Investigation of the Microbial Glyoxalase System

by

Uthaiwan Suttisansanee

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, 2011

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.

Uthaiwan Suttisansanee

ABSTRACT

The Glyoxalase system is composed of two metalloenzymes, Glyoxalase I and Glyoxalase II, that catalyze the conversion of toxic, metabolically produced α-ketoaldehydes, such as methyglyoxal, in the presence of a thiol cofactor, such as glutathione, into their corresponding nontoxic 2-hydroxycarboxylic acids, leading to detoxification of these cellular metabolites. Previous studies on the first enzyme in the Glyoxalase system, Glyoxalase I (GlxI), in yeast, protozoa, animals, human, plants and Gram-negative bacteria suggest two metal activation classes, Zn^{2+}-activation or non-Zn^{2+}-activation (but exhibiting selective Ni^{2+}/Co^{2+}-activation). However, there is a lack of knowledge concerning the Glyoxalase system from Gram-positive bacteria. Biochemical investigation on the metal specificity of *Clostridium acetobutylicum* GlxI suggests this enzyme is Ni^{2+}/Co^{2+}-activated and is inactive in the presence of Zn^{2+} ion. The active site geometry of this Ni^{2+}-activated enzyme possesses an octahedral coordination (PDB: 3HDP), while it becomes trigonal bipyramidal in its Zn^{2+}-bound inactive form (PDB: 2QH0). However, instead of forming an active site at the homodimeric interface similar to other dimeric enzymes in this family, a dimeric *C. acetobutylicum* GlxI forms an active site *within one subunit* (two active sites per dimer). Clearly various strategies are utilized in nature to provide the key active site structure for these enzymes.

The unique dimeric arrangement of *C. acetobutylicum* GlxI closely resembles the structure of methylmalonyl-CoA epimerase (MMCE), a metalloenzyme within the same βαββββ superfamily. Thus, close structural relatedness suggests that these two enzymes might be able to cross-function and possess enzymatic activity of the other one. However, experimental evidence suggests that GlxI exhibits no MMCE activity under a variety of experimental conditions. Analysis of the superimposed X-ray crystallographic structures also suggests that MMCE might not exhibit GlxI activity due to steric interference with the hemithioacetal, the GlxI substrate, at the entrance of the MMCE's active site.

From a similar perspective, the activity of fosfomycin resistance protein (FosA) activity from the same βαββββ superfamily was investigated in GlxI in order to understand the relatedness in terms of evolution of bacterial drug resistance mechanisms within the superfamily. However, GlxI exhibits no FosA activity and the structural investigation of these two metalloenzymes suggests that the variations in metal and substrate binding sites are large

such that the evolution of FosA has undergone several accumulated point mutations in order to develop a new biological function.

Interestingly, the metal activation classes of GlxI can be quite reliably predicted by the length of the amino acid sequence of the protein. The Zn^{2+}-activated enzymes possess a longer amino acid sequence than that of the enzymes in the Ni^{2+}/Co^{2+}-activation class. The analysis of the multiple sequence alignment and X-ray crystallographic structures suggests that the N-terminal arm, a long α-helix loop and at least two small loops that only exist in the Zn^{2+}-activated GlxI might be responsible for its metal specificity. However, the N-terminal arm was also reported to be significant for protein stability of the enzyme and its deletion would cause protein aggregation. The deletional mutagenesis of the long α-helix loop and two small loops on the Zn^{2+}-activated GloA3 from *Pseudomonas aeruginosa* and the enzymatic activity of these muteins suggests that the long α-helix is crucial and a major contributor to metal dependence in GlxI, while the two small loops affect the magnitude of the enzymatic activity.

The reaction mechanism of the Ni^{2+}/Co^{2+}-activated GlxI was also investigated in relatedness to the enzymatic flexibility upon binding of the transition state analogues, *S*-(*N*-hydroxy-*N*-methylcarbamoyl)glutathione and *S*-(*N*-hydroxy-*N*-phenylcarbamoyl)glutathione, which were synthesized as previously described. The inhibition studies suggested that the hydrophobicity of these inhibitors plays a significant role that affects the binding affinity in the catalytic pocket of GlxI. As well, the structural comparison of GlxI with and without bound inhibitor suggests that the loop movements of the enzyme act as the lids to cover and shield the active site from outer solvent as well as stabilize the GSH moiety upon binding of the inhibitor. The reaction mechanism of the $Ni^{2+/}Co^{2+}$-activated GlxI is proposed to be similar to that of the Zn^{2+}-activated enzyme, where one of the metal binding residues, glutamate, could act as a catalytic base to abstract and transfer the significant protons to form the product of the GlxI reaction, *S*-D-lactoylglutathione.

Some Actinomycetales and Streptomycetes produce a unique mycothiol as their major thiol instead of glutathione that is produced widely in eukaryotes, most Gram-negative bacteria and some Gram-positive bacteria. The Glyoxalase system using this different type of thiol cofactor, mycothiol, or mycoGlyoxlase system was hypothesized to possess a parallel pathway as the glutathione-dependent Glyoxalase system. The cofactor of the mycoGlyoxlase system, mycothiol, was isolated from non-pathogenic species, *Streptomyces jumonjinesis,* due to the

lack of commercial availability and its truncated des-*myo*-inositol mycothiol was synthesized based on previous reports. The investigation of the putative enzymes in the mycoGlyoxalase system from Gram-positive bacterium, *Streptomyces coelicolor* reveals three genes with annotations as putative dioxygenase, putative lyase and putative hydrolase. Based on sequence analysis, it was suggested that the putative dioxygenase and the putative lyase might exhibit mycoGlyoxalase I activity, while the putative hydrolase might behave as a mycoGlyoxalase II. An experimental investigation of the Glyoxalase I activity of the putative dioxygenase has indicated that this enzyme does in fact exhibit mycoGlyoxalase I activity, while the putative lyase might possess other functional bioactivity in the cells. Due to the difficulty of protein expression and purification as well as the lack of availability of the substrate mycothiol, the enzymatic activity of the putative hydrolase as a mycoGlyoxalase II requires more experimental research.

This thesis, thus, provides the key discoveries of the Glyoxalase system from Gram-positive microorganisms using the major thiol cofactor/cosubstrate that produced within that particular organisms as well as the relatedness of the proteins in the same βαβββ protein superfamily.

ACKNOWLEDGEMENTS

First of all, I would like to express my sincere thank to Dr. John Honek, my supervisor, for giving me the opportunity to continue my graduate study in his laboratory and work on the interesting projects. I really appreciate his guidance, advice, inspiration, time, support, patience and most of all, understanding. John, thank you very, very much. You will be my role model once I work.

I am also grateful for the advice to contribute the works in this thesis from the committee members, Dr. Guy Guillemette, Dr. Michael Palmer and Dr. Richard Manderville. As well, I owe my gratitude to Dr. Zhending Su, Dr. Elisabeth Daub, Dr. Nicole Sukdeo, Dr. Christine Hand and Kadia Mullings for sharing their knowledge and experiences on the Glyoxalase system and organic synthesis. Dr. Yanhong Ran, the visiting scientist from Jinan University (Guangdong, China) also kindly helped me on the FosA adduct synthesis and its assay. Special thanks go to Dr. Richard Smith for electrospray ionization MS analysis and Jan Venne for NMR experiments.

I would like to acknowledge our collaborators, Dr. Satyanarayana Lagishetty, Dr. Krishnamurthy N. Rao, Dr. Subramanyam Swaminathan, Dr. J. Michael Sauder and Dr. Stephen K. Burley from the New York Structural Genomics Research Consortium (NYSGXRC) for providing the DNA and protein of *Clostridium acetobutylicum* Glyoxalase I and performing the X-ray crystallographic analysis of Zn^{2+}- and Ni^{2+}-bound *C. acetobutylicum* Glyoxalase I with and without inhibitors (*S*-(*N*-hydroxy-*N*-methylcarbamoyl)glutathione and *S*-(*N*-hydroxy-*N*-phenylcarbamoyl)glutathione). The analysis of the X-ray crystallographic structures of *Pseudomonas aeruginosa* GloA2 and GloA3 were collaborated with Dr. Charles Bond from the University of Western Australia (WA, Australia). As well, the X-ray crystallographic structures of *E. coli* GlxI with bound hydroxamate inhibitor (TSI) with the resolution of 2.1 Å and with bound *N*-hydroxybenzamide (BHO) with the resolution of 1.6 Å were prepared by Dr. Susan Clugston (the previous Ph. D. student in our laboratory) as well as Dr. Molly He and Dr. Brian Matthews from University of Oregon (Eugene, Oregon). I also want to show my gratitude to Dr. Thomas Bobik from Iowa State University (IA), who kindly

provided the bacterial strain *E. coli* BL21 (DE3) CodonPlus RIL containing the plasmid pTA925 expression vector with inserted methylmalonyl-CoA mutase (MCM) coding gene and Dr. Gerald Newton from University of California (SanDiego, La Jolla, CA) for providing a disulfide form of mycothiol.

This thesis won't be completed without the affordable works from all the NSERC and forth year undergraduate students. I would like to thank David Ward, Kelvin Lau, Michele Cossette and Vincent Azhikannickal on their devoting time and hard work. The working experiences with you guys are valuable.

A friendly environment in the work place is warmly created by all the friends in the department and the Honek lab members, Dr. Christine Hand, Dr. Nicole Sukdeo, Dr. Pei Hang, Paula Walasek, Kadia Mullings, Meijun Lu, Danish Khan, Ronald Zahoruk, Ignace Moya and Cullen Myers. Nothing else would fit except "you guys are the best".

I also wish to acknowledge the Natural Sciences and Engineering Research Council of Canada (NSERC), University of Waterloo (Ontario, Canada) and the Government of Thailand for financial support during my graduate studies.

Finally, I would like to thank my family, Sacorn, Jewhiang, P'Poo, P'Pla, P'Tor and my cute little niece, Apo, for mental support and, actually, everything. Thanks for always be there.

And thank you, Glyoxalase I, for your existence. You are fun to work with.

DEDICATION

For

His Majesty King Bhumibol Adulyadej, Rama IX of Thailand

The father who gives

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

CHAPTER 1

CHAPTER 2

CHAPTER 3

CHAPTER 4

CHAPTER 6

CHAPTER 7

CHAPTER 8

CHAPTER 9

APPENDIX 1

APPENDIX 2

APPENDIX 3

APPENDIX 5

APPENDIX 6

APPENDIX 7

LIST OF TABLES

Page

CHAPTER 2

CHAPTER 3

CHAPTER 4

CHAPTER 5

CHAPTER 6

CHAPTER 7

APPENDIX 3

APPENDIX 4

LIST OF ABBREVIATIONS

ACN	Acetonitrile
ACTU	2-Chloro-1,1,3,3-tetramethyluronium hexachloroantimonate
Amp	Ampicillin
BAT	Bleomycin *N*-acetyltransferase
B-GSH	Benzyl-glutathione
BM	Bleomycin
BRP	Bleomycin resistance protein
BSA	Bovine serum albumin
CD	Circular dichroism
CLO GlxI	*Clostridium acetobutylicum* Glyoxalase I
Cp	Heat capacity
DCC	*N,N′*-dicyclohexylcarbodiimide
DIC	*N,N′*-diisopropylcarbodiimide
DIEA	Diisopropylethylamine
DPA	Dipicolinic acid
DSC	Differential scanning calorimetric
DTNB	5,5'-Dithiobis-(2-nitrobenzoic acid)
DTNP	Disulfide 2,2’-dithiobis(5-nitropyridine)
DTT	Dithiothreitol
EDTA	Ethylenediamine tetraacetic acid
ESI-MS	Electrospray ionization mass spectrometry
FosA	Fosfomycin resistance protein A
FM	Fosfomycin
FPLC	Fast peptide and protein liquid chromatography
GlcN-HCl	Glucosamine hydrochloride
GlxI	Glyoxalase I
GlxII	Glyoxalase II
GSH	Glutathione
HBTU	*O*-(benzotriazol-1-yl)-*N,N,N′,N′*-tetramethyluronium hexafluorophosphate

HIPC-GSH	*S*-(*N*-hydroxy-*N*-*p*-iodophenylcarbamoyl)-glutathione
HOBt	1-Hydroxybenzotriazole
HPLC	High performance liquid chromatography
ICP-MS	Inductively coupled plasma mass spectrometry
IPTG	Isopropyl-β-D-thiogalactopyranoside
Kan	Kanamycin
K_d	Dissociation constant
KPB	Potassium phosphate buffer
LB	Luria Bertani media
MC	Mitomycin C
mcGlx	MycoGlyoxalase
MG	Methylglyoxal
MCM	Methylmalonyl CoA mutase
MMCE	Methylmalonyl CoA epimerase
MRP	Mitomycin C resistance protein
MW	Molecular weight
NBC-GSH	*S*-*p*-nitrobenzyloxycarbonyl-glutathione
OD	Optical density
PAR	4-(2-Pyridylazo)resorcinol
PCR	Polymerase chain reaction
PDB	Protein data bank
PDO	Putative dioxygenase
PHL	Putative hydrolase
p*I*	Isoelectric point
PLA	Putative lyase
PMSF	Phenylmethanesulfonylfluoride
PS-HOBt(HL)	1-Hydroxybenzotriazole-6-sulfonamidomethyl polystyrene
Q-TOF	Quadrupole time of flight
r.m.s.d.	Root mean square deviation
SDL	*S*-D-lactoylglutathione
SDS PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis

TB	Tuberculosis
TCEP	Tris(2-carboxyethyl)phosphine
TFA	Trifluoroacetic acid
T_m	Denaturing temperature
TNTU	2-(5-Norbornene-2,3-dicarboximido)-1,1,3,3-tetramethyluronium tetrafluoroborate

Note: The standard three letter and one letter amino acid codes as well as the one letter DNA base codes have been used throughout this report.

CHAPTER 1: INTRODUCTION

1.1: Methylglyoxal

Methylglyoxal (MG), a highly reactive electrophilic α-ketoaldehyde, is a cytotoxic compound that is produced by a number of different enzymatic and non-enzymatic reactions and pathways (Figure 1.1) [2-5]. Among these, MG can be produced in cells by a side reaction of the glycolytic enzyme triose phosphate isomerase (TIM). It has been estimated that for every 100,000 turnovers of dihydroxyacetone phosphate (DHAP) to glyceraldehye-3-phosphate (G3P), a molecule of MG is formed. This occurs when deprotonation of dihydroxyacetone phosphate is followed by elimination of the phosphate group instead of reprotonation at an alternate carbon as in the "regular" mechanism of triose phosphate isomerase [3, 4]. Additionally, in some microorganisms the enzyme MG synthase uses DHAP as a substrate and converts it to MG [6]. This amounts to a glycolytic pathway "bypass", which is believed to be important for microorganisms growing under low phosphate conditions [2, 6]. Another enzymatic pathway for MG production is from the threonine dehydrogenase catalyzed conversion of aminoacetone to MG.

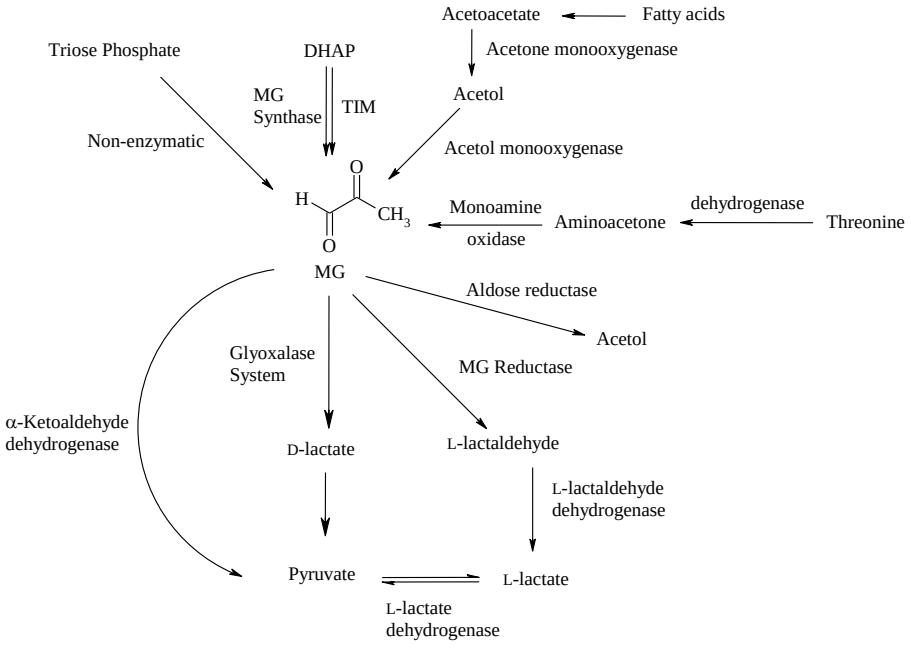


Figure 1.1: The overview of the biosynthesis and catabolism pathways of methylglyoxal based on the previous report by Kalapos, 1999 and Desai *et al*., 2010 [3, 4].

MG can be produced to an intracellular level as high as 0.4 mM/cell/day, which has been determined to lead to protein synthesis inhibition, adduct formation with proteins, DNA and RNA, and can promote Advanced Glycation End-products (AGEs) [7-9]. Glycation is a naturally occurring biochemical reaction in all tissues and body fluids, which has been found to impair the function of biomolecules. The early stage of the protein glycation reaction is a non-enzymatic reaction of the biomolecule with sugar molecules such as glucose and fructose glucose [7-9]. Adducts of sugar molecules (i.e., fructosyl-lysine and/or the attachment of fructosamine to the N-terminal amino acid residues) are then slowly converted to stable AGEs. MG, on the other hand, is highly reactive toward the free amino groups of arginine and lysine as well as the thiol moiety of cysteine residues in proteins, thus the formation of AGEs in the presence of MG can occur rapidly [7-9]. MG can also elevate AGEs formation in patients with diabetes due to the patient's high and variable cellular glucose levels [10, 11]. As well, AGEs have been shown to affect patients with Alzheimer's disease since the side products of the intermediate step in AGEs formation have been reported to produce β-amyloid proteins, which can contribute amyloid plaques in the brains of the patients [7, 12].

Since MG is a cytotoxic compound, most organisms develop several detoxification pathways as self-defense mechanisms that convert MG into non-toxic compounds (Figure 1.1). These pathways include the reactions of aldehyde reductase, α-ketoaldehyde dehydrogenase and the Glyoxalase system [13, 14]. Aldose reductase converts MG to hydroxyacetone and D-lactaldehyde, while MG reductase converts MG to L-lactaldehyde. The enzyme L-lactaldehyde dehydrogenase converts this L-lactaldehyde to L-lactate, which is then reversibly converted to pyruvate by L-lactate dehydrogenase. Similarly, MG can also be converted by MG dehydrogenase to pyruvate and then by pyruvate dehydrogenase to acetyl-CoA and formate. Among these, a normal catabolism of MG is believed to involve glutathione and the activities of glutathione-dependent enzymes such as the Glyoxalase system that has been reported to have a major function in the detoxification of MG [13, 14].

1.2: Glyoxalase System

One of the several pathways that appear to be involved in the detoxification of methylglyoxal is the Glyoxalase (Glx) system. This is a two-enzyme system consisting of Glyoxalase I (GlxI) and Glyoxalase II (GlxII) that convert α-ketoaldehydes into their corresponding 2-hydroxycarboxylic acids (D-lactate in the case of methylglyoxal), using an intracellular thiol as a cofactor/cosubstrate (Figure 1.2) [15]. The first enzyme, GlxI (*S*-D-lactoylglutathione methylglyoxal lyase (isomerizing), EC 4.4.1.5), converts a hemithioacetal, the product of a non-enzymatic reaction between MG and a thiol such as glutathione, to *S*-D-lactoylglutathione. This compound is then hydrolyzed by GlxII (*S*-2-hydroxyacylglutathione hydrolase, EC 3.1.2.6) to produce D-lactate and regenerate its corresponding thiol. The enzyme lactate dehydrogenase can convert D-lactate to pyruvate, forming a bypass process for the phosphorylating glycolytic pathway. There is evidence that the intermediate, *S*-lactoylglutathione, can control intracellular pH in bacteria through interaction with the KefGB potassium efflux system [16].

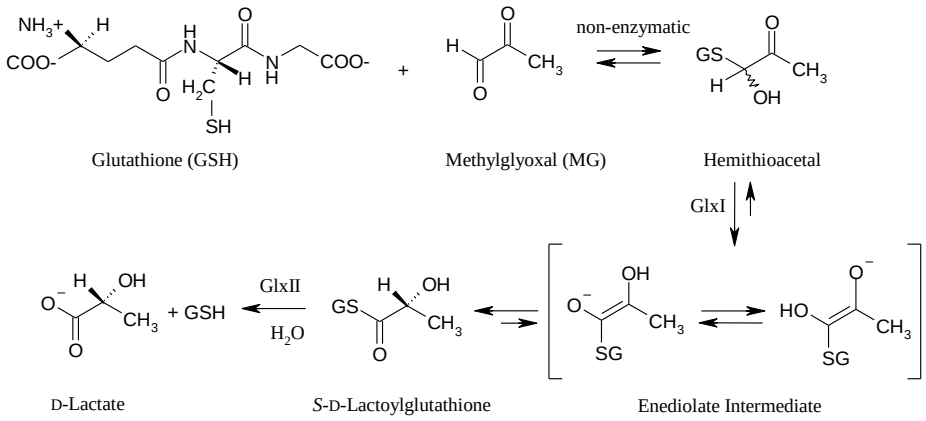


Figure 1.2: The two-enzyme Glyoxalase (Glx) system composed of Glyoxalase I (GlxI) and Glyoxalase II (GlxII).

1.2.1: Glyoxalase I

Glyoxalase I is a metalloenzyme that can be divided into two classes, a Zn^{2+}-activated class and a non-Zn^{2+}-activated class (being selectively Ni^{2+}/Co^{2+}-activated) [17-35]. Even though nothing is known regarding GlxI from archaea and Gram-positive bacteria, it is likely that almost all prokaryotes possess Ni^{2+}/Co^{2+}-activated GlxI, with the possible exception of pseudomonads [19, 28, 30, 32, 33]. It was found that GlxI from *Pseudomonas putida* is a

Zn^{2+}-activated enzyme [28, 30]. Interestingly, *Pseudomonas aeruginosa* has been reported to possess three genes coding for GlxI enzymes (*gloA1*, *gloA2* and *gloA3*) [32, 33]. The gene products, GloA1 and GloA2, are Ni^{2+}-activated, while GloA3 is a Zn^{2+}-activated enzyme [32, 33]. All Zn^{2+}-activated enzymes from pseudomonads have longer amino acid sequences, which are similar in length to that of *Homo sapiens* GlxI (also a Zn^{2+}-activated enzyme). Metal specificity for GlxI from eukaryotes seems to vary. The protozoan *Leishmania major* GlxI is Ni^{2+}/Co^{2+}-activated with no activity in the presence of Zn^{2+} ion [18, 35]. *Trypanosoma cruzi* GlxI is Ni^{2+}/Co^{2+}-activated with minor but measurable activity in the presence of Zn^{2+} [25]. On the other hand, the malarial parasite, *Plasmodium falciparum*, contains a GlxI that is Zn^{2+}-activated [17]. As well, yeast (*Saccharomyces cerevisiae*), human (*Homo sapiens*) and plant (*Brassica*) GlxI fall into the Zn^{2+}-activated class [20-24, 26, 29, 31, 34, 36]. Most of these Zn^{2+}-activated enzymes show low metal specificity. Yeast GlxI from *Schizosaccharomyces pombe* is also activated with other bivalent metals including Co^{2+}, Ni^{2+}, Fe^{2+}, Mn^{2+}, Mg^{2+} and Ca^{2+} (activity of metal reconstituted GlxI: $Mn^{2+} > Co^{2}+ > Ni^{2+} > Zn^{2+} > Mg^{2+} > Fe^{2+} > Ca^{2+}$) [34]. *H. sapiens* GlxI is activated by Mg^{2+}, Mn^{2+} and Co^{2+} (activity of $Mg^{2+}- > Zn^{2+}- > Co^{2+}- > Mn^{2+}$-activated enzyme) [31]. Higher plants such as *Brassica juncea* have a GlxI that can be activated by Mn^{2+}, Mg^{2+} and Ca^{2+} (activity of $Mn^{2+} > Mg^{2+} > Zn^{2+} > Ca^{2+}$ reconstituted GlxI) [23].

Analysis of previous metal titration, isothermal titration calorimetry (ITC) and nuclear magnetic resonance (NMR) experiments indicated that the dimeric GlxI from *E. coli* forms two active sites with 1 metal/dimer ratio for both Ni^{2+} and Co^{2+}-activation [19, 37, 38]. As well, it was found that the monomeric *P. falciparum* GlxI contains one metal per monomer (two active sites) [39]. These observations suggest that the metal binds tightly in the active pockets and only one active site may be functional at any one time, however, adding further complexity to this relatively "simple" enzyme.

GlxI is a member of the βαβββ superfamily of proteins, consisting of fosfomycin resistance protein (FosA), methylmalonyl-CoA epimerase (MMCE), extradiol dioxygenase (DIOX), mitomycin C resistance protein (MRP) and bleomycin resistance protein (BRP) [40-42]. The active site cleft is formed by arrangement of the βαβββ motif within one subunit as in the monomeric *Plasmodium falciparum* GlxI (a subunit fusion as determined by the secondary structural prediction and multiple sequence alignment [43]) or at the interface of a

homodimer as in the case of the *E. coli* and *H. sapiens* GlxI (PDB: 1F9Z and 1QIN, respectively) (Figure 1.3) [44]. A monomeric GlxI forms an active site from a single polypeptide (two active sites per dimer), while a homodimer requires active residues from both subunits to form a cleft for substrate binding.

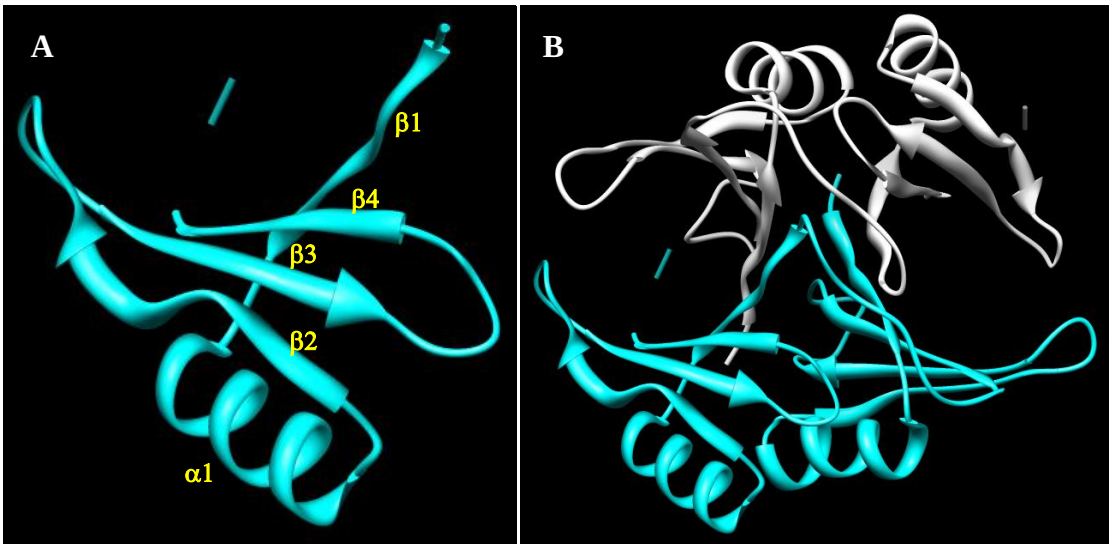


Figure 1.3: The ribbon structure of *E. coli* GlxI (PDB: 1F9Z) showing **(A)** the formation of the βαββββ topology and **(B)** the homodimeric structure created by two subunits (white and cyan) with two catalytic pocket at the dimeric interface. The 3D structures were generated by using the UCSF Chimera program (University of California, San Francisco) [1].

Recently, it has been determined that the protein's amino acid length and sequence could be used to predict the probable metal specificity of a putative GlxI given only its amino acid sequence. GlxI with shorter amino acid sequences (~130 amino acids in length) tend to be Ni^{2+}/Co^{2+}-activated (non-Zn^{2+}-activated), while longer ones (~180 amino acids in length) are likely to behave as Zn^{2+}-activated enzymes. The multiple sequence alignment as well as X-ray crystallographic structures indicates that the Zn^{2+}-activated GlxI possesses the long N-terminal arm that wraps around another subunit (domain swapping), thus holding a dimeric conformation together (Figure 1.4A). It also employs at least three more loops that are absent in the Ni^{2+}/Co^{2+}-activated enzyme. One out of three is a three-turned α-helix that blocks one side of the active pocket, hence preventing substrate from interacting with outer solvent (Figure 1.4B). Therefore, it would be interesting to investigate how the amino acid sequence

relates to metal activation profile of various GlxI and how the presence of these extra loops affects the type of metal binding in the catalytic pocket.

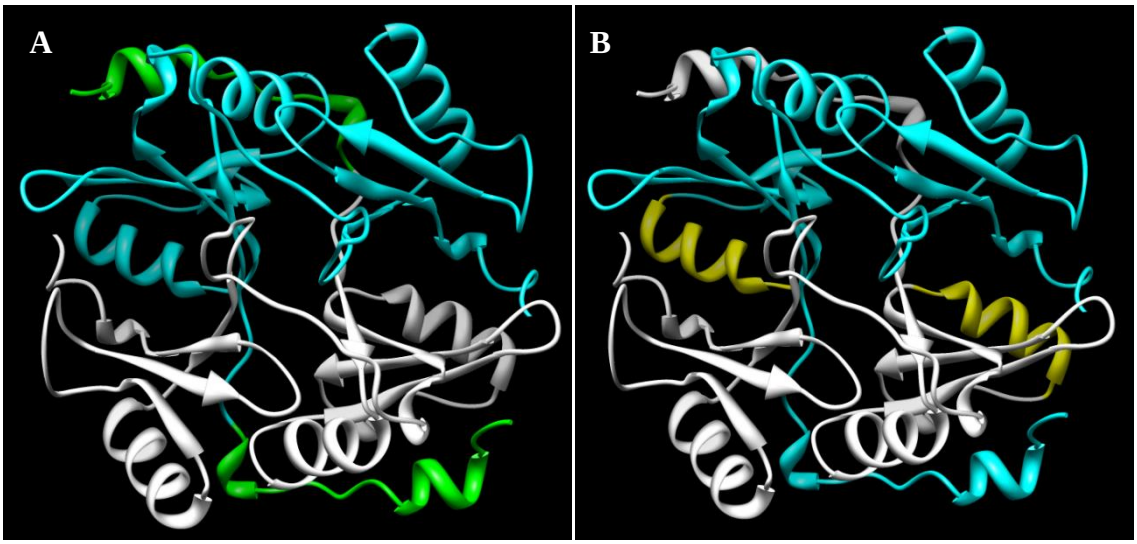


Figure 1.4: The ribbon structure of a homodimeric human GlxI (Zn^{2+}-activation, white and cyan, PDB: 1QIN) showing **(A)** the N-terminal arm (green) and **(B)** the three-turned α-helix (yellow) that are absent from the Ni^{2+}/Co^{2+}-activated enzymes. The 3D structures were generated by using the UCSF Chimera program (University of California, San Francisco) [1].

The difference between the active and inactive forms of GlxI may be explained by consideration of their active site geometry. The X-ray crystallographic structure of *E. coli* GlxI with bound Ni^{2+} in its active sites (active form, PDB: 1F9Z) possesses an octahedral geometry with four metal binding protein residues (two from each monomer including His^{5}, Glu^{56}, His^{74} and Glu^{122}) and two water molecules around the active metal (Figure 1.5A) [45]. Its corresponding inactive form with a bound Zn^{2+} ion (PDB: 1FA5) possesses a five-coordinated trigonal bipyramidal geometry with the same four metal binding protein residues but having only one water molecule bound to the zinc ion (Figure 1.5B) [45]. Similarly, the active Zn^{2+}-bound human GlxI possesses an octahedral geometry (PDB: 1QIN) [44]. It has been reported so far that only octahedral metal coordination yields active enzyme, regardless of the metal-activation class it belongs to (Zn^{2+}-activated, or Zn^{2+}-inactive but Ni^{2+}/Co^{2+}-activated). Hence, the active site geometry can be used to predict the enzyme activity of the GlxI family. This observation must be critical to the mechanism of GlxI, but its connection to the mechanism remains elusive, although hypothetical explorations have been put forth [46-50]. The proposed mechanisms of Zn^{2+}-activated human GlxI and its X-ray structure with a bound transition state

analogue (inhibitor) (PDB: 1QIN) suggest the displacement of two water molecules at the active site by the substrate [44]. This replacement, however, changes the coordination of the active site from six-coordinated environment to a five-coordinated site, indicating the loss of one metal binding residue. It was proposed that the displaced Glu^{172} residue acts as a catalytic base once the analogue binds to the active site and is responsible for deprotonation of the substrate [44, 46]. The mechanism of Ni^{2+}-activated GlxI, on the other hand, has yet to be investigated. The information derived from XAS experiments on *E. coli* GlxI suggests similar metal coordination where a five-coordinated metal geometry was detected after the addition of a transition-state inhibitor [47]. However, the displacement of the metal binding Glu residue under those conditions was not clear.

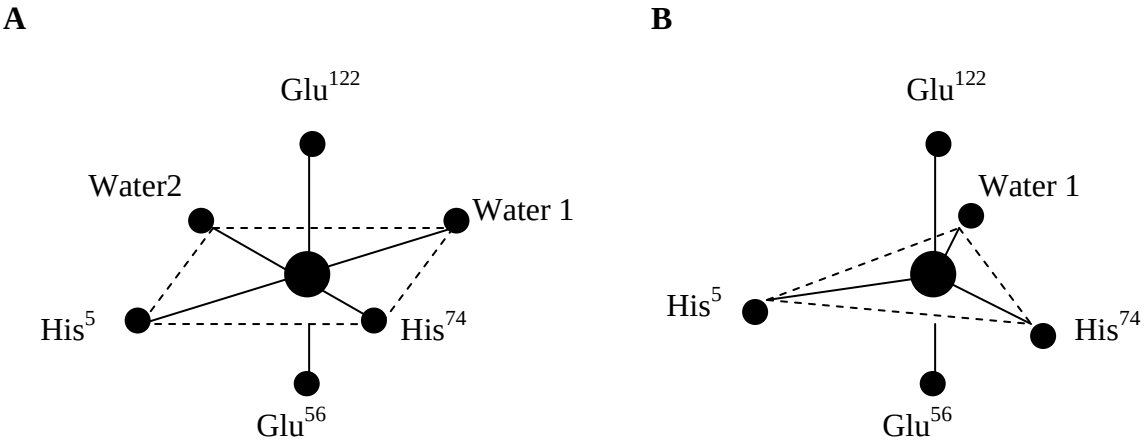


Figure 1.5: The metal coordination of **(A)** *E. coli* GlxI with bound active metal (such as Ni^{2+} or Co^{2+}) forming an octahedral geometry with four metal binding residues and two water molecules around the metal center and **(B)** *E. coli* GlxI with bound inactive metal (such as Zn^{2+}) forming a trigonal bipyramidal geometry with four metal binding residues and one water molecules [45].

1.2.2: Glyoxalase II

Glyoxalase II is a member of a superfamily of hydrolases and has been identified in humans, animals, plants, fungi and bacteria [47-62]. GlxII exists in both cytosol and mitochondria of eukaryotes, where the cytosolic and mitochondrial enzymes are encoded by separate genes [51-55]. Despite GlxI being detected only in the cytosol, one cytosolic form and one mitochondial form of GlxII have been found in yeast [36, 56, 57]. Interestingly, rat liver mitochondria possess multiple forms of GlxII, which have been found in both the intermembrane space and matrix [52, 58]. However, there is only one form of GlxII in the

cytosol and only the enzyme from the intermembrane space appears to resemble the cytosolic GlxII form [52, 58]. Plants, on the other hand, appear to possess only one mitochondrial GlxII form, but multiple forms of cytosolic GlxII [59, 60]. The isoelectric pH values (p*I*) of plant and yeast GlxII are acidic, while animal GlxII appear to have basic p*I* values [51-53, 56, 59-70]. The different p*I* values suggest significant amino acid compositional differences and evolution of the enzyme, which might lead one to believe that these enzymes probable arose in hypothetical ancestors before separation into the main kingdoms. Additionally, the detection of GlxII in mitochondria suggests the possible formation of glutathione thioesters from coenzyme A derivatives and the transport of *S*-D-lactoylglutathione from the cytosol to the mitochondria [53]. Since the mitochondrium contains D-α-hydroxyacid dehydrogenase, an enzyme that is responsible for the formation of pyruvate from D-lactate, it is possible that pyruvate can be formed directly in the mitochondria [53].

GlxII is a member of the metallo-β-lactamase superfamily, which includes metallo-β-lactamase, GlxII, rubredoxin:oxygen oxidoreductase (ROO) and ZiPD families [71, 72]. The proteins in this superfamily contain the conserved motif THxHxDH that serves as part of the metal binding ligands. GlxII are monomeric enzymes with a molecular weight between 23–30 kiloDaltons. The overall structure of GlxII can be divided into two parts according to the predominant secondary structure and folding: the β-sheet containing the N-terminal domain and the α-helix containing the C-terminal domain (Figure 1.6A). The N-terminal domain posseses a similar (protein) fold as a metallo-β-lactamase, which consists of a four-layered β-sandwich with two mixed β-sheets flanked by α-helices. The first half of the N-terminus has a ββββαβαββ motif and forms an angle of ~20° to the second half that possesses a βββββαβ topology. The C-terminal domain mainly consists of α-helices. The enzyme contains two metal binding sites, which are located in the N-terminal domain on the edge of the β-sandwich, while the substrate binding site lies at the interface of the two domains. For human GlxII, the Zn1 coordinates with Asp^{58} ($O_{\delta 1}$), His^{59} ($N_{\varepsilon 2}$) and His^{173} ($N_{\varepsilon 2}$), while Zn2 interacts with the residues His^{54} ($N_{\varepsilon 2}$), His^{56} ($N_{\delta 1}$) and His^{110} ($N_{\varepsilon 2}$) (PDB: 1QH5, Figure 1.6B) [73]. These metal ions are believed to be bridged by a hydroxide (HO^-) ion from a water molecule and the side chain of residual Asp^{134} ($O_{\delta 1}$), forming an octahedral geometry for both metal centers. Interestingly, *S*-(*N*-hydroxy-*N*-bromophenylcarbamoyl)glutathione (HBPC-GSH), an inhibitor of GlxI and a slow substrate of GlxII, bound to GlxII (PDB: 1QH5) generates a square

pyrimidal coordination arrangement around the metal, where the carbonyl oxygen of the substrate replaces the water molecule and binds to Zn1 [73]. However, the active site geometry is changed to the octahedral coordination upon the binding of GSH (inhibitor of GlxII), where another water molecule interacts with Zn1 and the sulfur from GSH interacts with Zn2 [73]. The difference in geometry of the active and inactive forms of GlxII (with respect to the substrate) suggests that the enzyme might employ similar rearrangements during catalysis as in GlxI (with respect to activated metal).

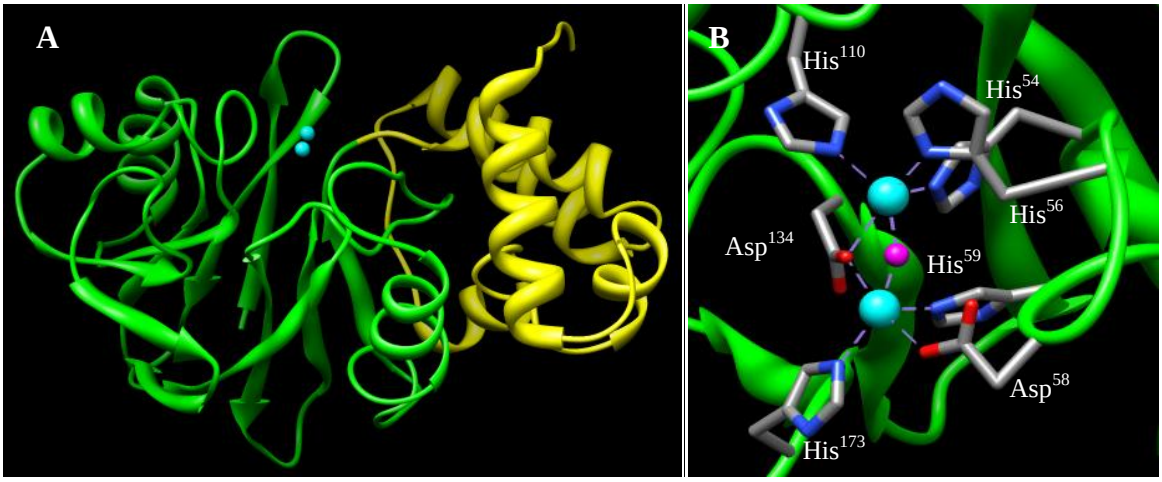


Figure 1.6: The ribbon structure of human GlxII (PDB: 1QH5) showing **(A)** the N-terminal domain with metallo-β-lactamase arrangment (green) and the C-terminal domain containing several α-helices (yellow) and **(B)** the active sites with bound binuclear Zn^{2+} atoms. The metals are shown as cyan spheres and the water molecule is a magenta sphere. The 3D structures were generated by using the UCSF Chimera program (University of California, San Francisco) [1].

GlxII is a metalloenzyme that can be activated by various metals depending on each particular organism. Its binuclear property with positive cooperativity is unique and conserved within GlxII, which differs from other metallo-β-lactamase family members that utilize a single metal ion [72]. A ratio of 1.5 moles of zinc and 0.7 moles of iron per mole GlxII has been found for the human GlxII [73]. The mitochondrial GlxII from *A. thaliana* employs variable metal centers including a predominant Fe^{3+}/Zn^{2+} and a minor Fe^{3+}/Fe^{2+} center [72]. The metal contents of approximately 0.58 moles of zinc and 0.61 moles of iron are found to be incorporated in the metal binding sites of this isolated enzyme [74]. The plant cell's cytosolic enzyme possesses an additional incorporated Mn^{2+} (0.4 moles zinc, 0.8 moles iron and 0.3 moles manganese), which is lacking in its mitochondrial counterpart [72, 75]. The cytosolic

enzyme appears to have a preference for binding one iron ion per enzyme, while the other metal binding site does not appear to have a metal preference [72]. The enzyme from *S. typhimurium* employs different metal cofactors such as Fe^{2+}, Fe^{3+}, Zn^{2+} and Mn^{2+} (0.21 moles zinc, 0.64 moles iron and 0.3 moles manganese), though with a clear predominance favoring iron ions [63]. Interestingly, the types of metal bound in the active site do not have a significant effect on enzyme activity [72]. On the other hand, *E. coli* GlxII binds on average to 1.7 moles of zinc per mole monomeric protein, while other metals are not detected from the "as isolated" enzyme [67]. Despite the existence of a Ni^{2+}/Co^{2+}-activated (class of) GlxI in *E. coli*, there does not appear to be a similar metal activation profile for its GlxII [67].

The Zn^{2+}-bound *E. coli* GlxII with *S*-D-lactoylglutathione possesses a k_{cat} of 53 s^{-1} and K_m of 184 ± 22 μM, and a similar enzymatic activity was observed with the Mn^{2+}-reconstituted enzyme [67]. The kinetics of *S. typhimurium* GlxII with bound Mn^{2+} on the same substrate exhibits a k_{cat} of 395 ± 11 s^{-1} and K_m of 241 ± 18 μM [63]. However, the metal content of this latter enzyme is only approximately 1 mole Mn^{2+} per mole monomeric enzyme, suggesting that the enzyme with only one metal atom is catalytically active and the second metal might not be necessary for catalysis [63]. These results are further supported by the discovery of the properties of the GlxII from *Salmonella enterica* serovar Typhimurium (YcbL), which possesses only one metal binding site, with optimal activity being observed in the presence of iron [68], and also by the human enzyme, which is active in its mono-metallated Zn^{2+} form [76].

Analysis of the X-ray crystallographic structure of human GlxII with bound *S*-(*N*-hydroxy-*N*-bromophenylcarbamoyl)glutathione (PDB: 1QH5) suggested several potential substrate binding residues (Lys^{143}, Tyr^{145}, Tyr^{175}, Arg^{249} and Lys^{252}). The carboxylate of the glycine group of the GSH moiety interacts with Arg^{249} and Lys^{252} from the C-terminal domain, while the amide nitrogen of the cysteine group interacts with Tyr^{175}. Lys^{143} lies within hydrogen bonding distance of the carboxyl oxygen of the γ-glutamate of the GSH moiety, which is also stabilized by interaction with the phenolic functionality of Tyr^{145}. This interaction, however, is not possible with GlxII from other organisms, since Tyr^{145} is substituted by Phe (except in the mitochondrial GlxII from *A. thaliana*, which possesses Ser), suggesting that this residue is not likely critical for substrate recognition [73, 77]. The

molecular recognition of the substrate is believed to occur through the GSH moiety and the interaction of the thioester group with the metal ions.

1.2.3: Glyoxalase III

A report has appeared that provides evidence for a single enzyme in *E. coli*, termed Glyoxlase III, which *directly* converts methylglyoxal to D-lactate without the need for an intracellular thiol cosubstrate [78]. Little information is available on this enzyme; however, one report indicates the possible role of RNA polymerase sigma factor (rpoS) in the bacterial regulation of this enzyme activity [79].

1.3: Intracellular Thiols

The cofactor/cosubstrate for the Glyoxalase I reaction, glutathione (GSH), is a tripeptide whose chemical structure is γ-L-glutamyl-L-cysteinyl-glycine (Figure 1.7A). GSH is mainly found in eukaryotes, most Gram-negative bacteria, some Gram-positive bacteria and yeast [80-83]. Its primary function is similar to other thiols, which is to serve as a cellular reductant, maintaining a reducing environment within the cell. It is also a contributor to the reduction of toxic peroxides and peroxy radicals. GSH is involved in many biological functions including aromatic metabolism, maintenance of cellular redox potentials and several detoxification mechanisms [82, 84]. Interestingly, the dissociation constant (K_d) of the hemithioacetal from the nonenzymatic reaction of MG and GSH in the Glx system is 3.1 mM [85, 86], which is in the range of the intracellular concentration of GSH (0.5–10 mM in mamalian cells) [87]. Therefore, the fraction of MG that exists as the hemithioacetal is markedly dependent upon the GSH concentration, especially because the flux of MG production intracellularly is in the low μmolar range [88, 89]. The net result is that the efficiency of the Glx system will decrease as the GSH concentration decreases, which may have important consequences under conditions of oxidative stress.

Other types of intracellular thiols are also produced, for example, protozoans such as *Leishmania* and *Trypanosoma* biosynthesize and accumulate trypanothione ($T(SH)_2$) in several reactions and pathways such as reductases, transferases and peroxidases (Figure 1.7C) [18, 25, 90-92]. Trypanothione is N^1,N^8-*bis*(*L*-γ-glutamyl-*L*-hemicystinyl-glycyl)spermidine that

contains two molecules of GSH linked together by the polyamine spermidine. Actinomycetes and mycobacteria, on the other hand, produce another distinct class of thiol, mycothiol (MSH) or 1-D-*myo*-inositol-2-(*N*-acetyl-L-cysteinyl)amido-2-deoxy-α-D-glucopyranoside [93-98]. Instead of forming a tripeptide linkage of cysteine with glutamate and glycine residues as in GSH, the cysteine residue in MSH is linked to a disaccharide containing inositol and *N*-glucosamine (Figure 1.7B). The functions of MSH and its metabolic pathways are believed to parallel those of GSH [98], thus GSH and GSH-dependent enzymes may serve as templates for the new discovery of MSH functionalities and pathways.

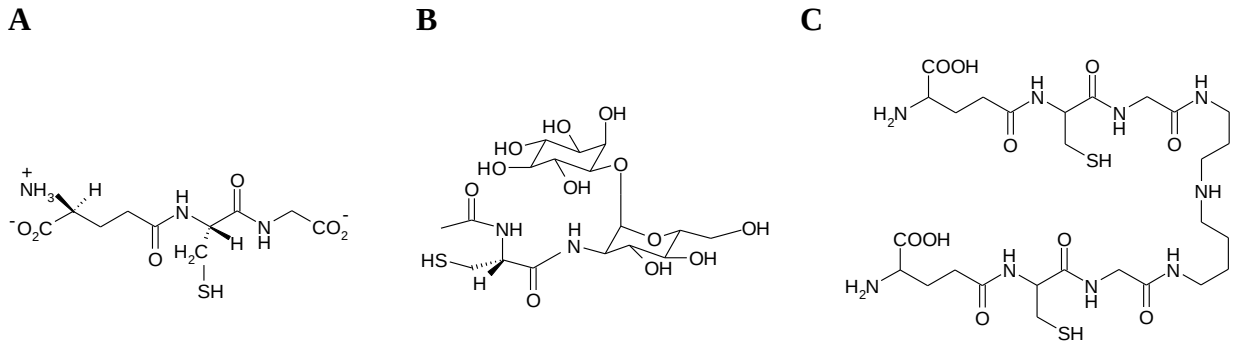


Figure 1.7: Different types of thiols produced in diverse groups of organisms including **(A)** glutathione, (**B**) mycothiol and (**C)** trypanothione**.**

The biosynthesis of MSH involves a series of enzymes that convert D-*myo*-inositol-3-phosphate (Ins(3)P_1) and UDP-*N*-acetylglucosamine (UDP-GlcNAc) to MSH (Figure 1.8) [95-98]. The first reaction is catalyzed by glycosyltransferase (MshA) and inositol kinase (MshA2), which converts these precursors to GlcNAc-Ins. The deacetylation by MshB then removes the acetyl group from this compound, thus forming GlcN-Ins. Its free amino group is attached to the cysteine moiety of the reactive Cys-AMP derivative by ligation of MshC in the presence of ATP to form Cys-GlcN-Ins. The MSH synthase, MshD, then transfers the acetyl group from acetyl-CoA to the amino group of the cysteine moiety to form MSH.

MSH is a cofactor/cosubstrate that plays a significant role in several enzymatic processes for the detoxification of thiol-reactive substances (such as formaldehyde, various electrophiles and antibiotics) and protects against oxygen toxicity. Several of these pathways remain hypothetical and have been inferred from analogy to GSH-dependent enzymes. However, additional experimentation is required to definitively confirm these pathways. MSH

can function as a detoxifying agent for thiol-reactive compounds (RX) by forming MSR conjugates, whose formation is catalyzed by the putative MSH *S*-transferase [98]. These MSR conjugates of antibiotics or toxins can be excreted from cells by some ATP-dependent transporters [98]. Mycothiol amidase (Mca), a zinc-dependent metalloenzyme, can also cleave the MSR conjugate to mercapturic acid that is secreted from cells and GlcN-Ins that can be reused in the MSH biosynthesis [95].

