, Jim; Novák, J.; Hillman, J. D.; Edison, A. S. *Biochemistry* **2003**, *42*, 10372-10384.
58. Greenidge, P. A.; Mérette, S. A. M.; Beck, R.; Dodson, G.; Goodwin, C.; Scully, M. F.; Spencer, J.; Wiser, J.; Deadman, J. J. *J. Med. Chem.* **2003**, *46*, 1293-1305.
59. Thirumoorthy, R.; Holder, J. R.; Bauzo, R. M.; Richards, N. G. J.; Edison, A. S.; Haskell-Luevano, C. *J. Med. Chem.* **2001**, *44*, 4114-4124.
60. Simon, C.; Barathieu, K.; Laguerre, M.; Schmitter, J.-M.; Fouquet, E.; Pianet, I.; Dufourc, E. J. *Biochemistry* **2003**, *42*, 10385-10395.
61. Imberty, A.; Perez, S. *Chem. Rev.* **2000**, *100*, 4567-4588.

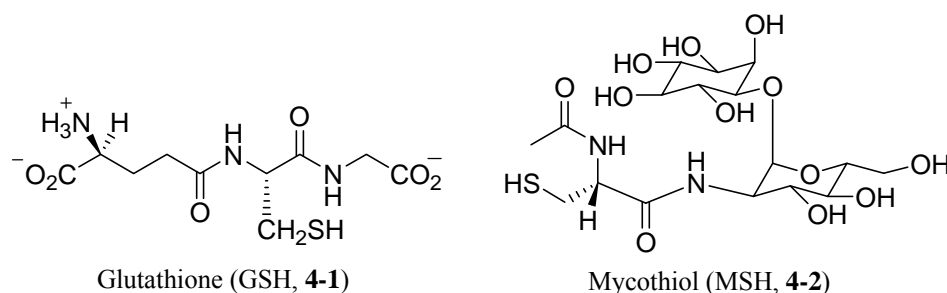
62. Rawat, M.; Newton, G. L.; Ko, M.; Martinez, G. J.; Fahey, R. C.; Av-Gay, Y. *Antimicrob. Agents Chemother.* **2002**, *46*, 3348-3355.
63. Nicholas, G. M.; Eckman, L. L.; Ray, S.; Hughes, R. O.; Pfefferkorn, J. A.; Barluenga, S.; Nicolaou, K. C.; Bewley, C. A. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 2487-2490.
64. Metaferia, B. B.; Ray, S.; Smith, J. A.; Bewley, C. A. *Bioorg. Med. Chem. Lett.* **2006**, doi:10.1016/j.bmcl.2006.1010.1031.
65. Fairlamb, A. H.; Blackburn, P.; Ulrich, P.; Chait, B. T.; Cerami, A. *Science* **1985**, *227*, 1485-1487.
66. Fairlamb, A. H.; Cerami, A. *Mol. Biochem. Parasitol.* **1985**, *14*, 187-198.
67. Steenkamp, D. J. *IUBMB Life* **2002**, *53*, 243-248.
68. Garrard, E. A.; Borman, E. C.; Cook, B. N.; Pike, E. J.; Alberg, D. G. *Org. Lett.* **2000**, *2*, 3539-3542.
69. Tromelin, A.; Moutiez, M.; Meziane-Cherif, D.; Aumercier, M.; Tartar, A.; Sergheraert, C. *Bioorg. Med. Chem. Lett.* **1993**, *3*, 1971-1976.
70. Kuriyan, J.; Kong, X. P.; Krishna, T. S.; Sweet, R. M.; Murgolo, N. J.; Field, H.; Cerami, A.; Henderson, G. B. *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 8764-8768.
71. Bailey, S.; Smith, K.; Fairlamb, A. H.; Hunter, W. N. *Eur J Biochem* **1993**, *213*, 67-75.
72. Hunter, W. N.; Bailey, S.; Habash, J.; Harrop, S. J.; Helliwell, J. R.; Aboagye-Kwarteng, T.; Smith, K.; Fairlamb, A. H. *J. Mol. Biol.* **1992**, *227*, 322-333.
73. Zhang, Y.; Bailey, S.; Naismith, J. H.; Bond, C. S.; Habash, J.; McLaughlin, P.; Papiz, M. Z.; Borges, A.; Cunningham, M.; Fairlamb, A. H.; et al. *J. Mol. Biol.* **1993**, *232*, 1217-1220.
74. Zhang, Y.; Bond, C. S.; Bailey, S.; Cunningham, M. L.; Fairlamb, A. H.; Hunter, W. N. *Protein Sci.* **1996**, *5*, 52-61.
75. Bond, C. S.; Zhang, Y.; Berriman, M.; Cunningham, M. L.; Fairlamb, A. H.; Hunter, W. N. *Structure Fold. Des.* **1999**, *7*, 81-89.



## Chapter 4: Efforts Towards the Identification of a Mycothiol-Utilizing Glyoxalase System

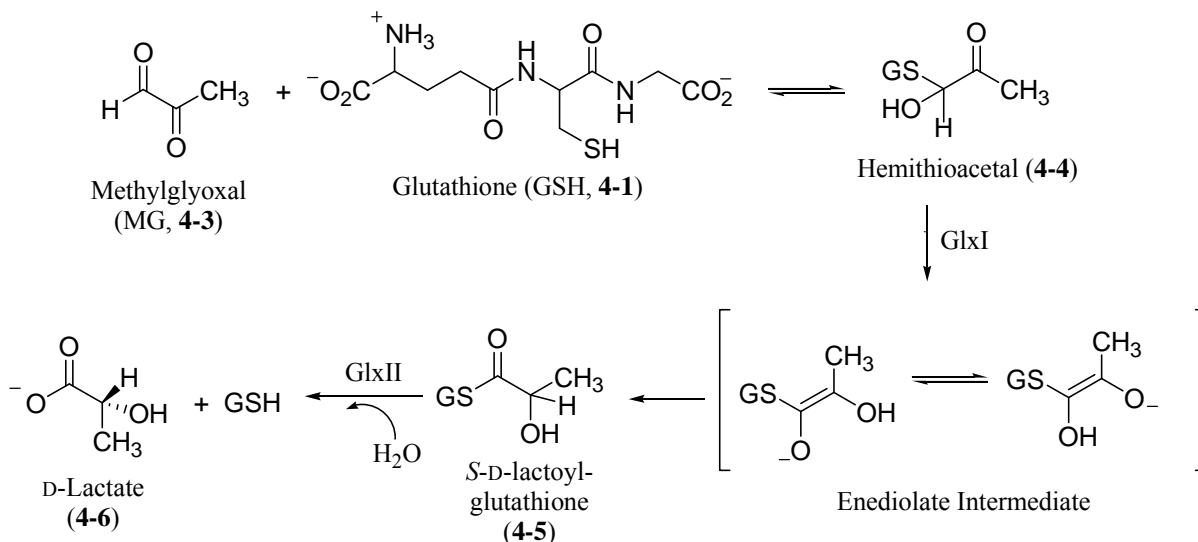
*Mycobacteria* are the causative agents of such diseases as tuberculosis (*Mycobacterium tuberculosis*),<sup>1</sup> leprosy (*Mycobacteria leprae*)<sup>2</sup> and Buruli ulcers (*Mycobacterium ulcerans*).<sup>3</sup> Combined, these diseases affect over 9 million people worldwide.<sup>1-3</sup> In the case of tuberculosis, over 1.69 million people succumbed to tuberculosis in 2004 alone. Drug resistant strains of *M. tuberculosis* have been documented in every country surveyed and strains of *M. tuberculosis* resistant to all major anti-tuberculin have emerged, threatening the efforts to control tuberculosis outbreaks.<sup>1</sup> The need for novel drug targets in *Mycobacteria* is essential for the future control of tuberculosis and other diseases caused by *Mycobacteria*.

Unlike humans and other eukaryotes, *Mycobacteria* and other *Actinomycetales* bacteria do not produce glutathione (GSH, **4-1**), but rather use the novel glycothiol mycothiol (MSH, **4-2**) as their major thiol.<sup>4</sup> Mutants of *M. tuberculosis* deficient in the biosynthesis of MSH are not viable<sup>5,6</sup> and MSH deficient mutants of *Mycobacteria smegmatis* are hypersensitive to toxic agents such as alkylating agents, free radicals, and antibiotics.<sup>7</sup> The MSH biosynthetic and metabolic pathways are therefore attractive targets for the development of novel therapeutics against *Mycobacteria*.<sup>8</sup>



**Figure 4.1:** Glutathione and mycothiol.

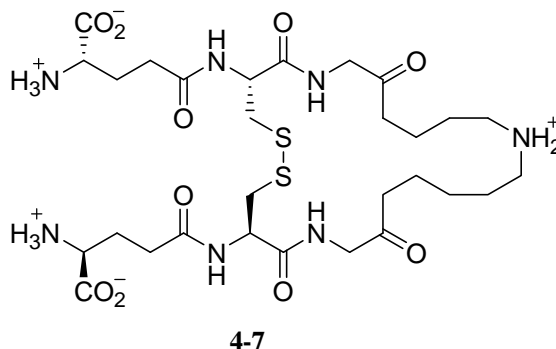
The known MSH-dependent enzymes parallel GSH pathways in function; hence the biochemistry of GSH may serve as a template for the discovery of new roles for MSH. Of particular interest to our laboratory is the possible existence of an MSH-dependent glyoxalase (Glx) system. The two enzyme Glx system (Figure 4.2) is responsible for the removal of cytotoxic methylglyoxal (MG, **4-3**) from the cell<sup>9</sup> and our laboratory has extensive experience in the study of the first step (GlxI) of bacterial Glx systems.<sup>10-15</sup> *Streptomyces coelicolor*, an organism known to produce MSH,<sup>4</sup> was chosen for our studies due to its non-pathogenicity, ease of handling and its complete genome having been sequenced.<sup>16</sup> The discovery of an MSH-dependent Glx system would provide a novel target for the development of anti-mycobacterial agents due to its essential role in the removal of cytotoxic MG.



**Figure 4.2:** The two-enzyme glyoxalase (Glx) system composed of GlxI and GlxII.

## 4.1. Introduction

The use of GSH as a template for the discovery of new thiol functionalities is not novel. The *Trypanosomatidae*, which include the causative agents of Chagas disease<sup>17</sup> and African sleeping sickness,<sup>18</sup> use trypanothione (TSH<sub>2</sub>, Figure 4.3, 4-7) as their major thiol.<sup>19,20</sup> The enzymes thus far discovered as TSH<sub>2</sub>-dependent parallel GSH-dependent enzymes, including peroxidase,<sup>21,22</sup> GlxI,<sup>23,24</sup> GlxII,<sup>25</sup> TSH<sub>2</sub>-S-transferase<sup>26,27</sup> and TSH<sub>2</sub> disulfide reductase.<sup>28-31</sup> It is expected that similar to TSH<sub>2</sub>, new MSH functions can be identified using GSH biochemistry as a guide.



**Figure 4.3:** Trypanothione (TSH<sub>2</sub>).

### 4.1.1. Parallel Functions of Glutathione and Mycothiol Utilizing Enzymes

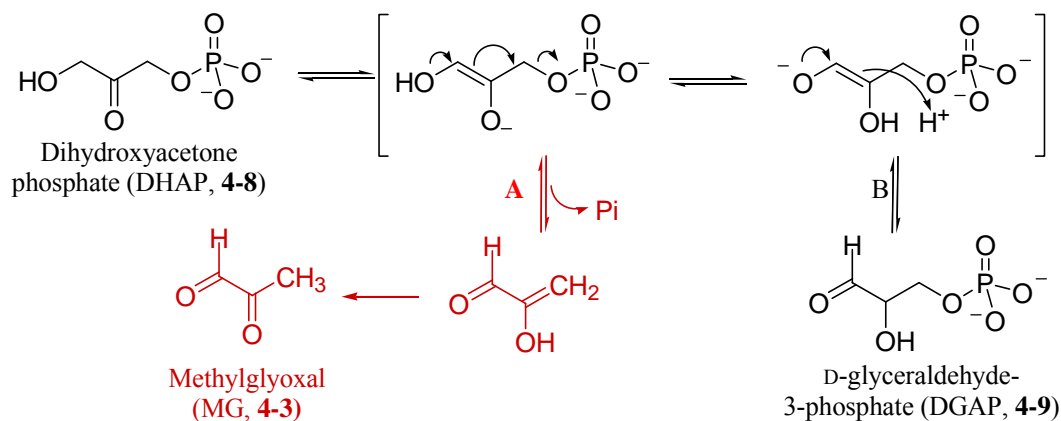
The four enzymes that have been identified to date as requiring MSH are maleyl pyruvate isomerase,<sup>32</sup> formaldehyde dehydrogenase,<sup>33-36</sup> MSH-S-conjugate amidase<sup>37,38</sup> and MSH disulfide (MSSM) reductase.<sup>39,40</sup> As seen in Table 4.1, all of these enzymes have counterparts in GSH biochemistry.<sup>41-44</sup> It is expected that many of the known GSH-dependent enzymes, such as GSH-S-transferase and the Glx system have MSH counterparts in the *Actinomycetales* bacteria.

**Table 4.1:** Comparison of known glutathione and mycothiol biochemistry.

<i>GSH Utilizing Enzyme</i>	<i>MSH-Utilizing Enzyme</i>
<b>Formaldehyde dehydrogenase</b> First step in the conversion of formaldehyde to formic acid <sup>45</sup> and is involved in the metabolism of <i>S</i> -nitroso-GSH <sup>46,47</sup> Found in animals, plants, certain bacteria <sup>45</sup>	<b>Formaldehyde dehydrogenase</b> First step in the conversion of formaldehyde to formic acid <sup>33-35</sup> and is involved in the metabolism of <i>S</i> -nitroso-MSH <sup>36</sup> Isolated from <i>Amycolaptopsis methanolica</i> , <i>Rhodococcus erythropolis</i> <sup>33-35</sup> and <i>Mycobacterium smegmatis</i> <sup>36</sup>
<b>GSH disulfide reductase</b> Reduces GSH disulfide to the free thiol form <sup>44</sup> Found in certain bacteria, fungi, plants protozoa and animals <sup>44</sup>	<b>MSH disulfide reductase</b> Reduces MSH disulfide to the free thiol form <sup>39,40,48</sup> Isolated from <i>M. tuberculosis</i> <sup>39,40,48</sup>
<b>GSH-S-conjugate metabolism</b> GSH-S-conjugates are converted to mercapturic acids, which are then excreted from the cell <sup>43</sup> Isolated from mammals, certain insects, plants, squid and bacteria <sup>49</sup>	<b>MSH-S-conjugate amidase</b> Hydrolyzes MSH-S-conjugates to form mercapturic acids which are then excreted from the cell <sup>37,38</sup> Isolated from <i>M. smegmatis</i> <sup>37,38</sup>
<b>Maleyl pyruvate isomerase</b> Converts maleyl pyruvate to fumaryl pyruvate during the metabolism of aromatic species <sup>41</sup> Identified in <i>Salmonella typhimurim</i> , <sup>50</sup> <i>Pseudomonas acidovorans</i> , <sup>51</sup> <i>Klebsiella pneumoniae</i> , <sup>52</sup> <i>Ralstonia</i> sp. Strain U2 <sup>53</sup> and <i>Morexella</i> OA3 <sup>54,55</sup>	<b>Maleyl pyruvate isomerase</b> Converts maleyl pyruvate to fumaryl pyruvate during the metabolism of aromatic species <sup>32</sup> Isolated from <i>Corynebacterium glutamicum</i> <sup>32</sup>

#### 4.1.2. A Brief Review of the Glutathione-Dependent Glyoxalase System

The two enzyme Glx system converts cytotoxic MG to D-lactate (**4-6**), using GSH as an essential cosubstrate (Figure 4.2). MG is produced during glycolysis when the deprotonation of dihydroxyacetone phosphate (DHAP, **4-8**) by triose phosphate isomerase (TIM) is followed by phosphate elimination rather than reprotonation (Figure 4.4).<sup>56,57</sup> While reprotonation occurs 10<sup>6</sup> times more frequently than phosphate elimination, the high cellular concentration of TIM can result in the accumulation of large concentrations of MG.<sup>56,57</sup> In microorganisms, MG is also produced directly from DHAP in a by-pass to phosphorylating glycolysis.<sup>58,59</sup> When organisms have limited access to phosphate, DHAP can accumulate and MG synthase, in conjunction with the Glx system, may serve as a pathway for the formation of pyruvate with the release of phosphate. The cellular concentration of MG must be tightly controlled as it can form adducts with proteins and nucleic acids, and can inhibit protein biosynthesis.<sup>60</sup>



**Figure 4.4:** The isomerization catalyzed by TIM: dihydroxyacetone phosphate (DHAP, 4-8) is converted to D-glyceraldehyde-3-phosphate (DGAP, 4-9) via a deprotonation and reprotonation mechanism. When inorganic phosphate (Pi) is eliminated (pathway A) methylglyoxal (MG, 4-3) is formed<sup>56</sup>

The true substrate of the first enzyme in the Glx pathway, GlxI, is the hemithioacetal (4-4) formed non-enzymatically by the reaction of MG with GSH (Figure 4.2).<sup>9</sup> This hemithioacetal is deprotonated and reprotonated by GlxI to yield *S*-D-lactoylglutathione (4-5), which is then hydrolyzed by GlxII to yield D-lactate (4-6) and GSH (4-1, Figure 4.2). GlxI has been characterized or identified in humans,<sup>61-63</sup> mice,<sup>64</sup> yeasts,<sup>65,66</sup> plants,<sup>67-74</sup> insects,<sup>9</sup> protozoa,<sup>75</sup> fungi<sup>76</sup> and in at least 42 strains of bacteria,<sup>9,10</sup> including *Escherichia coli*.<sup>14</sup> GlxII, however, has not been as thoroughly characterized. It has been identified in humans,<sup>45,77</sup> rats,<sup>78,79</sup> yeasts,<sup>80,81</sup> plants<sup>82-85</sup> and *Candida albicans*<sup>86</sup> and has been cloned from humans,<sup>87</sup> *Arabidopsis thaliana*,<sup>88</sup> *Saccharomyces cerevisiae*<sup>89</sup> and *Neisseria meningitidis*.<sup>90</sup> In addition, a single enzyme, termed GlxIII, has been identified in *E. coli*, as catalyzing the conversion of MG to D-lactate without the need for a thiol cofactor.<sup>91</sup> Due to its sensitivity to thiol blocking agents, it is believed that an active site thiol group may be involved in the reaction.<sup>91</sup> The identification of the Glx system in such a diverse sampling of organisms makes it appear likely that such a system would exist in the *Actinomycetales* bacteria.

#### 4.1.3. The Trypanothione-Dependent Glyoxalase System

The Glx system has also been identified in *Leishmania*<sup>24,25,92-94</sup> and *Trypanosoma*<sup>23,95</sup> and is reliant upon TSH<sub>2</sub> as an essential cosubstrate, in lieu of GSH. The TSH<sub>2</sub>-dependent GlxI enzyme is specific for the hemithioacetal formed from TSH<sub>2</sub> and MG<sup>24</sup> and forms mono- and bis-*S*-D-lactoyltrypanothione,<sup>95</sup> the substrates of TSH<sub>2</sub>-dependent GlxII. The TSH<sub>2</sub>-dependent GlxI enzyme from *Leishmania major* demonstrates high substrate specificity for TSH<sub>2</sub>-based hemithioacetals and is insensitive to GSH based inhibitors. It is thought that the structural differences between the human and *L. major* GlxI enzymes may be responsible for the specificity of the *L. major* enzyme. This specificity makes TSH<sub>2</sub>-dependent GlxI an attractive target for design of novel therapeutics against the pathogenic *Trypanosoma*.<sup>24</sup> The identification of a non-GSH dependent Glx system reinforces our belief that an MSH-dependent system exists and its probable divergent substrate specificity from that of humans, would make it an excellent target for the development of novel anti-microbial agents targeted against *Mycobacteria*. In addition, a biochemical comparison of the identified enzyme to the previously characterized GSH, and TSH<sub>2</sub> dependent Glx enzymes would be of interest as an example of convergent evolution.

#### 4.1.4. Methods Used for the Detection of the Glyoxalase System

The assays developed for the detection of GlxI and GlxII activity rely on spectrophotometric methods to monitor the formation and degradation of their respective products. *S*-D-Lactoyl-glutathione<sup>96-98</sup> and trypanothione,<sup>23,24</sup> formed by GlxI, display a characteristic absorbance at 240 nm, which can easily be monitored to quantify GlxI activity.<sup>24,99</sup> The formation of D-lactate, the product of GlxII, cannot be directly observed by absorbance; however, GlxII activity can still be monitored spectrophotometrically. The simplest method to quantify GlxII activity is to monitor the decrease in absorbance seen at 240 nm, which corresponds to the consumption of *S*-D-lactoylglutathione and *S*-D-lactoyltrypanothione.<sup>25,95,99,100</sup> The free thiol released by GlxII with each turnover can also be trapped by a thiol reactive reagent such as Ellman's reagent [5,5'-dithio-bis(2-nitrobenzoic acid), DTNB], which upon reaction, releases a chromophore having a characteristic absorbance at 412 nm.<sup>95,101</sup> Finally, the formation of D-lactate can be followed using D-lactate dehydrogenase (D-LDH). D-LDH converts D-lactate to pyruvate with the simultaneous reduction of NAD<sup>+</sup> to NADH (Eq. 4.1); the latter displays a characteristic absorbance at 340 nm.<sup>95,102</sup> To quantitatively monitor GlxII activity, the pyruvate formed by D-LDH must be removed from the equilibrium; therefore trapping agents, such as hydrazine, are often used in conjugation with basic pH (i.e., pH 9).<sup>102</sup> It is important to monitor for GlxII activity, as opposed to solely that of GlxI, when studying cell-free extracts as the activity of GlxII may be greater than that of GlxI, making the observation of a change at 240 nm, and the formation of the *S*-D-lactoyl intermediate, difficult. The use of an LDH assay would also be effective in the detection of a GlxIII enzyme, which converts MG directly to D-lactate without the need of GSH.<sup>91</sup>



Glx activity can be successfully detected using the described methods; therefore if an MSH-dependent Glx system, or a GlxIII enzyme, exists in *S. coelicolor*, it should be detectable using these techniques.

#### 4.1.5. Introduction to *Streptomyces* Bacteria

Unlike the *Mycobacteria*, the *Streptomyces* bacteria are of little clinical interest;<sup>\*</sup> however, they are industrially important producers of a large number of antibiotics.<sup>103</sup> Approximately two-thirds of the known antibiotics produced by microorganisms are attributable to the *Actinomycetales* bacteria and nearly 80 % of these antibiotics are produced by *Streptomyces* bacteria. A sampling of these antibiotics is shown Table 4.2. Similar to other *Actinomycetales* bacteria, *Streptomyces* have been shown to produce MSH as their major intracellular thiol (Table 4.3)<sup>4</sup> and the biosynthetic enzymes responsible for MSH production in *S. coelicolor* have been identified and their activities confirmed experimentally.<sup>104</sup> The complete genome of *S. coelicolor* has been sequenced<sup>16</sup> and is available for data mining.<sup>†</sup> *S. coelicolor* is an attractive target for our investigations into the presence of an MSH-dependent Glx system as it is known to produce MSH. The availability of its complete genome sequence would also help in the development of an over-expression system for any Glx type enzymes identified.

\* *Streptomyces somaliensis* is the only known human pathogen of this genus.

† [http://www.sanger.ac.uk/Projects/S\\_coelicolor/](http://www.sanger.ac.uk/Projects/S_coelicolor/)

**Table 4.2:** Useful antibiotics produced by *Streptomyces* bacteria.<sup>103</sup>

<i>Antibiotic</i>	<i>Producer</i>	<i>Application</i>	<i>Antibiotic</i>	<i>Producer</i>	<i>Application</i>
Actinomycin D	<i>S. spp.</i>	Antitumour	Mitomycin C	<i>S. caespitosus</i> <i>S. verticillus</i>	Antitumour
Avermectin	<i>S. avermitilis</i>	Antiparasitic	Neomycin	<i>S. fradiae</i>	Antibacterial
Bialaphos	<i>S. hygroscopicus</i>	Herbicidal	Novobiocin	<i>S. niveus</i>	Antibacterial
Bleomycin	<i>S. verticillus</i>	Antitumour	Oleandomycin	<i>S. antibioticus</i>	Antibacterial
Candicidin	<i>S. griseus</i>	Antifungal	Oxytetracyclin	<i>S. rimosus</i>	Antibacterial
Chloramphenicol	<i>S. venezuelae</i>	Antibacterial	Phleomycin	<i>S. verticillus</i>	Antitumour
Chlortetracycline	<i>S. aureofaciens</i>	Antibacterial	Pristinamycin	<i>S. pristinaespiralis</i>	Antibacterial
Cycloserine	<i>S. orchidaceus</i>	Antibacterial	Spectinomycin	<i>S. spectabilis</i>	Antibacterial
Daptomycin	<i>S. roseosporus</i>	Antibacterial	Spiramycin	<i>S. ambofaciens</i>	Antibacterial
Fosfomycin	<i>S. spp.</i>	Antibacterial	Streptogramins	<i>S. graminofaciens</i>	Antibacterial
Kanamycin	<i>S. kanamyceticus</i>	Antibacterial	Streptomycin	<i>S. griseus</i>	Antibacterial
Lincomycin	<i>S. lincolnensis</i>	Antibacterial	Tetracycline	<i>S. aureofaciens</i>	Antibacterial
Milbemycin	<i>S. hygroscopicus</i>	Antiparasitic	Thienamycin	<i>S. cattleya</i>	Antibacterial
Mithramycin	<i>S. argillaceus</i>	Antitumour	Tobramycin	<i>S. tenebrarius</i>	Antibacteria

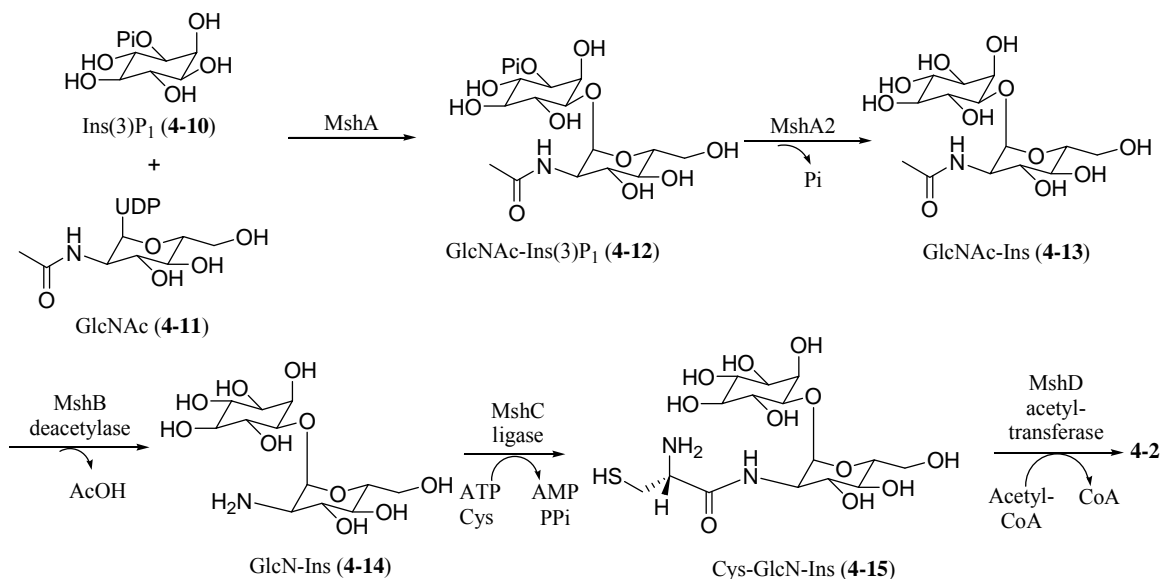
**Table 4.3:** Thiols produced by *Streptomyces* bacteria.<sup>105</sup>

<i>Strains</i>	<i>Thiol (μmol/g [residual dry weight])<sup>a</sup></i>					
	<i>MSH</i>	<i>Cys</i>	<i>GSH</i>	<i>H<sub>2</sub>S</i>	<i>CoA</i>	<i>Other</i>
<i>Streptomyces clavuligerus</i>	5.40	0.78	<0.01	1.3	0.8	1.55 <sup>b</sup>
<i>Streptomyces coelicolor</i>	2.80	0.56	<0.02	0.75	0.5	<0.02 <sup>c</sup>
<i>Streptomyces jumonjinensis</i>	6.70	1.30	<0.09	0.83	1.9	1.1 <sup>b</sup>
<i>Streptomyces lactamdurans</i>	5.90	1.30	<0.08	1.3	1.1	23.05 <sup>d</sup>
<i>Streptomyces lividans</i>	6.60	0.58	<0.06	1.3	1.0	0.51 <sup>b</sup>

<sup>a</sup>MSH: mycothiol, Cys: cysteine, GSH: glutathione, H<sub>2</sub>S: hydrogen sulfide, CoA: coenzyme A, adapted from Newton *et al.* (1996);<sup>4</sup> <sup>b</sup>δ-(L-α-amino-adipyl)-L-cysteiny-D-valine (ACV) and an unknown thiol; <sup>c</sup>ACV; <sup>d</sup>ACV and ergothioneine.

#### 4.1.6. The Synthesis of Mycothiol

MSH is not commercially available and the enzymes involved in its biosynthesis have not yet been fully characterized (Figure 4.5); hence, production of large quantities of MSH through the over-expression of its biosynthetic enzymes is not possible at this time. The synthetic preparation of MSH is not a trivial matter. The presence of the *myo*-inositol ring requires extensive protection and deprotection and the correct stereochemistry of the resulting product is essential for its use in enzymatic assays.<sup>106,107</sup> This latter criterion is further complicated by the reversal of the numbering conventions for *myo*-inositol published during the 1970s: what was previously considered an L- configuration of the *myo*-inositol ring is now labeled as D<sup>108</sup> (see *Chapter 5* for a more detailed discussion of *myo*-inositol stereochemistry). This reversal led to confusion in the initial structure determination of MSH: the solved structure was labeled as a D structure but was depicted as L.<sup>105,109</sup> This error was perpetuated in papers detailing partial syntheses of MSH, as well as enzyme and substrate specificity. This issue was resolved by the total synthesis of the bimane derivative of MSH (MSmB) and the characterization of MSH-*S*-conjugate amidase, an enzyme known to accept MSmB as a substrate.<sup>107</sup> MSH was unambiguously determined to have the original reported stereochemistry of 1-D-*myo*-inosityl-2-deoxy-2-(*N*-acetamido-L-cysteinamido)- $\alpha$ -D-glucopyranoside.<sup>107</sup>

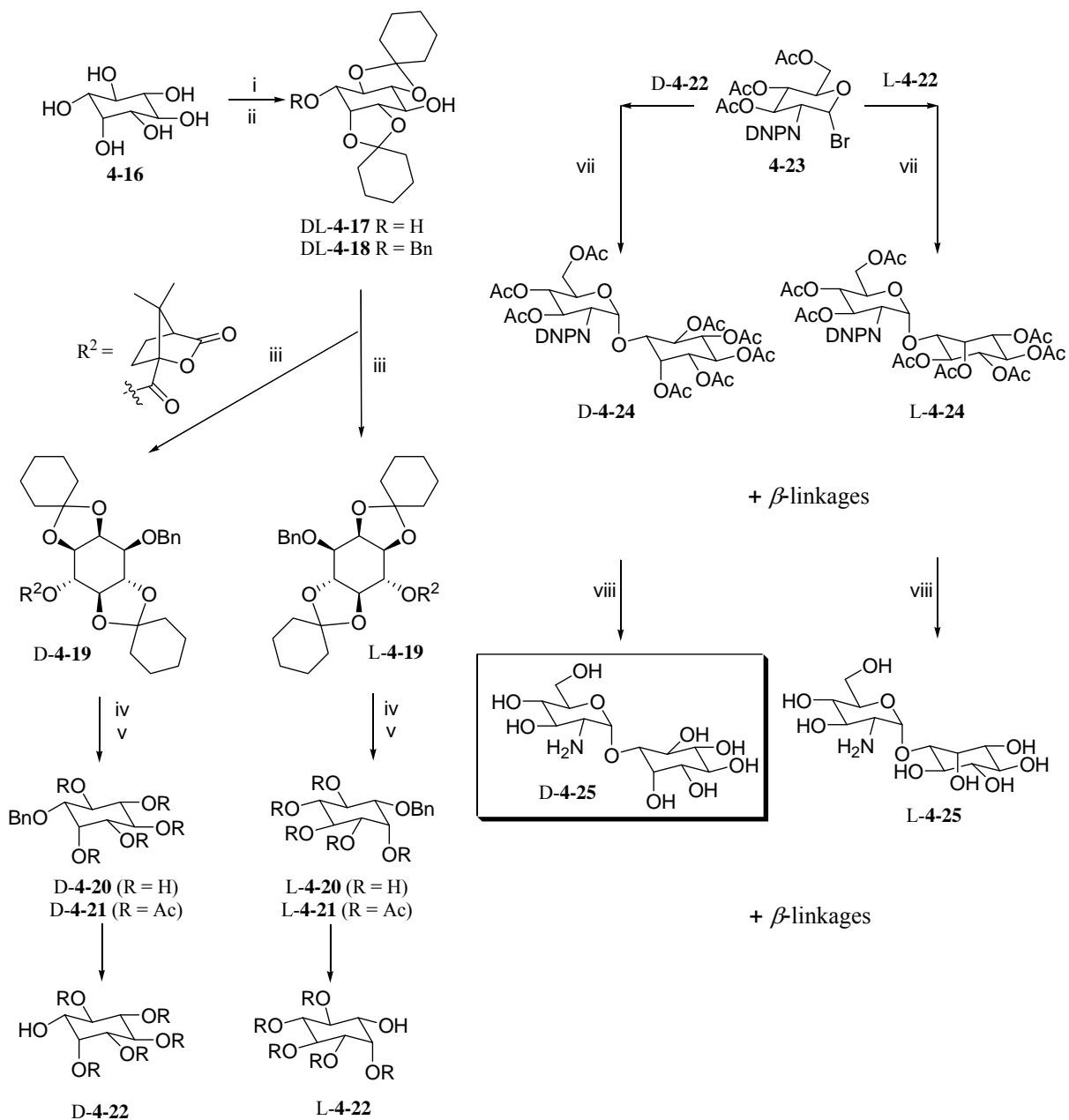


**Figure 4.5:** The biosynthesis of mycothiol. While the activity of the second enzyme in this pathway, MshA2, has been identified, the enzyme has not.

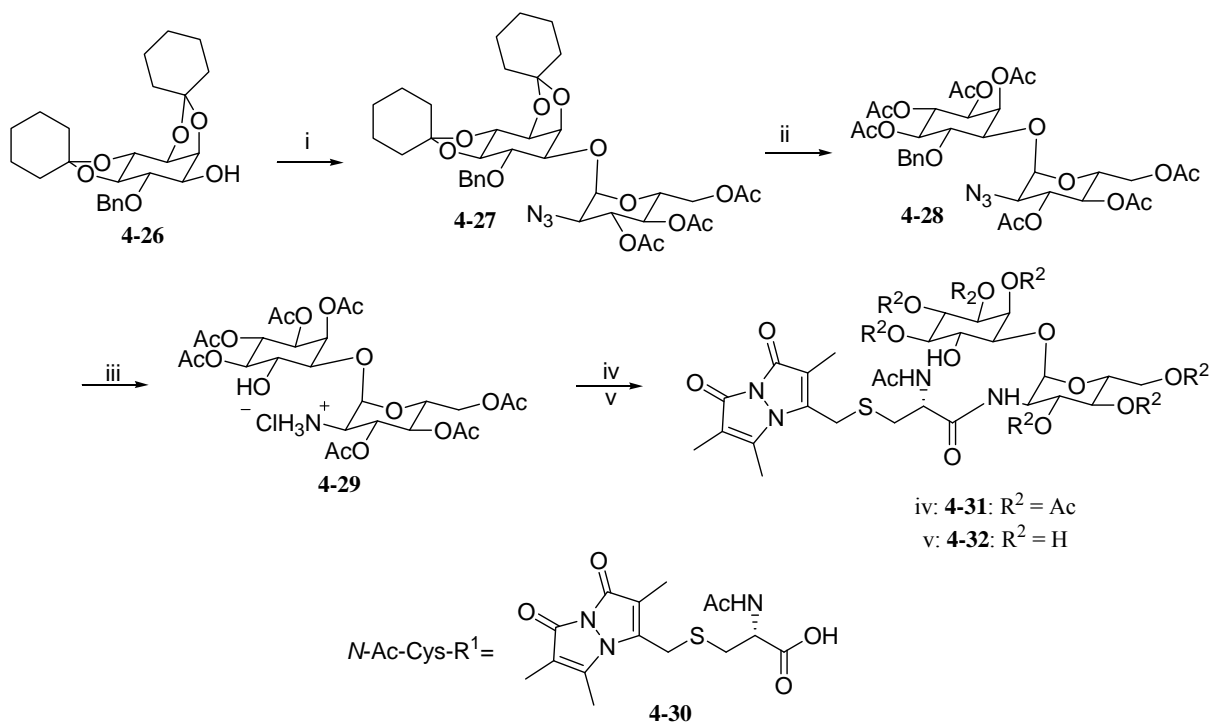
The reported syntheses of MSH are inefficient modes for the production of MSH. The first published synthesis of MSH by Jardine *et al.* (Figure 4.6) had an overall yield of less than 2 % over 6 steps<sup>110</sup> The key step, the coupling of cysteine to the disaccharide was accomplished using the cell-free extract of *M. smegmatis*<sup>111</sup> to yield a mixture of MSH and des-acetyl-MSH.<sup>110</sup> Unfortunately, this synthesis was published before the clarification of the stereochemistry of MSH, which casts doubt over the results of this method. The published synthesis of MSmB,<sup>107</sup> used to determine the stereochemistry of MSH (Figure 4.7), could be used as a basis for the synthesis of MSH. The alkylation of the cysteine by monobromobimane could be replaced by the protection of the sulfur atom followed by

deprotection as the final step. Unfortunately the reported overall yield of MSmB was less than 1 %.<sup>107</sup> The first total synthesis of MSH and the MSH disulfide (MSSM) was reported by Lee and Rosazza in 2004 (Figure 4.8),<sup>112</sup> and the resultant product was confirmed to be of the correct stereochemistry; however, the overall yield was, again, less than 1%. These syntheses, though significant accomplishments, are not useful methods for the production of the quantities of MSH required for enzyme identification and characterization, i.e. the amounts needed for the monitoring of protein fractions from chromatographic steps or detailed kinetic studies.