Another detoxification pathway involves oxidative and nitrosative stress protection through the reaction of a bifunctional formaldehyde dehydrogenase/nitrosothiol reductase (MscR), which is a NAD/mycothiol-dependent enzyme [99]. MscR is found to be more reactive as a nitrosothiol reductase than as a formaldehyde dehydrogenase [99]. MSH is spontaneously reactive toward formaldehyde to form the *S*-hydroxymethylmycothiol adduct, which is then converted to the MSH formate ester by formaldehyde dehydrogenase. This formate ester is predicted to eventually be converted to the carbonate ester, CO_2 and MSH by some unknown mechanisms as in the reaction of GSH-dependent formaldehyde dehydrogenase. As well, MscR can act like a nitrosothiol reductase that is reactive toward nitrosomycothiol (MSNO), which is converted to MSH sulphinamide (MS(NHOH)) in the presence of NADH [99]. This sulphinamide can subsequently be converted to nitrate and MSH *in vivo* through unknown reactions [98].

MSH is also a cofactor in the reaction of maleyl pyruvate isomerase, which converts maleyl pyruvate to fumaryl pyruvate during the metabolism of aromatic compounds [100]. As well, MSH can be oxidized to its disulfide form (MSSM) in the reduction of peroxides such as hydrogen peroxide (H_2O_2) or hydroperoxides (ROOH). The reaction was hypothesized to be catalyzed by antioxidant enzymes such as thiol peroxidase as in a GSH-dependent peroxidase [101, 102]. In order to maintain proper cellular levels of MSH, mycothiol reductase (Mtr) is able to reduce the disulfide bond of MSSM to its free thiol form in the presence of a FAD cofactor with the concomitant oxidation of NADPH [103]. In addition, the degradation of MSH to recover cysteine, which can be used in protein biosynthesis, could occur under limiting nutrient conditions. Once proper intracellular cysteine levels are sufficient, the cysteine could then be re-stored in the MSH form.

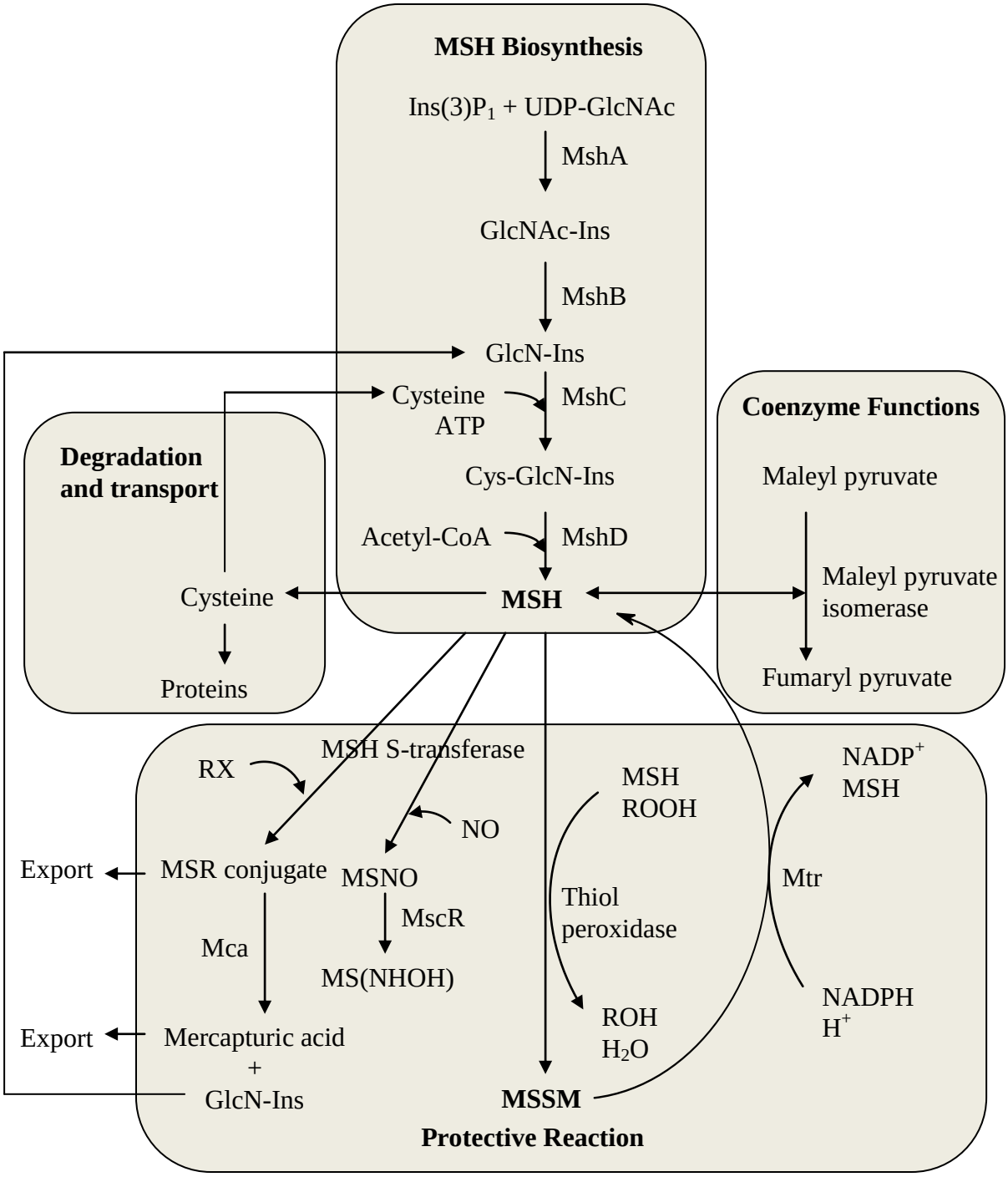


Figure 1.8: The overview of the biosynthesis and catabolic pathways of mycothiol based on the previous report by Rawat and Av-Gay, 2007 [98].

1.4: Possible Aspects for the Glyoxalase System in Gram-positive Bacteria and Evolution of the Glyoxalase Family

As mention previously, the Glyoxalase system has been widely investigated from various organisms; however, to our knowledge no detailed information regarding the enzymes from a Gram-positive bacterium has been reported. It was found that some Gram-positive bacteria such as *Streptococcus faecalis* and *Streptococcus agalactiae* produce glutathione as their major thiol, thus these organisms possibly utilize the Glx system as a pathway to detoxify cytotoxic methylglyoxal. Since several Gram-positive bacteria are pathogens that can infect both humans and animals, the enzymatic reaction and metal characteristics of GlxI from these organisms would allow us to extend our fundamental knowledge on the evolution of the enzyme from different organisms as well as support novel drug design, which might target the Glx system in infectious pathogens.

On the other hand, some Actinomycetales and Streptomycetes are reported to produce mycothiol [83]. These Gram-positive bacteria might possess a mycothiol-dependent Glyoxalase or a mycoGlyoxalase (mcGlx) system using a mycothiol cofactor as a parallel pathway to the GSH-dependent Glx system, which uses a glutathione cofactor. We hypothesized that the mcGlx system would consist of two enzymes, mycoGlyoxalase I (mcGlxI) and mycoGlyoxalase II (mcGlxII) (Figure 1.9). The first enzyme of the system, mcGlxI, could convert the hemithioacetal, the non-enzymatic product of the cytotoxic methylglyoxal and a mycothiol cofactor, to *S*-lactolymycothiol. This product is possibly hydrolyzed by the second enzyme, mcGlxII, to a non-toxic D-lactate and regenerates the cofactor. Since Actinomycetales are causative agents for various severe infectious diseases in humans, especially tuberculosis (TB), which is caused by *Mycobacterium tuberculosis* [104], MSH-dependent enzymes are possible targets for the development of novel pharmaceuticals against pathogenic mycobacteria.

GlxI is a member of the βαβββ superfamily of proteins, including fosfomycin resistance protein (FosA), methylmalonyl-CoA epimerase (MMCE), extradiol dioxygenase (DIOX), mitomycin C resistance protein (MRP) and bleomycin resistance protein (BRP). Among these, GlxI and MMCE are metalloenzymes with similar active sites and metal coordinations [105]. The enzymatic reaction of MMCE is to convert (2*S*)-methylmalonyl-CoA to its 2*R*-epimer and its activity can be optimized in the presence of Co^{2+} ion, similar to that of

GlxI in the Ni^{2+}/Co^{2+}-activation class [106-108]. Its active site also forms an octahedral metal coordination with four metal binding protein residues and two water molecules around the metal center [105]. Due to the metal characteristics and structural similarity of GlxI and MMCE, we hypothesized that these enzymes might be able to cross-function. This investigation would provide more supporting evidence and extend our knowledge on the evolution of the proteins within the βαβββ superfamily. These proteins are believed to have evolved from a common ancestor but due to gene fusions and duplication as well as accumulated point mutations, they possess different biological functions [30, 40, 41]. As well, to our knowledge, this is the first investigation of cross-functional activity on the enzymes within this same superfamily.

Fosfomycin resistance protein, FosA, is also a metalloenzyme in the βαβββ superfamily and has two metal binding sites, one for a divalent metal such as Mn^{2+} and another one for the monovalent metal K^{+} [109-111]. Similar to GlxI, FosA uses GSH as a cosubstrate in its reaction [112, 113]. It was thought that FosA might have evolved from an already existing gene in the same organism through gene modification and mutation. Thus, the investigation on the relatedness of GlxI and FosA in the same βαβββ superfamily would possibly lead to a better understanding of their evolution as well as knowledge concerning bacterial drug resistance mechanisms.

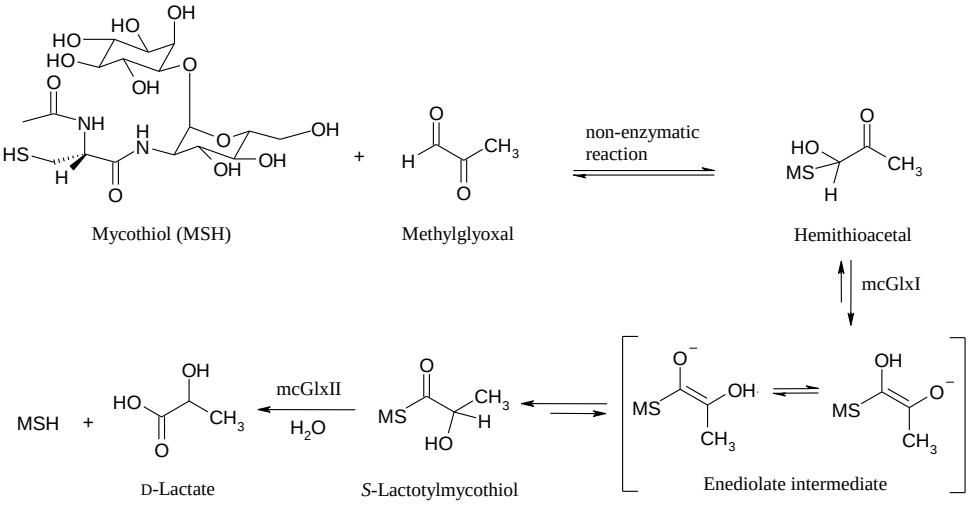


Figure 1.9: The putative two-enzyme mycoGlyoxalase (mcGlx) system composed of mycoGlyoxalase I (mcGlxI) and mycoGlyoxalase II (mcGlxII).

1.5: Summary of Experimental Objectives Achieved

Our interest was to investigate various aspects of the bacterial Glyoxalase systems. First, we investigated the GlxI reaction from Gram-positive bacteria using the hemithioacetal from the non-enzymatic reaction between MG and GSH as a substrate. The GlxI from *Clostridium acetobutylicum* was chosen as a representative of the enzyme from Gram-positive bacteria due to its structural availability and known genome sequence. As well, this bacterial strain is non-pathogenic to humans, thus is easy to handle. DNA and proteins of this bacterial GlxI were supplied by our collaborator, the New York Structural Genomics Research Consortium (NYSGXRC), and was additionally produced in our laboratory for investigations on protein characterization, metal profiles and enzymatic reactions in relation to its X-ray crystallographic structure. A comparison of this bacterial enzyme and the GlxI from other organisms was made in terms of enzyme evolution, metal activation classes, structural arrangement and size of the enzymes.

As mention previously, the metal activation class of a new GlxI can be predicted by the length of its amino acid sequence. We were therefore interested in investigating the possible factors that might influence the metal selection and enzymatic activity of the GlxI reaction. The analysis on the multiple sequence alignment of GlxI from different metal activation classes suggested three extra protein regions that only exist in the Zn^{2+}-activated GlxI and not in the Ni^{2+}/Co^{2+}-activated enzymes. Deletional mutagenesis of these regions was performed using the Zn^{2+}-activated GlxI from *Pseudomonas aeruginosa* (GloA3) as the starting protein in an attempt to convert this structure into a Ni^{2+}/Co^{2+}-activated enzyme (GloA1 and GloA2) from the same organism. This approach might provide an advantage since these proteins would be from the same organism. In addition, the X-ray crystallographic structures of GloA2 and GloA3 were being analyzed in collaboration with Dr. Charles Bond, University of Western Australia (WA, Australia).

Due to structural similarity of proteins within the $\beta\alpha\beta\beta\beta$ superfamily, the relationship of Glyoxalase I, methylmalonyl-CoA epimerase and fosfomycin resistance protein were also investigated. An investigation of whether GlxI can also catalyze the MMCE reaction was examined. Assay of the MMCE reaction requires a coupling reaction with methylmalonyl-CoA mutase. The gene coding for this protein was kindly supplied by Dr. Thomas Bobik, Iowa State University (IA, USA). Similarly, an investigation of whether a GlxI could catalyze

the fosfomycin-resistance reaction was investigated in *E. coli* GlxI by performing point mutagenesis in order to mimic the environment of the active site of FosA. The above enzymatic activities as well as structural analysis of these protein variants are discussed in terms of evolution within the βαββββ superfamily and possible drug resistance mechanisms.

We were also interested in the reaction mechanism of the Ni^{2+}/Co^{2+}-bound GlxI in comparison to those of the Zn^{2+}-activated enzyme. Two inhibitors, *S*-(*N*-hydroxy-*N*-methylcarbamoyl)glutathione and *S*-(*N*-hydroxy-*N*-phenylcarbamoyl)glutathione, were synthesized based on previous reports [114, 115] and the enzyme activity with bound inhibitors was investigated in terms of the effect of hydrophobicity of the catalytic pocket as well as structural flexibility upon binding of inhibitors. In addition, the X-ray crystallographic structures of *S*-(*N*-hydroxy-*N*-phenylcarbamoyl)glutathione bound GlxI were analyzed in collaboration with NYSGXRC and Dr. Charles Bond from University of Western Australia.

We then investigated the possible presence of the mycoGlyoxalase system in Gram-positive bacteria using the hemithioacetal from the non-enzymatic reaction between MG and MSH as a substrate. Due to the lack of commercial availability, MSH was isolated from the non-pathogenic bacterial source, *Streptomyces jumonjinesis*. However, its recovery yield and efficiency were low, thus the truncated mycothiol (tMSH) or des-*myo*-inositol mycothiol was synthesized since the disulfide form of this organic compound could be substituted for the real substrate in the reaction of a mycobacterial MSH disulfide reductase and it was hoped that the mycoGlyoxalase enzymes might also exhibit such promiscuity of substrate structure [116]. The putative mcGlxI and mcGlxII were cloned from *Streptomyces coelicolor*, the non-pathogenic Gram-positive bacterium that produced MSH as a major thiol, to a more accommodating expression vector and the protein was overproduced in *E. coli*. The protein was characterized and its metal activation and enzymatic activity were investigated. As well, the predicted structural aspects of this protein were compared to those of GlxI from other organisms.

CHAPTER 2: INVESTIGATION OF THE METALLOENZYME GLYOXALASE I FROM *CLOSTRIDIUM ACETOBUTYLICUM*

2.1: Introduction

Clostridia are Gram-positive bacteria that reside in many different environments, but mostly are found in soil. Most species in this genus possess a rod-like or a drumstick shape with a bulge at the terminal ends (Figure 2.1). These obligate anaerobic organisms can exist in two forms: a vegetative and a spore form. The vegetative or so called growth phase occurs under normal or nutrient-rich conditions, while spore formation, also termed the survival phase, occurs under environmental stress such as under conditions of limited nutrients or dessication. Clostridia can be useful as commercial producers of solvents (such as acetone and alcohol) or, on the other hand, are pathogens that cause neurotoxic infections in humans and animals. Significant diseases in humans due to *C. botulinum*, *C. difficile*, *C. tetani* and *C. perfringens* infections are botulism, diarrhea, tetanus and necrosis of the tissue adjacent to wounds, respectively [117]. Non-pathogenic strains such as *C. thermocellum* and *C. acetobutylicum* are useful in the solvent industry due to their capabilities to produce ethanol, acetone and butanol [118-120]. Some non-pathogenic Clostridia may have potential applications in medical treatments such as cancer, tumor detection and drug delivery [121-123].

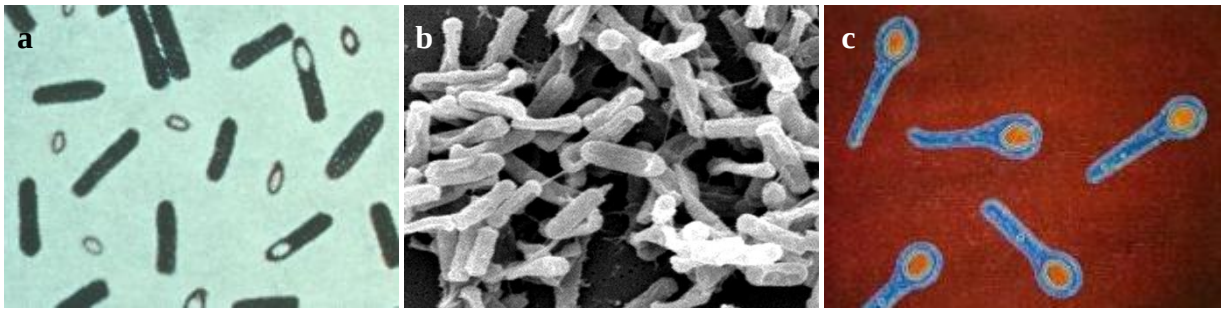


Figure 2.1: Spores of (a) *Clostridium botulinum*, (b) *Clostridium difficile* and (c) *Clostridium tetani* share a common drumstick shape. The images are ontained from Wikimedia Commons (http://commons.wikimedia.org).

Clostridium botulinum

Clostridium botulinum causes several well known diseases in humans including food-borne botulism, wound botulism, infant botulism as well as a rarely-found infection of

intestinal colonization, injection botulism and inhalation botulism [124, 125]. Neurotoxins produced by this species interfere with nerve impulses and affect muscle contraction, leading to muscular paralysis. This process involves binding of the neurotoxin at the recognized receptor on the nerve ending, followed by transportation of a part of the neurotoxin into the nerve cell, which eventually prevents the release of acetylcholine from the nerve ending that leads to muscle paralysis [117, 126]. This neurotoxin has a potential application as a biological weapon (only one gram can kill one million people). A much diluted form of the toxin is used in commercial cosmetics known as Botox (botulism neurotoxin) and also in some therapeutic applications such as in the treatment of painful muscle spasms [121].

Clostridium tetani

Clostridium tetani causes diseases with symptoms that are the opposite to those produced by *C. botulinum*. While *C. botulinum* neurotoxin can cause muscle paralysis, tetanospasmin (a neurotoxin produced in this species) rigidifies muscle and produces spasms [127]. Other symptoms are lockjaw (trismus), swallowing difficulty and stiff muscles as well as an effect on blood pressure and heart rate. Tetanospasmin can spread from a wound or cut by diffusion into adjacent muscle tissue, followed by transportation through blood vessels by the lymphatic system as well as passage through nerves [117, 127, 128]. This toxin acts on the inhibitory interneurons that normally moderate the activity of the motoneurons, thus preventing the release of the inhibitory mediators γ-aminobutyric acid and glycine and allowing uncontrolled stimulation of the muscle [117, 127, 128]. People in contact with animal waste, soil, illegal drug syringes and unsanitary birth conditions (neonatal tetanus) have a higher chance of being infected with tetanus. Since *C. tetani* is an obligate anaerobic bacterium that cannot survive when exposed to atmospheric oxygen, cleaning or opening wounds reduces the risk of the infection. Fortunately, tetanus is not frequently observed since the development of the tetanus vaccine.

Clostridium difficile

Clostridium difficile is harmful to humans as it is a cause of pseudomembranous colitis, an infection in the colon leading to diarrhea, fever, nausea and abdominal pain/tenderness [117, 126, 128]. Minute populations of this species can grow in the human intestine without harming

the host; however, its toxin can be overproduced under conditions of low antibiotic levels (immunocompromising conditions, illness and aging). *C. difficile* produces two types of toxin, enterotoxin A and cytotoxin B [117, 126, 128]. Toxin A can damage tissue by changing the permeability of the epithelial cells in the intestinal wall, causing a barrier formation between intestinal lumen and the surrounding tissue. This barrier causes a leakage of interstitial fluid that flows into the intestinal tract, a symptom of diarrhea. Toxin B is involved in the depolymerization of filamentous actin, which affects the structure of the cell cytoskeleton.

Clostridium perfringens

Clostridium perfringens produces enterotoxins that mainly cause food poisoning in humans. The toxins can be divided into A, B, C, D and E types [117, 126, 128]. The type A toxin is found in soil, dust and sewage and is responsible for gas gangrene and foodborne illness in humans. Type B, C, D and E toxins are only found in human and animal intestines, and mainly cause enterotoxemia in animals. Since its spore is not formed in food but in the intestine, cooking at high temperatures (> 60 °C) can reduce the risk of infection. The symptoms of infection are intense abdominal cramps, nausea, vomiting, gassy diarrhea and fever.

Clostridium thermocellum

Clostridium thermocellum is a non-pathogenic species that can degrade insoluble cellulose into small sugar molecules such as glucose and waste products including hydrogen, carbon dioxide, acetate and primarily ethanol *via* fermentation processes [118-120]. The cellulolytic enzyme system that is involved in this natural decomposition of biomass is called the cellulosome, and is composed of approximately twenty enzymes. *C. termocellum* can be found associated with many plants and farm animals such as the cow and the horse that require enzymes to decompose grass.

Clostridium acetobutylicum

Clostridium acetobutylicum is also known as the Weizmann organism because it was first used by Chaim Weizmann in 1916 [129]. The whole genome has been sequenced and reported in 2001 [130]. The genome contains genes coding for saccharolytic enzymes that can

break down sugar or other carbohydrates to produce many useful products including acetone, ethanol, butanol and hydrogen gas *via* substrate phosphorylation in fermentation processes. Many factories have been using the ABE method, a process that utilizes bacterial fermentation to produce solvents such as acetone, butanol and ethanol in a ratio of 3–6–1, respectively, from starch. In the first era of discovery for useful applications of this bacterium, acetone is the major product that was an important ingredient for smokeless gunpowder and trinitrotoluene industry. However, after World War I, the need of acetone dropped, causing most factories to shut down. Other products such as butanol can be used in the automobile industry as an alternative source of fuel. Hydrogen gas can be used as an alternative fuel, which produces no carbon dioxide or greenhouse gases upon its combustion. It can also be used to hydrogenate oils, changing non-saturated to saturated fats, which are important in the food industry.

Objective

Since many species of Clostridia can cause several severe diseases in most living organisms including humans and animals, an investigation on the Glyoxalase system, a detoxification pathway for cytotoxic methyglyoxal, would be of interest for fundamental knowledge of this enzyme system as well as to provide new information that might be exploited in new drug discovery against these pathogenic bacteria. Although several GlxI enzymes have been investigated from a number of biological sources, little is currently known concerning GlxI enzymes from Gram-positive microorganisms. The recently deposited but unpublished X-ray crystallographic structure of the *C. acetobutylicum* Glyoxalase I (CLO GlxI, AAK80149) obtained through a high throughput structural genomics initiative has led to interesting information concerning the possible structures of Gram-positive GlxI enzymes (PDB: 2QH0). The protein databank deposited enzyme structure was found to bind a zinc atom in its isolated form, produced from high throughput screening techniques. The metal coordination was observed to have a trigonal bipyramidal geometry around the metal with four metal binding ligands and one water molecule, indicating that this Zn^{2+}-bound enzyme might be in its inactive form. It has previously been reported by our laboratory that there are in fact two distinct metal-activation classes of GlxI, a Ni^{2+}/Co^{2+} metal-activated class (which is inactive with Zn^{2+}) and a Zn^{2+}-activated class [19, 32]. The analysis of multiple sequence alignment of CLO GlxI with those from various organisms including Gram-negative bacteria

(*E. coli*, *Y. pestis*, *N. meningitidis*, *P. aeruginosa* and *P. putida*), protozoa (*L. donovani* and *T. cruzi*) and humans (*H. sapiens*) has shown that CLO GlxI possesses four conserved metal binding residues (His^{5}, Glu^{52}, His^{75} and Glu^{124}), similar to those of other GlxI (Figure 2.2). The length of the enzyme is shorter and more in line with Ni^{2+}/Co^{2+}-activated GlxI enzymes than Zn^{2+}-activated GlxI enzymes, which seemed perplexing to us. From the analyses of X-ray crystallographic structure and multiple sequence alignment, we then hypothesized that CLO GlxI might be a likely Ni^{2+}/Co^{2+}-activated enzyme.

In an attempt to unambiguously determine the metal specificity of CLO GlxI, its DNA (*clo glxI*) was cloned into a pSGX3 (BC) expression vector and was transformed into an *E. coli* expression system (prepared by the New York Structural Genomics Research Consortium (NYSGXRC)). The conditions for protein purification and stability were optimized in our laboratory. The analysis of its metal activation profile, the metal reactivation of the apo-enzyme by metal titration, and the pH profile were all investigated for this clostridial enzyme. These results have indicated that CLO GlxI is indeed a Ni^{2+}/Co^{2+}-activated enzyme. In order to further investigate with respect to its active site geometry by undertaking X-ray crystallographic studies on the Ni^{2+}-active form of the enzyme, more protein was overproduced in *E. coli* cells and purified in our laboratory using the DNA vector containing *clo glxI* gene that was recloned from pSGX3 (BC) vector into an available pET-28b(+) expression vector. The X-ray crystallographic structure of Ni^{2+}-bound CLO GlxI exhibits an octahedral metal coordination with four metal binding protein residues and two water molecules around the metal center (PDB: 3HDP). These discoveries support our previous hypothesis on the relatedness of the active/inactive forms of GlxI with respect to metal selectivity and the metal coordination. Additionally, the overall structure of CLO GlxI is discussed and compared to other proteins with the $\beta\alpha\beta\beta\beta$ arrangement in order to gain a better understanding of the structural evolution of GlxI within this superfamily.

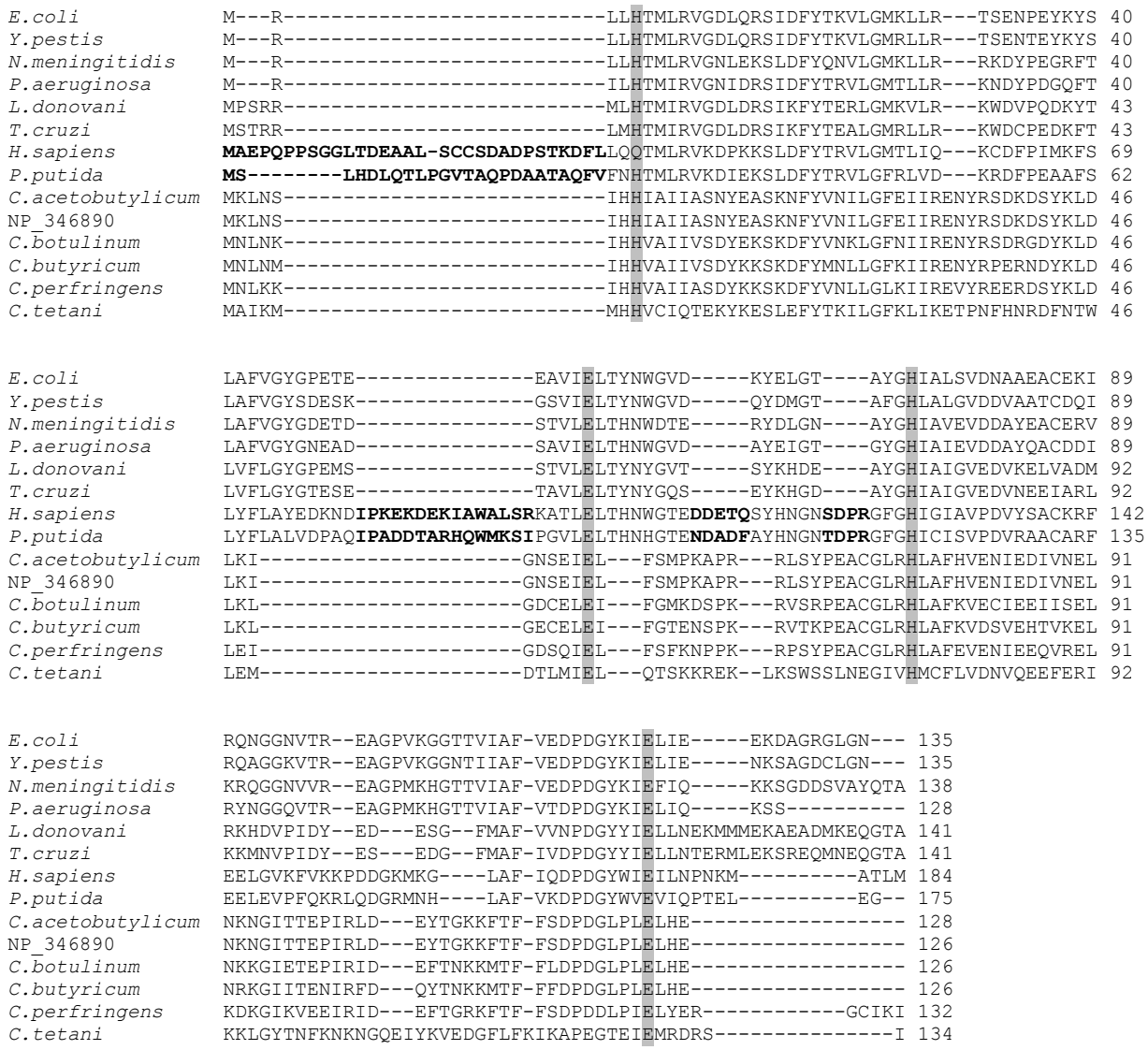


Figure 2.2: The multiple sequence alignment of Glyoxalase I from various organisms (organism name follows by National Center for Biotechnology Information (NCBI) accession number) including *E. coli* (NP_310387), *Y. pestis* (ZP_01887743), *N. meningitidis* (CAA74673), *P. aeruginosa* (AAG06912), *L. donovani* (AAU87880), *T. cruzi* (XP_818456), *H. sapiens* (AAB49495), *P. putida* (AAN69360) as well as putative Glyoxalase I from *C. botulinum* (YP_001920932), *C. butyricum* (ZP_02951467), *C. perfringens* (YP_697773) and *C. tetani* (NP_782738). There are two putative GlxI from *C. acetobutylicum*, AAK80149 and NP_346890. The one that is investigated in our experiments according to its metal specificity is AAK80149. The active residues are marked with a grey highlight and bold letters indicate the N-terminal arm and at least three loop regions that only exist in the Zn^{2+}-activated GlxI. The alignment was created using CLC Free Workbench (version 3.0.1) with the accurate alignment algorithm (http://www.clcbio.com).

2.2: Reagents, Materials and Instrumentation

Reagents and Materials

The 100 base pair DNA ladder and all restriction enzymes including T4 DNA ligase, *Nde*I and *BamH*I were obtained from New England Biolabs (Beverly, MA), while PWO DNA polymerase was provided by Roche Diagnostics (Laval, QC). DNA purification was performed using QIAprep Spin Miniprep Kit and QIAquick Gel Extraction Kit (QIAGEN, Mississauga, ON). Agarose was obtained from EM Science (Gibbstown, NJ).

Microbiological reagents acquired from EMD Pharmaceuticals (Durham, NC) were as follows: trypton, yeast extract and isopropyl-β-D-thiogalactopyranoside (IPTG). The antibiotic, kanamycin (Kan), was obtained from Sigma Chemical Company (St. Louis, MO).

Acid reagents such as hydrochloric acid (HCl), glacial acetic acid (CH_3CO_2H), phosphoric acid (H_3PO_4) and trichloroacetic acid (TCA) were obtained from EMD Pharmaceuticals (Durham, NC), while ethylenediamine tetraacetic acid (EDTA) was acquired from Fisher Scientific (Nepean, ON). Base reagent, sodium hydroxide (NaOH), was obtained from EM Science (Gibbstown, NJ).

Salts were provided from EMD Pharmaceuticals (Durham, NC) including potassium chloride (KCl), sodium chloride (NaCl), potassium phosphate monobasic (KH_2PO_4) and potassium phosphate dibasic (K_2HPO_4).

Metals were purchased from J.T. Baker Chemical Co. (Phillipsburg, NJ) including zinc chloride (assay, 99.3%), cobalt chloride hexahydrate (assay, 100.4%; specification, 98–102%), manganese chloride tetrahydrate (assay, 98.8%) and cadmium chloride (assay, 99.4%). Magnesium chloride hexahydrate (assay, 99.2%) and calcium chloride dihydrate (assay, 75.5%) were obtained from BDH Chemicals Ltd. (Poole, England). Nickel chloride hexahydrate (99.9999% pure) and cupric chloride hydrate (redox titration, 99.0%) were supplied from Sigma Aldrich (St. Louis, MO).

General buffers such as tris(hydroxymethyl)aminomethane (Tris), 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), 3-(*N*-morpholino)-2-hydroxypropanesulfonic acid (MOPSO) and 3-(*N*-morpholino)propanesulfonic acid (MOPS) were obtained from BioShop Canada Inc. (Burlington, ON).

Proteins such as *Saccharomyces cerevisiae* GlxI (grade IV, 1590 units per mg protein), bovine serum albumin (BSA) and protein standards for gel permeation chromatography (blue

dextran, bovine serum albumin, carbonic anhydrase, cytochrome *c* and aprotinin with molecular weights of 2000, 66, 29, 12.4 and 1.35 kDa, respectively) were obtained from Sigma Chemical Company (St. Louis, MO). Hen egg white lysozyme was supplied by ICN Biochemicals, Inc. (Costa Mesa, CA). Low molecular weight calibration kit (phophorylase b, albumin, ovalbumin, carbonic anhydrase, trypsin inhibitor, α-lactalbumin with molecular weights of 97, 66, 45, 30, 20.1, and 14.4 kDa, respectively) and chemicals (10–15% acrylamide gradient gel and buffer stripes) for sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) as well as thrombin protease were obtained from GE Healthcare (Piscataway, NJ). Proteinase K was obtained from Roche Diagnostics (Mannheim, Germany).

Dyes such as Coomassie brilliant blue R, ethidium bromide and bromophenol blue were supplied from Sigma Chemical Company (St. Louis, MO), while Coomassie brilliant blue G-250 was obtained from Eastman Kodak Co. (Rochester, NY).

General reagents such as glycerol were obtained from EMD Pharmaceuticals (Durham, NC), β-mercaptoethanol, methylglyoxal (MG), glutathione (GSH), iodoacetamide and 4-(2-pyridylazo)resorcinol, monosodium salt hydrate (PAR) were supplied from Sigma Chemical Company (St. Louis, MO). Imidazole was purchased from BioShop Canada Inc. (Burlington, ON).

The chromatographic columns for protein purification such as HisTrap™ HP affinity column (1 mL), HiTrap Benzamidine FF affinity column (1 mL) and Superdex™ 75 HR 10/30 column were obtained from GE Healthcare (Piscataway, NJ). Chelex® 100 chelating ion exchange resin was received from Bio-Rad laboratories (Hercules, CA).

Dialysis tubing used in all experiments was Spectra/Por® molecular porous membrane with 12–14 kDa molecular weight cut-off (Spectrum laboratories, Inc., Rancho Dominguez, CA). The Acrodisc® 25 mm Syringe filter with 0.2 μm Supor® membrane for protein filtration propose was obtained from PALL (East Hills, NY), while 0.22 μm GV filter membrane for buffer filtration was obtained from Millipore (Billerica, MA). The Vivaspin (2 mL) ultrafiltration spin columns with 10 kDa molecular weight cut-off polyethersulfone (PES) membrane, utilized to concentrate protein and exchange buffers, was obtained from Sartorius AG (Goettingen, Germany).

Water used in all experiments was Milli-Q water (18 MΩ-cm conductivity).

Equipment

All polymerase chain reactions were performed using a Techne® (TC-512) Thermal Cycler from Techgene, Inc. (Princeton, NJ). Cell growth for protein expression and induction was performed using an incubator shaker (series 25) from New Brunswick Scientific Co., Inc. (Edison, NJ). A Sonicator™ cell disruptor (model W225) with converter model #2 and a standard tapered microtip with the output control set at 5 (Heat Systems-Ultrasonics, Inc., Plainview, NY) was used for cell wall disruption of crude cell extracts in protein purification protocols.

The large volume centrifugation for protein purification was performed using a Avanti® J-E Centrifuge from Beckman Coulter, Inc. (Palo Alto, CA). A small volume (microcentrifuge scale) was centrifuged using a tabletop Biofuge A microcentrifuge (Heraeus Sepatech GmbH, Germany).

A high performance liquid chromatography (HPLC) system from Waters Corporation (Chicago, IL) with a Waters 625 LC system and a 994 programmable photodiode array detector was used for protein purification, quaternary structural investigation and methylmalonyl-CoA epimerase assays.

The identification of protein samples was performed using a Micromass Q-TOF Ultima™ Global electrospray mass spectrometer (ESI-MS), which was supplied by the Waterloo Chemical Analysis Facility, University of Waterloo. Protein visualization on SDS-PAGE with Coomassie Brilliant Blue staining was performed on a PhastSystem™ instrumentation (Pharmacia, Etobicoke, ON).

Circular dichroism (CD) measurements were performed on a Jasco J-715 spectropolarimeter from Jasco Inc. (Easton, MD) with a Jasco Spectropolarimetry power supply (PS-150J) and a Jasco J-700 for Windows Standard Analysis program (version 1.20.00).

Differential Scanning Calorimetry (DSC) experiments were performed on a MicroCal VP-DSC microcalorimeter (MicroCal Inc., Northampton, MA). The sample was degassed and the temperature was adjusted in an evacuated chamber, ThermoVac for Sample Degassing and Thermostat (MicroCal Inc.). Data processing and analysis was performed utilizing a MicroCal Origin software program, version 5.0 (MicroCal Inc.).

Spectrophotometric scanning for optical density readings at the specific wavelength for DNA and protein concentration measurements was performed using a Varian Cary 3

spectrophotometer (Mississauga, ON) with associated software package (designated version 3.04). The enzymatic assay was performed on a SpectraMax 190 96-well UV-visible spectrophotometer using a Soft Max Pro software program (Molecular Devices, Sunnyvale, CA).

2.3: Experimental Protocols

2.3.1: Theoretical Calculations

Calculations of the theoretical molecular weight and the isoelectric point (p*I*) of CLO GlxI were performed on the web based software, ProtParam tool from the ExPASy Proteomics Server (http://ca.expasy.org/tools/protparam.html) [131]. Dimerization and interface calculations were performed on "Protein interfaces, surfaces and assemblies service (PISA)", at the European Bioinformatics Institute (http://www.ebi.ac.uk/msd-srv/prot_int/pistart.html), authored by E. Krissinel and K. Henrick [132] and 3D comparison studies are performed at Dali server (Protein Structure Database Searching by DaliLite v. 3, http://ekhidna.biocenter.helsinki.fi/dali_server) [133].

2.3.2: DNA Cloning and Protein Purification

His-tagged CLO GlxI–high-throughput protocols

The pSGX3-CLO GlxI plasmid was supplied by NYSGXRC and was transformed by heat shock into *E. coli* BL21 (DE3) for protein expression purposes. *E. coli* cells containing the desired plasmids were inoculated overnight in Luria Bertani (LB) broth (5 mL) containing kanamycin (30 μg/mL) in 37 ºC shaker (220 rpm). The culture was then transferred into 1 L LB containing Kan (30 μg/mL) and allowed to grow in a shaker (220 rpm) at 25 ºC until OD_{600} reached 0.6. Proteins were then induced by adding 1 mM IPTG and shaken at the same speed and temperature for 4 hours. The culture was then centrifuged at 6,000×*g* for 10 min. Cell pellet was resuspended in a minimal volume of wash buffer (50 mM Tris (pH 8.0), 0.5 M KCl and 20 mM imidazole) before centrifugation at 6,000×*g* for an additional 10 min. The cell pellet was then collected and flash frozen in liquid nitrogen before storing at -80 ºC.

The frozen pellet was resuspended in wash buffer (1 g cell pellet/3 mL wash buffer) with the addition of 1 mg/mL hen egg white lysozyme and 1 mM

phenylmethanesulfonylfluoride (PMSF). Cells were disrupted by sonication at 4 ºC for 25 × 10 sec. The suspension was centrifuged at 48,384×*g* for 15 min. The supernatant was filtered through 0.2 μm GHP membrane filters and loaded onto a HisTrap affinity column. The column was then washed with wash buffer to eliminate non-specific binding of *E. coli* proteins, while His-tagged proteins were collected through a utilization of elution buffer (50 mM Tris (pH 8.0), 0.5 M KCl and 500 mM imidazole). To these eluted proteins, EDTA (4 mM) was added in order to prevent protein aggregation. Dialysis against 1 L of Tris buffer (50 mM Tris (pH 8.0), 150 mM KCl and 10% glycerol) was then performed overnight at 4 °C using a 12–14 kDa cut-off dialysis membrane.

His-tagged CLO GlxI was further purified by gel permeation chromatography utilizing Tris buffer (50 mM Tris (pH 8.0), 150 mM KCl and 10% glycerol) with a flow rate of 0.5 mL/min on a Superdex75 HR 10/30 column. The purified protein was then dialyzed against MOPS buffer (50 mM MOPS (pH 6.0), 150 mM KCl and 10% glycerol) overnight at 4 °C. The protein concentration was calculated by Bradford assay using BSA as a standard [134].

An alternative protocol for the protein purification was performed on a gradient HisTrap affinity column. Instead of eluting His-tagged protein with 100% elution buffer at one time, a gradient of 0–60% eluting buffer over a 60 min interval with the flow rate of 0.5 mL/min was able to isolate the purified His-tagged CLO GlxI at a retention time of 72–86 min (25–60% of eluted buffer). This is a facile one-step purification, which was found to be less time consuming than the previous protocol. The flow chart of protein purification was summarized in Appendix 2 (Figure A2.1).

Non His-tagged CLO GlxI

All DNA manipulations and purifications were performed according to the protocols outlined by Sambrook and Russell [135]. The pSGX3-CLO GlxI plasmid and proteins were supplied from NYSGXRC; however, in attempt to produce more proteins, non His-tagged CLO GlxI was generated by recloning *clo glxI* into pET-28b(+) expression vector (Novagen, Mississauga, ON) using PCR and *Nde*I and *BamH*I restriction enzymes. The primers (Sigma-Genosys Canada, Oakville, ON) were designed as in Table 2.1. PCR reagents and plasmids were prepared as in Table 2.2, and thermal cycles were set as in Table 2.3. The PCR plasmids as well as pET-28b(+) vector were digested with *Nde*I and *BamH*I restriction enzymes. The

resulting desired plasmid fragments were purified by gel extraction using a QIAquick Gel Extraction Kit followed by ligation using T4 DNA ligase at 16 °C overnight. The circular DNA was heat shock transformed into competent *E. coli* DH5α cells followed by cell growth on Kan (30 μg/mL) containing agar plates overnight at 37 ºC. A single colony was picked and inoculated overnight in LB (5 mL) containing Kan (30 μg/mL) at 37 ºC in a rotary shaker (220 rpm). The plasmid was purified using a QIAprep Spin Miniprep Kit and sent for sequencing when appropriate (Molecular Biology Core Facility, University of Waterloo). The summary for DNA cloning is shown as a flow chart in Appendix 1 (Figure A1.1).

Table 2.1: The forward and reverse primers for pET-28b-*clo glxI* DNA cloning

Primer name	***Sequence of primer (5' to 3')***
T7.For[†]	TAATACGACTCACTATAGG
CLO.Rev[†]	GGCGAATTCGGATCCTACTTTTCTAATAATTC[¥]

[†] The melting temperatures (T_m) for T7.For and CLO.Rev primers are 45 and 59 ºC, respectively.
[¥] The grey highlights indicate *BamH*I restriction site and the underlined letters are coding for the C-terminus of the protein.

Table 2.2: PCR recipe for CLO GlxI encoding gene used for DNA amplification

Plasmids/Reagents	***Stock Concentration***	***Amount Used (μL)***	***Final Concentration***
DNA template	50 ng/μL	1	50 ng
T7.For	20 μM	1	0.4 μM
CLO.Rev	20 μM	1	0.4 μM
dNTPs	10 mM	1	0.2 mM
pfu Buffer	10×	5	
ddH_2O		40	
pfu polymerase		1	
Total		50.0	

Table 2.3: PCR programming set for DNA cloning of CLO GlxI

Steps	***Temperature (ºC)***	***Time***	
Initial Denaturation	95	5 min	
Denaturation	95	30 sec	Repeat 30 cycles
Annealing	55	1 min	
Elongation	72	1 min	
Final Extension	72	10 min	
Hold	4		

The pET28b-CLO GlxI plasmid was heat shock transformed into *E. coli* BL21 (DE3) for protein purification purposes. Protein expression, induction and purification using a HisTrap affinity column were utilized as stated previously. Eluted His-tagged proteins were dialyzed against Tris buffer (50 mM Tris (pH 8.0) and 500 mM KCl) overnight at 4 °C.

To the dialyzed protein, thrombin protease (1 unit thrombin/100 μg protein) and $CaCl_2$ (2.5 mM) were added before incubating overnight at 4 °C in order to cleave the N-terminal His-tag. Imidazole (20 mM) was then added to the reaction before loading onto the combined HisTrap HP and HiTrap Benzamidine FF affinity columns. Purified CLO GlxI was collected in the wash buffer before being dialyzed at 4 °C overnight against 50 mM HEPES (pH 7.0), 200 mM KCl, 10 mM Met and 10% glycerol using a 12–14 kDa cut-off dialysis membrane. The concentration of dialyzed proteins was calculated by the Bradford assay using BSA as a standard. The flow chart of the protein purification is summarized in Appendix 2 (Figure A2.2).

2.3.3: Apo-enzyme Preparation

Buffers utilized for the purified enzyme were exchanged with a Chelex-treated buffer, which were prepared prior to use with the protein by running buffer through a Chelex 100 resin column to remove all metal ions. The determination of metal concentration in aqueous protein samples and buffer control samples were performed using a reported 4-(2-pyridylazo)resorcinol (PAR) assay [136-138]. This assay is very sensitive and can be used to measure nanomole levels of Zn^{2+}, Ni^{2+}, Co^{2+} and Cu^{2+} depending on their corresponding absorption wavelengths (490 nm for Ni^{2+}/Zn^{2+} and 514 nm for Co^{2+}/Cu^{2+}). Two of these metallochromic indicators can form a complex with one atom of metal, which can be detected as a color change (yellow to orange color with Zn^{2+}) (Figure 2.3).

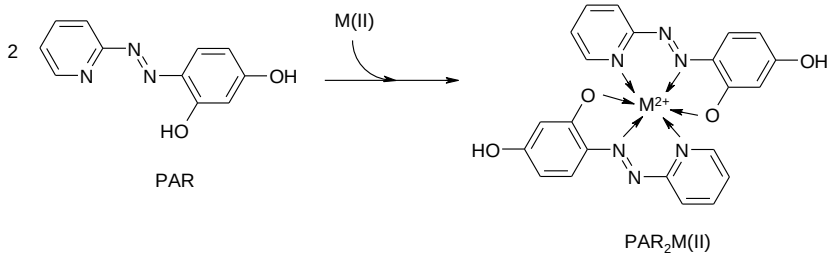


Figure 2.3: The colorimetric assay using 4-(2-pyridylazo)resorcinol (PAR) to quantitate Co^{2+}, Ni^{2+}, Cu^{2+} and Zn^{2+} in the sample [136-138].

In order to perform the PAR assay, the protein was incubated with Proteinase K (50 μg/mL) in Chelex-treated buffer to the total volume of 50 μL before heating at 56 °C for 30 min. The reaction was mixed with a dye buffer (50 μL) containing 5 mM iodoacetamide and 0.2 mM PAR. The absorption at 490 nm was then measured using a microplate reader. The concentration of zinc was calculated from a standard plot (0–16 μM $ZnCl_2$).

2.3.4: Molecular Mass Determination

The molecular mass of a denatured protein was determined by ESI-MS and SDS-PAGE, while the molecular mass of its native form was determined by gel permeation chromatography (Superdex75 HR 10/30 column). As well, disulfide bond formation between subunits (if any) was investigated in the presence of DTT (20 mM). A standard curve was prepared using Sigma protein standards containing BSA (66 kDa), carbonic anhydrase (29 kDa), cytochrome *c* (12.4 kDa) and apotinin (6.5 kDa).

2.3.5: Protein Stability

In order to investigate the effect of protein concentration, pH of buffer, type of buffer, ionic strength, additives and metal reconstitution on the secondary structure of His-tagged CLO GlxI, circular dichroism (CD) was performed using a Jasco J-715 spectropolarimeter with a cuvette of 1 mm path length. The CD spectrum was scanned between 190–250 nm with 0.2 nm step resolution, 100 nm/min speed, 25 accumulations, 1 sec response, 1 nm band width, 50 mdeg sensitivity and 500 μm slit width. All experiments were performed under nitrogen gas.

The melting temperature (T_m) was calculated using differential scanning calorimetry (DSC) and CD analysis as described above. The DSC experiments were performed on a MicroCal VP-DSC microcalorimeter with cell volumes of 0.5 mL and self-contained pressurizing system of 0–30 p.s.i. for scanning solutions above boiling points to prevent any degassing during heating. The sample and the reference solutions were degassed and the temperature was adjusted in an evacuated chamber at room temperature and carefully loaded into the cell of the calorimeter to eliminate bubbling effects. The experiment with His-tagged CLO GlxI (0.12 mg/mL) was run against MOPS buffer over a temperature range of 10–80 °C with a scanning rate of 1 °C/min. Several up- and down-scans were performed for reference to obtain a stable baseline. The reference was replaced by protein sample in a sample cell, which

was continued running using the same parameters as before. To test protein stability, the sample was allowed to run another round of up- and down-scans. The Origin scientific plotting software package supplied by MicroCal was used for baseline subtraction and T_m calculation by integration of the heat capacity (C_p) versus temperature (t) curve. For consistency and comparability, the integration was always done between the two points; one from where the curve starts to deviate and the other where it returns to the baseline.

2.3.6: Metal Analysis

The metal analysis of the protein sample was performed as previously described for *E. coli* GlxI [19]. The sample was prepared by exchanging buffer with Chelex-treated water, which was prepared prior to use with the protein by passing Milli-Q water through a column containing Chelex 100 resin. All containers were presoaked in acid solution (10% v/v nitric acid in water) and rinsed several times with Chelex-treated water. The protein was diluted with Chelex-treated water to a final volume of 14 mL. The sample was prepared in duplicate using Chelex-treated water as reference. The metal content was quantitated by inductively coupled plasma mass spectrometry (ICP-MS) (ETL Chemspec Analytical Ltd., Waterloo, ON). The following elements were analyzed: Al, Sb, As, Ba, Be, Bi, B, Cd, Ca, Cr, Co, Cu, Fe, Pb, Mg, Mn, Mo, Ni, P, K, Se, Si, Ag, Na, Sr, Tl, Sn, Ti, W, U, V, Zn and Zr.