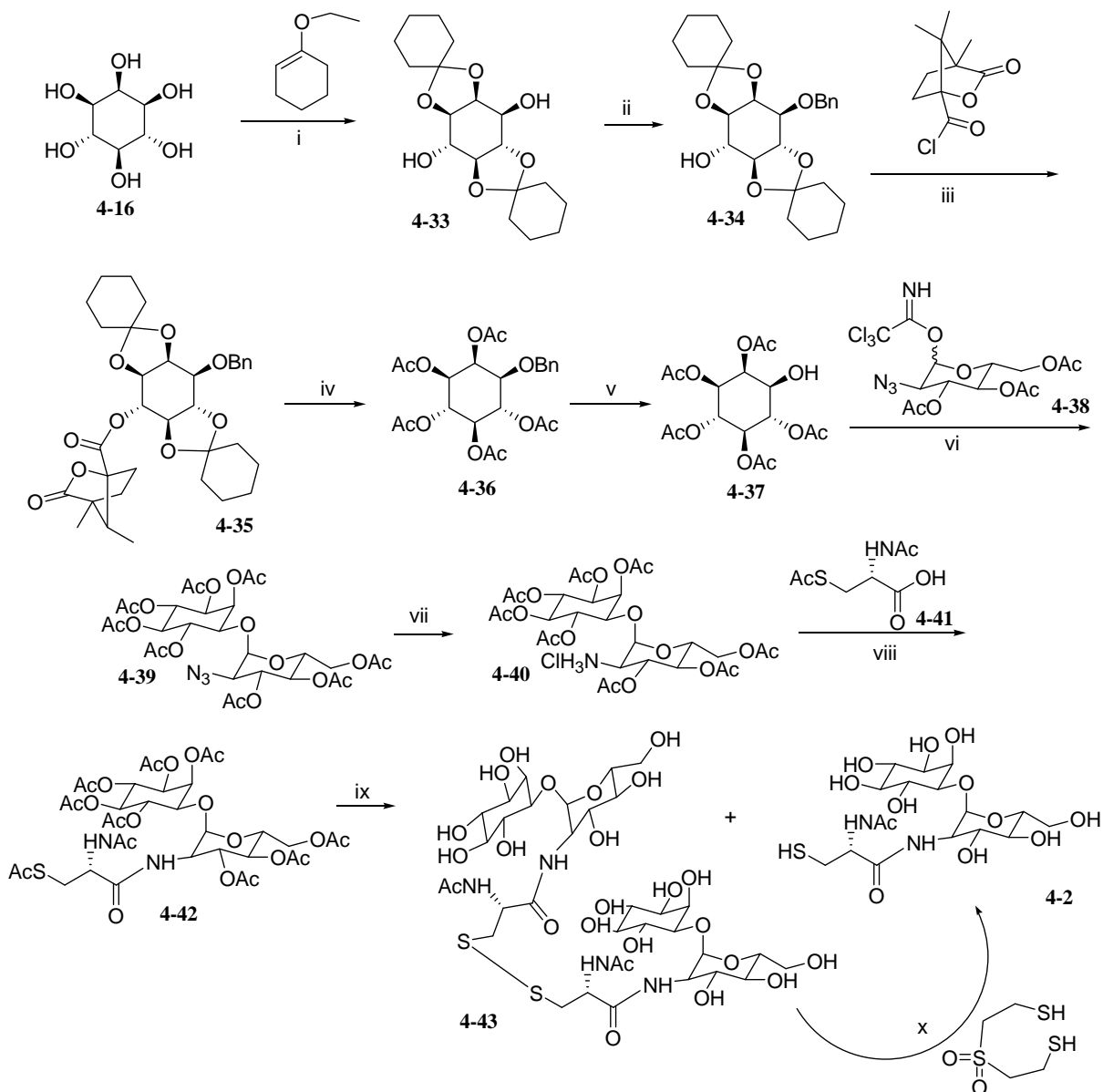




**Figure 4.6:** The synthesis of MSH as reported by Jardine *et al.*<sup>110</sup> Reagents: (i) 1-ethoxycyclohexene, *p*-toluenesulphonic acid and DMF; (ii) NaH, BnBr and toluene; (iii) camphanic acid chloride and NEt<sub>3</sub> in CH<sub>2</sub>Cl<sub>2</sub>; (iv) (a) KOH in ethanol; (b) 80 % acetic acid (v/v); (v) acetic anhydride and pyridine; (vi) H<sub>2</sub> and Pd/C; (vii) AgOTf, 2,6-*tert*-dibutylpyridine and CH<sub>2</sub>Cl<sub>2</sub>; (viii) Amberlite IR400 (OH). The addition of cysteine was achieved enzymatically through the addition of the boxed structure to crude cell-free extract of *Mycobacterium smegmatis*.



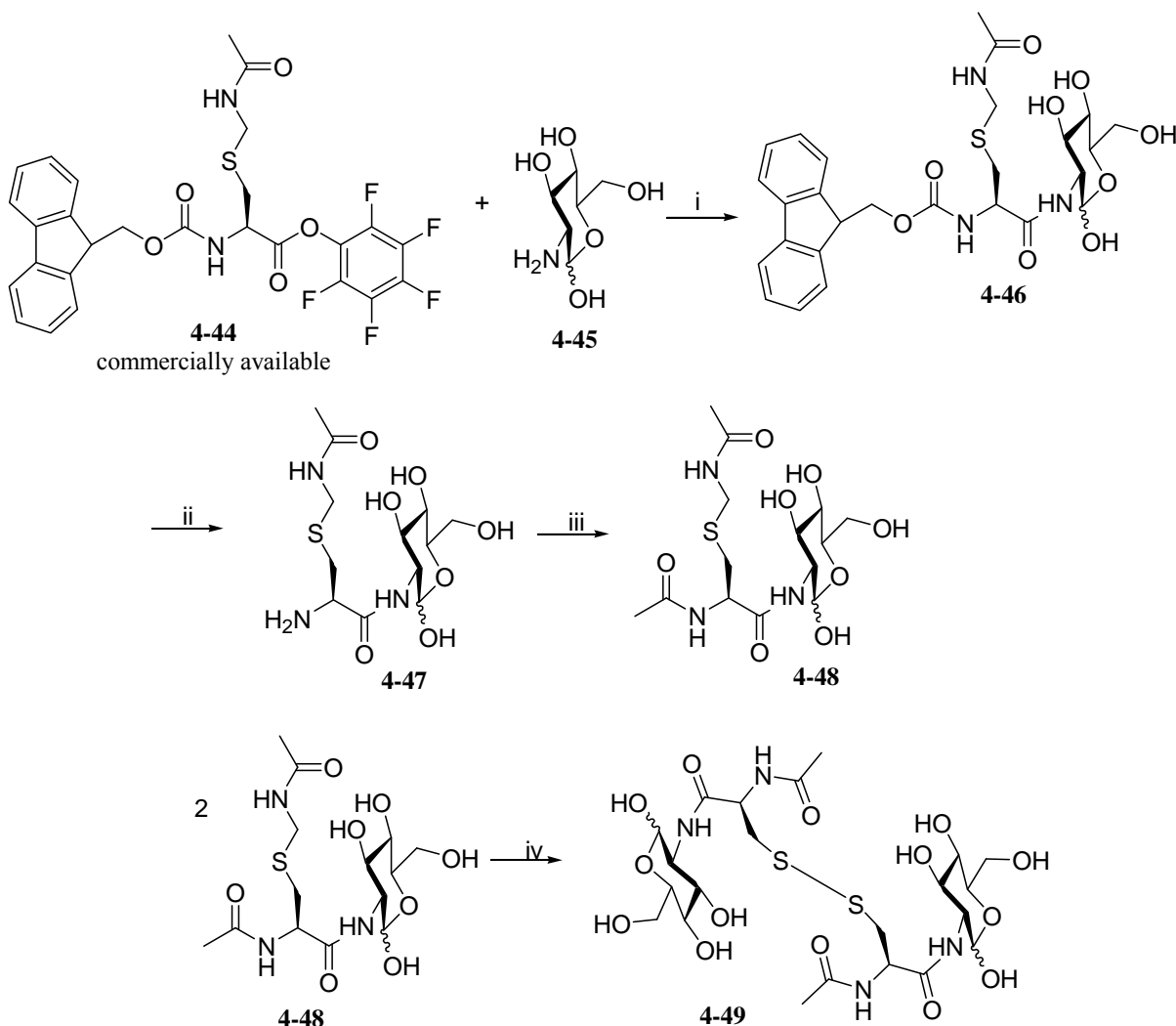
**Figure 4.7:** Synthesis of the bimane derivative of mycothiol as described by Nicholas *et al.*<sup>107</sup> Reagents: (i) 2,4,6-tri-*O*-acetyl-2-azido-2-deoxy- $\alpha$ -D-glucopyranosyl chloride, AgOTf, 2,6-diisopropyl-4-methyl-pyridine, CH<sub>2</sub>Cl<sub>2</sub>; (ii) 1. ethyleneglycol, (+)-camphor sulfonic acid, CH<sub>3</sub>CN, 2. Ac<sub>2</sub>O, pyridine; (iii) Pd-C, H<sub>2</sub>, EtOAc; (iv) DEPC, iPr<sub>2</sub>EtN, DMF, *N*-Ac-Cys-R<sup>1</sup> (shown in above figure); (v) Mg(OMe)<sub>2</sub>, MeOH (dry).



**Figure 4.8:** Synthesis of mycothiol as reported by Lee and Rosazza.<sup>112</sup> Reagents: (i) *p*-TsOH, CH<sub>2</sub>Cl<sub>2</sub>; (ii) BnBr, NaH, PhMe; (iii) DMAP, TEA, CH<sub>2</sub>Cl<sub>2</sub>; (iv) 1. KOH, EtOH; 2. AcOH, H<sub>2</sub>O then Ac<sub>2</sub>O, pyridine; (v) H<sub>2</sub>, Pd-C, EtOAc; (vi) TMSOTf, molecular sieves, CH<sub>2</sub>Cl<sub>2</sub>; (vii) H<sub>2</sub>, Pd-C, EtOAc, HCl; (viii) HATU, HOAt, collidine, DMF; (ix) Mg(OMe)<sub>2</sub>, MeOH; (x) H<sub>2</sub>O.

#### 4.1.7. The Use of Truncated Mycothiol as a Probe for Enzymatic Activity

Interest in identifying MSH-dependent enzymes preceded the publication of synthetic methods for the production of MSH. To avoid the complicated protection and deprotection strategies required by the inclusion of the *myo*-inositol moiety, an oxidized, truncated MSH analogue, des-*myo*-inositol mycothiol, was used to successfully identify MSSM reductase activity in *M. tuberculosis*. The truncated disulfide substrate provided sufficient features of MSSM to act as an active analogue. The synthesis of this truncated version reported an overall yield of ~ 9 % over four steps,<sup>113</sup> making it an attractive tool for initial attempts to identify MSH-dependent enzyme activity.

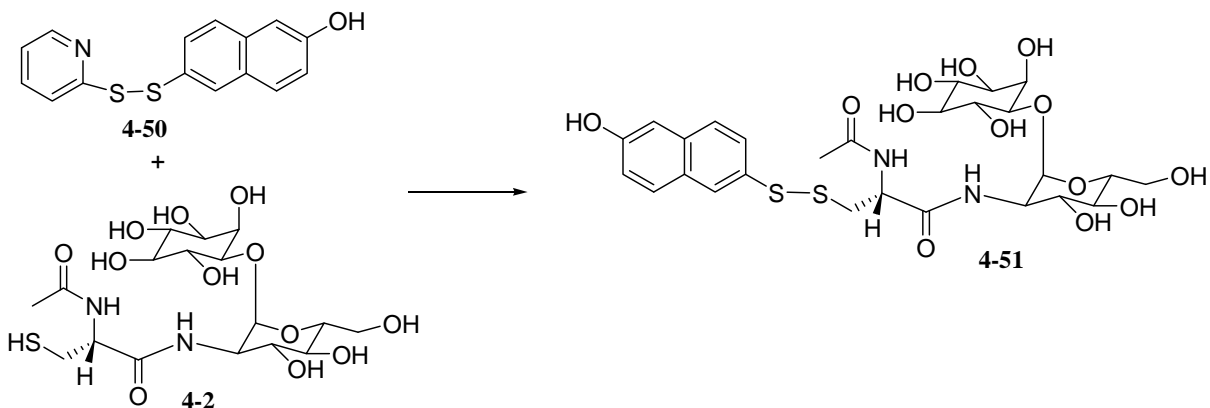


**Figure 4.9:** Synthesis of des-*myo*-inositol mycothiol as reported by Patel and Blanchard. Reagents: (i) HOBt,  $\alpha$ -D-glucosamine, DMF; (ii) 5% piperidine in DMF, (iii) acetic anhydride,  $K_2CO_3$ ,  $H_2O$ ; (iv)  $Tl(CF_3CO_2)_3$ , anisole, TFA.<sup>113</sup>

#### 4.1.8. Isolation of Mycothiol from Cell-Free Extracts

While a truncated version of MSH would be useful for identification, the accurate determination of an enzyme's kinetic parameters and full characterization requires the enzyme's true substrates, i.e. MSH. The synthetic methods for preparation of MSH are less than satisfactory for the production of sufficient quantities for enzyme characterization. MSH; therefore, MSH must be isolated from cell-free extracts. In the first protocol published for the isolation of MSH, it was purified directly from cell-free extracts using HPLC techniques.<sup>114</sup> This method is not necessarily useful for the isolation of large quantities of MSH, as the amount of cell-free extract to be processed would be prohibitively large for the average academic laboratory.<sup>115</sup> In addition, thiols can be easily oxidized to symmetric and mixed disulfides or to sulfinic acids and can be modified through addition reactions to aldehydes and activated double bonds. It is therefore desirable to remove the thiol from contaminants as soon as possible or to reversibly convert it to a more stable form, such as a disulfide of known composition.<sup>115</sup>

The second published protocol for the isolation of MSH addresses these issues by using a hydrophobic, thiol reactive agent, 2-*S*-(2'-thiopyridyl)-6-hydroxynaphthyl disulfide (Figure 4.10), to tag MSH for purification.<sup>115</sup> This hydrophobic tag increases the solubility of MSH in organic solvents and allows for its retention on a hydrophobic solid support resin, which partially purifies and concentrates the tagged MSH for easier purification by HPLC using C<sub>18</sub> reverse phase chromatography.<sup>115</sup> The tag can then be easily removed to return the free thiol.



**Figure 4.10:** The isolation of mycothiol using 2-*S*-(2'-thiopyridyl)-6-hydroxynaphthyl disulfide as described by Steenkamp and Vogt.<sup>115</sup>

#### 4.1.9. Plan of Action

The MSH-dependent enzymes identified to date parallel GSH-dependent enzymes in function. These enzymes can serve as targets for the development of novel therapeutics aimed at the pathogenic *Mycobacteria*. The extensive experience of our laboratory in studying bacterial GlxI enzymes has led us to focus on the possibility of an MSH-dependent Glx system and *S. coelicolor* was selected as our organism of interest. Initial assays were performed to determine if the endogenous levels of MSH would be sufficient to detect Glx-type activity. A control experiment using *E. coli*, an organism known to contain a GSH-dependent Glx system, was also performed to determine if Glx activity could be observed in this fashion. It was determined that extraneous metabolites must be removed from the system, followed by the readdition of the pure substrates and cofactors of the enzyme system of interest. The synthesis of des-*myo*-inositol mycothiol was undertaken, as was the isolation of MSH from cell-free extracts of *S. jumonjinensis*, with the aim of identifying Glx type activity in *S. coelicolor*.

#### 4.2. Materials and Methods

All HPLC purifications were performed using either a Waters 600 gradient pump system with a 996 photodiode array detector or a Waters 625 LC system with a 994 programmable photodiode array. All HPLC purifications used Milli-Q water and HPLC grade methanol or acetonitrile. All aqueous solvents were filtered through a 0.2  $\mu\text{m}$  membrane filter (Pall Life Sciences, East Hills, NY) and all organic solvents were filtered through a 0.45  $\mu\text{m}$  membrane filter (Millipore Corp., Billerica, MA) and all solvents were degassed before use. Before injection, all samples were filtered through a syringe filter: 0.2  $\mu\text{m}$  PTFE membrane (Millipore) for organic solvents, 0.2  $\mu\text{m}$  polyethersulfone membrane for aqueous samples (VWR International, Mississauga, ON).

All nuclear magnetic resonance (NMR) spectra were obtained using Bruker 300 MHz spectrometers using deuterated solvents manufactured by Cambridge Isotopes Laboratories (Andover, MA). Mass spectra were recorded using a Micromass Q-TOF Ultima Global instrument (electrospray ionization, ESI) using 1:1 water:acetonitrile with 0.2 % formic acid unless otherwise noted, or using a JEOL HX110 mass spectrometer (electron impact ionization, EI).

*N*-(9-Fluorenylmethoxycarbonyl)-*S*-acetamidylmethyl-cysteine-pentafluorophenyl ester (FmocCys(Acm)-OPfP), and *N*-acetyl-*S*-triphenylmethyl-cysteine (AcCys(Trt)-OH) were purchased from Bachem Bioscience Inc. (King of Prussia, PA). *N,N'*-Diacetyl cysteine was obtained from Toronto Research Chemicals (North York, ON). *O*-Benzotriazole-*N,N,N',N'*-tetramethyl-uronium-hexafluoro-phosphate (HBTU) was obtained from Alfa Aesar (Ward Hill, MA) and ninhydrin was purchased from Pierce Chemicals (Rockford, IL).

The following were obtained from Sigma-Aldrich Inc.:  $\alpha$ -D-glucosamine hydrochloride (GlcN-HCl), 6-hydroxynaphthyl disulfide, 2-thiopyridyl disulfide (trade name: 2-Aldrithiol), D- and L-LDH, MG, molybdophosphoric acid, NAD<sup>+</sup>, nickel chloride hexahydrate, piperidine, and reduced GSH.

Ammonium chloride, cobalt chloride hexahydrate, cadmium chloride, dextrose, ethylenediaminetetraacetic acid (EDTA) disodium salt, hydrazine hydrochloride, manganous chloride tetrahydrate, sodium phosphate monobasic, and zinc chloride were purchased from J. T. Baker (Phillipsburgh, NJ). The Amberlyst A21 resin was supplied by Mallinckrodt Inc. (Phillipsburgh, NJ).

Dimethylformamide (DMF) and trifluoroacetic acid (TFA) are from Caledon Laboratories Ltd. (Georgetown, ON). DMF was dried over 4 Å molecular sieves overnight before use. Acetic acid, acetic anhydride, HPLC grade ethanol, hydrochloric acid, perchloric acid, potassium phosphate, and sulfuric acid were purchased from Fisher Scientific Co. (Ottawa, ON). Bacto™ peptone and malt extract were purchased from Becton, Dickinson and Company (Franklin Lakes, NJ). Calcium chloride, iodine, potassium carbonate and vanillin were obtained from BDH Laboratories. Glycine was supplied by Schwarz/Mann Biotech (Cleveland, OH).

The following chemicals were purchased from EMD Chemicals Inc. (Gibbstown, NJ): agar, anhydrous magnesium sulfate, glucose, HPLC grade acetonitrile and methanol, magnesium chloride hexahydrate, potassium phosphate monobasic, sodium bicarbonate, sodium chloride, sucrose, triethyl amine (TEA), tryptone, and yeast extract. Hydroxybenzotriazole (HOBt) was purchased from NovaBiochem, a brand of EMD Biosciences, Inc.

The presence of hydroxyl groups was tested using a vanillin stain: 2.6 g of vanillin in a solution of 90 mL of ethanol, 2 mL of glacial acetic acid and 3.4 mL of sulfuric acid. The presence of amines was tested using a ninhydrin stain: 0.2 % ninhydrin in ethanol. TLC plates were also visualized using phosphomolybdic acid: 0.4 % in ethanol.

Cell-free extracts were obtained from suspended cells by sonication with a Sonicator cell disruptor model W225, converter model # 2 and a standard tapered microtip, with the output control set at 5, from Heat Systems-Ultrasonics, Inc. (Plainsview, NY). Centrifugation of cell cultures over 20 mL were performed using a Beckman Avanti J-25I centrifuge (Mississauga, ON). Cell cultures were grown in either a Gyrotory Model G76 water bath shaker or a Series 25 controlled environment incubator shaker, both of New Brunswick Scientific (Edison, NJ).

Spectrophotometric assays were performed using a Varian (Mississauga, ON) Cary 3 spectrophotometer and version 3.04 of the control and data collection software package.

NMR peak assignments were aided by the use of the ACD/HNMR DB and ACD/CNMR DB software, both version 2.51 (Advanced Chemical Development Inc. Toronto, ON) and Silverstein and Webster (1998).<sup>116</sup>

#### 4.2.1. Data Mining of the *Streptomyces coelicolor* Genome

The genome of *S. coelicolor* has been sequenced and is available for sequence searching.<sup>†</sup> To investigate the presence of the known MSH-dependent enzymes in *S. coelicolor*, a series of sequence searches were completed using the amino acid sequences of MSH-dependent formaldehyde dehydrogenase, MSH-S-conjugate amidase, MSH-dependent maleyl pyruvate isomerase and MSSM reductase. Data mining using known GSH-dependent enzymes was also performed. All sequence searches were performed using the BLAST server available through the Sanger Centre.<sup>‡</sup> Multiple sequence alignments were performed using ClustalW available through the European Bioinformatics Institute.<sup>§117</sup>

#### 4.2.2. Media and Buffers Used

##### **Double Strength Germination Media**

1 % yeast extract

1 % casamino acids

0.01 M CaCl<sub>2</sub>, prepared and autoclaved separately as a 5 M solution; add 2 μL per mL of medium

##### **TES Buffer**

0.05 M *N*-(tris(hydroxymethyl)methyl)-2-aminoethanesulfonic acid (TES)

pH 8

##### **Yeast Extract-Malt Extract (YEME) Medium**

3 g yeast extract

5 g bacto-peptone

3 g oxoid malt extract

10 g glucose

340 g sucrose

Distilled water to a final volume of 1 L

After autoclaving the above solution add 2 mL of 2.5 M MgCl<sub>2</sub>•6 H<sub>2</sub>O to yield a final concentration of 5 mM.

##### **M9 Media**

12.8 g sodium phosphate dibasic heptahydrate

3.0 g potassium phosphate dibasic

0.5 g sodium chloride

1.0 g ammonium chloride

Distilled water up to 1 L

---

<sup>‡</sup> [http://www.sanger.ac.uk/cgi-bin/blast/submitblast/s\\_coelicolor](http://www.sanger.ac.uk/cgi-bin/blast/submitblast/s_coelicolor)

<sup>§</sup> <http://www.ebi.ac.uk/clustalw/#>

**Phosphate Buffered Saline (PBS)**

100 mM potassium phosphate

0.8 % sodium chloride

pH 7.2

**International Streptomyces Project (ISP) Medium 1**

5 g tryptone

3 g yeast extract

Distilled water to 1 L

pH 7.0 to 7.2

**ISP Medium 2**

4 g yeast extract

10 g malt extract

4 g dextrose

20 g agar

Distilled water up to 1 L

pH 7.3 – 7.5

**Lysis Buffer 1**

30 mM potassium phosphate

0.1 mM EDTA disodium salt

10 mM MgCl<sub>2</sub>

pH 7.2

**Lysis Buffer 2**

0.25 M perchloric acid

2 mM EDTA

40 % acetonitrile in water

**Glx Assay Buffer 1**

27 mM MG

50 mM potassium phosphate

pH 6.6

**Glx Assay Buffer 2**

27 mM MG

50 mM potassium phosphate

1.8 mM GSH

pH 6.6

**LDH Assay Solution**

0.17 M glycine

0.17 M hydrazine hydrochloride

pH 9.5



## LDH

The LDH enzymes were used as supplied in a pH 6, ammonium sulfate (L-: 2.1 M, D-: 3.2 M) stock solution

### Activities

D-LDH: 1228 units/mL\*\*

L-LDH: 6000 units/mL

## NAD<sup>+</sup> Stock Solution

60 mg of  $\beta$ -NAD<sup>+</sup> in 2 mL of milli-Q water

### 4.2.3. Growth of *Streptomyces coelicolor*

Spores of *S. coelicolor* A3(2) were a generous gift from Professor Sir David Hopwood of the Department of Molecular Microbiology at the John Innes Centre in Colney, Norwich, UK. The spores were germinated as detailed by Kieser *et al.*<sup>103</sup> The spores were pelleted on a bench-top centrifuge (3000 rpm for 2 min), re-suspended in 5 mL of TES buffer, transferred to a 125 mL erlenmeyer flask and heat shocked for 10 minutes at 50 °C and cooled under cold water. To the flask was added 5 mL of autoclaved double strength Germination Medium and the resulting mixture was shaken at 37 °C at 300 rpm for 3 hr. The suspension was decanted into a sterile Falcon tube and centrifuged at 5 800 rpm for 10 minutes to pellet the cells. The pellet was resuspended in autoclaved Milli-Q water and this was added to 200 mL of YEME media and shaken at 30 °C for approximately 49 hours. The culture was centrifuged at 4 °C for 15 min at 8 000 rpm, the supernatant was decanted and the pellets were washed with 100 mL of M9 or PBS media and centrifuged for 15 min at 10 000 rpm. The resultant loose pellet was frozen in liquid nitrogen and the cells were stored at -80 °C.

### 4.2.4. Initial Assay for Glyoxalase Activity in *Streptomyces coelicolor*

Frozen *S. coelicolor* cells were thawed on ice with the addition of a small amount of Lysis Buffer 1 to cover the cells. The cells were diluted up to 2 mL for 0.6 g of cells and 3 mL for 1.2 g of cells (wet weight). The cells were sonicated thirty times for 10 sec with a 30 sec of wait time between each interval. The cells were kept on ice at all times. The cells were centrifuged in a small chemical tabletop centrifuge (speed setting 5) yielding a yellow supernatant and a fairly small pellet. Half of the supernatant was then put through a 3 000 molecular weight cut-off Centricon filter and the other half was used immediately. Aliquots of both the unfiltered and filtered cell-free extract were incubated with MG (final MG concentration = 0.3  $\mu$ M) for 30 min. A cocktail of metal ions, copper chloride, cadmium chloride, manganous chloride and nickel chloride, each at a concentration of 500 mM, was prepared and used as noted below.

To assay for GlxI activity the following combinations were placed in a 1 mL quartz cuvette and the change in absorbance at 240 and 340 nm was observed. These assays were performed using unfiltered cell-free extract and unfiltered cell-free extract incubated with MG:

---

\*\* 1 unit = the amount of enzyme that would reduce 1  $\mu$ mole of pyruvate to lactate per minute at pH 7.0 and 25 °C.

1. 950  $\mu\text{L}$  of Glx Assay Buffer 1, 50  $\mu\text{L}$  of cell-free extract
2. 950  $\mu\text{L}$  of Glx Assay Buffer 1, 50  $\mu\text{L}$  of cell-free extract, 1  $\mu\text{L}$  of metal cocktail
3. 950  $\mu\text{L}$  of Glx Assay Buffer 2, 50  $\mu\text{L}$  of cell-free extract
4. 950  $\mu\text{L}$  of Glx Assay Buffer 2, 50  $\mu\text{L}$  of cell-free extract, 1  $\mu\text{L}$  of metal cocktail.

To assay for GlxII and overall Glx system activity the following solutions were mixed in a 1 mL quartz cuvette and the change in absorbance at 340 nm observed. These assays were performed using unfiltered cell-free extract and unfiltered cell-free extract incubated with MG:

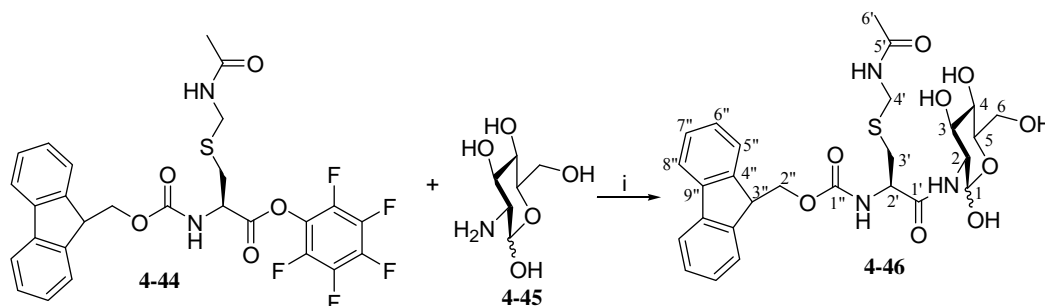
5. 895  $\mu\text{L}$  of LDH assay buffer, 50  $\mu\text{L}$  cell-free extract, 56  $\mu\text{L}$  of  $\text{NAD}^+$  stock solution and 2  $\mu\text{L}$  of either L- or D-LDH stock solution
6. 895  $\mu\text{L}$  of LDH assay buffer, 50  $\mu\text{L}$  cell-free extract, 56  $\mu\text{L}$  of  $\text{NAD}^+$  stock solution, 2  $\mu\text{L}$  of either L or D-LDH stock solution and 1  $\mu\text{L}$  of metal cocktail

As a control, Assays 1 – 6 were repeated using filtered cell-free extract to monitor background activity. In addition, Assays 5 and 6 were repeated without the addition of any LDH to monitor background activity.

#### 4.2.5. Control Assays for Glyoxalase Activity in *Escherichia coli*

To determine if it is possible to monitor Glx activity using the assay described above, Assays 1 – 6 were repeated using cell-free extracts of *E. coli* MG1655, with the exception that a 500  $\mu\text{L}$  nickel chloride solution was used in lieu of the metal cocktail. Approximately 1.6 g of cells were thawed and suspended in 4 mL of Lysis Buffer 1. The cells were sonicated ten times for 10 sec with 30 sec between each interval and the mixture centrifuged. Approximately half of the supernatant was put through a 3 000 molecular weight cut off Centricon filter.

#### 4.2.6. Synthesis of *Des-myo-Inositol Mycothiol*



**Figure 4.11:** Synthesis of *N*- $\alpha$ -Fmoc-*S*-acetamidomethyl-L-cysteinyl-2-amino-2-deoxy- $\alpha$ -D-glucopyranoside (Fmoc-Cys(Acm)GlcN). (i) HOBt, DMF.

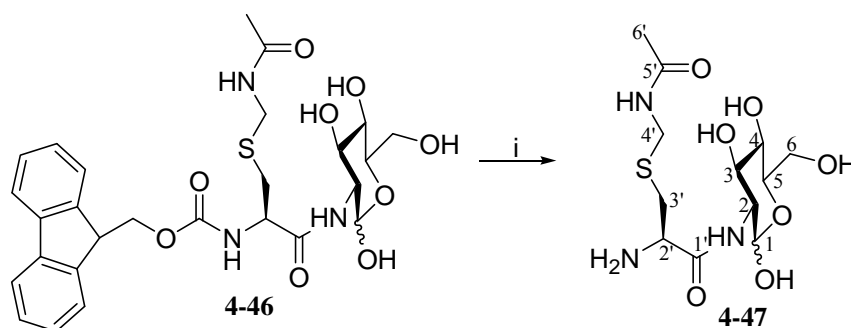
This procedure is an adaptation of that published by Patel and Blanchard.<sup>113</sup> GlcN-HCl (3.72 g, 17 mmol) was dissolved in a 1 M solution of NaOH (17 mL, 17 mmol) and allowed to stir for 5 min at room temperature. This solution was then frozen in liquid nitrogen and dried on the lyophilizer overnight. Fmoc-Cys(Acm)-OPfP (4-44, 2 g, 3.4 mmol) was dissolved in 10 mL of dry DMF and HOBt was added (0.930 g, 6.1 mmol). This was immediately transferred to the dried GlcN solid. GlcN did not dissolve

immediately and vigorous stirring was required to break up the GlcN solid. This mixture was allowed to stir for 3 hr at room temperature under argon. The solvent was then removed by a rotary evaporator attached to a vacuum pump. The product was precipitated with 100 mL of water and isolated by filtration followed by thorough rinsing with water to remove any residual DMF, GlcN, pentafluorophenol, NaCl, and HOBt. The product was either collected from the filter manually or dissolved in MeOH and dried by rotary evaporation. In both cases, the sample was also dried *in vacuo* overnight to remove any residual water. Initial NMR data indicated the presence of unreacted Fmoc-Cys(Acm)-OH; therefore, the sample was dissolved in methanol and stirred with Amberlyst A21, a weakly basic macroreticular resin, for 30 minutes. The resin was removed by filtration and the sample dried by rotary evaporation to give a 56 % yield of **4-46**.

$^1\text{H}$  NMR (300 MHz, MeOH- $d_4$ )  $\delta$  7.79 (d, 2H, H8''), 7.75 (d, 2H, H5''), 7.39 – 7.26 (m, 4H, H6'', H7''), 5.10 (d, 1H, H1), 4.47 – 4.22 (m, 6H, H2', H4', H2'', H3''), 3.79 – 3.61 (m, 6H, H2, H3, H4, H5, H6), 3.1 – 2.8 (m, 2H, H3'), 2.0 (s, 3 H, H6').

$^{13}\text{C}$  NMR (75 MHz, DMSO- $d_6$ )  $\delta$  172.0/171.5 (C1'/C5'), 156.5 (C1''), 144.1 (C9''), 141.0 (C4''), 128.1 (C7''), 127.5 (C6''), 125.7 (C5''), 120.5 (C8''), 91.0 (C1), 73.1/72.0/71.5 (C3/C4/C5), 66.2 (C2''), 61.0 (C6), 55.0 (C2), 54.5 (C2'), 46.9 (C3''), 32.5 (C3'), 23.5 (C6'), C4' is expected to appear at  $\sim \delta$  39, which would be hidden by the DMSO- $d_6$  solvent peak.

ESI-MS (pos):  $m/z = 576$  (MH) $^+$ , 598 (MNa) $^+$ , 614 (MK) $^+$ , calculated (MH) $^+ = 576$ .



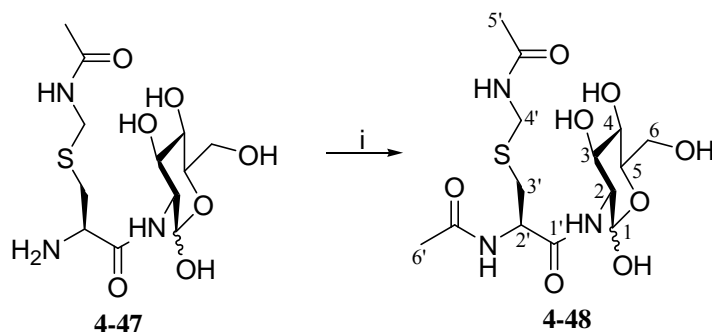
**Figure 4.12:** Synthesis of *S*-acetamidomethyl-L-cysteinyl-2-amino-2-deoxy- $\alpha,\beta$ -D-glucopyranoside (NH $_2$ -Cys(Acm)GlcN). (i) 5 % piperidine in DMF.

The Fmoc-Cys(Acm)-GlcN (**4-46**, 0.55 g, 0.9 mmol) was dissolved in 5 mL of a solution of 5 % piperidine in DMF. This was allowed to stir for 30 min and the solvent was removed with a rotary evaporator attached to a vacuum pump. The cleaved dibenzofulvene was precipitated with 30 mL of water. The precipitate was removed by filtration and the filtrate was dried *in vacuo* to give a 99 % yield of **4-47**.

$^1\text{H}$  NMR (300 MHz, D $_2$ O)  $\delta$  5.0 (d, 1H, H1), 4.2 (m, 2H, H4'), 3.8 – 3.3 (m, 7H, H2, H3, H4, H5, H6, H2'), 3.0 – 2.8 (m, 2H, H3'), 2.0 (s, 3H, H6').

$^{13}\text{C}$  NMR (75MHz, D $_2$ O)  $\delta$  174.2/172.3 (C1'/C5'), 90.9 (C1), 76.0/75.3/73.7 (C3/C4/C5), 60.0 (C2), 63.6 (C6), 54.2 (C2'), 35.1 (C4'), 32.0 (C3'), 22.1 (C6').

ESI-MS (pos):  $m/z = 354$  (MH) $^+$ , calculated (MH) $^+ = 354$ .



**Figure 4.13:** Synthesis of *N*-acetyl-*S*-acetamidomethyl-*L*-cysteinyl-2-amino-2-deoxy- $\alpha,\beta$ -D-glucopyranoside (AcCys(Acm)GlcN). (i) Ac<sub>2</sub>O, K<sub>2</sub>CO<sub>3</sub>.

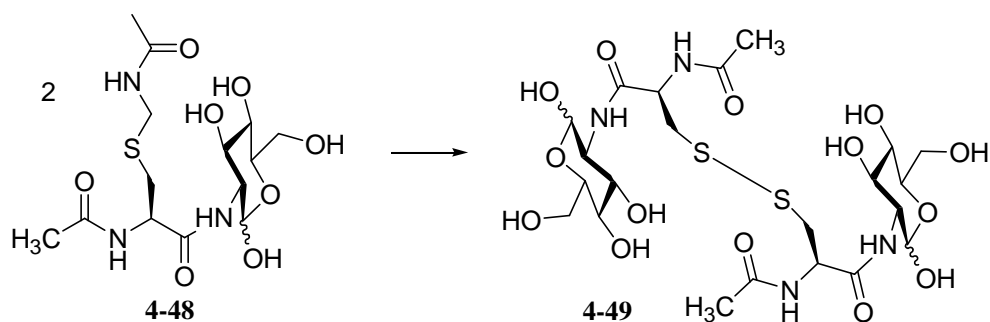
NH<sub>2</sub>-Cys(Acm)-GlcN (**4-47**, 0.17 g, 0.47 mmol) was dissolved in 1.2 mL of water, potassium carbonate (0.13 g, 0.94 mmol, 2 equiv.) was added and the mixture was stirred until the potassium carbonate was dissolved. Acetic anhydride (56  $\mu$ L, 0.06 g, 1.3 equiv.) was added dropwise over 30 min and the solution was allowed to stir at room temperature for 45 minutes and then dried *in vacuo*. The product was dissolved in 5 mL of ethanol to precipitate the potassium carbonate, which was removed by filtration, and the filtrate was dried to an oil. The sample was purified by preparative HPLC using a Waters  $\mu$ Bondapak C<sub>18</sub> reverse phase radial compression column (25  $\times$  100 mm) and the gradient shown in Table 4.4. The product eluted at  $\sim$  10 min with a yield of 25 % of **4-48**.

**Table 4.4:** Solvent gradient used to purify *N*-acetyl-*L*-cysteinyl-2-amino-2-deoxy- $\alpha,\beta$ -D-glucopyranoside.

<i>Time</i>	<i>Flow Rate</i>	<i>% Water</i>	<i>% Methanol</i>
0	5 mL/min	100	0
14	5 mL/min	100	0
17	5 mL/min	0	100
30	5 mL/min	0	100
35	5 mL/min	100	0

<sup>1</sup>H NMR (300 MHz, D<sub>2</sub>O)  $\delta$  5.0 (d, 1H, H1), 4.5 (m, 1H, H2'), 4.3 (m, 2H, H4'), 3.9 – 3.2 (m, 6 H, H2, H3, H4, H5, H6), 3.0 (d, 1H, H3'), 2.8 (d, 1H, H3'), 1.8 (s, 6H, H5', H6').

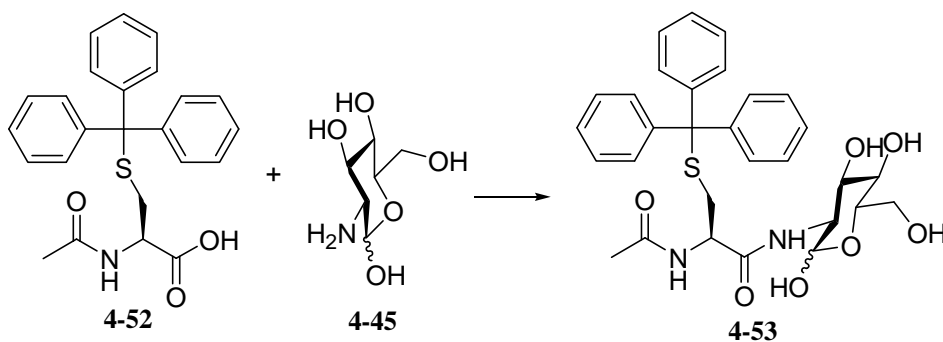
ESI-MS (pos):  $m/z = 434$  (MK)<sup>+</sup>, calculated (MK)<sup>+</sup> = 434.



**Figure 4.14:** Proposed synthesis of *N*-acetyl-L-cysteinyl-2-amino-2-deoxy- $\alpha,\beta$ -D-glucopyranoside disulfide.

This procedure was adapted from that of Kamber *et al.*<sup>118</sup> for the removal of the acetamidomethyl protecting group from thiols using iodine, with the simultaneous formation of a disulfide. AcCys(Acm)-GlcN (**4-48**, 20 mg, 0.051 mmol) was dissolved in 7 mL of DMF and 3 mL of a 0.1 N solution of iodine (6 equiv.) was added. This mixture was allowed to stir for 1 min at room temperature and was then diluted with 10 mL of chloroform and 10 mL of water. The water layer was washed repeatedly with chloroform until the chloroform layer was colourless. The water layer was dried *in vacuo* and the residue was purified by HPLC using the conditions described above. The predicted product (**4-49**) was not visible by <sup>1</sup>H NMR nor ESI-MS.

#### 4.2.7. Alternative Synthesis of *Des-myo-Inositol*

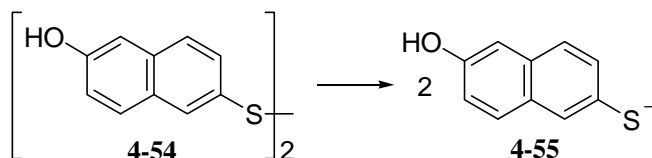


**Figure 4.15:** Synthesis of *N*-acetyl-S-trityl-L-cysteinyl-2-amino-2-deoxy- $\alpha,\beta$ -D-glucopyranoside (AcCys(Trt)GlcN).