2.3:7: Glyoxalase I Enzymatic Assay

Calibration of Methylglyoxal

Commercial methylglyoxal was transferred to a round bottle containing a boiling chip before heating with a heat gun in order to eliminate contaminants and MG polymers. The purified MG was collected by distillation at 90 °C using room pressure and stored at 4 °C. The MG concentration was calculated according to the developed protocol by Bergmeyer, 1984 [139]. The purified MG was diluted 50–150 fold (0.2–1.0 mM) with 50 mM KPB (pH 6.6) containing GSH (0.5–1.0 mM) and 0.1 units of commercial yeast GlxI before measuring the absorbance at 240 nm. The concentration of MG was calculated using the Beer-Lambert Law (ε_{240} = 2860 $M^{-1}cm^{-1}$) [85, 86]. The experiment was repeated in triplicate with different concentrations of MG.

Calibration of Glutathione

Various concentrations of glutathione were prepared by dissolving various amounts of solid GSH into 50 mM KPB (pH 6.6; previously degassed by bubbling argon through the solution for 30 minutes). The pH of the GSH solution (2–10 μL) was adjusted by mixing with 50 mM KPB (pH 8.0) to a final volume of 990 μL. Excess 5,5'-dithiobis-(2-nitrobenzoic acid) or DTNB (10 μL of approx. 50 mg/mL) was added into the solution and incubated (in the dark) at room temperature for 20 min before measuring the absorbance at 412 nm. The concentration of GSH was calculated using the Beer-Lambert Law (ε_{412} = 14150 $M^{-1}cm^{-1}$) [140]. The experiment was repeated in triplicate with different concentrations of GSH.

Enzymatic Assay for Glyoxalase I

The activity of enzyme assay for GlxI reaction was performed utilizing a well established protocol [19]. This method utilizes a prepared equilibrium mixture of the substrate hemithioacetal (dissociation constant of 3.1 mM), which is a non-enzymatic product of MG and GSH. The previous determined equilibrium equation is as follows; K_d = [MG][GSH]/[MG-GSH] = ([MG_0]-[MG-GSH])([GSH_0]-[MG-GSH])/[MG-GSH], where [MG] is the concentration of MG, [GSH] is the concentration of GSH, [MG-GSH] is the concentration of hemithioacetal, [MG_0] is the initial concentration of MG and [GSH_0] is the initial concentration of GSH [85, 86]. The concentration of GSH was maintained at 0.1 mM in order to avoid enzymatic inhibition [141].

The hemithioacetal was prepared by incubating MG and GSH in KPB (pH 6.6) at room temperature for 15 min. The enzymatic activity was measured as an increase in absorbance at 240 nm (ε_{240} = 2860 $M^{-1}cm^{-1}$) for the formation of *S*-D-lactoylglutathione using a 96-well UV-visible plate on a plate reader.

2.3.8: Metal Characterization of Glyoxalase I Activity

Metal Activation

His-tagged CLO GlxI (10 μg) was incubated with 10 equivalents of metals ($ZnCl_2$, $NiCl_2$, $CoCl_2$, $CuCl_2$, $MnCl_2$, $MgCl_2$, $CaCl_2$ and $CdCl_2$) in 50 mM MOPS (pH 6.0) at 4 °C overnight. The enzymatic assay was performed in triplicate at room temperature using a substrate MG-GSH concentration of 0.5 mM.

Metal Titration

Apo (metal-free) His-tagged CLO GlxI (15 μg) in KPB (pH 6.6) was incubated with 0–5 equivalents of metals ($NiCl_2$ and $CoCl_2$) at 4 °C overnight before performing the enzymatic assays with 0.5 mM substrate.

pH Profile

Apo His-tagged CLO GlxI was incubated with 50 equivalents of $NiCl_2$ overnight at 4 °C before diluting with KPB buffer with pH range from 5.8–8.0 to a final concentration of 15 μg of enzyme. This pH is within the valid range of the dissociation constant (3.1 mM) for the substrate formation [85, 86]. The enzymatic assay was performed in triplicate using KPB at the particular pH with 0.04–1 mM substrate being utilized. The effect of buffer ionic strength was performed separately on the apo-enzyme (15 μg) with 5 equivalents of $NiCl_2$ in KPB (pH 7.0) in the presence of 0.05–1 M KCl.

2.3.9: Kinetic Studies

The Michaelis–Menten constant, K_m, and maximum velocity, V_{max}, were determined using Ni^{2+}-reconstituted non His-tagged CLO GlxI (795 ng) with ten substrate concentrations (0.04–1.0 mM) in KPB and MOPS buffer with pH of 6.0 and 7.0. The assay was performed using a 96-well UV-visible plate on a plate reader. The initial rate was fitted by the Michaelis–Menten equation with least squares fit parameter using a GraphPad Prism version 5.00 for Windows (GraphPad Software, San Diego California USA, www.graphpad.com). The kinetics on commercial yeast GlxI (2–9 ng) was performed in parallel as a control.

2.3.10: Enzymatic Assay for Methylmalonyl-CoA Epimerase Activity

The activity of methylmalonyl-CoA epimerase on CLO GlxI was performed as previously reported by Bobik and Rasche, 2003 [107]. This is a coupled assay utilizing methylmalonyl-CoA mutase and methylmalonyl-CoA epimerase along with the commercial (*RS*)-methylmalonyl-CoA as substrate (detailed assay in Chapter 5). The (2*R*)-methylmalonyl-CoA was converted to succinyl CoA by methylmalonyl-CoA mutase (coenzyme B12 dependent), leaving its *S*-form, the specific substrate for methylmalonyl-CoA epimerase, to convert to its *R*-form. The assay reaction was monitored using a reverse phase HPLC

instrument with a μBondapak^{TM} C18 column (3.9 × 300 mm) (Waters, Milford, MA). A bacterial strain *E. coli* BL21 (DE3) RIP containing the gene for the methylmalonyl-CoA mutase was kindly supplied by Dr. Thomas Bobik from the Iowa State University, IA.

2.3.11: Structural Investigation

Apo-CLO GlxI was reconstituted with $NiCl_2$ (5 equivalents) before incubating overnight at 4 °C. Excess metal was removed by dialyzing against Chelex-treated HEPES buffer (50 mM HEPES (pH 7.0), 200 mM KCl, 10 mM Met and 10% glycerol). Ni^{2+}-reconstitution was quantitated by the PAR assay and metal analysis was performed utilizing ICP-MS before performing X-ray crystallography in collaboration with our colleagues at NYSGXRC.

Structural visualizations and molecular images were produced by the UCSF Chimera package from the Resource for Biocomputing, Visualization and Informatics at the University of California, San Francisco (supported by NIH P41 RR-01081) [1] and the PyMOL Molecular Graphics System, version 1.2r3pre, Schrödinger, LLC (http://www.pymol.org/).

2.4: Results and Discussion

2.4.1: Amino Acid Sequence Analysis

The nucleic acid sequence coding for *C. acetobutylicum* GlxI (Glyoxalase I, AAK80149) is in the region between 2284462 to 2284848 on the minus strand of the *C. acetobutylicum* ATCC 824 plasmid, consisting of 384 nucleic acids (Figure 2.4). The protein consists of 128 amino acids with a predicted molecular weight of 14608.8 Daltons and a predicted p*I* of 6.31 (Figure 2.4). His-tagged CLO GlxI was generated to possess a replacement of the first Met residue at the N-terminus with three extra amino acids, MSL, as well as an addition of a His_6-tag at the C-terminus (EGHHHHHH), which changes the molecular mass of this enzyme to 15818.1 Da and a predicted p*I* to 6.54. This His-tag was used for purification purposes and was left uncleaved. The non His-tagged CLO GlxI, on the other hand, was engineered to leave three extra N-terminal GSH amino acids after cleavage by thrombin protease. The molecular weight was changed to 15090.3 Da with a predicted p*I* of 6.49.

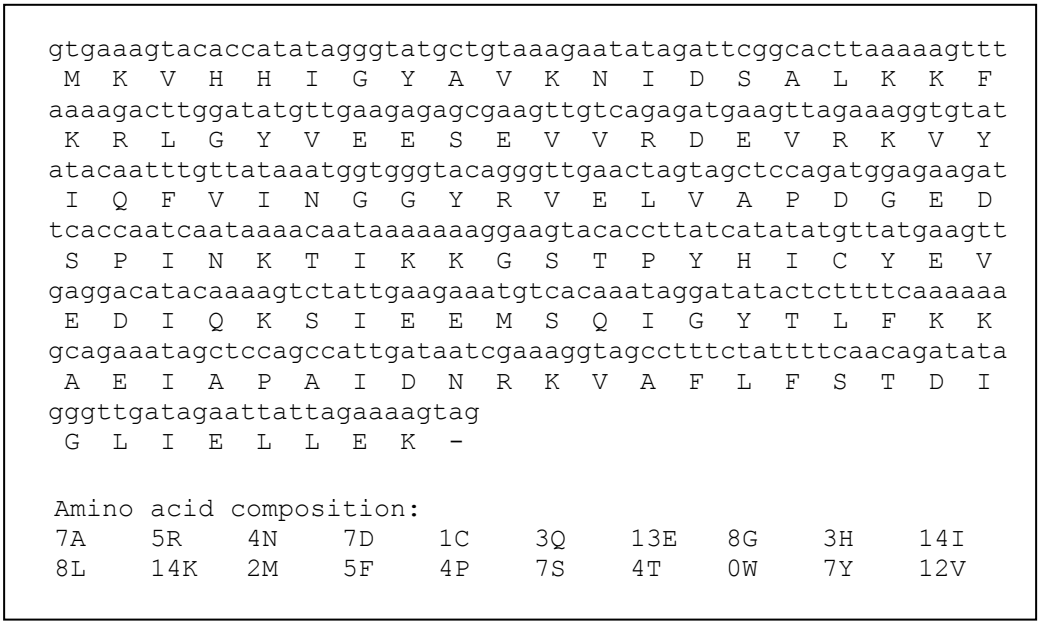


Figure 2.4: The DNA sequence (384 nucleic acids), amino acid sequence (128 residues) and amino acid composition of *C. acetobutylicum* Glyoxalase I.

2.4.2: Protein Characterization

The purification of recombinant CLO GlxI was readily accomplished making use of a two-step protocol composed of a HisTrap affinity step followed by gel permeation chromatography. The His_6-tag was left uncleaved at the C-terminus. However, the presence of the His-tag most likely is the cause of enhanced protein aggregation for this form of the enzyme since oligomers may occur due to interaction with Ni^{2+} ions in solution. The non-His-tagged enzyme was then cloned and purified in an attempt to increase recovery yield and stability. The detailed purifications are as follows.

His-tagged CLO GlxI

In attempt to optimize protein expression, growth and induction temperatures were varied. Two growth and induction temperatures were examined at 25 °C and 37 °C. Higher levels of overexpressed proteins were observed at 37 °C with induction by 1 mM IPTG for three hours (Figure 2.5). However, more soluble proteins were found when induction and expression were accomplished at 25 °C (Figure 2.6). A cell pellet of ~3.8 g was harvested from 1 L of LB culture. The bacterial cell wall was then disrupted by sonication and the resulting cell lysate was centrifuged. The supernatant was loaded onto a HisTrap column to

eliminate *E. coli* proteins (Figure 2.7). The eluted His-tagged proteins from this affinity column were approximately 3.56 mg/mL (12.67 mg total proteins) in concentration and were dialyzed against MOPS buffer (50 mM MOPS (pH 8.0), 150 mM KCl and 10% glycerol) overnight at 4 °C. However, precipitation was observed, suggesting that the proteins were not stable under these buffer conditions. An addition of 4 mM EDTA and two-fold dilution with dialysis buffer were employed to decrease protein aggregation. As well, the pH of the buffer was found to play a significant role, where a higher recovery yield of proteins in MOPS buffer pH 6.0 (~15 mg) was collected, compared to proteins in buffer pH of 8.0 (~8 mg). Further purification by gel permeation chromatography indicated that the eluted fractions from the HisTrap affinity column contained mainly His-tagged CLO GlxI and trace amounts of *E. coli* proteins (Figure 2.8). The purity of proteins from each purification step was visualized by SDS-PAGE (Figure 2.9A).

The molecular weight (MW) determination of the purified His-tagged CLO GlxI was performed utilizing SDS-PAGE, ESI-MS and gel permeation chromatography. The enzyme was observed as a single band on the SDS-PAGE with an approx. MW of 16 kDa. The mass spectrometric experiments provided a more precise MW of 15685.5 Da, corresponding to the mass of the enzyme without the N-terminal Met residue (calculated MW of 15686.9 Da) (Figure 2.9B). Gel permeation chromatography using a Superdex75 HR 10/30 column or a Superose6 10/300 GL column showed a MW of approx. 30.3 kDa and 31.8 kDa in buffer containing DTT and without DTT, respectively (Appendix 3). These results indicated that the native His-tagged CLO GlxI forms a dimer and the dimeric interaction does not involve disulfide bond formation.

Due to disadvantages of gel permeation chromatography (time consuming, labor and protein loss), a gradient HisTrap affinity column was developed to improve the efficiency of the protein purification. Instead of using 100% eluting buffer containing 500 mM imidazole to elute all His-tagged proteins at one time as previously described, a gradient of 0–60% elution buffer was employed. A trace amount of His-tagged CLO GlxI was eluted with 10% eluting buffer along with most of the *E. coli* proteins, while the purer enzyme was eluted after 25% eluting buffer (Figure 2.10–2.11). The purified His-tagged CLO GlxI was collected with 60% of eluting buffer. This protocol eliminated some precipitates, while its recovery yield was slightly decreased.

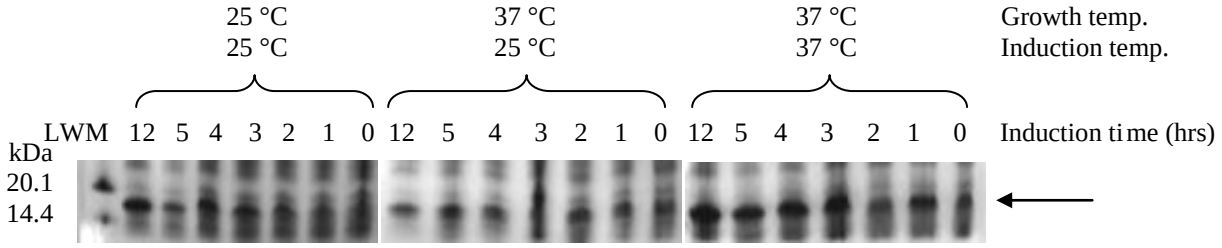


Figure 2.5: The SDS-PAGE showing induction tests of His-tagged CLO GlxI under different growth and induction temperatures (25 and 37 °C). The arrow indicates the presence of His-tagged CLO GlxI.

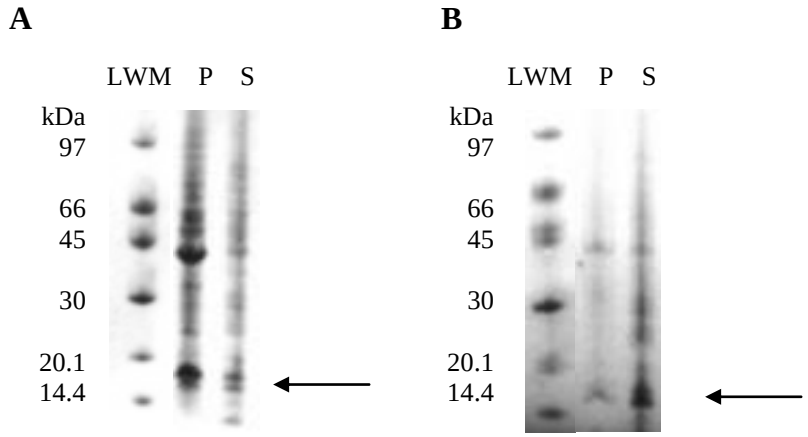


Figure 2.6: The SDS-PAGE for solubility test of His-tagged CLO GlxI with **(A)** 37 °C and **(B)** 25 °C growth and induction temperatures: LWM = low molecular weight marker, P = pellet after sonication and S = supernatant after sonication. The arrow indicates the presence of His-tagged CLO GlxI.

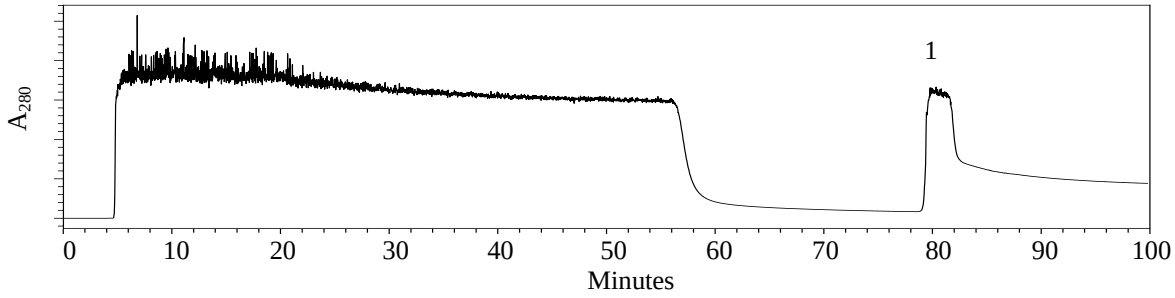


Figure 2.7: The chromatogram of His-tagged CLO GlxI purification using a HisTrap affinity column with a flow rate of 0.5 mL/min and a buffer containing 20–500 mM imidazole. The His-tagged proteins were visualized as a peak labeled number 1, which eluted with buffer containing 500 mM imdazole (80–83 min).

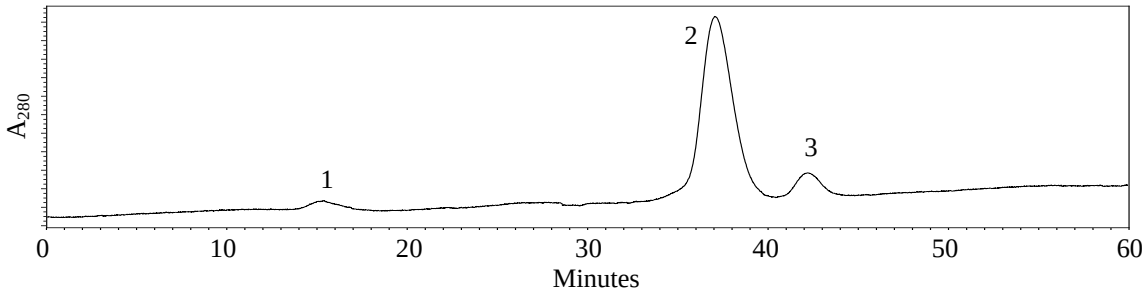


Figure 2.8: The gel permeation chromatograms of His-tagged CLO GlxI (3.65 mg/mL) using a Superose6 10/300 GL column in 50 mM Tris (pH 8.0) and 150 mM KCl with 0.5 mL/min flow rate. Peak 2 contained His-tagged CLO GlxI as indicated by SDS-PAGE. The molecular mass of approx. 27.5 kDa was calculated utilizing a standard curve.

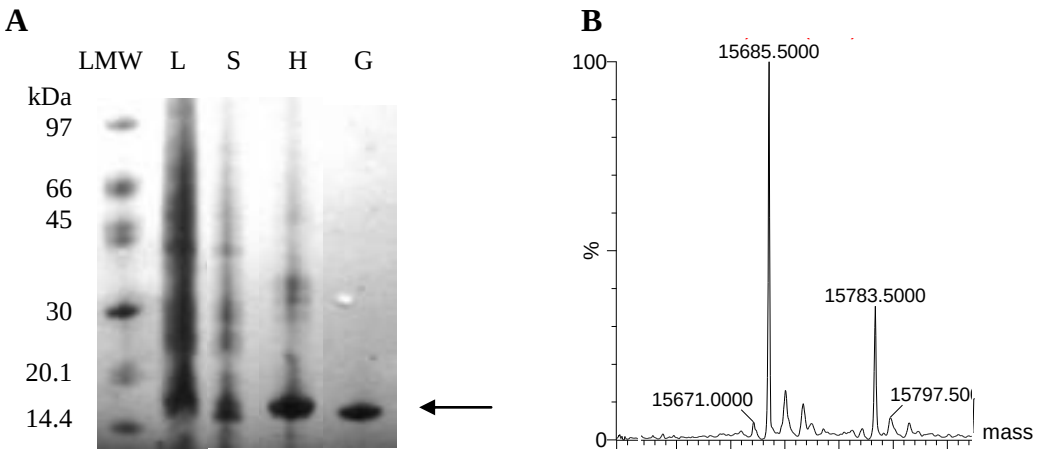


Figure 2.9: (A) The SDS-PAGE of His-tagged CLO GlxI purification: LMW = low molecular weight marker, L = lysate, S = supernatant after sonication, H = HisTrap fraction and G = gel filtration fraction. The arrow indicates the His-tagged CLO GlxI. **(B)** The electrospray mass spectrum of His-tagged CLO GlxI shows the molecular weight of a single subunit at 15685.50 Da (calculated MW is 15686.9 Da) and its adduct with $2K^+$ and Na^+ atoms at 15783.5 Da.

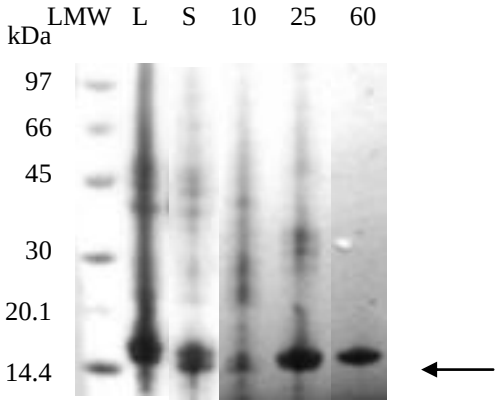


Figure 2.10: The SDS-PAGE of His-tagged CLO GlxI purification using a gradient HisTrap affinity column: LMW = low molecular weight marker, L = lysate, S = supernatant after sonication, 10 = 10% eluted fraction, 25 = 25% eluted fraction and 60 = 60% eluted fraction. The arrow indicates the His-tagged CLO GlxI.

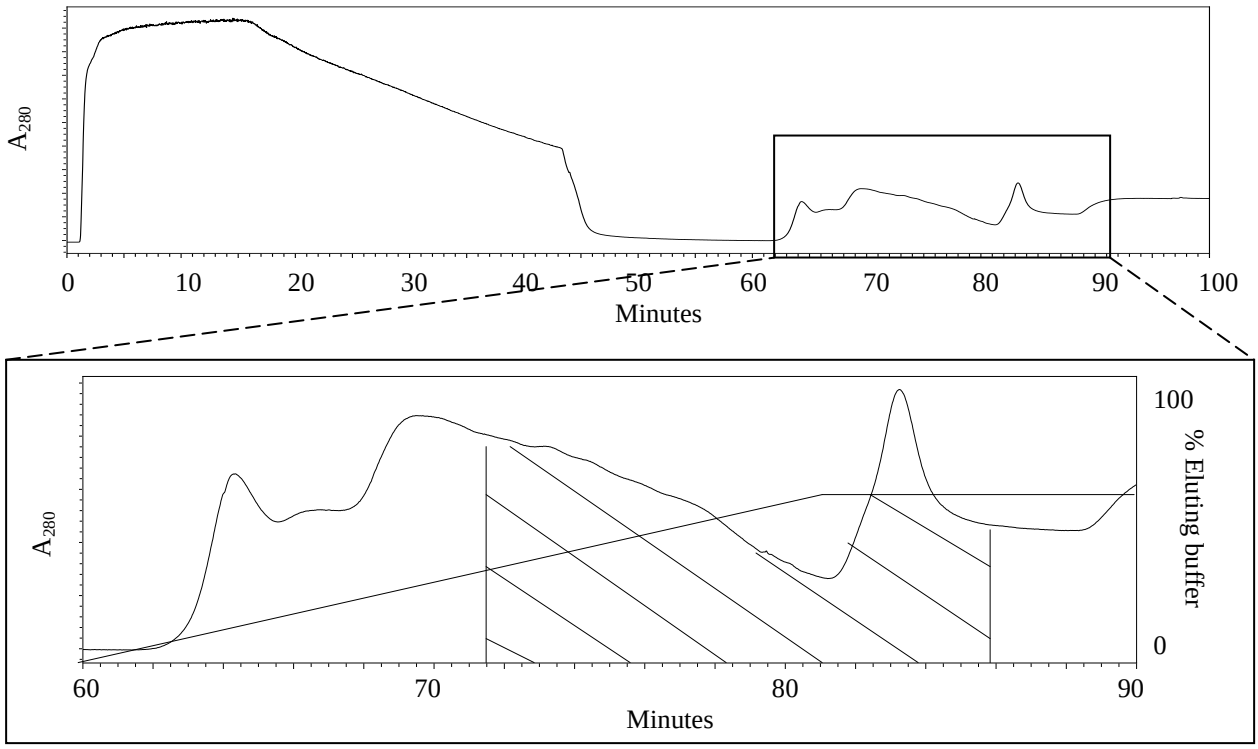


Figure 2.11: The chromatogram of gradient HisTrap affinity column purification with the flow-through fraction at 0–55 min and the elution fraction containing His-tagged proteins at 60–100 min (peak 1). The shaded area indicated the CLO GlxI was eluted with 25–60% eluting buffer.

Non His-tagged CLO GlxI

Non His-tagged CLO GlxI was generated by utilization of the polymerase chain reaction (PCR) with *Nde*I and *BamH*I restriction enzymes. The gene of interest was cloned into a pET-28b expression vector and heat shock transformed into an *E. coli* system to express the protein with the N-terminal His_6-tag followed by a thrombin protease cleavage site. The produced cell pellet of ~5 g was harvested from a 1 L culture. After protein purification using HisTrap affinity columns, protein precipitation was observed when dialyzing against Tris buffer pH of 8.0, but no precipitation occurred when using the buffer pH of 6.0. All buffers involved in purification of CLO GlxI were then adjusted to this pH. Thrombin protease was added to the dialyzed proteins (2.24 mg/mL or 8.97 mg total proteins from 2.58 g cell) to cleave off the N-terminal His-tag. Purified CLO GlxI was collected from the combined HisTrap and Benzamidine affinity columns (0.7 mg/mL, 5.2 mg total enzyme). The SDS-PAGE suggested a molecular weight of approx. 15 kDa, while ESI-MS provided a more precise molecular weight of 15089 Da (calculated MW of 15090.3 Da) (Figure 2.12). Post-

translational modification was not observed for this protein. The gel permeation chromatography using a Superdex75 HR 10/30 column indicated that the quaternary structure of the non His-tagged enzyme is a dimer with the molecular weight of approximately 32 kDa, suggesting that the Histag does not interfere with the dimeric formation of the enzyme.

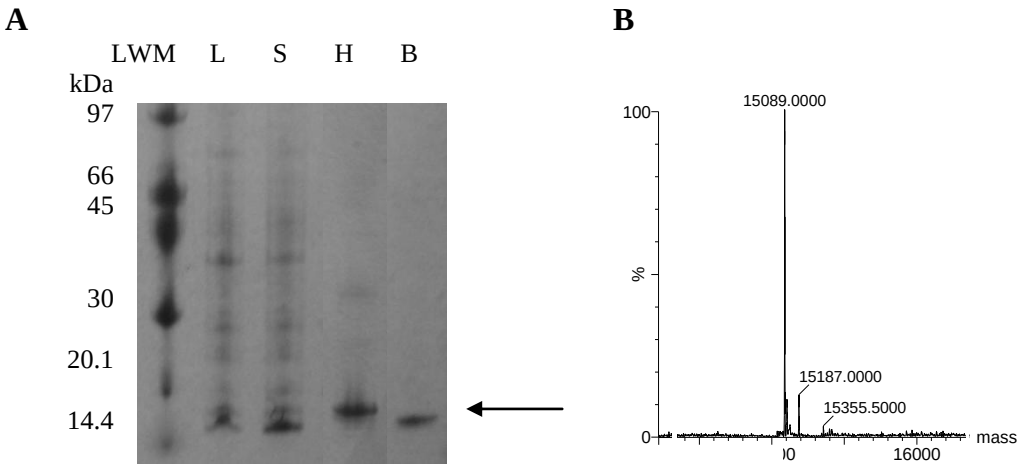


Figure 2.12: (A) SDS-PAGE results for CLO GlxI purification: LMW = low molecular weight marker, L = lysate, S = supernatant after sonication, H = HisTrap fraction and B = Benzamidine and HisTrap fraction. The arrow indicates a band of CLO GlxI. **(B)** The electrospray mass spectrum of CLO GlxI shows the molecular weight of a single subunit at 15089 Da (calculated MW is 15090.3 Da).

2.4.3: Protein Stability

The purpose of this investigation was to find the optimum conditions for protein stability and increase the yield of soluble protein for X-ray crystallography. Protein precipitation of His-tagged CLO GlxI was observed through the use of CD spectroscopy and gel permeation chromatography. Investigations toward the effects of protein concentration, temperature, pH, buffer, additives (salt and glycerol) and metals were employed. Among all these, only protein concentration, temperature and pH of buffer created a significant effect on protein stability as discussed in the following details.

The secondary structure as a function of protein concentration as monitored by CD spectroscopy indicated that the protein at low concentrations (<7.80 μM or <0.12 mg/mL) exhibited a negative maximum at 208 nm with a shoulder at 215 nm, which is a characteristic of a predominantly β-sheet containing protein (Figure 2.13). These results are consistent with the X-ray crystallographic structure of CLO GlxI (PDB: 2QH0), which consists of a βαβββ arrangement. A minor change in shape and intensity of the molar ellipticity [Θ] at 215 nm on

the series of proteins with concentration lower than 0.12 mg/mL was observed ($[\Theta]_{215}$ is approx. -18 000 deg.cm^2.dmol^{-1}), suggesting that these proteins are stable and there is no significant change in their secondary structures. However, a dramatic decrease in intensity at $[\Theta]_{215}$ and a shift of the negative maximum at 215 nm to 222 nm were detected with higher protein concentrations, suggesting either protein aggregation or a change in the quaternary structure (i.e. monomeric protein at low concentration and dimeric formation at higher protein concentration). Besides, the quality of CD signal of these samples is significantly affected by protein concentrations (more than 7.8 μM) at wavelengths below 215 nm, which cause saturated high tension (HT) voltage in the instrument. The analysis of the results of gel permeation chromatography indicated that there was no effect on the quaternary structure for protein concentrations of 0.6-9.6 μM (Figure 2.14). However, the formation of higher protein oligomers was clearly observed at higher protein concentration (data not shown), thus the alteration in the CD spectra with concentrations indicated likely protein aggregation and that soluble protein was found at concentrations lower than 7.8 μM.

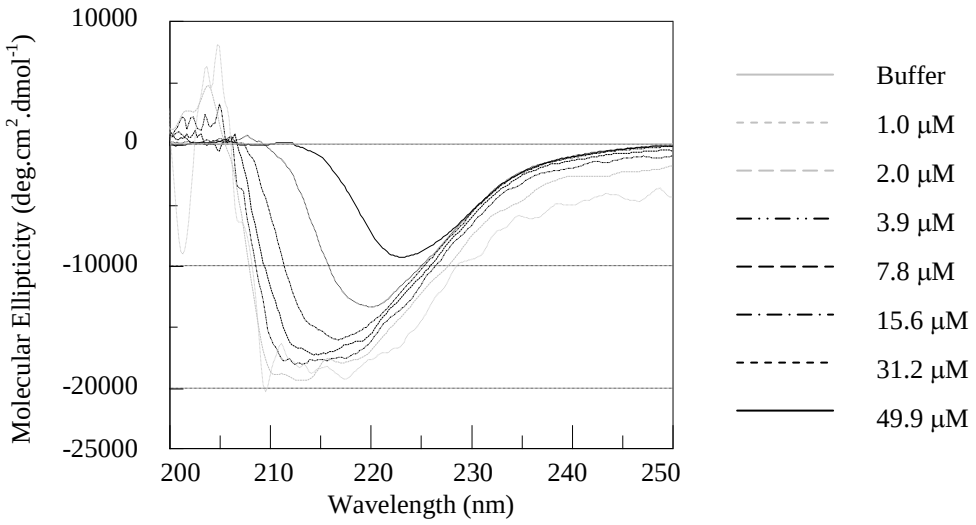


Figure 2.13: The CD spectra of various concentrations of His-tagged CLO GlxI (49.9, 31.2, 15.6, 7.8, 3.9, 2.0 and 1.0 μM) in MOPS buffer (50 mM MOPS (pH 8.0), 150 mM KCl and 10% glycerol) scanned between 200–250 nm. The CD analysis indicated that the protein was stable when the concentration was less than 7.8 μM (0.12 mg/mL).

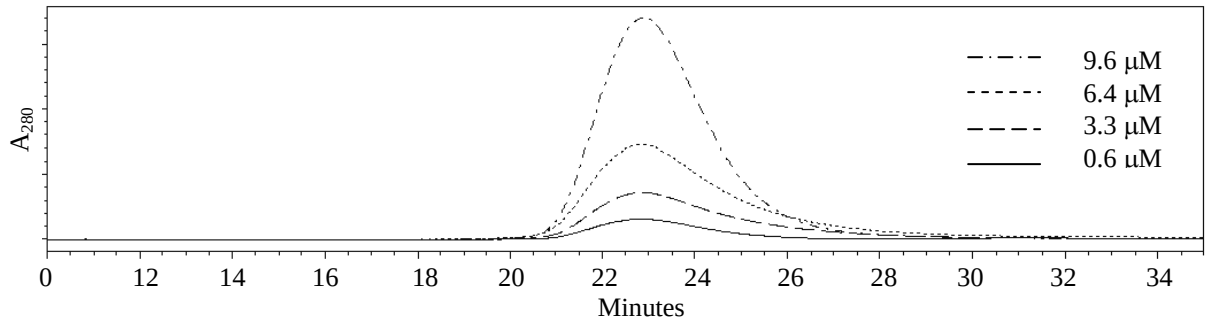


Figure 2.14: The gel permeation chromatograms of CLO GlxI using a Superdex75 HR 10/30 column and a flow rate of 0.5 mL/min with MOPS buffer (pH 6.0) exhibit a peak of dimeric His-tagged CLO GlxI measured at various protein concentrations (9.6, 6.4, 3.2 and 0.6 μM). The molecular weights of the samples were calculated using a protein standard plot indicated that 9.6 μM His-tagged CLO has a MW of approx. 30.3 kDa, 6.4 μM His-tagged CLO has MW of approx. 31.5 kDa, 3.2 μM His-tagged CLO has MW of approx. 31.5 kDa and 0.6 μM His-tagged CLO has MW of approx. 30.3 kDa.

Since the decrease in intensity at $[\Theta]_{215}$ and a shift of the negative band at 215 nm to 222 nm indicated protein aggregation, the decrease in $[\Theta]_{208}/[\Theta]_{215}$ ratio suggested decreased stability of the protein toward increased temperatures (Figure 2.15). The melting temperature (T_m) of His-tagged CLO GlxI (0.12 mg/mL) determined from the plot of $[\Theta]_{208}$ versus temperature was approx. 61.3 °C (Figure 2.16). However, the DSC experiments on this enzyme (0.70 mg/mL and 0.12 mg/mL) in MOPS buffer pH of 8.0 and 6.0 exhibited aggregation at 23 °C. It was also possible that after the protein unfolded, it quickly aggregated such that DSC could not be used to determine its T_m. These results also implied that the thermal denaturation of the protein is an irreversible process.

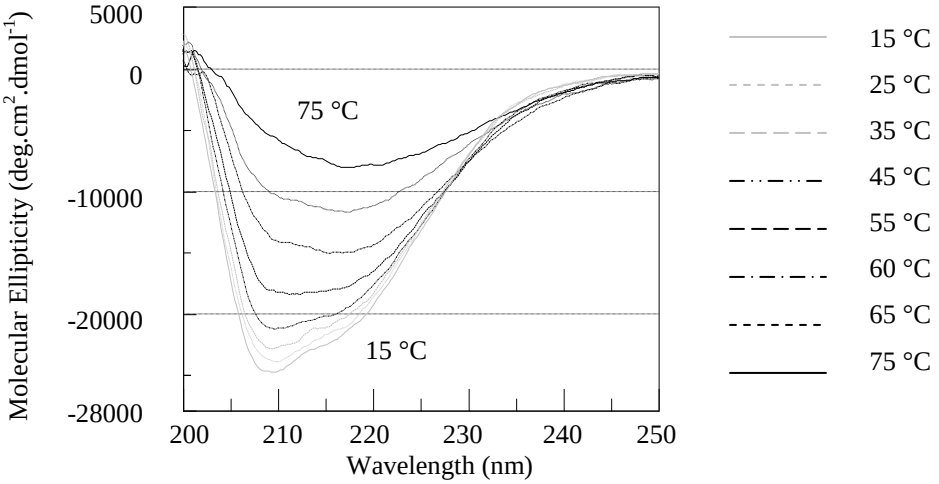


Figure 2.15: The thermal stability of His-tagged CLO GlxI (0.12 mg/mL) in MOPS buffer (50 mM MOPS (pH 6.0), 150 mM KCl and 10% glycerol) measured by CD experiments under different temperatures (15–75 °C). The CD analysis indicated that the protein stability decreased with increasing temperature.

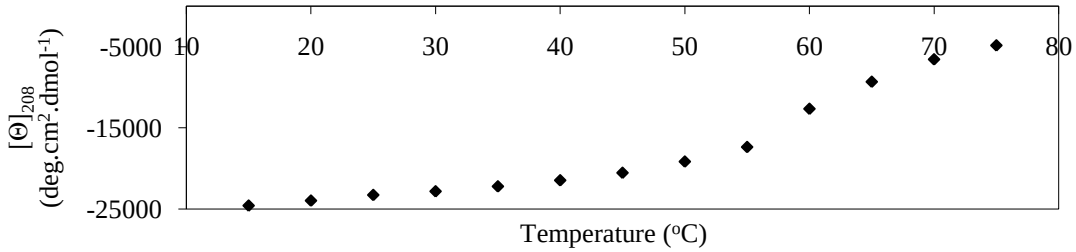


Figure 2.16: The melting temperature (T_m) of the His-tagged CLO GlxI (0.12 mg/mL) in MOPS buffer (50 mM MOPS (pH 6.0), 150 mM KCl and 10% glycerol) determined from the plot of $[\theta]_{208}$ versus temperatures (15–75 °C). The T_m determined from this graph was 61.3 °C (the T_m was calculated using GraphPad Prism program with the non linear regression fit with dose-response analysis).

The effect of pH on the secondary structure of the protein was similarly monitored. Despite there being no observable change in the quaternary structure (Figure 2.18), the negative maximum band at 208 nm ($[\Theta]_{max}$ = -15 000 deg.cm^2.dmol^{-1}) with the shoulder at 222 nm of the sample in buffer pH 5.0–6.0 was shifted to a negative maximum at 215 nm, where $[\Theta]_{208}$ decreased to -11 000 deg.cm^2.dmol^{-1} when the sample was incubated in buffer pH 8–10 (Figure 2.17). The decrease in intensity and the shift of the negative maximum band suggested that the protein is more stable at pH values lower than 7.0. However, the protein in buffer solutions of pH 4.5 or at pH 7.0 possessed negative maxima at 211 nm, which was likely caused by the structural change from β-sheet to aggregate form. In buffer at pH of 6.5, there was considerable precipitation due to the proximity to the p*I* of the protein, which was indicated by a marked decrease in intensity at 208 nm. Therefore, a buffer of pH of 6.0 was chosen as the most suitable buffer for storage of the enzyme.

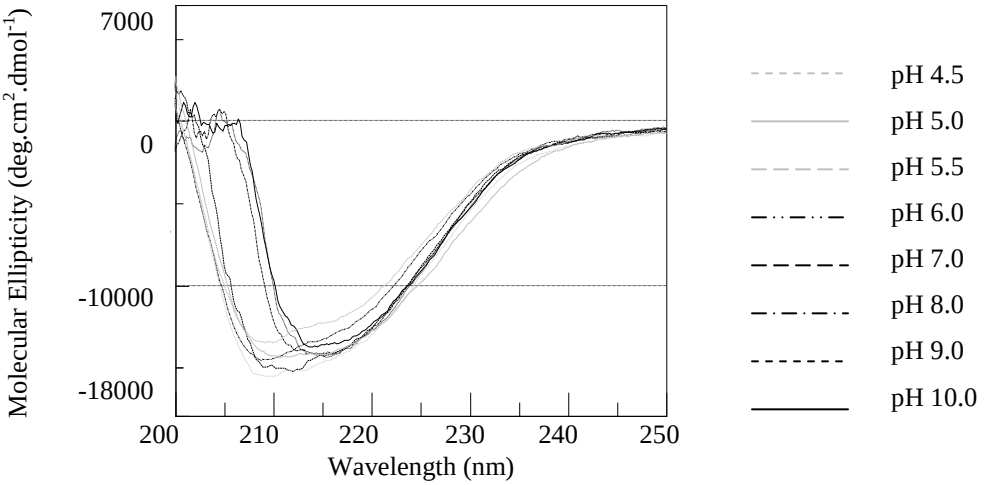


Figure 2.17: The CD spectra of His-tagged CLO GlxI (0.10 mg/mL) in MOPS buffer (50 mM MOPS, 150 mM KCl and 10% glycerol) adjusted to different pH values (pH 4.5, 5.0, 5.5, 6.0, 7.0, 8.0, 9.0 and 10.0). The shift of negative maxima at 208 to 215 nm indicated the decrease in protein stability.

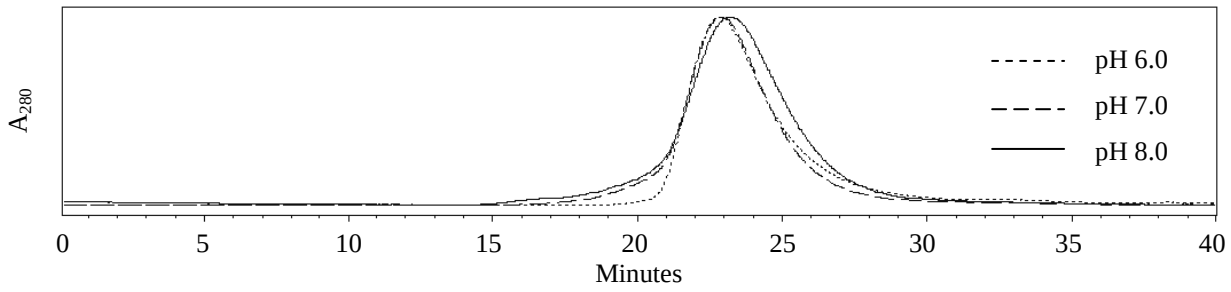


Figure 2.18: The gel permeation chromatograms of His-tagged CLO GlxI (0.15 mg/mL) using a Superdex75 HR 10/30 column and a 0.5 mL/min flow rate indicated no change in its quaternary structure in MOPS buffer pH 6.0 (MW~31.5 kDa), MOPS buffer pH 7.0 (MW~31.5 kDa) and MOPS buffer pH 8.0 (MW~29.2 kDa).

Additives such as salt or glycerol possessed only minor effects on the secondary structure of the protein. The ratio of $[\Theta]_{208}/[\Theta]_{215}$ decreased slowly, suggesting that aggregation occurs over time (Figure 2.19 and 2.21). The plot of the KCl concentration (or glycerol) versus the decrease in $[\Theta]_{208}$ per day suggested that the enzyme can be stored in buffer pH of 6.0 containing 200 mM KCl with the presence of 10% glycerol in order to optimize protein stability (Figure 2.20 and 2.22). As well, it was found that the protein in Tris buffer and KPB possessed a CD spectrum similar to those of the stable protein with negative maxima at 208 nm and a shoulder at 215 nm, indicating a predominantly β-sheet and a minority of α-helix secondary structure (Figure 2.23). The protein in HEPES, MOPS and MOPSO buffers possessed negative maxima at 217 nm similar to those of the unstable protein. Therefore, the protein should be stored in either Tris or KPB buffer.

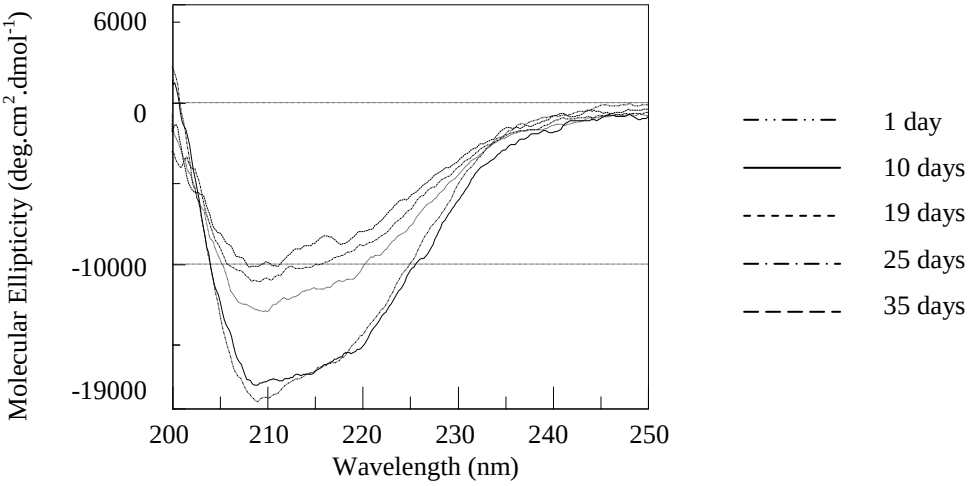


Figure 2.19: The CD spectra of His-tagged CLO GlxI (0.05 mg/mL) in MOPS buffer (50 mM MOPS (pH 6.0) and 200 mM KCl) were measured between 200–250 nm with different time periods. The decrease in intensity of negative maximum at 208 nm suggested the decrease in protein stability that occurred with time.

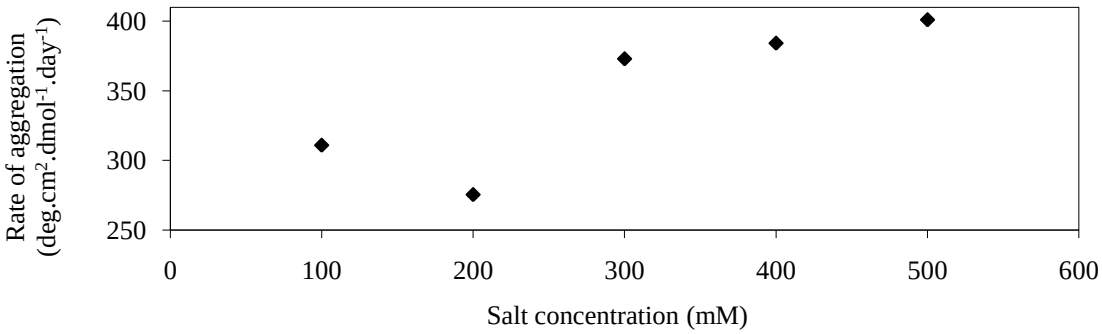


Figure 2.20: The decreasing rate of $[\theta]_{208}$ of His-tagged CLO GlxI (0.05 mg/mL) in MOPS buffer (pH 6.0) containing various concentrations of KCl (100, 200, 300, 400 and 500 mM) indicated the slowest rate of aggregated formation in buffer containing 200 mM KCl.

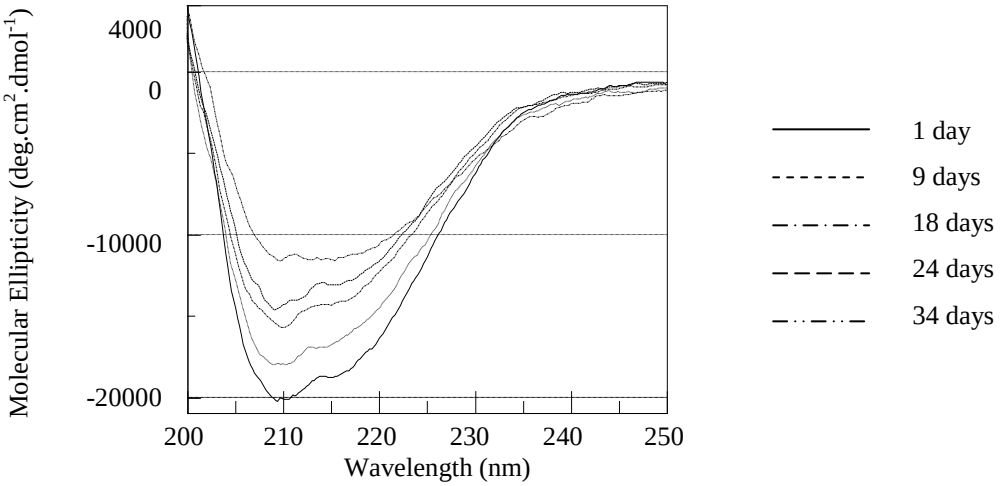


Figure 2.21: The CD spectra of His-tagged CLO GlxI (0.05 mg/mL) in MOPS buffer (50 mM MOPS (pH 6.0) and 10% glycerol) measured between 200–250 nm at different time periods. Similarly to the effect of salt, the decrease in intensity of negative maximum at 208 nm suggested a decrease in protein stability that occurred with time.

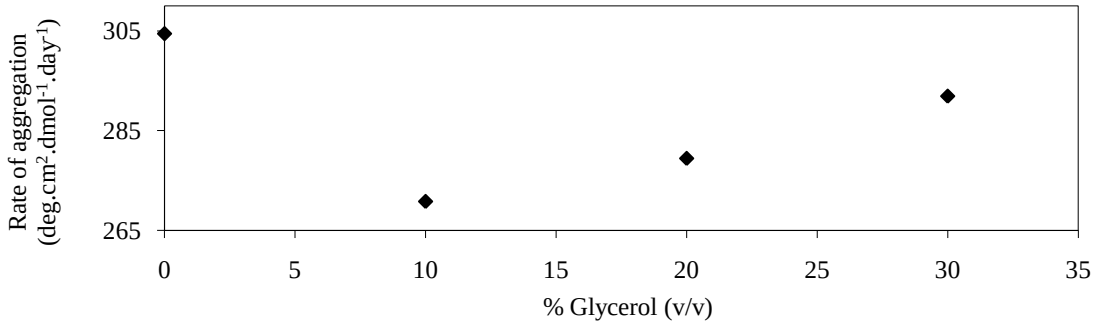


Figure 2.22: The decreasing rate of $[\theta]_{208}$ of His-tagged CLO GlxI (0.05 mg/mL) solutions in MOPS buffer (pH 6.0) containing various concentrations of glycerol (0, 10, 20 and 30% v/v) with the slowest rate of aggregation occurring with 10% (v/v) glycerol.