*N*-acetyl-S-trityl-cysteine (**4-52**, 0.489 g, 1.21mmol) was dissolved in 10 mL of dry DMF followed by the addition of GlcN-HCl (0.290 g, 1.35 mmol, 1.12 equiv.), TEA (350  $\mu$ L, 2.5 mmol, 2.1 equiv.), and HBTU (0.504 g, 1.33 mmol, 1.10 equiv.). The reaction was allowed to stir overnight under argon. The mixture was diluted with 35 mL of saturated aqueous sodium chloride resulting in a large amount of precipitate. The aqueous layer was neutralized (to pH  $\sim$ 7) using HCl, which eliminated most of the cloudiness. The aqueous mixture was washed three times with 25 mL of dichloromethane and the organic layers were pooled, washed with 2 N HCl, water, 5 % sodium bicarbonate, and water, and then dried over anhydrous magnesium sulfate, which was later removed by filtration. The product was dried on the rotary evaporator and then dried overnight with a vacuum pump to give an oil. This oil was taken up in a minimum amount of 85:15 chloroform:methanol and purified by silica gel chromatography using the method described by Still *et al.*<sup>119</sup> (85:15

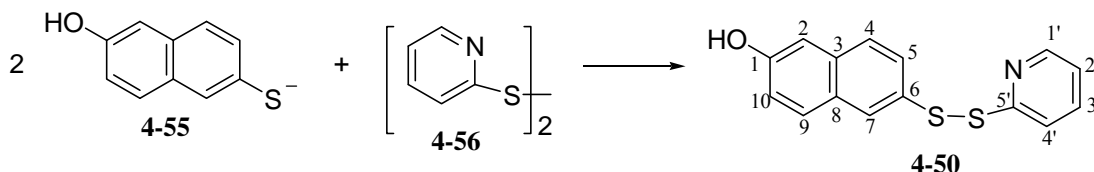
chloroform:methanol) to give a yield of 53 % of **4-53**. The sample was contaminated with TEA and it was noted that racemization at the  $\alpha$ -carbon had likely occurred, as the  $^1\text{H}$  NMR spectrum revealed multiple peaks for the acetyl group and the anomeric carbon, indicative of the formation of diastereomers. The multiplets expected for the other protons in the molecule, GlcN etc, yielded a convoluted spectrum. This synthesis was halted.

#### 4.2.8. Synthesis of 2-S-(2'-thiopyridyl)-6-hydroxynaphthyl disulfide



**Figure 4.16:** Reduction of 6-hydroxynaphthyl disulfide with sodium borohydride.

This procedure was modified from that of Steenkamp and Vogt.<sup>115</sup> 6-Hydroxynaphthyl disulfide (**4-54**, 0.198 g, 0.56 mmol) was added to 100 mL of a 1:1 acetonitrile: 0.1 M  $\text{NaHCO}_3$  solution with a pH of 9, which had been degassed for 10 min. Sodium borohydride (0.6 g, 0.02 mol, 28 equiv.) was added, the system was flushed three times with argon and the mixture was allowed to stir at room temperature for 2 hr under argon. As the disulfide was reduced, the solution became a clear yellow. The reaction was quenched with glacial acetic acid (2 mL) and the pH was adjusted to approximately 4.55. This solution was then degassed for 5 min before use in the next step.



**Figure 4.17:** Synthesis of 2-S-(2'-thiopyridyl)-6-hydroxynaphthyl disulfide.

2-Thiopyridyl disulfide (**4-56**, 0.474 g, 2.06 mmol, 1.8 equiv) was dissolved in a 1:1 acetonitrile:water solution that had been degassed for 10 min. The 2-thio-6-hydroxynaphthol solution was added dropwise to 2-thiopyridyl disulfide over 20 minutes and allowed to stir at room temperature for 1.5 hours. The mixture was diluted with 150 mL of 4 °C water and the precipitate isolated by filtration. The very fine precipitate was dissolved from the filter paper using methanol and the solvent was removed by rotary evaporation. It was essential to avoid heating the product to prevent isomerization of the disulfide. The mixed disulfide was dissolved in a minimal amount of 95:5 chloroform:methanol, filtered and the filtrate was purified by silica gel chromatography using the method described by Still *et al.*<sup>119</sup> (95:5 chloroform: methanol, mixed disulfide  $R_f$ : 0.32, 6-hydroxynaphthyl disulfide  $R_f$ : 0.18, 2-thiopyridyl disulfide  $R_f$ : 0.62). The collected precipitate was redissolved in 95:5 chloroform:methanol and purified by silica gel chromatography to give an overall yield of 25 % of **4-50**. The product was most stable when stored dry.

$^1\text{H}$  NMR (300 MHz,  $\text{MeOH-d}_4$ )  $\delta$  8.4 (d, 1H, H1'), 7.9 (d, 1H, H7), 7.8 (d, 1H, H4') 7.7 (dd, 1H, H2') 7.6 (t, 2H, H4, H9), 7.5 (dd, 1H, H2), 7.2 (m, 1H, H3'), 7.0 (m, 2H, H5, H10)

$^{13}\text{C}$  NMR (75 MHz, MeOH- $d_4$ )  $\delta$  157.5 (C5'), 150.4 (C1), 139.3 (C3', C1'), 136.0 (C3), 130.5 (C9), 130.2 (C6), 129.7 (C8), 128.9 (C10), 128.7 (C7), 127.6 (C5), 122.6 (C4), 121.5 (C2'), 120.4 (C4'), 110.0 (C2).

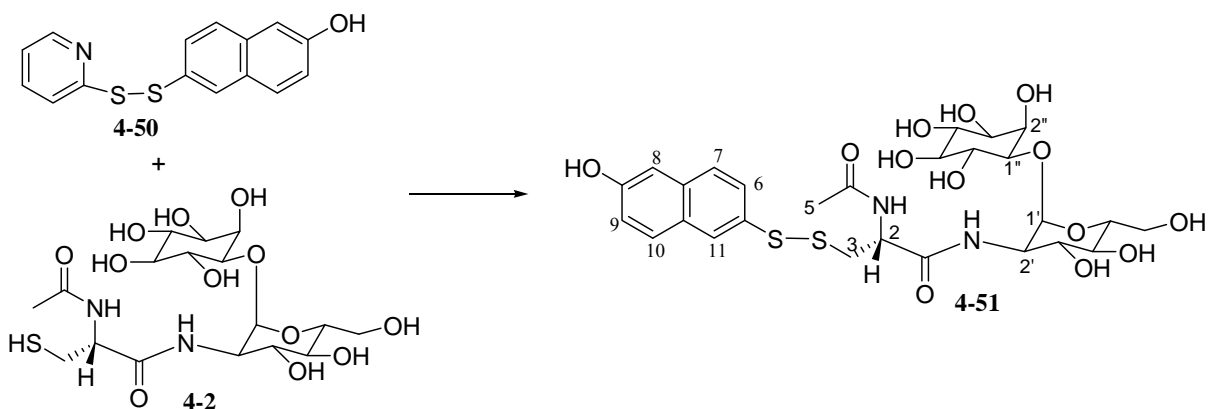
EI-MS (pos)  $m/z = 285$ , calculated  $(M)^+ = 285$ .

#### 4.2.9. Growth of *Streptomyces jumonjinensis*, the Chosen Source of Mycothiol

*S. jumonjinensis* was obtained from the American Tissue Culture Collection (ATCC, number 29864) as a freeze dried pellet. The pellet was taken up in 1 mL of ISP medium 1. A few drops of this suspension were added to slants of ISP medium 2, which were then incubated at 27 °C and showed growth after three days. The remainder of this suspension was divided between four flasks: two flasks with 50 mL of YEME media and two flasks containing 50 mL of ISP medium 1. After three days of shaking at 26 °C the ISP medium cultures were very cloudy while the YEME cultures did not show a large amount of growth. The ISP medium 1 culture was centrifuged at 8 000 rpm for 10 min at 4 °C, and the pellet collected, frozen and stored at -80 °C.

A second freeze dried pellet was rehydrated in the same fashion and after shaking for three days in ISP medium 1, the culture was evenly divided among six 4 L flasks each with 500 mL of ISP medium 1. These flasks were shaken at 26 °C for five days and the culture was centrifuged at 12 000 rpm for 20 minutes, the pellet collected, frozen and stored at -80 °C.

#### 4.2.10. Isolation of Mycothiol from *Streptomyces jumonjinensis*



**Figure 4.18:** Formation of 2-S-6-hydroxynaphthyl-mycothiolsulfide.

This procedure was adapted from that of Steenkamp and Vogt which used *M. smegmatis* as the source of MSH.<sup>115</sup> Frozen *S. jumonjinensis* cells (4.029 g of wet weight) were thawed and suspended in approximately 30 mL of Lysis Buffer 2 over three 15 mL centrifuge tubes. Each tube was sonicated with thirty 10 sec pulse sets with a 10 sec rest between each set. These tubes were then centrifuged and the supernatant was distributed into two 15 mL centrifuge tubes as a thick yellow liquid. The pH of the solutions was adjusted to approximately 4.7 using potassium carbonate and dilute hydrochloric acid. The cell-free extract was then degassed for 10 min and added dropwise over 20 min to 3.65 mL of a degassed 4.4 mM (saturated) solution of 2-S-(2'-thiopyridyl)-6-hydroxynaphthyl disulfide in 60 % (v/v) acetonitrile:water (0.016 mmoles). The solution was allowed to stir under argon

at room temperature for 4.5 hours. The reaction mixture was diluted with 56 mL of water and then filtered through a Waters classic Sep-pak, which had been conditioned with methanol and water. The Sep-pak was rinsed with water followed by 1:1 acetonitrile:water, which gave a yellow product in approximately 1 mL. This fraction was concentrated using rotary evaporation to remove the acetonitrile. The sample was then purified by analytical HPLC using a Waters  $\mu$ Bondapak C<sub>18</sub> reverse-phase radial compression column (8 × 100 mm) and the gradient shown in Table 4.5. The elution was monitored at 220 and 280 nm and tagged MSH eluted at 28 min. The 2-thiopyridyl disulfide eluted at 36 min, 2-*S*-(2'-thiopyridyl)-6-hydroxynaphthyl disulfide at 41 min and 6-hydroxynaphthyl disulfide at 44 min. The use of ~ 4 g of wet weight cells yielded approximately 1.5 to 2 mg of tagged MSH.

**Table 4.5:** Solvent gradient used to purify 2-*S*-6-hydroxynaphthyl-mycothioliol disulfide.

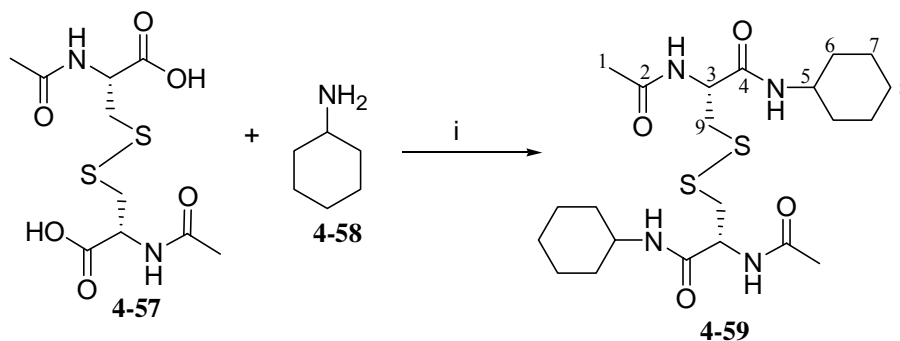
Time	Flow Rate	% of 0.1 % TFA <sup>a</sup> in Water	% of Acetonitrile
0	1 mL/min	100	0
10	1 mL/min	100	0
50	1 mL/min	0	100

<sup>a</sup>TFA = trifluoroacetic acid

<sup>1</sup>H-NMR (500 MHz, D<sub>2</sub>O)  $\delta$  8.0 (s, 1H, H11), 7.8 (dd, 2H, H7, H10), 7.6 (d, 1H, H8), 7.2 (m, 2H, H6, H9), 5.0 (s, 1H, H1'), 4.5 (m, 1H, H2), 4.1 (s, 1H, H2''), 3.9 – 3.7 (m, 7H, H2', H3', H5', H6' (2H), H4'', H6''), 3.6 (m, 4H, H4', H1'', H3'', H5''), 3.2 (t, 1H, H3a), 2.9 (t, 1H, H3b), 2.1 (s, 3H, H5).<sup>††</sup>

ESI-MS (pos):  $m/z$  (MH)<sup>+</sup> = 661, (MLi)<sup>+</sup> = 667, (MNa)<sup>+</sup> = 683, (M + 2 Na<sup>+</sup> + CF<sub>3</sub>COO<sup>-</sup>)<sup>+</sup> = 819, (M + 4 Na<sup>+</sup> + 3 CF<sub>3</sub>COO<sup>-</sup>)<sup>+</sup> = 1091, (M + 5 Na<sup>+</sup> + 4 CF<sub>3</sub>COO<sup>-</sup>)<sup>+</sup> = 1227, (M + 6 Na<sup>+</sup> + 5 CF<sub>3</sub>COO<sup>-</sup>)<sup>+</sup> = 1363, (M + 7 Na<sup>+</sup> + 6 CF<sub>3</sub>COO<sup>-</sup>)<sup>+</sup> = 1499, calculated (MH)<sup>+</sup> = 661.

#### 4.2.11. Synthesis of a Mycothiol Analogue



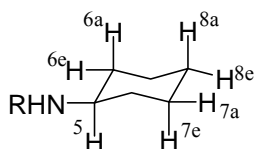
**Figure 4.19:** Synthesis of a proposed mycothiol analogue. Reagents: (i) HBTU, TEA, DMF.

*N*-*N'*-Diacetyl-cystine (101 mg, 0.31 mmol) was dissolved in 0.6 mL of dry DMF. HBTU (260 mg, 0.69 mmol, 2.1 equiv.) was dissolved in 1.2 mL of dry DMF and added to the amino acid followed by cyclohexyl amine (74  $\mu$ L, 0.65 mmol, 1.0 equiv.) and TEA

<sup>††</sup> The pattern seen for the protons belonging to MSH is in full agreement with those found by Spies and Steenkamp (1994), Sakuda *et al.* (1994) and Newton *et al.* (1995) for the bimane derivative of MSH.



(180  $\mu$ L, 1.29 mmol, 2.1 equiv.). The mixture was allowed to stir at room temperature for 1 hr. The reaction mixture was diluted with 4 mL of water, which was washed three times with 3 mL of ethyl acetate. The organic layer was then washed with 2 M HCl, water, 5 % sodium bicarbonate, and water and dried over magnesium sulfate. The mixture was filtered and dried *in vacuo*. The product was taken up in 1:1 water:methanol, loaded onto a preconditioned Sep-pak and eluted with 100 % methanol. The product was then dried by rotary evaporation to give a 13 % overall yield of the *N-N'*-diacetyl-cystinyl-cyclohexylamide (**4-59**) product.



$^1\text{H-NMR}$  (300 MHz MeOH- $d_4$ )  $\delta$  4.67 (m, 1H, H3), 3.62 (m, 1H, H5), 3.2 – 3.0 (m, 1H, H9a), 3.0 – 2.8 (m, 1H, H9b), 2.0 (s, 3H, H1), 1.8 – 1.5 (m, 5H, H6e, H6a, H7a), 1.3 – 1.0 (m, 5H, H7e, H8a, H8e).

$^{13}\text{C-NMR}$  (75 MHz, MeOH- $d_4$ )  $\delta$  171.7/171.5 (C2/C4), 52.4 (C3), 51.0 (C5), 32.0/31.9 (C6/C9), 25.1 (C7), 24.6 (C8), 21.0 (C1)

ESI-MS (pos):  $m/z$  (MH) $^+$  = 487, calculated (MH) $^+$  = 487

### 4.3. Results and Discussion

MSH is the major thiol of the *Actinomycetales* bacteria and is essential for the viability of *M. tuberculosis*, the causative agent of tuberculosis. The MSH-dependent enzymes identified to date parallel GSH biochemistry, implying that the functions of GSH may serve as a template for the discovery of new MSH functionalities. GSH is an essential cofactor for the Glx system, which converts cytotoxic MG to D-lactate. Our efforts towards the identification of an MSH-dependent Glx system in *S. coelicolor* have created a solid foundation for the possible future identification of this system in the *Actinomycetales* bacteria.

#### 4.3.1. Sequence Searching of the *Streptomyces coelicolor* Genome

Sequence searches of the genome of *S. coelicolor* for homologous sequences to known MSH-dependent enzymes were performed using the sequences listed in Table 4.6. The sequences returned for MSH-dependent formaldehyde dehydrogenase and MSSM reductase are annotated as likely having these functions, indicating that these enzymes are likely present in *S. coelicolor* (Table 4.6 and Alignments 1 and 3). The searches performed using the MSH-*S*-conjugate amidase and MSH-dependent maleyl pyruvate isomerase sequences returned sequences with identities of 49 % and 30 %, respectively, indicating that these enzymes are likely present in *S. coelicolor* as well (Table 4.6 and Alignments 2 and 4).

**Table 4.6:** Results for the sequence searches of the *Streptomyces coelicolor* genome for homologous sequences to known mycothiol dependent enzymes.

<i>Query Organism</i>	<i>Accession #</i>	<i>S. coelicolor Sequence</i>	<i>Putative Role</i>	<i>% Identity</i>	<i>% Similar</i>	<i>E value</i>
<b>Formaldehyde dehydrogenase</b>						
<i>A. methanolica</i>	P80094	NP_625045	Zn-dependent alcohol dehydrogenase, class III <sup>a</sup>	74 %	83 %	$7.7 \times 10^{-146}$
<b>Mycothiol-S-conjugate amidase</b>						
<i>M. tuberculosis</i>	NP_215598	NP_629119	None <sup>b</sup>	49 %	62 %	$5.7 \times 10^{-70}$
<b>Mycothiol disulfide reductase</b>						
<i>M. tuberculosis</i>	AF002193	NP_627648	pyridine nucleotide-disulphide oxidoreductase	28 %	43 %	$6.6 \times 10^{-37}$
<b>Mycothiol dependent maleyl pyruvate isomerase</b>						
<i>C. glutamicum</i>	NP_602215	NP_626223	None <sup>b</sup>	30 %	43 %	$9.0 \times 10^{-17}$

<sup>a</sup>The same class to which the formaldehyde dehydrogenases belong; <sup>b</sup>This peptide sequence, identified from the genome, has not yet been assigned a putative function.

A similar search of the *S. coelicolor* genome was performed using a sampling of known GlxI and GlxII protein sequences (Table 4.7, Alignments 4 – 13). The sequences returned by the search using GlxI returned three sequences including a putative lyase with similarity to *S*-D-lactoylglutathione lysases (GlxII) (NP\_625147), as well as a hypothetical protein (NP\_625386) and a putative dioxygenase (NP\_626233), both of which may belong to the glyoxalase superfamily. A comparison of the *E. coli*, *L. major* and *S. coelicolor* sequences (Alignment 9) demonstrated that three of the four ligands responsible for the binding of the catalytic metal in *E. coli* and *L. major* were not present in the *S. coelicolor* sequence (NP\_625386). This does not preclude the presence of a GlxI type enzyme in *S. coelicolor*, however, it does suggest that the enzyme may be structurally dissimilar to other known GlxI enzymes and may not be a metalloenzyme.

The searches of the *S. coelicolor* genome using various GlxII enzyme sequences returned a hypothetical protein (NP\_626224) and a hydrolase (NP\_625789), both with similarity to the glyoxalases. While the identity of the returned sequences was not high, this does not prohibit the presence of an MSH-dependent Glx system as the TSH<sub>2</sub>-dependent GlxII enzyme in *L. donovani* demonstrated only 35 % sequence identity with the human GlxII enzyme.<sup>25</sup> The similarity between the *L. major* GlxI sequence and other eukaryotic GlxI sequences was only between 32.4 and 39.4 %.<sup>23</sup> The structure of MSH, a pseudo-disaccharide coupled to *N*-acetyl-cysteine, is very different from that of GSH, a linear tripeptide, and TSH<sub>2</sub>, which consists of two GSH molecules linked by spermidine; hence, it

is possible that sequences of the MSH-dependent Glx enzymes would differ dramatically from those of GSH and TSH<sub>2</sub>-dependent enzymes. This idea is supported by the lack of homology between the MSH-dependent maleyl pyruvate isomerase sequence from *C. glutamicum* and the sequences from GSH-dependent maleyl pyruvate isomerase from *K. pneumoniae* or *Ralstonia* U2.<sup>32</sup>

### Alignment 1:

Alignment of the peptide sequence of MSH-dependent formaldehyde dehydrogenase from *Amycolatopsis methanolica* (Assession No. P80094) with the best matching sequence from *S. coelicolor*, a putative oxidoreductase predicted to be a zinc-dependent alcohol dehydrogenase (Accession No. NP\_625045): 74 % identity and 83 % similarity (E value =  $7.7 \times 10^{-146}$ ).

Query = P80094, Sbjct = NP\_625045

```
Query:      2 QTVRGVIARSKGAPVELTDIVIPDPGPGSEVTALITCAVCHTDLTYREGGINDEFPPFLLG 61
             Q VRGVIA K PV + IV+PDPGP E I C VCHTDL Y++GGINDEFPPFLLG
Sbjct:      3 QEVRGVIAPGKDEPVRIETIVVPDPGPGGEAVVKIQACGVCHTDLHYKQGGINDEFPPFLLG 62

Query:      62 HEAAGTVESVGEVDSVQPGDYVVLNWRVAVCGQCRACKRGRPQYCFSTFNATQKMTLTD- 120
             HEAAG VESVG GV V PGD+V+LNWRVAVCG+CRAC+RGRP YCF T NA Q+MTL
Sbjct:      63 HEAAGVVESVGAGVTDVAPGDFVILNWRVAVCGKCRACRRGRPWFYCFDTHNAEQRMTLAST 122

Query:     121 GTELTALGIGAFADKTLVHAGQCTKVDPAADPAVAGLLGCGVMAGLGAAVNTGAVSRGD 180
             G EL+PALGIGAF+A+KTLV AGQCTKVDPA VAGLLGCGVMAG+GAA+NTG V RGD
Sbjct:     123 GQELSPALGIGAFAEKTLVAAGQCTKVDPAVPAEVAGLLGCGVMAGIGAAINTGNVGRGD 182

Query:     181 SVAVIGCGAVGDAVIAGARLAGANKIIAIVDRDAKLEWATELGATHTVFNATETDVVEAVQ 240
             +VAVIGCG VGDA IAG+ LAGA+K+IAVD D +KLE A +GATHTVN+ TD VEA++
Sbjct:     183 TVAVIGCGGVGDAAIAGSYLAGASKVIAVDIDDRKLETARTMGATHTVNSRGTDPVEAIR 242

Query:     241 ALTGGFGADVVIDAVGRPETWKQAFYARDLAGTVVVLVGVPTPDMRLEMPLLDFFSRGGAL 300
             LTGGFGADVVI+AVGRPET++QAFYARDLAGTVVVLVGVPTP+M+LE+PLLD F RGG+L
Sbjct:     243 ELTGGFGADVIEAVGRPETYRQAFYARDLAGTVVVLVGVPTPEMKLELPLLDVFGRRGSL 302

Query:     301 KSSWYGDCLPERDFPVLIDLHLQGRPLDKFVTERISLDDVEKAFHTMHAGEVLRVSVV 359
             KSSWYGDCLP RDFP+L+DLHLQGR L L KFVTE I LD+VE+AF MHAG+VLRVSVV
Sbjct:     303 KSSWYGDCLPSRDFPMLVDLHLQGRDLGKVFVTEIRLDEVEQAFDRMHAGDVLRVSVV 361
```

## Alignment 2:

The alignment of the peptide sequence of MSH-S-conjugate amidase from *M. tuberculosis* (Accession No. NP\_215598) and the best matching sequence from *S. coelicolor*, a conserved hypothetical protein (Accession No. NP\_629119): 49 % identity and 62 % similarity (E value =  $5.7 \times 10^{-70}$ ).

Query = NP\_215598, Sbjct = NP\_629119

```
Query:      3 ELRLMAVHAHPDDESSKGAATLARYADEGHRVLVVTLTGGERGEILNPAMDLPD-VHGRI 61
              +LRLMAVHAHPDDESSKGAAT+A+Y EG VLVVT TGGERG ILNP + + I
Sbjct:      4 QLRLMAVHAHPDDESSKGAATMAKYVSEGVDVLVVTCTGGERGSILNPKLQGDAYIEENI 63

Query:     62 AEIRRDEMTKAAEILGVEHTWLG FVDSXXXXXXXXXXXXXXXXXCFARVPLEVSTEALVRVVR 121
              E+RR EM +A EILGV WLG FVDS CFA ++ + LVR +R
Sbjct:     64 HEVRRKEMDEAREILGVGQEWLGFVDSGLPEGDPLPPLPEGCFALVDKAAGELVRKIR 123

Query:    122 EFRPHVMTTYDENG GYPHPDHIRCHQVSVAAYEAAGDFCRFPDA--GEPWTVSKLYYVHG 179
              FRP V+TTYDENG GYPHPDHI H++++ A+E A D ++P++ G + K+YY G
Sbjct:    124 SFRPQVITTYDENG GYPHPDHIMTHKITMVAFEGAADTEKYPESEYGTAYQPLKVYYNQG 183

Query:    180 FLRERMQMLQDEFARHGQRGPFEQWLAYWDPDHDFLT SRVTTRVECSKYFSQRDDALRAH 239
              F R R + L G P+E WL W + + +TT V C+ +F RD AL AH
Sbjct:    184 FNRPRTEALHHALLDRGLESPYEDWLKRWS-EFERKERTLTTHVPCADFFEIRDKALIAH 242

Query:    240 ATQIDPNAEFFAAPLAWQERLWPTEEFELARSRIPARPPETELFAGI 286
              ATQIDP +F P+ Q+ +WPTEE+ELA+S + PE +LFAGI
Sbjct:    243 ATQIDPEGGWFRVPM EIQKEVWPTEEYELAKSLVETSLPEDDLFAGI 289
```

### Alignment 3:

The alignment of the translated peptide sequence of MSSM reductase from *M. tuberculosis* (Accession No. AF002193) and the best matching sequence from *S. coelicolor*, a putative pyridine nucleotide-disulfide oxidoreductase (Accession No. NP\_627648): 28 % identity and 43 % similarity (E value =  $6.6 \times 10^{-37}$ ).

Query = AF002193; Sbjct = NP\_627648

```
Query:      1 METYDIAIIGTSGNSILDERYASKRAAIC--EQGTFGGTCLNVGCIPTKMFVYAAEV-- 56
      ++ D+ +IG G G + ER A + + E GG C CIP+KM + +
Sbjct:      2 VQEIDVVVIGMGVGEHVAERLAAEGLDVGVEAELVGGECPYVVCIPSKMMIRGGNLLA 61

Query:     57 -AKTIRG-ASRYGIDAHIXXXXXXXXXXXXXXFRIDPIALSGEDYRRCAPNIDVYRTHTRF 114
      A+ I G A + + D +A+ R R R
Sbjct:     62 EARRIPGMAGQAQVTPDFAPVASRIREEATDDWNDQVAVD----RFTGKGGQFVRGRARL 117

Query:    115 -GP--VQADGRYLLRTDAGEEFTAEQ-VVIAAGSRPVIPPAILASGVVDYHTSDTVMRIAE 170
      GP V+ DGR EFTA + VVIA GSRP IPP V Y T+ + + +
Sbjct:    118 AGPKRVEIDGR-----EFTARRGVVIATGSRPQIPPVPLDAVPYWTNREAIAVKD 168

Query:    171 LPEHIVIVSGSFIAAEFAHFVFSALGVRVTLVIRGSCLLRHCDTICERFTRIASTKW-EL 229
      P ++++G G + E A F+ G VT+V LL + + + +
Sbjct:    169 SPRSLMVLGGGAVGVELAQAFARFGTAVTVVEAAERLLPAEEPETSALLADVLGAEGITV 228

Query:    230 RTHRNVDGQQRGSGVALRLDDGCTINADLLLIVATGRVSNADLLDAEQAGVDVEDGRVI- 288
      RT + G L L+ G + A+ LLVATGR + L G+D D R +
Sbjct:    229 RTGAQANRARHDGDTFTLTLEGGEELTAERLLVATGRRAGLAGLGLTVGLD-PDARALH 287

Query:    289 VDEYQRTSARGVFALGDVSSPYLLKHVANHEARVVQHNLCDWEDTQSMIVTDHRYVPAA 348
      VDE R +A G++ +GDV+ HVA ++A + +L Q D+R +P
Sbjct:    288 VDEQLR-AAPGLWGVGDVTGRGAFTHVAMYQAEI AVRAIL-----GQPGPGADYRALPRV 341

Query:    349 VFTDPQIAAVGLTENQAVAKGLDISVKIQDYGDVAYGWAME-DTSGIVKLITERGSGRLL 407
      FTDP++ +VGLTE +A KGL + A GW + G++KL+ + G L+
Sbjct:    342 TFTDPEVGSVGLTEARAREKGLRVRTGTAQVPSSARGWIHKAGNEGLIKLVEDVDRGVLV 401

Query:    408 GAHIMGYQASSLIQPLIQAMSFGLTAAEMARGQYWIHPALPEVVENALLGLR 459
      GA G ++ L+ A+ G + R + +P VE+AL LR
Sbjct:    402 GATSAGPMGGEVLYGLVAVQ-GEVPVDRLRHMMYAYPTFHRAVEDALRALR 452
```

#### Alignment 4:

The alignment of the peptide sequence of MSH-dependent maleyl pyruvate isomerase from *C. glutamicum* (Accession No. NP\_602215) with the best matching sequence from *S. coelicolor*, a hypothetical protein with no hypothesized function (Accession No. NP\_626223): 30 % identity and 43 % similarity ( $9.0 \times 10^{-17}$ ).

Query = NP\_602215; Sbjct: NP\_626223

```
Query:   13 LTLARLGTSHYSRQLSLVDNAEFGEHSLLEGWTRSHLIAHVAYNAIALCNLMHWANTGEE 72
          L  R  T      +  +DNA  E S L GW+R H++AH+A NA AL N++      E
Sbjct:   8 LASVRDATDRLLTAVGKLDNAFVTESSRLPGWSRGHVLHLARNADALVNVL-----EG 61

Query:   73 TPMYVSPPEARNEEIIAYGSTLNPDALRNLHEHSVARLDVAWRETSEDAWSHEVLTAQGRT 131
          PMYVS EAR+ +I  +   DA L ++ E S AR      ++ W+ V   G T
Sbjct:   62 RPMYVSGEARDADIERDAPRPLDAQLADVRE-SAAFQDVGGAAPAD--WTRTVELRNGVT 118

Query:   132 VPASETLMWRSREVIHAVDLGAVATFGDIPEVILRTLAAEITQKWTSQGAGEGLVLLDE 191
          AS  + R  EV +H VDLGA      D+P              +  ++          L D
Sbjct:   119 DSASRVPFRRWAEVELHHVDLGAGYELEDLPAEFTERETDFLAARFAGHPDVPPTRLTDG 178

Query:   192 PSSTRYPAPGQDEVVVGSLAGIVRYAAGR-GSDGVTSSSTGEVPE-PP 238
          +          EV V+G  A ++ + AGR      +T   G +P  PP
Sbjct:   179 TRAWTTGRESDAPEVTVTGPPADLLGWLAGRRAGSALTVEGGALPTLPP 227
```

#### Alignment 5:

The alignment of the peptide sequence of GlxI from humans (Accession No. NP\_006699) with the best matching sequence from *S. coelicolor*, a putative lyase (Accession No. NP\_625147): 34 % identity and 57 % similarity (E value =  $1.6 \times 10^{-5}$ ) to a portion of the human enzyme.

Query = NP\_006699; Sbjt = NP\_625147

```
Query:   128 IGIAPDVYSACKRFEELGVKFKVPDDGKMKGLAFIQDPDGYWIEI 174
          + I V D+      + +  LG +F+  P D K +      +++DPDGY IE+
Sbjct:   87 LNIRVADIERCHREWSALGAEFLTPPIDRKAELRCYL RDPDGYLIEV 133
```

**Alignment 6:**

The alignment of the peptide sequence of GlxI from yeast (*S. cerevisiae*, Accession No. CAC16163) with the best matching sequence from *S. coelicolor*, a putative dioxygenase believed to belong to the glyoxalase/bleomycin resistance protein/dioxygenase superfamily (Accession No. NP\_626233): 23 % identity and 43 % similarity (E value = 0.12) to a portion of the *S. cerevisiae* enzyme.

Query = CAC16163; Sbjct = NP\_626233

```
Query: 182 KFNHTMIRIKNPTRSLEFYQNVLGMKLLRTSEHESAKFTLYFLGYGVPKTDSVFSCEVSL 241
      + +H ++ +++P + FY+ LG++ LR +E+ + T+ F V D E++L
Sbjct: 12 RLDHIVLWVRDPVAAAGFYEKNLGLEPLRITEYAAG--TVSFPS--VRLND-----ETIL 62

Query: 242 ELTHNWGTENDPNFHYHNGNSEPQGY--GHICISCD--DAGALCKEIE---VKYGDKIQW 294
      +L + + G G+ HIC+S D AL +E + D + +
Sbjct: 63 DLAPHAMADR---MRMVPGADASAGHPVNHICVSLSPHDFDALRTRLEEQSIPVSD-LSY 118

Query: 295 SPKFNQGRMKNIAFLKDPDGYSIEVVPY 322
      +G + + DPDG IE Y
Sbjct: 119 DSYGARGMARRSFYFGDPDGNIIEARHY 146
```



### Alignment 7:

The alignment of the peptide sequence of GlxI from *E. coli* (Accession No. 1FA8A) with the best matching sequence from *S. coelicolor*, a conserved hypothetical protein believed to belong to the glyoxalase/bleomycin resistance protein/dioxygenase superfamily (Accession No. NP\_625386): 29 % identity and 49 % similarity (E value = 0.00061) to a portion of the *E. coli*.

Query = 1FA8A; Sbjct = NP\_625386

```
Query:      6 TMLRVGDLQRSIDFYTKVLGMKLLRT-SENPEYKYSLAFVGYGPETEEAVIELTYNWGVD 64
              T+LR  D  RS  FY + LG+ + R      PE + ++ F+G G      +E++  G
Sbjct:      9 TLLRPTDPDRSRAFYGEQLGLAVHREFGTGPE-RGTVYFLGNG-----FLEVS---GRS 58

Query:     65 KYELGTAYGHIALSVDNAAEACEKIRQNGGNVTREAGPVKGGTTVIA-FVEDPDGYKIEL 123
              L      + + V++AA A E++  G + R  PVK      ++  ++ DPDG +I L
Sbjct:     59 DTPLTPGLA-LWIQVEDAAAHEELGAKGVEIVRP--PVKEPWGLVEMWIADPDGTRIVL 115

Query:     124 IE 125
              +E
Sbjct:     116 VE 117
```

### Alignment 8:

The alignment of the peptide sequence of GlxI from *L. major* (Accession No. AAT98624) with the best matching sequence from *S. coelicolor*, a conserved hypothetical protein believed to belong to the glyoxalase/bleomycin resistance protein/dioxygenase superfamily (Accession No. NP\_625386): 31 % identity and 50 % similarity (E value =  $8.2 \times 10^{-7}$ ) to a portion of the *L. major* enzyme.

Query = AAT98624; Sbjct = NP\_625386

```
Query:      9 TMIRVGDLDRSIKIFYTERLGMKVLKWDVPEDKYTLVFLGYG-PEMSS-TVLELTNYGV 66
              T++R  D  DRS  FY E+LG+ V R++      ++ T+ FLG G  E+S  +  LT  +
Sbjct:      9 TLLRPTDPDRSRAFYGEQLGLAVHREFGTGPERGTVYFLGNGFLEVSGRSDTPLTPGLAL 68

Query:     67 TSYKHDEAYGHIAIGVEDVKELVADMRKHDVPIDYEDESGFMA-FVVDPDGYYIELL 122
              D A H  +G + V E+V      P+  ++  G +  ++ DPDG  I L+
Sbjct:     69 WIQVEDAAAHEELGAKGV-EIVRP-----PV--KEPWGLVEMWIADPDGTRIVLV 116
```

### Alignment 9:

Sequence alignment of the GlxI amino acid sequences from *E. coli* (Accession No. 1FA8A) and *L. major* (Accession No. AAT98624) with the best matching peptide sequence from *S. coelicolor*.<sup>‡‡</sup> The residues in bold correspond to the ligands known to bind the catalytic metal in *E. coli*.

```
EcoliGlxI      ---MRLLHTMLRVGDLQRSIDFYTKVLGMKLLRTSENPEYKYSLAFVGYGPETEEAVIEL 57
LmajorGlxI    MPSRRMLHTMIRVGDLDRSIKFYTERLGMKVLRKWDVPEDKYTLVFLGYGPEMSSTVLEL 60
Scoelicolor   MVHVLSSRTLLRPTDPDRSRAFYGEQLGLAVHREFGTGPERGTVYFLGNG-----FLEV 54
               :*::*  * :**  ** : ** : *           : :: ** *          .:::

EcoliGlxI      TYNWGVDKYELGTAYGHIALSVDNAAEACEKIRQNGGNVTREAGPVKGGTTVIAFVEDPD 117
LmajorGlxI    TYNGVTSYKHDEAYGHIAIGVEDVKELVADMRKHDVPIDYED-----ESGFMAFVDDPD 115
Scoelicolor   S---GRSDTPLTPGL-ALWIQVEDAAAHEELGAKGVEIVRPP-VKEPWGLVEMWIADPD 109
               :  *  .      .      : : *::.      .:  .:  :           .  :: ***

EcoliGlxI      GYKIELIEEK----DAGRGLGN---- 135
LmajorGlxI    GYYIELLNEKTMMEKAEADMKEQGTA 141
Scoelicolor   GTRIVLVEVP-----AEHPLRYRPGI 130
               *  * *::      *      :
```

### Alignment 10:

The alignment of the peptide sequence of GlxII from humans (Accession No.CAA62483) with the best matching sequence from *S. coelicolor*, a possible hydrolase (Accession No. NP\_625789): 28 % identity and 45 % similarity (E value =  $9.5 \times 10^{-5}$ ) to a portion of the human protein.

Query = CAA62483; Sbjct = NP\_625789

```
Query:      88 KITHLSTLQVGSLNVKCLATPCHTSGHICYFVSKPGGSEPPAVFTGDTLFVAGCGK--FY 145
             ++T  + L +  L +      P HT G ++ + +      PP +F+GD LF      G+  F
Sbjct:     116 ELTDGARLALAGLELTVAHAPGHTKGSVTFLPE-AADIPPVMFSGDLLFAGSIGRTDFP 174

Query:     146 EGTADEMCKALLEVLGRLPPDTRVYCGH--EYTINNLK----FARHVEPGNAA 192
             G+ D+M ++L  V   L   T V  GH  + TI  +   + R V  G  A
Sbjct:     175 GSMDDMLESLARVCLPLDDSTVVLSGHPQTTIQERATNPYLRQVAAGQGA 227
```

<sup>‡‡</sup> The symbols below the alignment indicate the similarity of the sequences: “\*” indicates that the residues in that column are identical; “.” indicates that conserved substitutions are observed and “:” indicates a semi-conserved substitution.

**Alignment 11:**

The alignment of the peptide sequence of GlxII from yeast (*S. cerevisiae*, Accession No. CAA58065) with the best matching sequence from *S. coelicolor*, a possible hydrolase (Accession No. NP\_626224): 27 % identity and 40 % similarity (E value =  $1.6 \times 10^{-5}$ ) to a portion of the yeast enzyme.