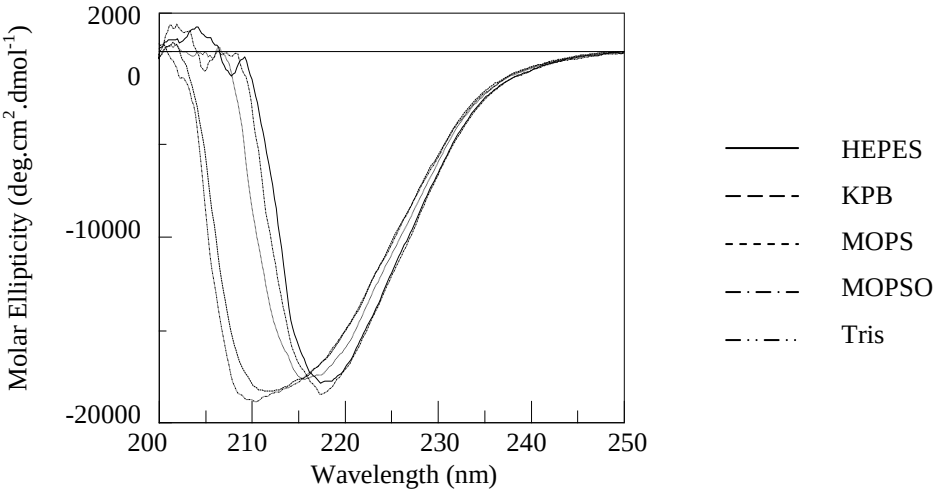


Figure 2.23: The CD spectra of His-tagged CLO GlxI (0.05 mg/mL) in HEPES, KPB, MOPS, MOPSO and Tris buffer (pH 7.0) measured between 200–250 nm. The shift of negative maxima at 208 to 217 nm suggested that Tris and KPB were suitable buffers to stabilize the enzyme.

For the effect of metals on the structure/stability of the protein, apo-His-tagged CLO GlxI was incubated with two equivalents of metals ($ZnCl_2$ and $NiCl_2$) overnight at 4 °C. It was found that while protein concentration and pH of the buffer were the major causes of aggregation, metal reconstitution with Zn^{2+} and Ni^{2+} atoms exhibited no significant change in either the secondary or quaternary structure of the enzyme (Figure 2.24 and 2.25).

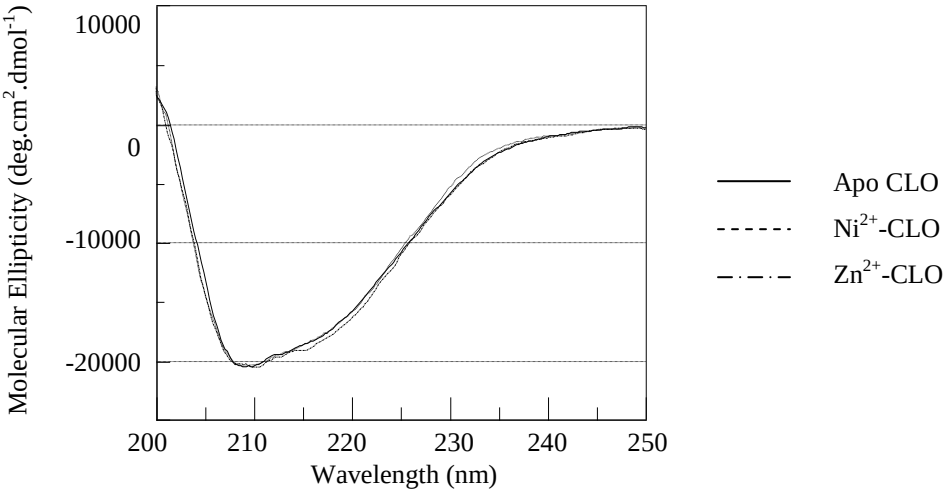


Figure 2.24: The CD spectra of His-tagged CLO GlxI (0.05 mg/mL) incorporated with different metals (Ni^{2+} and Zn^{2+}) or no metal (apo-form) in MOPS buffer (50 mM MOPS (pH 6.0) and 10% glycerol) were measured between 200–250 nm. Neither change in shape nor shift of negative maxima was observed under these conditions, suggesting that metal ions had no effect on the secondary structure of the enzyme.

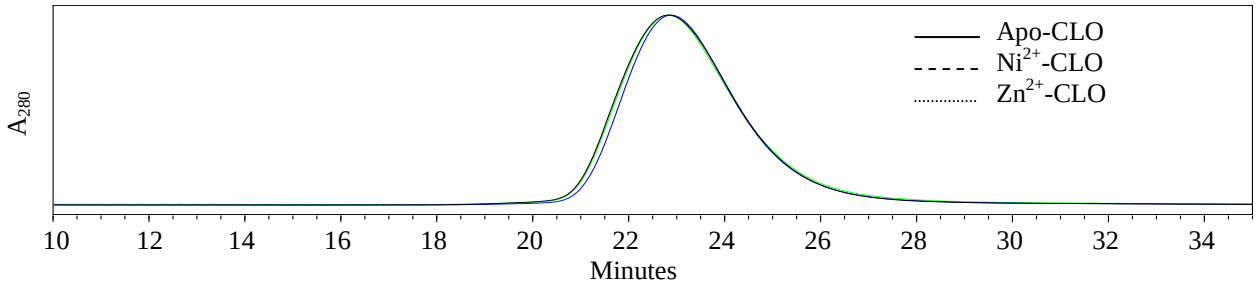


Figure 2.25: The gel permeation chromatograms using a Superdex75 HR 10/30 column and a 0.5 mL/min flow rate of the apo-His-tagged CLO GlxI (MW~30 kDa), Ni^{2+}-reconstituted His-tagged CLO GlxI (MW~27 kDa) and Zn^{2+}-reconstituted His-tagged CLO GlxI (MW~32 kDa) in MOPS buffer (50 mM MOPS (pH 6.0) and 10% glycerol) indicated that metals had no effect on the quaternary structure of the enzyme.

2.4.4: Metal Characterization

Metal analysis of the purified protein solution (0.0367 mg/mL) as determined by ICP-MS indicated the presence of lead, sodium and zinc (Table A4.1, Appendix 4). Lead and sodium are commonly found in the background, indicating that only zinc is the likely metal ion associated with this enzyme when isolated under the studied conditions. The zinc/enzyme ratio was found to be 0.042, suggesting that the purified (native) enzyme is almost in its apo-form as isolated from these expression systems (Table 2.4).

Although ICP-MS analysis and X-ray crystallographic studies of CLO GlxI indicated that zinc is the metal ion associated with this enzyme when isolated under the high throughput crystallography conditions, analysis of the metal activation profiles of the enzyme, however, indicated that the enzyme was activated by Ni^{2+}, Co^{2+}, Cd^{2+} and Mn^{2+} (activity of enzyme: $Ni^{2+} > Co^{2+} > Cd^{2+} \sim Mn^{2+}$) (Figure 2.26A). The enzymatic activity of Co^{2+}-reconstituted CLO GlxI was approx. 40%, while Mn^{2+}- and Cd^{2+}-activated forms were approx. 5% of that of the Ni^{2+}-activated enzyme. The activities with Ca^{2+} and Mg^{2+} ions were negligible. No activity was observed with the apo- and native enzyme as well as Zn^{2+}-reconstituted enzyme, even though Zn^{2+} was the metal detected from the high throughput studies. These experiments indicated that the Gram-positive CLO GlxI is a member of the Ni^{2+}/Co^{2+}-activated Glyoxalase I enzymes similar to several Gram-negative bacteria such as *E. coli*, *Y. pestis* and *N. meningitides* [19, 32]. This enzyme also exhibits no activity with Zn^{2+}, the metal ion that actually co-crystallized with the enzyme. It is likely that almost all prokaryotes (even though nothing is known for the archaea) especially bacteria (both Gram-positive and Gram-negative

bacteria) possess Ni^{2+}/Co^{2+}-activated GlxI with the possible exception of pseudomonads [19, 28, 30, 32, 33].

Analysis of the metal titration experiments indicated that full enzymatic activity was reached at a metal/dimeric enzyme ratio of approximately 1 for both Ni^{2+} and Co^{2+}-activated enzyme (Figure 2.26B). Similar results were observed with *E. coli* GlxI (one metal per dimer) and *P. falciparum* GlxI, where one metal per monomer (two active sites) were found [39, 43]. This suggests that only one active site may be functional, which is possibly due to a conformational change in one active site when bound to metal ion, a situation which does not permit the second active site to bind metal ion. Alternatively, the dimeric enzyme may be asymmetric such that only one active site can bind metal ion. Metal reconstitution of the apo-enzyme with Ni^{2+} or Zn^{2+} atoms neither showed a major conformational change in the CD spectra of the enzyme nor a change in the gel permeation chromatographic profile. These results suggest that metal binding does not have a major affect on the secondary and quaternary structures or that these techniques are not sensitive enough to detect the structural change.

Despite evidence for only one active site binding metal ion, two metal atoms have been detected in the crystallographic structure (one metal atom per active site). It was reported earlier that the X-ray crystallographic structure of apo-GlxI from *E. coli* possesses a slight asymmetry between the two active sites [45]. Three water molecules at one active site form a tighter cluster (distances from nearby water molecules to the center water molecule are 2.6 and 3.2 Å) than another (4.2 and 4.4 Å) (PDB: 1FA8). The asymmetric active sites of the apo-enzyme were confirmed by ^{15}N-^{1}H HSQC NMR experiments, which suggest the possibility that one active site possesses higher affinity for metal binding than another [38]. Addition of metal to the first active site occurs quickly and induces a subtle protein conformational change [38]. The binding of metal to the second unoccupied active site requires additional time, and both active sites of metal-reconstituted enzyme seem to be more symmetric (distances from nearby water molecules to the metal center are 2.1 and 2.2 Å for both active sites in Ni^{2+}-bound *E. coli* GlxI (PDB: 1F9Z), and 2.0 and 2.1 Å for its Zn^{2+}-bound form (PDB: 1FA5)). Our experiments also agree with the time-dependent metal binding data. The metal per enzyme ratio calculating by the PAR assay (one metal per dimer) was found to be lower than that detected by ICP-MS (two metals per dimer) (Table A4.2, Appendix 4), which took placed approx. two weeks after performing the PAR assay. Since metal titration analysis indicated

that only one active site is functional at a time, it is possible that the first metal binding in one active site might induce a conformational change such that it possesses higher affinity to accept the hemithioacetal substrate than another empty active site (more details in Chapter 6). The mystery of substrate selectivity is, however, still left unknown.

With Ni^{2+} ion as the activating metal, the enzymatic activity of CLO GlxI is optimal in KPB buffer pH of 7.0 (Figure 2.26C). Since the p*I* of the enzyme is approx. 6.5, the metal binding residues (two His and two Glu residues) in lower pH buffer are likely to be protonated, causing the observed decrease in activity. Salt was also found to produce a minor effect on enzymatic activity. It was also found that V_{max}/K_m of the enzyme (4.80 ± 0.02 μmol/min/mg/mM) was optimal in buffer containing 0.4–0.8 M KCl, suggesting that the enzyme possesses highest activity under these buffer conditions (Figure 2.26D). However, this effect seems to involve protein stability rather than any effects on the catalytic reaction itself.

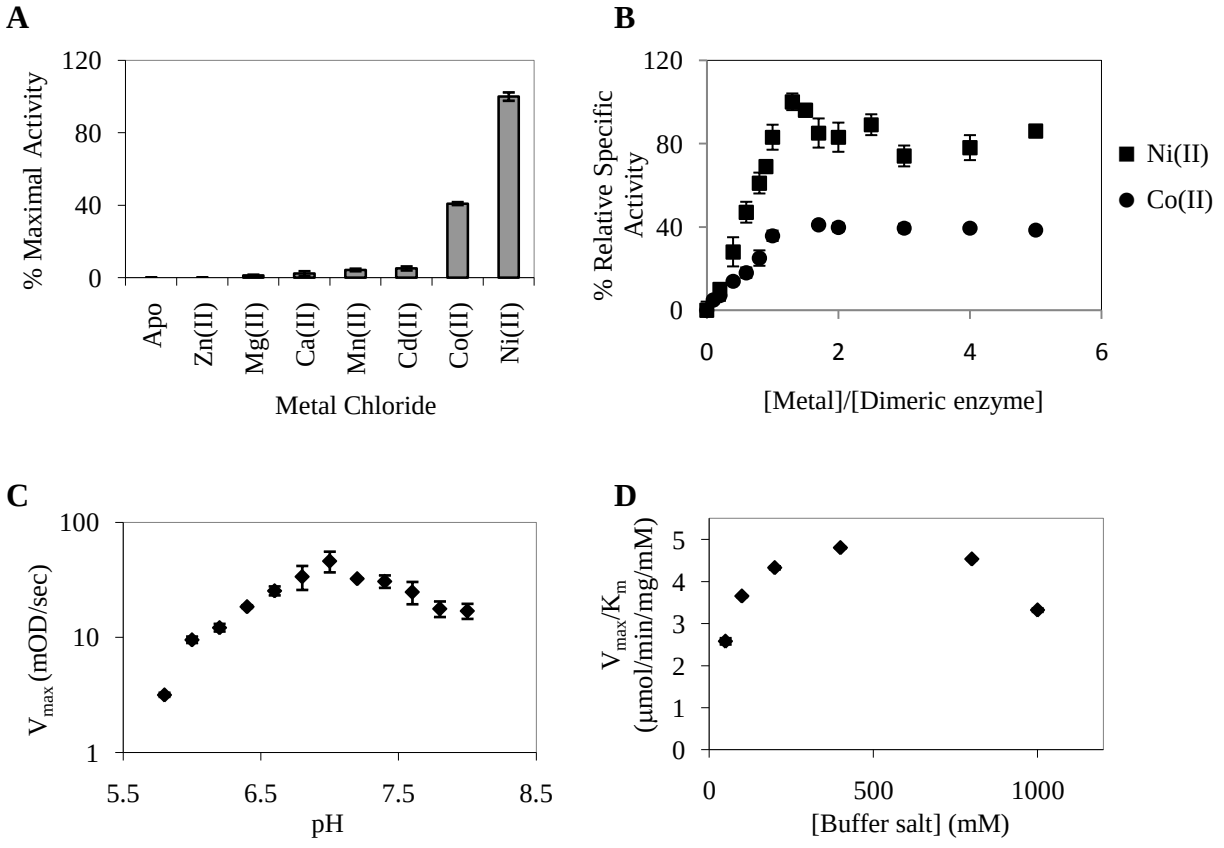


Figure 2.26: The diagram shows metal characterization of *C. acetobutylicum* GlxI (15 μg) including **(A)** metal specificity, **(B)** metal titration with (■) $NiCl_2$ and (●) $CoCl_2$, **(C)** the pH profile with $NiCl_2$ (log plot of V_{max}) and **(D)** the effect of buffer ionic strength on Ni^{2+}-activated enzyme.

Table 2.4: The calculation for the amount of metals contained in *C. acetobutylicum* GlxI using ICP-MS.

Metals	*MW (g/mol)*	*[Metal] in CLO1 (mg/L)*	*[Metal] in CLO2 (mg/L)*	*Ave. [metal] (mg/L)*	*[Metal] (mM)*	*[Metal]/[protein] ratio*
Zinc	65.39	0.006	0.006	0.006	0.09	0.042
Lead	106.40	0.002	0.002	0.002	0.02	0.009
Sodium	22.99	2.800	2.600	2.700	117.44	53.383

2.4.5: Structural Investigation

The CLO GlxI monomer consists of two repeating domains, domain I and domain II, with a α/β fold (Figure 2.27A). The overall structure has an 8-stranded β-sheet generated by edge-to-edge packing of two internally duplicated βαβββ modules. The N-terminal module or domain-I consists of β1α1β2β3β4 topology (residues 2–60) followed by the C-terminal module or domain-II with β5α2β6β7β8 arrangement (residues 72–140) connected by a short α–helix (residues 61–71) (Figure 2.27B). The core of the CLO GlxI enzyme creates a large cleft and harbors the active site with the two central strands (β1 and β5), while the outer strands (β4 and β8) cover from the side. Although the two internally duplicated βαβββ modules share structural similarity (root mean square deviation (r.m.s.d.) of 1.69 Å, 2–70 residues of domain-I to 71–130 residues of domain-II), they show little sequence identity and similarity (12% and 18%, respectively). The Zn^{2+} and Ni^{2+} ions are bound deep in the active site cleft, coordinating to the side chains of His^{5} (Nε2), Glu^{52} (Oε1), His^{75} (Nε2) and Glu^{124} (Oε1) (Figure 2.27C-D). A water molecule completes the near trigonal bipyramidal coordination in the active site of an inactive Zn^{2+}-bound CLO GlxI (Figure 2.27F), whereas its Ni^{2+}-bound form possesses octahedral geometry with two water molecules and the same four protein ligands (Figure 2.27E). The metal binding ligands come from topologically equivalent positions in the two modules; His^{5} and Glu^{52} from the first and forth β-strands of domain-I and His^{75} and Glu^{124} from equivalent positions in the first and fourth β-strands of domain-II. Despite their different metal coordinations, the overall structures of Ni^{2+} and Zn^{2+} bound CLO GlxI are well superimposed with r.m.s.d. of 0.434 Å for all 129 Cα pairs (Figure 2.28A).

It was previously found that *E. coli* GlxI possesses two different active site geometries depending on the specific metal bound to the enzyme [44, 45]. Investigation of the active site in *E. coli* GlxI indicates that the Ni^{2+}-activated form possesses octahedral geometry with four

active residues (His^{5}, Glu^{56}, His^{74} and Glu^{122}) and two water molecules around the metal center, while its catalytically inactive form with Zn^{2+} atom has a trigonal bipyramidal coordination with only one water molecule coordinated to the metal ion (PDB: 1F9Z and 1FA5, respectively) [45]. These data support our hypothesis that CLO GlxI is active in the presence of Ni^{2+} but is inactive with Zn^{2+} ions. Besides, the enzymatic activity of the GlxI family can still be predicted by the active site geometry.

PISA (Protein Interfaces, Surfaces and Assemblies), an interactive tool for the exploration of macromolecular interfaces [132], predicted the probable quaternary structure assembly of CLO GlxI as a dimer, which corresponds to the result from gel permeation chromatography. It shows that the solvent accessible area at the interface is 20.6% of the total surface with 10 hydrogen bonds and two salt bridges, providing ionic interactions that stabilize the dimeric interface. The dimeric formation of CLO GlxI is unique among small subunit proteins in the GlxI family, all of which form an active site at the interface of a dimer. In contrast, two monomers of CLO GlxI pack back-to-back by the stacking of their β sheets with their active sites opening outward in the opposite directions. The active site is thus placed away from the dimeric interface and all the active site residues are contributed by a single monomer. This type of active site formation is mostly found in methylmalonyl-CoA epimerase (MMCE), a member of the βαββββ superfamily. A list of closely resembling structures in the PDB identified through Dali search [133] indicated that most of these structures belong to dioxygenase family members, with the closest structure being MMCE from *P. shermanii* with r.m.s.d. of 2 Å (PDB: 1JC5) (Table 2.5, Figure 2.28B).

Interestingly, the analysis of multiple sequence alignment (Figure 2.2) suggests that GlxI from other species of Clostridia may have similar structures as CLO GlxI, where one active site is formed within one subunit. By using the web based Geno3D structure prediction software (http://geno3d-pbil.ibcp.fr) [142], it was found that metal binding residues are located at the active site of the predicted structure of the putative GlxI from *C. acetobutylicum* (NP_346890) when using CLO GlxI as template (Figure 2.29). These residues are not conserved when using *E. coli* GlxI as template.

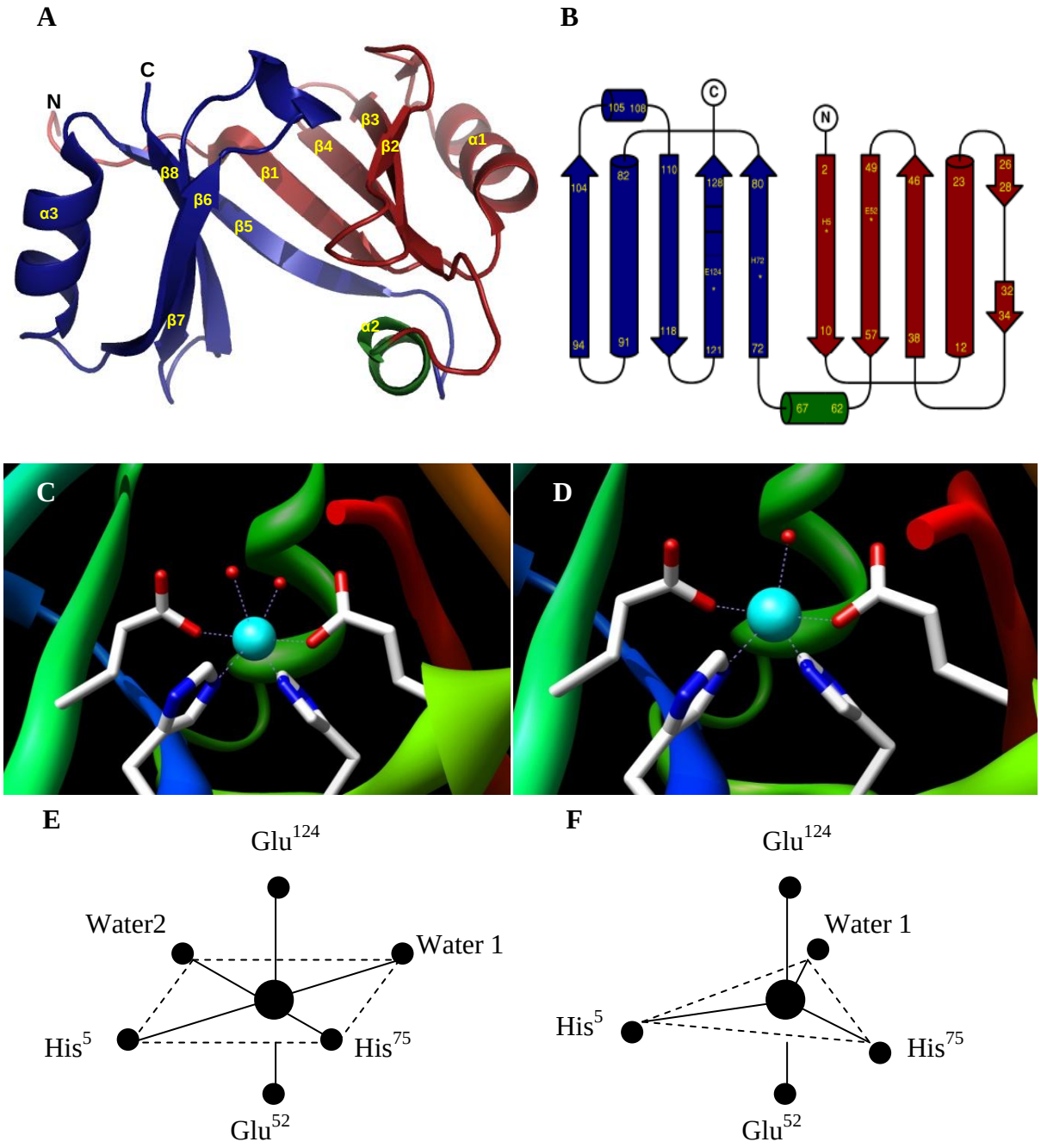


Figure 2.27: (A) Ribbons representation and **(B)** topological representation of CLO GlxI showing the N-terminal βαββ module in firebrick, the connecting α-helix in green and the C-terminal module in blue. **(C)** The representation of the active site of CLO GlxI in an active octahedral geometry with the bound Ni^{2+} (cyan) (PDB: 3HDP) and **(D)** with the Zn^{2+} atom (cyan), forming an inactive trigonal bipyramidal geometry (PDB: 2QH0). Two oxygens of water molecules (red spheres) and the residues His^{5}, Glu^{52}, His^{75}, Glu^{124} (CPK sticks) are involved in metal binding and catalysis. The 3D structures were generated by using the UCSF Chimera program (University of California, San Francisco) [1]. **(E)** The metal coordination of CLO GlxI with bound active metal (such as Ni^{2+} or Co^{2+}) forming an octahedral geometry with four metal binding protein residues and two water molecules around the metal center and **(F)** CLO GlxI with bound inactive metal (such as Zn^{2+}) forming a trigonal bipyramidal geometry with the same four metal binding residues but with only one water molecules.

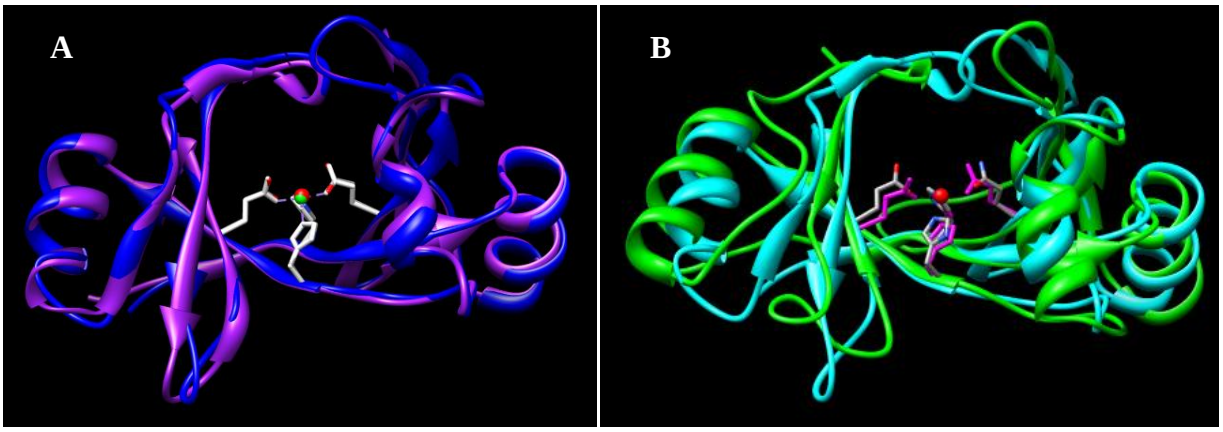


Figure 2.28: (A) The overlapped structures of Ni^{2+}-bound CLO GlxI (purple) and Zn^{2+}-bound CLO GlxI (blue) (PDB: 3HDP and 2QH0, respectively). The metal binding residues are colored in CPK for Zn^{2+}-bound CLO GlxI and in white for Ni^{2+}-bound CLO GlxI. Metal atoms (Zn^{2+} and Ni^{2+}) are overlapped and were colored in red and green, respectively. **(B)** The overlapped structures of CLO GlxI (cyan, PDB: 2QH0) and *P. shermanii* MMCE (green, PDB: 1JC5). The metal binding residues of CLO GlxI are colored in purple, while those of MMCE are in CPK. Zinc atom is colored in red. The 3D structures were generated by using the UCSF Chimera program (University of California, San Francisco) [1].

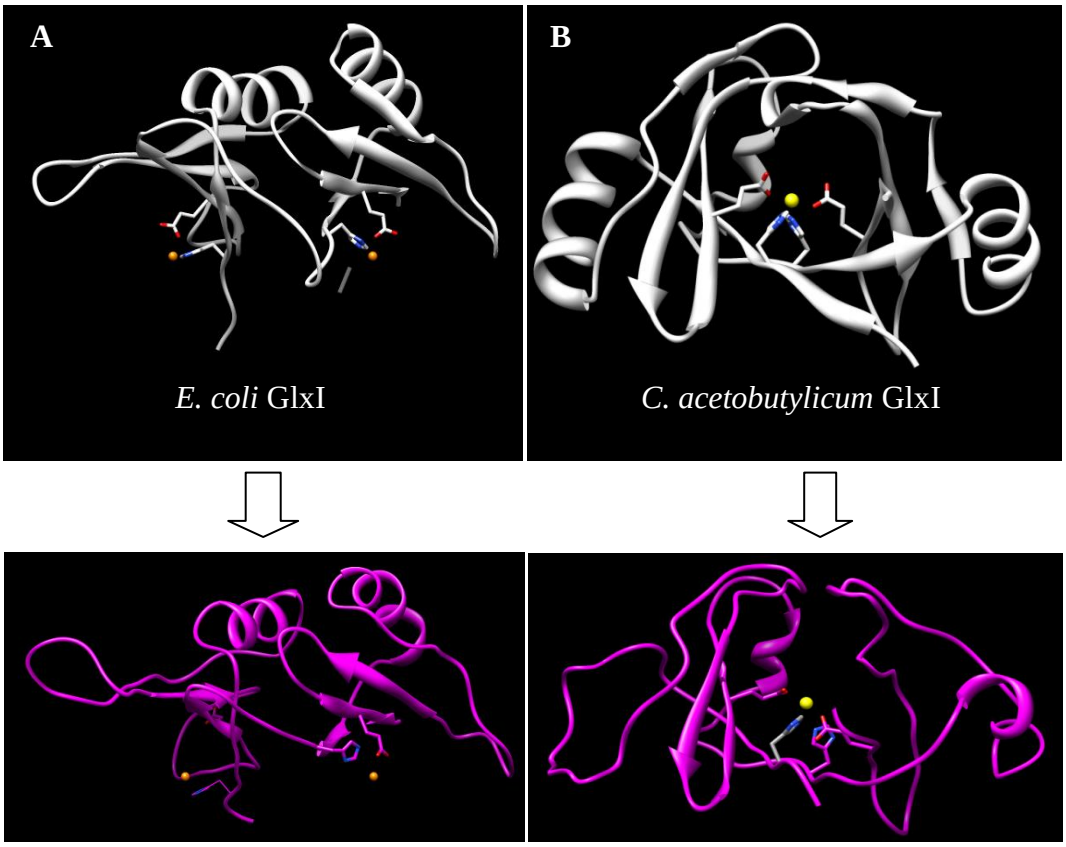


Figure 2.29: The predicted 3D structures of a putative *C. acetobutylicum* GlxI (NP_346890, purple) using **(A)** *E. coli* GlxI (PDB: 1F9Z) and **(B)** CLO GlxI (PDB: 2QH0) as templates using a web based Geno3D structure prediction software (http://geno3d-pbil.ibcp.fr) [142]. The 3D structures were generated by using the UCSF Chimera program (University of California, San Francisco) [1].

Table 2.5: Comparison of *C. acetobutyliucm* GlxI with the homologue structures in the protein data bank through Dali search [133].

	Parameter				
PDB	***Z-score***	***RMSD***	***No. Ca pairs***	***% Sequence ID***	***Ref***
1JC4	16.2	2.0	125	20	[1, 105]
1CJX	10.1	2.5	109	14	[143]
1MPY	10.0	2.9	110	11	[144]
1HAN	8.0	3.1	105	13	[145]
1FRO_N-terMod 6.6	2.5	62	18	63	[44]
1FA5_N-terMod 8.0	2.1	61	20	14	[45]

2.4.6: Enzyme Kinetics

In the kinetic studies, it was found that non His-tagged CLO GlxI possessed more catalytic activity (k_{cat}/K_m) in KPB pH of 7.0 than at pH 6.6 (2.6×10^3 $M^{-1}sec^{-1}$ in KPB pH 7.0 and 2.3×10^3 $M^{-1}sec^{-1}$ in KPB pH of 6.6) (Table 2.6). These data matched the pH profile, as it was previously determined that the optimized pH was found to be at pH 7.0. However, its highest enzymatic activity was found in MOPS buffer pH of 6.6 (3.7×10^3 $M^{-1}sec^{-1}$). In all cases, the activities of yeast GlxI were found to be greater than those of CLO GlxI, while their K_m values for the hemithioacetal substrate (MG-GSH) were similar. This low activity is also observed in most plants (such as maize, *Brassica juncea* and *Aloe vera* [22, 23, 146]) and nematode parasites (such as *Trichostrongylus colubriformis, Ascaris lumbricoides* and *Ascaridia galli* [147]). These results suggest the possibility that more than one GlxI in this Gram-positive bacterium may exist (similar to the case of *P. aeruginosa* with three GlxI as stated previously). Another putative GlxI from *C. acetobutylicum* (NP_346890) was found from the protein data base in the National Center for Biotechnology Information (NCBI). The multiple sequence alignment with other GlxI shows that this enzyme possesses four conserved metal binding residues and it may behave as a Ni^{2+}-activated enzyme (shorter amino acids sequence) (Figure 2.2). As well, putative GlxI from other species of Clostridia also possess shorter amino acid sequences, which are more in line with the Ni^{2+}-activated enzymes (Figure 2.2). Therefore, it is possible that the GlxI enzymes from all Clostridia may be Ni^{2+}-activated enzymes.

It is interesting to note that investigations as to the detoxification of methylglyoxal in other Clostridia such as *Clostridium beijerinckii* and *Clostridium difficile* have identified the glycerol dehydrogenase in these microorganisms as the main contributor to MG detoxification

and Glyoxalase I and II may play a lesser role in this regard [148]. As well, the amount of glutathione produced by γ-glutamylcysteine synthetase and glutathione synthetase (amino acid sequences in *C. acetobutylicum* had been identified by BLAST search [149]) in Clostridia (0.24 μmol/g residual dry weight in *Clostridium perfringens*) is found to be lower than those in Gram-negative bacteria (4.5 μmol/g in *E. coli*), animals (16–25 μmol/g in rat liver) and plants (4.8 μmol/g in spinach leaves) [83]. It is also possible that another thiol in Clostridia might participate in the Glyoxalase reaction and that our data is for glutathione, which is used by the enzyme. Besides, it was found earlier that human GlxI undergoes posttranslational modifications including N-terminal Ala acetylation, the disulfide formation of Cys^{18}-Cys^{19} and the mixed disulfide formation of Cys^{138} and glutathione [150]. The first two factors do not interfere with the enzyme, while the latter strongly inhibits enzymatic activity *in vitro* [150]. The Cys^{138} can also interact with Cys^{60} that is located in another monomeric domain but within the same catalytic pocket, thus forming a disulfide bond that completely closes the catalytic pocket [150]. In the case of low activity in CLO GlxI, the cysteine residue (Cys^{79}) that is located close to the active site of the enzyme might be able to form a disulfide bond with glutathione (Figure 2.30). This is the only Cys residue in the CLO GlxI, suggesting that it is less likely for the enzyme to form the disulfide bond that can quench enzymatic activity. Formation of a disulfide interaction with another subunit is unlikely, since this Cys residue lies in the interior of the enzyme. However, it is possible that this Cys^{79} might interact with metal and disturb the geometry of the activated reconstituted enzyme.

Since the structural comparison, multiple sequence alignment and the Dali search indicated that CLO GlxI possesses a similar structure to that of MMCE (PDB: 1JC5), it was possible that this enzyme possesses dual activity as Glyoxalase I and methylmalonyl-CoA epimerase, or that the protein is actually an MMCE with a low, but measurable alternate Glyoxalase I activity. We therefore evaluated the MMCE activity of the CLO GlxI protein, using the HPLC assay of MMCE enzymes reported by Bobik and Rasche, 2003 [107], where the activity of 162 μmol/min/mg of *Pyrococcus horikoshii* MMCE (0.02 μg) was detected [108]. However, no MMCE activity of the Co^{2+}-bound CLO GlxI (1–10 μg) and Ni^{2+}-bound CLO GlxI (30 μg) was observed, indicating that this enzyme does not possess methylmalonyl-CoA epimerase activity under the studied conditions. Despite its structure, these results

indicate that the enzyme is very specific to the MG-GSH substrate, and that structure alone cannot predict enzymatic activity (detailed discussion in Chapter 5).

Table 2.6: The summary of enzyme kinetics of **(A)** the commercial yeast GlxI (2–9 ng) and **(B)** *C. acetobutylicum* GlxI (795 ng) in KPB and MOPS buffer pH of 6.6 and 7.0. The enzymatic assay was performed using 0.04–1 mM MG-GSH at room temperature.

A

	Yeast GlxI			
Buffer	V_{max} ***(μmol/min/mg)***	K_m ***(mM)***	k_{cat} (s^{-1})	k_{cat}/K_m $(s^{-1}M^{-1})$
KPB (pH 6.6)	1146 ± 37	0.25 ± 0.03	712	2.8×10^6
KPB (pH 7.0)	1306 ± 97	0.31 ± 0.07	811	2.6×10^6
MOPS (pH 6.6)	1492 ± 60	0.22 ± 0.05	927	4.2×10^6
MOPS (pH 7.0)	1093 ± 79	0.19 ± 0.03	679	3.6×10^6

B

	CLO GlxI			
Buffer	V_{max} ***(μmol/min/mg)***	K_m ***(mM)***	k_{cat} (s^{-1})	k_{cat}/K_m $(s^{-1}M^{-1})$
KPB (pH 6.6)	2.42 ± 0.09	0.54 ± 0.02	1.22	2.3×10^3
KPB (pH 7.0)	2.74 ± 0.12	0.53 ± 0.01	1.38	2.6×10^3
MOPS (pH 6.6)	3.34 ± 0.44	0.45 ± 0.18	1.68	3.7×10^3
MOPS (pH 7.0)	2.81 ± 0.58	0.42 ± 0.09	1.42	3.4×10^3

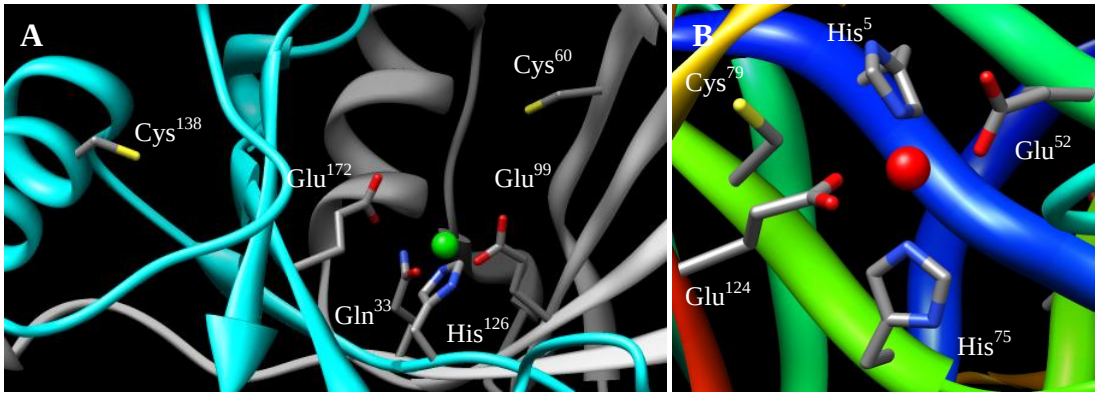


Figure 2.30: The ribbon structure of **(A)** human GlxI (PDB: 1QIN) and **(B)** CLO GlxI (PDB: 2QH0) showing the active sites and the Cys residues that locate within the catalytic pockets. This Cys residue (Cys^{139}) in human GlxI was found to form a mixed disulfide interaction with glutathione, causing a reduction in enzymatic activity [150]. The Cys residue (Cys^{79}) in CLO GlxI is predicted to either form a disulfide bond with glutathione or interact with the active metal, thus disturbing the active site geometry and lowering the activity of the enzyme. The green sphere in human GlxI is the Zn^{2+} atom and the red spheres CLO GlxI represent the Ni^{2+} atom. The 3D structures were generated by using the UCSF Chimera program (University of California, San Francisco) [1].

2.4.7: Suggested Structural Evolution of Glyoxalase I

It has been previously proposed that the variations of domain swapping of the Glyoxalase I family are the result of gene duplication and fusion of a one-module ancestor (Figure 2.31). A two-motif pseudosymmetric metallomonomer from *P. putida* with one active site was believed to have evolved from this single domain precursor by forming a linker between the two domains [30]. This evolutionary pathway was suggested from the discovery of the monomeric GlxI from *P. putida* that forms upon the treatment of the Zn_2 dimeric protein by EDTA (apo-monomer) or glutathione (Zn monomer) [30]. It was suggested that the Zn monomer is an intermediate between the Zn_2 dimer and the apo-monomer and contains non-covalently bound glutathione in the active site, in which its linker is rearranged in order to improve GSH binding [30]. This Zn monomeric enzyme accumulated some enzymatic activity ($k_{cat} \sim 115 \pm 40\ s^{-1}$) but was significantly lower than its fully dimeric two-domain structure ($k_{cat} \sim 500 \pm 100\ s^{-1}$) [30]. Upon glutathione removal, the Zn monomer regains its structure as a dimer.

Due to remarkable structural and noticeable sequence similarity of the two domains, it is possible that the small dimeric GlxI (four domains of two polypeptides) with two active sites was a second generation of the gene duplication-fusion event. The dimeric interaction of small GlxI, so far, is found to be able to form two possible arrangements; two intermolecular domain-domain interfaces (*E. coli* and human GlxI) or two intramolecular domain-domain interfaces (CLO GlxI). The dimeric GlxI from *C. acetobutylicum* is formed with a convenient back-to-back interaction of two monomers with one active site forming within one subunit. The dimeric formation of *E. coli* and human GlxI, if evolving from the monomeric two-domain GlxI (i.e. monomeric *P. putida* GlxI), requires the unfolding process of a monomer to form a new dimeric interaction, creating the active sites at the interface of a dimer. These processes are far more complicated than the formation of dimeric CLO GlxI; therefore, CLO GlxI might be considered an intermediate between the monomeric two-domain and the dimeric two-domain GlxI enzyme structures. In addition, the interface interaction observed in the homodimeric human GlxI is slightly different from that observed in the *E. coli* GlxI. The human GlxI possesses an N-terminal long loop that extends into another subunit (a domain-swapped dimer of a two-domain protein), while the active site of *E. coli* GlxI is formed by the interface interaction of a dimer (a non domain-swapped dimer of a two-domain protein). Large

GlxI (monomeric four-domain from a large single polypeptide) proteins are believed to be the result of gene duplication-fusion events from the small GlxI. Yeast and *P. falciparum* GlxI, for example, possess the N-terminus (two domains) that resembles its C-terminal domain (two subunits), forming two active sites within a four-domain of one polypeptide with two possible arrangements that needed to be identified by some structure determination methods [39, 151].

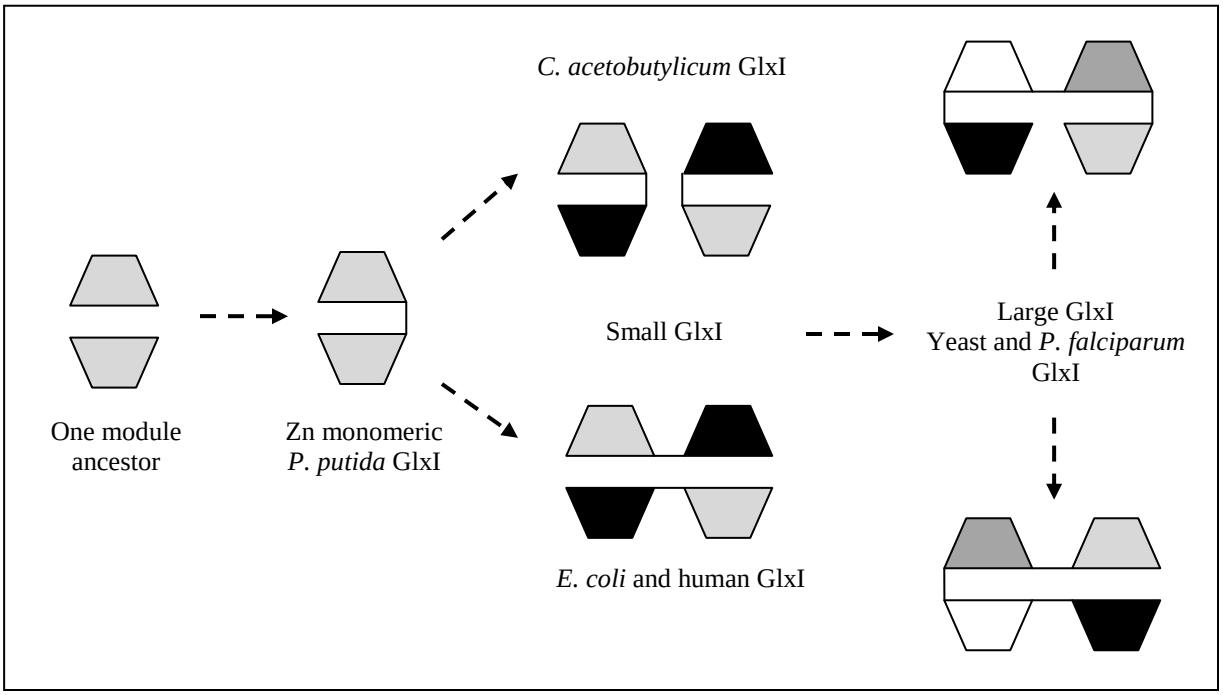


Figure 2.31: The suggested scheme represents the possible structural evolution toward the active site formation within GlxI family based on the previous report by Saint-Jean *et al*., 1998 [30].

2.5: Conclusions

Previous studies in yeast, protozoa, animals, human, plants and Gram-negative bacteria divide GlxI into two classes according to metal activation, the Zn^{2+}- activated class and the non Zn^{2+}-activated class, which exhibits selective Ni^{2+}/Co^{2+}-activation. However, there is a lack of knowledge concerning the Glyoxalase system from Gram-positive bacteria. The X-ray crystallographic structure determined as a result of high throughput proteomics investigations indicates that *Clostridium acetobutylicum* GlxI possesses a Zn^{2+} atom in its active site, forming trigonal bipyramidal coordination with four metal binding residues and one water molecule

(PDB: 2QH0). From previous knowledge on the active site geometry of GlxI, this Zn^{2+}-bound CLO GlxI was predicted by our group to be in its inactive form.

The clone of this CLO GlxI was supplied to our research group by the NYSGXRC consortium and the overproduced protein was further purified by HisTrap affinity chromatography and gel permeation chromatography. Metal characterization indicated that CLO GlxI is a Ni^{2+}/Co^{2+}-activated enzyme with optimized activity in buffer pH of 7. Analysis of data from various metal titration experiments also suggested that only one active site of the dimeric enzyme is functional. As well, the stability of the enzyme under various conditions was investigated. It was found that the enzyme is stable at protein concentrations less than 7.8 μM in Tris buffers (pH 6.0) containing 200 mM KCl and 10% (v/v) glycerol. The T_m of the enzyme in MOPS buffer (pH 6.0) containing 150 mM KCl and 10% glycerol is approx. 61.3 °C. Under these studied conditions, non His-tagged CLO GlxI was cloned and purified by utilizing HisTrap and Benzamidine affinity chromatography. The apo-enyzme was reconstituted with $NiCl_2$ and was sent to NYSGXRC for X-ray crystallography in order to compare with its Zn^{2+}-bound form. The active site of Ni^{2+}-bound CLO GlxI was found to form an octahedral coordination with two water molecules and four metal binding ligands. CLO GlxI exhibits a unique dimeric quaternary structure (two intramolecular domain-domain interfaces) among other small subunit GlxI. Hence, the GlxI from *C. acetobutylicum* is the first enzyme from Gram-positive bacteria that has been fully characterized as to its three-dimensional structure and its detailed metal activation profile.

CHAPTER 3: DELETIONAL MUTAGENESIS OF ZINC-ACTIVATED GLYOXALASE I FROM *PSEUDOMONAS AERUGINOSA*

3.1: Introduction

It has previously been observed that the metalloenzyme Glyoxalase I (GlxI) can be divided into two classes according to its metal activation profile, a Zn^{2+}-activated enzyme class and a non Zn^{2+}-activated class, which is instead activated by Ni^{2+}/Co^{2+} ions [17-35]. The metal activation characteristics can be predicted by active site geometry around the metal center as well as by analysis of the length of its amino acid sequence [45]. In both classes, the metal coordination of the catalytically active enzyme (e.g. *E. coli* GlxI with Ni^{2+}/Co^{2+} atoms or human GlxI with Zn^{2+} atoms bound) has octahedral geometry with four metal binding residues and two water molecules around the active metal [44-46]. On the other hand, its inactive form (*E. coli* with bound Zn^{2+} for example) employs a five-coordinated trigonal bipyramidal arrangement with only one water molecule bond to the metal center [45]. In addition, the Zn^{2+}-activated GlxI enzymes tend to possess a longer amino acid sequence compared to those of the Ni^{2+}/Co^{2+}-activated GlxI enzyme class. The analysis of multiple sequence alignments as well as the X-ray crystallographic structures indicate that the Zn^{2+}-activated GlxI possesses a long N-terminal arm (region A in Figure 3.1 and 3.2) that wraps around another subunit, thus appearing to hold the dimeric structure together [44, 46]. This N-terminal arm also seems to be significant for protein stability for the Zn^{2+}-activated enzyme, since its deletion can cause protein aggregation [44, 46]. However, even though the Ni^{2+}/Co^{2+}-dependent enzyme does not possess this extra N-terminal helix, its dimerization and protein stability are still maintained. The Zn^{2+}-activated GlxI also contains at least three other additional short regions (regions B, C and D in Figure 3.1 and 3.2A) that are absent in the Ni^{2+}/Co^{2+}-activated form. Region B is a three-turned α-helix that blocks one side of the catalytic pocket of the Zn^{2+}-activated enzyme, hence creating a more packed active site compared to that of the Ni^{2+}/Co^{2+}-activated enzyme, and it possibly also protects the bound substrate from outer solvent (Figure 3.2B). We hypothesize that this region might play a significant role in determining the metal preference of each GlxI family member, since it locates next to the active site. Regions C and D are the small loops that locate at the outer layer of the catalytic pocket; hence, they might be expected to play a minor role in metal specificity and enzymatic activity. Therefore, an investigation on

the influence of these extra amino acid sequences on metal specificity and enzymatic activity could lead to valuable knowledge on the relatedness of structure and function of GlxI that could be significant in many aspects as well as for future drug discovery efforts to target this metalloenzyme in different organisms.

```
                                A
H.sapiens       MAEPQPPSGGLTDEAALSCCSDADPSTKDFLLQQTMLRVKDPKKSLDFYTRVLGMTLIQKCDFPIMKFSL 70
P.aer GloA1     -----------------------------MRILHTMIRVGNIDRSIDFYTRVLGMTLLRKNDYPDGQFTL 41
P.aer GloA2     -----------------------------MRILHSMLRVADLEAALEFYTRALDMRLLRRRDYPEGRFTL 41
P.aer GloA3     MSFNT--------EVQPGICMEPDAITQEYVFNHTMLRVKDPKRSLDFYSRVLGMRLLRRLDFEEGRFSL 62
P.putida        MSLND-------LNTLPGVTAQADPATAQFVFNHTMLRVKDIEKSLDFYTRVLGFKLVDKRDFVEAKFSL 63
                                             *
                                B                C          D
H.sapiens       YFLAYEDKNDIPKEKDEKIAWALSRKATLELTHNWGTEDDETQSYHNGNSDPRGFGHIGIAVPDBYSACK 140
P.aer GloA1     AFVGYGNEAD----------------SAVIELTHNWG-----VDAYEIGT----GYGHIAIEVDDAYQACD 87
P.aer GloA2     AFVGYQDERA----------------AAALELTHNWD-----RDGYTQGD----GYGHLAIEVEDAAVTCA 87
P.aer GloA3     YFLAMTRGEEVPDAVDERQRYTFGRQSVLELTHNWGSESDDSQ-YHNGNQDPRGFGHICFSVPDLVAACE 131
P.putida        YFLALVDPATIPADDDARHQWMKSIPGVLELTHNHGTERDADFAYHHGNTDPRGFGHICVSVPDVVAACE 133
                                          *                      *

H.sapiens       RFEELGVKFVKK--PDDGKMKGLAFIQDPDGYWIEILNPNKMATLM----- 184
P.aer GloA1     DIRYNGGQVTREAGPMKHGTTVIAFVTDPDGYKIELIQKSS---------- 128
P.aer GloA2     RARALGYRVTREAGLMQHGRSVIAFLEDPDGYKVELIQKGTQFD--------131
P.aer GloA3     RFETLGVNFVK---PLDRGMKNVAFISDPDGYWVEIVQASLNGEMGRG----176
P.putida        RFEALQVPFQKR--LSDGRMNHLAFIKDPDGYWVEVIQPTPL--------- 173
                                        *
```

Figure 3.1: Multiple sequence alignment of multi-expressed GlxI from *Pseudomonas aeruginosa* including GloA1 (AAG06912), GloA2 (AAG04099) and GloA3 (AAG08496) with GlxI from *H. sapiens* (AAB49495) and *P. putida* (AAN69360) (organism name follows by National Center for Biotechnology Information (NCBI) accession number). The metal binding residues are marked with an asterisk. The loop regions in the Zn^{2+}-activated GlxI are highlighted in grey and marked with bold capital letters (A, B, C and D). The alignment was created using CLC Free Workbench (version 3.0.1) with the accurate alignment algorithm (http://www.clcbio.com).

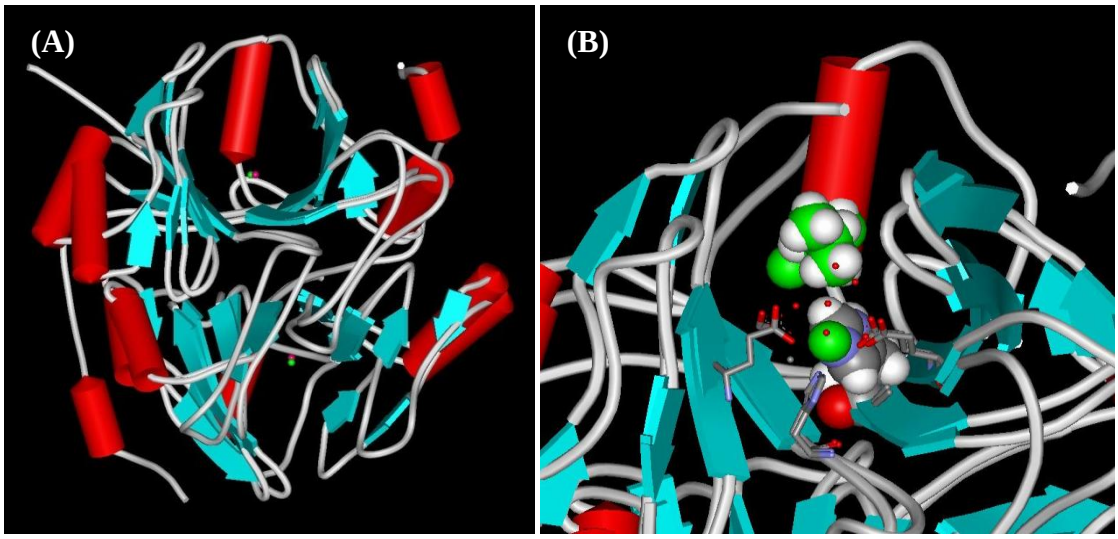


Figure 3.2: (A) The superimposed structures of the dimeric *E. coli* GlxI (PDB: 1F9Z) and human GlxI (PDB: 1QIN) showing the extra amino acid sequences (non-overlapped sections) that only exist in the Zn^{2+}-activated

enzyme (human GlxI). **(B)** The close-up view shows the close interactions of amino acid in the helix B and one of the metal binding ligands in human GlxI (PDB: 1QIN).

Analysis of the multiple sequence alignment of GlxI from both metal activation classes indicated that one of the metal binding residues in human GlxI (Gln^{34}) is not conserved and is replaced by His in GlxI from other organisms (Figure 3.1). It was possible that this amino acid might play a crucial role on metal selectivity, although other Zn^{2+}-activated GlxI, such as the *Saccharomyces cerevisiae* and the *Pseudomonas putida* enzymes, also did not have a glutamine as a metal binding residue, but rather utilize histidine as the *E. coli* enzyme does in that position. Mutagenesis of His^5 to Gln^5 (H5Q) in *E. coli* GlxI (relevant to Gln^{34} in human enzyme) resulted in a mutein that was activated in the presence of Zn^{2+}, Ni^{2+}, Co^{2+}, Mn^{2+} and Cd^{2+} (the activity of Co^{2+}> Ni^{2+}> Zn^{2+}> Mn^{2+}> Cd^{2+}), while the wild-type enzyme was inactive with Zn^{2+} [37]. However, not only did this mutation reduce metal specificity, but it also greatly decreased enzymatic activity [37]. Besides, as mentioned above, the metal binding residues of the Zn^{2+}-activated GlxI from *Pseudomonas putida* and *Saccharomyces cerevisiae* are similar to those of *E. coli* enzyme, a Ni^{2+}/Co^{2+}-activated enzyme, suggesting that this mutation alone is less important in determining the metal specificity of the enzyme than in the magnitude of the enzymatic reaction rate.

Additionally, Dr. Nicole Sukdeo, a previous Ph. D. student in our laboratory, attempted to utilize insertional mutagenesis in order to introduce the loop regions of B, C and D amino acids from human GlxI into the *E. coli* enzyme (unpublished data). The analyses of enzymatic activity indicated that the H5Q mutant with the insertion of helix B possesses a low but detectable activity in the presence of Zn^{2+}. However, the activity of this Ni^{2+}-reconstituted mutein remained at ~600 μmol/min/mg (specific activity of wild-type *E. coli* GlxI was 676 μmol/min/mg [19]), suggesting that the enzyme possesses lower metal specificity as well as stability and is functional even with a large insertion of 15 amino acids. Similar results were observed for the muteins with the combined insertions of BC and BD loops. However, their enzymatic activity in the presence of Zn^{2+} ion was drastically reduced in the muteins with a single insertion of these loops. From these preliminary data, it seemed worthwhile to make another attempt to determine how the amino acid sequence relates to metal activation of GlxI and how these extra loops affect the metal selection of the enzyme.

The Gram-negative bacterium, *Pseudomonas aeruginosa*, was chosen as the source of the GlxI to be investigated for exploring the effect of various amino acid regions on the metal activation profile of this enzyme. This organism contains three *glxI* genes, *gloA1*, *gloA2* and *gloA3*, that code for GloA1, GloA2 and GloA3, respectively [33]. GloA1 and GloA2 are Ni^{2+}/Co^{2+}-activated enzymes, while GloA3 is Zn^{2+}-activated [33]. Since GloA2 possesses higher enzymatic activity than GloA1, it was chosen as the representative for the Ni^{2+}-activation class, while GloA3 was chosen as the representative of the Zn^{2+}-activation class. By selecting GlxI from the same organism to study the structural differences between metal activation classes, it is more likely that the resulting chimeric enzymes might have a better chance of exhibiting clear structure-function activity differences. In an attempt to mimic the Ni^{2+}/Co^{2+}-activated GloA2 and its metal preference, deletional mutagenesis of the loop regions B, C and D in the Zn^{2+}-activated GloA3 was undertaken, utilizing the polymerase chain reaction and various restriction enzymes. The desired mutant genes were induced and expressed in *E. coli* cells. The resulting muteins were purified and characterized in terms of protein stability, metal activation and enzyme kinetics. Their structural investigations were also analyzed in comparison to the wild-type GloA3 and the GloA2 enzymes. The contribution that each of these protein regions makes to the metal specificity of the enzyme was analyzed and discussed in this chapter.

3.2: Reagents, Materials and Instrumentation

All reagents, materials and equipment used in this chapter are listed below otherwise they are included in the previous chapters. The designed primers for DNA cloning were received from Sigma-Genosys Canada (Oakville, ON). The restriction enzymes including *Nde*I and *Eco*RI were obtained from New England Biolabs (Beverly, MA). The antibiotic ampicillin (Amp) was obtained from Sigma Chemical Company (St. Louis, MO). Ammonium sulfate ($(NH_4)_2SO_4$) was received from BHD (Toronto, ON). Protein standards for gel permeation chromatography including bovine serum albumin (BSA), ovalbumin, carbonic anhydrase, myoglobulin and vitamin B12 (molecular weights of 66, 44, 29, 17 and 1.35 kDa, respectively) as well as a protein purification column, Bio-Scale® Mini UNOsphere™ Q Cartridges (1 mL), were obtained from Bio-Rad laboratories (Hercules, CA). A Superdex™

75 HR 10/30 column, a Q-Sepharose Fast Flow HR 10/30 column, a Mono-Q HR 5/5 column and a HiTrap Phenyl HP column (5 mL) were obtained from GE Healthcare (Piscataway, NJ).

3.3: Experimental Protocols

3.3.1: DNA Cloning and Manipulation

All DNA manipulations and purifications were performed according to the protocols outlined by Sambrook and Russell [135]. The DNA cloning of *gloA2* and *gloA3* into the pET-22b(+) expression vector was performed by Dr. Nicole Sukdeo as previously reported [33]. A single deletional mutation of GloA3 was performed by Kadia Mullings, a previous M.Sc. student in our laboratory, utilizing QuikChange PCR protocols and the following primers.

bgloA3 (+) 5′-CGCCATGACCCGTGGCGAAGAACAGTCGGTCCTCGAGCTGAACC-3′

(−) 5′-GGGTCAGCTCGAGGACCGACTGTTCTTCGCCACGGGTCATGGCG-3′

cgloA3 (+) 5′-CCCACAACTGGGGCCAGTACCACAACGGC-3′

(−) 5′-GCCGTTGTGGTACTGGCCCCAGTTGTGGG-3′

dgloA3 (+) 5′-GCCAGTACCACAACGGCAACGGGTTCGGCCACATCTGC-3′

(−) 5′-GCAGATGTGGCCGAACCCGTTGCCGTTGTGGTACTGGC-3′

Double and triple deletional mutations were generated by a two-stage PCR method. For example, the introduction of the gene coding for delbcGloA3 (deletion of regions B and C from GloA3, Figure 3.1) into the pET-22b(+) expression vector was first performed by two polymerase chain reactions utilizing the pET-22b-*delbgloA3* plasmid as template. One PCR employed T7.For and cgloA3.Rev as primers to generate the gene with a *Nde*I restriction site and a deleted region C on the *delbgloA3* gene. Another PCR reaction used cgloA3.For and T7.Rev as primers to generate a *EcoR*I restriction site on *delcgloA3*. The recipe of reagents and thermal cycles for these first two polymerase chain reactions was prepared and set as previously described (Table 2.2 and 2.3 in Chapter 2). Both PCR products were isolated and purified using a QIAquick Gel Extraction Kit. The third PCR linked these two purified PCR products using T7.For and T7.Rev as primers to generate the *delbcgloA3* gene with *Nde*I and *EcoR*I restriction sites. The recipe for the reagents of the third PCR are shown in Table 3.1 and the thermal cycles were set as in Table 2.3 in Chapter 2. T7 promoter and terminator primers

were purchased as T7.For; 5′-TAATACGACTCACTATAGG-3′ and T7.Rev; 5′-GCTAGTTATTGCTCAGCGG-3′.

Table 3.1: PCR recipe for GloA3 mutated encoding gene used for DNA amplification

Plasmids/Reagents	***Stock Concentration***	***Amount Used (μL)***	***Final Concentration***
Product from 1st PCR	50 ng/μL	1	50 ng
Product from 2nd PCR	50 ng/μL	1	50 ng
T7.For	20 μM	1	0.4 μM
T7.Rev	20 μM	1	0.4 μM
dNTPs	10 mM	1	0.2 mM
pfu Buffer	10×	5	
ddH_2O		39	
pfu polymerase		1	
Total		50.0	

The third PCR plasmid as well as the pET-22b(+) expression vector was digested with *Nde*I and *EcoR*I restriction enzymes. The desired produced plasmid sections were isolated and purified by utilizing a QIAprep Spin Miniprep Kit. The gene *delbcgloA3* was cloned into a pET-22b(+) expression vector by overnight ligation with T4 DNA ligase at 16 °C. The circular DNA was heat shock transformed into competent *E. coli* DH5α followed by overnight growth on LB agar plates containing ampicillin (100 μg/mL) at 37 ºC. A single colony was picked and inoculated overnight in LB (5 mL) containing Amp (100 μg/mL) at 37 ºC in a 220 rpm shaker. The plasmid was purified using a QIAprep Spin Miniprep Kit and sent for sequencing when appropriate. The summary for DNA cloning is shown as a flow chart in Appendix 1 (Figure A1.2).

Other double and triple deletional GloA3 variants were prepared similarly. The pET-22b-*deldgloA3* plasmid was used as a template with T7 and bgloA3 primers to generate the *delbdgloA3* plasmid. The plasmid containing *delcdgloA3* was generated by utilizing the pET-22b-*deldgloA3* template with T7 and cgloA3 primers. The triple deletional GloA3 was generated by combining the first PCR product of pET-22b-*delbgloA3* with T7.For and cgloA3.Rev and the second PCR product of pET-22b-*deldgloA3* with cgloA3.For and T7.Rev using the T7 primers. The desired plasmid was cloned back into the pET-22b(+) expression vector and heat shock transformed into competent *E. coli* DH5α. The plasmid was isolated and purified as previously described.

3.3.2: Protein Expression, Induction and Purification

E. coli BL21 (λDE3) containing the GloA3 mutant gene was inoculated in TB (10 mL) containing Amp (100 μg/mL) overnight at 37 °C. The culture was then transferred into 1 L TB with or without an additional 1 mM $NiCl_2$ or $ZnCl_2$ at 37 ºC before inducing by the addition of 0.5 mM IPTG for 4 hours at 25 °C. The cell pellets were harvested by centrifugation (6,000×*g* for 10 min) and stored at -80 ºC.

The protein purification of GloA2 and GloA3 was performed as previously reported by Sukdeo and Honek, 2007 [33]. The cell pellet was resuspended in 20 mM Tris (pH 7.5) containing 10 mM DTT, 10% (v/v) glycerol and 1 mM PMSF before disrupting the cell wall by sonication. The cell lysate was then centrifuged at 48,384×*g* for 20 min. The supernatant was loaded onto a 10/30 Q-Sepaharose Fast Flow column. The proteins were collected using a gradient of 0–1 M KCl in 20 mM Tris buffer (pH 7.5) containing 10 mM DTT and 10% (v/v) glycerol with 1 mL/min flow rate over a 100 min interval. The fractions containing the desired protein were monitored by SDS-PAGE and dialyzed against 20 mM Tris (pH 7.5) containing 10 mM DTT and 10% (v/v) glycerol overnight at 4 °C.

The dialyzed sample was loaded onto a Mono-Q HR 5/5 column. The desired proteins were eluted using the same buffers as the previous step. The fractions containing desired protein were monitored by SDS-PAGE and were then dialyzed against 50 mM KPB (pH 6.6) containing 10 mM DTT and 10% glycerol overnight at 4 °C.

To the dialyzed sample, solid $(NH_4)_2SO_4$ was slowly added to a final concentration of 1.7 M before filtering through a 0.2 μm filter. The sample was then loaded onto a Phenyl HP column (5 mL) and eluted using a gradient of 1–0 M $(NH_4)_2SO_4$ in 50 mM KPB (pH 6.6) containing 10 mM DTT and 10% (v/v) glycerol over a 100 min period (1 mL/min flow rate). The fractions containing desired protein were monitored by SDS-PAGE and were then dialyzed against buffer containing 20 mM Tris (pH 7.5) and 150 mM NaCl overnight at 4 °C.

The protein was further purified by gel permeation chromatography, which was performed using a buffer containing 20 mM Tris (pH 7.5) and 150 mM NaCl with 0.5 mL/min flow rate on a Superdex75 HR 10/30 column. The purified enzyme was identified by SDS-PAGE and ESI-MS based analyses. The protein concentration was calculated by the Bradford assay using BSA as a standard. The flow chart for the protein purification is summarized in Figure A2.3, Appendix 2.

Changes were made to the purification steps for samples with low protein stability. A Q-Sepharose Fast Flow HR 10/30 column was replaced by a UNOsphere Q cartridge (1 mL) for the double and triple deletional mutants. The purification utilizing a Mono-Q HR 5/5 column was skipped for deldGloA3, delcdGloA3 and delbdGloA3. The double deletional delbcGloA3 was purified with only anion exchange chromatography using UNOsphere Q cartridges and gel permeation chromatography using a Superdex75 HR 10/30 column. The triple deletional delbcdGloA3 was purified by the addition of $(NH_4)_2SO_4$ to a concentration of 1.275 M, followed by isolation using a HiTrap Phenyl HP column and a Superdex75 HR 10/30 column.

3.3.3: Enzymatic Assay

The enzyme assay was performed as described previously in Chapter 2 using the hemithioacetal, the non-enzymatic product from the reaction of methylglyoxal (MG) and glutathione (GSH), as substrate. The metal-reconstituted enzymes were prepared by incubating 5 equivalents of metals ($NiCl_2$ or $ZnCl_2$) to the purified enzymes overnight at 4 °C. The preliminary kinetics was measured using these enzymes with 0.02–1 mM substrate in 50 mM KPB (pH 6.6). From these data, wild-type GloA3 and its selected variants with relatively high ratio of activity with Ni^{2+} compared to Zn^{2+} were chosen and prepared in the apo-forms (see Chapter 2 for preparation). These apo-enzymes were then titrated with $NiCl_2$ and $ZnCl_2$ and the enzymatic assay was repeated in order to achieve optimal activity.

3.3.4: Structural Investigation

The secondary structures of GloA3 and its mutants were examined by circular dichroism (CD) analysis. The experiments were performed using a Jasco J-715 spectropolarimeter with the set parameters as previously described in Chapter 2. The purified samples were incubated with an addition of 5 equivalents of metals ($NiCl_2$ or $ZnCl_2$) overnight at 4 °C in 50 mM Tris (pH 7.5) and 150 mM NaCl to the final protein concentration of 0.3 mg/mL prior to the measurement.

The quaternary structural investigations were performed by gel permeation chromatography utilizing Tris buffer (20 mM Tris (pH 7.5) and 150 mM NaCl) with a flow rate of 0.5 mL/min on a Superdex75 HR 10/30 column. The molecular mass of the protein

sample was calculated using Bio-Rad protein standards including BSA (66 kDa), ovalbumin (44 kDa), carbonic anhydrase (29 kDa), myoglobulin (17 kDa) and vitamin B12 (1.35 kDa) (example of calculation in Appendix 3).

The predicted 3D structure of GloA3 based on the structure of human GlxI (PDB: 1QIN) was investigated using a web based software, Geno3D Automatic comparative molecular modeling of protein (http://geno3d-pbil.ibcp.fr) [142]. Structural visualizations and molecular images were generated by the UCSF Chimera package from the Resource for Biocomputing, Visualization and Informatics at the University of California, San Francisco (supported by NIH P41 RR-01081) [1]. The unpublished X-ray crystallographic structures of *E. coli* GlxI with bound hydroxamate inhibitor (TSI) with the resolution of 2.1 Å and with bound *N*-hydroxybenzamide (BHO) with the resolution of 1.6 Å were prepared by Dr. Susan Clugston (a previous Ph. D. student in our laboratory) as well as Dr. Molly He and Dr. Brian Matthews from University of Oregon (Eugene, Oregon). The pdb files for both inhibitors bound *E. coli* GlxI are shown in Appendix 8.

Additionally, Ni^{2+}-bound native GloA2 (expressed in the *E. coli* system in the presence of 1 mM $NiCl_2$) and Zn^{2+}-bound native GloA3 (expressed in the *E. coli* system in the presence of 2 mM $ZnCl_2$) that were purified as previously described were sent to Dr. Charles Bond from University of Western Australia (WA, Australia) for X-ray crystallographic analysis.

3.4: Results and Discussion

3.4.1: Protein Characterization

The gene of interest was cloned into the pET-22b(+) expression vector and was expressed in an *E. coli* expression system under standard conditions. The proteins were purified using five purification-steps including anion exchange chromatography (Q-Sepharose Fast Flow HR 10/30 column (or UNOsphere Q 1 mL cartridge) and Mono-Q HR 5/5 columns), ammonium sulfate precipitation, Phenyl Sepharose purification (5 mL HiTrap Phenyl HP column) and gel permeation chromatography (Superdex75 HR 10/30 column). The first two columns for anion exchange chromatography roughly eliminated most of the *E. coli* proteins through the flow-through fraction in basic buffer containing no salt (Figure 3.3 and 3.4). After these two columns, the desired proteins were partially purified but also showed some

contamination with larger *E. coli* proteins (high molecular weight), while smaller proteins were washed away through dialysis (dialysis tubing having a molecular weight cut-off of 12–14 kDa). Next, ammonium sulfate precipitation and Phenyl Sepharose purification were found to eliminate large *E. coli* proteins that precipitated with the high concentration of salt (Figure 3.5). The desired enzyme was achieved following gel permeation chromatography (Figure 3.6) as monitored by SDS-PAGE (Figure 3.7A). The identity of the protein was confirmed by ESI-MS (Figure 3.7B) experiments.

Since the purification using the Mono-Q column (after the Q-Sepharose column) only resulted in small improvements in purification, this step was skipped for most proteins. All these three anion exchange columns (Q-Sepharose, UNOsphere Q and Mono-Q columns) possess strong ion exchange groups such as quaternary amines ($-N^+(CH_3)_3$). The major difference between a Q-Sepharose column and a Mono-Q column is based on their matrix beds. The Mono-Q column has smaller and better matrices to separate more highly purified proteins. Its matrix is composed of polystyrene/divinyl benzene that can form a particle size of 10 μm. The matrix of a Q-Sepharose resin, on the other hand, consists of 6% agarose, creating a spherical bed form with a size of 45–165 μm. As well, the size (bed volume) of a Mono-Q column that was utilized was only 1 mL, thus having the advantage of being useful for performing small scale purifications (with economical uses of reagents and reduced purification times). However, the Mono-Q column can only bind a maximum of 50 mg protein, while a UNOsphere Q column can bind approximately 180 mg of protein, even though both columns were the same size. Therefore, the UNOsphere Q column was later replaced by the Q-Sepharose and Mono-Q columns, thus limiting protein loss from the multiple purification steps.

Most proteins were purified as listed by the protocols stated above. However, the double deletional delbcGloA3 protein was purified by only a two-step purification, the anion exchange chromatography using UNOsphere Q cartridges and gel permeation chromatography using a Superdex75 column. The proteins isolated by anion exchange chromatography were quite pure as determined by SDS-PAGE analysis. These proteins, however, were precipitated in the presence of high salt concentrations (1.7 M $(NH_4)_2SO_4$), so this purification step as well as the Phenyl Sepharose purification step were not used. The purified protein after gel permeation chromatography was achieved since a single band of protein was visualized by

SDS-PAGE analyses utilizing Coomassie brilliant blue (data not shown). The triple deletional delbcdGloA3 protein was the most difficult variant to express and isolate. The initial purification by anion exchange chromatography was unsuccessful, since this mutant was eluted from the column in the flow-through fraction (in buffer containing no salt). The initial purification, thus, then changed to $(NH_4)_2SO_4$ precipitation followed by Phenyl Sepharose purification. However, this protein was unstable in buffer containing a high concentration of salt (1.7 M $(NH_4)_2SO_4$), so only 1.275 M $(NH_4)_2SO_4$ was applied. Analysis of the Phenyl Sepharose purification indicated that the desired protein was eluted from the column at a low concentration of salt. After the purification using gel permeation chromatography, the purified delbcdGloA3 protein was obtained (data not shown).

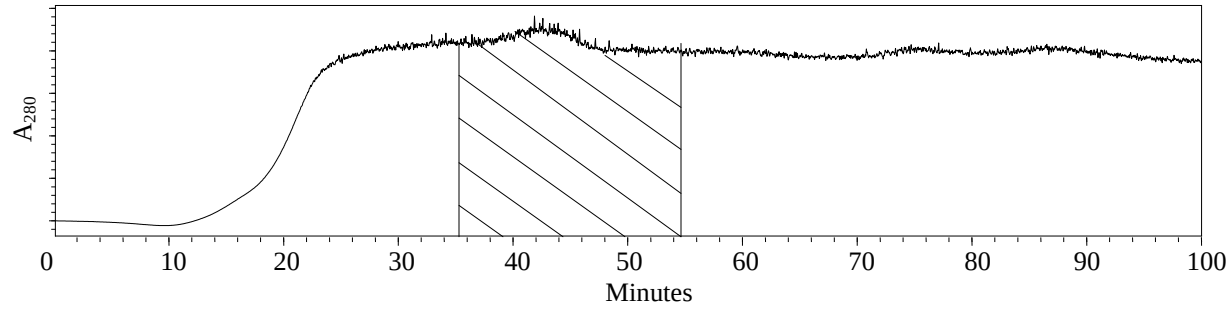


Figure 3.3: The HPLC chromatogram of the delbGloA3 purification by anion exchange chromatography using a 10/30 Q-Sepaharose Fast Flow column and a gradient buffer of 0–1 M KCl in 20 mM Tris buffer (pH 7.5) containing 10 mM DTT and 10% (v/v) glycerol over a 100 min interval with a flow rate of 1 mL/min. Fractions containing delbGloA3 protein were monitored by SDS–PAGE, which corresponded to the fractions eluted between 35–55% KCl (shade area).

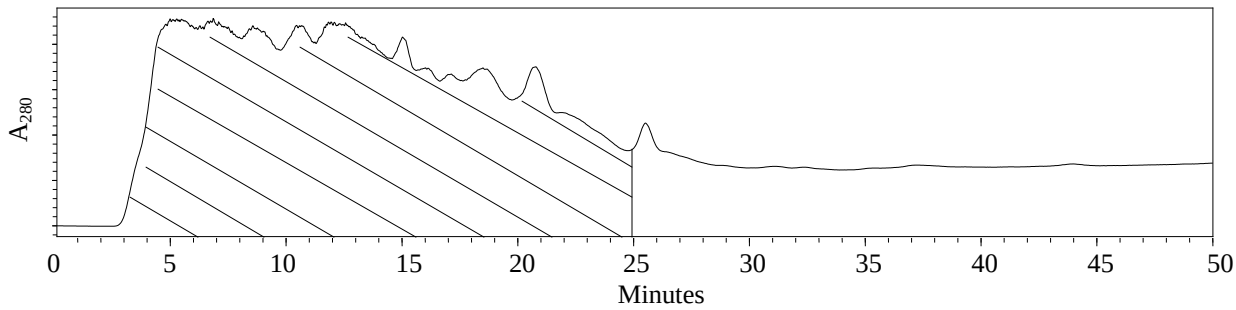


Figure 3.4: The HPLC chromatogram of delbGloA3 purification by anion exchange chromatography using a Mono-Q HR 5/5 column and a gradient buffer of 0–1 M KCl in 20 mM Tris buffer (pH 7.5) containing 10 mM DTT and 10% (v/v) glycerol over a 100 min interval with a flow rate of 1 mL/min. Fractions containing delbGloA3 protein were monitored by SDS–PAGE, which corresponding to the fractions eluted between 4–25% KCl (shade area).

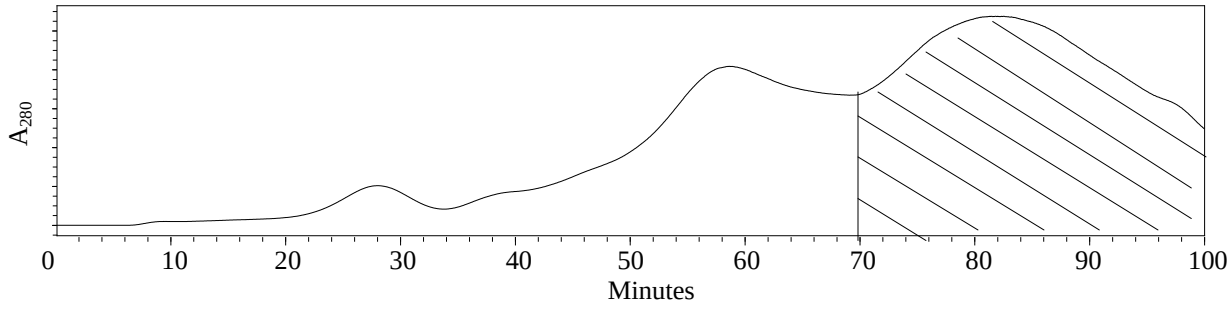


Figure 3.5: The HPLC chromatogram of delbGloA3 purification by Phenyl Sepharose chromatography using the HiTrap Phenyl HP column (5 mL) and a gradient of 1–0 M $(NH_4)_2SO_4$ in 50 mM KPB (pH 6.6) over a 100 min interval with a flow rate of 1 mL/min. Fractions containing delbGloA3 protein were monitored by SDS–PAGE, which corresponding to the fractions eluted with 0–30% $(NH_4)_2SO_4$ (shade area).

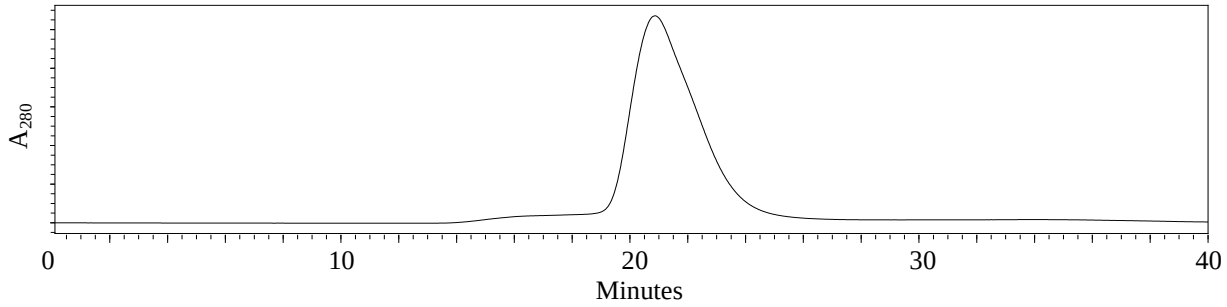


Figure 3.6: The HPLC chromatogram of delbGloA3 purification by gel permeation chromatography using a Superdex75 HR 10/30 column and a buffer containing 20 mM Tris (pH 7.5) and 150 mM KCl with a flow rate of 0.5 mL/min. The peak containing purified delbGloA3 protein was monitored by SDS-PAGE.

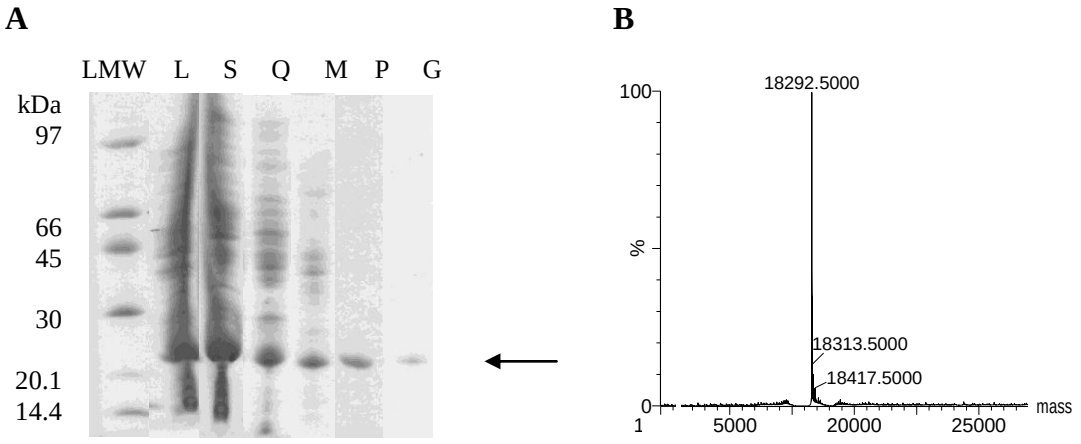


Figure 3.7: (A) The SDS-PAGE of the delbGloA3 purification: LMW = low molecular weight marker, L = lysate, S = supernatant after sonication, Q = Q-Sepharose eluted fraction, M = Mono-Q eluted fraction, P = HiTrap Phenyl HP eluted fraction and G = gel permeation fraction. The arrow indicates the presence of delbGloA3. **(B)** The electrospray mass spectrum of delbGloA3 showed the molecular weight of a single subunit at 18292.5 Da (calculated MW is 18292.5 Da).

The protein characterization of the GloA2 and GloA3 proteins was similar to the previous report by Sukdeo and Honek, 2007 [33]. The recovery yield, molecular weight and predicted p*I* of the proteins are shown in Table 3.2 and Figure 3.8. Post-translational modification was observed as a cleavage of the first N-terminal Met residue in all GloA3 samples. All GloA3 mutants have a similar predicted p*I* that varies from 4.9–5.4 among the proteins. The recovery yield after purification was in the following order, GloA3 > delcGloA3 > deldGloA3 > delbGloA3 > delbdGloA3 > delcdGloA3 > delbcGloA3 > delbcdGloA3. These data might imply that the recovery yield was related to protein stability. Wild-type GloA3, the most stable form among all samples, had the highest recovery yield (232 mg purified protein per 1 L culture or 29 mg per g cell pellet). The delbGloA3 mutein possessed the lowest recovery yield among the single deletional mutants (including delbGloA3, delcGloA3 and deldGloA3), which is possibly due to its largest deletional segment (15 amino acids). All purified double deletional mutations had similar recovery yields (25–47 mg purified protein per L culture or ~7 mg per g cell pellet). However, these variants had lower expression levels than those of wild-type GloA3 and its single deletional mutations. The most unstable mutein was the triple deletional protein, delbcdGloA3, which possessed the lowest recovery yield (5 mg purified protein per 1 L culture or 1.5 mg per g cell pellet). This triple deletional mutant and delbcGloA3 were difficult to purify and only gave low recovery yields because they could not tolerate high salt concentrations (protein precipitation being observed).

Table 3.2: The summary of calculated molecular mass, recovery yield and theoretical p*I* for wild-type GloA3 and its variants

Sample	***[#]Molecular mass (Da)***	***Recovery yield (mg)/L culture***	***Recovery yield (mg)/g cell pellet***	***[#]Theoretical pI***
GloA3	20083.5	232	29	5.00
delbGloA3	18292.5	52	6	4.95
delcGloA3	19462.9	111	20	5.39
deldGloA3	19587.0	85	10	4.99
delbcGloA3	17672.0	25	6	5.35
delbdGloA3	17796.0	47	8	4.94
delcdGloA3	18966.4	33	7	5.38
delbcdGloA3	17106.4	10	1.5	5.34

[#] Molecular mass and p*I* of the monomeric enzyme was theoretical calculated from amino acid sequences using the web-based software, ProtParam tool from ExPASy Proteomics Server (http://ca.expasy.org/tools/protparam.html).

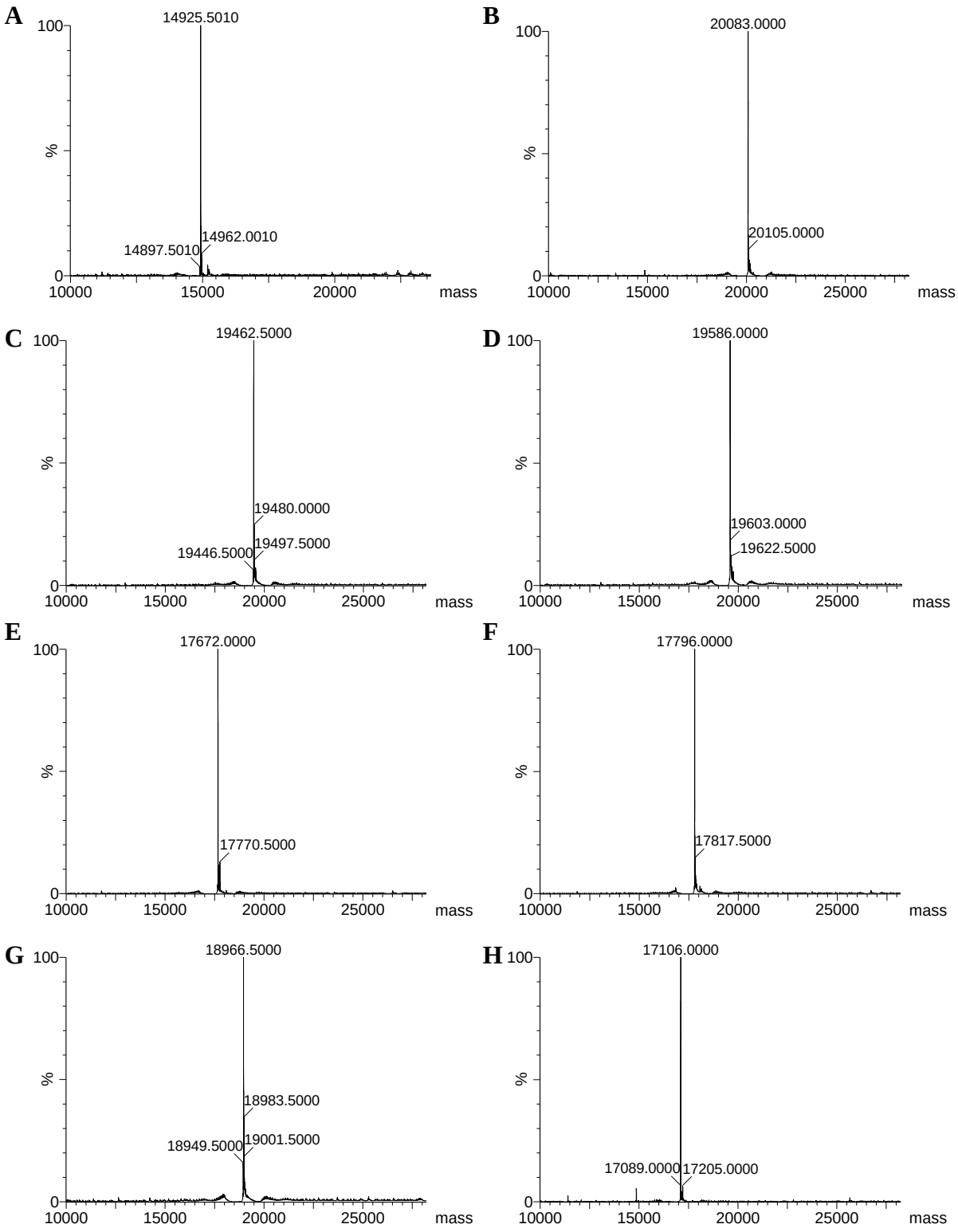


Figure 3.8: The electrospray mass spectra of (**A**) GloA2 with the molecular weight of a single subunit at 14925.5 Da (calculated MW is 14925.8 Da), (**B**) GloA3 with the molecular weight of a single subunit at 20083.0 Da (calculated MW is 20083.5 Da), (**C**) delcGloA3 with the molecular weight of a single subunit at 19462.5 Da (calculated MW is 19462.9 Da), (**D**) deldGloA3 with the molecular weight of a single subunit at 19586.0 Da (calculated MW is 19587.0 Da), (**E**) delbcGloA3 with the molecular weight of a single subunit at 17672.0 Da

(calculated MW is 17672.0 Da), (**F**) delbdGloA3 with the molecular weight of a single subunit at 17796.0 Da (calculated MW is 17796.0 Da), (**G**) delcdGloA3 with the molecular weight of a single subunit at 18966.5 Da (calculated MW is 18966.4 Da) and (**H**) delbcdGloA3 with the molecular weight of a single subunit at 17106.0 Da (calculated MW is 17106.4 Da).

3.4.2: Effect of Deletional Mutagenesis on Zinc-Activated Glyoxalase I Activity

The preliminary kinetic data of the wild-type GloA3 enzyme and its variants were performed using the purified enzymes that were induced and expressed in the presence of 1 mM metal ($NiCl_2$ or $ZnCl_2$). These metal-bound native enzymes were then incubated with an additional 5 equivalents of the particular metal overnight at 4 °C to ensure metal incorporation into the active sites of the enzymes. The kinetic data of all samples are summarized in Table 3.3. The native enzymes (the enzymes that were overexpressed by IPTG induction without any additional metals) were purified and enzymatic assays were done in parallel as controls. These native enzymes were found to exhibit some activities even without additional metals, suggesting that these enzymes were able to pick up some metals from their environment. The wild-type GloA3, a Zn^{2+}-activated enzyme, had observable activity in the presence of both Ni^{2+} (k_{cat} of 669 s^{-1} and K_m of 0.13 ± 0.01 mM) and Zn^{2+} (k_{cat} of 556 s^{-1} and K_m of 0.21 ± 0.01 mM). Thus, its relative ratio of k_{cat}/K_m of the Ni^{2+}-bound native GloA3 to its Zn^{2+}-bound native form was approximately 2. These data correspond to the previously reported kinetics (Zn^{2+}-bound GloA3: k_{cat} of 787 s^{-1} and K_m of 0.29 ± 0.05 mM) [33]. Interestingly, delbGloA3 possessed highest activity in the presence of Ni^{2+} (k_{cat} of 1661 s^{-1} and K_m of 1.06 ± 0.09 mM) among all wild-type enzymes and its muteins, while its Zn^{2+}-bound counterpart and its native form only resulted in a trace of activity. The relative ratio of k_{cat}/K_m of the Ni^{2+}-bound native delbGloA3 to its Zn^{2+}-bound native form was 493, the highest among single deletional mutants. The Ni^{2+}-bound native delcGloA3 had similar activity to that of wild-type GloA3 (k_{cat} of 616 s^{-1} and K_m of 0.71 ± 0.03 mM); however, the activity of its Zn^{2+}-bound form was drastically decreased (k_{cat} of 71 s^{-1} and K_m of 0.56 ± 0.05 mM). The difference in activity of the Ni^{2+} and Zn^{2+}-activation causes the relative ratio of k_{cat}/K_m for the Ni^{2+}-bound native delcGloA3 to its Zn^{2+}-bound counterpart to be only 7. On the other hand, deldGloA3 possessed similar enzymatic activity in the presence of Ni^{2+} (k_{cat} of 244 s^{-1} and K_m of 0.72 ± 0.01 mM) and Zn^{2+} (k_{cat} of 167 s^{-1} and K_m of 0.41 ± 0.19 mM). Both metal bound forms, however, contained lower activity than those of the wild-type enzyme.

Since it was previously reported that the non-Zn^{2+}-activated GlxI enzymes (being Ni^{2+}/Co^{2+}-activated) exhibit no enzyme activity in the presence of Zn^{2+} [19, 32], delbGloA3 appears to have lost its characteristic of being Zn^{2+}-activated and has become more like the Ni^{2+}/Co^{2+}-activated GlxI enzymes, which is totally different from the original wild-type GloA3. Thus, helix B might play a crucial role in the metal selection and activation of GlxI enzymes. Besides, the activity loss in delcGloA3 in the presence of Zn^{2+} might also indicate that region C of the protein can also play a minor part in determination of the metal activation profile of the enzyme. On the other hand, the lower activity of both Ni^{2+}- and Zn^{2+}-bound native deldGloA3 suggests that loop D might be insignificant in the control of metal preference but is still important for enzymatic activity. The reason behind the activity loss might be the close location of this loop with one of the metal binding residues, His^{118}. Thus, deletion of this loop might affect the metal binding affinity but not metal selectivity.

Among all three double deletional mutants, delbcGloA3 possessed the lowest enzymatic activity in the presence of either Ni^{2+} (k_{cat} of 1.7 s^{-1} and K_m of 0.22 ± 0.03 mM) or Zn^{2+} (k_{cat} of 0.07 s^{-1} and K_m of 0.28 ± 0.01 mM). On the other hand, delbdGloA3 exhibited high activity with Ni^{2+} (k_{cat} of 268 s^{-1} and K_m of 0.98 ± 0.09 mM), but contained only trace activity in the presence of Zn^{2+} (k_{cat} of 0.73 s^{-1} and K_m of 1.93 ± 0.05 mM). Thus, this variant exhibited a relatively high ratio of k_{cat}/K_m for the Ni^{2+}-bound native enzyme compared to its Zn^{2+}-bound counterpart (721 times). The activity of the delcdGloA3 protein was observable in the presence of both Ni^{2+} (k_{cat} of 216 s^{-1} and K_m of 0.38 ± 0.02 mM) and Zn^{2+} (k_{cat} of 83 s^{-1} and K_m of 0.80 ± 0.08 mM). Even though the activities of this mutein were lower than those of wild-type GloA3, the relative ratio of k_{cat}/K_m for the Ni^{2+}-bound native enzyme to its Zn^{2+}-bound counterpart was observed to be higher (6 times). The triple deletional variant, delbcdGloA3, possessed a dramatic decrease in activity for both Ni^{2+}- and Zn^{2+} (k_{cat} of 10 s^{-1} and K_m of 0.97 ± 0.01 mM for Ni^{2+}-bound native enzyme and k_{cat} of 0.02 s^{-1} and K_m of 0.5 ± 0.05 mM for Zn^{2+}-bound form). Its relative ratio of k_{cat}/K_m for the Ni^{2+}-bound native enzyme to its Zn^{2+}-bound form, however, was greater (334 times) than that of wild-type GloA3.

From these data, it is clear that these deletional combinations affected both metal activation profile and enzymatic activity. The combination of helix B and loop D seemed to be significant for metal activation in comparison to other double and triple deletional mutations. The drastic decrease in activity of this variant as well as its preference toward Ni^{2+}-activation

was the result of the deletion of helix B, which seems to control metal selection, and deletional loop D, which might have a predominant effect on the catalytic activity. The drastically decreased activity of the delbcGloA3 protein might be the combined result of the deletion of a large helix B and a loop C. Since the single deletions of these two regions gave lower activity than those of wild-type GloA3 and deldGloA3, the decrease in activity of their combinations was expected. Among all double mutations, delcdGloA3 possessed the highest activity with a preference toward Ni^{2+}-activation. This variant contained the deletion of loop C that gave a relatively smaller effect on metal selection comparing to that of helix B. As well, the deletion of loop D influenced catalytic activity. The triple mutant, however, exhibited relatively low activity for both metals but had an observable preference toward Ni^{2+}-activation.

Table 3.3: The preliminary data for kinetics of wild-type GloA3 and its variants prepared from *E. coli* expressions systems that were grown and expressed in the presence and in the absence of 1 mM metals ($NiCl_2$ and $ZnCl_2$). The purified enzymes from cells grown in the presence of metals were incubated with 5 equivalents of those particular metals overnight at 4 °C before performing the enzymatic assay using 0.02–1 mM substrate (MG-GSH) in 50 mM KPB (pH 6.6) at room temperature.

Enzyme	***Metal chloride***	V_{max} ***(μmol/min/mg)***	K_m ***(mM)***	k_{cat} (s^{-1})	k_{cat}/K_m $(M^{-1}.s^{-1})$	[*] ***Relative*** k_{cat}/K_m
GloA3	-	589 ± 42	0.24 ± 0.04	395	1.7×10^6	0.6
	Zn^{2+}	831 ± 45	0.21 ± 0.01	556	2.6×10^6	1
	Ni^{2+}	1000 ± 67	0.13 ± 0.01	669	5.3×10^6	2
delbGloA3	-	2.6 ± 0.3	0.71 ± 0.09	1.6	2188	0.7
	Zn^{2+}	4.6 ± 0.5	0.88 ± 0.16	2.8	3182	1
	Ni^{2+}	2724 ± 174	1.06 ± 0.09	1661	1.6×10^6	493
delcGloA3	-	145 ± 9	2.23 ± 0.18	94	4.2×10^4	0.3
	Zn^{2+}	109 ± 9	0.56 ± 0.05	71	1.3×10^5	1
	Ni^{2+}	950 ± 43	0.71 ± 0.03	616	8.7×10^5	7
deldGloA3	-	183 ± 41	1.67 ± 0.49	120	7.2×10^4	0.2
	Zn^{2+}	258 ± 61	0.41 ± 0.19	168	4.1×10^5	1
	Ni^{2+}	374 ± 14	0.72 ± 0.01	244	3.4×10^5	0.8
delbcGloA3	-	0.030 ± 0.006	0.06 ± 0.01	0.02	277	1.1
	Zn^{2+}	0.116 ± 0.005	0.28 ± 0.01	0.07	241	1
	Ni^{2+}	2.9 ± 0.6	0.22 ± 0.03	1.7	7977	33
delbdGloA3	-	0.043 ± 0.003	0.05 ± 0.02	0.03	568	1.5
	Zn^{2+}	1.237 ± 0.023	1.93 ± 0.05	0.73	381	1
	Ni^{2+}	451 ± 45	0.98 ± 0.09	268	2.7×10^5	721
delcdGloA3	-	58 ± 2	0.64 ± 0.01	36	5.7×10^4	0.6
	Zn^{2+}	132 ± 14	0.80 ± 0.08	83	1.0×10^5	1
	Ni^{2+}	341 ± 18	0.38 ± 0.02	216	5.8×10^5	6
delbcdGloA3	-	0.014 ± 0.001	0.55 ± 0.07	0.01	15	0.5
	Zn^{2+}	0.027 ± 0.002	0.50 ± 0.05	0.02	31	1
	Ni^{2+}	17.39 ± 0.04	0.97 ± 0.01	10.0	1.0×10^4	334

[*]Relative k_{cat}/K_m was calculated from the k_{cat}/K_m of each specific sample relative to that of its Zn^{2+}-bound form.

From analysis of these preliminary data, it was found that the delbGloA3 protein exhibited the highest activity in its Ni^{2+}-bound form, followed in catalytic activity by wild-type GloA3, delcGloA3, delbdGloA3, deldGloA3 and delcdGloA3, respectively (Figure 3.9A). The Ni^{2+}-bound form of the other variants only exhibited trace activity. Among the samples that exhibited high activity in the presence of Ni^{2+}, only delbGloA3 and delbdGloA3 possessed trace activity with Zn^{2+}. Analyses of this data suggested that the two variants had a high preference for Ni^{2+}/Co^{2+}-activation. Besides, delbdGloA3 protein exhibited the highest relative ratio of activity of the Ni^{2+}-bound native enzyme to its Zn^{2+}-bound counterpart among all mutants (Figure 3.9B). However, even though this variant exhibited activity in the presence of Ni^{2+}, its absolute activity was decreased by 20-fold compared to that of the wild-type GloA3. Other muteins including delbGloA3, delbcdGloA3 and delbcGloA3 also exhibited observable higher relative activity for the Ni^{2+}-bound native enzyme to the Zn^{2+}-bound counterpart (relative activity ratio of delbdGloA3 > delbGloA3 > delbcdGloA3 > delbcGloA3). Although the triple deletional mutation delbcdGloA3 had low activity in the presence of both Ni^{2+} and Zn^{2+}, its relative ratio of activity for the Ni^{2+}-bound form to its Zn^{2+}-bound counterpart was high. It was suggested that only mutations with a deleted B region exhibited the high activity ratio of the Ni^{2+}-bound form to its Zn^{2+}-bound form, while others did not. Thus, helix B might be significant for metal selection and activation. Loop C is possibly an important contributor to both metal preference as well as enzymatic activity, and loop D might only affect enzymatic activity.

Nevertheless, these activities might be the result of other metals that enzymes accept from their environment during their expression and purification processes. In order to confirm that only one type of metal was incorporated into the active sites of the enzymes, the selected samples with high relative k_{cat}/K_m ratio of Ni^{2+}-/Zn^{2+}-bound native forms including wild-type GloA3, delbGloA3, delbdGloA3 and delbcdGloA3 were prepared as apo-enzymes. Their metal titration profiles were then investigated, which eventually led to optimization of enzymatic activities.