Query = CAA58065; Sbjct = NP\_626224

```
Query:   26 NYSYLLSTEDRRNSWLIDPAEPLVSPKLSAEKKSIDAIVNTTHHHYDHSGGNLALYSIL 85
          N +YLL          LID A  +   L          I ++V TH H DH          + +
Sbjct:   33 NNAYLLRCRATDEQLLIDAAN--DAHTLLGTIGDDGIASVVTTHRHGDHWQALAEVVAAT 90

Query:   86 -CQENSG-HDIKIIGGSKSSPGVTEVP-DNLQQYHLGNLRVTCIRTPCHTKDSICYIYIKD 142
          + ++G HD + I          P T+V D+          +G + +T          HT S+   D
Sbjct:   91 GARTHAGRHD AEGI-----PVPTDVLIDDGDTIRVGQVELTARHIAGHTPGSVVLVYDD 144

Query:  143 LETGEQCIFTGDTLFIAGCGRFFEGTGRDMDMALNQIMLRVGETNWNKVKIYPGH 198
          G  +FTGD LF  G G  F+   D  +L  +   V +T  ++  +YPGH
Sbjct:  145 PH-GHPHVFTGDCLFPGGVGNTFKDP--DAFASLIHDVQTKVFDLTPDETWVYPGH 197
```

**Alignment 12:**

The alignment of the peptide sequence of GlxII from *T. brucei brucei* (Accession No. CAD37800) with the best matching sequence from *S. coelicolor*, a possible hydrolase (Accession No. NP\_626224): 27 % identity and 47 % similarity (E value =  $1.2 \times 10^{-5}$ ) to a portion of the *T. brucei brucei* enzyme.

Query = CAD37800; Sbjct = NP\_626224

```
Query:   64 LRTILSTHKHHDHSGGNISLQKKNAMGAFRIIGGAN--EPIPGVTEK-VREGDHFSIGE 120
      + +++++TH+H DH      +L + + A GA R   G + E IP T+  + +GD  +G+
Sbjct:   67 IASVVTTHRHGDHWQ---ALAEVVAATGA-RTHAGRHDAEGIPVPTDVLIDDGDTIRVGQ 122

Query:  121 LKVDVLDAPCHTSGHVLYKVYHPQKAENGIALFTGDTMFVGGIGAFFEGDAVLMCSALRK 180
      +++          HT G V+  VY          + FTGD +F GG+G  F+ D      S +
Sbjct:  123 VELTARHIAGHTPGSVVL-VYDDPHGHPHV--FTGDCLFPGGVGNTFK-DPDAFASLIHD 178

Query:  181 VYNLNGACESSACDATDVQKRDNHTYIFPGH 211
      V          ++   D          + T+++PGH
Sbjct:  179 V-----QTKVFDTLP-----DETWVYPGH 197
```

**Alignment 13:**

The alignment of the peptide sequence of GlxII from *L. infantum* (Accession No. ABC41261) with the best matching sequence from *S. coelicolor*, a possible hydrolase (Accession No. NP\_626224): 26 % identity and 47 % similarity (E value =  $2.0 \times 10^{-6}$ ) to a portion of the *L. infantum* enzyme.

Query = ACB41261; Sbjct = NP\_626224

```
Query:   70 STILSTHKHWDHSGGNAKLKAELEAMNSTVPPVVVGGANDS--IPAVTKP-VREGDRVQV 126
          ++++++TH+H DH   A++ A  A           G +D+ IP T  + +GD ++V
Sbjct:   68 ASVVTTHRHGDHWQALAEVVAATGARTHA-----GRHDAEGIPVPTDVLIDDGDTIRV 120

Query:  127 GDLSVEVIDAPCHTRGHVLYKVQHPQ-HPNDGVALFTGDTMFIAGIGAFFEGDEKDMCRA 185
          G + +           HT G V+   P  HP+   +FTGD +F  G+G  F+  + D  +
Sbjct:  121 GQVELTARHIAGHTPGSVVLVYDDPHGHPH----VFTGDCLFPGGVGNTFK--DPDAFAS 174

Query:  186 MEKVYHIHKGNDYALDKVTFIFPGH 210
          +  ++ +           L  T+++PGH
Sbjct:  175 L--IHDVQTKVFDLTPDETWVYPGH 197
```

**Alignment 14:**

Sequence alignment of the GlxII amino acid sequences from yeast (*S. cerevisiae*), *T. brucei brucei* and *L. infantum* (Accession Nos. CAA68065, CAD37800, and ABC41261 respectively) with the best matching peptide sequence from *S. coelicolor* (Accession No. NP\_626224).<sup>††</sup>

```

TbruceiGlxII      -----MEVVVKSIGTAFTVAVIPVLKD---NYSYVIHDKATNTLAAVDVSVSDIDP 47
LinfantumGlxII   -----MRNYCTKTFGSAFSVTVPVPTLKD---NFSYLINDHTTHTLAAVDVNADYKP 48
YeastGlxII       -----MKFLLQQIRNMHVKPIKMRWLTGGVNYSYLLSTEDRRNSWLIDPAEPLEV 50
Scoelicolor      MTYSGEVTVGGPADVHELKDLMITKVAVGPMN---NNAYLLRCRATDEQLLIDAANDAHT 57
                  :   :           :           :*   :*:   .           :*           .

TbruceiGlxII      VIDYVRR----LGGVDRTTDLRTLILSTHKHHDHSGGNI SLQKKNAMG---AFRIIGGAN 100
LinfantumGlxII   ILTYIEEHLKQQGNADVITYTFSTILSTHKHWDHSGGNAKLKAELEAMNSTVPVVVGGAN 108
YeastGlxII       SPKLSAE-----EKSIDAIVNTHHHYDHS GGNLALYSILCQENSGHDIKIIGGS- 100
Scoelicolor      LLGTIGD-----DGIASVVTTHRHGDH---WQALAEVVAATGARTHAGRHDAEG 103
                  :   : : : . * : * * * * * * * * * * * * * * * * * * * * * *

TbruceiGlxII      EPIPGVTEKVVREGDHF SIGELKVDVLDAPCHTSGHVLYKVYHPQKAENGIALFTGDTMFV 160
LinfantumGlxII   DSIPAVTKPVREGDRVQVGDLSVEVIDAPCHTRGHVLYKVQHPQHPNDGVALFTGDTMFI 168
YeastGlxII       KSSPGVTEVPDNLQQYHLGNLRVTCIRTPCHTKDSICYIYKDLLETGEQ--CIFGTGDTLFI 158
Scoelicolor      IPVP-TDVLIDDGDTIRVQVELTARHIAGHTPGSVVLVYDDPHGHPH---VFTGDCLFP 159
                  . * .           :   :   : * : :   . * * . :           . .           : * * * * * * *

TbruceiGlxII      GGIGAFFEGDAVLMCSALRKVYNLNGACESSACDATDVQKRDNHTYIFPGHEYTVNFLRF 220
LinfantumGlxII   AGIGAFFEGDEKDMCRAMEKVYHIHKGN-----YALDKVTFIFPGHEYTSGFMTF 219
YeastGlxII       AGCGRFFEGTGRDMDMALNQIMLRVGETN-----WNKVKIYPGHEYTKGNVSF 207
Scoelicolor      GGVGNTFK-DPDAFASLIHDVQTKVFDTL-----DETWVYPGHGNDT----- 201
                  . * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *

TbruceiGlxII      SRDALP-ASHPDVSFVEAQLRRYTESVAGNVPTVPSTLAEKRQNLFLRTCDEAFVRVMN 279
LinfantumGlxII   SEKTFPDRASDDLAFIQAQRAKYAAAVKTDGPSVPSLAEKRQNLFLRVADPAFVAKMN 279
YeastGlxII       IRAKIYSDIGQNKEFDALQYCKSNECTTG----HF TLRDELGYNPFMRLDDRAVRLAVG 263
Scoelicolor      -----TLGAERPHLPEWRARGW----- 218
                  :           :           : * * * * * * * * * * * * * * * *

TbruceiGlxII      KGETAVKLMDFLYNTCP----- 296
LinfantumGlxII   QG-NAHALMMYLYNACD----- 295
YeastGlxII       DTAGTYPRSVVMQELRKLKNAM 285
Scoelicolor      -----
                  :           :           :

```

**Table 4.7:** Results for the sequence searches of the *Streptomyces coelicolor* genome for homologous sequences to known glyoxalase enzymes.

<b>Query Organism</b>	<b>Accession #</b>	<b><i>S. coelicolor</i> Sequence</b>	<b>Putative Role</b>	<b>% Identity</b>	<b>% Similar<sup>a</sup></b>	<b><i>E</i> value (<math>\times 10^{-5}</math>)</b>
<b>GlxI</b>						
Human	NP_006699	NP_625147	Lyase	34 %	57 %	1.6
Yeast	CAC16163	NP_626233	Dioxygenase <sup>b</sup>	23 %	43 %	12000
<i>E. coli</i>	1FA8A	NP_625386	None <sup>c</sup>	29 %	49 %	61
<i>L. major</i>	AAT98624	NP_625386	None <sup>c</sup>	31 %	50 %	0.082
<b>GlxII</b>						
Human	CAA62483	NP_625789	Hydrolase <sup>d</sup>	28 %	45 %	9.8
Yeast	CAA58065	NP_626224	None <sup>b,c</sup>	27 %	40 %	1.6
<i>T. brucei</i>	CAD37800	NP_626224	None <sup>c</sup>	27 %	47 %	1.2
<i>L. infantum</i>	ABC41261	NP_626224	None <sup>c</sup>	26 %	47 %	0.2

<sup>a</sup>% identity and similarity to portions of the query enzymes; <sup>b</sup>May belong to the glyoxalase superfamily; <sup>c</sup>This peptide sequence, identified from the genome, has not yet been assigned a putative function; <sup>d</sup>Similarity to the glyoxalases

#### 4.3.2. Choice of Organism

*S. coelicolor* was chosen for our investigations into the presence of an MSH-dependent Glx system, as it is known to produce MSH as its major thiol<sup>4</sup> and the genes responsible for biosynthesis of MSH having been identified.<sup>104</sup> The *S. coelicolor* genome has been sequenced<sup>16</sup> and is available for homologous sequence searching,<sup>†</sup> which allows for the development of an over-expression system for a Glx enzyme should it be identified. The isolation of an MSH-dependent Glx system would represent the first MSH-dependent enzyme identified in *Streptomyces* bacteria.

#### 4.3.3. Preliminary Assays for the Identification of a Mycothiol Dependent Glyoxalase System

Cell-free extracts of *S. coelicolor* were prepared using sonication in Lysis Buffer 1. The thawed *S. coelicolor* cells were very clumpy and the initial five to ten sonication intervals distributed the cells more evenly in the buffer. EDTA (0.1 mM), a chelating agent, was included in the lysis buffer to inhibit any metal dependent proteases present in the extract. The inclusion of EDTA in the lysis buffer complicates the assays used to identify Glx type activity as all of the identified GlxI enzymes are metalloenzymes and are maximally activated by either nickel (most eukaryotes), or zinc (*E. coli* and *L. major*). The GlxII enzymes are also metalloenzymes; however, the *in vivo* metal preferences of this enzyme have not yet been identified. It is likely that, in addition to inhibiting metal dependent proteases, EDTA could remove the essential metal from the putative MSH-dependent Glx enzymes. A cocktail of common metals was added to the various assay mixtures in hopes of reconstituting the metal in the putative MSH-dependent Glx metalloenzymes.

To fully test for the presence of an MSH-dependent glyoxalase system, multiple possibilities must be explored. With the exception of the GlxIII enzyme reported in *E. coli*, all reported Glx systems use two enzymes to convert cytotoxic MG to D-lactate. The first

enzyme in this pathway, GlxI, converts the hemithioacetal of MG and GSH or TSH<sub>2</sub> to *S*-D-lactoylglutathione or *S*-D-lactoyltrypanothione, both of which have a characteristic absorbance at 240 nm. Assays to monitor GlxI activity, therefore, typically observe changes in absorption at this wavelength. In addition, MG and the thiol of interest are usually incubated for at least 15 minutes before the assay to allow for the formation of the hemithioacetal. It is likely that should an MSH-dependent GlxI enzyme exist, the resultant *S*-D-lactoylmycothiol would also have a characteristic absorbance at 240 nm; therefore, Assays 1 – 6 monitored the change in absorbance at this wavelength. Assay Mixture 2 included GSH, as well as MG, in hopes that an MSH-dependent Glx system would accept GSH in lieu of MSH, as MSH was unavailable at the time of this experiment.

Unlike the product of GlxI, the product of GlxII, D-lactate, does not demonstrate absorbance at a unique wavelength. Traditionally, GlxII activity is monitored by observing a decrease in absorbance at 240 nm, which corresponds to the conversion of the *S*-D-lactoylglutathione, or *S*-D-lactoyltrypanothione, to D-lactate. The activity of GlxII can also be monitored using D-LDH which converts D-lactate to pyruvate, with the simultaneous reduction of NAD<sup>+</sup> to NADH, the latter having a specific absorbance at 340 nm. An increase at 340 nm would therefore correspond directly to the formation of D-lactate. A pyruvate trapping agent, such as hydrazine, and a pH of ~9 is often used to force the equilibrium towards the formation of pyruvate. It is possible that an MSH-dependent Glx system could convert MG to L-lactate, rather than D-lactate; therefore both possibilities were investigated in Assays 5 and 6.

Unfortunately these assays were inconclusive. Assays 1 – 4 did not show a significant change in absorption at 240 nm. If *S*-D-lactoylmycothiol was formed, it is possible that it was consumed by MSH-dependent GlxII too quickly for an increase in absorbance to be noted. Assays 5 – 6 were difficult to interpret due to the large amount of background absorbance in these assays. The unfiltered lysate contained a large number of metabolites and other enzymes which could produce substrates for D and L-LDH.

#### 4.3.4. Comparative Assays Performed in *Escherichia coli*

The assays aimed at the detection of Glx activity were repeated using cell-extracts from *E. coli*, an organism known to possess a Glx system. Assays 1 – 4, which test for GlxI activity by observing a change in absorbance at 240 nm, performed using unfiltered cell-free extract did not show a change in absorbance at 240 nm. This result is not surprising as GlxII would be present in the assay mixture and would very quickly convert any *S*-D-lactoylglutathione formed to D-lactate. Assays 5 and 6, which used a coupled assay with D-LDH, did show a change in absorbance at 340 nm with the unfiltered cell-free extract. This result is consistent with the reduction of NAD<sup>+</sup> to NADH, indicating the formation of D-lactate and GlxII activity; however, an increase in absorbance was also seen with L-LDH. The unfiltered cell-free extract would contain a variety of metabolites, including L-lactate, which could account for the increase in absorbance at 340 nm.

Assays 1 – 4 were repeated using the filtered cell-free extract from *E. coli* and no increase in absorbance was noted at 240 nm; however, assays 5 and 6 did show an increase in absorbance at 340 nm when D-LDH was used. In addition, no increase was seen with the use of L-LDH. Combined, these results indicate the presence of a Glx system. The lack of absorbance noted with Assays 1 – 4 is likely due to the fast consumption of *S*-D-lactoylglutathione to D-lactate. The increase in absorbance at 340 nm in the absence of



GSH is attributed to the activity of GlxIII, which converts the MG directly to D-lactate without the need for GSH.

Based upon the results of the *E. coli* assays, it is likely that the accurate detection of MSH-dependent Glx activity will require the removal of all small molecules from cell-free extracts followed by the readdition of MG and MSH. We therefore turned our attention to the isolation of MSH and the synthesis of des-*myo*-inositol mycothiol, a truncated version of MSH shown to be useful in the identification of MSSM reductase activity.

#### 4.3.5. Synthesis of Des-*myo*-Inositol-Mycothiol

The synthesis of des-*myo*-inositol mycothiol has been reported<sup>113</sup> and was followed closely with the exceptions noted below. The coupling of Fmoc-Cys(Acm)-OPfp (**4-44**) and D-glucosamine (**4-45**) did not go to completion, as some starting amino acid remained. The activated ester of the remaining starting material would have been cleaved during the reaction to give Fmoc-Cys(Acm)-OH, which is also insoluble in water and would precipitate with the product during purification. To separate the amino acid from the product, the precipitate was dissolved in methanol and stirred with a weakly basic anion exchange macroporous resin to bind the amino acid. The solution was then filtered to remove the resin affording Fmoc-Cys(Acm)-GlcN (**4-46**) in 56 % yield.

As detailed by the original protocol, the Fmoc group was removed with 5 % piperidine in DMF (99 % yield of **4-47**) and acetylation of the  $\alpha$ -amino group was performed using acetic anhydride in water with a pH of  $\sim 9$ , maintained by potassium carbonate. This method facilitated the selective acetylation of the  $\alpha$ -amino group, as opposed to esterification of the hydroxyl groups of D-glucosamine, to yield AcCys(Acm)-GlcN (**4-48**).<sup>113</sup>

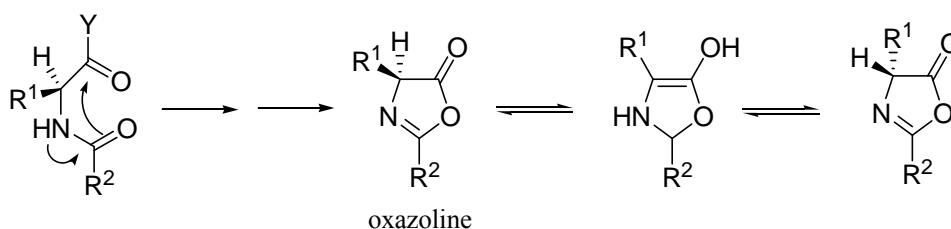
In an attempt to avoid the use of highly toxic thallium trifluoroacetate, the oxidative deprotection of the thiol was attempted using iodine,<sup>118</sup> one of the most simple, less expensive, less toxic oxidants.<sup>120</sup> Iodine has been shown to be successful in the removal of *S*-acetamidomethyl and *S*-trityl groups from cysteine with the simultaneous formation of cystine.<sup>118</sup> Unfortunately, the use of iodine in this instance did not appear to remove the *S*-acetamidomethyl group and the resultant <sup>1</sup>H NMR spectrum was convoluted. Catalytic amounts of iodine are known to catalyze the acetylation of D-glucose in the presence of acetic anhydride.<sup>120,121</sup> By <sup>1</sup>H NMR spectroscopy, it appeared that there was a small amount of residual acetic anhydride remaining after the purification of AcCys(Acm)-GlcN; therefore, it is possible that a mixture of products was formed with varying levels of acetylation. The use of iodine as an alternative was rejected and future experiments will use thallium trifluoroacetate, as described by Patel and Blanchard,<sup>113</sup> for the deprotection of cysteine with the simultaneous formation of the disulfide of des-*myo*-inositol mycothiol.

#### 4.3.6. Alternative Synthesis of Des-*myo*-Inositol Mycothiol

AcCys(Trt)-OH (**4-52**) was used as a starting point in lieu of FmocCys(Acm)-OPfp in an attempt to simplify the production of des-*myo*-inositol mycothiol. The use of a pre-acetylated starting material eliminates two steps from the published protocol<sup>113</sup> and removes the possibility of the esterification of the hydroxyl groups of D-glucosamine during the acetylation of the  $\alpha$ -amino group. The *S*-triphenylmethyl thioether protecting group was chosen to increase the organic solubility of the molecule. Solely water soluble compounds cannot be purified with standard silica gel chromatography and even removing the solvent is

time consuming. Increasing the product's solubility in organic solvents opens more options for purification and makes it easier to handle.

During the coupling of AcCys(Trt)-OH to D-glucosamine using HBTU racemization at the  $\alpha$ -carbon occurred as expected. Racemization can occur during the base catalyzed coupling of an *N*-acyl-protected, carboxyl-activated amino acid through the mechanism shown in Figure 4.20. This racemization can be minimized with the use of nonpolar solvents, a minimum amount of base and low reaction temperatures.<sup>122</sup> Due to the nature of the components involved in these reactions, such as D-glucosamine, nonpolar solvents are not an option. In addition, the low solubility of D-glucosamine in DMF eliminates the use of low reaction temperatures. Although it was known a racemic mixture was likely, this method was initially an attractive option as it was thought that the mixture could be easily prepared in large quantities and would be useful in testing for activity during the expected purification of MSH-dependent enzymes. The racemic mixture was very difficult to purify, therefore this method was abandoned and all further efforts were focused on the published method for the production of des-*myo*-inositol mycothiol.



**Figure 4.20:** The mechanism for the racemization of amino acids that can occur during the base-catalyzed coupling reaction of an *N*-protected, carboxyl-activated amino acid. Racemization takes place in the intermediate oxazolone that forms readily from an *N*-acyl-protected amino acid ( $R^2$  = alkyl, aryl). Adapted from Greene and Wuts (1999)<sup>122</sup>

#### 4.3.7. Synthesis of 2-*S*-(2'-Thiopyridyl)-6-hydroxynaphthyl disulfide and Isolation of Mycothiol

The synthesis of 2-*S*-(2'-thiopyridyl)-6-hydroxynaphthyl disulfide (**4-50**) was unsuccessful when the protocol detailed by Steenkamp and Vogt<sup>115</sup> was followed exactly. The initial step, the reduction of 6-hydroxynaphthyl disulfide (**4-55**) by sodium borohydride in 1:1 acetonitrile:water, failed. A review of the literature regarding the reactivity of sodium borohydride indicated that the optimum pH for this reagent is 9.<sup>123</sup> The use of a 1:1 acetonitrile:0.1 M sodium bicarbonate solution with a pH of 9 as the solvent resulted in the successful reduction of the 6-hydroxynaphthyl disulfide. Also unlike the published protocol, all solvents were degassed before use and all reactions proceeded under argon to reduce the chance of oxidation of any free thiols in solution.<sup>115</sup>

The published protocol used *M. smegmatis* as the source of MSH<sup>115</sup> the largest known producer of this intracellular thiol.<sup>4</sup> *M. smegmatis* is considered a class two biohazard, as it can cause tuberculosis-like symptoms in marine mammals. The facilities available at the time of these experiments were not equipped to handle such an organism; therefore, *S. jumonjinensis*, a non-pathogenic organism was used as the source of MSH for this work, as *S. jumonjinensis* produces the largest amount of MSH of the non-pathogenic *Streptomyces* bacteria (Table 4.8).<sup>4</sup>

**Table 4.8:** A comparison of the amount of mycothiol produced by *Streptomyces* bacteria to that produced by *Mycobacterium smegmatis*.

<b>Organism</b>	<b>MSH (<math>\mu\text{mol/g dry weight}</math>)</b>
<i>Mycobacterium smegmatis</i>	19.0
<i>Streptomyces clavuligerus</i>	5.40
<i>Streptomyces coelicolor</i>	2.80
<b><i>Streptomyces jumonjinensis</i></b>	<b>6.70</b>
<i>Streptomyces lactamdurans</i>	5.90
<i>Streptomyces lividans</i>	6.60

The isolation of MSH from *S. jumonjinensis* followed the protocol detailed by Steenkamp and Vogt fairly closely except that the reagent solution and the cell-free extract were degassed for 10 minutes before use and a different reverse-phase C<sub>18</sub> column was used for HPLC purification due to availability. There is a discrepancy in the published report of the purification of MSH. The text of the paper states that tagged-MSH elutes at 35.5 min; however, the chromatogram included in the manuscript labels the peak at 35.5 minutes as excess reagent and the peak at 23 min as the tagged MSH. The elution profile of this work is in better agreement with the chromatogram as tagged MSH eluted at 28 minutes, followed by excess reagents. The amount of MSH isolated by this method varies, with a maximum yield of 1.5 mg of the tagged MSH being isolated from 4 g of wet weight *S. jumonjinensis* cells.

#### 4.3.8. Synthesis of Mycothiol Analogue

The *N,N'*-diacetyl-cystinyl-cyclohexyl amide (**4-59**) was synthesized as a possible MSH analogue using HBTU as a coupling agent to give an overall yield of only 13 %. The *myo*-inositol moiety is not essential for the detection of MSSM reductase activity,<sup>113</sup> indicating that it may not be essential for the binding of MSH to its cognate enzymes. This analogue could be used to determine if the hydroxyl groups of the D-glucosamine of MSH are essential for binding or if the cyclohexyl scaffold is sufficient.

#### 4.4. Conclusions and Future Work

Ground work towards the identification of a MSH-dependent Glx system in *S. coelicolor* was completed. Preliminary assays for GlxI and GlxII activity in *S. coelicolor* were inconclusive. Similar assays for Glx activity in *E. coli*, an organism known to contain the Glx system, indicated that, for accurate detection of Glx activity, all small molecules must be removed followed by the readdition of MG and the thiol cosubstrate. Subsequent work therefore focused on the synthesis of des-*myo*-inositol mycothiol, a simplified MSH analogue used to detect MSSM reductase activity, as well as the synthesis of 2-*S*-(2'-thiopyridyl)-6-hydroxynaphthyl disulfide and its use in isolating MSH from cell-free extracts of *S. jumonjinensis*.

The obvious next steps of this project would be the removal of the acetamidomethyl group from AcCys(Acm)-GlcN to yield des-*myo*-inositol mycothiol disulfide, followed by reduction to the free thiol. Once purified, this thiol would be used to assay cell-free extracts of *S. coelicolor*, which had been passed through a 3 000 molecular weight cut-off filter, for MSH-dependent Glx activity. If activity was not detected, the assays would then be repeated

using MSH as an MSH-dependent Glx system might not accept the hemithioacetal of MG and des-*myo*-inositol mycothiol as a substrate. If activity is detected, the enzymes responsible for this activity will be purified and the resulting enzymes characterized. The identification of an MSH-dependent Glx system would be the first example of an MSH-utilizing enzyme in *S. coelicolor*. In addition, because the Glx system is an essential detoxification pathway, it would be a new target for the development of novel therapeutics aimed at *Mycobacteria*.

#### 4.5. References

1. World Health Organization, *Tuberculosis*; 2006, <http://www.who.int/mediacentre/factsheets/fs104/en/>
2. World Health Organization, *Leprosy*; 2005, <http://www.who.int/mediacentre/factsheets/fs101/en/index.html>
3. World Health Organization, *Buruli ulcer disease*; 2005, <http://www.who.int/mediacentre/factsheets/fs199/en/index.html>
4. Newton, G. L.; Arnold, K.; Price, M. S.; Sherrill, C.; Delcardayre, S. B.; Aharonowitz, Y.; Cohen, G.; Davies, J.; Fahey, R. C.; Davis, C. *J. Bacteriol.* **1996**, *178*, 1990-1995.
5. Newton, G. L.; Ta, P.; Bzymek, K. P.; Fahey, R. C. *J. Biol. Chem.* **2006**, *In press*.
6. Buchmeier, N. A.; Newton, G. L.; Koledin, T.; Fahey, R. C. *Mol. Microbiol.* **2003**, *47*, 1723-1732.
7. Rawat, M.; Newton, G. L.; Ko, M.; Martinez, G. J.; Fahey, R. C.; Av-Gay, Y. *Antimicrob. Agents Chemother.* **2002**, *46*, 3348-3355.
8. Gammon, D. W.; Hunter, R.; Steenkamp, D. J.; Mudzunga, T. T. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 2045-2049.
9. Thornalley, P. J. *Biochem. Soc. Trans.* **2003**, *31*, 1343-1348.
10. Clugston, S. L.; Daub, E.; Honek, J. F. *J. Mol. Evol.* **1998**, *47*, 230-234.
11. Clugston, S. L.; Honek, J. F. *J. Mol. Evol.* **2000**, *50*, 491-495.
12. Clugston, S. L.; Yajima, R.; Honek, J. F. *Biochem. J.* **2004**, *377*, 309-316.
13. Davidson, G.; Clugston, S. L.; Honek, J. F.; Maroney, M. J. *Biochemistry* **2001**, *40*, 4569-4582.
14. He, M. M.; Clugston, S. L.; Honek, J. F.; Matthews, B. W. *Biochemistry* **2000**, *39*, 8719-8727.
15. Sukdeo, N.; Clugston, S. L.; Daub, E.; Honek, J. F. *Biochem. J.* **2004**, *384*, 111-117.
16. Bentley, S. D.; Chater, K. F.; Cerdeño-Tárraga, A. M.; Challis, G. L.; Thomson, N. R.; James, K. D.; Harris, D. E.; Quail, M. A.; Kieser, H.; Harper, D.; Bateman, A.; Brown, S.; Chandra, G.; Chen, C. W.; Collins, M.; Cronin, A.; Fraser, A.; Goble, A.; Hidalgo, J.; Hornsby, T.; Howarth, S.; Huang, C. H.; Kieser, T.; Larke, L.; Murphy, L.; Oliver, K.; O'Neil, S.; Rabinowitsch, E.; Rajandream, M. A.; Rutherford, K.; Rutter, S.; Seeger, K.; Saunders, D.; Sharp, S.; Squares, R.; Squares, S.; Taylor, K.; Warren, T.; Wietzorrek, A.; Woodward, J.; Barrell, B. G.; Parkhill, J.; Hopwood, D. A. *Nature* **2002**, *417*, 141-147.
17. World Health Organization, *Chagas Disease*; 2004, <http://www.who.int/tdr/diseases/chagas/default.htm>
18. World Health Organization, *African trypanosomiasis (sleeping sickness)*; 2006, <http://www.who.int/mediacentre/factsheets/fs259/en/index.html>

19. Fairlamb, A. H.; Blackburn, P.; Ulrich, P.; Chait, B. T.; Cerami, A. *Science* **1985**, *227*, 1485-1487.
20. Fairlamb, A. H.; Cerami, A. *Mol. Biochem. Parasitol.* **1985**, *14*, 187-198.
21. Lopez, J. A.; Carvalho, T. U.; de Souza, W.; Flohe, L.; Guerrero, S. A.; Montemartini, M.; Kalisz, H. M.; Nogoceke, E.; Singh, M.; Alves, M. J. M.; Colli, W. *Free Radic. Biol. Med.* **2000**, *28*, 767-772.
22. Flohe, L.; Hecht, H. J.; Steinert, P. *Free Radic. Biol. Med.* **1999**, *27*, 966-984.
23. Greig, N.; Wyllie, S.; Vickers, T. J.; Fairlamb, A. H. *Biochem. J.* **2006**, *400*, 217-223.
24. Vickers, T. J.; Greig, N.; Fairlamb, A. H. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 13186-13191.
25. Padmanabhan, R. K.; Mukherjee, A.; Madhubala, R. *Biochem. J.* **2006**, *393*, 227-234.
26. Vickers, T. J.; Fairlamb, A. H. *J. Biol. Chem.* **2004**, *279*, 27246-27256.
27. Vickers, T. J.; Wyllie, S.; Fairlamb, A. H. *J. Biol. Chem.* **2004**, *279*, 49003-49009.
28. Kuriyan, J.; Kong, X. P.; Krishna, T. S.; Sweet, R. M.; Murgolo, N. J.; Field, H.; Cerami, A.; Henderson, G. B. *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 8764-8768.
29. Bailey, S.; Smith, K.; Fairlamb, A. H.; Hunter, W. N. *Eur. J. Biochem.* **1993**, *213*, 67-75.
30. Zhang, Y.; Bailey, S.; Naismith, J. H.; Bond, C. S.; Habash, J.; McLaughlin, P.; Papiz, M. Z.; Borges, A.; Cunningham, M.; Fairlamb, A. H.; et al. *J. Mol. Biol.* **1993**, *232*, 1217-1220.
31. Zhang, Y.; Bond, C. S.; Bailey, S.; Cunningham, M. L.; Fairlamb, A. H.; Hunter, W. N. *Protein Sci.* **1996**, *5*, 52-61.
32. Feng, J.; Che, Y.; Milse, J.; Yin, Y.-J.; Liu, L.; Rückert, C.; Shen, X.-H.; Qi, S.-W.; Kalinowski, J.; Liu, S.-J. *J. Biol. Chem.* **2006**, *271*, 10778-10785.
33. Misset-Smits, M.; van Ophem, P. W.; Sakuda, S.; Duine, J. A. *FEBS Lett.* **1997**, *409*, 221-222.
34. Norin, A.; van Ophem, P. W.; Piersma, S. R.; Persson, B.; Duine, J. A.; Jörnvall, H. *Eur. J. Biochem.* **1997**, *248*, 282-289.
35. van Ophem, P. W.; Van Beeumen, J.; Dune, J. A. *Eur. J. Biochem.* **1992**, *206*, 511-518.
36. Vogt, R. N.; Steenkamp, D. J.; Zheng, R.; Blanchard, J. S. *Biochem. J.* **2003**, *374*, 657-666.
37. Newton, G. L.; Av-Gay, Y.; Fahey, R. C. *Biochemistry* **2000**, *39*, 10739-10749.
38. Steffek, M.; Newton, G. L.; Av-Gay, Y.; Fahey, R. C. *Biochemistry* **2003**, *42*, 12067-12076.
39. Patel, M. P.; Blanchard, J. S. *Biochemistry* **1999**, *38*, 11827-11833.
40. Patel, M. P.; Blanchard, J. S. *Biochemistry* **2001**, *40*, 5119-5126.
41. Lack, L. *J. Biol. Chem.* **1961**, *236*, 2835-2840.
42. Uotila, L.; Koivusalo, M. In *Functions of Glutathione: Biochemical, Physiological, Toxicological and Clinical Aspects*; A. Larsson, S. Orrenius, A. Homgren, B. Mannervick, Eds.; Raven Press: New York, NY, USA, 1983, pp 175-186.
43. Pickett, C. B.; Lu, A. Y. *Ann. Rev. Biochem.* **1989**, *58*, 743-764.
44. Schirmer, H.; Krauth-Siegel, R. L.; Schulz, G. E. In *Glutathione: Chemical, Biochemical and Medical Aspects*; D. Dolphin, R. Poulson, O. Avramovic, Eds.; John Wiley and Sons: New York, 1989; Vol. Part A, pp 553-596.
45. Uotila, L. *Biochemistry* **1973**, *12*, 3944-3951.
46. Liu, L.; Hausladen, A.; Zeng, M.; Que, L.; Heitman, J.; Stamler, J. S. *Nature* **2001**, *410*, 490-494.

47. Jensen, D. E.; Belka, G. K.; Du Bois, G. C. *Biochem. J.* **1998**, *331*, 659-668.
48. Patel, M. P.; Blanchard, J. S. *J. Am. Chem. Soc.* **1998**, *120*, 11538-11539.
49. Vuilleumier, S. *J. Bacteriol.* **1997**, *179*, 1431-1441.
50. Goetz, F. E.; Harmuth, L. J. *FEMS Microbiol. Lett.* **1992**, *97*, 45-49.
51. Hareland, W. A.; Crawford, R. L.; Chapman, P. J.; Dagley, S. *J. Bacteriol.* **1975**, *121*, 272-285.
52. Robson, N. D.; Parrott, S.; Cooper, R. A. *Microbiology* **1996**, *142*, 2115-2120.
53. Zhou, N. Y.; Fuenmayor, S. L.; Williams, P. A. *J. Bacteriol.* **2001**, *183*, 700-708.
54. Crawford, R. L.; Hutton, S. W.; Chapman, P. J. *J. Bacteriol.* **1975**, *121*, 794-799.
55. Crawford, R. L.; Frick, T. D. *Appl. Environ. Microbiol.* **1977**, *34*, 170-174.
56. Richard, J. P. *Biochem. Soc. Trans.* **1993**, *21*, 549-553.
57. Richard, J. P. *Biochemistry* **1991**, *30*, 4581-4585.
58. Inoue, Y.; Kimura, A. *Adv. Microb. Physiol.* **1995**, *37*, 177-227.
59. Hopper, D. J.; Cooper, R. A. *Biochem. J.* **1972**, *128*, 321-329.
60. Fraval, H. N.; McBrien, D. C. *J. Gen. Microbiol.* **1980**, *117*, 127-134.
61. Cameron, A. D.; Ridderstrom, M.; Olin, B.; Mannervik, B. *Structure* **1999**, *7*, 1067-1078.
62. Cameron, A. D.; Olin, B.; Ridderstrom, M.; Mannervik, B.; Jones, T. A. *Embo J* **1997**, *16*, 3386-3395.
63. Cameron, A. D.; Ridderstrom, M.; Olin, B.; Kavarana, M. J.; Creighton, D. J.; Mannervik, B. *Biochemistry* **1999**, *38*, 13480-13490.
64. Hovatta, I.; Tennant, R. S.; Helton, R.; Marr, R. A.; Singer, O.; Redwine, J. M.; Ellison, J. A.; Schadt, E. E.; Verma, I. M.; Lockhart, D. J.; Barlow, C. *Nature* **2005**, *438*, 662-666.
65. Inoue, Y.; Kimura, A. *J. Biol. Chem.* **1996**, *271*, 25958-25965.
66. Wood, V.; Gwilliam, R.; Rajandream, M. A.; Lyne, M.; Lyne, R.; Stewart, A.; Sgouros, J.; Peat, N.; Hayles, J.; Baker, S.; Basham, D.; Bowman, S.; Brooks, K.; Brown, D.; Brown, S.; Chillingworth, T.; Churcher, C.; Collins, M.; Connor, R.; Cronin, A.; Davis, P.; Feltwell, T.; Fraser, A.; Gentles, S.; Goble, A.; Hamlin, N.; Harris, D.; Hidalgo, J.; Hodgson, G.; Holroyd, S.; Hornsby, T.; Howarth, S.; Huckle, E. J.; Hunt, S.; Jagels, K.; James, K.; Jones, L.; Jones, M.; Leather, S.; McDonald, S.; McLean, J.; Mooney, P.; Moule, S.; Mungall, K.; Murphy, L.; Niblett, D.; Odell, C.; Oliver, K.; O'Neil, S.; Pearson, D.; Quail, M. A.; Rabbinowitsch, E.; Rutherford, K.; Rutter, S.; Saunders, D.; Seeger, K.; Sharp, S.; Skelton, J.; Simmonds, M.; Squares, R.; Squares, S.; Stevens, K.; Taylor, K.; Taylor, R. G.; Tivey, A.; Walsh, S.; Warren, T.; Whitehead, S.; Woodward, J.; Volckaert, G.; Aert, R.; Robben, J.; Grymonprez, B.; Weltjens, I.; Vanstreels, E.; Rieger, M.; Schafer, M.; Muller-Auer, S.; Gabel, C.; Fuchs, M.; Dusterhoft, A.; Fritzc, C.; Holzer, E.; Moestl, D.; Hilbert, H.; Borzym, K.; Langer, I.; Beck, A.; Lehrach, H.; Reinhardt, R.; Pohl, T. M.; Eger, P.; Zimmermann, W.; Wedler, H.; Wambutt, R.; Purnelle, B.; Goffeau, A.; Cadieu, E.; Dreano, S.; Gloux, S.; Lelaure, V.; Mottier, S.; Galibert, F.; Aves, S. J.; Xiang, Z.; Hunt, C.; Moore, K.; Hurst, S. M.; Lucas, M.; Rochet, M.; Gaillardin, C.; Tallada, V. A.; Garzon, A.; Thode, G.; Daga, R. R.; Cruzado, L.; Jimenez, J.; Sanchez, M.; del Rey, F.; Benito, J.; Dominguez, A.; Revuelta, J. L.; Moreno, S.; Armstrong, J.; Forsburg, S. L.; Cerutti, L.; Lowe, T.; McCombie, W. R.; Paulsen, I.; Potashkin, J.; Shpakovski, G. V.; Ussery, D.; Barrell, B. G.; Nurse, P. *Nature* **2002**, *415*, 871-880.
67. Theologis, A.; Ecker, J. R.; Palm, C. J.; Federspiel, N. A.; Kaul, S.; White, O.; Alonso, J.; Altafi, H.; Araujo, R.; Bowman, C. L.; Brooks, S. Y.; Buehler, E.; Chan, A.; Chao, Q.; Chen,

- H.; Cheuk, R. F.; Chin, C. W.; Chung, M. K.; Conn, L.; Conway, A. B.; Conway, A. R.; Creasy, T. H.; Dewar, K.; Dunn, P.; Etgu, P.; Feldblyum, T. V.; Feng, J.; Fong, B.; Fujii, C. Y.; Gill, J. E.; Goldsmith, A. D.; Haas, B.; Hansen, N. F.; Hughes, B.; Huizar, L.; Hunter, J. L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C. J.; Koo, H. L.; Kremenetskaia, I.; Kurtz, D. B.; Kwan, A.; Lam, B.; Langin-Hooper, S.; Lee, A.; Lee, J. M.; Lenz, C. A.; Li, J. H.; Li, Y.; Lin, X.; Liu, S. X.; Liu, Z. A.; Luros, J. S.; Maiti, R.; Marziali, A.; Militscher, J.; Miranda, M.; Nguyen, M.; Nierman, W. C.; Osborne, B. I.; Pai, G.; Peterson, J.; Pham, P. K.; Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.; Salzberg, S. L.; Schwartz, J. R.; Shinn, P.; Southwick, A. M.; Sun, H.; Tallon, L. J.; Tambunga, G.; Toriumi, M. J.; Town, C. D.; Utterback, T.; Van Aken, S.; Vaysberg, M.; Vysotskaia, V. S.; Walker, M.; Wu, D.; Yu, G.; Fraser, C. M.; Venter, J. C.; Davis, R. W. *Nature* **2000**, *408*, 816-820.
68. Parani, M.; Parida, A. *Unpublished data*, Accession number AAK06838.
69. Booker, J. P. *Unpublished data*, Accession number Q39366.
70. Veena; Reddy, V. S.; Sopory, S. K. *Plant J.* **1999**, *17*, 385-395.
71. Skipsey, M.; Andrews, C. J.; Townson, J. K.; Jepson, I.; Edwards, R. *Arch. Biochem. Biophys.* **2000**, *374*, 261-268.
72. Usui, Y.; Nakase, M.; Hotta, H.; Urisu, A.; Aoki, N.; Kitajima, K.; Matsuda, T. *J. Biol. Chem.* **2001**, *276*, 11376-11381.
73. Espartero, J.; Sanchez-Aguayo, I.; Pardo, J. M. *Plant. Mol. Biol.* **1995**, *29*, 1223-1233.
74. Johansen, K. S.; Svendsen, I. I.; Rasmussen, S. K. *Plant Science* **2000**, *155*, 11-20.
75. Iozef, R.; Rahlfs, S.; Chang, T.; Schirmer, H.; Becker, K. *FEBS Lett.* **2003**, *554*, 284-288.
76. Castro, N. S.; Bastos, K. P.; Jesuino, R. S. A.; Flleipe, M. S. S.; Pereira, M.; Soares, C. M. A. *Unpublished data*, Accession number AAP03992.
77. Allen, R. E.; Lo, T. W.; Thornalley, P. J. *Eur. J. Biochem.* **1993**, *213*, 1261-1267.
78. Talesa, V.; Uotila, L.; Koivusalo, M.; Principato, G.; Giovannini, E.; Rosi, G. *Biochim. Biophys. Acta.* **1988**, *955*, 103-110.
79. Talesa, V.; Uotila, L.; Koivusalo, M.; Principato, G.; Giovannini, E.; Rosi, G. *Biochim. Biophys. Acta* **1989**, *993*, 7-11.
80. Murata, K.; Saikusa, T.; Watanabe, K.; Fukuda, Y.; Makoto, S.; Kimura, S. *J. Ferment. Technol.* **1986**, *64*, 1-4.
81. Inoue, T.; Kimura, A. *J. Ferment. Bioeng.* **1992**, *73*, 271-276.
82. Norton, S. J.; Principato, G. B.; Talesa, V.; Lupattelli, M.; Rosi, G. *Enzyme* **1989**, *42*, 189-196.
83. Norton, S. J.; Talesa, V.; Yuan, W. J.; Principato, G. B. *Biochem. Int.* **1990**, *22*, 411-418.
84. Talesa, V.; Rosi, G.; Contenti, S.; Mangiabene, C.; Lupattelli, M.; Norton, S. J.; Giovannini, E.; Principato, G. B. *Biochem. Int.* **1990**, *22*, 1115-1120.
85. Sexena, M.; Bisht, R.; Roy, S. D.; Sopory, S. K.; Bhalla-Sarin, N. *Biochem. Biophys. Res. Comm.* **2005**, *336*, 813-819.
86. Talesa, V.; Rosi, G.; Bistoni, F.; Marconi, P.; Norton, S. J.; Principato, G. B. *Biochem. Int.* **1990**, *21*, 397-403.
87. Ridderstrom, M.; Saccucci, F.; Hellman, U.; Bergman, T.; Principato, G.; Mannervik, B. *J. Biol. Chem.* **1996**, *271*, 319-323.
88. Ridderstrom, M.; Mannervik, B. *Biochem J* **1997**, *322*, 449-454.
89. Bito, A.; Haider, M.; Hadler, I.; Breitenbach, M. *J. Biol. Chem.* **1997**, *272*, 21509-21519.
90. Kizil, G.; Wilks, K.; Wells, D.; Ala'Aldeen, D. A. *J. Med. Microbiol.* **2000**, *49*, 669-673.
91. Misra, K.; Bannerjee, A. B.; Ray, S.; Ray, M. *Biochem. J.* **1995**, *305*, 999-1003.