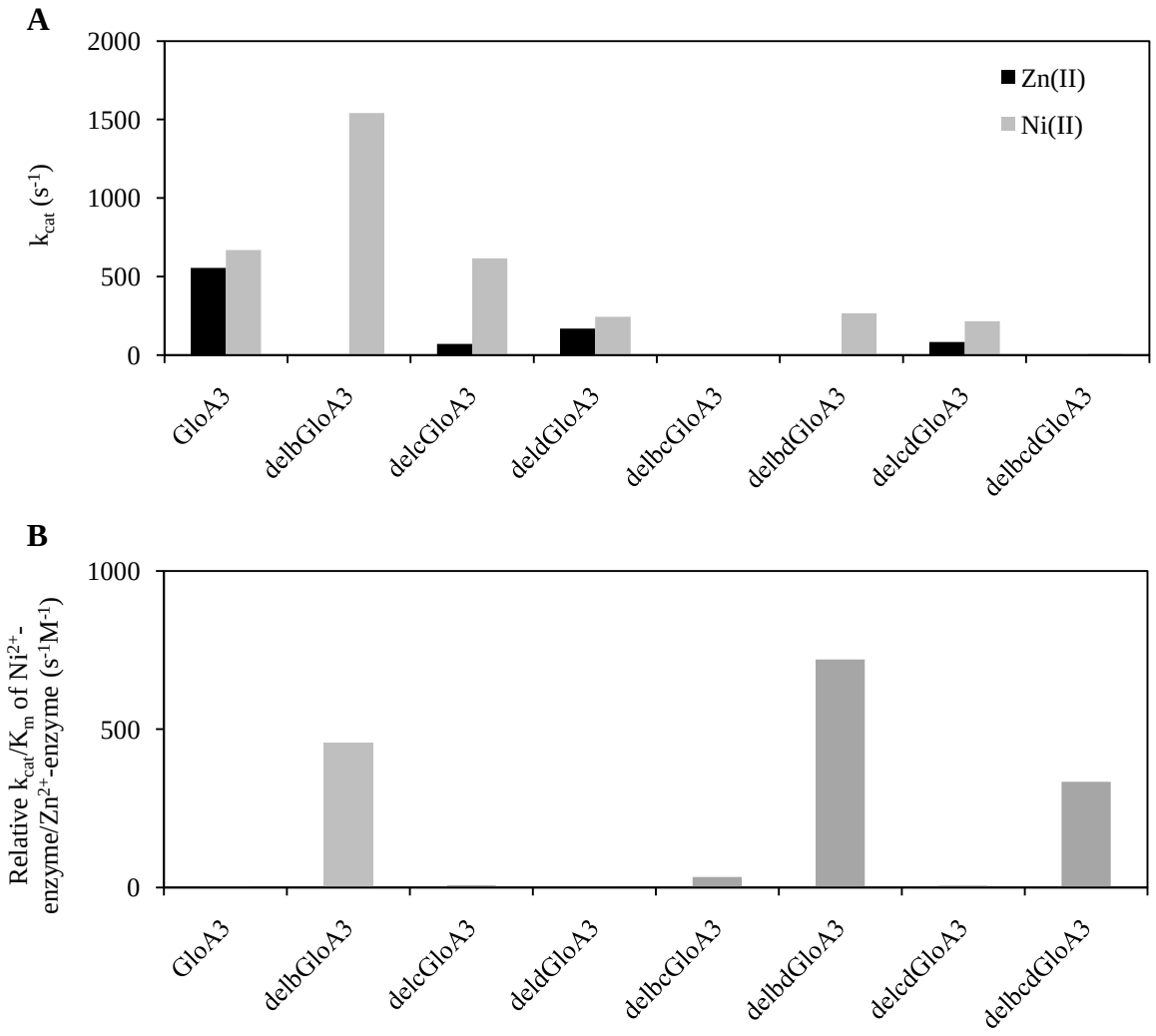


Figure 3.9: The comparison of **(A)** k_{cat} of Zn^{2+}- and Ni^{2+}-bound native enzymes and **(B)** the relative ratio of k_{cat}/K_m of Ni^{2+}-bound native enzymes to the Zn^{2+}-bound forms used for further investigation on selected samples that might change their characteristics from being Zn^{2+}-activated enzyme to Ni^{2+}/Co^{2+}-activation. The enzymes were produced and expressed in the presence of 1 mM metals ($NiCl_2$ and $ZnCl_2$). The purified enzymes were incubated with an additional 5 equivalents of the particular metals before performing enzymatic assays using 0.02–1 mM substrate (MG-GSH) in 50 mM KPB (pH 6.6) at room temperature.

The structural investigations of the human GlxI (Zn^{2+}-activation) have suggested that its active sites were more packed and less opened than those of the *E. coli* enzyme (Ni^{2+}/Co^{2+}-activation) due to the extra helices A and B and loops C and D (Figure 3.2) [44, 46]. These structures, therefore, support strong interactions between metal and enzyme, which in turn causes the metal removal from the Zn^{2+}-bound form to be more difficult than with Ni^{2+}/Co^{2+}-activated enzyme. EDTA was previously reported to be inefficient in removing metals from GloA3 (Zn^{2+}-activation) and a stronger metal chelator such as dipicolinic acid (DPA) was required for preparation of the apo-enzyme [33]. It was found that the metal removal from the

selected GloA3 variants (delbGloA3, delbdGloA3 and delbcdGloA3) was not harsh compared to the effort required to remove metal from the wild-type GloA3, supporting our previous hypothesis that these variants might have preference toward being Ni^{2+}/Co^{2+}-activation. Analyses of the metal titration profiles of wild-type GloA2 (with Ni^{2+}) and GloA3 (with Ni^{2+} and Zn^{2+}) indicated that the metal per dimeric enzyme ratio was approximately 1 (Figure 3.10A-B), which is similar to what is found for GlxI from *E. coli* GlxI and *Plasmodium falciparum* GlxI [39, 43]. These results also suggested that the metal is tightly bound to the enzyme and only one active site in the dimer is functional. Higher concentrations of metals did not seem to disturb the enzymatic activity of either GloA2 (0–200 equivalents per dimeric enzyme or 0–7 μM $NiCl_2$) or GloA3 (0–200 equivalents per dimeric enzyme or 0–5 μM $NiCl_2$ and $ZnCl_2$). The enzymatic activity of Ni^{2+}-reconstituted GloA2 (V_{max} of 479 ± 13 μmol/min/mg, K_m of 0.17 ± 0.01 mM and k_{cat} of 238 s^{-1}) was slightly increased from that of the Ni^{2+}-bound native enzyme (the purified GloA2 that was grown and expressed with additional 1 mM $NiCl_2$ and was assayed with an additional 5 equivalents of metal to give a V_{max} of 377 ± 7 μmol/min/mg, K_m of 0.12 ± 0.02 mM and k_{cat} of 188 s^{-1}). Analyses of these data suggested that the metal-bound native enzyme might accept other metals in its active site, while the Ni^{2+}-reconstituted form only contained Ni^{2+} atoms. On the other hand, the activity of the Zn^{2+}-reconstituted GloA3 was reduced to only 41% (V_{max} of 337 ± 39 μmol/min/mg, K_m of 0.27 ± 0.02 mM and k_{cat} of 226 s^{-1}) of that of its Zn^{2+}-bound native enzyme (V_{max} of 831 ± 45 μmol/min/mg, K_m of 0.21 ± 0.01 mM and k_{cat} of 556 s^{-1}) (Table 3.4). The decrease in activity of holoGloA3 with Zn^{2+} was also previously reported, suggesting that the Zn^{2+} atom could inhibit enzymatic activity, possibly through interaction with residues close to the active site [33]. Thus, the activity of the purified GloA3 that was grown and expressed in the presence of 2 mM $ZnCl_2$ (V_{max} of 1102 ± 6 μmol/min/mg, K_m of 0.48 ± 0.02 mM and k_{cat} of 738 s^{-1}) and was proven to contain ~1 mole metal per mole dimeric enzyme by ICP-MS (Table A4.3 in Appendix 4) was then used for comparison with other GloA3 muteins.

Interestingly, analysis of the metal titration of the Ni^{2+}-reconstituted delbGloA3 indicated that maximum activity was achieved in the presence of 5 equivalents of metal (~0.7 μM) (Figure 3.10C), while that of its Zn^{2+}-reconstituted counterpart only required 1 mole of metal (~0.8 μM) per mole of dimeric enzyme (Figure 3.10D). The loss in activity of Zn^{2+}-reconstituted delbGloA3 was observed when more than 5 equivalents of metal (> 4.1 μM) were

added into the apo-enzyme. Thus, high concentrations of Zn^{2+} could inhibit enzymatic activity, possibly *via* severe interactions of metal-enzyme complex that can cause protein precipitation. Metal inhibition of GlxI has been previously observed for wild-type GloA3 [33]. The activity of Zn^{2+}-reconstituted delbGloA3 (V_{max} of 7.8 ± 2.2 μmol/min/mg, K_m of 1.37 ± 0.49 mM and k_{cat} of 4.8 s^{-1}, Table 3.4) was elevated two-fold higher than that of its Zn^{2+}-bound native enzyme (V_{max} of 4.6 ± 0.5 μmol/min/mg, K_m of 0.88 ± 0.16 mM and k_{cat} of 2.8 s^{-1}, Table 3.3). Analyses of these data suggested that the amount of Zn^{2+} in the latter enzyme (with 5 equivalents of metal added before assay) might be too high such that it could produce a loss of catalytic activity. On the other hand, a high concentration of Ni^{2+} (up to 5000 equivalents or ~0.7 μM) neither affected enzymatic activity nor initiated observable protein aggregation. A slight loss in activity of the Ni^{2+}-reconstituted delbGloA3 (V_{max} of 2255 ± 24 μmol/min/mg, K_m of 0.98 ± 0.16 mM and k_{cat} of 1375 s^{-1}, Table 3.4) compared to its Ni^{2+}-bound native form (V_{max} of 2724 ± 174 μmol/min/mg, K_m of 1.06 ± 0.09 mM and k_{cat} of 1661 s^{-1}, Table 3.3) was observed. It is possible that some of the enzyme might become inactivated during the process of apo-enzyme preparation. The apo-enzyme might be unstable and the addition of Ni^{2+} might not fully recover the enzymatic activity. However, the relative ratio of k_{cat}/K_m of the Ni^{2+}-reconstituted delbGloA3 to its Zn^{2+}-reconstitution form (404 times) was still higher than that of wild-type GloA3 (only 2–3 times).

On the other hand, the metal titration of Ni^{2+}-reconstituted delbdGloA3 required high concentrations of metal to obtain maximum activity (at least 50 moles of metal (~140 μM) per mole dimeric enzyme, Figure 3.10E). Neither protein precipitation nor activity loss was observed in these cases. The activity of Ni^{2+}-reconstituted delbdGloA3 (V_{max} of 282 ± 52 μmol/min/mg, K_m of 1.06 ± 0.22 mM and k_{cat} of 167 s^{-1}, Table 3.4), however, was approximately 2-fold lower than that of its Ni^{2+}-bound native enzyme (V_{max} of 451 ± 45 μmol/min/mg, K_m of 0.98 ± 0.09 mM and k_{cat} of 268 s^{-1}, Table 3.3). Its Zn^{2+}-reconstituted counterpart only required 1 equivalent of metal (~14 μM) to reach the maximum activity (Figure 3.10F), similar to the wild-type GloA3. Its kinetics (V_{max} of 1.53 ± 0.17 μmol/min/mg, K_m of 1.02 ± 0.21 mM and k_{cat} of 0.89 s^{-1}, Table 3.4) were also similar to that of the Zn^{2+}-bound native delbdGloA3 (V_{max} of 1.24 ± 0.02 μmol/min/mg, K_m of 1.93 ± 0.05 mM and k_{cat} of 0.73 s^{-1}, Table 3.3). However, higher concentrations of Zn^{2+} (more than 3 equivalents of

metal or >42 μM) could initiate protein aggregation and, thus, activity loss. Due to the decrease in activity of its Ni^{2+}-reconstituted enzyme, the relative ratio of k_{cat}/K_m of Ni^{2+}- to Zn^{2+}-reconstituted forms was reduced to 181-fold, in comparison to the previous relative ratio of k_{cat}/K_m of Ni^{2+}- to Zn^{2+}-bound native enzymes, which was 721 times (Table 3.3).

The metal titration of Ni^{2+}-reconstituted delbcdGloA3 was similar to that of delbGloA3, in which the maximum activity was achieved with the addition of 5 equivalents of metal (~1 μM $NiCl_2$) per dimeric enzyme (Figure 3.10G). Its Zn^{2+}-reconstituted counterpart required 2 equivalents of metal (~29 μM) to reach its optimal activity (Figure 3.10H). Higher concentrations of Ni^{2+} (up to 5000 equivalents of metal or ~0.4 mM) did not initiate protein precipitation nor change in catalytic activity. However, clearly observable protein precipitation was observed in the presence of high concentrations of Zn^{2+} (more than 2 equivalents or >29 μM), which is significant for activity loss in the Zn^{2+}-reconstitution. The activities of the Ni^{2+}-reconstituted delbcdGloA3 (V_{max} of 24 ± 2 μmol/min/mg, K_m of 1.22 ± 0.19 mM and k_{cat} of 14 s^{-1}, Table 3.4) were slightly higher than those of its Ni^{2+}-bound native delbcdGloA3 enzyme (V_{max} of 17.39 ± 0.04 μmol/min/mg, K_m of 0.97 ± 0.01 mM and k_{cat} of 10 s^{-1}, Table 3.3), while the activities of its Zn^{2+}-reconstituted counterpart (V_{max} of 0.04 ± 0.001 μmol/min/mg, K_m of 0.43 ± 0.03 mM and k_{cat} of 0.02 s^{-1}, Table 3.4) was similar to those of the Zn^{2+}-bound native form (V_{max} of 0.03 ± 0.002 μmol/min/mg, K_m of 0.5 ± 0.05 mM and k_{cat} of 0.02 s^{-1}, Table 3.3). These activities, thus, caused an effect on the relative ratio of k_{cat}/K_m of the Ni^{2+}- to Zn^{2+}-reconstituted enzymes that was reduced to 212-fold (Table 3.4) in comparison to the relative k_{cat}/K_m ratio of Ni^{2+}- to Zn^{2+}-bound native forms (334 times, Table 3.3).

Analyses of these metal titration profiles suggested that all selected GloA3 variants possessed lower metal per dimeric enzyme ratio for bound Zn^{2+} atoms (approximately 1 mole Zn^{2+} atom per mole dimeric enzyme) to produce maximal catalytic activity than those for Ni^{2+} addition (more than 5 moles Ni^{2+} per mole enzyme). Analyses of these data suggested that Zn^{2+} strongly interacts with the enzymes. Ni^{2+}, on the other hand, only forms marginal metal-enzyme interactions, which might possibly cause a lower binding affinity. These muteins, however, were favorable toward being Ni^{2+}-activated, where the activities of the Ni^{2+}-reconstituted enzymes were more than 200 times greater than that of their Zn^{2+}-reconstituted forms. It was also confirmed that helix B was significantly involved in metal activation of GlxI. The mutein, delbGloA3, was highly favorable toward Ni^{2+}-activation with higher

enzymatic activity in the presence of Ni^{2+} than GloA1 (V_{max} of 571 ± 28 μmol/min/mg, K_m of 32 ± 2 μM and k_{cat} of 271 s^{-1}) and GloA2 (V_{max} of 497 ± 8 μmol/min/mg, K_m of 21 ± 0 μM and k_{cat} of 247 s^{-1}) [33], the actual two Ni^{2+}-activated GlxI isolated from *P. aeruginosa*.

Table 3.4: The summary of kinetics of wild-type GloA3 and its selected muteins including delbGloA3, delbdGloA3 and delbcdGloA3 proteins showing the different catalytic activities in the presence of Ni^{2+} and Zn^{2+} ions. The proteins were overproduced and expressed without additional metals and were purified and prepared in the apo-forms. Metals ($NiCl_2$ or $ZnCl_2$) were then added into the apo-enzymes and were incubated overnight at 4 °C before performing enzymatic assays using 0.02–1 mM substrate (MG-GSH) in 50 mM KPB (pH 6.6) at room temperature.

Enzyme	***Metal chloride***	***Equivalents of metal****	***V_{max} (μmol/min/mg)***	***K_m (mM)***	***k_{cat} (s^{-1})***	***k_{cat}/K_m ($M^{-1}.s^{-1}$)***	***[#]Relative k_{cat}/K_m***
GloA3	[†]Zn^{2+}	5	337 ± 39	0.27 ± 0.02	226	8.4×10^5	1
	[$]$Zn^{2+}$	5	831 ± 45	0.21 ± 0.01	556	2.6×10^6	3.2
	[£]Zn^{2+}	-	1102 ± 6	0.48 ± 0.02	738	1.5×10^6	1.8
delbGloA3	Zn^{2+}	5	7.8 ± 2.2	1.37 ± 0.49	4.8	3472	1
	Ni^{2+}	5	2255 ± 24	0.98 ± 0.16	1375	1.4×10^6	404
delbdGloA3	Zn^{2+}	2	1.53 ± 0.17	1.02 ± 0.21	0.89	873	1
	Ni^{2+}	50	282 ± 52	1.06 ± 0.22	167	1.6×10^5	181
delbcdGloA3	Zn^{2+}	2	0.04 ± 0.001	0.43 ± 0.03	0.02	53	1
	Ni^{2+}	5	24 ± 2	1.22± 0.19	14	1.1×10^4	212

[†] The purified holoGloA3 that assayed with 5 equivalents of metal
[$] The purified GloA3 that was produced with 1 mM $ZnCl_2$ and assayed with additional 5 equivalents of metal
[£] The purified GloA3 that was produced with 2 mM $ZnCl_2$ and assayed without additional metal
* The equivalents of moles metal per 1 mole of dimeric enzymes
[#] Relative k_{cat}/K_m was calculated from the k_{cat}/K_m of each specific sample relative to that of its Zn^{2+}-reconstituted form.

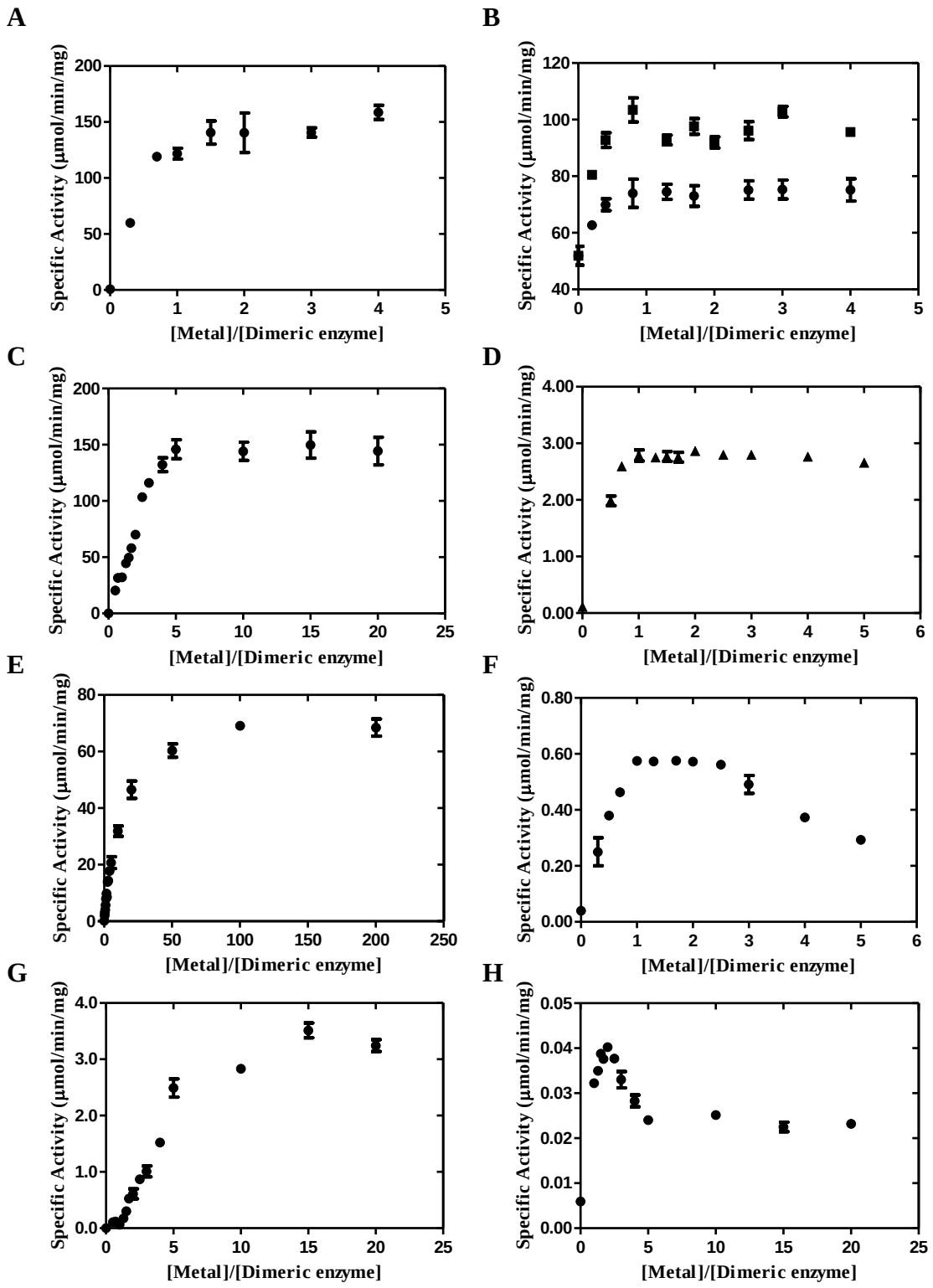


Figure 3.10: Metal titration profiles for **(A)** GloA2 with Ni^{2+}, **(B)** GloA3 with Ni^{2+} (■) and Zn^{2+} (●), **(C)** delbGloA3 with Ni^{2+}, **(D)** delbGloA3 with Zn^{2+}, **(E)** delbdGloA3 with Ni^{2+}, **(F)** delbdGloA3 with Zn^{2+}, **(G)** delbcdGloA3 with Ni^{2+} and **(H)** delbcdGloA3 with Zn^{2+}. The apo-enzymes for GloA3 variants were incubated with metals overnight at 4 °C, while GloA2 and GloA3were incubated with metals for 15 min. The assay was performed using 0.5 mM MG-GSH in 50 mM KPB (pH 6.6) at room temperature.

3.4.3: Structural Investigation of GloA2, GloA3 and Deletional Mutants

The secondary structures of the proteins were investigated through circular dichroism analysis. The CD spectra of GloA2 under different protein concentrations (5.0–0.6 mg/mL) in 50 mM Tris (pH 7.5), 150 mM NaCl and 10% (v/v) glycerol indicated a change in shape as well as a shift of negative maxima from 208 and 220 nm to 223 nm (Figure 3.11A). These changes suggested protein aggregation and formation of oligomers at higher protein concentrations (>1.3 mg/mL). The stable form of GloA2 could be found at lower protein concentrations (<1.3 mg/mL) with the predicted secondary structure content of 26% α-helix, 21% β-sheet and 53% random coil. Similarly, the CD spectra of GloA3 at lower protein concentrations (<1.3 mg/mL) resulted in negative bands at 208 and 222 nm with the predicted secondary structural content of 28% α-helix, 23% β-sheet and 49% random coil (Figure 3.11B). However, these negative maxima shifted to 226 nm at higher protein concentrations (>1.3 mg/mL). The shape and negative maxima in the CD spectra of GloA2 and GloA3 suggested that their secondary structures might be similar.

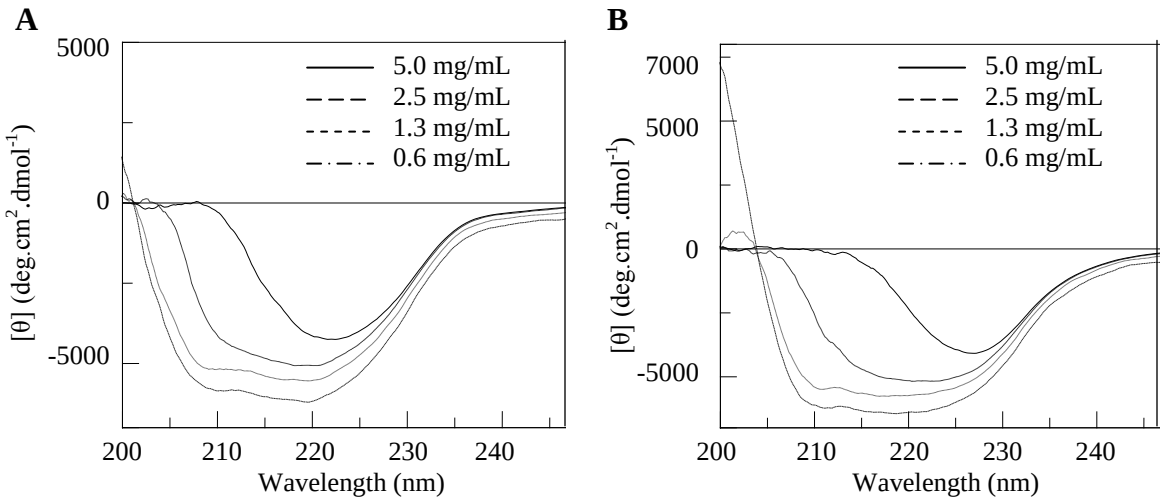


Figure 3.11: The CD spectra of **(A)** Ni^{2+}-bound GloA2 (5–0.6 mg/mL) and **(B)** Zn^{2+}-bound GloA3 (5–0.6 mg/mL) in 50 mM Tris (pH 7.5), 150 mM NaCl and 10% (v/v) glycerol scanning between 200–250 nm. Note that the quality of CD signal is significantly affected by the samples with high protein concentrations (> 2.5 mg/mL) at wavelengths below 210 nm, which cause saturated high tension (HT) voltage in the instrument.

The secondary structures of the GloA3 variants were similar to that of the wild-type protein, where negative maxima at 208 and 222 nm were observed (Figure 3.12). The CD spectra of all variants were unaltered with the addition of Ni^{2+} and Zn^{2+} atoms, suggesting that

the presence of metals have no effect on their secondary structures (data not shown). Interestingly, the variants with a deleted helix B (delbGloA3, delbcGloA3, delbdGloA3 and delbcdGloA3) contained similar predicted secondary structural content as those of wild-type GloA3 under the same studied conditions (9% α-helix, 43% β-sheet and 48% random coil, Table 3.5). These data suggested that the deletion of helix B does not initiate a secondary structural change, even though this is the largest deletion of all three (helix B, loop C and loop D). On the other hand, other variants with deleted C and D segments resulted in a slightly higher content of α-helix and lower content of β-sheet than that of wild-type GloA3 (Table 3.5). Since loops C and D were predicted to connect the N-terminal and the C-terminal domains as well as locate close to the metal binding residues, the changes in their secondary structures might explain the activity loss. These secondary structural predictions, thus, are consistent with the previous kinetic data, where it was suggested that helix B was significant for metal selection and slightly affected the catalytic activity. Loops C and D, on the other hand, were mainly crucial for enzymatic activity, in which a drastic decreased activity was observed with the variants with deleted C and D regions. Nevertheless, the secondary structure estimated from the CD spectra was only approximate, particularly for the predominantly β-structural containing protein.

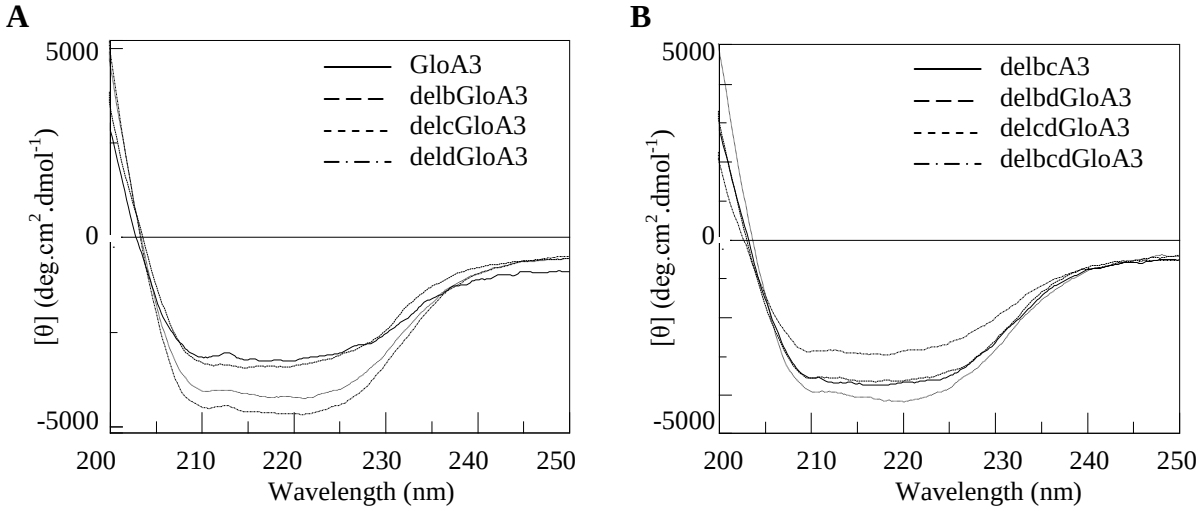


Figure 3.12: The CD spectra of the Ni^{2+}-bound enzymes (0.3 mg/mL) including **(A)** GloA3, delbGloA3, delcGloA3 and deldGloA3 and **(B)** delbcGloA3, delbdGloA3, delcdGloA3 and delbcdGloA3 in 50 mM Tris (pH 7.5) and 150 mM NaCl scanning between 200–250 nm.

Table 3.5: The predicted secondary structural contents of wild-type GloA3 and its variants (0.3 mg/mL) in 50 mM Tris (pH 7.5) and 150 mM NaCl using a web-based program, K2D-Prediction of percentages of protein secondary structure from CD spectra [152, 153].

	Content of the secondary structure		
Sample	***α-helix (%)***	***β-sheet (%)***	***Random coil (%)***
GloA3	9	43	48
delbGloA3	9	43	48
delcGloA3	14	37	48
deldGloA3	18	33	49
delbcGloA3	9	43	48
delbdGloA3	9	43	48
delcdGloA3	12	40	48
delbcdGloA3	9	43	48

The secondary structures of GlxI from other organisms measured by CD analysis were also previously reported. It was found that the predicted secondary structure of GloA2 and GloA3 were similar to GlxI from another Pseudomonad species, *P. putida* in the Zn_2 dimeric form, which possessed approximately 18% α-helix, 50% β-sheet and 33% random coil [30]. Human GlxI, on the other hand, contained 23% α-helix, 26% β-sheet and 51% random coil [44]. It was suggested that higher apparent β-sheet content in *P. putida* might be the result of non hydrogen-bonded extended regions that contribute to the β-sheet content estimated from the sample, whereas only hydrogen-bonded extended regions in the X-ray crystallographic structure were considered to be β-sheet [30]. Moreover, by comparing the secondary structures of this *P. putida* Zn_2 dimeric GlxI with those of its Zn monomeric form (with predicted ~54% α-helix, ~18% β-sheet and ~28% random coil), it was suggested that the predicted secondary structural content within the dimer (two active sites at the dimeric interface) and the monomer (one active site within one subunit) were quite different such that higher α-helix and lower β-sheet contents were observed in a monomeric protein [30]. These observations were consistent with the secondary structural prediction performed on the dimeric GlxI from *Clostridium acetobutylicum* (CLO GlxI, Chapter 2), in which its monomer likely possessed similar protein folding to the *P. putida* Zn monomeric enzyme. The CLO GlxI with an active site formed within one subunit (two active sites per dimer) also contained high α-helix and low β-sheet content (57% α-helix, 8% β-sheet and 34% random coil). Thus, different types of structural arrangement could possibly affect CD interpretation and analysis of the secondary structures,

even though X-ray crystallographic analysis indicated that the secondary structural contents of these proteins were similar. These observations, however, could be useful for the prediction of protein folding as well as its quaternary structure.

Despite their differences in the secondary structure, the quaternary structure of wild-type GloA3 and its variants were unaltered, in which all samples were shown to form homodimers (Figure 3.13, Table 3.6). A difference of less than 8 kDa between calculated molecular weight from gel permeation chromatography and their theoretical values was observed. This error was possibly produced from a slight shift in elution volume that might cause a large change in a plot of log molecular weight (example of calculation in Appendix 3). Similar to their secondary structures, metal-incorporations into the protein of Ni^{2+} and Zn^{2+} were found to have no effect toward dimerization of wild-type GloA3 and its variants. Besides, the buffer systems including Tris (purification buffer) and KPB (assay buffer) did not alter the quaternary structure (dimer) of the selected protein samples (wild-type GloA3, delbGloA3, delbdGloA3 and delbcdGloA3).

Table 3.6: The theoretical and calculated molecular mass of wild-type GloA3 and its variants (0.3 mg/mL) in 50 mM Tris (pH 7.5) and 150 mM NaCl by gel permeation chromatography using a Superdex75 HR 10/30 column with a flow rate of 0.5 mL/min

Sample	***Calculated molecular mass (kDa) from gel permeation chromatography***	***†Theoretical molecular mass (kDa) of homodimeric protein***	***The difference between calculated and theoretical molecular mass (kDa)***
GloA3	32.4	40.4	-8.0
delbGloA3	34.6	36.6	-2.0
delcGloA3	35.4	38.9	-3.5
deldGloA3	32.4	39.2	-6.8
delbcGloA3	35.7	35.3	+0.4
delbdGloA3	35.9	35.6	+0.3
delcdGloA3	29.7	37.9	-8.2
delbcdGloA3	36.2	34.4	+1.8

† Molecular mass of the dimeric enzyme was theoretical calculated from amino acid sequences using the web-based software, ProtParam tool from ExPASy Proteomics Server (http://ca.expasy.org/tools/protparam.html).

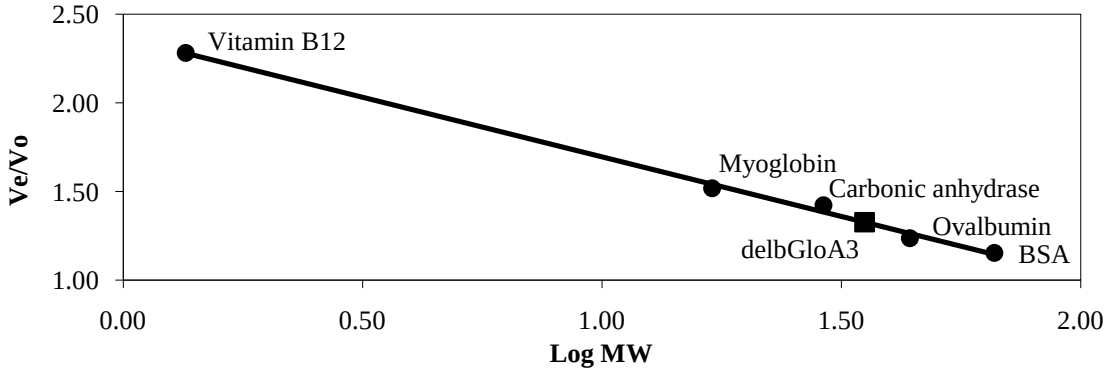


Figure 3.13: An example of a gel permeation chromatographic profile of (■) delbGloA3 (1 mg/mL) that fits into the Bio-Rad protein standard plot using a Superdex75 HR 10/30 column and a buffer containing 20 mM Tris (pH 7.5) and 150 mM KCl with a flow rate of 0.5 mL/min. The protein standards (●) consist of BSA (66 kDa), ovalbumin (44 kDa), carbonic anhydrase (29 kDa), myoglobulin (17 kDa) and vitamin B12 (1.35 kDa).

It was previously reported that *E. coli* GlxI is activated in the presence of Ni^{2+}, Co^{2+}, Cd^{2+} and Mn^{2+}, while its Zn^{2+}-bound form is inactive [19]. Only active enzyme would form an octahedral coordinated active site with four metal binding residues (His^{5}, Glu^{56}, His^{74} and Glu^{122}) and two water molecules around the metal center, while the inactive enzyme forms a trigonal bipyramidal metal-coordinated arrangement with only one water molecule bound [45]. The metal binding residues interact with the metal center *via* $His^{5}N_{\varepsilon 2}$, $Glu^{56}O_{\varepsilon 1}$, $His^{74}N_{\varepsilon 2}$ and $Glu^{122}O_{\varepsilon 1}$ with the ligand distance within 2.5 Å (Table 3.7). It was previously suggested that the appropriate metal coordination is a prerequisite for the active GlxI [45]. Apo-*E. coli* GlxI was found to form a six-liganded coordination with four binding residues and two water molecules around the water molecule in the center (PDB: 1FA8). Addition of the activating metals maintains the octahedral environment of the enzyme, while the inactive metals alter the metal coordination to a trigonal bipyramidal geometry. In the active enzyme, two water molecules are replaced by two oxygens from the acetal group of GSH moiety upon binding of the substrate (hemithioacetal) [46]. Thus, the octahedral environment in the active site is likely maintained (more details in Chapter 6). The distance between two oxygens of the GSH moiety (~2.38 Å measured from an unpublished structure of *E. coli* GlxI with bound transition state analogue of hydroxamate, TSI, and 2.56 Å measured from unpublished structure of *E. coli* GlxI with bound *N*-hydroxybenzamide, BHO) is comparable to the distance between two water molecules that locate close to the active metal (2.81 Å in Ni^{2+}-bound *E. coli* GlxI, PDB: 1F9Z). The trigonal bipyramidal geometry might not support this mode of substrate binding, thus resulting in activity loss.

It is interesting to investigate how metals influence active site geometries of the enzyme. The superimposed structure of *E. coli* GlxI with different incorporated metals suggested a possible explanation for the relationship between metal coordination geometry and catalytic activity. The overlapped structures of Ni^{2+}-bound *E. coli* GlxI (PDB: 1F9Z) and its Co^{2+}-bound form (PDB: 1FA6) show very good superimposition of the metal centers as well as metal binding residues (Figure 3.14A). However, the side chain of $Glu^{122}O_{\varepsilon 1}$ in the Co^{2+}-bound enzyme was tilted from the orientation found in the Ni^{2+}-bound form by 0.55 Å, while its $Glu^{122}O_{\varepsilon 2}$ was shifted by 0.25 Å. These shifts might affect the position of the water molecule (W1) in the Co^{2+}-bound enzyme such that it is tilted by 0.17 Å from its position in the Ni^{2+}-bound form. This shift shortens the distance between $Glu^{122}O_{\varepsilon 2}$ and W1 in the Co^{2+}-bound enzyme by 0.27 Å and generates a greater distance between the two water molecules bound to the metal center (2.94 Å). This distance is greater than the distance of two oxygens on the GSH moiety that replace water molecules in the active site (~2.38 Å). This might affect the binding affinity of the substrate in the active site and possibly cause a decrease in activity in the Co^{2+}-bound enzyme comparing to its Ni^{2+}-bound form.

Similarly, analyses of the superimposed structures of Ni^{2+}-bound *E. coli* GlxI (PDB: 1F9Z) and its Cd^{2+}-bound form (PDB: 1FA7) indicated that the side chain of the metal binding residue Glu^{122} is tilted, where $Glu^{122}O_{\varepsilon 1}$ and $Glu^{122}O_{\varepsilon 2}$ are shifted by 0.82 and 0.42 Å, respectively, away from their locations in the Ni^{2+}-bound form (Figure 3.14B). The distance between the two metal-bound water molecules in the active site of the Cd^{2+}-bound form is 2.96 Å, which is larger than that between the two water molecules in the Ni^{2+}- and Co^{2+}-bound enzymes and also larger than the distance between two oxygens of the GSH moiety. As well, the coordination distances between water molecules and the active metal of the Cd^{2+}-bound enzyme are greater (Table 3.7), which is consistent with the increased radius of the metal (the atomic radius of nickel is 124 pm, of cobalt is 125 pm and of cadmium is 151 pm). The increased distance between two water molecules and the distance between water molecules and the metal center might lower the substrate binding affinity as well as create an unsuitable environment for the bound substrate.

On the other hand, the inactive Zn^{2+}-bound *E. coli* GlxI (PDB: 1FA5) possesses a trigonal bipyramidal metal coordination with four metal binding residues and one water molecule. The superimposed structures of Ni^{2+}- and Zn^{2+}-bound enzymes (PDB: 1F9Z and

1FA5, respectively) are nicely overlapped (root mean square deviation of 0.16 Å for 128 C_α). There is no major movement of residues in the core of the enzyme. Only residues at the outer surface, the connecting loop (residues 63–68) and the loops that act as lids to cover the active site (residues 34–39 and 100–110) are flexible. Even though the distance between the metal binding residues and the metal center are similar to those in the Ni^{2+}-bound form (Table 3.7), slight shifts of the side chains of two of the metal binding residues, $Glu^{122}O_{\varepsilon 1}$ and $Glu^{56}O_{\varepsilon 1}$ (shifted 0.60 Å and 0.32 Å, respectively, away from the positions in the Ni^{2+}-bound form) are observed (Figure 3.14C). These shifts may possibly limit the catalytic space as well as influence the location of the water molecules. While the metals in Ni^{2+}-, Co^{2+}- and Cd^{2+}-bound enzymes are nicely overlapped, the Zn^{2+} atom is buried deeper into the catalytic pocket. It is possible that the Zn^{2+} atom (having atomic radius of 134 pm) can form stronger interactions with the metal binding residues and is pulled deeper into the catalytic pocket compared to Ni^{2+} and Co^{2+} atoms. Based on previous isothermal titration calorimetry studies in our laboratory, Zn^{2+} is found to bind more tightly to the enzyme than other metals [37]. As well, the distances between metal binding residues and the metal center are shorter than those of the Co^{2+}- and Cd^{2+}-bound enzyme (Table 3.7). The interaction of the Zn^{2+} atom and the enzyme (deeply buried in the active site) might create more space for metal binding residues to rotate, thus in turn limiting space for water molecules to enter the active site. This factor might create an environment that is suitable for a trigonal bipyramidal geometry, which allows only one water molecule to fit into the active site. On the other hand, the larger Cd^{2+} atom might be too large to fit properly into the active pocket. Therefore, it would be impossible for Cd^{2+} to be buried as deep as the Zn^{2+} atom and so creating limited space for metal binding residues to rotate (Figure 3.14D). Thus, the Cd^{2+}-bound enzyme is able to form an octahedral metal coordination with two water molecules.

Additionally, the superimposed structures of the Ni^{2+}-bound enzyme and its Zn^{2+}-bound form also allows us to hypothesize how an octahedral arrangement might look like if a Zn^{2+}-bound enzyme could bind two water molecules in the active site. However, even though this enzyme possesses two water molecules around the Zn^{2+} center, the bond distances and bond angles would not fit an octahedral geometry (Table 3.7 and 3.8). Thus, the Zn^{2+} atom in *E. coli* GlxI only creates a suitable environment for the trigonal bipyramidal coordination with only one water molecule around the metal center.

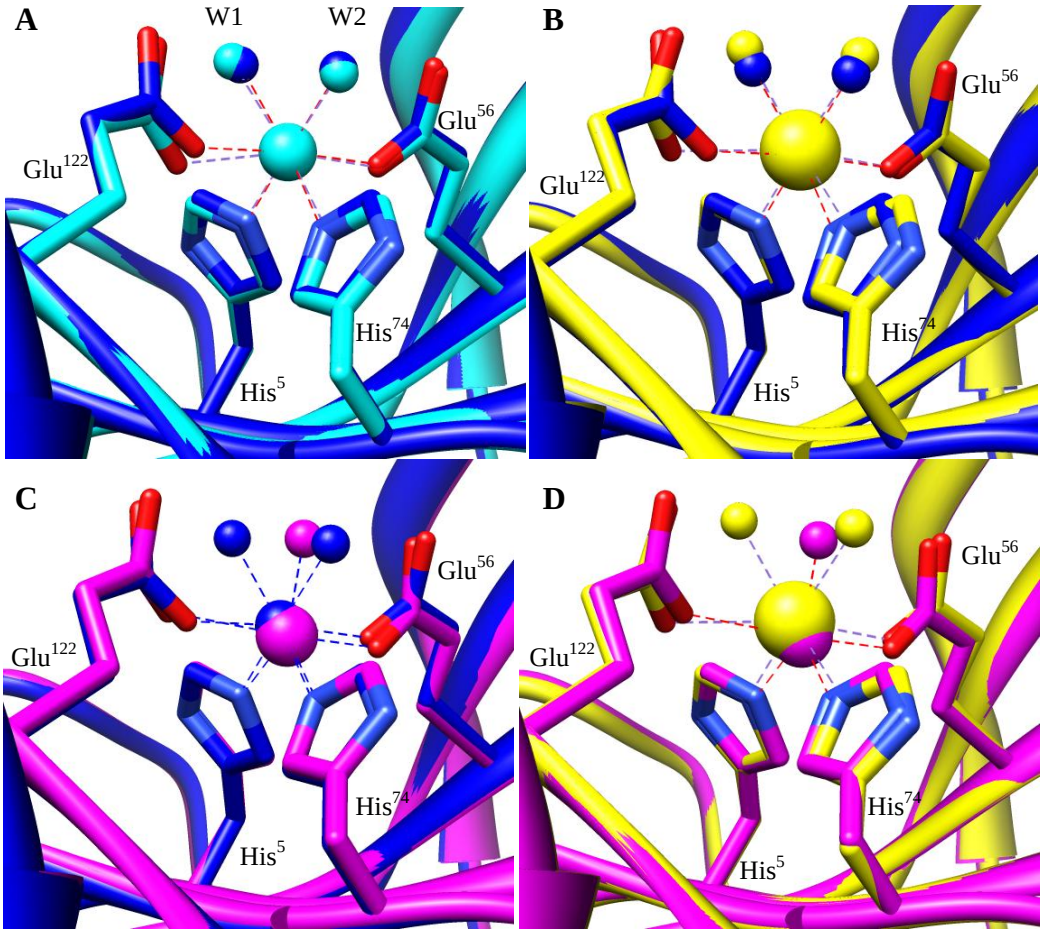


Figure 3.14: The overlapped structures of *E. coli* GlxI with different incorporated metals; **(A)** Ni^{2+}-bound GlxI (blue, PDB: 1F9Z) and Co^{2+}-bound GlxI (cyan, PDB: 1FA6), **(B)** Ni^{2+}-bound GlxI (blue) and Cd^{2+}-bound GlxI (yellow, PDB: 1FA7), **(C)** Ni^{2+}-bound GlxI (blue) and Zn^{2+}-bound GlxI (magenta, PDB: 1FA5) and **(D)** Cd^{2+}-bound GlxI (yellow) and Zn^{2+}-bound GlxI (magenta). The 3D structures were generated by using the UCSF Chimera program (University of California, San Francisco) [1].

Table 3.7: The coordination distance of the metal center and the metal binding ligands in metal-bound *E. coli* GlxI [45].

	Ligand distance (Å)									
	Ni^{2+}		***Co^{2+}***		***Cd^{2+}***		***Zn^{2+}***		****Zn^{2+}***	
Ligands	***Site 1***	***Site 2***	***Site 1***	***Site 2***	***Site 1***	***Site 2***	***Site 1***	***Site 2***	***Site 1***	***Site 2***
$His5N_{\varepsilon 2}$	2.2	2.1	2.3	2.2	2.4	2.3	2.2	2.1	2.1	2.1
$Glu56O_{\varepsilon 1}$	2.1	2.1	2.1	2.1	2.4	2.3	2.1	2.1	2.1	2.1
$His74N_{\varepsilon 2}$	2.3	2.3	2.4	2.3	2.6	2.5	2.2	2.2	2.1	2.1
$Glu122O_{\varepsilon 1}$	2.1	2.1	2.2	2.4	2.8	2.4	2.1	2.4	2.2	2.5
W1	2.1	2.1	2.3	2.2	2.4	2.4	2.0	2.1	2.2	2.1
W2	2.2	2.2	2.4	2.4	2.4	2.6			2.5	2.5

* The liganded distances were measured based on the superimposed structures of Ni^{2+}- and Zn^{2+}-bound *E. coli* GlxI using the two water molecules in the Ni^{2+}-bound enzyme and the metal binding ligands from the Zn^{2+}-bound form.

Table 3.9: The bond angles of the water molecules and the metal binding ligands, Glu^{122} and Glu^{56}, in metal-bound *E. coli* GlxI based on the previous report by He *et al*., 2000 [45].

	Bond angles for octahedral geometry (°)								
		Ni^{2+}		***Co^{2+}***		***Cd^{2+}***		****Zn^{2+}***	
Ligands	***Ideal***	***Site 1***	***Site 2***	***Site 1***	***Site 2***	***Site 1***	***Site 2***	***Site 1***	***Site 2***
$Glu122O_{\varepsilon1}$-Me- $His74N_{\varepsilon2}$	90	81	82	84	90	87	90	96	91
$His5N_{\varepsilon2}$-Me- $Glu122O_{\varepsilon1}$	90	99	95	91	79	79	95	94	89
$Glu56O_{\varepsilon1}$-Me- $His5N_{\varepsilon2}$	90	79	82	81	85	85	84	78	86
W1-Me- $His5N_{\varepsilon2}$	90	88	88	85	87	80	81	91	90
W1-Me- $Glu122O_{\varepsilon1}$	90	93	96	103	108	104	100	90	91
$Glu122O_{\varepsilon1}$-Me- $Glu56O_{\varepsilon1}$	180	179	176	172	163	164	179	164	171
W1-Me- $His5N_{\varepsilon2}$	180	163	165	161	163	161	161	150	161
$His74N_{\varepsilon2}$-Me-W1	90	91	92	90	91	92	89	93	90
$His74N_{\varepsilon2}$-Me- $Glu56O_{\varepsilon1}$	90	99	96	98	96	93	91	100	98
$His74N_{\varepsilon2}$-Me- $His5N_{\varepsilon2}$	90	100	99	105	104	102	102	116	109
W2-Me- $His5N_{\varepsilon2}$	90	92	93	93	90	93	95	85	87
W2-Me- $Glu56O_{\varepsilon1}$	90	97	98	101	101	104	103	98	102
W1-Me-W2	90	79	79	78	77	78	78	69	75
W2-Me- $Glu122O_{\varepsilon1}$	90	83	84	79	77	79	76	67	70
W2-Me- $His74N_{\varepsilon2}$	180	161	162	155	158	157	160	155	155

* The ligand distances were measured based on the superimposed structures of Ni^{2+}- and Zn^{2+}-bound *E. coli* GlxI (PDB: 1F9Z and 1FA5, respectively) by using the two water molecules in the Ni^{2+}-bound enzyme and the metal binding ligands from its Zn^{2+}-bound form.

	Bond angles for the trigonal bipyramidal geometry (°)		
		Zn^{2+}	
Ligands	***Ideal***	***Site 1***	***Site 2***
$Glu122O_{\varepsilon1}$-Me- $His74N_{\varepsilon2}$	90	90	92
$His5N_{\varepsilon2}$-Me- $Glu122O_{\varepsilon1}$	90	88	88
$Glu56O_{\varepsilon1}$-Me- $His5N_{\varepsilon2}$	90	83	88
W1-Me- $His5N_{\varepsilon2}$	90	87	93
W1-Me- $Glu122O_{\varepsilon1}$	90	95	86
$Glu122O_{\varepsilon1}$-Me- $Glu56O_{\varepsilon1}$	180	169	173
W1-Me- $His5N_{\varepsilon2}$	120	143	131
$His74N_{\varepsilon2}$-Me-W1	120	101	112
$His74N_{\varepsilon2}$-Me- $Glu56O_{\varepsilon1}$	90	99	98
$His74N_{\varepsilon2}$-Me- $His5N_{\varepsilon2}$	120	116	116

Human GlxI, on the other hand, is in the Zn^{2+}-activation class [50]. The analyses of optical and electron paramagnetic resonance (EPR) spectra as well as X-ray absorption studies on human GlxI with active metal suggested the octahedral metal coordination [31, 44, 46].

Unfortunately, only X-ray crystallographic structures of the human enzyme with inhibitors are available. The enzyme with a bound reactive intermediate analogue of *S*-(*N*-hydroxy-*N*-*p*-iodophenylcarbamoyl)glutathione (HIPC-GSH; PDB: 1QIN) suggested an octahedral metal coordination with four metal binding residues (Gln^{33}, Glu^{99}, His^{126} and Glu^{172}) and two oxygens from HIPC-GSH around the metal center (Figure 3.15A). Analysis of the X-ray crystallographic structure generated a proposal for a possible mechanism of GlxI, where the dislocation of Glu^{172} was clearly observed with an elongated distance between $Glu^{172}O_{\varepsilon 1}$ and Zn^{2+} atom (more details in Chapter 6). Other available structures including human GlxI with a substrate-like inhibitor of benzyl-glutathione (B-GSH, PDB: 1FRO) and a product-like compound of *S*-*p*-nitrobenzyloxycarbonyl-glutathione (NBC-GSH, PDB: 1QIP) only indicated one water molecule in the active site (Figure 3.15B). The distance between this water molecule and the metal center is ~2.1 Å. The second water molecule is located ~2.8 Å away from the metal, suggesting the possible movement of water molecule upon the entrance of the substrate and the exit of the product.

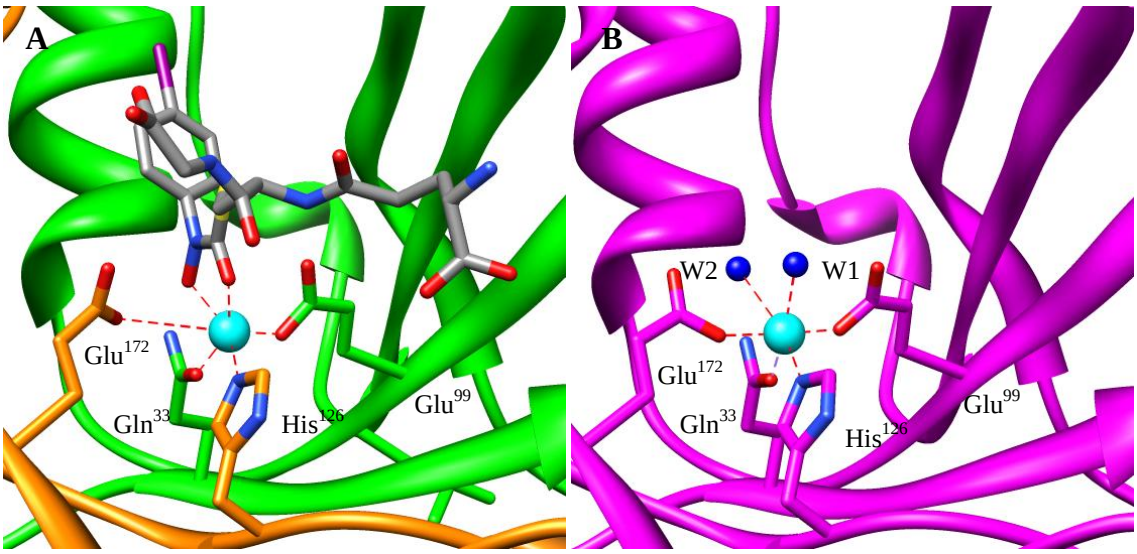


Figure 3.15: The ribbon structures of human GlxI with **(A)** bound *S*-(*N*-hydroxy-*N*-*p*-iodophenylcarbamoyl)-glutathione (HIPC-GSH; PDB: 1QIN) and **(B)** bound *S*-*p*-nitrobenzyloxycarbonyl-glutathione (NBC-GSH, PDB: 1QIP) showing the catalytic pocket with metal binding residues (Gln^{33}, Glu^{99}, His^{126} and Glu^{172}), Zn^{2+} atom (cyan), inhibitor (CPK) and water molecules (blue). The 3D structures were generated by using the UCSF Chimera program (University of California, San Francisco) [1].

Since the helix B (residues 82–94 in human GlxI) appears likely to contribute to the metal activation in GlxI, it is interesting to investigate the residues in this helix that might

interact with the metal binding residues, which in turn would create a suitable environment for a specific metal to activate the enzyme. However, it was found that only Ile^{88} and Leu^{92} are pointing toward the catalytic pocket, especially Leu^{92} that is located closed to the metal binding residues (Figure 3.2 and 3.16A). Other charged residues are facing outside and interacting with solvent. The bulky side chain of Leu^{92} might interfere with the metal binding residues, thus, creating an environment that is suitable to form the octahedral metal coordination with both Ni^{2+} and Zn^{2+} atoms.

In order to investigate the effect of helix B on GloA3, the quaternary structural prediction of GloA3 was performed with a web-based program, Geno3D: Automatic comparative molecular modeling of protein [142], using human GlxI (PDB: 1QIN) as a model. It was found that the catalytic pocket of GloA3 becomes more packed due to the presence of helix B (residues 72–84). There are three residues (Asp^{75}, Gln^{78} and Phe^{82}) in this helix that point toward the active site (Figure 3.16B). The bulky side chain of Phe^{82} (relevant to Leu^{92} in human GlxI) that lies in a close distance to the active site possibly interferes with the metal binding residues, which in turn affects metal selection of this enzyme and allow activity in the presence of Ni^{2+} and Zn^{2+}. Deletion of the helix B would eliminate this interaction and the enzyme becomes inactive with Zn^{2+} atom.

It is interesting to further investigate how the helix B affects metal selection in the GlxI. One possibility is the close interaction between the bulky residue in the helix B and the metal binding residues that might initiate metal preference for the enzyme. The presence of helix B also increases hydrophobicity of the catalytic pocket, which in turn emphasizes the polar interaction between metal binding ligands and the active metal [44, 46]. It was also found that the increase in hydrophobicity of the catalytic pocket elevates the enzymatic activity (more details in Chapter 6). Thus, the effect of hydrophobicity on the metal selection as well as other possible factors including the catalytic space and rotation of the metal binding residues should also be taken into consideration.

Nevertheless, these are hypothetical interpretations of metal activation in GloA3 and its variants. Further experiments to pursue the reasons behind these possibilities and to answer these questions should be conducted. Besides, the actual X-ray crystallographic structures of GloA2 and GloA3 with different bound metals that are being investigated by collaboration with Dr. Charlie Bond from University of Western Australia (WA, Australia) are necessary as

part of the experimental support of our hypothesis. Thus, until then it is suggested that the hydrophobicity of bulky residue(s) on the helix B affect metal selection and activation of GlxI.

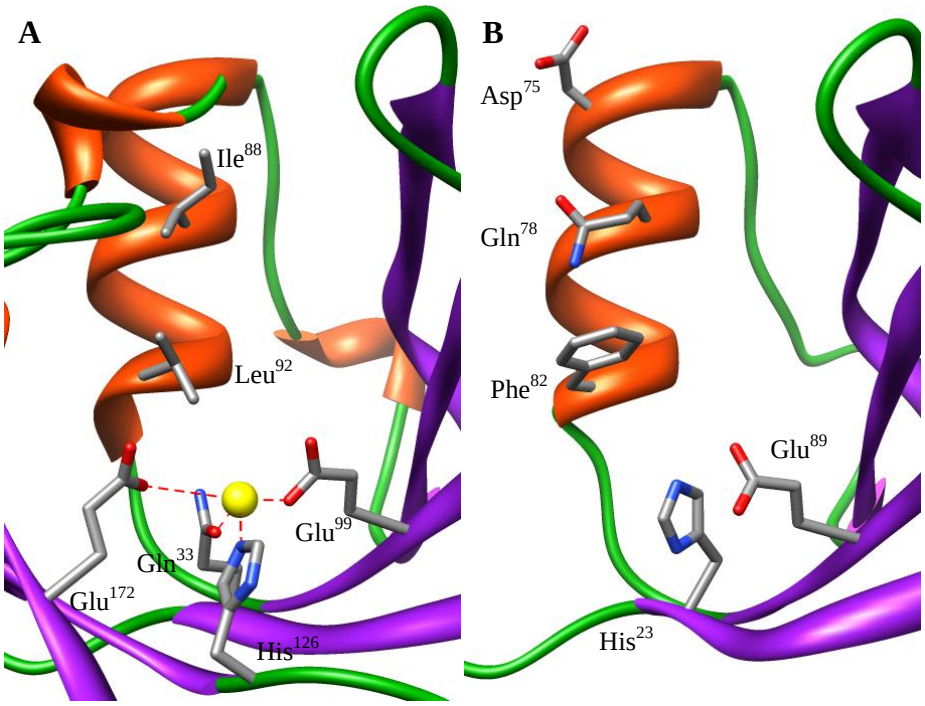


Figure 3.16: The X-ray crystallographic structure of **(A)** human GlxI (PDB 1QIN) showing the two residues, Ile^{88} and Leu^{92}, in the helix that only exist in Zn^{2+}-activation class, which might be responsible for metal activation of GlxI. **(B)** Analysis of the predicted structure of GloA3 by Geno3D web-based program [142] suggested the bulky residues of Asp^{75}, Gln^{78} and Phe^{82} in the helix B that might increase hydrophobicity of the catalytic pocket and limit the space of the active site. The 3D structures were generated by using the UCSF Chimera program (University of California, San Francisco) [1].

3.5: Conclusions

The GlxI from *P. aeruginosa*, GloA2 and GloA3, were expressed in the presence of $NiCl_2$ and $ZnCl_2$, respectively. The enzymatic activities of both enzymes were examined under select conditions. GloA3 is a Zn^{2+}-activated enzyme with activity being observed in the presence of both Zn^{2+} and Ni^{2+} atoms. On the other hand, GloA2 is subject to Ni^{2+}/Co^{2+}-activation, but no activity is found with Zn^{2+}. The initial structural investigations of both proteins used circular dichroism for the secondary structures and gel permeation chromatography for the quaternary structures. Both proteins contained similar secondary structural contents as what had been reported previously on GlxI from *P. putida* and *H. sapiens*

[30, 44]. Both GloA2 and GloA3 are homodimeric with possibly two active sites. However, the metal titration profiles suggested that only one active site is functional.

We were also interested in an investigation of the factors that influence metal activation of GlxI. Analysis of the multiple sequence alignments of both Zn^{2+}-activated GlxI and Ni^{2+}/Co^{2+}-activated enzymes from various organisms suggest that two extra helices (A and B) and at least two loop regions (C and D), that only exist in the Zn^{2+}-activation (Figure 3.2), may be important in some way to metal activation of the enzyme. Helix A was previously reported to be significant for dimerization as well as protein stability [44, 46]. Thus, deletional mutagenesis of helix B, loop C and loop D on the Zn^{2+}-activated GloA3 were performed in attempts to investigate metal activation, metal titration and kinetics as well as structural analysis in comparison with the Ni^{2+}/Co^{2+}-activated GloA2. Analysis of the experimental results suggested that helix B is responsible for metal activation selectivity, while loops C and D may be more critical for enzymatic activity.

The initial structural investigations using circular dichroism and gel permeation chromatography on GloA3 and its variants were analyzed and indicate that the deletions of helix B, loop C and loop D do not affect the quaternary structure of the enzymes, in which all muteins appear to maintain their dimeric structures. The secondary structures of the muteins, in which the B segment was deleted, were similar to that of wild-type GloA3. In contrast, muteins with deletions of segments C or D showed a change in the secondary structure, such that the content of β-sheet was reduced. This change might cause the activity loss in these variants. Besides, metals have no effect on the secondary and quaternary structures of any of the proteins.

Analyses of the X-ray structures of *E. coli* GlxI (Ni^{2+}/Co^{2+}-activation) with different bound metals suggested that the smaller size of Ni^{2+} and Co^{2+} atoms might promote their octahedral coordination in constrast with the larger Zn^{2+} atom in the molecular structure of the *E. coli* enzyme. Even though it possesses a larger ionic radius, Zn^{2+} seems to fit deeper into the active site of the enzyme. This might limit the space for water molecules and hence produce a more suitable environment for trigonal bipyramidal metal coordination. Additionally, analysis of the X-ray crystallographic structure of human GlxI (Zn^{2+}-activation) and a predicted structure of GloA3 suggests that the residue(s) in the helix B locates in close proximity to the metal binding residues, thus possibly affecting metal selection of the GlxI.

This helix B also blocks one side of the catalytic pocket, thus creating a packed active site as well as increasing the hydrophobicity of the catalytic pocket. However, in order to investigate these possibilities, further experimental studies are required.

CHAPTER 4: INVESTIGATION OF POSSIBLE FOSFOMYCIN-RESISTANCE ACTIVITY BY AN *ESCHERICHIA COLI* GLYOXALASE I (E56A) MUTANT

4.1: Introduction

Fosfomycin (FM), (1*R*, 2*S*)-epoxypropylphosphonic acid, is a natural product that contains the rare carbon-phosphorus bond in its chemical structure [154]. Its biosynthetic pathway involves four main steps as shown in Figure 4.1 [155-158]. First, the reversible reaction of phosphoenolpyruvate (PEP) isomerization by PEP phophomutase generates phosphonopyruvate (PnPy). The reaction is favorable since the product PnPy is removed by decarboxylation by the enzyme PnPy decarboxylase. This reaction converts the acidic functional group of PnPy to an aldehyde and forms phosphonoacetaldehyde (PnAA) and CO_2 as products. PnPy decarboxylase is a member of the α-ketodecarboxylase family, which requires thiamine pyrophosphate and a divalent metal for catalysis. The product PnAA is then reduced to 2-hydroxyethylphosphonic acid (HEP) by PnAA reductase, an iron-dependent enzyme that requires NADPH as cofactor. The addition of a methyl group to HEP by HEP C-methyltransferase generates (*S*)-2-hydroxypropylphosphonic acid (HPP). This enzyme employs a B_{12}-binding domain as well as *S*-adenosylmethionine (AdoMet) and methylcobalamin (MeClb) binding sites. The cofactor AdoMet generates a radical HEP before inserting a methyl group from MeCbl into the Cα of the primary alcohol. The final product in the biosynthesis of fosfomycin (FM) is the oxidative reaction of HPP by HPP dehydrogenase or HPP epoxidase. This enzyme is an iron-dependent enzyme, which converts the secondary alcohol on HPP to the epoxide functional group in FM.

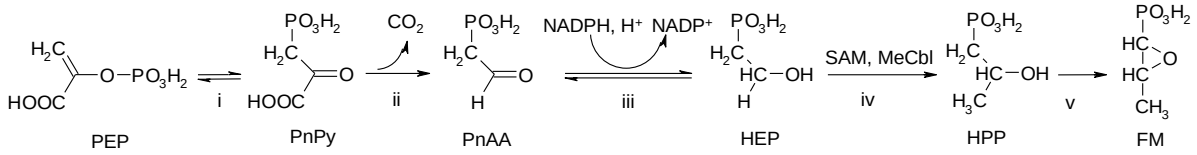


Figure 4.1: Biosynthetic pathway of fosfomycin (FM) based on several previous reports [155-158]. (i) PEP phophomutase, (ii) PnPy decraboxylase, (iii) PnAA reductase, (iv) HEP C-methyltransferase and (v) HPP epoxidase. PEP, Phosphoenolpyruvate; PnPy, phosphonopyruvate; PnAA, phosphonoacetaldehyde; HEP, 2-hydroxyethylphosphonic acid; HPP, 2-hydroxypropylphosphonic acid; SAM, *S*-adenosylmethionine; MeCbl, methylcobalamin.

Fosfomycin is an antibiotic useful in the treatment of lower urinary tract and gastrointestinal infections due to its stability and broad spectrum of activity [155-158]. This antibiotic affects both Gram-positive and Gram-negative bacteria by interfering with cell wall biosynthesis or peptidoglycan formation [158]. Since prokaryotic peptidoglycan possesses a unique structure, it has the advantage of selectively targeting a microbial pathway with little side effect on the patient (almost no toxicity in humans has been found) [159]. Peptidoglycan consists of linear glycan chains, composed of alternating units of *N*-acetylglucosamine and *N*-acetylmuramic acid, interlinked by short peptides. The first stage of peptidoglycan biosynthesis includes the formation of monomeric building blocks of *N*-acetyglucosamine-*N*-acetylmuramyl pentapeptide by a series of Mur enzymes (MurA-MurF) [159]. The first enzyme, MurA, transfers the enolpyruvate functional group of phosphoenolpyruvate (PEP) to position 3 of UDP-*N*-acetylglucosamine (UDP-Glc-NAc), thus forming UDP-*N*-acetylglucosamine enolpyruvate (UDP-Glc-NAc-EP) (Figure 4.2A). The reaction mechanism is believed to involve an addition-elimination process, which initially produces a tetrahedral ketal intermediate that is thought to be covalently stabilized by several residues from MurA [159, 160]. Among these, Cys^{115} is a significant residue that interacts with C-2 of PEP to form the *O*-phosphothioketal intermediate [161, 162]. Fosfomycin is a PEP analogue, and thus functions as an inhibitor that covalently reacts with Cys^{115} of MurA (Figure 4.2B). Upon analysis of the structure of the inhibited enzyme, it has been found that the antibiotic is tightly packed between the enzyme and UDP-Glc-NAc and forms H-bonding interactions with several enzyme residues [159, 160]. However, fosfomycin does not interact with other PEP-utilizing enzymes, which makes it an effective antimicrobial agent that interferes with the biosynthesis and assembly of peptidoglycan.

Fosfomycin is mainly produced by Streptomycetes and Pseudomonads [109, 163]. These organisms thus possess self-resistance mechanisms such that this antibiotic does not harm themselves. It was found that the introduction of *fomA* and *fomB* resistance genes that were isolated from *Streptomyces wedmorensis* and placed into and expressed by *E. coli* results in *E. coli* being capable of producing higher levels of fosfomycin resistance [164, 165]. The investigation on characterization of the FomA and FomB proteins (encoded from *fomA* and *fomB* genes, respectively) suggested that once fosfomycin is synthesized in the cytosol, it is

rapidly converted into inactive fosfomycin monophosphate by FomA in the presence of ATP and Mg^{2+}, thus preventing interruption of cell wall biosynthesis in the fosfomycin-producing organisms (Figure 4.3) [164, 165]. FomB then subsequently converts fosfomycin monophosphate into fosfomycin diphosphate, a high energy compound that can be exported outside the cell [164, 165]. Another self-resistance mechanism was discovered and found to be the *fosC* gene that codes for the FosC protein in *Pseudomonas pyringae* [166]. The enzyme FosC also catalyzes phosphorylation of fosfomycin into the inactive fosfomycin monophosphate by acquiring ATP and Mg^{2+}, the reaction that is similar to that of FomA [166]. However, their amino acid sequences possess low sequence identity (25.8%) and no *fomB* gene is found within *Pseudomonas pyringae*, suggesting that FosC and FomA are two independent fosfomycin phosphotransferase enzymes [166].

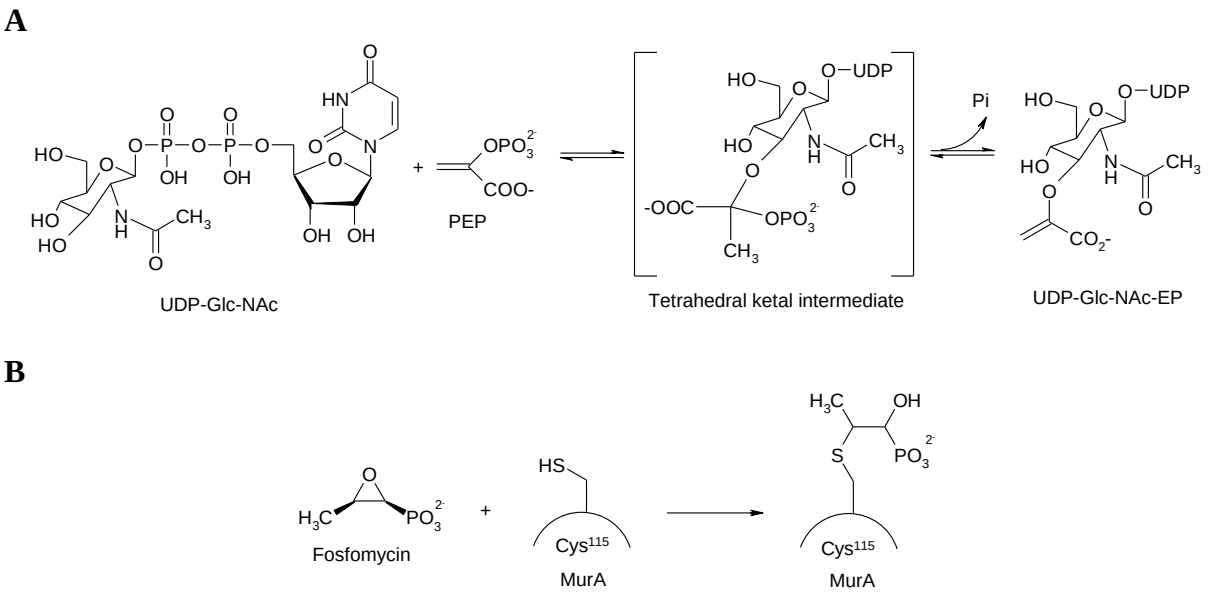


Figure 4.2: (A) The first step in prokaryotic cell wall biosynthesis by UDP-*N*-acetylglucosamine (UDP-GlcNAc) enolpyruvyltransferase (MurA) [159, 160] and **(B)** the inhibition of MurA by fosfomycin at protein residue Cys^{115} [161, 162].

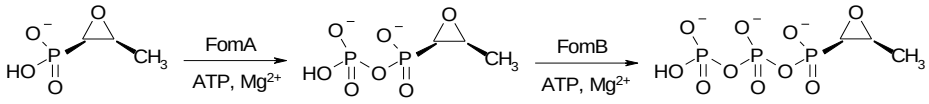


Figure 4.3: The self-resistance mechanism of FomA and FomB in fosfomycin-producing organisms.

Antibiotic resistance in bacteria is nowadays a crucial problem in the treatment of infectious disease. After providing fosfomycin treatment, bacteria evolve resistance mechanisms in an attempt to fight back the antibiotic and to survive under these stressful conditions. These defense mechanisms have been discovered in clinical plasmid isolates of fosfomycin-treated patients [110]. The mechanisms include the development of mutations in the MurA gene, the phosphonate uptake pathway, as well as evolution of resistance proteins that destroy the C-O bond of the epoxide moiety. The fosfomycin-resistant proteins including FosA, FosB and FosX are metalloenzymes that utilize different active sites for bound metal ions and cofactors (Figure 4.4). FosA catalyzes the epoxide cleavage reaction using glutathione (GSH) and Mn^{2+}/K^{+} ions as cofactors. Similarly, FosB requires another type of thiol (L-CySH) and the activating metal Mg^{2+} ion, while FosX only requires H_2O with Mn^{2+} for the reaction. The amino acid sequence of FosA is more compatible with FosB (~20% sequence identity and ~40% sequence similarity) than with FosX (~10% sequence identity and ~20% sequence similarity).

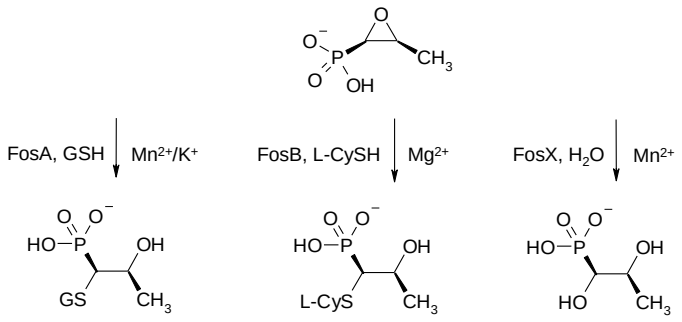


Figure 4.4: Reactions catalyzed by the various fosfomycin-resistance proteins, FosA, FosB and FosX that add GSH, L-Cys and H_2O, respectively, to fosfomycin [112, 167]. These metalloenzymes employ different activating metals.

A well-characterized FosA has been found to be a structural member of the βαβββ superfamily of proteins and is the most active among the fosfomycin-resistance proteins [112, 167]. FosA acts as a GSH transferase that transfers GSH to fosfomycin, forming an inactive 1-(*S*-glutathionyl)-2-hydroxypropylphosphonic acid (GS-fosfomycin). Its overall structure is similar to Glyoxalase I (GlxI) and is believed to share a common ancestor [41]. Both are metalloenzymes that utilize GSH as cofactor and form an active site geometry around the metal center as an octahedral coordination [44-46, 109, 168]. However, FosA employs three water

molecules and three metal binding ligands that coordinate to the metal center *via* vicinal oxygen [109, 169], while GlxI occupies two water molecules and four metal binding residues around the active metal [45]. FosA also employs two metal binding sites, one for divalent metal Mn^{2+} and another one for monovalent metal K^+ [110, 111, 170]. GlxI, on the other hand, only occupies one divalent metal binding site [44-46].

It was found that several analogues of phosphonic acid are capable of inhibiting FosA activity, including phosphonoformate (**4-1**, $K_i = 0.41 \pm 0.09$ μM), acetylphosphonate (**4-2**, $K_i = 18 \pm 2$ μM), phosphonoacetate (**4-3**, $K_i = 10 \pm 1$ μM) and 2-carboxyethylphosphonate (**4-4**, K_i >100 μM) (Figure 4.5) [171]. The most effective compound **4-1** with the lowest K_i has been viewed as a minimal transition state analogue. The inhibitor is believed to form a tetrahedral geometry with the Mn(II) centre with its carboxylate oxygen lying almost in the same position as the oxirane oxygen of the substrate [171]. This compound, however, only binds approximately six-fold tighter to the enzyme than the substrate, which suggested that the addition of the GSH group might play a significant role in the inhibitor-enzyme interaction [171]. Expanding the chelate ring by methylene groups of **4-2**, **4-3** and **4-4** seems to decrease binding affinity of the inhibitor to the enzyme. All these inhibitors, however, are not effective toward enhancement of fosfomycin bioactivity *in vitro* (liquid culture and agar plate) so bacterial resistance to fosfomycin was still observed upon addition of inhibitor under the studied conditions [171]. It was suggested that drug transport across the bacterial membrane might be a major contributor to the lack of observable activity of these inhibitors [171]. Additional knowledge on drug delivery of these compounds and possibly creation of prodrug versions of these compounds might allow them to survive bacterial degradation and further investigation in this area appears warranted.

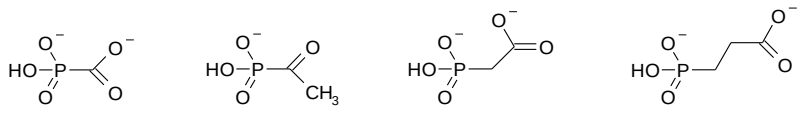


Figure 4.5: The fosfomycin analogues that act as FosA inhibitors including phosphonoformate (**4-1**), acetylphosphonate (**4-2**), phosphonoacetate (**4-3**) and 2-carboxyethylphosphonate (**4-4**) [171].

Recently, FosA has been detected in the *Pseudomonas aeruginosa* genome and shares 60% sequence identity with the plasmid encoded protein from transposon Tn2921 [172]. From

this information, it is possible that genomically encoded FosA might have evolved from a protein with similar folding that already existed in an ancestral organism. Glyoxalase I with its high structural similarity might be a possible candidate, which might have been a tetradentated progenitor of the tridentated FosA that evolved to develop FosA-like activity by a process of gene modifications to accomplish these more complicated reactions (two-substrate reaction) and therefore providing biological functional variety. In an attempt to investigate the relatedness between these two proteins, the deletion of one metal binding ligand in GlxI was performed according to the multiple sequence alignment in order to prepare the tridentated enzyme with a similar active site coordination around the metal ion as that of FosA (Figure 4.6). The GlxI from the Gram-negative bacteria *Escherichia coli* was chosen as the model of study due to its thorough characterization in our laboratory and its available X-ray crystallographic structure, which was determined by our group in collaboration with the laboratory of Dr. Brian Matthews [45]. Even though *E. coli* GlxI and *Pseudomonas aeruginosa* FosA share high overall structural similarity, they only show 10% sequence identity and 16% sequence similarity. The mutagenesis of *E. coli* GlxI-E56A was performed by polymerase chain reaction (PCR). The protein was expressed in *E. coli* and purified as previously reported [19]. The GlxI and FosA activities of the mutant were examined in comparison to the reported kinetics on the wild-type enzymes under similar conditions. The results on enzymatic activity as well as structural comparison of GlxI and FosA will be discussed in terms of possible evolution of function. Since an inhibitor for both FosA and MurA has yet to be investigated, studying and understanding the FosA system might lead to new knowledge on the evolution of bacterial drug resistance mechanisms.

Figure 4.6: The multiple sequence alignment of Glyoxalase I from various organisms (organism name follows by National Center for Biotechnology Information (NCBI) accession number) including *E. coli* (NP_310387), *Y. pestis* (ZP_01887743), *N. meningitidis* (CAA74673), *P. aeruginosa* (AAG06912), *L. donovani* (AAU87880), *T. cruzi* (XP_818456), *H. sapiens* (AAB49495), *P. putida* (AAN69360) with *Serratia marcescens* FosA (Q56415). The metal binding residues on Glyoxalase I and FosA are marked with a grey highlight. The alignment was created using CLC Free Workbench (version 3.0.1) with the accurate alignment algorithm (http://www.clcbio.com).

4.2: Reagents, Materials and Instrumentation

All reagents, materials and equipment used in this chapter are listed below otherwise they are included in the previous chapters. Fosfomycin disodium salt was received from Sigma Aldrich (St. Louis, MO). The synthesis of GS-fosfomycin adduct was performed on a Biotage Microwave Synthesizer (Uppsala, Sweden), while its purification and FosA assay were achieved using a column containing Sephadex G-10 resin (GE Healthcare, Piscataway, NJ). The identification of organic compounds and purified protein were performed using an electrospray ionization mass spectrometer (ESI-MS) with a Micromass Q-TOF Ultima™ Global supplied by the Waterloo Chemical Analysis Facility, University of Waterloo.

4.3: Experimental Protocols

4.3.1: Protein Expression, Induction and Purification

E. coli Glyoxalase I mutant, GlxI-E65A, was generated by Kadia Mullings, a previous Masters student in our laboratory, by utilization of QuikChange® Site-Directed Mutagenesis PCR on the plasmid containing wild type *E. coli glxI* (pET22b-*glxI*). The desired plasmid was heat shock transformed into *E. coli* BL21 (DE3) for purification purposes. *E. coli* cells were then inoculated overnight in LB (10 mL) containing ampicillin (100 μg/mL) in a 37 ºC shaker (220 rpm). The culture was transferred into 1 L LB and allowed to grow in a 37 ºC shaker (220 rpm) until an OD_{600} of 0.6 was attained. Protein production was induced by the addition of IPTG to a concentration of 1 mM at 25 °C for 4 hours. The cell pellet was harvested by centrifugation at 6,000×*g* for 10 min before flash freezing in liquid nitrogen and storing at -80 ºC.

The cell pellet was resuspended in 20 mM Tris buffer (pH 7.5) containing 10% (v/v) glycerol and 1 mM PMSF before performing sonication to disrupt the cell walls. Cell lysate was centrifuged at 48,384×*g* for 20 min and the supernatant was loaded onto an UNOsphere Q cartridge (1 mL). The proteins were collected utilizing a gradient of 0–1 M KCl in 20 mM Tris buffer (pH 7.5) and 10% (v/v) glycerol with a 1 mL/min flow rate over a 100 min interval. The fractions containing desired protein (based on molecular weight) were identified by SDS-PAGE and dialyzed against buffer containing 20 mM Tris buffer (pH 7.5) and 150 mM NaCl overnight at 4 °C. The dialyzed sample was further purified by gel permeation chromatography (Superdex™ 75 HR 10/30 column) using an eluting buffer solution of 20 mM Tris (pH 7.5) and 150 mM NaCl with a 0.5 mL/min flow rate. The presence of purified enzyme was confirmed by the analyses of SDS-PAGE and ESI-MS experiments. The apo-enzyme was prepared as previously described in Chapter 2. The protein concentration was determined by Bradford assay using BSA as a standard [134].

4.3.2: Initial Structural Investigation

The initial structural investigation of *E. coli* GlxI-E56A was performed using gel permeation chromatography and circular dichroism (CD). The former was to examine the quaternary structure and to aid in the calculation of the molecular weight of the protein under native conditions using a standard curve as previously described for that of CLO GlxI

(Appendix 3). Circular dichroism analysis was used to investigate the secondary structure of the protein and was performed using a Jasco J-715 spectropolarimeter with the parameters as previously described in Chapter 2. The samples were prepared by incubating the apo-enzyme with 10 equivalents of metals in 20 mM Tris (pH 7.5) and 150 mM NaCl to a final concentration of 0.3 mg/mL.

As well, the predicted structure of *E. coli* GlxI-E56A based on the structure of wild type *E. coli* GlxI (PDB: 1FA8) was determined using the web based software, Geno3D Automatic comparative molecular modeling of protein (http://geno3d-pbil.ibcp.fr) [142]. Structural visualizations and molecular images were generated by the UCSF Chimera package from the Resource for Biocomputing, Visualization and Informatics at the University of California, San Francisco (supported by NIH P41 RR-01081) [1].

4.3.3: Synthesis of Fosfomycin-glutathione Adduct

Fosfomycin (250 mM) and glutathione (250 mM) were mixed in 1 mL Milli-Q water (adjusted pH to 8.0) that had been degassed under argon for 30 min prior to addition of the reactants. The sample was submitted to reaction in a Biotage Microwave Synthesizer under the set-up conditions of very high absorption level, 2 min pre-stirring time, 150 °C reaction temperature, 10 min reaction time using a vial type of 0.5–2 mL. The reaction mixture was then filtered and loaded onto a Sephadex G-10 26/40 column. The eluted fraction was collected every 1 mL using Milli-Q water with gravity flow rate as the eluent. The product was detected at 240 nm using a 96-well UV plate on a plate reader. The identification of the purified compound was performed using ESI-MS analysis. Dr. Yanhong Ran, a visiting scientist from Jinan University (Guangdong, China), helped with this protocol.

4.3.4: Enzymatic Assay

Glyoxalase I Assay

The assay was performed as described previously in Chapter 2 using the hemithioacetal, the product from the non-enzymatic reaction of methyglyoxal and glutathione, as substrate in 50 mM KPB (pH 6.6) at room temperature. The apo-enzyme (0.01 mg/mL) in metal-free Tris buffer (20 mM Tris (pH 7.5) and 150 mM NaCl, preparation was described in Chapter 2) was incubated with different divalent metals (5 equivalents of $NiCl_2$, $CoCl_2$, $CuCl_2$,

$ZnCl_2$, $MnCl_2$, $CdCl_2$, $MgCl_2$ and $CaCl_2$) overnight at 4 °C before examining its metal specificity. Additionally, Ni^{2+}- and Co^{2+}-titration profiles were performed similarly by titrating the metals (0–20 equivalents) to the apo-enzyme (0.01 mg/mL). For kinetics, Ni^{2+}-reconstituted enzyme (0.15–0.35 μg) and Co^{2+}-reconstituted enzyme (0.25–0.50 μg) were incubated with 0.02–0.70 mM substrate.

FosA Activity Assay

The assay for FosA function was performed as previously reported using Ellman's reagent (5, 5'-dithiobis-(2-nitrobenzoic acid) or DTNB) [109]. Fosfomycin (10 mM) was mixed with glutathione (15 mM) in buffer containing 50 mM HEPES (pH 8.0) and 100 mM KCl that had been degassed under argon for 30 min prior to the reaction. The metal reconstituted GlxI-E56A (4 μM) that was prepared by incubating the apo-enzyme with metal chloride (50 μM) overnight at 4 °C was added to the GSH-fosfomycin mixture (final volume of 100 μL) before incubating at room temperature for 3 hours. The reaction was then quenched by an addition of methanol (300 μL). The free thiol of GSH was quantitated by mixing 5 μL of the quenched reaction, 10 μL of 18 mM DTNB solution and 185 μL HEPES buffer before monitoring at 412 nm using a plate reader (ε_{412} = 14150 $M^{-1}cm^{-1}$). The metal activation was investigated using reconstituted mutant with $NiCl_2$, $CoCl_2$, $CuCl_2$, $ZnCl_2$, $MnCl_2$, $CdCl_2$, $MgCl_2$ and $CaCl_2$. The control experiments were performed in parallel without the addition of the enzyme.

An alternative assay was performed using the analysis of ESI-MS experiments. The apo-GlxI-E56A was incubated with 10 equivalents of metals ($NiCl_2$ and $MnCl_2$) overnight at 4 °C. Metal-reconstituted GlxI-E56A (14.4 μM) was mixed with fosfomycin (150 mM) and GSH (150 mM) in 150 mM sodium borate (pH 8.0) to a final volume of 373 μL and incubated at 30 °C for 3 hours. The reaction was then quenched by the addition of 500 μL of methanol and fractionated on a Sephadex G-10 26/40 column with a gravity flow rate. Products were isolated for analysis by collection of every 1 mL fraction and monitored at 240 nm using the plate reader. The purified GS-fosfomycin adduct was identified by ESI-MS. Another attempt to achieve the FosA activity on the GlxI mutant was to change the reaction buffer to HEPES (pH 8.0) with additional 100 mM KCl present. The Ni^{2+}-reconstituted mutant (14.4 μM) was

added into the GSH-fosfomycin mixture and the assay was carried out under similar conditions as before. The control experiments were performed in parallel without the addition of the enzyme.

4.4: Results and Discussion

4.4.1: Protein Purification and Characterization

The gene for *E. coli* mutant, *glxI*-E56A, was cloned into the *E. coli* BL21 (DE3) expression system. The mutant consists of 135 amino acids with a predicted molecular weight of 14861.8 Da and a predicted p*I* of 5.06 (Figure 4.7). The proteins were purified using a two-step purification protocol including anion exchange chromatography using a UNOsphere Q cartridge (Figure 4.8) and gel permeation chromatography using a Superdex75 10/30 column (Figure 4.9). The purified protein of approximately 23 mg was recovered from 5 g cell pellet that was harvested from 1 L cell culture. The SDS-PAGE and ESI-MS analyses confirmed the identity of the purified enzyme (Figure 4.10). The molecular weight that was estimated from analysis of the SDS-PAGE was approximately 15 kDa, while ESI-MS gave a more precise molecular weight of 14861.5 Da (compared to the theoretical calculated molecular mass of 14861.8 Da). No post-translational modification of the purified mutein was observed.

```
atgcgtcttcttcataccatgctgcgcgttggcgatttgcaacgctccatcgatttttat
 M  R  L  L  H  T  M  L  R  V  G  D  L  Q  R  S  I  D  F  Y
accaaagtgctgggcatgaaactgctgcgtaccagcgaaaacccggaatacaaatactca
 T  K  V  L  G  M  K  L  L  R  T  S  E  N  P  E  Y  K  Y  S
ctggcgtttgttggctacggcccggaaaccgaagaagcggtgattgccctgacctacaac
 L  A  F  V  G  Y  G  P  E  T  E  E  A  V  I  A  L  T  Y  N
tggggcgtggataaatacgaactcggcactgcttatggtcacatcgcgcttagcgtagat
 W  G  V  D  K  Y  E  L  G  T  A  Y  G  H  I  A  L  S  V  D
aacgccgctgaagcgtgcgaaaaaatccgtcaaaacgggggtaacgtgacccgtgaagcg
 N  A  A  E  A  C  E  K  I  R  Q  N  G  G  N  V  T  R  E  A
ggtccggtaaaaggcggtactacggttatcgcgtttgtggaagatccggacggttacaaa
 G  P  V  K  G  G  T  T  V  I  A  F  V  E  D  P  D  G  Y  K
attgagttaatcgaagagaaagacgccggtcgcggtctgggcaactaa
 I  E  L  I  E  E  K  D  A  G  R  G  L  G  N  -

 Amino acid composition:
 11A      7R    6N    7D    1C    2Q    13E   16G   2H    7I
 13L      8K    3M    3F    4P    4S    9T    1W    8Y    10V
```

Figure 4.7: The DNA sequence (405 nucleic acids), amino acid sequence (135 residues) and amino acid composition of *E. coli* GlxI-E56A showing the mutation of the glutamate at position 56 to alanine (grey highlight).

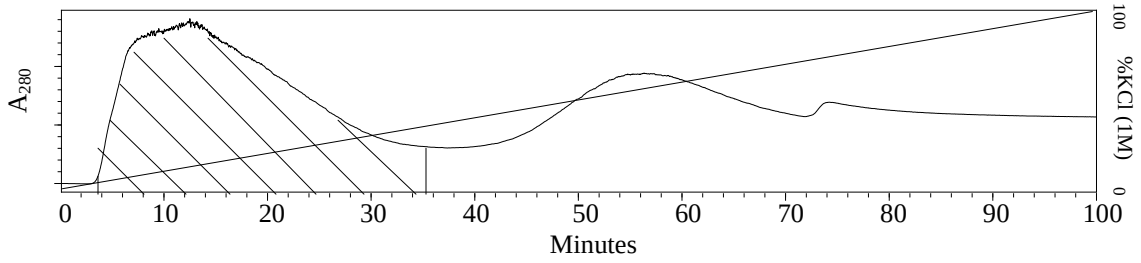


Figure 4.8: The HPLC chromatogram of *E. coli* GlxI-E56A purification by anion exchange chromatography using a UNOsphere Q cartridge and a gradient buffer of 0–1 M KCl in 20 mM Tris buffer (pH 7.5) containing 10 mM DTT and 10% (v/v) glycerol over a 100 min interval with a flow rate of 1 mL/min. Fractions containing GlxI-E56A were identified by SDS–PAGE analysis, which corresponded to 5–35% KCl (shade area).

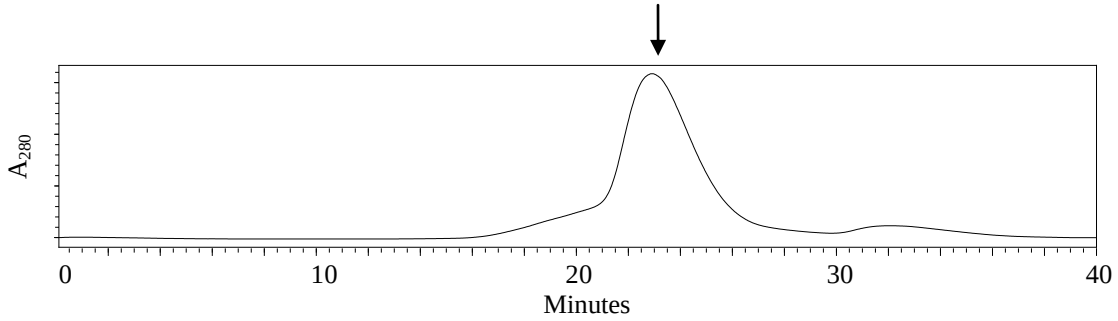


Figure 4.9: The HPLC chromatogram of *E. coli* GlxI-E56A purification by gel permeation chromatography using a Superdex75 HR 10/30 column with 20 mM Tris (pH7.5) and 150 mM NaCl and a flow rate of 0.5 mL/min. The arrow indicates the eluted peak of the mutein, which was identified by SDS–PAGE analysis.

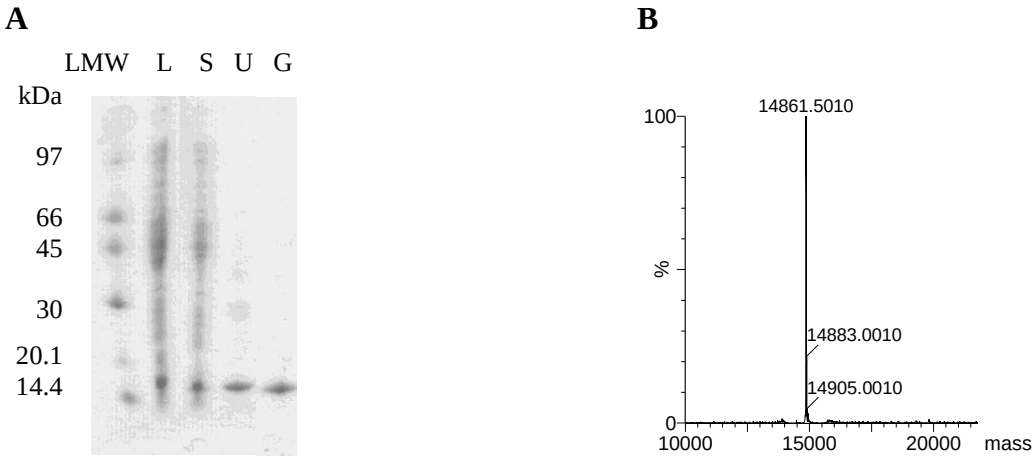


Figure 4.10: (A) The SDS-PAGE of the purification of *E. coli* GlxI-E56A: LMW = low molecular weight marker, L = lysate, S = supernatant after sonication, U = UNOsphere Q fraction and G = gel permeation fraction. The arrow indicates the presence of GlxI-E56A. **(B)** The electrospray mass spectrum of *E. coli* GlxI-E56A shows the molecular weight of a single subunit at 14861.5 Da (calculated MW is 14861.8 Da).

4.4.2: Glyoxalase I Function

It was found that the mutein GlxI-E56A enzyme had the same metal specificity activation of its Glyoxalase I activity as wild type GlxI has, where activity was found with the Ni^{2+}- and Co^{2+}-reconstituted enzyme (Figure 4.11). The maximum activity was achieved with Ni^{2+}-reconstitution and approximately 60% activity was observed with Co^{2+}-reconstitution. The enzyme reconstitution with other divalent metals such as Mn^{2+} and Cd^{2+} only produced trace GlxI activities. Analyses of metal titration profiles indicated that the enzyme needed at least 10 equivalents of metal to reach its maximum activity (Figure 4.12), in contrast to wild type *E. coli* GlxI, in which maximum activity was obtained with only 1 mole of metal ion per mole of dimeric enzyme [19]. By comparing the enzymatic activity of Ni^{2+}- and Co^{2+}-reconstituted GlxI-E56A, it was found that the addition of 5 equivalents of metal ion slightly differentiated their activities (V_{max} of 19.6 ± 2.6 μmol/min/mg for Ni^{2+}-reconstituted GlxI-E56A and 13.6 ± 0.7 μmol/min/mg for Co^{2+}-reconstitution) (Table 4.1). However, GlxI-E56A with 10 equivalents of metal presented in the reconstitution experiments produced a greater effect on the activity, which can be clearly observed between Ni^{2+}- and Co^{2+}-activated mutants (V_{max} of 23.3 ± 1.9 μmol/min/mg for Ni^{2+}-reconstitution and 14.7 ± 1.4 μmol/min/mg for Co^{2+}-reconstitution) (Table 4.1). The activity of Ni^{2+}-reconstituted GlxI-E65A was decreased to only 3.4% of that of wild type *E. coli* GlxI (V_{max} of 676 ± 17 μmol/min/mg) [19]. These results are likely explained by the lack of one of the metal binding ligands by the mutation that was created.

The results observed with *E. coli* GlxI-E56A corresponded to a previous report on mutagenesis of human GlxI (hGlxI-E99N) and *Pseudomonas putida* GlxI (pGlxI-E93D) (Glu^{99} in human GlxI and Glu^{93} in *P. putida* GlxI correspond to Glu^{56} in *E. coli* GlxI according to multiple sequence alignment) [50, 173]. It was found that the mutagenesis of Glu^{99} in human GlxI decreased the activity of the enzyme to only 1.5% compared to its wild type counterpart [50]. As well, the activity of the mutant pGlxI-E93D from *P. putida* was decreased to <0.02% compared to its wild type enzyme [173]. Metal analysis of hGlxI-E99N suggested that the mutein only contained 0.3 mol zinc per subunit enzyme [50]. Compared to the mutagenesis of Glu^{172} in the human GlxI (corresponding to Glu^{122} in *E. coli* GlxI), the activity of this mutant was undetectable (less than 10^{-8} of that of wild type enzyme) [50]. Interestingly, the metal analysis of this mutein indicated that the metal content remained as 1 mol of zinc per mol

enzyme subunit [50]. These results suggested that Glu^{172} in human GlxI is important for the enzymatic GlxI mechanism as it has been proposed to serve, as the general base involved in proton abstraction and transfer from the substrate hemithioacetal (more details in Chapter 6) [44, 46]. Thus, these data could suggest that Glu^{172} in human GlxI is significant for the reaction mechanism, while another metal binding ligand, Glu^{99}, is crucial for metal binding affinity as well as for the reaction mechanism since Glu^{99} is also proposed to act as a general base for proton transfer of the *R*-enantiomer of the hemithioacetal substrate [44, 46].

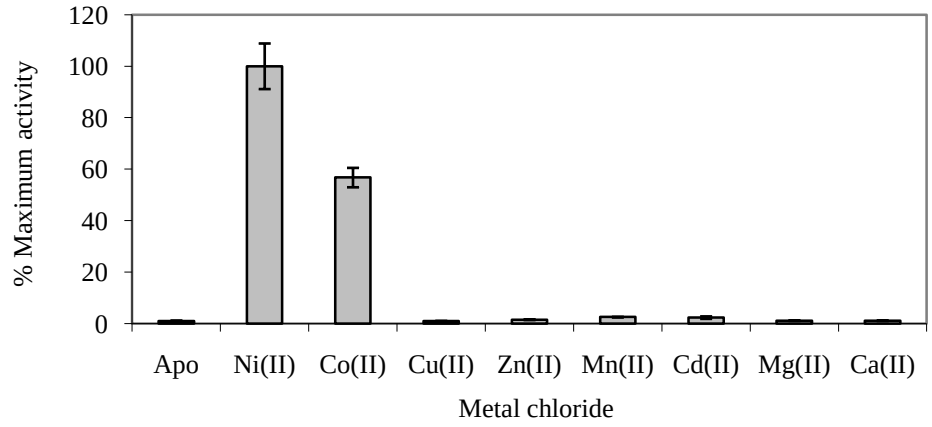


Figure 4.11: The metal activation profile of *E. coli* GlxI-E56A (0.5 μg in 200 μL assay) using 0.5 mM MG-GSH in KPB (pH 6.6) at room temperature. The specific activity of Ni^{2+}-reconstituted enzyme was 8.15 ± 0.72 μmol/min/mg.