92. Padmanabhan, R. K.; Mukherjee, A.; Singh, S.; Chattopadhyaya, S.; Gowri, V. S.; Myler, P. J.; Srinivasan, N.; Madhubala, R. *Biochem. Biophys. Res. Comm.* **2005**, *337*, 1237-1248.
93. Silva, M. S.; Ferreira, A. E. N.; Tomás, A. M.; Cordeiro, C.; Freire, A. P. *FEBS J.* **2005**, *272*, 2388-2398.
94. Ariza, A.; Vickers, T. J.; Greig, N.; Armour, K. A.; Dixon, M. J.; Eggleston, I. M.; Fairlamb, A. H.; Bond, C., S. *Mol. Microbiol.* **2005**, *59*, 1239-1248.
95. Irsch, T.; Krauth-Siegel, R. L. *J. Biol. Chem.* **2004**, *279*, 22209-22217.
96. Bergmeyer, H. U.; Gawehn, K., Eds. *Methods of Enzymatic Analysis*; Verlag Chemie: Weinheim, 1974; Vol. 1.
97. Mannervick, B.; Aronsson, A.-C.; Tibbelin, G. *Meth. Enzymol.* **1982**, *90*, 535-541.
98. Oray, B.; Norton, S. J. *Meth. Enzymol.* **1982**, *90*, 542-546.
99. Raker, E. *J. Biol. Chem.* **1951**, *190*, 685-696.
100. Oray, B.; Norton, S. J. *Meth. Enzymol.* **1982**, *90*, 547-551.
101. Ridderstrom, M.; Succucci, F.; Hellman, U.; Bergman, T.; Principato, G.; Mannervick, B. *J. Biol. Chem.* **1996**, *271*, 319-323.
102. Bergmeyer, H. U.; Gawehn, K., Eds. *Methods of Enzymatic Analysis*; Verlag Chemie: Weinheim, 1974; Vol. 3.
103. Kieser, T.; Bibb, M. J.; Buttner, M. J.; Chater, K. F.; Hopwood, D. A. *Practical Streptomyces Genetics*; The John Innes Foundation: Crowes, 2000.
104. Park, J. H.; Cha, C. J.; Roe, J. H. *J. Microbiol.* **2006**, *44*, 121-125.
105. Newton, G. L.; Bewley, C. A.; Dwyer, T. J.; Horn, R.; Aharonowitz, Y.; Cohen, G.; Davies, J.; Faulkner, D. J.; Fahey, R. C. *Eur. J. Biochem.* **1995**, *230*, 821-825.
106. Nicholas, G. M.; Eckman, L. L.; Kovác, P.; Otero-Quintero, S.; Bewley, C. A. *Bioorg. Med. Chem.* **2003**, *11*, 2641-2647.
107. Nicholas, G. M.; Kovác, P.; Bewley, C. A. *J. Am. Chem. Soc.* **2002**, *124*, 3492-3493.
108. Irvine, R. F.; Schell, M. J. *Nature Rev. Mol. Cell. Biol.* **2001**, *2*, 327-338.
109. Spies, H. S.; Steenkamp, D. J. *Eur. J. Biochem.* **1994**, *224*, 203-213.
110. Jardine, M. A.; Spies, H. S. C.; Nkambule, C. M.; Gammon, D. W.; Steenkamp, D. J. *Bioorg. Med. Chem.* **2002**, *10*, 875-881.
111. Bornemann, C.; Jardine, M. A.; Spies, H. S. C.; Steenkamp, D. J. *Biochem. J.* **1997**, *325*, 623-629.
112. Lee, S.; Rosazza, J. P. N. *Org. Lett.* **2004**, *6*, 365-368.
113. Patel, M.; Blanchard, J. S. *J. Am. Chem. Soc.* **1998**, *120*, 11538-11539.
114. Unson, M. D.; Newton, G. L.; Davis, C.; Fahey, R. C. *J. Immunol. Meth.* **1998**, *214*, 29-39.
115. Steenkamp, D. J.; Vogt, R. N. *Anal. Biochem.* **2004**, *325*, 21-27.
116. Silverstein, R. W.; Webster, F. X. *Spectrometric Identification of Organic Compounds*; Sixth ed.; John Wiley & Sons, Inc.: New York, 1998.
117. Chenna, R.; Sugawara, H.; Koike, T.; Lopez, R.; Gibson, T. J.; Higgins, D. G.; Thompson, J. D. *Nucleic Acids Res.* **2003**, *31*, 3497-3500.
118. Kamber, B.; Harmann, A.; Eisler, K.; Riniker, B.; Rink, H.; Sieber, P.; Rittel, W. *Helv. Chim. Acta* **1980**, *63*, 899-914.
119. Still, W. C.; Kahn, M.; Mitra, A. *J. Org. Chem.* **1978**, *43*, 2923-2925.
120. Togo, H.; Iida, S. *Synlett* **2006**, *14*, 2159-2175.
121. Kartha, K. P. R.; Field, R. A. *Tetrahedron* **1997**, *53*, 11753-11766.



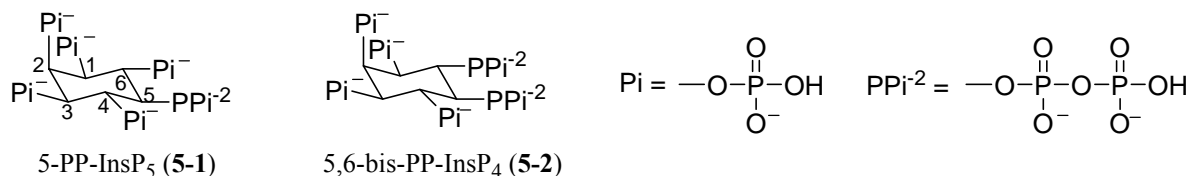
122. Greene, T. W.; Wuts, P. G. M. *Protective Groups in Organic Synthesis*; Third ed.; John Wiley & Sons, Inc.: New York, 1999.

123. Lundbland, R. L.; Noyes, C. M. *Chemical Reagents for Protein Modification*; CRC Press: Boca Raton, 1984.

## Chapter 5: Phosphorylation by Inositol Pyrophosphates

Phosphorylation is a major component of molecular energy production and is essential for a diverse range of processes including the control of cellular pathways, hormone action and signal transduction.<sup>1</sup> Adenosine and guanosine triphosphate (ATP and GTP) are the most commonly used phosphate donors in the cell; however, a number of other biological compounds are capable of phosphate donation. Phosphoenol pyruvate and 1,3-bisphosphoglycerate, for example, are known to be thermodynamically superior phosphate donors to ATP,<sup>1</sup> and pyrophosphate alone has been shown to fulfill the role of ATP in some reactions in certain parasitic protozoa.<sup>2-5</sup> In addition, inorganic polyphosphate has been shown to act as a phosphate donor through the polyphosphate kinase catalyzed conversion of polyphosphate and adenosine diphosphate (ADP) to ATP.<sup>6</sup> The driving force for these reactions is the instability of the phosphoanhydride linkage caused by competing resonance and electrostatic effects, as well as solvation effects.<sup>1</sup>

Recently, 5-diphosphoinositol pentakisphosphate (5-PP-InsP<sub>5</sub>, Figure 5.1, **5-1**) has been shown to phosphorylate proteinaceous serine residues in mammalian and yeast cell extracts.<sup>7</sup> Snyder and coworkers have shown that this novel activity proceeds without enzyme participation and is specific to eukaryotes, with no activity being observed in bacterial extracts. The exact role that this process plays in these organisms is currently unknown, though 5-PP-InsP<sub>5</sub> and bis-diphosphoinositol tetrakisphosphate (bis-PP-InsP<sub>4</sub>, Figure 5.1, **5-2**) have been linked to cellular chemotaxis,<sup>8</sup> apoptosis<sup>9,10</sup> and endocytosis<sup>11</sup> and it has been proposed this non-enzymatic phosphorylation pathway may represent a novel signalling mechanism.<sup>7</sup>



**Figure 5.1:** Representative structures of hexaphosphorylated *myo*-inositol (InsP<sub>6</sub>), 5-diphospho-*myo*-inositol pentakisphosphate (5-PP-InsP<sub>5</sub>), and 5,6-bis-diphospho-*myo*-inositol tetrakisphosphate (5,6-bis-PP-InsP<sub>4</sub>).

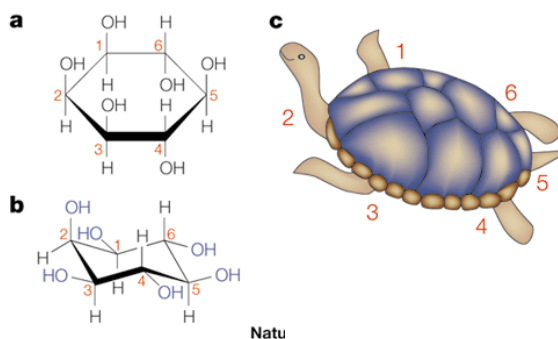
### 5.1. Introduction to Inositol Phosphates and Pyrophosphates

#### 5.1.1. Numbering

Inconsistencies in the literature and changes in nomenclature conventions have led to a large amount of confusion regarding the numbering of the inositol phosphates (InsPs). *Myo*-Inositol itself is a meso compound, with a plane of symmetry running from C<sup>2</sup> to C<sup>5</sup> (see Figure 5.1 for numbering); once substituted outside of this plane, this derivative is chiral and can be classified as D or L depending on the direction in which the ring is numbered. To clarify the numbering of *myo*-inositol derivatives, Agranoff developed a simple mnemonic for numbering *myo*-inositol derivatives: Agranoff likens *myo*-inositol to a turtle with the axial hydroxyl (C<sup>2</sup>) being its head, the opposite hydroxyl (C<sup>5</sup>) its tail, and all other hydroxyl groups are the flippers of the turtle (Figure 5.2). Modern biochemical nomenclature uses the D numbering system, therefore the turtle is “right-flipper” and the “front-right flipper” is labelled C<sup>1</sup> and the ring is labelled counter-clockwise. If the L convention is in use, the

“front-left flipper” of the turtle is labelled C<sup>1</sup> and the ring is numbered clockwise. Biological journals tend to use the D numbering system exclusively, while chemical journals often use the L numbering system. To further complicate matters, the official designation of the D and L numbering systems of the inositol rings were swapped in the 1970s; therefore, D labelled structures in earlier papers are actually L structures using today’s conventions.<sup>12</sup>

In keeping with the current literature describing the inositol pyrophosphates (InsPPs) all structures and diagrams in this work will follow the D numbering system for *myo*-inositol derivatives. Accepted naming conventions use the abbreviation “Ins” for *myo*-inositol and unless otherwise stated, it is assumed that the D numbering convention is used with this abbreviation.<sup>13</sup> For example, D-*myo*-inositol-3-phosphate would be abbreviated as Ins(3)P<sub>1</sub>, while D-*myo*-inositol-3,4-bisphosphate would be abbreviated to Ins(3,4)P<sub>2</sub>. A standard method for the abbreviation of the InsPPs has yet to be adopted. Throughout this chapter, if the position of the pyrophosphate group is known, or relevant, it be included in the abbreviation, for example the structures shown in Figure 5.1 are abbreviated as 5-PP-InsP<sub>5</sub> and 5,6-bis-PP-InsP<sub>4</sub>.



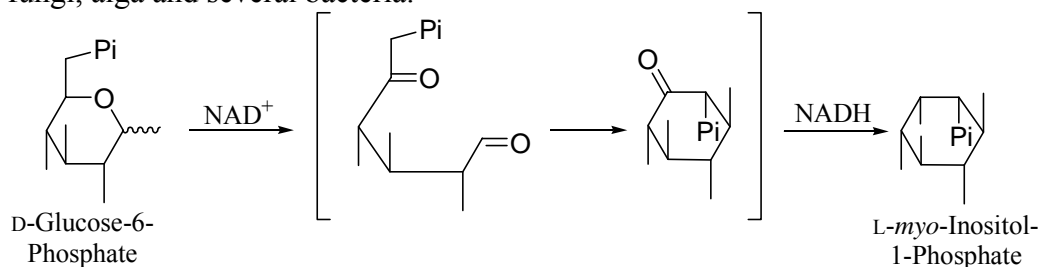
**Figure 5.2:** (a) A Haworth projection of *myo*-inositol; (b) The predicted thermodynamically preferred chair conformation of *myo*-inositol; (c) Agranoff’s turtle, which is used as an aid in determining the numbering of *myo*-inositol and its derivatives. This model follows the D numbering convention, with the head of the turtle corresponding to C<sup>2</sup> and the “right front flipper” of the turtle corresponding to C<sup>1</sup>. This figure is credited to Irvine and Schell.<sup>12</sup>

### 5.1.2. Distribution and Structure

InsPPs have been detected in mammalian cells,<sup>14-16</sup> *Saccharomyces cerevisiae*,<sup>17</sup> *Dictyostelium discoideum*,<sup>18</sup> *Entamoeba histolytica*,<sup>19</sup> *Polysphondylium palladum*<sup>20</sup> and *Phreatamoeba bamuthi*.<sup>21</sup> The structures of the InsPPs vary with species and have not been fully elucidated. In mammals, the predominant PP-InsP<sub>5</sub> has been identified as 5-PP-InsP<sub>5</sub>,<sup>22</sup> however, the identity of the major bis-PP-InsP<sub>4</sub> is unknown.<sup>23</sup> The structural identification of the InsPPs in mammals has been hampered by their low concentration which is in the  $\mu$ M range. In contrast, the concentration of InsPPs in *D. discoideum* is almost 300 times that of mammals<sup>24</sup> (0.05 – 0.25 mM),<sup>18</sup> and the structures of both predominant InsPPs have been identified: 5-PP-InsP<sub>5</sub> and 5,6-bis-PP-InsP<sub>4</sub>, with 10-25% of the pool of PP-InsP<sub>5</sub> existing as 6-PP-InsP<sub>5</sub>.<sup>22,25</sup> In addition, the predominant PP-InsP<sub>5</sub> in *P. palladum* has been identified as 1,5-bis-PP-InsP<sub>4</sub>, or its enantiomer, 3,5-bis-PP-InsP<sub>4</sub>.<sup>20</sup>

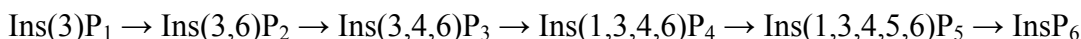
### 5.1.3. Biosynthesis

The biosynthesis of the InsPs begins with the conversion of D-glucose-6-phosphate to Ins(3)P<sub>1</sub>,\* catalyzed by L-*myo*-inositol-1-phosphate synthase,<sup>26,27</sup> which is found in animals, plants, fungi, alga and several bacteria.<sup>27</sup>



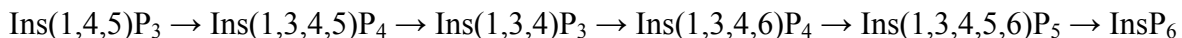
**Figure 5.3:** The biosynthesis of L-*myo*-inositol-1-phosphate (D-*myo*-inositol-3-phosphate) catalyzed by L-*myo*-inositol-1-phosphate synthase. All components are shown using Haworth projections and Pi represents a phosphate group.<sup>28</sup>

Ins(3)P<sub>1</sub> is further phosphorylated to yield a wide variety of InsPs which differ in the number and placement of the phosphate groups.<sup>12</sup> InsP<sub>6</sub> was originally thought to be the end point of InsP biosynthesis, but is now known to be further phosphorylated to the InsPPs. The biosynthetic pathway of InsP<sub>6</sub> has not been elucidated with the biosynthetic intermediates being fully identified in only two organisms: *Spirodela polyrhiza*<sup>29</sup> and *D. discoideum*.<sup>30</sup> The latter is a known producer of InsPPs and its biosynthesis of InsP<sub>6</sub> is as follows:<sup>30</sup>

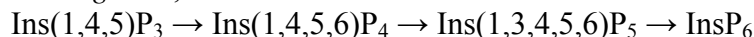


In addition, partial pathways for the biosynthesis of InsP<sub>6</sub> have been identified in humans, *S. cerevisiae*, *Drosophila melanogaster* and *Arabidopsis thaliana*, commencing with Ins(1,4,5)P<sub>3</sub><sup>31</sup>

Human:<sup>31</sup>



*S. cerevisiae*, *D. melanogaster*, and *A. thaliana*:<sup>31</sup>



Once formed, InsP<sub>6</sub> can be further phosphorylated to PP-InsP<sub>5</sub> by the InsP<sub>6</sub> kinases (InsP<sub>6</sub>Ks) using ATP as a source of phosphate. InsP<sub>6</sub>Ks have been identified in mammals,<sup>32-34</sup> *Xenopus* oocytes,<sup>35</sup> and *S. cerevisiae*.<sup>32</sup> Mammals have been shown to produce three unique isoforms of InsP<sub>6</sub>K which vary in cellular localization and expression level.<sup>32,34</sup> InsP<sub>6</sub>K has a high specificity for InsP<sub>6</sub> and does not further phosphorylate PP-InsP<sub>5</sub> to bis-PP-InsP<sub>4</sub>.<sup>32,33,35,36</sup> A separate enzyme, PP-InsP<sub>5</sub> kinase, has been identified which catalyzes the formation of bis-PP-InsP<sub>4</sub> using PP-InsP<sub>5</sub> as a substrate and ATP as a phosphate source.<sup>23</sup>

\* Ins(3)P<sub>1</sub> is equivalent to L-*myo*-inositol-1-phosphate

#### 5.1.4. Biological Role

The InsP<sub>6</sub>Ks and the InsPPs have been implicated in a number of biological processes; however, a definitive role has yet to be determined. The InsP<sub>6</sub>Ks are believed to be involved in the regulation of apoptosis, cell death and cellular growth.<sup>9,10,37</sup> Unfortunately, the exact involvement of the InsPPs in these processes has yet to be determined.

The InsP<sub>6</sub> and PP-InsP<sub>5</sub> kinases have been shown to transfer a phosphate group from their cognate InsPP to ADP, forming ATP.<sup>23,33</sup> *In vitro* studies indicate that the rate of ATP formation by these kinases is similar to the rate of InsPP formation. The *in vivo* direction of the enzymatic reaction will therefore likely be defined by the relative concentrations of ATP and the InsPPs.<sup>33</sup> It has been suggested that the ability of the InsPPs to donate a phosphate group to ATP may indicate that they can act as energy sources for specific cellular functions.<sup>23</sup>

Recently the ability of 5-PP-InsP<sub>5</sub> to phosphorylate proteinaceous serine residues in eukaryotic cell extracts has been noted.<sup>7</sup> Preliminary studies of the hydrolysis of the pyrophosphate moieties of the InsPPs have indicated that these compounds may have comparable or even greater energies of hydrolysis than ATP.<sup>18,38</sup> To date, the exact role this phosphorylation plays in cellular physiology is unknown but it has been hypothesized that it may represent a novel signalling mechanism in these organisms.<sup>7</sup>

## 5.2. Objectives

The InsPPs have been linked to a number of important biological functions; however, the exact mechanism of action is unknown. The ability of the InsPPs to non-enzymatically phosphorylate proteinaceous serine residues in eukaryotes may be crucial to their cellular role. To better understand the driving forces behind these reactions, we performed a detailed theoretical study of the phosphorylation of methanol, as a model of serine, by the InsPPs. Our goal was to determine the effect of InsPP structure and overall charge on the thermodynamics of these reactions. This information should contribute to our understanding of this novel cellular phosphorylation process.

## 5.3. Computational Studies of Phosphate Transfer

Phosphate transfer is an essential process in biological systems and hence, many studies have been performed on these reactions to further understand the behaviour of the compounds involved and the mechanisms of these reactions.

### 5.3.1. The Modelling of Pyrophosphate

Pyrophosphate has been thoroughly studied on its own and as a model compound for ATP. These investigations have focused on the conformational properties, ionization states, solvation, and hydrolysis of pyrophosphate, as well as the effects of complexed metal ions on pyrophosphate behaviour.<sup>39-51</sup> These studies have been performed using Hartree-Fock (HF) and Density Functional Theory (DFT) techniques in both the gas and aqueous phases.<sup>39,42-45,51</sup> In addition, methyl pyrophosphate and its reactions have been studied to gain insight into phosphorylation.<sup>44,52</sup> A major finding of these studies, through the comparison of theoretical and experimental results, is that the accurate modelling of phosphate transfer requires the consideration of solvation, as gas phase approaches do not

produce results in agreement with experimental data.<sup>39,42,46,49,53</sup> Additionally the solvation energies of the substrates and products contribute substantially to the overall thermodynamics of these reactions.<sup>46,48-50</sup>

At biologically relevant pH, phosphates are highly charged species with a large amount of intramolecular electrostatic repulsion. In the aqueous phase, water molecules provide hydrogen bonding and electrostatic shielding, minimizing the electrostatic repulsions felt by the phosphate and pyrophosphate groups. These interactions and shielding are not present in the gas phase, which can exaggerate the contribution of intramolecular interactions to the calculated thermodynamic stability of these compounds.<sup>54</sup> The thorough study of pyrophosphates and phosphate donation is further complicated by the role that counter-ions are known to play in the stability of these compounds.<sup>46,48-52</sup> The accurate placement and coordination of these ions adds another layer of complexity to the design of theoretical studies of these systems.

### 5.3.2. Modelling of the Inositol Phosphates

To date, extensive modelling studies of the InsPPs have not been conducted. Preliminary investigations into the bond energies of 1, 4-bis-PP-InsP<sub>4</sub><sup>†</sup> using molecular mechanics indicate that these compounds may have a comparable or even greater energy of hydrolysis than ADP.<sup>18</sup> Given the importance that serine phosphorylation by the InsPPs may play in cellular physiology, a more detailed study of the thermodynamics of these reactions appeared warranted.

## 5.4. Computational Chemistry

There are multiple methods available to determine the thermodynamic properties of a system. The higher level HF and DFT calculations were initially used for our studies of the InsPPs, followed by lower level molecular mechanics methods; however, in all cases, these techniques proved unsatisfactory. An introduction to HF and DFT theory can be found in *Chapter 2*, while an introduction to molecular mechanics can be found in *Chapter 3*. The semi-empirical methods were found to be better suited to this study and were used for all subsequent calculations.

### 5.4.1. Semi-Empirical Theory

Semi-empirical calculations were developed to model systems where the time and computational demands of an HF or DFT calculation would be prohibitive, but a level of theory higher than that of molecular mechanics was desired. One of the most time-consuming aspects of an HF calculation is the determination and manipulation of the integrals involved. An obvious route to reduce the time required to complete these calculations would be to omit or approximate some of these integrals. Semi-empirical techniques do this by explicitly including only the valence electrons in the calculation, while grouping the core electrons with the nucleus, as the core electrons have little effect on chemical bonding and other thermodynamic phenomena.<sup>55</sup> The approximations made during semi-empirical calculations can lead to artificially large enthalpies; however, the relative magnitudes of the thermodynamic values are often realistic<sup>56</sup> and these elevated effects often cancel in the determination of overall reaction enthalpies.

---

<sup>†</sup> 1,4-bis-PP-InsP<sub>4</sub> or its enantiomer 3,6-bis-PP-InsP<sub>4</sub>

A semi-empirical calculation takes the same general form as an HF calculation in that they both have a Hamiltonian and a wavefunction;<sup>57</sup> however, they are parameterized to reduce the amount of error inherent to the approximation of the integrals. These parameters are obtained by fitting the equations to experimental data or the results of *ab initio* calculations.<sup>57</sup> Semi-empirical methods are therefore a compromise between the speed of molecular mechanics and the accuracy of HF techniques. Due to the inclusion of parameterization, the accuracy of a semi-empirical calculation is dependent upon the parameters in use; however, semi-empirical calculations are not as sensitive as molecular mechanics methods to parameterization and have been successful in the description of organic chemistry.<sup>57</sup> If the molecule to be studied is similar to the compounds used to develop the parameter set, the results should be reasonable.<sup>57</sup> A number of semi-empirical methods have been developed with the most popular methods in recent literature being AM1 and PM3, which differ in their parameterization.<sup>58</sup> The parameters of AM1 were developed using chemical knowledge and intuition and those of PM3 were derived from an automated parameterization procedure.<sup>58</sup> While some of the parameters of AM1 and PM3 differ greatly, they predict various thermodynamic and structural properties to approximately the same level of accuracy.<sup>58</sup> The results obtained do not require zero-point energy correction as it is implicit in the parameterization.<sup>57</sup>

#### 5.4.2. Plan of Action

The ability of the InsPPs to phosphorylate proteinaceous serine residues may be critical to cellular biochemistry. To compare the thermodynamics of phosphate donation by the InsPPs to other model compounds, we performed geometry optimizations on pyrophosphate and a number of pyrophosphate derivatives, as well as ATP, triphosphate and methyl triphosphate to act as benchmarks for our study. Geometry optimizations were performed on a variety of protonation states of InsP<sub>6</sub>, 5-PP-InsP<sub>5</sub> and 5,6-PP-InsP<sub>4</sub>, and the  $\Delta H_f$  values for each structure were determined. The overall thermodynamics ( $\Delta H_{rxn}$ ) of phosphorylation of methanol by the various phosphate group donors was calculated to determine what trends, if any, exist, and if charge and InsPP structure have an effect on the thermodynamics of phosphorylation by the InsPPs.

### 5.5. Materials and Methods

#### 5.5.1. Initial Hartree-Fock, Density Functional Theory and Molecular Mechanics Geometry Optimizations

All structures were drawn and visualized in Maestro (Schrödinger Inc., Portland, OR) and all calculations were performed on an SGI O<sub>2</sub> workstation (Violin) or the University of Waterloo's multi-CPU Origin 3800 system (Flexor). For complete details of these systems, please see the *Materials and Methods* section of *Chapter 2*.

All HF and DFT geometry optimizations and vibrational frequency calculations were performed using the 6-31++G\*\* basis set with and without implicit solvation, both as implemented by Jaguar version 6.0 (Schrödinger Inc.). All DFT calculations used the B3LYP density functional, also as implemented by Jaguar. If convergence was a problem, a similar tautomer whose geometry had been successfully optimized, was modified and used as a starting structure. If a negative frequency was returned the conformation is not a minimum energy point on the potential energy surface. The structure was then modified through bond

and/or torsion angle rotation to give a new starting structure or a similar optimized tautomer was modified and used as a new starting structure.

All molecular mechanics calculations were performed using the OPLS2003 force field and the GB/SA solvation model as implemented by MacroModel (Schrödinger Inc.).

### 5.5.2. *Semi-Empirical Geometry Optimizations*

All final geometry optimizations and vibrational frequency calculations were performed using the PM3 semi-empirical method<sup>59,60</sup> with SM5.2 implicit water solvation as implemented by AMPAC 8 and the AMSOL module<sup>61</sup> (SemicheM Inc. Kansas City, KS). All structures were drawn and visualized using the AMPAC Gui 8. All optimizations minimized the heat of formation of the molecule using the TRUST model<sup>62</sup> and a restricted HF wavefunction. Only those structures that returned non-negative frequencies were used for further analysis. The starting structures for these calculations were drawn in the AMPAC Gui or were modified from previously optimized structures with a similar charge. If a negative frequency was returned, the bond and/or torsional angles of the structure were modified to give a new starting structure. If a negative frequency was returned despite these modifications, a minimum energy structure from a similar tautomer was modified and used as the new starting structure and the result was checked for negative frequencies.

When a negative frequency could still not be obtained the structure's geometry was optimized using molecular mechanics methods with the OPLS2003 force field and the GB/SA water solvation model as implemented by MacroModel. All structures optimized by MacroModel were drawn and visualized using Maestro. The resultant structure was then imported into the AMPAC Gui as a pdb file and its geometry optimized as detailed above.

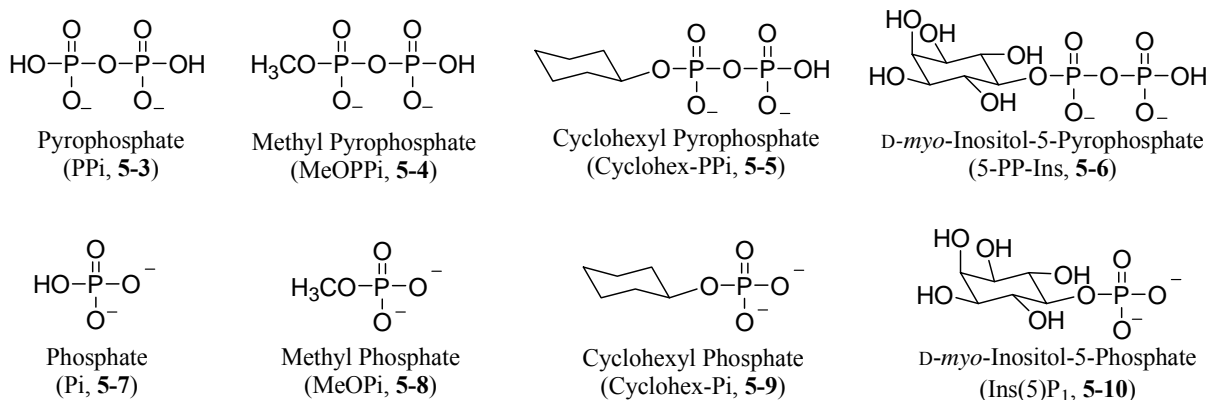
All AMPAC calculations were performed on an IBM Z Pro Intel Xeon two processor work station. All Maestro and MacroModel operations were run on either Violin or Flexor.

### 5.5.3. *Geometry Optimizations of Model Phosphate Donors*

Geometry optimizations were performed on the model compounds pyrophosphate (PPi), methyl pyrophosphate (MeOPPi), cyclohexyl pyrophosphate (cyclohex-PPi) and D-*myo*-inositol-5-pyrophosphate (5-PP-Ins, Figure 5.4, (**5-3** to **5-6**)). The pKa values of PPi are known (Table 5.1) and it is expected to have a charge of -2 or -3 at physiological pH. The pKa values of the other model PPi compounds are unknown but are expected to be of a similar charged state under physiological conditions. The model PPi derivatives were therefore modelled with charges of -2 and -3. The starting conformations of cyclohex-PPi and 5-PP-Ins were such that the pyrophosphate group was in the equatorial position as shown in Figure 5.4.

Geometry optimizations were also performed on the expected products of phosphate donation: phosphate (Pi), methyl phosphate (MeOPi), cyclohexyl phosphate (cylcohex-Pi), and Ins(5)P<sub>1</sub> (Figure 5.4, (**5-7** to **5-10**)). The pKa values of Pi and MeOPi are known and these compounds are expected to have a charge of -1 or -2 at physiological pH. The pKa values of cyclohex-Pi and Ins(5)P<sub>1</sub> are unknown, therefore based upon the Pi and MeOPi values, all of the model monophosphates were modelled with charges of -2 and -3 (Table 5.1).





**Figure 5.4:** Compounds used to model phosphate transfer.

**Table 5.1:** The pKa values of pyrophosphate and its cleavage products phosphate and methyl phosphate<sup>1,63,64</sup>

Structure	pKa <sub>1</sub>	pKa <sub>2</sub>	pKa <sub>3</sub>	pKa <sub>4</sub>	Expected Charge <sup>a</sup>
Pyrophosphate	0.85	1.96	6.68	9.39	-2 / -3
Methyl phosphate	1.1	6.36	--	--	-1 / -2
Phosphate	2.15	6.82	12.38		-1 / -2

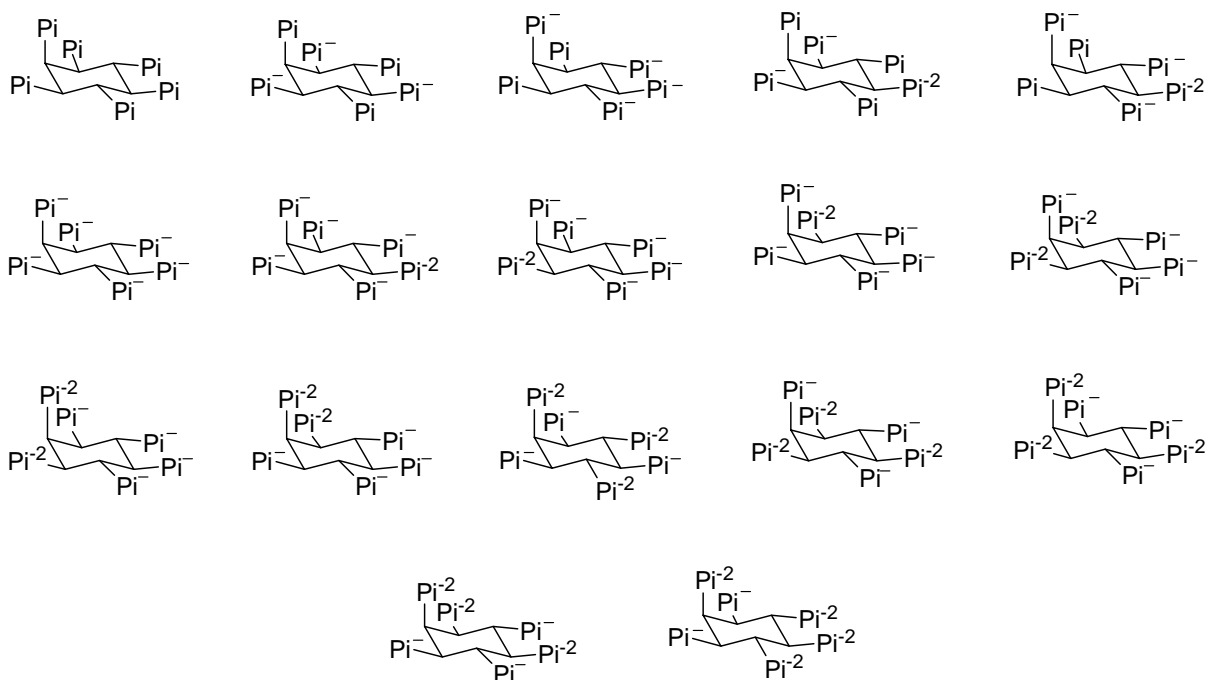
<sup>a</sup>at physiological pH

#### 5.5.4. Geometry Optimizations of the Inositol Pyrophosphates

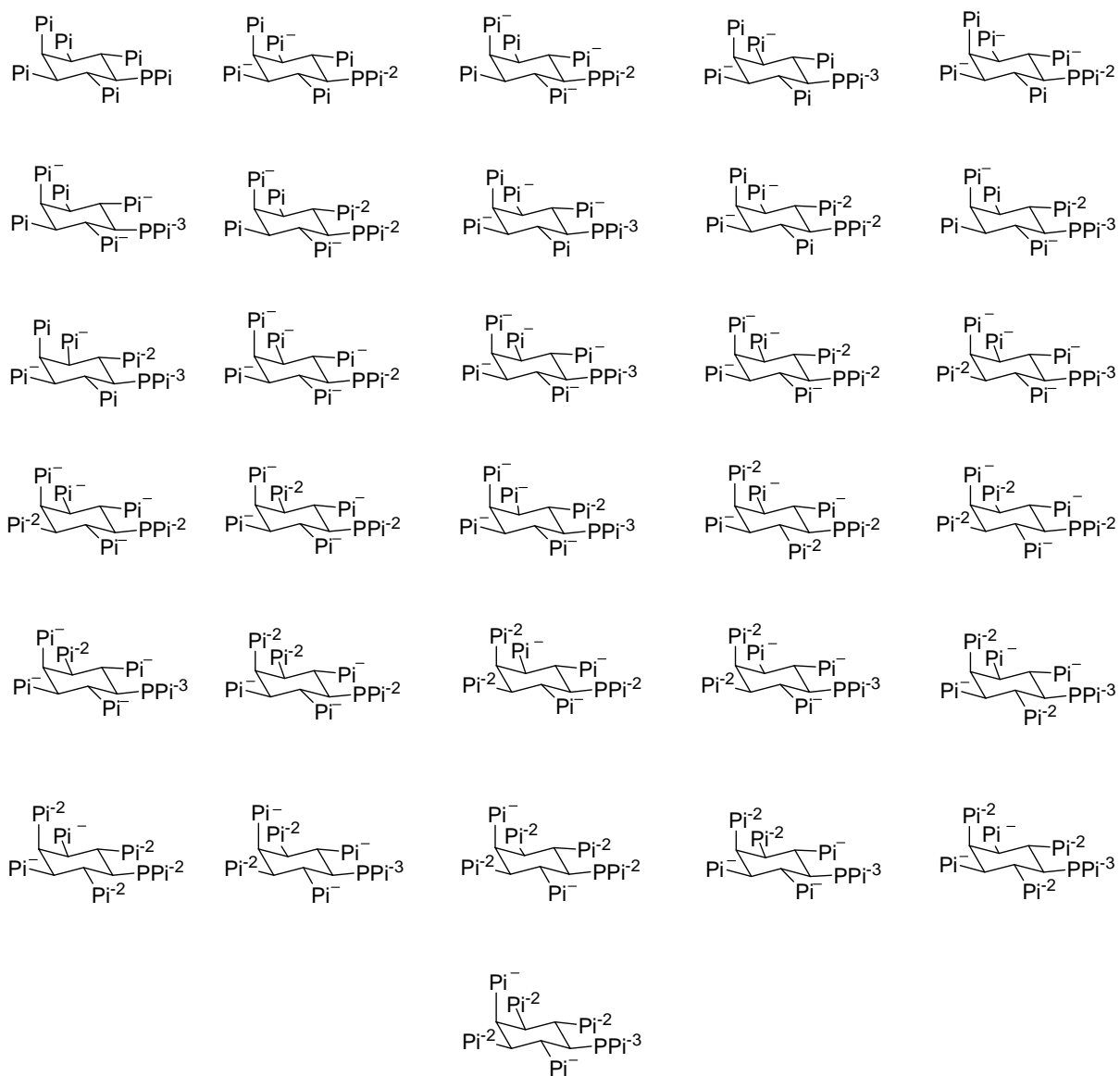
InsP<sub>6</sub>, 5-PP-InsP<sub>5</sub> and 5,6-bis-PP-InsP<sub>4</sub> were optimized with the 5 equatorial/ 1 axial (5 eq/1 ax) starting conformation as shown in Figure 5.5, Figure 5.6 and Figure 5.7. The pKa values of the InsPPs are unknown; therefore, the known pKa values of InsP<sub>6</sub> (Table 5.2) were used to estimate the protonation states of the 5-PP-InsP<sub>5</sub> and 5,6-bis-PP-InsP<sub>4</sub> at physiological pH. Additional charged states not expected to be present at physiological pH were also modelled to determine the effects of increased charge on the *myo*-inositol ring on the thermodynamics of phosphate donation.

**Table 5.2:** The pKa values known for InsP<sub>6</sub><sup>65</sup>

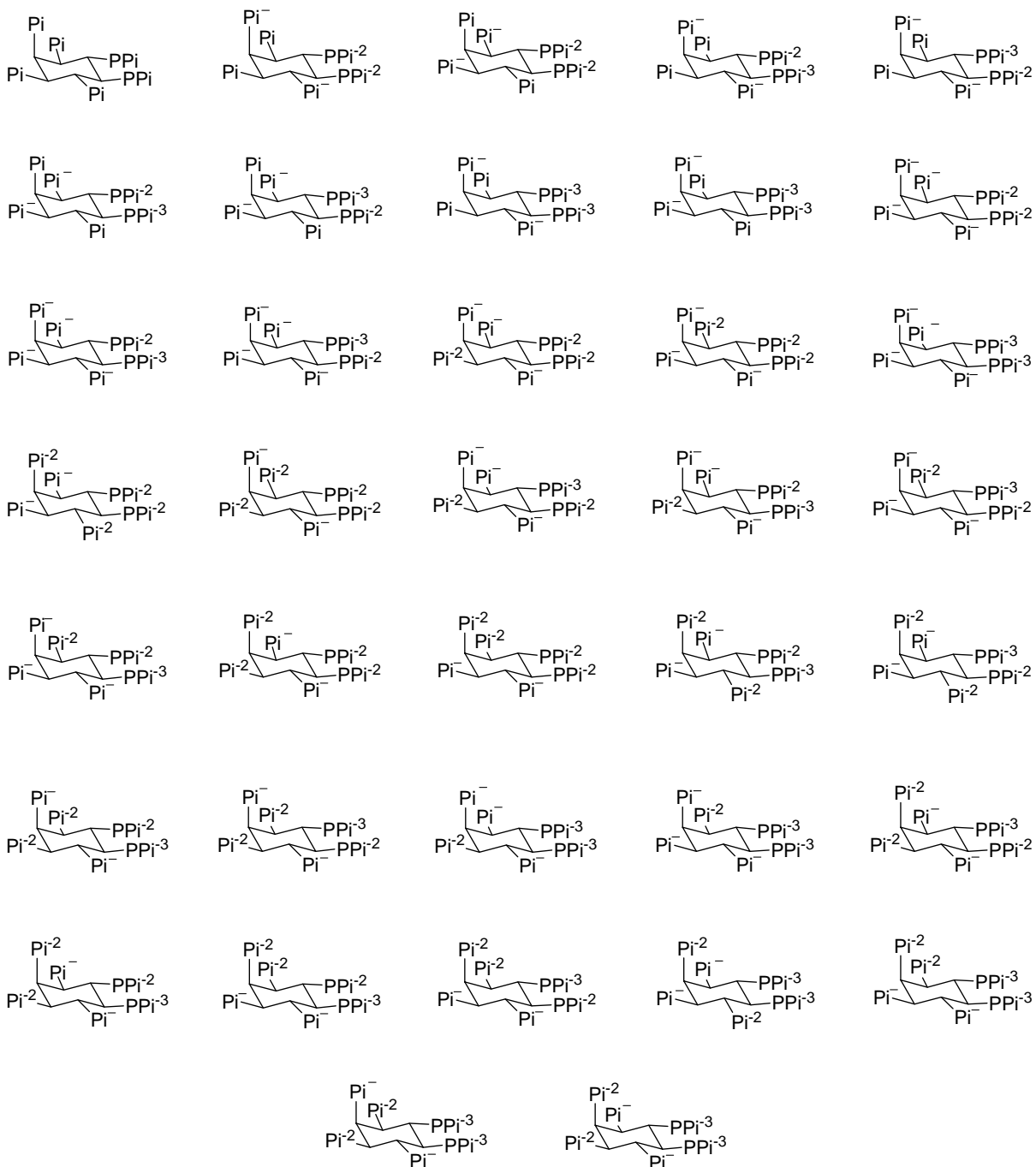
pKa	Carbon Number			
	C <sup>1</sup> /C <sup>3</sup>	C <sup>2</sup>	C <sup>4</sup> /C <sup>6</sup>	C <sup>5</sup>
pKa <sub>1</sub>	1.5	1.1	2.1	1.7
pKa <sub>2</sub>	5.7/12.0	6.9	10.0	7.6



**Figure 5.5:** InsP<sub>6</sub> ionization states modelled in this study; where Pi represents a phosphate group.



**Figure 5.6:** 5-PP-InsP<sub>5</sub> ionization states modelled in this study; where Pi represents a phosphate group and PPi represents a pyrophosphate group.

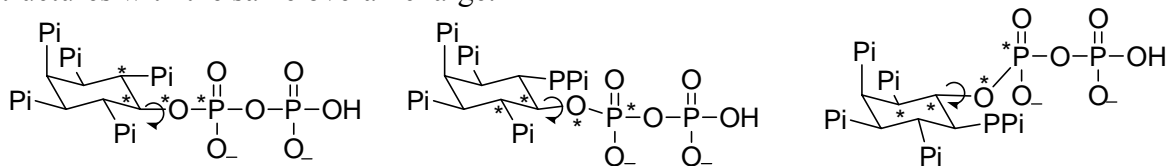


**Figure 5.7:** 5,6-Bis-PP-InsP<sub>4</sub> ionization states modelled in this study; where Pi represents a phosphate group and PPI represents a pyrophosphate group.

### 5.5.5. Dihedral Angle Driving

In some cases, the calculated  $\Delta H_f$  values of the optimized geometries were not in agreement with other structures with the same overall charge. These structures were subjected to dihedral angle driving ( $0^\circ$  to  $360^\circ$  in  $10^\circ$  increments) of the pyrophosphate group to overcome any energy barriers encountered during the initial minimization (see Figure 5.8 for angles) in hopes of finding the global minimum or at the very least, a lower energy local

minimum structure. The lowest energy structure was then minimized as detailed above. In the case of 5,6-bis-PP-InsP<sub>4</sub>, the dihedral angle for the pyrophosphate group at position 5 on the inositol ring was first scanned, followed by the pyrophosphate group at position 6. In most cases, these dihedral angle drives resulted in structures in better agreement with other structures with the same overall charge.



**Figure 5.8:** During dihedral angle scans, the angles marked by asterisks were altered systematically if required.

### 5.5.6. Geometry Optimizations of Other Important Compounds

Methanol was optimized in its neutral state as a model compound of serine. ATP, triphosphate, and methyl triphosphate were modelled with the protonation state expected at physiological pH and therefore had overall charges of -2 and -3. ADP was modelled with charges of -2 and -3 and AMP was modelled with overall charges of -1 and -2.

### 5.5.7. Calculation of the Thermodynamics of Phosphate Group Transfer

The geometry optimizations of the various compounds studied returned  $\Delta H_f$  values which were used to calculate the overall  $\Delta H_{rxn}$  values for the phosphorylation of methanol to form methyl phosphate.

## 5.6. Results and Discussion

The newly discovered ability of 5-PP-InsP<sub>5</sub> to phosphorylate proteinaceous serine residues in eukaryotes has been hypothesized to be a novel signalling mechanism and may be critical for biological function. Preliminary modelling studies of the InsPPs suggest that the phosphate donating ability of these compounds may be comparable to that of ATP. Our theoretical study into the thermodynamics of these reactions gives insight into the effect of InsPP structure and overall charge on the ability of the InsPPs to donate a phosphate group and compares this ability to that of pyrophosphate, its derivatives and ATP.

### 5.6.1. Selection of Modelling Technique

Initial attempts to model 5-PP-InsP<sub>5</sub> and 5,6-bis-PP-InsP<sub>4</sub> using HF and DFT techniques *in vacuo* and with implicit solvation did not yield satisfactory results. These calculations were very lengthy, even when parallelized;<sup>‡</sup> and the large number of negative charges present at physiological pH in 5,6-bis-PP-InsP<sub>4</sub>, combined with the inherent flexibility of the molecule, resulted in the inability of the geometry optimization algorithm to converge to a minimum energy structure for many tautomers. In addition, the preliminary calculations performed using implicit water solvation gave results which were not in agreement with experimental data (Table 5.3). Phosphorylation by pyrophosphate is known to be a favourable process; however, the results shown in Table 5.3 indicate that phosphorylation of methanol is unfavourable. The inability of these methods to return results

<sup>‡</sup>There were as many as 816 basis functions for InsP<sub>6</sub>, 902 for 5-PP-InsP<sub>5</sub> and 940 for 5,6-bis-PP-InsP<sub>4</sub> to be evaluated, depending on the protonation state of the molecule.

in agreement with experimental data, combined with the time constraints and the difficulty in finding a minimum energy structure for certain 5,6-bis-PP-InsP<sub>4</sub> charged states, resulted in the abandonment of this method as a viable option for the modelling of phosphorylation by the InsPPs.

**Table 5.3:** Calculated  $\Delta H_{\text{rxn}}$  values for phosphorylation of methanol by pyrophosphate.

<i>Pi Donor</i> <sup>a</sup>	→	<i>Product</i>	Calculated $\Delta H_{\text{rxn}}$ Values (kcal/mol)			
			<i>HF</i> <sup>b</sup>		<i>DFT</i> <sup>c</sup>	
			<i>Gas</i>	<i>Water</i>	<i>Gas</i>	<i>Water</i>
<b>Forming MeOPi(-1)<sup>d</sup></b>						
PPi(-2) <sup>e</sup>	→	Pi(-1)	-55.2	41.4	-57.7	59.7
PPi(-3)	→	Pi(-2)	-121.11	114.2	-122.4	113.0
<b>Forming MeOPi(-2)</b>						
PPi(-3)	→	Pi(-1)	-124.9	116.1	-122.4	119.2

<sup>a</sup>Pi = phosphate; <sup>b</sup>HF = Hartree-Fock methods; <sup>c</sup>DFT = density functional theory methods, <sup>d</sup>MeOPi = methyl phosphate; <sup>e</sup>PPi = pyrophosphate.

As an alternative to the HF and DFT methods, the InsPPs were subsequently modelled using molecular mechanics methods with the OPLS2003 force field. OPLS2003 is an enhanced version of the OPLS-AA force field, which, though traditionally used to model amino acids, has been parameterized<sup>66</sup> and successfully used for modelling carbohydrates.<sup>67-70</sup> Previous studies had shown that AMBER\*, another well-known force field used to model carbohydrates, could not predict the 5 eq / 1 ax conformation of InsP<sub>6</sub>, and hence was not used in this study.<sup>71</sup> The OPLS2003 force field successfully predicted the 5 eq/1 ax conformation of InsP<sub>6</sub>; however, the magnitude of the  $\Delta H_{\text{rxn}}$  values determined for the phosphorylation of methanol by pyrophosphate called these results into question (Table 5.4, see *Appendix 2* for detailed results).

**Table 5.4:** Calculated  $\Delta H_{\text{rxn}}$  values for the phosphorylation of methanol by pyrophosphate using molecular mechanics methods.

<i>Pi Donor</i> <sup>a</sup>	→	<i>Product</i>	Calculated $\Delta H_{\text{rxn}}$ Values (kcal/mol)	
			<i>Gas</i>	<i>Water</i>
<b>Forming MeOPi(-1)<sup>b</sup></b>				
PPi(-2) <sup>c</sup>	→	Pi(-1)	-262.4	-193.0
PPi(-3)	→	Pi(-2)	-330.8	-210.6
<b>Forming MeOPi(-2)</b>				
PPi(-3)	→	Pi(-1)	-315.2	-200.7

<sup>a</sup>Pi = phosphate; <sup>b</sup>MeOPi = methyl phosphate; <sup>c</sup>PPi = pyrophosphate

As noted in *Chapter 3*, MacroModel returns information regarding the quality of the bend, stretch and torsion parameters of the force field used to model a particular structure. Of the force fields available, OPLS2003 seemed best suited for the study of the InsPPs: MacroModel returned only one type of low parameter the O-P-O-P torsional angle, and personal communications with Schrödinger Inc. indicated that despite the rating of this parameter, its value was reasonable. The lower the quality of the parameter in use, the less experimental or *ab initio* data that has been utilized in fitting the force field equation, resulting in a generalized parameter, which can lead to inaccurate conformational energy differences and geometries.<sup>72</sup> Despite the assurances of Schrödinger Inc., concern over the accuracy of these results caused this method to be rejected as well.

Finally, semi-empirical methods were attempted as they can provide greater accuracy over the molecular mechanics methods and greater speed and fewer computational demands than the HF and DFT methods. The PM3 semi-empirical method was chosen as it has been used extensively to successfully model phosphorous<sup>41,73</sup> and phosphate groups.<sup>40,74,75</sup> In addition, the PM3 method accurately predicted the conformation of InsP<sub>6</sub>, while the AM1 method did not.<sup>71</sup> All further discussion of phosphate donation by the InsPPs focuses on the results obtained using semi-empirical methods.

### 5.6.2. Exclusion of Counter-Ions

Counter-ions are known to affect the thermodynamics of phosphate group transfer from pyrophosphates and their derivatives.<sup>46,48-52</sup> In addition, the transfer of a phosphate group from 5-PP-InsP<sub>5</sub> to proteinaceous serine residues in eukaryotic cellular extracts is known to require the presence of magnesium (II) ions, though the exact role these ions play is unknown. Studies of InsP<sub>6</sub> indicate that in the cytosol, InsP<sub>6</sub> likely exists in complex with five magnesium (II) ions and 22 water molecules; however, the exact location and coordination pattern of the magnesium (II) ions is unknown. A search of the Cambridge Crystallographic Database for crystal structures of InsP<sub>6</sub> in coordination with biologically relevant ions, such as sodium (I) and magnesium (II), returned structures involving sodium (I) ions and hydrogen bonded water molecules.<sup>§76-79</sup> The use of these structures as a starting point in an implicitly solvated study would likely require the removal of the explicitly defined water molecules, invalidating the use of the structure as a starting point. The water molecules contribute hydrogen bonding and possibly electrostatic interactions integral to the conformation of the molecule. In addition, there is no data available regarding the coordination of counter-ions and water molecules by the InsPPs; therefore, any inclusion of counter-ions or water molecules in our calculations would be arbitrary with a large amount of uncertainty in their placement. This uncertainty would add a prohibitive layer of complexity to this preliminary study and therefore all counter-ions and water molecules were omitted from our geometry optimization.

While magnesium (II) is required for phosphorylation of proteinaceous serine residues by 5-PP-InsP<sub>5</sub> in eukaryotic cell extracts, it is not anticipated that the omission of magnesium (II) ions from these calculations would dramatically alter the *relative*  $\Delta H_{\text{rxn}}$  values for phosphate transfer as experimental studies of ATP hydrolysis in the presence of magnesium (II) ions have indicated that the  $\Delta G^{\circ}_{\text{obs}}$  of hydrolysis ranges from ~31 to ~35 kJ/mol, a change of only 1 kcal/mol, depending on the concentration of magnesium (II)<sup>80</sup> and therefore any trends that we observe in our data should be valid for this preliminary investigation.

### 5.6.3. Use of Methanol as a Model for Serine

The phosphorylation of eukaryotic proteins by 5-PP-InsP<sub>5</sub> is specific to serine residues, with no phosphorylation of threonine observed. Serine is a primary alcohol therefore methanol was used as a model. Serine itself was not used as the inclusion of  $\alpha$ -amino and carboxyl groups would add extra electrostatic interactions which may affect the overall thermodynamics of phosphorylation.

---

<sup>§</sup> A search conducted November 4, 2006 returned five structures for InsP<sub>6</sub> with CCDC Deposition Codes NAMIHP01, NAMIHP02, NAMIHP10, QQQCQV and QQQCQS.

#### 5.6.4. Geometry Optimizations of Inositol Pyrophosphates using Semi-Empirical Techniques

The non-enzymatic phosphorylation of serine residues in eukaryotic cells specifically required 5-PP-InsP<sub>5</sub> as a phosphate donor;<sup>7</sup> therefore, this regioisomer was chosen for these studies. The phosphorylation of protein residues by bis-PP-InsP<sub>4</sub> has not been observed; however, it is expected that the increased electrostatic and steric repulsion caused by the additional pyrophosphate group would make phosphate donation by bis-PP-InsP<sub>4</sub> even more thermodynamically favourable. Because the structure of bis-PP-InsP<sub>4</sub> in mammals and *S. cerevisiae* is unknown, 5,6-bis-PP-InsP<sub>4</sub>, the predominant bis-PP-InsP<sub>4</sub> in the eukaryote *D. discoideum*,<sup>22,81</sup> was used in this investigation.

To date, the conformational preferences of the InsPPs have not been determined. It is likely that, similar to InsP<sub>6</sub>,<sup>71,82,83</sup> the most stable conformation of the InsPPs at physiological pH would have the maximum number of equatorial substituents, with the pyrophosphate groups in particular being equatorial. 5-PP-InsP<sub>5</sub> and 5,6-PP-InsP<sub>4</sub> were therefore modelled using a 5 eq / 1 ax starting conformation.

When the calculated  $\Delta H_f$  values for optimized InsPP structures with identical charges were compared, it was apparent that certain structures had a  $\Delta H_f$  value unexpectedly higher than other conformations of the same charge. During geometry optimization, energy barriers can be encountered which funnel geometry optimizations towards local minima of higher energy (see *Chapter 2* for an introduction to geometry optimization). To overcome these potential energy barriers, a dihedral angle drive of the pyrophosphate groups was performed on these anomalous structures and the geometries of the lowest energy conformations optimized. In almost all cases, these drives resulted in lower energy conformations whose  $\Delta H_f$  values were in better agreement with other tautomers of similar charge. The average change in the  $\Delta H_f$  values of these structures was 7.7 kcal/mol, with the greatest change being 18.1 kcal/mol and the lowest being 0.1 kcal/mol. The best estimate in the error present in our calculated  $\Delta H_f$  values would therefore be  $\pm 7.7$  kcal/mol. All final energy values can be found in *Appendix 2*.

#### 5.6.5. Phosphorylation of Methanol by Model Compounds

Initial calculations focused on the phosphorylation of methanol by the model compounds PPi, MeOPPi, Cyclohex-PPi and 5-PP-Ins (**5-3** to **5-6**) to serve as comparison points and to indicate any trends or factors which may be important for the accurate modeling of phosphorylation by the InsPPs. 5-PP-Ins was used as a model for the *myo*-inositol framework, omitting the steric and electrostatic contributions of the multiple phosphate moieties to the overall thermodynamics of phosphorylation. The  $\Delta H_{rxn}$  values determined for these preliminary calculations (Table 5.5) indicate that the thermodynamics of phosphorylation may be affected by the overall charge of the molecules involved. Phosphorylation is predicted to be thermodynamically more favourable when the starting structure has a higher charge. For example, phosphorylation of methanol by PPi(-3) to form MeOPi(-1) was calculated to be more favourable than phosphorylation by PPi(-2). In addition, the formation of MeOPi(-2) was predicted to be favoured over the formation of MeOPi(-1), when the same starting ionization states are involved. These results indicate that the overall thermodynamics of phosphorylation may be affected by the ionization states of both the reactants and products of these reactions.



**Table 5.5:**  $\Delta H_{rxn}$  values determined for the phosphorylation of methanol to form methyl phosphate.

<i>Pi Donor</i> <sup>a</sup> <i>Forming MeOPi(-1)</i> <sup>b</sup>	→	<i>Product</i>	$\Delta H_{rxn}$ <sup>c</sup>	<i>Pi Donor</i>	→	<i>Product</i> <i>Forming MeOPi(-2)</i>	$\Delta H_{rxn}$
PPi(-2) <sup>d</sup>	→	Pi(-1)	-8.9	PPi(-3)	→	Pi(-1)	-13.9
PPi(-3)	→	Pi(-2)	-12.9				
MeOPPi(-2) <sup>e</sup>	→	MeOPi(-1)	-7.2	MeOPi(-3)	→	MeOPi(-1)	-15.2
Cyclohex-PPi(-2) <sup>f</sup>	→	Cyclohex-Pi(-1) <sup>g</sup>	-7.3	Cyclohex-PPi(-2)	→	Cyclohex-Pi(-1)	-14.8
Cyclohex-PPi(-3)	→	Cyclohex-Pi(-2)	-14.3				
InsPPi(-2) <sup>h</sup>	→	InsPi(-1) <sup>i</sup>	-2.0	InsPPi(-3)	→	InsPi(-2)	-6.2
InsPPi(-3)	→	InsPi(-2)	-14.8				

<sup>a</sup>Pi = phosphate; <sup>b</sup>MeOPi = methyl phosphate; <sup>c</sup> $\Delta H_f$  in kcal/mol; <sup>d</sup>PPi = pyrophosphate; <sup>e</sup>MeOPPi = methyl pyrophosphate; <sup>f</sup>Cyclohex-PPi = cyclohexyl pyrophosphate; <sup>g</sup>Cyclohex-Pi = cyclohexyl phosphate; <sup>h</sup>InsPPi = *myo*-inositol-5-pyrophosphate; <sup>i</sup>InsPi = *myo*-inositol-5-phosphate (Ins(5)P<sub>1</sub>).

### 5.6.6. Phosphorylation of Methanol by the Inositol Pyrophosphates

The effects of reactant and product ionization state on the thermodynamics of phosphorylation by pyrophosphate and the model compounds discussed above indicate that protonation state may also play a role in the thermodynamics of phosphorylation of serine by the InsPPs. A variety of ionization states were therefore modelled, in hopes of obtaining a complete picture of the factors affecting the thermodynamics of phosphorylation. Because the pKa values of the InsPPs are unknown, the pKa values of InsP<sub>6</sub> were used to estimate the protonation states of the InsPPs likely to be present at physiological pH.

Selected calculations producing protonation states of InsP<sub>6</sub> consistent with those found at physiological pH are shown in Table 5.6. In all cases, the transfer of a phosphate group from 5-PP-InsP<sub>5</sub> was calculated to be thermodynamically favourable. Similar to the results found for PPi, the charge of the products formed affects the overall thermodynamics of these reactions. For example, the formation of MeOPi(-2) was predicted to be favoured over the formation of MeOPi(-1) when comparing the same starting protonation states. In general, increasing the negative charge around the *myo*-inositol ring also increased the predicted thermodynamic favourability of phosphorylation. This result is likely due to the increased electrostatic repulsion felt by the pyrophosphate group of 5-PP-InsP<sub>5</sub>. In addition, under physiological conditions, the phosphorylation of methanol, and by extension serine, by 5-PP-InsP<sub>5</sub> was predicted to be thermodynamically favoured over phosphorylation by PPi. There are subtle variations between the predicted  $\Delta H_{rxn}$  values for tautomers of 5-PP-InsP<sub>5</sub> with the same charge; however, this observation may be due to conformational variations caused by the differences in charge placement around the ring, as well as differences in hydrogen bonding patterns. A more encompassing set of calculations can be found in Appendix 2.

**Table 5.6:**  $\Delta H_{\text{rxn}}$  values calculated for the phosphorylation of methanol by 5-PP-InsP<sub>5</sub> ionization states expected at physiological pH.

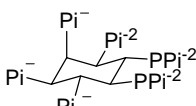
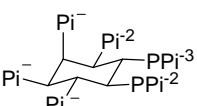
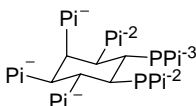
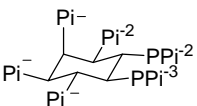
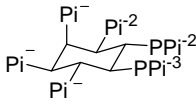
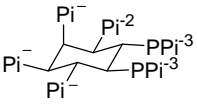
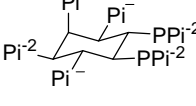
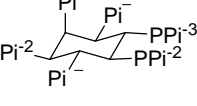
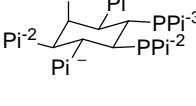
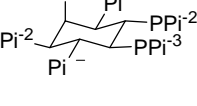
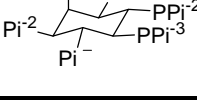
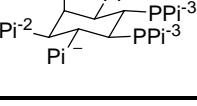
5-PP-InsP <sub>5</sub> Forming MeOPi(-1)	→	InsP <sub>6</sub>	$\Delta H_{\text{rxn}}^a$	5-PP-InsP <sub>5</sub> Forming MeOPi(-2)	→	InsP <sub>6</sub>	$\Delta H_{\text{rxn}}^a$
<b>pH 5.7</b>							
	→		-15.7		→		-32.1
	→		-14.9		→		-31.8
<b>pH 6.8</b>							
	→		-24.9		→		-38.3
	→		-15.9		→		-33.9
<b>pH 7.6</b>							
	→		-22.6	ND <sup>b</sup>			
	→		-19.9	ND			

<sup>a</sup>in kcal/mol; <sup>b</sup>Not determined as further deprotonation is improbable due to the increasing negative charge on InsP<sub>6</sub>.

Results for the phosphorylation of methanol by 5,6-bis-PP-InsP<sub>4</sub> to yield 5-PP-InsP<sub>5</sub> with protonation states consistent with the known pK<sub>a</sub> values of InsP<sub>6</sub> are shown in Table 5.7 and Table 5.8. Consistent with the pyrophosphate model compounds and 5-PP-InsP<sub>5</sub>, the charge states of the product affects the overall thermodynamics of the phosphorylation: the formation of MeOPi(-2) was predicted to be favoured over MeOPi(-1) for the same phosphorylating agent. Similar to 5-PP-InsP<sub>5</sub>, an increase in the charge around the *myo*-inositol ring increase the calculated thermodynamic favourability of the phosphorylation of methanol by 5,6-bis-PP-InsP<sub>4</sub>. A comparison of the  $\Delta H_{\text{rxn}}$  values determined for the phosphorylation of methanol at physiological pH indicates that when forming the same ionization state of MeOPi, 5,6-bis-PP-InsP<sub>4</sub> was calculated to be a thermodynamically

superior phosphorylating agent to 5-PP-InsP<sub>5</sub> and the pyrophosphate model compounds. Steric and electrostatic repulsions caused by the adjacent pyrophosphate groups are likely the cause of this increased favourability. Similar to the results for 5-PP-InsP<sub>5</sub>, there are variations between the  $\Delta H_{\text{rxn}}$  values of tautomers of the same charge. As discussed for PP-InsP<sub>5</sub>, this result may be due to conformational variations caused by the different charge placements around the ring, as well as changes to the hydrogen bonding patterns. A more encompassing set of calculations can be found in *Appendix 2*. The highly charged and sterically congested structure of 5,6-bis-PP-InsP<sub>4</sub> in an overall charged state of -8 can be appreciated in Figure 5.1, where the electrostatic potential is mapped onto the electron density of this compound.

**Table 5.7:**  $\Delta H_{\text{rxn}}$  values calculated for the phosphorylation of methanol by 5,6-bis-PP-InsP<sub>4</sub> ionization states expected at pH 5.7.

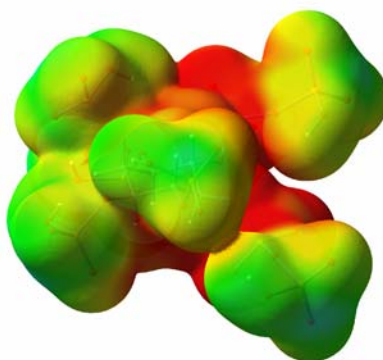
5,6-PP-InsP <sub>4</sub> Forming MeOPi(-1)	$\Delta H_{\text{rxn}}$	5,6-PP-InsP <sub>4</sub> Forming MeOPi(-2)	$\Delta H_{\text{rxn}}$
	-29.7		-35.0
	-18.6		-35.8
	-19.5		-52.9
	-17.9		-35.7
	-18.8		-35.7
	-17.6		-53.2

<sup>a</sup>In kcal/mol.

**Table 5.8:**  $\Delta H_{\text{rxn}}$  values calculated for the phosphorylation of methanol by 5,6-bis-PP-InsP<sub>4</sub> ionization states expected at pH 6.8 – 7.6.

5,6-bis-PP-InsP <sub>4</sub> Forming MeOPi(-1)	$\Delta H_{\text{rxn}}^a$	5,6-bisPP-InsP <sub>4</sub> Forming MeOPi(-2)	$\Delta H_{\text{rxn}}$
	-30.6		-39.6
	-14.9		-46.6
	-33.2		-53.8
	-28.8		-48.5
	-30.6		-48.0
	-30.1		-50.8

<sup>a</sup>in kcal/mol.



**Figure 5.9:** Electrostatic potential (using a range of -0.8 (red) to -0.5 (blue) kcal/mol) mapped onto the electron density isosurface (0.0004 electrons/au<sup>3</sup>) for 5,6-bis-PP-InsP<sub>4</sub> with a charge of -8.

### 5.6.7. Phosphorylation of Methanol by Adenosine Triphosphate and Other Model Compounds

Preliminary calculations of the  $\Delta H_{\text{rxn}}$  values for phosphorylation of methanol by ATP, ADP, AMP and the model compounds triphosphate and methyl triphosphate were performed. Comparisons of these results to those of the InsPPs indicate that, in some cases, phosphorylation of methanol by the 5-PP-InsP<sub>5</sub> and 5,6-bis-PP-InsP<sub>4</sub> is more favourable than phosphorylation by ATP and ADP.

**Table 5.9:**  $\Delta H_{\text{rxn}}$  values calculated for the phosphorylation of methanol by ATP, ADP and model compounds.

<i>Pi Donor</i> <sup>a</sup>	→	<i>Product</i>	$\Delta H_{\text{rxn}}$ <sup>c</sup>	<i>Pi Donor</i>	→	<i>Product</i>	$\Delta H_{\text{rxn}}$
Forming MeOPi(-1) <sup>b</sup>				Forming MeOPi(-2)			
ATP(-3) <sup>d</sup>	→	ADP(-2) <sup>e</sup>	-2.2	ATP(-4)	→	ADP(-2)	-10.7
ATP(-4)	→	ADP(-3)	-11.8				
ADP(-2)	→	AMP(-1) <sup>f</sup>	-5.4	ADP(-3)	→	AMP(-1)	-4.3
ADP(-3)	→	AMP(-2)	-7.4				
MeOPPPPi(-3) <sup>g</sup>	→	MeOPPi(-2) <sup>h</sup>	-10.0	MeOPPPPi(-4)	→	MeOPPi(-2)	-22.3
MeOPPPPi(-4)	→	MeOPPi(-3)	-14.3				
PPPi(-3) <sup>i</sup>	→	PPi(-2) <sup>j</sup>	-7.2	PPPi(-4)	→	PPi(-2)	-18.1
PPPi(-4)	→	PPi(-3)	-13.0				

<sup>a</sup>Pi = phosphate; <sup>b</sup>MeOPi = methyl phosphate; <sup>c</sup> $\Delta H_{\text{rxn}}$  is in kcal/mol; <sup>d</sup>ATP = adenosine triphosphate; <sup>e</sup>ADP = adenosine diphosphate; <sup>f</sup>AMP = adenosine monophosphate; <sup>g</sup>MeOPPPPi = methyl triphosphate; <sup>h</sup>MeOPPi = methyl pyrophosphate; <sup>i</sup>PPPi = triphosphate; <sup>j</sup>PPi = pyrophosphate.

## 5.7. Conclusions and Future Work

It is evident from this study that increasing the overall charge of the phosphate donor, as well as the steric bulk near the pyrophosphate group, tends to increase the predicted thermodynamic favourability of phosphorylation. A specific value for the transfer potential of a phosphate group to methanol by 5-PP-InsP<sub>5</sub> and 5,6-PP-InsP<sub>4</sub> cannot be easily determined due to the ensemble of tautomeric and charged states that would exist under physiological conditions. It is clear that phosphorylation would be a thermodynamically favourable process and biological systems may be able to control the phosphate transfer potential of the InsPPs by tuning the steric and electrostatic environments of the pyrophosphate moiety.

It is tempting to comment on the effect of the position of the negative charges on the overall  $\Delta H_{\text{rxn}}$  values of phosphate transfer; however, the limitations of the methods used precludes the drawing of certain conclusions. The use of implicit solvation simplifies the calculation but it eliminates any thermodynamic contribution of hydrogen bonding which may occur between the surrounding water molecules and compound in question. Further analysis is warranted to determine the thermodynamic contribution of each phosphate on the inositol ring to the overall thermodynamics of each 5-PP-InsP<sub>5</sub> and 5,6-bis-PP-InsP<sub>4</sub> structure. The inclusion of at least one solvation sphere of explicitly defined water molecules should be a future approach to account for the effects of hydrogen bonding with the solvation sphere. The contribution of counter-ions to the overall thermodynamics of these reactions would be an obvious route towards a more quantitative comparison of the thermodynamics of phosphorylation by the InsPPs to that of ATP and ADP.

## 5.8. References:

1. Voet, D.; Voet, J. G. *Biochemistry*; John Wiley and Sons: Hoboken, 2004.
2. Reeves, R. E.; Guthrie, J. D. *Biochem. Biophys. Res. Commun.* **1975**, *66*, 1389-1395.
3. Bruchhaus, I.; Jacobs, T.; Denart, M.; Tannich, E. *Biochem. J.* **1996**, *316*, 57-63.
4. Mertens, E.; Ladrör, U. S.; Lee, J. A.; Miretsky, A.; Morris, A.; Rozario, C.; Kemp, R. G.; Muller, M. *J. Mol. Evol.* **1998**, *47*, 739-750.
5. Chi, A.; Kemp, R. G. *J. Biol. Chem.* **2000**, *275*, 35677-35679.
6. Kornberg, A.; Rao, N. N.; Ault-Riche, D. *Ann. Rev. Biochem.* **1999**, *68*, 89-125.
7. Saiardi, A.; Bhandari, R.; Resnick, A. C.; Snowman, A. M.; Snyder, S. H. *Science* **2004**, *306*, 2101-2105.
8. Luo, H. R.; Huang, C. F.; Chen, J.-C.; Saiardi, A.; Iijima, M.; Ye, K.; Hunag, Y.; Nagata, E.; Devreotes, P.; Snyder, S. H. *Cell* **2003**, *114*, 559-572.
9. Morrison, B. H.; Bauer, J. A.; Kalvakolanu, D. V.; Lindner, D. J. *J. Biol. Chem.* **2001**, *276*, 24965-24970.
10. Saiardi, A.; Resnick, A. C.; Snowman, A. M.; Wendland, B.; Snyder, S. H. *Proc. Natl. Acad. Sci.* **2005**, *102*, 1911-1914.
11. Saiardi, A.; Sciambi, C.; McCaffery, J. M.; Wendland, B.; Snyder, S. H. *Proc. Natl. Acad. Sci. USA* **2005**, *99*, 14206-14211.
12. Irvine, R. F.; Schell, M. J. *Nature Rev. Mol. Cell. Biol.* **2001**, *2*, 327-338.
13. McCarthy, A. A.; Peterson, N. A.; Knijff, R.; Baker, E. N. *J. Mol. Biol.* **2004**, *335*, 1131-1141.
14. Menniti, F. S.; Miller, R. N.; Putney, J. W., Jr.; Shears, S. B. *J. Biol. Chem.* **1993**, *268*, 3850-3856.
15. Shears, S. B.; Ali, N.; Craxton, A.; Bembenek, M. E. *J. Biol. Chem.* **1995**, *270*, 10489-10497.
16. Glennon, M. C.; Shears, S. B. *Biochem. J.* **1993**, *293*, 583-590.
17. Saiardi, A.; Caffrey, J. J.; Snyder, S. H.; Shears, S. B. *FEBS Lett.* **2000**, *468*, 28-32.
18. Stephens, L.; Radenberg, T.; Theil, U.; Vogel, G.; Khoo, K.-H.; Dell, A.; Jackson, T. R.; Hawkins, P. T.; Mayr, G. W. *J. Biol. Chem.* **1993**, *268*, 4009-4015.
19. Martin, J.-B.; Bakker-Grunwald, T.; Klein, G. *Eur. J. Biochem.* **1993**, *214*, 711-718.
20. Laussmann, T.; Hansen, A.; Reddy, K. M.; Reddy, K. K.; Falck, J. R.; Vogel, G. *FEBS Lett.* **1998**, *426*, 145-150.
21. Martin, J.-B.; Bakker-Grunwald, T.; Klein, G. *J. Eukaryot. Microbiol.* **1995**, *42*, 183-191.
22. Albert, C.; Safrany, S. T.; Bembenek, M. E.; Reddy, K. M. *Biochem. J.* **1997**, *327*, 553-560.
23. Huang, C. F.; Voglmaier, S. M.; Bembenek, M. E.; Saiardi, A.; Snyder, S. H. *Biochemistry* **1998**, *37*, 14998-15004.
24. Laussmann, T.; Pikzack, C.; Thiel, U.; Mayr, G. W.; Vogel, G. *Eur. J. Biochem.* **2000**, *267*, 2447-2451.
25. Laussman, T.; Reddy, K. M.; Reddy, K. K.; Falck, J. R.; Vogel, G. *Biochem. J.* **1997**, *322*, 31-33.
26. Chen, L.; Zhou, C.; Yang, H.; Roberts, M. F. *Biochemistry* **2000**, *39*, 12415-12423.
27. Majumder, A. L.; Johnson, M. D.; Henry, S. A. *Biochim Biophys Acta* **1997**, *1348*, 245-256.
28. Loewus, M. W. *J. Biol. Chem.* **1977**, *252*, 7221-7223.
29. Bearly, C. A.; Hanke, D. E. *Biochem. J.* **1996**, *314*, 227-233.

30. Stephens, L.; Irvine, R. F. *Nature* **1990**, *346*, 580-583.
31. Verbsky, J. W.; Chang, S.-C.; Wilson, M. P.; Mochizuki, Y.; Majerus, P. W. *J. Biol. Chem.* **2005**, *280*, 1911-1920.
32. Saiardi, A.; Erdjument-Bromage, H.; Snowman, A. M.; Tempst, P.; Snyder, S. H. *Curr. Biol.* **1999**, *9*, 1323-1326.
33. Voglmaier, S. M.; Bembenek, M. E.; Kaplin, A. E.; Dormán, G.; Olszewski, J. D.; Prestwich, G. D.; Snyder, S. H. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 4305-4310.
34. Saiardi, A.; Nagata, E.; Luo, H. R.; Snowman, A. M.; Snyder, S. H. *J. Biol. Chem.* **2001**, *276*, 39179-39185.
35. Schell, M. J.; Letcher, A. J.; Brearley, C. A.; Biber, J.; Murer, H.; Irvine, R. F. *FEBS Lett.* **1999**, *461*, 169-172.
36. Saiardi, A.; Caffrey, J. J.; Snyder, S. H.; Shears, S. B. *J. Biol. Chem.* **2000**, *275*, 24686-24692.
37. Nagata, E.; Luo, H. R.; Saiardi, A.; Bae, B.-I.; Suzuki, N.; Snyder, S. H. *J. Biol. Chem.* **2005**, *280*, 1634-1640.
38. Laussmann, T.; Eujen, R.; M., W.; Thiel, U.; Vogel, G. *Biochem. J.* **1996**, *315*, 715-720.
39. Saint-Martin, H.; Ortega-Blake, I.; Les, A.; Adamowicz, L. *Biochim. Biophys. Acta* **1994**, *1207*, 12-23.
40. Hart, J. C.; Burton, N. A.; Hillier, I. H.; Harrison, M. J.; Jewsbury, P. *Chem. Commun.* **1997**, 1431-1432.
41. Rzepa, H. S.; Yi, M. *Chem. Commun.* **1989**, 1502-1504.
42. Colvin, M.; Evleth, E. M.; Akacem, Y. *J. Am. Chem. Soc.* **1995**, *117*, 4357-4362.
43. Ma, B.; Meredith, C.; Henry F. Schaefer, I. *J. Phys. Chem.* **1995**, *99*, 3815-3822.
44. Hwang, M.-J.; Chu, P.-Y.; Chen, J.-C.; Chao, I. *J. Comput. Chem.* **1999**, *20*, 1702-1715.
45. Wang, X.-B.; Vorpapel, E. R.; Yang, X.; Wang, L.-S. *J. Phys. Chem. A* **2001**, *105*, 10468-10474.
46. Saint-Martin, H.; Ruiz-Icent, L. E.; Ramirez-Solis, A.; Ortega-Blake, I. *J. Am. Chem. Soc.* **1996**, *118*, 12167-12173.
47. Forián, J.; Warshel, A. *J. Phys. Chem. B* **1998**, *102*, 719-734.
48. McCarthy, W. J.; Smith, D. M. A.; Adamowicz, L.; Saint-Martin, H.; Ortega-Blake, I. *J. Am. Chem. Soc.* **1998**, *120*, 6113-6120.
49. Saint-Martin, H.; Vicent, L. E. *J. Phys. Chem. A* **1999**, *103*, 6862-6872.
50. McCarthy, W. J.; Adamowicz, L.; Saint-Martin, H. *Rev. Soc. Quim. Méx.* **2002**, *46*, 145-158.
51. Ma, B.; Meredith, C.; Schaefer, H. F., III *J. Phys. Chem.* **1994**, *98*, 8216-8223.
52. Franzini, E.; Fantucci, P.; de Gioia, L. *J. Mol. Catalysis A* **2003**, *204-205*, 409.
53. Dejaegere, A.; Liang, X.; Karplus, M. *J. Chem. Soc., Faraday Trans* **1994**, *90*, 1763 - 1770.
54. Boström, J.; Norrby, P.-O.; Liljefors, T. *J. Comput. Aided Mol. Des.* **1998**, *12*, 383-396.
55. Leach, A. R. *Molecular Modelling: Principles and Applications*; Second ed.; Prentice Hall: Harlow, Essex, England, 2001.
56. Toney, M. D. *Biochemistry* **2001**, *40*, 1378-1384.
57. Young, D. C. *Computational Chemistry: A Practical Guide for Applying Techniques for Real World Problems*; Wiley Interscience: New York, 2001.
58. Cramer, C. J. *Essentials of Computational Chemistry*; Second ed.; John Wiley and Sons Ltd.: West Sussex, 2004.

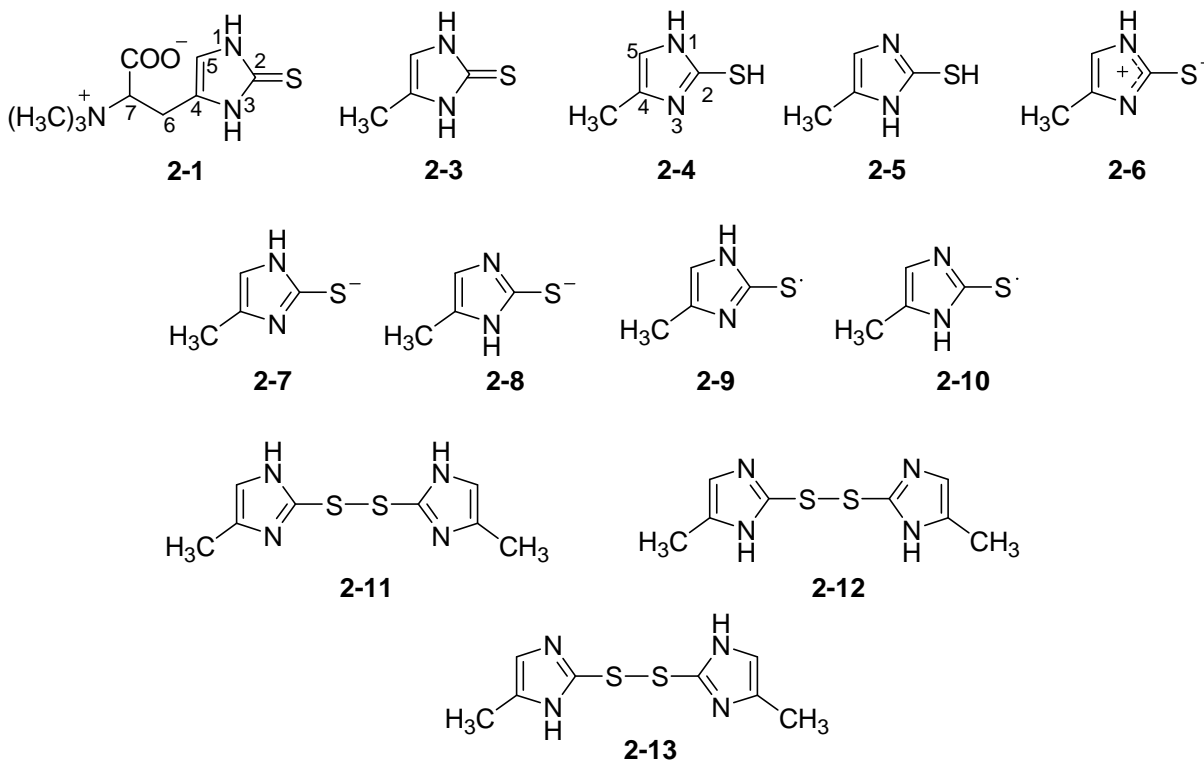
59. Stewart, J. J. P. *J. Comput. Chem.* **1989**, *10*, 209-220.
60. Stewart, J. J. P. *J. Comput. Chem.* **1989**, *10*, 221-264.
61. Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. B* **1998**, *102*, 3257-3271.
62. Dewar, M. J. S.; Holder, A. J.; Dennington, R. D., II; Liotard, D. A.; Truhlar, D. G.; Keith, T. A.; Millam, J. M.; Harris, C. D. *AMPAC 8 User Manual*; Semichem Inc.: Kansas City, 2004.
63. Bunton, C. A.; Llewellyn, D. R.; Oldham, K. G.; Vernon, C. A. *J. Chem. Soc.* **1958**, 3574 - 3587.
64. Saha, A.; Saha, N.; Ji, L.-N.; Zhao, J.; Gregan, F.; Sajadi, S. A. A.; Song, B.; Sigel, H. *J. Biol. Inorg. Chem.* **1996**, *1*, 231-238.
65. Costello, A. J. R.; Glonek, T.; Mycers, T. C. *Carbohydr. Res.* **1976**, *46*, 159.
66. Damm, W.; Frontera, S.; Tirado-Rives, J.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1997**, *119*, 1955-1970.
67. Vishnyakov, A.; Widmalm, G.; Kowalewski, J.; Laaksonen, A. *J. Am. Chem. Soc.* **1999**, *121*, 5403-5412.
68. Conrad, P. B.; de Pablo, J. J. *J. Phys. Chem. A* **1999**, *103*, 4049-4055.
69. Umemura, M.; Hayashi, S.; Nakagawa, T.; Urakawa, H.; Kajiwara, K. *J. Mol. Struct. (THEOCHEM)*. **2003**, *636*, 215-228.
70. Starikov, E. B.; Bräsicke, K.; Knapp, E. W.; Saenger, W. *Chem. Phys. Lett.* **2001**, *336*, 504-510.
71. Bauman, A. T.; Chateuaneuf, G. M.; Boyd, B. R.; Brown, R. E.; Murthy, P. P. N. *Tetrahedron Lett.* **1999**, *40*, 4489.
72. Shrodinger Inc.
73. Corrie, J. E. T.; Barth, A.; Munasinghe, V. R. N.; Trentham, D. R.; Hutter, M. C. *J. Am. Chem. Soc.* **2004**, *126*, 8546.
74. Brandt, W.; Dessoy, M. A.; Fulhorst, M.; Gao, W.; Zenk, M. H.; Wessjohann, L. A. *ChemBioChem* **2004**, *5*, 311.
75. Yang, H.; Feng, J.; Liu, Y.; Yang, Y.; Wu, J.; Zhang, Z.; Shen, G.; Hyu, R. *J. Raman Spec.* **2005**, *36*, 824.
76. Guruzzo, E.; Rogers, D.; Williams, D. J. *Eur. Cryst. Meeting* **1979**, *5*, 32.
77. Fennessey, J. P.; Nowacki, W. *Helv. Chim. Acta* **1968**, *51*, 1475.
78. Blank, G. E.; Pletcher, J.; Sax, M. *Acta Crystallogr.* **1975**, *B31*, 2584-2592.
79. Truter, M. R.; Tate, M. E. *J. Chem. Soc. B* **1970**, 70.
80. Guynn, R. W.; Veech, R. L. *J. Biol. Chem.* **1973**, *248*, 6966-?
81. Laussmann, T.; Reddy, K. M.; Reddy, K. K.; Falck, J. R.; Vogel, G. *Biochem. J.* **1997**, *322*, 31-33.
82. Barrientos, L. G.; Murthy, P. P. N. *Carbohydr. Res.* **1996**, *296*, 39-54.
83. Paton, G.; Noailly, M.; Mossoyan, J. C. *J. Phys. Org. Chem.* **1999**, *12*, 401-407.



## Appendix 1: Modelling of Ergothioneine and Ovothiol Supplementary Information

Please see *Chapter 2* for a full description of the methods used as well as a discussion of these results.

### A1.1: Studies of Ergothioneine and 2-Thiol-4-Methyl-Imidazole



**Figure A1.1:** Structures of 2-thiol-4-methyl-imidazole (TMI) and ergothioneine (ESH)

#### A1.1.1: Crystal Structure Data

ESH was recrystallized from water/ethanol and its crystal structure was determined by Dr. Nicholas J. Taylor. The coordinates were deposited in the Cambridge Crystallographic Database, deposition FOLRII.

**Table A1.1:** Bond lengths of ergothioneine determined using X-ray crystallography in this work.<sup>a</sup>

<b><i>Bond</i></b>	<b><i>Length in Å</i></b>	<b><i>Bond</i></b>	<b><i>Length in Å</i></b>
S-C <sup>2</sup>	1.6931(11)	C <sup>7</sup> -C <sup>8</sup>	1.5524(16)
N <sup>1</sup> -C <sup>2</sup>	1.3518(15)	C <sup>7</sup> -N <sup>11</sup>	1.5365(14)
N <sup>1</sup> -C <sup>5</sup>	1.3807(15)	C <sup>8</sup> -O <sup>9</sup>	1.2522(15)
N <sup>3</sup> -C <sup>2</sup>	1.3541(14)	C <sup>9</sup> -O <sup>10</sup>	1.2512(15)
N <sup>3</sup> -C <sup>4</sup>	1.3870(15)	N <sup>11</sup> -C <sup>12</sup>	1.5004(16)
C <sup>4</sup> -C <sup>5</sup>	1.3524(17)	N <sup>11</sup> -C <sup>13</sup>	1.5049(15)
C <sup>4</sup> -C <sup>6</sup>	1.4952(16)	N <sup>11</sup> -C <sup>14</sup>	1.5107(15)
C <sup>6</sup> -C <sup>7</sup>	1.5238(16)		

<sup>a</sup>See Figure A1.1 for numbering.**Table A1.2:** Bond angles of ergothioneine as determined by X-ray crystallography in this work.<sup>a</sup>

<b><i>Bonds</i></b>	<b><i>Angle (in degrees)</i></b>	<b><i>Bonds</i></b>	<b><i>Angle (in degrees)</i></b>
N <sup>1</sup> -C <sup>2</sup> -N <sup>3</sup>	105.55(10)	C <sup>6</sup> -C <sup>7</sup> -C <sup>8</sup>	111.39(10)
N <sup>1</sup> -C <sup>2</sup> -S	126.55(9)	C <sup>6</sup> -C <sup>7</sup> -N <sup>11</sup>	110.79(8)
N <sup>1</sup> -C <sup>5</sup> -C <sup>4</sup>	107.22(11)	C <sup>7</sup> -C <sup>8</sup> -O <sup>9</sup>	115.52(11)
C <sup>2</sup> -N <sup>1</sup> -C <sup>5</sup>	110.34(10)	C <sup>7</sup> -C <sup>8</sup> -O <sup>10</sup>	116.25(9)
C <sup>2</sup> -N <sup>3</sup> -C <sup>4</sup>	110.41(10)	C <sup>7</sup> -N <sup>11</sup> -O <sup>12</sup>	109.65(9)
N <sup>3</sup> -C <sup>2</sup> -S	127.89(9)	C <sup>7</sup> -N <sup>11</sup> -O <sup>13</sup>	108.62(8)
N <sup>3</sup> -C <sup>4</sup> -C <sup>5</sup>	106.48(10)	C <sup>7</sup> -N <sup>11</sup> -O <sup>14</sup>	113.39(9)
N <sup>3</sup> -C <sup>4</sup> -C <sup>6</sup>	123.71(11)	C <sup>12</sup> -N <sup>11</sup> -C <sup>13</sup>	108.68(10)
C <sup>4</sup> -C <sup>6</sup> -C <sup>7</sup>	112.37(9)	C <sup>12</sup> -N <sup>11</sup> -C <sup>14</sup>	108.35(9)
C <sup>5</sup> -C <sup>4</sup> -C <sup>6</sup>	129.75(12)	C <sup>13</sup> -N <sup>11</sup> -C <sup>14</sup>	108.04(9)

<sup>a</sup>See Figure A1.1 for numbering

### *A1.1.2: Geometry Optimizations of the Ergothioneine Crystal Structure*

A series of geometry optimizations of the ESH crystal structure were performed using Jaguar (Schrodinger Inc.) and Hartree-Fock (HF) and Density Functional Theory (DFT) methods with a variety of basis sets in the presence and absence of diffuse functions. The DFT calculations used the B3LYP functional as implemented by Jaguar unless otherwise stated. Full procedural details can be found in *Chapter 2*.

**Table A1.3:** Bond distances (in Å) returned by geometry optimizations of the crystal structure of ergothioneine<sup>a</sup> using density functional theory methods in the gas phase.<sup>b</sup>

<i>Basis set</i>	**	++	$C^2-S$	$N^1-C^2$	$C^2-N^3$	$N^1-C^5$	$N^3-C^4$	$C^4-C^5$	$C^4-C^6$
<b>B3LYP</b>									
6-31G	**	++	1.68	1.37	1.37	1.39	1.40	1.36	1.50
	**	None	1.68	1.37	1.37	1.39	1.40	1.36	1.50
6-311g	none	None	1.73	1.38	1.38	1.40	1.41	1.36	1.50
	**	++	1.67	1.37	1.37	1.39	1.40	1.35	1.50
	**	None	1.68	1.37	1.37	1.39	1.40	1.36	1.50
cc-pVDZ	None	none	1.72	1.37	1.38	1.40	1.41	1.36	1.50
	Incl. <sup>c</sup>	++ <sup>d</sup>	1.68	1.37	1.37	1.39	1.40	1.36	1.50
cc-pVTZ(-f)	Incl.	None	1.67	1.37	1.37	1.38	1.40	1.36	1.50
	Incl.	++	ND <sup>e</sup>	ND	ND	ND	ND	ND	ND
	Incl.	None	1.67	1.37	1.37	1.39	1.40	1.35	1.49
<b>B3PW91</b>									
cc-pVTZ(-f)	**	++	1.67	1.36	1.36	1.38	1.39	1.35	1.49
<b>Crystal<sup>a</sup></b>	--	--	1.69	1.35	1.35	1.38	1.39	1.35	1.50

<sup>a</sup>Obtained in this work; <sup>b</sup>See Figure A1.1 for numbering; <sup>c</sup>Polarization functions are included in this basis set;

<sup>d</sup>Convergence required the use of the GVB-DIIS convergence scheme; <sup>e</sup>Not determined

**Table A1.4:** Bond angles (in degrees) returned by a geometry optimization of the ergothioneine crystal structure<sup>a</sup> using density function theory methods in the gas phase.<sup>b</sup>

<i>Basis set</i>	**	++	<i>S-C<sup>2</sup>-N<sup>1</sup></i>	<i>S-C<sup>2</sup>-N<sup>3</sup></i>	<i>C<sup>2</sup>-N<sup>3</sup>-C<sup>4</sup></i>	<i>C<sup>2</sup>-N<sup>1</sup>-C<sup>5</sup></i>	<i>N<sup>1</sup>-C<sup>2</sup>-N<sup>3</sup></i>	<i>N<sup>3</sup>-C<sup>4</sup>-C<sup>5</sup></i>	<i>N<sup>1</sup>-C<sup>5</sup>-C<sup>4</sup></i>	<i>N<sup>3</sup>-C<sup>4</sup>-C<sup>6</sup></i>
<b>B3LYP</b>										
6-31G	**	++	128.80	127.95	111.91	111.78	103.24	106.01	107.06	123.29
	**	None	128.89	128.05	112.02	112.00	103.06	105.97	106.94	122.43
6-311g	None	None	128.73	127.91	111.75	111.88	103.36	106.93	106.93	122.23
	**	++	128.56	128.14	111.78	111.75	103.30	106.06	107.11	123.42
	**	None	128.60	128.08	111.84	111.84	103.32	106.05	106.96	123.08
cc-pVDZ	None	None	128.52	127.93	111.64	111.66	103.55	106.09	107.07	123.55
	Incl. <sup>c</sup>	++ <sup>d</sup>	128.69	127.97	111.86	111.73	103.33	106.05	107.02	123.17
cc-pVTZ(-f)	Incl.	None	128.75	128.06	112.01	111.84	103.24	106.01	107.06	123.29
	Incl.	None	128.76	127.99	111.85	111.64	103.25	106.04	107.18	123.02
	Incl.	++	ND <sup>e</sup>	ND	ND	ND	ND	ND	ND	ND
<b>B3PW91</b>										
cc-pVTZ(-f)	Incl.	None	128.81	127.97	111.92	111.71	105.23	105.93	107.21	123.37
<b>Crystal<sup>a</sup></b>	--	--	126.55	127.89	110.41	110.34	105.55	106.48	107.22	123.71

<sup>a</sup>Obtained in this work; <sup>b</sup>See Figure A1.1 for numbering; <sup>c</sup>Polarization functions are included in this basis set; <sup>d</sup>Convergence required the use of the GVB-DIIS convergence scheme; <sup>e</sup>Not determined

**Table A1.5:** Bond lengths (in Å) returned after geometry optimizations of the ergothioneine crystal structure<sup>a</sup> using density functional theory methods with implicit water solvation.<sup>b</sup>

<i>Basis set</i>	**	++	<i>C</i> <sup>2</sup> - <i>S</i>	<i>N</i> <sup>1</sup> - <i>C</i> <sup>2</sup>	<i>C</i> <sup>2</sup> - <i>N</i> <sup>3</sup>	<i>N</i> <sup>1</sup> - <i>C</i> <sup>5</sup>	<i>N</i> <sup>3</sup> - <i>C</i> <sup>4</sup>	<i>C</i> <sup>4</sup> - <i>C</i> <sup>5</sup>	<i>C</i> <sup>4</sup> - <i>C</i> <sup>6</sup>
<b>B3LYP</b>									
6-31G	**	++	1.71	1.36	1.36	1.39	1.39	1.36	1.50
	**	None	1.71	1.36	1.36	1.39	1.39	1.36	1.50
	None	None	1.75	1.36	1.36	1.40	1.40	1.37	1.50
6-311G	**	++ <sup>c</sup>	1.70	1.35	1.35	1.38	1.39	1.36	1.50
	**	None <sup>d</sup>	1.71	1.35	1.35	1.40	1.40	1.37	1.50
	None	None	1.75	1.36	1.36	1.39	1.40	1.36	1.50
cc-pVDZ	Incl.	++	N/D	N/D	N/D	N/D	N/D	N/D	N/D
	Incl.	None	1.70	1.35	1.36	1.38	1.39	1.36	1.49
cc-pVTZ(-f)	Incl.	++	1.70	1.35	1.35	1.38	1.39	1.35	1.49
	Incl	None	1.71	1.35	1.35	1.38	1.40	1.36	1.49
<b>B3PW91</b>									
cc-pVTZ(-f)			1.70	1.35	1.35	1.38	1.39	1.35	1.49
<b>Crystal<sup>a</sup></b>	--	--	1.69	1.35	1.35	1.38	1.39	1.35	1.50

<sup>a</sup>Obtained in this work; <sup>b</sup>See Figure A1.1 for numbering; <sup>c</sup>The starting structure for this optimization was the conformation obtained from the optimization of the crystal structure using HF/6-311++G\*\* methods and implicit water solvation; <sup>d</sup>The starting structure for this optimization was the conformation obtained from the optimization of the crystal structure using HF/6-311G\*\* methods and implicit water solvation;

**Table A1.6:** Bond angles (in degrees) returned by geometry optimizations of the ergothioneine crystal structure<sup>a</sup> using density functional theory methods with implicit water solvation<sup>b</sup>

<i>Basis set</i>	**	++	<i>S-C<sup>2</sup>-N<sup>1</sup></i>	<i>S-C<sup>2</sup>-N<sup>3</sup></i>	<i>C<sup>2</sup>-N<sup>3</sup>-C<sup>4</sup></i>	<i>C<sup>2</sup>-N<sup>1</sup>-C<sup>5</sup></i>	<i>N<sup>1</sup>-C<sup>2</sup>-N<sup>3</sup></i>	<i>N<sup>3</sup>-C<sup>4</sup>-C<sup>5</sup></i>	<i>N<sup>1</sup>-C<sup>5</sup>-C<sup>4</sup></i>	<i>N<sup>3</sup>-C<sup>4</sup>-C<sup>6</sup></i>
<b>B3LYP</b>										
6-31G	**	++	127.57	127.18	111.06	110.61	105.24	105.88	107.20	123.14
	**	None	127.68	127.30	111.24	110.77	105.02	105.80	107.18	122.50
6-311G	None	None	127.42	127.33	110.98	110.68	105.25	105..88	107.21	122.81
	**	++ <sup>c</sup>	127.15	127.46	110.96	110.57	105.36	105.86	107.23	123.04
	**	None <sup>d</sup>	127.22	127.38	110.95	110.58	105.39	105.93	107.14	122.76
cc-pVDZ	None	None	127.22	127.29	110.77	110.55	105.49	106.02	107.17	123.13
	Incl.	++	N/D	N/D	N/D	N/D	N/D	N/D	N/D	N/D
cc-pVTZ(-f)	Incl.	None	127.62	127.39	111.21	110.76	104.99	105.74	107.30	122.87
	Incl.	++	127.67	126.98	110.89	110.48	105.36	106.00	107.27	123.12
	Incl.	None	127.68	127.15	111.02	110.67	105.17	105.93	107.22	122.56
<b>B3PW91</b>										
cc-pVTZ(-f)	Incl.	None	128.81	127.97	111.92	111.71	103.21	105.94	107.21	123.25
<b>Crystal<sup>a</sup></b>	--	--	126.55	127.89	110.41	110.34	105.55	106.48	107.22	123.71

<sup>a</sup>Obtained in this work; <sup>b</sup>See Figure A1.1 for numbering; <sup>c</sup>The starting structure for this optimization was the conformation obtained from the optimization of the crystal structure using HF/6-311++G\*\* methods and implicit water solvation; <sup>d</sup>The starting structure for this optimization was the conformation obtained from the optimization of the crystal structure using HF/6-311G\*\* methods and implicit water solvation;

**Table A1.7:** Bond lengths (in Å) returned by geometry optimizations of the ergothioneine crystal structure<sup>a</sup> using Hartree-Fock methods in the gas phase.<sup>b</sup>

<i>Basis set</i>	**	++	$C^2-S$	$N^1-C^2$	$C^2-N^3$	$N^1-C^5$	$N^3-C^4$	$C^4-C^5$	$C^4-C^6$
<b>B3LYP</b>									
3-21G	*	++	1.68	1.36	1.36	1.40	1.41	1.33	1.50
	*	None	1.68	1.34	1.35	1.40	1.41	1.33	1.50
	None	None	1.74	1.34	1.35	1.40	1.41	1.34	1.50
6-31G	**	++	1.69	1.34	1.34	1.39	1.39	1.33	1.50
	**	None	1.69	1.34	1.34	1.39	1.40	1.33	1.50
	None	None	1.74	1.35	1.35	1.40	1.40	1.34	1.50
6-311G	**	++	1.68	1.34	1.34	1.39	1.39	1.33	1.50
	**	None	1.68	1.34	1.34	1.39	1.39	1.33	1.50
	None	None	1.73	1.34	1.34	1.39	1.40	1.34	1.50
cc-pVDz	Incl	++	1.69	1.34	1.34	1.39	1.40	1.34	1.50
		none	1.68	1.34	1.34	1.38	1.39	1.33	1.49
cc-pVTZ(-f)	Incl	++	1.68	1.34	1.34	1.39	1.40	1.32	1.50
	Incl	none	1.68	1.34	1.34	1.39	1.39	1.33	1.49
<b>Crystal<sup>a</sup></b>	--	--	1.69	1.35	1.35	1.38	1.39	1.35	1.50

<sup>a</sup>Obtained in this work; <sup>b</sup>See Figure A1.1 for numbering

**Table A1.8:** Bond angles (in degrees) returned by the geometry optimizations of the ergothioneine crystal structure<sup>a</sup> using Hartree-Fock methods in the gas phase.<sup>b</sup>

<i>Basis set</i>	**	++	<i>S-C<sup>2</sup>-N<sup>1</sup></i>	<i>S-C<sup>2</sup>-N<sup>3</sup></i>	<i>C<sup>2</sup>-N<sup>3</sup>-C<sup>4</sup></i>	<i>C<sup>2</sup>-N<sup>1</sup>-C<sup>5</sup></i>	<i>N<sup>1</sup>-C<sup>2</sup>-N<sup>3</sup></i>	<i>N<sup>3</sup>-C<sup>4</sup>-C<sup>5</sup></i>	<i>N<sup>1</sup>-C<sup>5</sup>-C<sup>4</sup></i>	<i>N<sup>3</sup>-C<sup>4</sup>-C<sup>6</sup></i>
<b>B3LYP</b>										
3-21G	*	++	128.47	127.53	111.34	111.34	103.99	106.29	107.04	122.90
	*	None	128.42	127.49	111.37	111.41	104.09	106.19	106.94	123.10
	None	None	128.21	127.09	111.20	111.22	104.70	106.06	106.82	123.15
6-31G	**	++	128.31	127.18	111.32	111.18	104.51	106.02	106.97	122.64
	**	None	128.24	127.37	111.39	111.30	104.40	105.97	106.94	122.67
	None	None	128.20	127.28	111.34	111.31	104.52	105.95	106.88	122.71
6-311G	**	++	128.04	127.35	111.20	111.12	104.61	106.05	107.02	122.84
	**	None	128.16	127.28	111.30	111.21	104.56	106.00	106.93	122.67
	None	None	128.13	127.26	111.25	111.19	104.61	105.97	106.98	122.81
cc-pVDZ	Incl.	++	128.15	127.23	111.26	111.12	104.62	106.02	106.98	122.83
	Incl.	None	128.24	127.26	111.36	111.20	104.50	105.94	106.99	122.39
cc-pVTZ(-f)	Incl.	++	128.23	127.23	111.35	111.01	104.53	105.87	107.22	122.70
	Incl.	None	128.26	127.14	111.25	111.00	104.60	106.02	107.11	122.55
<b>Crystal<sup>a</sup></b>	--	--	126.55	127.89	110.41	110.34	105.55	106.48	107.22	123.71

<sup>a</sup>Obtained in this work; <sup>b</sup>See Figure A1.1 for numbering



**Table A1.9:** Bond lengths (in Å) returned by geometry optimizations of the ergothioneine crystal structure<sup>a</sup> using Hartree-Fock methods with implicit water solvation.<sup>b</sup>

<i>Basis set</i>	**	++	$C^2-S$	$N^1-C^2$	$C^2-N^3$	$N^1-C^5$	$N^3-C^4$	$C^4-C^5$	$C^4-C^6$
6-31G	**	++	1.72	1.33	1.33	1.38	1.39	1.34	1.50
	**	None	1.71	1.33	1.33	1.38	1.39	1.34	1.50
	none	None	1.76	1.33	1.34	1.39	1.40	1.34	1.50
6-311G	**	++	1.71	1.32	1.33	1.38	1.39	1.34	1.50
	**	None	1.71	1.32	1.33	1.38	1.39	1.34	1.50
	None	None	1.75	1.33	1.33	1.39	1.40	1.34	1.49
cc-pVDZ	Incl.	++	1.72	1.33	1.33	1.38	1.39	1.34	1.50
	Incl.	None	1.71	1.32	1.33	1.38	1.38	1.34	1.49
cc-pVTZ(-f)	Incl.	++	1.71	1.32	1.32	1.38	1.39	1.33	1.49
	Incl.	None	1.71	1.32	1.33	1.38	1.39	1.33	1.49
<b>Crystal<sup>a</sup></b>	--	--	1.69	1.35	1.35	1.38	1.39	1.35	1.50

<sup>a</sup>Obtained in this work; <sup>b</sup>See Figure A1.1 for numbering

**Table A1.10:** Bond angles (in degrees) returned after the geometry optimization of the ergothioneine crystal structure<sup>a</sup> using Hartree-Fock methods with implicit water solvation<sup>b</sup>

<i>Basis set</i>	**	++	<i>S-C<sup>2</sup>-N<sup>1</sup></i>	<i>S-C<sup>2</sup>-N<sup>3</sup></i>	<i>C<sup>2</sup>-N<sup>3</sup>-C<sup>4</sup></i>	<i>C<sup>2</sup>-N<sup>1</sup>-C<sup>5</sup></i>	<i>N<sup>1</sup>-C<sup>2</sup>-N<sup>3</sup></i>	<i>N<sup>3</sup>-C<sup>4</sup>-C<sup>5</sup></i>	<i>N<sup>1</sup>-C<sup>5</sup>-C<sup>4</sup></i>	<i>N<sup>3</sup>-C<sup>4</sup>-C<sup>6</sup></i>
<b>B3LYP</b>										
6-31G	**	++	126.95	126.55	110.39	110.06	106.50	105.97	107.07	123.07
	**	None	126.93	126.70	110.44	110.13	106.67	105.94	107.12	123.17
	None	None	127.06	126.74	110.59	110.30	106.20	105.92	106.98	123.12
6-311G	**	++	126.87	126.53	110.33	110.04	106.59	105.96	107.08	123.00
	**	None	126.94	126.50	110.33	110.06	106.56	106.02	107.04	123.06
	None	None	127.07	126.56	110.47	110.22	106.37	105.96	106.98	122.63
cc-pVDZ	Incl.	++	126.92	126.50	110.38	110.03	106.58	105.93	107.08	123.24
	Incl.	None	127.06	126.61	110.51	110.13	106.33	105.94	107.08	122.95
cc-pVTZ(-f)	Incl.	++	126.91	126.55	110.36	109.97	106.54	105.99	107.15	122.79
	Incl.	None	127.02	126.45	110.34	110.01	106.54	106.02	107.09	122.55
<b>Crystal<sup>a</sup></b>	--	--	126.55	127.89	110.41	110.34	105.55	106.48	107.22	123.71

<sup>a</sup>Obtained in this work; <sup>b</sup>See Figure A1.1 for numbering

### A1.1.3: Geometry Optimizations of 2-Thiol-4-Methyl-Imidazole

All structures were optimized using Gaussian 03 both *in vacuo* and with implicit solvation. The B3LYP functional was used for all calculations with the 6-311G basis set with the addition of polarization and diffuse functions. Full details of these calculations can be found in *Chapter 2*.

**Table A1.11:** Energies, bond lengths and angles of 2-thiol-4-methylimidazole, its tautomers and ergothioneine determined by geometry optimization.<sup>a,b</sup>

<i>Tautomer</i>	<i>Phase</i>	<i>Energy (kcal/mol)</i>	$N^1-C^2-N^3$	$N^1-C^2$	$C^2-N^3$	$C^2-S$
<b>2-3</b>	Gas	-416525.75	103.2	1.37	1.37	1.67
	Water	-416545.39	104.7	1.35	1.36	1.71
<b>2-4</b>	Gas	-416518.04	111.9	1.36	1.31	1.77
	Water	-416532.12	111.8	1.36	1.32	1.77
<b>2-5</b>	Gas	-416517.20	111.2	1.31	1.37	1.77
	Water	-416531.86	110.0	1.32	1.37	1.77
<b>2-6</b>	Gas	-416525.76	103.2	1.37	1.37	1.67
	Water	-416545.41	104.7	1.35	1.36	1.71
ESH <sub>xray</sub> <sup>c</sup>	Gas	-668215.49	103.2	1.37	1.37	1.68
	Water	-668261.72	104.7	1.36	1.40	1.68

<sup>a</sup>Angles in degrees, bond lengths in Å; <sup>b</sup>See Figure A1.1 for numbering <sup>c</sup>X-ray structure, this work<sup>1</sup> optimized using 6-311++G(d,p).

**Table A1.12:** Energies, bond lengths and angles of the thiolate and thiyl radicals of 2-thiol-4-methylimidazole determined by geometry optimization.<sup>a,b</sup>

<i>Species</i>	<i>Phase</i>	<i>Energy (kcal/mol)</i>	$N^1-C^2-N^3$	$N^1-C^2$	$C^2-N^3$	$C^2-S$
<b>2-7</b>	Gas	-416184.52	108.4	1.38	1.34	1.73
	Water	-416251.60	109.5	1.37	1.34	1.75
<b>2-8</b>	Gas	-416183.10	108.5	1.34	1.39	1.73
	Water	-416251.37	109.4	1.33	1.38	1.75
<b>2-9</b>	Gas	-416134.63	108.6	1.40	1.35	1.68
	Water	-416148.87	108.7	1.39	1.36	1.68
<b>2-10</b>	Gas	-416134.52	108.5	1.35	1.40	1.68
	Water	-416149.09	108.7	1.36	1.39	1.68

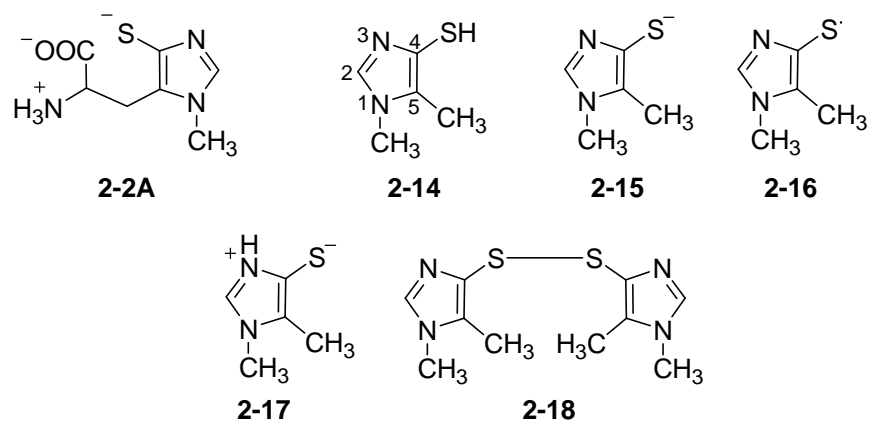
<sup>a</sup>Angles in degrees, bond lengths in Å; <sup>b</sup>See Figure A1.1 for numbering

**Table A1.13:** Energies and selected bond lengths and angles of the various tautomers of the disulfide of 2-thiol-4-methylimidazole determined by geometry optimization.<sup>a,b</sup>

<i>Form</i>	<i>Phase</i>	<i>Energy</i>	<i>N<sup>1</sup>-C<sup>2</sup>-N<sup>3</sup></i>	<i>Ring A/Ring B</i>		<i>C-S</i>	<i>S-S</i>	<i>C<sup>A2</sup>-S<sup>A</sup>-S<sup>B</sup>-C<sup>B2</sup></i>
				<i>N<sup>1</sup>-C<sup>2</sup></i>	<i>C<sup>2</sup>-N<sup>3</sup></i>			
<b>2-11</b>	Gas	-832280.12	110.8/110.8	1.37/1.37	1.32/1.32	1.76/1.76	2.16	-80.7
	Water	-832299.34	111.8/110.9	1.36/1.36	1.32/1.76	1.77/1.76	2.15	83.8
<b>2-12</b>	Gas	-832280.36	110.9/110.9	1.32/1.32	1.38/1.38	1.75/1.75	2.18	-73.3
	Water	ND <sup>c</sup>	ND	ND	ND	ND	ND	ND
<b>2-13</b>	Gas	-832283.43	112.1/110.5	1.36/1.37	1.32/1.32	1.76/1.76	2.14	-84.1
	Water	-832299.29	111.7/110.9	1.36/1.33	1.32/1.36	1.77/1.76	2.15	-84.0

<sup>a</sup>Energy in kcal/mol, angles in degrees, bond lengths in Å; <sup>b</sup>See Figure A1.1 for numbering; <sup>c</sup>ND = not determined

### A1.2: Studies of Ovothiol A and 4-thiol-*N*<sup>1</sup>-methyl-5-methylimidazole



**Figure A1.2:** Ovothiol A (OSH<sub>A</sub>) and 4-thiol-*N*<sup>1</sup>-methyl-5-methylimidazole (TMMI), used to model OSH and its forms.

### A1.2.1: Geometry Optimizations of 4-Thiol-*N*<sup>1</sup>-Methyl-5-Methyl-Imidazole

All structures were optimized using Gaussian 03 both *in vacuo* and with implicit solvation. The B3LYP functional was used for all calculations with the 6-311G bases set with the addition of polarization and diffuse functions. Full details of these calculations can be found in *Chapter 2*.

**Table A1.14:** Energies and selected bond lengths and angles for ovoidiol A and 4-thiol-*N*<sup>1</sup>-methyl-5-methylimidazole, its thiolate, thiyl radical and zwitterion determined by geometry optimization.<sup>a,b</sup>

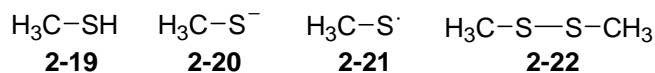
<i>Species</i>	<i>Phase</i>	<i>Energy (kcal/mol)</i>	<i>N<sup>1</sup>-C<sup>2</sup>-N<sup>3</sup></i>	<i>N<sup>1</sup>-C<sup>2</sup></i>	<i>C<sup>2</sup>-N<sup>3</sup></i>	<i>C<sup>4</sup>-S</i>
<b>2-14</b>	Gas	-441174.55	112.3	1.37	1.31	1.78
	Water	-441185.67	111.9	1.36	1.32	1.78
<b>2-15</b>	Gas	-440829.10	112.6	1.37	1.31	1.75
	Water	-440899.53	111.9	1.36	1.32	1.77
<b>2-16</b>	Gas	-440789.69	112.9	1.39	1.30	1.70
	Water	-440801.20	112.6	1.38	1.30	1.65
<b>2-17</b>	Gas	-441166.64	107.0	1.34	1.34	1.67
	Water	-441189.96	107.7	1.33	1.34	1.74
<b>OSH<sub>A</sub> thiolate</b>	Gas	-618596.11	112.8	1.36	1.31	1.75
	Water	ND	ND	ND	ND	ND

<sup>a</sup>Angles in °, bond lengths in Å; <sup>b</sup>See Figure A1.3

### A1.3: Studies of Methyl Mercaptan and Other Biologically Relevant Compounds

All structures were optimized using Gaussian 03 both *in vacuo* and with implicit solvation. The B3LYP functional was used for all calculations with the 6-311G bases set with the addition of polarization and diffuse functions. Full details of these calculations can be found in *Chapter 2*.

#### A1.3.1: Geometry Optimizations of Methyl Mercaptan and its Derivatives



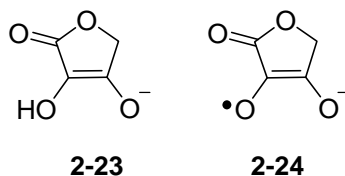
**Figure A1.3:** Methylmercaptan structures used to model glutathione and its various forms.

**Table A1.15:** Energies and C-S bond lengths for methylmercaptan, its thiolate, thiyl radical and disulfide determined by geometry optimization.<sup>a,b</sup>

<i>Species</i>	<i>Phase</i>	<i>Energy (kcal/mol)</i>	<i>C-S (Å)</i>	<i>S-S (Å)</i>
<b>2-19</b>	Gas	-275302.12	1.84	--
	Water	-275306.64	1.84	--
<b>2-20</b>	Gas	-274946.62	1.85	--
	Water	-275014.20	1.86	--
<b>2-21</b>	Gas	-274946.52	1.81	--
	Water	-274907.33	1.81	--
<b>2-22</b>	Gas	-549849.30	1.83	2.09
	Water	-549852.16	1.83	2.10

<sup>a</sup>Bond lengths in Å, <sup>b</sup>See Figure A1.3

### A1.3.2: Geometry Optimizations of the Ascorbate Model Compound



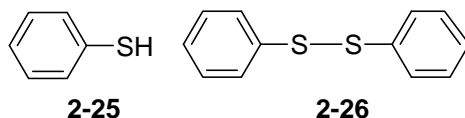
**Figure A1.4:** The compounds used to model ascorbic acid and its radical.

**Table A1.16:** Energies of the model compounds of ascorbate and its radical determined by geometry optimization.<sup>a</sup>

<i>Species</i>	<i>Phase</i>	<i>Energy (kcal/mol)</i>
<b>2-23</b>	Gas	-285696.31
	Water	-285759.82
<b>2-24</b>	Gas	-285319.30
	Water	-285379.69

<sup>a</sup>See Figure A1.4

### A1.3.3: Geometry Optimizations of Thiophenol and Diphenyldisulfide



**Figure A1.5:** Thiophenol and diphenyldisulfide are used as models for aromatic thiols.

**Table A1.17:** Energies and bond lengths of thiophenol and diphenyldisulfide determined by geometry optimization.<sup>a,b</sup>

<i>Species</i>	<i>Phase</i>	<i>Energy (kcal/mol)</i>	<i>C-S</i>	<i>S-S</i>
<b>2-25</b>	Gas	-395618.20	1.81	--
	Water	-395626.00	1.79	--
<b>2-26</b>	Gas	-790479.64	1.80/1.79	2.13
	Water	-790487.28	1.81	2.08

<sup>a</sup>Bond lengths in Å; <sup>b</sup>See Figure A1.5

### A1.3.4: Geometry Optimizations of other Biologically Relevant Molecules

**Table A1.18:** Energies of other biologically relevant molecules determined by geometry optimization.

<i>Species</i>	<i>Phase</i>	<i>Energy (kcal/mol)</i>
<b>HO<sup>•</sup></b>	Gas	-47546.95
	Water	-47555.28
<b>HO<sup>•</sup></b>	Gas	-47587.29
	Water	-47679.52
<b>H<sub>2</sub>O<sub>2</sub></b>	Gas	-95128.80
	Water	-95143.20
<b>H<sub>2</sub>O</b>	Gas	-47976.17
	Water	-47985.63
<b>H<sub>3</sub>O<sup>+</sup></b>	Gas	-48140.08
	Water	-48230.14

### A1.4: Reaction Enthalpy Calculations

To determine the thermodynamics of a number of biologically relevant oxidation and reduction reactions, the sum of the thermal and electronic energies (ZPE corrected) were taken as  $\Delta G^\circ$  for each compound.

**Table A1.19:**  $\text{RSSR} + 2 \text{CH}_3\text{SH} \rightarrow 2 \text{RSH} + \text{H}_3\text{CSSCH}_3$

<i>Reactant</i>	<i>Product</i>	<i>Phase</i>	<i><math>\Delta G^\circ</math> (kcal/mol)</i>
<b>2-11</b>	<b>2 2-3</b>	Gas	-16.43
		Water	-30.32
	<b>2 2-4</b>	Gas	-1.00
		Water	-3.78
<b>2-12</b>	<b>2 2-3</b>	Gas	-16.19
		Water	ND <sup>a</sup>
	<b>2 2-5</b>	Gas	-0.32
		Water	ND
<b>2-13</b>	<b>2 2-3</b>	Gas	-13.12
		Water	-30.37
<b>2-18</b>	<b>2-4 + 2-5</b>	Gas	3.15
		Water	3.47
	<b>2 2-14</b>	Gas	0.03
		Water	2.18
<b>2 2-15<sup>b</sup></b>	Gas	363.11	
	Water	85.46	
<b>2 2-17</b>	Gas	15.84	
	Water	-6.40	

<sup>a</sup>ND = not determined; <sup>b</sup>Balanced equation becomes  $(\text{TMMI-S})_2 + 2\text{CH}_3\text{SH} + 2\text{H}_2\text{O} \rightarrow 2 \text{TMMI}^- + \text{H}_3\text{CSSCH}_3 + 2 \text{H}_3\text{O}^+$

**Table A1.20:**  $2 \text{TMI}_{\text{thiol}} \rightarrow 2 \text{TMI}_{\text{thione}}$ 

<i>Reactant</i>	<i>Product</i>	<i>Phase</i>	$\Delta G^\circ$ (kcal/mol)
2 2-4	2 2-3	Gas	-15.43
		Water	-26.53
2 2-5	2 2-3	Gas	-16.27
		Water	-27.06
2-4 + 2-5	2 2-3	Gas	-17.11
		Water	-26.80

**Table A1.21:**  $2 \text{TMMI} \rightarrow 2 \text{TMMI}_{\text{zwitterion}}$ 

<i>Reactant</i>	<i>Product</i>	<i>Phase</i>	$\Delta G^\circ$ (kcal/mol)
2 2-14	2 2-17	Gas	15.81
		Water	-8.58
2 2-15 <sup>a</sup>	2 2-17	Gas	-347.27
		Water	-91.85

<sup>a</sup>Balanced equation becomes  $2 \text{OS}^- + 2 \text{H}_3\text{O}^+ \rightarrow 2 \text{OSH}_{\text{zwitterion}} + 2 \text{H}_2\text{O}$

**Table A1.22:**  $\text{TMMI}_{\text{zwitterion}} + \text{OH}^- \rightarrow \text{TMMI}^- + \text{H}_2\text{O}$ 

<i>Reactant</i>	<i>Product</i>	<i>Phase</i>	$\Delta G^\circ$ (kcal/mol)
2-17	2-15	Gas	-51.33
		Water	-21.68



**Table A1.23:**  $2 \text{RSH} + \text{H}_2\text{O}_2 \rightarrow \text{RSSR} + 2 \text{H}_2\text{O}$ 

<i>Reactant</i>	<i>Product</i>	<i>Phase</i>	$\Delta G^\circ$ (kcal/mol)
2 2-4	2-11	Gas	-67.58
		Water	-63.17
2 2-5	2-12	Gas	-69.49
		Water	ND <sup>a</sup>
2-4 + 2-5	2-13	Gas	-71.73
		Water	-63.38
2 2-3	2-11	Gas	-52.15
		Water	-36.63
	2-12	Gas	-52.39
		Water	ND
	2-13	Gas	-55.46
		Water	-36.58
2 2-14	2-18	Gas	-68.61
		Water	-69.13
2 2-15 <sup>b</sup>	2-18	Gas	-431.69
		Water	-152.41
2 2-17	2-18	Gas	-84.42
		Water	-60.55
2 2-19	2-22	Gas	-68.58
		Water	-75.97

<sup>a</sup>ND = not determined; <sup>b</sup>Balanced equation becomes  $2 \text{TMMI} + 2 \text{H}_3\text{O}^+ + \text{H}_2\text{O}_2 \rightarrow (\text{TMMI-S})_2 + 4 \text{H}_2\text{O}$ .

**Table A1.24:**  $\text{RSH} + \text{HO}^\bullet \rightarrow \text{RS}^\bullet + \text{H}_2\text{O}$ 

<i>Reactant</i>	<i>Product</i>	<i>Phase</i>	$\Delta G^\circ$ (kcal/mol)
2-4	2-9	Gas	-45.80
		Water	-47.10
2-5	2-10	Gas	-46.53
		Water	-47.58
2-14	2-16	Gas	-44.35
		Water	-45.88
2-19	2-21	Gas	-35.35
		Water	-31.04

**Table A1.25:**  $RS^{\bullet} + HO^{\bullet} + H_3O^+ \rightarrow RS^{\bullet} + 2 H_2O$

<i>Reactant</i>	<i>Product</i>	<i>Phase</i>	$\Delta G^{\circ}$ (kcal/mol)
<b>2-7</b>	<b>2-9</b>	Gas	-215.40
		Water	-83.12
<b>2-8</b>	<b>2-10</b>	Gas	-216.71
		Water	-83.57
<b>2-15</b>	<b>2-16</b>	Gas	-225.89
		Water	-87.52
<b>2-20</b>	<b>2-21</b>	Gas	-224.04
		Water	-78.97

**Table A1.26:**  $2 RSH + 2 HO^{\bullet} \rightarrow RSSR + 2 H_2O$

<i>Reactant</i>	<i>Product</i>	<i>Phase</i>	$\Delta G^{\circ}$ (kcal/mol)
<b>2 2-3</b>	<b>2-11</b>	Gas	-87.04
		Water	-69.27
	<b>2-12</b>	Gas	-87.28
		Water	ND <sup>a</sup>
	<b>2-13</b>	Gas	-90.35
		Water	-69.21
<b>2 2-4</b>	<b>2-11</b>	Gas	-102.48
		Water	-95.80
<b>2 2-5</b>	<b>2-12</b>	Gas	-104.39
		Water	ND
<b>2-5 + 2-4</b>	<b>2-13</b>	Gas	-106.62
		Water	-72.85
<b>2-14</b>	<b>2-18</b>	Gas	-103.50
		Water	-101.77
<b>2-15<sup>b</sup></b>	<b>2-18</b>	Gas	-466.58
		Water	-185.04
<b>2-17</b>	<b>2-18</b>	Gas	-119.32
		Water	-93.19
<b>2-19</b>	<b>2-22</b>	Gas	-103.47
		Water	-99.58

<sup>a</sup>ND = not determined; <sup>a</sup>Balanced equation becomes  $2 TMMI^{\bullet} + 2 HO^{\bullet} + 2 H_3O^+ \rightarrow (TMMI-S)_2 + 4 H_2O$

Table A1.27:  $RS^{\bullet} + 2-23 \rightarrow RSH + 2-24$

<i>Reactant</i>	<i>Product</i>	<i>Phase</i>	$\Delta G^{\circ}$ (kcal/mol)
<b>2-9</b>	<b>2-3</b>	Gas	-14.11
		Water	-16.39
	<b>2-4</b>	Gas	-6.40
		Water	-3.12
<b>2-10</b>	<b>2-3</b>	Gas	-14.22
		Water	-16.17
	<b>2-5</b>	Gas	-5.67
		Water	-2.64
<b>2-16</b>	<b>2-14</b>	Gas	-7.85
		Water	-4.34
	<b>2-15<sup>a</sup></b>	Gas	173.69
		Water	37.3
	<b>2-17</b>	Gas	-163.86
		Water	-8.63
<b>2-21</b>	<b>2-20</b>	Gas	-19.85
		Water	-19.18

<sup>a</sup>Balanced equation modified to  $TMMI^{\bullet} + 2-23 + H_2O \rightarrow TMMI + 2-24 + H_3O^+$

### A1.5: Electron Density Determination

**Table A1.28:** Mullikan charges and spin densities for 2-thiol-4-methyl-imidazole and 4-thiol-*N*<sup>1</sup>-methyl-5-methylimidazole.

<i>Compound</i>	<i>Atom</i>	<i>Gas Phase</i>		<i>Water Phase</i>	
		<i>Charge</i>	<i>Spin Density</i>	<i>Charge</i>	<i>Spin Density</i>
<b>2-9</b>	N <sup>1</sup>	-0.127	0.031	-0.118	0.021
	N <sup>1</sup> - <u>H</u>	0.329	-0.001	0.423	-0.001
	C <sup>2</sup>	0.170	0.003	0.219	0.052
	N <sup>3</sup>	0.010	0.111	-0.119	0.069
	C <sup>4</sup>	0.068	0.001	0.050	0.072
	C <sup>5</sup>	-0.033	0.297	0.032	0.313
	C <sup>5</sup> - <u>H</u>	0.175	-0.019	0.269	-0.018
	C <sup>6</sup>	-0.572	-0.001	-0.651	-0.007
	C <sup>6</sup> - <u>H</u>	0.174	0.005	0.178	0.008
	C <sup>6</sup> - <u>H</u>	0.174	0.005	0.178	0.008
	C <sup>6</sup> - <u>H</u> (in plane)	0.136	-0.001	0.165	-0.001
	S	-0.503	0.571	-0.627	0.483
	<b>2-10</b>	N <sup>1</sup>	-0.042	0.137	-0.178
C <sup>2</sup>		0.126	0.003	0.182	0.053
N <sup>3</sup>		-0.064	0.012	-0.048	-0.001
N <sup>3</sup> - <u>H</u>		0.329	-0.001	0.426	0.000
C <sup>4</sup>		0.019	0.247	0.031	0.268
C <sup>5</sup>		0.169	0.023	0.134	0.090
C <sup>5</sup> - <u>H</u>		0.184	-0.005	0.261	-0.009
C <sup>6</sup>		-0.754	-0.017	-0.775	-0.015
C <sup>6</sup> - <u>H</u>		0.169	0.014	0.188	0.016
C <sup>6</sup> - <u>H</u>		0.169	0.014	0.188	0.016
C <sup>6</sup> - <u>H</u> (in plane)		0.168	0.000	0.183	0.000
S		-0.472	0.571	-0.592	0.486
<b>2-16</b>		N <sup>1</sup>	0.022	0.015	0.053
	C <sup>2</sup>	0.176	0.058	0.157	0.114
	C <sup>2</sup> - <u>H</u>	0.187	-0.006	0.284	-0.009
	N <sup>3</sup>	-0.022	0.000	-0.177	-0.029
	C <sup>4</sup>	0.029	0.064	0.072	0.105
	C <sup>5</sup>	0.275	0.182	0.325	0.204
	C <sup>6</sup>	-0.767	-0.018	-0.812	-0.018
	C <sup>6</sup> - <u>H</u>	0.153	0.012	0.184	0.014
	C <sup>6</sup> - <u>H</u>	0.154	0.012	0.185	0.014
	C <sup>6</sup> - <u>H</u> (in plane)	0.164	-0.008	0.160	-0.006
	C <sup>7</sup>	-0.299	-0.002	-0.320	-0.002
	C <sup>7</sup> - <u>H</u>	0.181	0.003	0.203	0.003
	C <sup>7</sup> - <u>H</u>	0.182	0.004	0.204	0.003
C <sup>7</sup> - <u>H</u> (in plane)	0.137	-0.001	0.164	-0.001	
S	-0.572	0.684	-0.681	0.602	

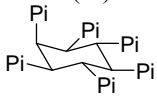
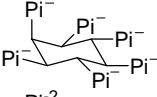
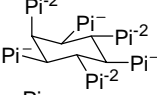
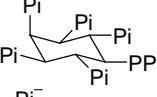
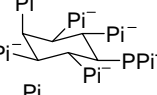
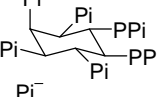
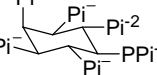
## Appendix 2: Phosphorylation by Inositol Pyrophosphates

Please see *Chapter 5* for a full introduction and discussion of this topic.

### A2.1: Hartree-Fock and Density Functional Techniques

All HF and DFT geometry optimizations and vibrational frequency calculations were performed using the 6-31++G\*\* basis set with and without implicit solvation, both as implemented by Jaguar version 6.0 (Schrödinger Inc.). All DFT calculations used the B3LYP density functional, also as implemented by Jaguar. Full procedural details can be found in *Chapter 5*.

**Table A2.1:**  $\Delta G^\circ_f$  values calculated using HF and DFT techniques

Structure	$\Delta G^\circ_f$ (kcal/mol)			
	HF		DFT	
	Gas Phase	Implicit Solvation	Gas Phase	Implicit Solvation
MeOH	-72176.1	-72195.6	-72604.7	-72623.6
MeOPi(0)	-427350.9	-427403.6	-428855.7	-428902.2
MeOPi(-1)	-427015.8	-427189.0	-428531.9	-428695.0
MeOPi(-2)	-426556.5	-427076.7	-428080.1	-428585.2
PPi(0)	-758049.1	-758138.2	-760456.9	-760531.5
PPi(-2)	-757333.1	-757760.4	-759755.9	-760194.5
PPi(-3)	-756804.1	-757722.7	-759239.4	-760134.3
PPi(-4)	-756177.8	-757773.6	-758621.3	-760185.7
Pi(0)	-402880.6	-402934.5	-404209.6	-404265.5
Pi(-1)	-402548.6	-402725.6	-403886.5	-404053.4
Pi(-2)	-402085.5	-402615.1	-403434.6	-403949.8
Pi(-3)	-401497.5	-402605.1	-402858.0	-403943.6
	-2559771.3	-2559960.8	-2568649.9	-2568800.5
	-2557088.7	-2559372.8	-2566006.2	-2568245.2
	-2554876.1	-2559865.4	ND <sup>a</sup>	-2568754.7
	-2914937.9	-2915133.7	-2924888.3	-2925054.0
	-2911681.7	-2914603.4	-2921684.2	-2924549.4
	-3270105.7	-32070302.8	-3281131.7	-3281272.3
	-3266223.3	ND	ND	ND

<sup>a</sup>ND = not determined

**Table A2.2:**  $\Delta G^{\circ}_{rxn}$  values for the phosphorylation of methanol by pyrophosphate as calculated using HF and DFT methods.

<i>Pi Donor</i> <sup>a</sup>	→	<i>Product</i>	$\Delta G^{\circ}_{rxn}$ <sup>c</sup>	<i>Pi Donor</i>	→	<i>Product</i>	$\Delta G^{\circ}_{rxn}$
<i>Forming MeOPi(-1)</i> <sup>b</sup>				<i>Forming MeOPi(-2)</i>			
<b>HF</b>							
<b>Gas Phase</b>							
PPi(-2) <sup>c</sup>	→	Pi(-1)	-55.2	PPi(-3)	→	Pi(-1)	-124.9
PPi(-3)	→	Pi(-2)	-121.11		→		
<b>Implicit Solvation</b>							
PPi(-2) <sup>c</sup>	→	Pi(-1)	41.4	PPi(-3)	→	Pi(-1)	116.1
PPi(-3)	→	Pi(-2)	114.2		→		
<b>DFT</b>							
<b>Gas Phase</b>							
PPi(-2) <sup>c</sup>	→	Pi(-1)	-57.7	PPi(-3)	→	Pi(-1)	-122.4
PPi(-3)	→	Pi(-2)	-122.4		→		
<b>Implicit Solvation</b>							
PPi(-2) <sup>c</sup>	→	Pi(-1)	59.7	PPi(-3)	→	Pi(-1)	119.2
PPi(-3)	→	Pi(-2)	113.0		→		

<sup>a</sup>Pi = phosphate; <sup>b</sup>MeOPi = methyl phosphate; <sup>c</sup>in kcal/mol

**Table A2.3:** Calculated  $\Delta G^{\circ}_{rxn}$  values determined for the phosphorylation of methanol by InsPPs using HF and DFT techniques.

<i>Pi Donor</i> <sup>b</sup>	→	<i>Product</i>	$\Delta G^{\circ}_{rxn}$ <sup>a</sup> (kcal/mol)			
			<i>HF</i>		<i>DFT</i>	
			<i>Gas Phase</i>	<i>Implicit Solvation</i>	<i>Gas Phase</i>	<i>Implicit Solvation</i>
<i>Forming MeOPi(-1)</i> <sup>c</sup>						
	→		-8.2	-35.1	-12.6	-25.1
	→		-246.6	237.2	-249.1	232.8
	→		-6.9	-38.9	-7.5	-60.2
	→		-298.1	ND <sub>a</sub>	ND	ND

<sup>a</sup>in kcal/mol; <sup>b</sup>Pi = phosphate; <sup>c</sup>MeOPi = methyl phosphate; <sup>d</sup>ND = Not determined

## A2.2: Molecular Mechanics Methods

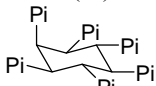
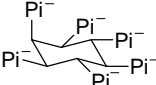
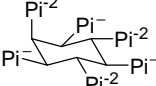
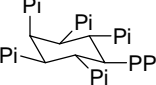
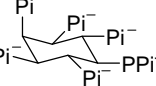
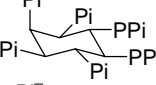
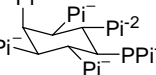
All molecular mechanics calculations were performed using the OPLS2003 force field and the GB/SA solvation model as implemented by MacroModel (Schrödinger Inc.). Full details can be found in *Chapter 5*.

**Table A2.4:** The quality ratings of the parameters of the force fields available in MacroModel.

<i>Force Field</i>	<i>Stretch</i>			<i>Bend</i>			<i>Torsion</i>		
	<i>High</i>	<i>Med</i>	<i>Low</i>	<i>High</i>	<i>Med</i>	<i>Low</i>	<i>High</i>	<i>Med</i>	<i>Low</i>
MM2*	50	28	0	114	0	42	106	36	92
MM3*	-- <sup>a</sup>	--	--	--	--	--	--	--	-
MMFF	24	28	0	36	54	1	54	42	18
MMFFs	24	28	0	36	54	1	52	42	18
AMBER*	46	6	0	90	0	1	114	0	0
OPLS	--	--	--	--	--	--	--	--	--
AMBER94	--	--	--	--	--	--	--	--	--
OPLS2001	52	0	0	91	0	0	90	0	0
OPLS2003	52	0	0	91	0	0	90	0	6 <sup>b</sup>

<sup>a</sup>Missing force field constants; <sup>b</sup>Personal communications with Schrödinger Inc. indicated that these parameters, though rated low, were reasonable values.

**Table A2.5:** Calculated  $\Delta H_f$  values determined using molecular mechanics methods.

<i>Structure</i>	<i><math>\Delta H_f</math> (kcal/mol)</i>	
	<i>Gas</i>	<i>Implicit Solvation</i>
MeOH <sup>a</sup>	3.4	-2.2
MeOPi(0) <sup>b</sup>	-66.3	-82.5
MeOPi(-1)	-79.8	-169.2
MeOPi(-2)	-40.0	-320.9
PPi(0) <sup>c</sup>	-63.8	-92.1
PPi(-2)	74.8	-168.4
PPi(-3)	167.4	-312.4
PPi(-4)	302.7	-539.9
Pi(0) <sup>d</sup>	-72.3	-91.4
Pi(-1)	-104.4	-194.3
Pi(-2)	-80.2	-355.9
Pi(-3)	29.4	-569.0
	-532.1	580.4
	40.8	-1101.8
	1038.4	-1535.5
	-516.6	-583.1
	388.4	-1088.3
	-525.5	-592.6
	782.0	-1047.2

<sup>a</sup>MeOH = methanol; <sup>b</sup>MeOPi = methyl phosphate; <sup>c</sup>PPi = pyrophosphate; <sup>d</sup>Pi = phosphate.

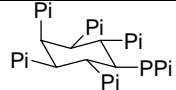
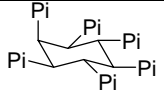
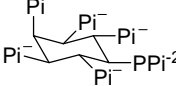
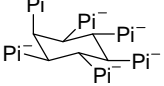
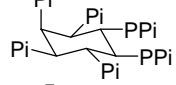
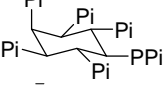
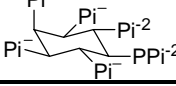
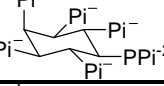


**Table A2.6:** Calculated  $\Delta H_{\text{rxn}}$  values for the phosphorylation of methanol by pyrophosphate using molecular mechanics methods

<i>Pi Donor</i> <sup>a</sup>	→	<i>Product</i>	$\Delta H_{\text{rxn}}$ Values (kcal/mol)	
			<i>Gas Phase</i>	<i>Implicit Solvation</i>
<b>Forming MeOPi(-1)<sup>b</sup></b>				
PPi(-2) <sup>c</sup>	→	Pi(-1)	-262.4	-193.0
PPi(-3)	→	Pi(-2)	-330.8	-210.6
<b>Forming MeOPi(-2)</b>				
PPi(-3)	→	Pi(-1)	-315.2	-200.7

<sup>a</sup>Pi = phosphate; <sup>b</sup>MeOPi = methyl phosphate; <sup>c</sup>PPi = pyrophosphate

**Table A2.7:** Calculated  $\Delta H_{\text{rxn}}$  values for the phosphorylation of methanol by the InsPPs as determined by molecular mechanics methods

<i>Pi Donor</i> <sup>a</sup>	→	<i>Product</i>	$\Delta H_{\text{rxn}}$ Values (kcal/mol)	
			<i>Gas Phase</i>	<i>Implicit Solvation</i>
<b>Forming MeOPi(-1)<sup>b</sup></b>				
	→		-84.9	-77.3
	→		-430.8	-180.5
	→		-60.9	-70.9
	→		-476.5	-208.2

<sup>a</sup>Pi = phosphate; <sup>b</sup>MeOPi = methyl phosphate;

### A2.3: Results from Semi-Empirical Methods

All final geometry optimizations and vibrational frequency calculations were performed using the PM3 semi-empirical method<sup>58, 59</sup> with SM5.2 implicit water solvation as implemented by AMPAC 8 and the AMSOL module<sup>60</sup> (Semichem Inc. Kansas City). Full details can be found in *Chapter 5*.

**Table A2.8:** Calculated  $\Delta H_f$  values determined using semi-empirical techniques for the various  $\text{InSP}_6$  charged states studied.

<i>Structure</i>	$\Delta H_f$ (kcal/mol)	<i>Structure</i>	$\Delta H_f$ (kcal/mol)	<i>Structure</i>	$\Delta H_f$ (kcal/mol)
	-1789.9		-2042.5		-2125.3
	-1856.0		-2042.6		-2123.7
	-1843.9		-2085.7		-2122.1
	-1900.8		-2083.4		-2154.3
	-1996.1		-2083.4		
	-2038.3		-2118.8		

**Table A2.9:**  $\Delta H_f$  values calculated for the various 5-PP-InsP<sub>3</sub> charged states studied.

<i>Structure</i>	$\Delta H_f$ (kcal/mol)	<i>Structure</i>	$\Delta H_f$ (kcal/mol)	<i>Structure</i>	$\Delta H_f$ (kcal/mol)
	-2073.8		-2275.4		-2361.1
	-2138.3		-2314.0		-2352.2
	-2115.2		-2312.5		-2391.9
	-2131.3		-2321.3		-2392.1
	-2179.9		-2320.7		-2399.8
	-2180.8		-2360.6		-2398.7
	-2184.1		-2356.3		-2397.5
	-2176.3		-2363.1		-2393.2
	-2222.6		-2358.7		-2427.7
	-2217.6		-2358.5		-2437.9

**Table A2.10:**  $\Delta H_f$  values calculated for the various 5, 6-bis-PP-InsP<sub>4</sub> charged states studied.

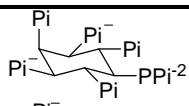
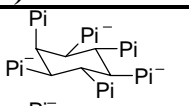
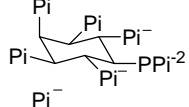
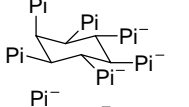
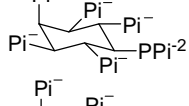
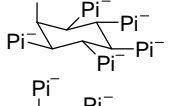
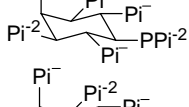
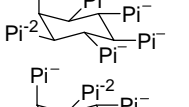
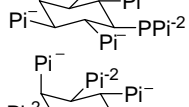
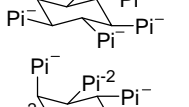
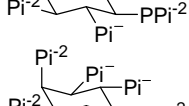
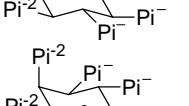
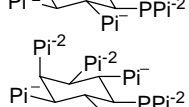
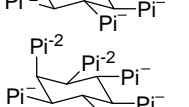
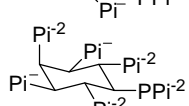
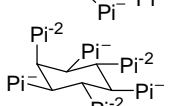
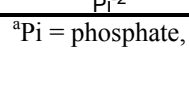
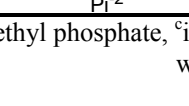
<i>Structure</i>	$\Delta H_f$ (kcal/mol)	<i>Structure</i>	$\Delta H_f$ (kcal/mol)	<i>Structure</i>	$\Delta H_f$ (kcal/mol)
	-2415.9		-2584.7		-2673.0
	-2419.3		-2627.6		-2672.4
	-2470.6		-2626.3		-2653.5
	-2461.6		-2635.5		-2653.7
	-2471.5		-2633.6		-2660.6
	-2465.9		-2634.8		-2661.1
	-2522.2		-2633.7		-2653.6
	-2507.9		-2632.8		-2672.0
	-2552.0		-2626.1		-2701.1
	-2589.3		-2615.3		-2708.2
	-2588.4		-2664.1		-2694.7
	-2597.0		-2663.9		-2687.42

**Table A2.11:**  $\Delta H_{rxn}$  values determined for the phosphorylation of methanol by various pyrophosphates.

<i>Pi Donor</i> <sup>a</sup> <i>Forming MeOPi(-1)</i> <sup>b</sup>	→	<i>Product</i>	$\Delta H_f^c$	<i>Pi Donor</i> <i>Forming MeOPi(-2)</i>	→	<i>Product</i>	$\Delta H_f$
PPi(-2) <sup>d</sup>	→	Pi(-1)	-8.9	PPi(-3)	→	Pi(-1)	-13.9
PPi(-3)	→	Pi(-2)	-12.9				
MeOPPi(-2) <sup>e</sup>	→	MeOPi(-1)	-7.2	MeOPi(-3)	→	MeOPi(-1)	-15.2
Cyclohex-PPi(-2) <sup>f</sup>	→	Cyclohex-Pi(-1) <sup>g</sup>	-7.3	Cyclohex-PPi(-2)	→	Cyclohex-Pi(-1)	-14.8
Cyclohex-PPi(-3)	→	Cyclohex-Pi(-2)	-14.3				
InsPPi(-2) <sup>h</sup>	→	InsPi(-1) <sup>i</sup>	-2.0	InsPPi(-3)	→	InsPi(-2)	-6.2
InsPPi(-3)	→	InsPi(-2)	-14.8				

<sup>a</sup>Pi = phosphate; <sup>b</sup>MeOPi = methyl phosphate; <sup>c</sup> $\Delta H_f$  in kcal/mol; <sup>d</sup>PPi = pyrophosphate; <sup>e</sup>MeOPPi = methyl pyrophosphate; <sup>f</sup>Cyclohex-PPi = cyclohexyl pyrophosphate; <sup>g</sup>Cyclohex-Pi = cyclohexyl phosphate; <sup>h</sup>InsPPi = *myo*-inositol-5-pyrophosphate; <sup>i</sup>InsPi = *myo*-inositol-5-phosphate (Ins(5)P<sub>1</sub>).

**Table A2.12:**  $\Delta H_{rxn}$  values calculated for the phosphorylation of methanol by 5-PP-InsP<sub>5</sub>.

<i>Pi<sup>a</sup> Donor</i> <i>Forming MeOPi(-1)</i> <sup>b</sup>	→	<i>Product</i>	$\Delta H_{rxn}^c$	<i>Pi Donor</i> <i>Forming MeOPi(-2)</i>	→	<i>Product</i>	$\Delta H_{rxn}$
	→		-9.8		→		ND <sup>d</sup>
	→		-11.3		→		ND
	→		-14.4		→		ND
	→		-14.9		→		ND
	→		-15.7		→		ND
	→		-16.3		→		ND
	→		-15.9		→		ND
	→		-24.9		→		ND
	→		-20.4		→		ND

<sup>a</sup>Pi = phosphate, <sup>b</sup>MeOPi = methyl phosphate, <sup>c</sup>in kcal/mol, <sup>d</sup>Not determined as inappropriate charge states would result

**Table A2.13:**  $\Delta H_{rxn}$  values calculated for the phosphorylation of methanol by 5-PP-InsP<sub>5</sub>.

<i>Pi<sup>d</sup></i> Donor <i>Forming MeOPi(-1)<sup>b</sup></i>	→	<i>Product</i>	$\Delta H_{rxn}^c$	<i>Pi</i> Donor <i>Forming MeOPi(-2)</i>	→	<i>Product</i>	$\Delta H_{rxn}$
	→		-18.0		→		-18.3
	→		-14.5		→		-24.1
	→		-18.0		→		-30.1
	→		-19.2		→		-31.8
	→		-19.9		→		-32.1
	→		-22.6		→		-33.9
	→		-20.3		→		-33.9
					→		-38.3
					→		-39.2

<sup>a</sup>Pi = phosphate, <sup>b</sup>MeOPi = methyl phosphate, <sup>c</sup>in kcal/mol

**Table A2.14:**  $\Delta H_{rxn}$  values calculated for the phosphorylation of methanol by 5, 6-bis-PP-InsP<sub>5</sub> (Part A).

<i>Pi<sup>d</sup></i> Donor <i>Forming MeOPi(-1)<sup>b</sup></i>	→	<i>Product</i>	$\Delta H_{rxn}$ <sup>c</sup>	<i>Pi Donor</i> <i>Forming MeOPi(-2)</i>	→	<i>Product</i>	$\Delta H_{rxn}$
	→		-16.1		→		ND <sup>d</sup>
	→		-5.6		→		ND
	→		-17.1		→		ND
	→		-17.9		→		ND
	→		-29.7		→		ND
	→		-23.7		→		ND
	→		-21.3		→		ND
	→		-28.8		→		ND
	→		-30.6		→		ND

<sup>a</sup>Pi = phosphate, <sup>b</sup>MeOPi = methyl phosphate, <sup>c</sup>in kcal/mol, <sup>d</sup>Not determined as inappropriate charges states would result.

**Table A2.15:**  $\Delta H_{rxn}$  values calculated for the phosphorylation of methanol by 5, 6-bis-PP-InsP<sub>5</sub> (Part B).

<i>Pi<sup>d</sup></i> Donor Forming MeOPi(-1) <sup>b</sup>	→	Product	$\Delta H_{rxn}$ <sup>c</sup>	<i>Pi</i> Donor Forming MeOPi(-2)	→	Product	$\Delta H_{rxn}$
	→		-3.1		→		-15.8
	→		-3.9		→		-7.8
	→		1.5		→		-34.1
	→		-6.3		→		-35.8
	→		-16.9		→		-33.9
	→		-18.4		→		-40.2
	→		-19.5		→		-38.1
	→		-17.6		→		-46.6
	→		-21.7		→		-48.0
	→		-21.5		→		-53.8
	→		-19.4		→		-50.8
	→		-20.5				
	→		-33.2				
	→		-30.1				

<sup>a</sup>Pi = phosphate, <sup>b</sup>MeOPi = methyl phosphate, <sup>c</sup>in kcal/mol.



**Table A2.16:**  $\Delta H_{rxn}$  values calculated for the phosphorylation of methanol by 5, 6-bis-PP-InsP<sub>5</sub> (Part C).

<i>Pi<sup>a</sup></i> Donor <i>Forming MeOPi(-1)<sup>b</sup></i>	→	<i>Product</i>	$\Delta H_{rxn}^c$	<i>Pi</i> Donor <i>Forming MeOPi(-2)</i>	→	<i>Product</i>	$\Delta H_{rxn}$
	→		-12.8		→		-24.7
	→		-12.0		→		-13.4
	→		-5.5		→		-35.0
	→		-11.9		→		-35.7
	→		-17.8		→		-35.0
	→		-19.3		→		-40.5
	→		-22.0		→		-38.7
	→		-21.7		→		-48.5
	→		-20.3		→		-28.2
	→		-21.1				

<sup>a</sup>Pi = phosphate, <sup>b</sup>MeOPi = methyl phosphate, <sup>c</sup>in kcal/mol

**TableA2.17:**  $\Delta H_{\text{rxn}}$  values calculated for the phosphorylation of methanol by 5, 6-bis-PP-InsP<sub>5</sub> (Part D).

<b>Pi<sup>a</sup> Donor</b>	<b>→</b>	<b>Product</b>	<b><math>\Delta H_{\text{rxn}}^c</math></b>	<b>Pi Donor</b>	<b>→</b>	<b>Product</b>	<b><math>\Delta H_{\text{rxn}}</math></b>
<b>Forming MeOPi(-1)<sup>b</sup></b>				<b>Forming MeOPi(-2)</b>			
	→		-8.3		→		-20.0
	→		-7.4		→		-24.0
	→		-26.7		→		-24.2
	→		-20.3		→		-33.0
	→		-23.3		→		-34.5
					→		-38.8
					→		-39.1
					→		-38.52
					→		-39.6
					→		-52.9
					→		-53.2
					→		-53.8
					→		-50.8

<sup>a</sup>Pi = phosphate, <sup>b</sup>MeOPi = methyl phosphate, <sup>c</sup>in kcal/mol.

**Table A2.18:**  $\Delta H_{\text{rxn}}$  values calculated for the phosphorylation of methanol by ATP, ADP and model compounds.

<b>Pi<sup>a</sup> Donor</b>	<b>→</b>	<b>Product</b>	<b><math>\Delta H_{\text{rxn}}</math><sup>c</sup></b>	<b>Pi Donor</b>	<b>→</b>	<b>Product</b>	<b><math>\Delta H_{\text{rxn}}</math></b>
<b>Forming MeOPi(-1)<sup>b</sup></b>				<b>Forming MeOPi(-2)</b>			
ATP(-3) <sup>c</sup>	→	ADP(-2) <sup>d</sup>	-2.2	ATP(-4)	→	ADP(-2)	-10.7
ATP(-4)	→	ADP(-3)	-11.8				
ADP(-2)	→	AMP(-1) <sup>e</sup>	-5.4	ADP(-3)	→	AMP(-1)	-4.3
ADP(-3)	→	AMP(-2)	-7.4				
MeOPPPi(-3) <sup>f</sup>	→	MeOPPi(-2) <sup>g</sup>	-10.0	MeOPPPi(-4)	→	MeOPPi(-2)	-22.3
MeOPPPi(-4)	→	MeOPPi(-3)	-14.3				
PPPi(-3) <sup>h</sup>	→	PPi(-2) <sup>i</sup>	-7.2	PPPi(-4)	→	PPi(-2)	-18.1
PPPi(-4)	→	PPi(-3)	-13.0				

<sup>a</sup>Pi = phosphate; <sup>b</sup>MeOPi = methyl phosphate; <sup>c</sup>in kcal/mol