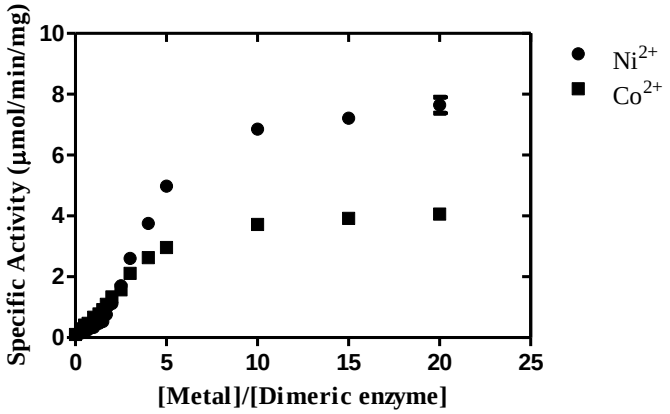


Figure 4.12: The metal titration profile ($NiCl_2$ and $CoCl_2$) of *E. coli* GlxI-E56A (0.5 μg in 200 μL assay) using 0.5 mM MG-GSH in KPB (pH 6.6) at room temperature. The maximum activity of the enzyme was achieved with 10 equivalents of reconstituted metals.

Table 4.1: Summary of enzyme kinetics of *E. coli* GlxI-E56A (0.15–0.5 μg) with 5 and 10 equivalents of reconstituted metals ($NiCl_2$ and $CoCl_2$) using 0.02–0.70 mM MG-GSH in KPB (pH 6.6) at room temperature.

	Ni^{2+}-reconstituted enzyme				***Co^{2+}-reconstituted enzyme***			
Metal equiv.	***V_{max} (μmol/min/mg)***	***K_m (mM)***	***k_{cat} (s^{-1})***	***k_{cat}/K_m ($s^{-1}M^{-1}$)***	***V_{max} (μmol/min/mg)***	***K_m (mM)***	***k_{cat} (s^{-1})***	***k_{cat}/K_m ($s^{-1}M^{-1}$)***
5	19.6 ± 2.6	0.60 ± 0.07	9.7	1.6×10^4	13.6 ± 0.7	0.85 ± 0.08	6.7	7.9×10^3
10	23.3 ± 1.9	0.68 ± 0.08	11.5	1.7×10^4	14.7 ± 1.4	0.93 ± 0.17	7.3	7.8×10^3

4.4.3: Fosfomycin Resistance Function

Although as mentioned above, the normal (Glyoxalase I) activities of GlxI from various source organisms mutated at one of the glutamate metal binding ligands were reported to be low to non-existent [37, 50, 173], to date, no group to our knowledge has ever reported a study on possible alternate activities of these mutated enzymes. Hence, our work appears to be the first that has explored possible FosA activity of a GlxI containing this key mutation. FosA transfers GSH to fosfomycin in order to form an inactive 1-(*S*-glutathionyl)-2-hydroxypropylphosphonic acid (GS-fosfomycin). This product was first synthesized and purified in attempt to use as a control in the assay reaction. The enzyme reactions using the mutein GlxI-E65A were performed by two approaches including the utilization of Ellman's (DTNB) reagent to detect the decrease of the thiol cosubstrate and the analysis of ESI-MS experiments on the formation of the purified product.

The synthesis of the GS-fosfomycin adduct was performed using a Biotage Microwave Synthesizer. Previous methods reported to make this compound listed long reaction times [111]. Although these were attempted, a method to speed up reaction times for chemical reactions is that of microwave assisted organic synthesis. This method utilizes the selective heating of a reaction solution by microwaves in a pressure vessel to reduce the overall reaction time. The reaction simply involved mixing fosfomycin and glutathione in degassed alkaline water (pH 8.0) and activating the reaction at high temperature (150 °C for 10 min). The product was isolated using a Sephadex G-10 column. The GS-fosfomycin adduct was identified by analysis of the ESI-MS, which indicated a successful synthesis of the compound. However, the purification using the Sephadex G-10 column was not efficient such that GSH was able to co-elute with the product. The mass at 444 Da suggested the presence of the GS-fosfomycin adduct (calculated mass of 444.4 Da) (data not shown). The mass of 307 Da

indicated the presence of GSH (calculated mass of 307.3 Da) and an additional mass of 384 Da indicated the likely presence of a $2K^+$ adduct of GSH (calculated mass of 385.5 Da).

The epoxide cleavage of fosfomycin and the transfer of GSH to the molecule catalyzed by the mutant *E. coli* GlxI-E56A was monitored by utilizing Ellman's (DTNB) reagent as a previous report by Beharry and Palzkill, 2005 [109]. The reaction mixture containing glutathione and fosfomycin was incubated with the mutein for 3 hours and then quenched by the addition of methanol. The unreacted thiol in the solution was then monitored by the reaction with DTNB, thus the enzyme kinetics was examined in terms of the decrease in substrate thiol level. The free thiol cleaves the disulfide bond of DTNB, yielding the products, 2-nitro-5-thiobenzoate (NTB) and the mixed disulfide (Figure 4.13). The NTB is ionized under alkaline conditions to give the NTB^{2-} dianion, which can be spectrophotometrically quantitated by monitoring the absorbance at 412 nm. However, in this experiment the absorbance of the control experiment and the quenched reaction mixture were similar, suggesting that there was either no or low fosfomycin adduct formation by the GlxI-E56A mutant under studied conditions. No divalent metal ($NiCl_2$, $CoCl_2$, $CuCl_2$, $ZnCl_2$, $MnCl_2$, $CdCl_2$, $MgCl_2$ and $CaCl_2$) could activate the mutant enzyme. It is possible that the level of unreacted free thiol was too high for proper comparison to the product levels, thus the decrease in absorbance at 412 nm was unnoticeable. Therefore, the DTNB assay might have a limitation to its use to detect low FosA activity by the mutein.

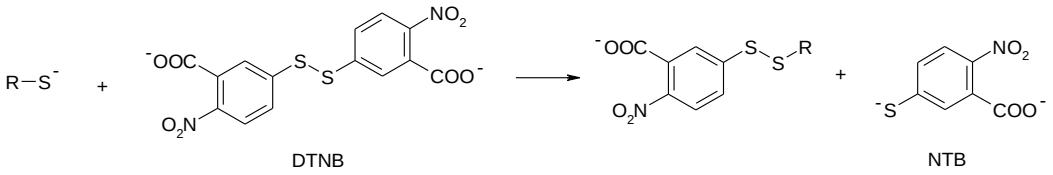


Figure 4.13: The cleavage of 5, 5'-dithiobis-(2-nitrobenzoic acid) (DTNB) by thiol giving the product of 2-nitro-5-thiobenzoate (NTB), which under alkaline conditions is ionized into the thiolate anion that can be monitored at 412 nm.

An alternative method to determine whether the mutein actually produced the adduct was to identify the GS-fosfomycin adduct by ESI-MS experiments. The product of the quenched reaction was isolated by utilizing a Sephadex G-10 column and the eluted fractions were collected every 1 mL. The product-containing fraction was then monitored at 240 nm and confirmed by ESI-MS, in which the presence of a mass of 444 Da indicated the presence

of the GS-fosfomycin adduct. Analysis of the ESI-MS spectra indicated unresolved purification using the Sephadex G-10 column, where GSH coeluted with the compound whose mass corresponded to GS-fosfomycin adduct. The detection of the product in ESI-MS of the control experiment (without addition of the enzyme) and the enzymatic reaction mixture was similar, suggesting that small amounts of the GS-fosfomycin adduct could form without addition of the enzyme (data not shown). These results also suggested that the mutein possibly possessed little or no detectable fosfomycin conjugating activity. Neither activity was detected in HEPES buffer with additional KCl nor in the sodium borate buffer. The divalent metals (Ni^{2+} and Mn^{2+}) did not appear to alter the outcome of the enzymatic reaction.

4.4.4: Structural Investigation

The secondary and quaternary structure of *E. coli*-E56A GlxI was examined using CD analysis and gel permeation chromatography, respectively. The CD spectrum of GlxI-E56A exhibited two negative maxima at 218 and 222 nm (Figure 4.14), indicating the present of 32% α-helix, 16% β-sheet and 52% random coil (estimated by K2D program calculation of protein secondary structure [152, 153]). This CD spectrum is similar to that of reported wild type *E. coli* GlxI, suggesting that the mutation of E56A did not affect the native protein secondary structure [174]. The quaternary structure investigated by the gel permeation chromatography showed that the enzyme remains in its dimeric arrangement with an estimated molecular weight of approximately 27 kDa (calculated molecular weight of a dimer is 29.7 kDa), suggesting that the mutation E56A did not alter the protein's quaternary structure (Figure 4.15).

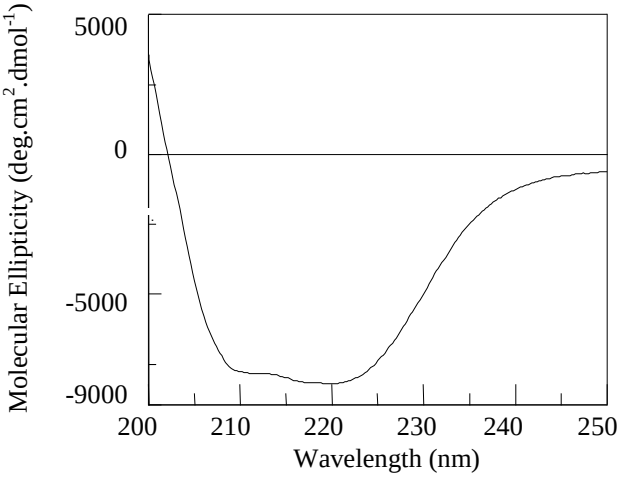


Figure 4.14: The CD spectrum of *E. coli* GlxI-E56A (0.1 mg/mL) in 20 mM Tris (pH 7.0) and 150 mM KCl scanning between 200–250 nm and analysis indicates a mixture of α-helix and β-sheet content.

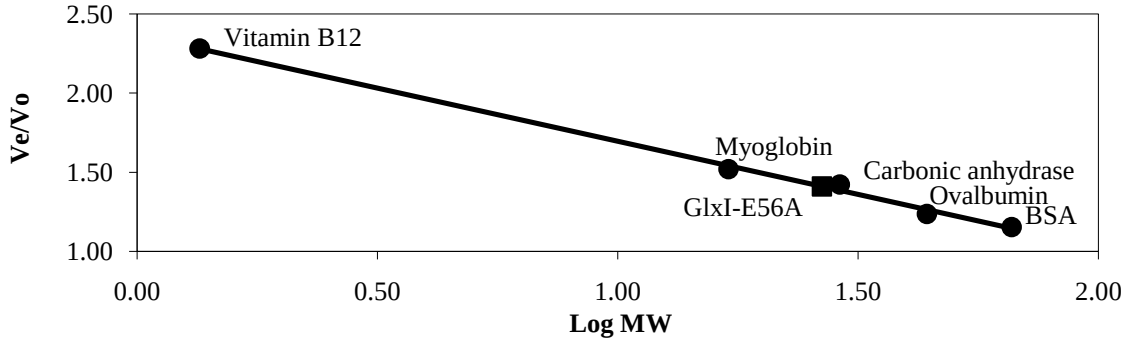


Figure 4.15: The gel permeation chromatographic profile of (■) *E. coli* GlxI-E56A (0.5 mg/mL) that fits into the Bio-Rad protein standard plot using a Superdex75 HR 10/30 column and a buffer containing 20 mM Tris (pH 7) and 150 mM KCl with a flow rate of 0.5 mL/min. The protein standards (●) consisted of BSA (66 kDa), ovalbumin (44 kDa), carbonic anhydrase (29 kDa), myoglobulin (17 kDa) and vitamin B12 (1.35 kDa).

Analysis of the X-ray crystallographic structure of *E. coli* GlxI (PDB: 1FA8) indicated that the enzyme is a homodimer with two intermolecular domain-domain interactions (Figure 4.16A). Each monomer consists of two domains, the N-terminal domain (residues 3–60) with a β1α1β2β3β4 arrangement and the C-terminal domain (residues 72–126) containing a topology of a β5α2β6β7β8 motif. Both domains are connected by an extended loop (residues 61–71). A series of 8 β-sheets (4 from each subunit) curves up to form a catalytic pocket at the dimeric interface. The α-helices (α1 and α2) that locate at the outer layer interact with outside solvent and keep the core of the enzyme relatively hydrophobic. The active site lies in the bottom of the catalytic pocket and forms an octahedral geometry around the active metal with four metal binding residues (His^{5}, Glu^{56}, His^{74} and Glu^{122}) and two water molecules. The hydrophobic catalytic pocket is formed by the presence of Leu^{41}, Phe^{43} and Val^{54} from one subunit and Ala^{76}, Val^{103}, Ile^{110}, Phe^{112} and Ile^{124} from another subunit. The last nine amino acids at the C-terminus are absent in the enzyme structure without bound inhibitor but are observed in the hydroxamate analogue bound enzyme (unpublished structure, more details in Chapter 6). This C-terminus (residues 126–135) as well as the loop that connects β3 and β4 (residues 46–54) partially blocks the active site, thus creating a wall that prevents the substrate from contacting the solvent. In this C-terminus, the side chain of Ala^{129} is pointing toward the active site, supporting its hydrophobic pocket. Other charged residues Asp^{128} and Arg^{131} are pointing out, thus they are in favor to interact with solvent.

The 3D structural prediction of *E. coli* GlxI-E56A using Ni^{2+}-bound GlxI (PDB: 1F9Z) as a template on a web-based program (Geno3D automatic comparative molecular modeling of

protein [142]) showed that their overall structures, not unexpectedly, are highly superimposed with a root mean square deviation (r.m.s.d.) of 0.711 Å (124 Cα pairs) (Figure 4.16B). However, the shift of the metal binding residues can be observed by their side chain binding distances from the Ni^{2+} atom (Table 4.2). These measured distances indicated that His^{74} moves slightly forward to the active metal (the difference in distance of Ni^{2+}–$His^{74}/N_{\varepsilon 2}$ between Ni^{2+}-bound enzyme and the mutein is 0.247 Å), while larger movements are observed with His^{5} and Glu^{122}, where both residues shift away from the metal (the differences in distance of Ni^{2+}–$His^{5}/N_{\varepsilon 2}$ and Ni^{2+}–$Glu^{122}/O_{\varepsilon 1}$between Ni^{2+}-bound enzyme and the mutein are 0.976 and 0.504 Å, respectively). The shifts of these residues as well as the loss of one metal binding ligand are possibly responsible for the decrease in metal binding affinity, although it is expected that there is large error in the coordinates generated for the homology modeled structure. As well, Glu^{122} is a significant residue that is believed to be involved in the catalytic reaction (Chapter 6), and its new position in the mutant might cause the observed decrease in activity. In addition, the predicted structure of GlxI-E56A was also performed using *Pseudomonas aeruginosa* FosA (PDB: 1LQK) as a template. However, the structure is not well solved and thus not shown here.

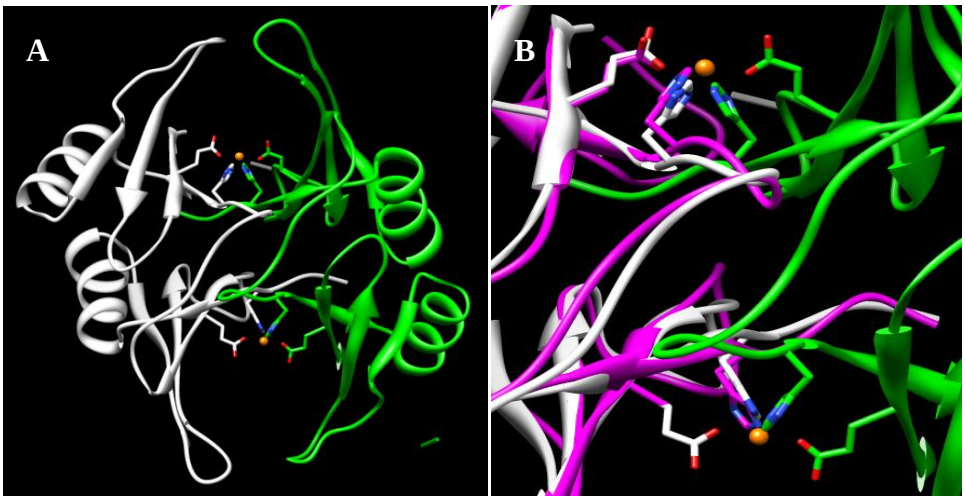


Figure 4.16: (A) The ribbon structure of Ni^{2+}-bound *E. coli* GlxI (PDB:1F9Z) showing the metal binding residues (His^{5}, Glu^{56}, His^{74} and Glu^{122}) in stick form and the Ni^{2+} atom as an orange sphere at the dimeric interface of a homodimer, where each color represents one subunit (white and green). **(B)** The superimposed structures of Ni^{2+}-bound *E. coli* GlxI (white and green, PDB:1F9Z) and the predicted structure of *E. coli* GlxI-E56A (magenta) by Geno3D web-based program [142] showed the shifts of the metal binding residues that possibly cause the decrease in metal binding affinity and enzymatic activity. The 3D structures were generated by using the UCSF Chimera program (University of California, San Francisco) [1].

Table 4.2: The distances of the Ni^{2+} atom and the metal binding residues of Ni^{2+}-bound *E. coli* GlxI (PDB: 1F9Z) and the predicted structure of GlxI-E56A by Geno3D web-based program [142].

Interactions between Ni^{2+} atom and metal binding residues	***Ni^{2+}-bound GlxI***	***Predicted GlxI-E56A***
Ni^{2+} – His5/$N_{\varepsilon 2}$	2.032 Å	3.008 Å
Ni^{2+} – Glu56/$O_{\varepsilon 1}$	2.163 Å	-
Ni^{2+} – His74/$N_{\varepsilon 2}$	2.172 Å	1.925 Å
Ni^{2+} – Glu122/$O_{\varepsilon 1}$	2.162 Å	2.666 Å

Pseudomonas aeruginosa FosA, a member of the βαββββ protein superfamily, is also a homodimer that forms two active sites at the dimeric interface (PDB: 1LQP, [169]) (Figure 4.17A). One subunit of this enzyme consists of two domains, the N-terminal domain (residues 1–52) and the C-terminal domain (residues 64–134) that are connected by an extended loop (residues 53–63). The topology of the N-terminal domain is β1α1β2β3β4, which is quite different from that of the C-terminal domain, containing a β5α2β6β7β8α3β9 motif. These 9 β-strands curve up to form a catalytic pocket. The first β-sheet of the N-terminal domain of one subunit and the first β-sheet of the C-terminal domain of another subunit provide a platform for the active site. The C-terminal α-helix, α3 (residues 144–127), locates almost in the middle of the catalytic pocket, thus creating a wall that protects the bound substrates, fosfomycin and GSH, from the outside environment. This helix resembles the helix B in GlxI (residues 81–95 in human GlxI), which is the extra helix that only exists in the Zn^{2+}-activated enzyme (more details in Chapter 3). The last β-sheet in the C-terminus (β9) contains a short amino acid sequence (residues 131–134) that lies in an antiparallel fashion with a series of β-sheets in the N-terminal domain of another subunit.

It was found earlier by electron paramagnetic resonance spectroscopy that the active site of FosA is formed by the binding of three metal binding residues (His^{7}, His^{64} and Glu^{110}) and three water molecules around the active metal, Mn^{2+} atom, thus generating an octahedral geometry [110, 111, 175, 176]. However, later the X-ray crystallographic structure of FosA with bound fosfomycin suggested that the enzyme forms a trigonal bipyramidal geometry around the Mn^{2+} center (PDB: 1LQP) [169]. The equatorial ligands consist of two His residues ($His^{7}$and His^{64}) and one phosphonate oxygen of fosfomycin, while the axial ligands contain Glu^{110} and the oxirane oxygen of the substrate. Mutations in these metal-binding residues cause decreases in metal binding affinity as well as enzymatic activity, thus lowering antibiotic

resistance [109, 111]. The activation constant (K_{act}) of Mn^{2+} (K_{act} = 0.08 ± 0.03 μM) lies within the intracellular concentration range of Mn^{2+} in typical microorganisms, suggesting that Mn^{2+} is the preferred metal *in vivo* [111]. Other divalent metals including Fe^{2+}, Co^{2+}, Zn^{2+} and Mg^{2+} are also able to activate the enzyme in the laboratory (the activity of enzyme with Mn^{2+} > Fe^{2+} > Co^{2+} > Zn^{2+} >Mg^{2+}) [110].

In addition, it was found that FosA requires a monovalent metal, K^+, to achieve its optimal activity, while the activity of the enzyme without K^+ is significantly decreased [111]. This monovalent metal locates in the K^+-loop (residues 91–99) that connects β6 and β7. The residues in this K^+-loop form six coordinated interactions with the K^+-atom *via* three main chain carbonyl oxygens of Ser^{94}, $Glu^{95}$and Gly^{96}, two hydroxyl groups of Ser^{94} and Ser^{98} and one amide oxygen of Asn^{92} (Figure 4.17B). The deletion or disruption in this K^+-loop causes a dramatic decrease in enzymatic activity (k_{cat} of K^+-deletional mutant is approximately 5% of that of wild type enzyme [113]). This K^+ atom can also be replaced by other monovalent metals with similar radii [111]. It was found that the replacement of K^+ atom with Tl^+, Rb^+ or NH_4^+ decreases the enzymatic activity to approximately 70% of that of K^+-bound enzyme, while the enzyme with bound Na^+, Li^+ or Cs^+ only exhibits approximately 9% activity [111]. The activation constant of K^+ atom is much higher than that of the Mn^{2+} atom (K_{act} of K^+ = 6.2 ± 1.6 mM) [111], suggesting that the monovalent metal has a lower binding affinity and is only partially involved in catalysis. The loss of the metal binding residues (His^7, His^{64} and Glu^{110}) by mutagenesis also affected the binding of the K^+ atom such that K_{act} of K^+ atom is increased (K_{act} of K^+ in E113A was increased by 32-fold, while those of H7A and H67A were increased by 5–10 fold) [111]. There is no direct interaction between Mn^{2+} and K^+ atoms (the distance between Mn^{2+} and K^+ is 6.527 Å). These two metals are bridged *via* two metal binding residues, K^+-binding residue Ser^{98} and Mn^{2+}-binding residues Glu^{110}. Thus, it was suggested that K^+ ion might support a balance of charges of the active metal and lower the activation barrier for the addition of the anionic nucleophile to the drug molecule [111, 113].

The binding site for the substrate of the FosA reaction, fosfomycin with K_m^{fos} of 48 ± 3 μM and K_d^{fos} of 62 ± 9 μM under studied conditions [111], is formed by the amino acid residues Thr^9, Lys^{90}, Ser^{94}, Tyr^{100} and Arg^{119} (Figure 4.17C) [113]. The residue Thr^9 interacts with the oxirane ring, thus it was suggested that this residue might be involved in epoxide opening *via* a protonation or polarization of the oxirane oxygen [113]. Other fosfomycin

binding residues (Lys^{90}, Ser^{94}, Tyr^{100} and Arg^{119}) are within H-bonding distance to the phosphonate oxygens. The residue Ser^{94} in the K^+-loop is able to bridge fosfomycin and the K^+ atom, suggesting that the K^+-loop and the K^+ atom serve to orient the phosphonate of fosfomycin in the active site (distance of K^+ atom and phosphonate is 4.4 Å) [113].

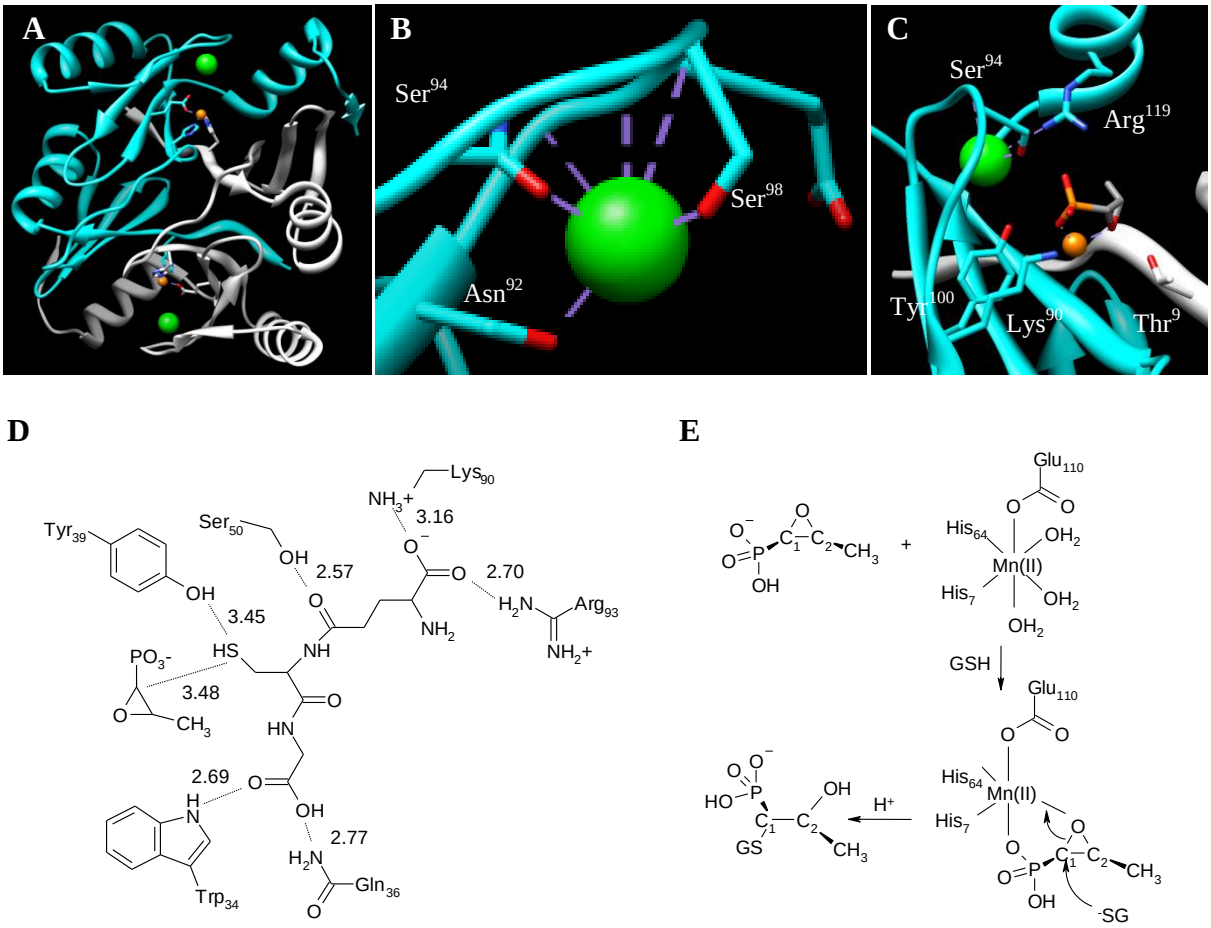


Figure 4.17: The ribbon structure of *Pseudomonas aeruginosa* FosA (PDB: 1LQP) showing **(A)** the active sites at the dimeric interface, **(B)** the K^+-binding site and **(C)** the fosfomycin binding site. Two subunits are represented as cyan and white, while the Mn^{2+} atom and K^+ atom are shown as the orange and green spheres, respectively. The metal and fosfomycin binding residues are shown in stick structures. **(D)** The GSH binding site and **(E)** the proposed reaction mechanism of FosA suggest the interactions of protein residues and the substrates, GSH and fosfomycin, and the epoxide ring opening upon interaction with GSH, respectively. The 3D structures were generated by using the UCSF Chimera program (University of California, San Francisco) [1].

The K^+- and fosfomycin-binding sites seem to occupy half of the catalytic pocket (the C-terminal domain of one subunit), thus donating another half (the N-terminal domain of another subunit) to establish a binding site for GSH, the co-substrate of the FosA reaction. The

GSH binding site is formed by 6 electrostatic or H-bonding interactions with Trp^{34}, Gln^{36}, Tyr^{39}, Ser^{50}, Lys^{90} and Arg^{93} as well as second-sphere van der Waals interactions with Cys^{48}, Tyr^{62}, His^{64}, Gln^{91}, Arg^{119} and Tyr^{128} (Figure 4.17D) [113]. It was found that Tyr^{39} is likely involved in the ionization of GSH during the catalytic reaction [113]. This amino acid is conserved among FosA family members and the mutagenesis of this residue (Y39F with k_{cat} of $14 \pm 2\ s^{-1}$) causes a 14-fold decrease in activity and a 161-fold increase in K_d of GSH [113]. The hydrogen bonding interaction of Tyr^{39} and the thiol group of the cysteine residue in GSH, assisted by the second-sphere interactions of Tyr^{128} and Arg^{119}, lowers its pK_a, which in turn promotes the deprotonation of GSH and initiates the nucleophilic addition to the C1 position of fosfomycin. Other mutagenesis experiments on Trp^{34}, Gln^{36}, Ser^{50}, Lys^{90} and Arg^{93} were also found to reduce the catalytic activity as well as the GSH binding affinity of the enzyme [113]. There is no direct contact of GSH with Mn^{2+} and K^+ atoms. However, two residues, Lys^{90} and Arg^{93} in the first-sphere interactions with GSH, are in the K^+-loop, thus it was suggested that the K^+ atom and the K^+-loop are involved in proper orientation of these stabilizing residues to assist GSH binding. The interaction of GSH and fosfomycin occurs through Lys^{90}, which bridges the phosphonate oxygen of fosfomycin and the carboxyl group of the glycyl residue of GSH *via* its side chain amine. As well, the Arg^{119} residue that directly interacts with fosfomycin, is in a second-sphere interaction with GSH (interacts with GSH *via* Tyr^{39}). These interactions might stabilize and pull the two substrates to locate next to each other, thus optimizing the catalytic reaction.

The reaction mechanism of FosA, as proposed so far is as follows (Figure 4.17E) [110, 111, 168, 175, 176]. First, it is believed that the active site of the enzyme forms an octahedral geometry around the metal center with three metal binding residues (His^7, His^{64} and Glu^{110} in *P. aeruginosa* FosA) and three water molecules. Upon the binding of fosfomycin, its two oxygens (oxirane oxygen and one of the phosphonate oxygens) replace the water molecules, changing the geometry of the active site into the trigonal bipyramidal coordination. The second substrate, GSH, which is ionized by the aid of the enzyme residue Tyr^{39}, then subsequently attacks the fosfomycin at the C1 position, causing the epoxide to ring open and results in the formation of the product GS-fosfomycin that diffuses from the active site.

The predicted binding sites for metals (K^+ and Mn^{2+}) and substrates (fosfomycin and GSH) in *E. coli* GlxI (unpublished structure with bound hydroxamate inhibitor) was

investigated by structural comparison with *P. aeruginosa* FosA (PDB: 1LPQ). The structures of GlxI and FosA are superimposed with the r.m.s.d. of 1.058 Å (61 Cα pairs) (Figure 4.18). Both share overall structural similarity; however, differences can be clearly observed in the connecting loops as well as the C-terminus. The loop connecting β2 and β3 of GlxI (residues 35–38) is longer than that of FosA (residues 34–36). This loop in GlxI also bends toward the active site, thus acting as a lid to protect a substrate from outer solvent. Similarly, the loop connecting β3 and β4 in GlxI (residues 46–53) is longer than the one in FosA (residues 43–45). This long loop curves up from the base of the catalytic pocket, thus partially blocking the active site along with the C-terminus upon binding of the inhibitor to GlxI. On the other hand, the K^+-loop in FosA that connects β6 and β7 (residues 91–99) is shorter but bends toward the active site more than that of GlxI (residues 102–110). Despite possessing several charged residues that can interact with the K^+-atom like the ones in FosA, the loop in GlxI only contains one residue (Thr^{108}) that possibly interacts with K^+ ion. Besides, FosA possesses a longer C-terminus (residues 116–134) that forms the α-helix (α3) and β-sheet (β9) that completely block one side of the active site, thus creating a wall that prevents the substrates from contacting the outside environment. This C-terminus is more defined than the one in GlxI (residues 126–135), in which it is only resolved upon the binding of the inhibitor. Despite the structural similarity between GlxI and FosA, limited space in the active site due to the long loops that act as lids to cover the active site (residues 35–38 and 102–110) and the lack of the long C-terminus in *E. coli* GlxI (residues 144–127) might prevent the binding of the two substrates, fosfomycin and GSH, and thus might account for the loss of detectable fosfomycin-conjugating activities by GlxI-E56A.

One residue in the C-terminus of FosA that binds fosfomycin (Arg^{119}) is missing in GlxI. Moreover, the multiple sequence alignment of *E. coli* GlxI-E56A and FosA from various organisms suggests that GlxI employs none of the fosfomycin-binding residues (Thr^9, Lys^{90}, Ser^{94}, Tyr^{100} and Arg^{119} in FosA) (Figure 4.19). It also suggests that the residue Arg^{98}, a GSH-binding residue in GlxI, might be able to act as fosfomycin binding ligand (Figure 4.19). However, this residue resides in the β-strand (β7) that is located distantly from the active site. Even though the K^+- and fosfomycin-binding sites are missing in GlxI, the GSH binding site is conserved. The multiple sequence alignment also shows that the conserved Tyr^{39} in FosA, the critical residue for ionization of GSH, is replaced by Phe^{43} in GlxI (Figure 4.19). The absence

of this residue in GlxI as well as its weaker affinity for fosfomycin and the K^+-atom might account for the absence of detectable fosfomycin-conjugating activity by GlxI-E56A.

Figure 4.18: The superimposed structures of *Pseudomonas aeruginosa* FosA (yellow and magenta, PDB: 1LQP) and *E. coli* GlxI (white and cyan, unpublished structure of hydroxamate inhibited enyzme) showing the active site metals including K^+ (green) and Mn^{2+} (orange) in FosA and Ni^{2+} (blue) in GlxI. These 3D strcutures were generated by UCSF Chimera program (University of California, San Francisco) [1].

```
FosA_pa         MLTGLNHLILAVADLPASIAFYRDLLGFRLEARWDQG------AYLELG----SLWLCLS  50
FosA_tn         MLQSLNHLTLAVSDLQKSVTFWHELLGLTLHARWNTG------AYLTCG----DLWVCLS  50
FosA_kp         MLSGLNHLTLAVSQLAPSVAFYQQLLGMTLHARWDSG------AYLSCG----DLWLCLS  50
FosA_ct         MLSGLNHLILAVRDVARSLAFYQSLPGLRLHARWDGG------AYLSCG----SLWLCLS  50
FosA_so         MLSGLNHLILAVRDVERSLEFYRQTLGFHLHARWQQG------AYLTLG----ELWLCLS  50
E56A            M--RLLHTMLRVGDLQRSIDFYTKVLGMKLLRTSENPEYKYSLAFVGYGPETEEAVIALT  58

FosA_pa         REPQ---YGGPAADYTHYAFGIAAADFARFAAQLRAHGVREWKQ-NRSEGDSF---YFLD  103
FosA_tn         YDEARQYVPPQESDYTHYAFTVAEEDFEPLSQRLEQAGVTIWKQ-NKSEGASF---YFLD  106
FosA_kp         LDPQRRITPPEESDYTHYAFSISEADFASFAARLEVAGVAVWKL-NRSEGASH---YFLD  106
FosA_ct         LDEN---AGSPAADYTHYAFTVSEADFPQVVETLTAAGVTPWKT-NRSEGDSW---YFLD  103
FosA_so         LDDTRAACAPR--DYTHYAFSITSADFPAMVERLRQAGVRQWKS-NRSEGESL---YFLD  104
E56A            YNWGVDKYELGTA-YGHIALSVDNA--AEACEKIRQNGGNVTREAGPVKGGTTVIAFVED  115

FosA_pa         PDGHRLLAHVGDLRSRLAACRQAPYAGMRFAD---  135
FosA_tn         PDGHKLLLHVGSLAARLAACREKPYAGMVFTSDEA  141
FosA_kp         PDGHKLLLHVGSLAQRLAACREQPYKGMVFF--DQ  139
FosA_ct         PDGHQLLLHVGSLATRLAACRAAPYKGMRFYD---  135
FosA_so         PDGHQLLIHSGDLASRLAACRAAPYQGMAFSEDLK  139
E56A            PDGYKIELIEEKDAGR----------GLG-----N  135
```

Figure 4.19: The multiple sequence alignment of the *E. coli* Glyoxalase I mutein (E56A) with fosfomycin-resistance protein from various organisms (organism name follows by National Center for Biotechnology Information (NCBI) accession number) including genomically encoded FosA from *Pseudomonas aeruginosa* (FosA_pa, 1LQP), plasmid encoded FosA from *Pseudomonas aeruginosa* (FosA_tn, 1NPB), FosA from *Klebsiella pneumoniae* (FosA_kp, YP_002240656), FosA from *Cronobacter turicensis* (FosA_ct, YP_003210620) and FosA from *Serratia odorifera* (FosA_so, ZP_06639975). The predicted divalent metal

(Mn^{2+}) binding residues are marked with grey highlight, while the predicted K^+ binding residues are marked in yellow. The predicted residue that interacts with both Mn^{2+} and K^+ atoms is highlighted in blue. The cyan highlight emphasizes the predicted glutathione binding residues and the magenta highlight represents the predicted fosfomycin binding residues. The predicted residue that interacts with both substrates is highlighted in green, while the predicted residue that binds both K^+ atom and fosfomycin is highlighted in red. The alignment was created using CLC Free Workbench (version 3.0.1) with the accurate alignment algorithm (http://www.clcbio.com).

4.4.5: Structural Similarity among Drug Resistance Proteins in the βαβββ Superfamily

There are five classes of proteins in the βαβββ superfamily including Glyoxalase I, methylmalonyl Co-A epimerase, dioxygenase, fosfomycin resistance protein, bleomycin resistance protein and mitomycin resistance protein that share high structural similarity but low sequence identity and similarity [30, 40, 41]. Drug resistance proteins such as bleomycin and mitomycin C resistance proteins that contain no metal binding ligands are believed to have evolved by modification of the precursor gene along with various mutations to create a hydrophobic cavity, a suitable environment for sequestering the somewhat hydrophobic antibiotics. In contrast to other resistance proteins in this superfamily, fosfomycin resistance protein acts as a metalloenzyme that breaks the epoxide ring of fosfomycin, thus inactivating the drug to prevent its interference with bacterial cell wall biosynthesis. This enzyme forms an active metal center that interacts with the incoming substrates, catalyzes the reaction and then releases the product from its active site. Bleomycin and mitomycin resistance proteins, on the other hand, are neither enzymes nor metalloproteins. These proteins reversibly bind the drugs in their hydrophobic pockets, thus preventing these drugs from their commonly biological function and preventing the antibiotic-producing organism from inhibiting its own cellular biochemistry.

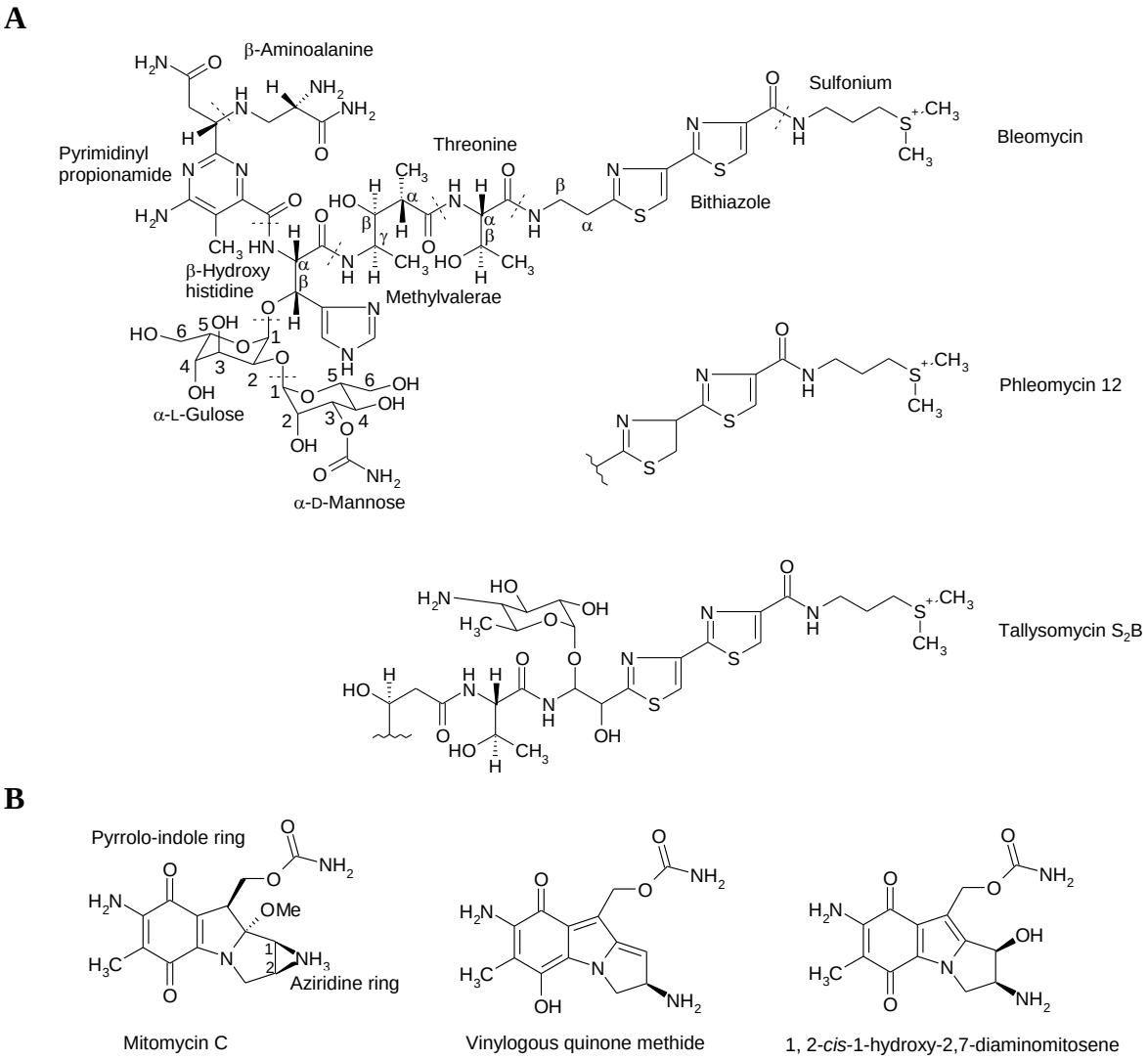


Figure 4.20: (A) The structures of bleomycin A2 and its analogues (phleomycin 12 and tallysomycin S_2B). The dashed lines indicate each molecular unit of bleomycin including β-aminoalanine, pyrimidinyl propionamide, β-hydroxy histidine, α-L-gulose, α-D-mannose, threonine, methylvalerae, biothiazole and γ-aminopropyldimethylsulfonium. **(B)** The structure of mitomycin C, its reactive vinylogous quinone methide that forms after enzymatic or chemical reduction and its reoxidized product, the reactive quinone methide, 1, 2-*cis*-1-hydroxy-2, 7-diaminomitosenes.

Bleomycin (BM) is a large glycopeptide antitumor compound with a molecular mass of approximately 1.4 kDa. It binds to DNA and RNA in a sequence-specific fashion. It also binds metal ions, which catalyze the function of reactive oxygen species (ROS); the later cause strand breaks in the nucleic acids [177-179]. Its biosynthetic pathway involves a hybrid nonribosomal peptide synthetase and polyketide synthase using nine amino acids, two methyl groups from *S*-adenosyl methionine and carboxylic acid precursors [180, 181]. BM consists of

a metal-binding domain (β-aminoalanine, pyrimidinyl propionamide and β-hydroxy histidine), a sugar domain (α-L-gulose and α-D-mannose), linker peptide (threonine and methylvalerae) and the C-terminal tail (biothiazole and γ-aminopropyldimethylsulfonium) (Figure 4.20A). BM is able to bind various types of metals including Fe^{2+}, Co^{2+}, Co^{3+}, Zn^{2+}, Cu^{2+} and Cd^{3+} [182, 183]. These metals are crucial for activity of BM and BM derivatives that are incapable of binding metal ion are found to be inactive [177, 179]. BM forms a square pyramidal coordination geometry around the active metal with an axial ligand that is formed by the primary amine nitrogen of the β-aminoalanine moiety and equatorial ligands that consist of two nitrogen atoms from an imidazole moiety and an amide from the β-hydroxyl histidine, one from pyrimidine N1 and the other from the secondary amide from β-aminoalanine [178, 184]. Investigations on BM bound to DNA suggest that another axial ligand is molecular oxygen, thus forming an octahedral geometry around the metal center [178, 184]. The disaccharide moiety of BM plays a role in cellular uptake and cell recognition, while the linker unit is significant for DNA cleavage efficiency [178, 179]. The C-terminal γ-aminopropyldimethylsulfonium and biothiazole moieties (DNA binding domain) support the binding affinity of BM and its target DNA as well as likely contributing to DNA recognition and selectivity [178, 179].

BM is mainly produced from *Streptomyces*, the genus of microorganisms that has been found to produce many useful antibiotics [185]. These organisms thus possess self-resistance mechanisms such that the antibiotic cannot harm the antibiotic producer. It was earlier found that *Streptomyces verticillus*, a BM producing organism, contains two genes, *blmA* and *blmB* that are involved in inactivation of BM [185]. The former gene codes for a bleomycin-binding protein (BLMA) and the later codes for a bleomycin *N*-acetyltransferase (BAT) [185, 186]. BLMA binds BM in its binding pocket through hydrogen bonding and hydrophobic interactions [186, 187]. BAT, on the other hand, modifies BM by acetylating the α-amine functionality of the β-aminoalanine moiety in the presence of acetyl-coenzyme A [188]. This acetylated BM is incapable of inducing DNA damage *in vitro* [189]. Streptomyces strains that do not produce BM, such as *Streptoalloteichus hindustanus* and *Streptomyces flavoviridis*, also possess the gene that codes for the bleomycin-binding protein [190, 191]. *Streptomyces flavoviridis* produces phleomycin, a BM analogue that lacks a double bond on one of the thiazole rings (Figure 4.20A). On the other hand, *Streptomyces verticillus* ATCC21890, a BM

producing organism, contains a gene that codes for a phleomycin-binding protein (SvP) [192]. These binding proteins are able to enhance cellular protection against both BM and phleomycin, suggesting a close relation between the antibiotics as well as their binding proteins [192, 193]. *Streptoalloteichus hindustanus* produces the antibiotic tallysomycin, another BM analogue containing an extra 4-amino-4,6-dideoxy-L-talose at the N-terminus of the bithiazole moiety (Figure 4.20A). The bleomycin-binding protein from this organism (ShBle) was found to be able to bind BM, tallysomycin and phleomycin [191, 194]. Even though the binding proteins for BM and BM analogues originated from different organisms, they share high sequence similarity and identity (~64% identity and ~72% similarity).

Other organisms that do not produce BM nor its analogues are also found to employ BM resistance mechanisms such as *Streptomyces caespitosus, Streptomyces lavendulae, Staphylococcus aureus* and *Klebsiella pneumonia* in order to survive under such stressed conditions [195-197]. *Streptomyces caespitosus* and *Streptomyces lavendulae* produce mitomycin C (MC), an antitumor antibiotic that is able to crosslink to DNA and eventually leads to cell apoptosis [198-200]. MC is a small compound that consists of the pyrrolo-indole ring system and an aziridine ring with a molecular mass of approximately 334 Da (Figure 4.20B). The inactive mitomycin C can be activated by enzymatic or chemical reduction at its indole moiety, thus resulting in the loss of the methoxide substituent. This process initiates the opening of the aziridine ring. This compound then becomes an unstable alkylating agent containing a vinylogous quinone methide that is reactive toward DNA by covalently cross-linking DNA strands (Figure 4.20B). The electrophilic reduced product of this reaction on MC can also react with water molecules and produce a mixture of 1, 2-*cis*-1-hydroxy-2, 7-diaminomitosenes (1, 2-*cis*-1-hydroxy-2, 7-DAMS, Figure 4.20B) and 1, 2-*trans*-1-hydroxy-2, 7-DAMS, while its nucleophilic pathway leads to the formation of 2, 7-diaminomitosene [201, 202]. To prevent damage from this drug, *Streptomyces lavendulae*, the MC producing organism, contains the *mcrA* and *mrd* genes that code for resistance proteins against MC [203, 204]. The gene *mrcA* codes for MCRA, a flavoenzyme that functions as a hydroquinone oxidase to reoxidize the reductively reactive MC to its nontoxic prodrug form [205, 206]. The *mrd* gene codes for a mitomycin binding protein (MRD) that can reversibly bind the prodrug form of MC as well as 1-hydroxy-2, 7- DAMS [207]. *Streptomyces lavendulae* also contains another MC resistance gene, *mct*, which codes for a membrane-associated MC transport

protein (Mct) [208]. This protein together with MRD functions as a MC export system that secretes MC outside the cells [208]. MRD is also able to sequester Cu(II)-BM with a K_d of 5.1 nM [195]. On the other hand, cells with the transfected *blmA* gene are susceptible to mitomycin C [209]. It is possible that MRD and BLMA are able to replace one another since both share high structure similarity, even though they employ low sequence identity and similarity (2.5% sequence identity and 11.6% sequence similarity).

Methicillin-resistant *Staphylococcus aureus* contains a *blmS* gene that codes for a bleomycin-binding protein (BLMS) [197, 210]. Methicillin, a β-lactam antibiotic, is a small compound with a molecular mass of approximately 380 Da. As well, the transposon Tn5, that was originally isolated from *Klebsiella pneumonia* and has been transformed into *E. coli* cells, contains a *ble* gene that codes for a bleomycin-binding protein (BLMT) [196]. These proteins (BLMS and BLMT) share low sequence identity and similarity (10% identity and 24% similarity). When comparing the amino acid sequences of bleomycin binding proteins from BM and BM analogue-producing organisms (BLMA, ShBle and SvP), only approximately 7% sequence identity and 14% sequence similarity are found. However, it has been suggested that *Staphylococcus aureus* and *Klebsiella pneumonia* recruit BM-binding genes from BM-producing organisms [193]. This information suggests that bleomycin binding proteins are linked to the biosynthesis of the BM family of antibiotics.

The bleomycin-resistance protein from *Streptomyces verticillus* (BLMA) is a homodimeric protein with a molecular weight of approximately 14 kDa (Figure 4.21A) [187]. This acidic protein contains many negatively charged residues that counter the positive charges on the bound BM. Each subunit consists of two domains, the N-terminal domain (residues 1–52) and the C-terminal domain (residues 62–122) (PDB: 1JIF). The N-terminal domain has a β1α1β2β3β4 topology and the C-terminal domain possesses a β5α3β6β7β8 arrangement. These two domains are connected by an α2-helix loop (residues 53–62). Even though these two domains share structural similarity, they possess low sequence identity and similarity (3.3% identity and 17.6% similarity). The binding pocket is created by a series of eight β-strands (four from each subunit) that curve up to provide a large cavity and a long groove that is located at the dimeric interface. There are two binding pockets per dimeric protein that are able to accept two molecules of bleomycin [187]. The groove surface is relatively hydrophobic and is formed by several aromatic residues (Phe^{33}, Phe^{38}, His^{49}, Trp^{65} and His^{117}). The

negatively charged residues (Asp^{32}, Asp^{45}, Glu^{67} and Glu^{122}) that lie along the cleft between the two subunits are found to be significant contributors to ligand recognition [187]. The two α-helices (α1 and α3) that are located at the outer layer of the binding pocket employ charged residues that can interact with solvent. The loop that links β3 and β4 (residues 43–46) is a type-II-β-turn and the loop that connects β6 and β7 (residues 100–103) is considered as a type-I-β-turn. This type-I-β-turn acts as a lid that covers the binding pocket from above. The interactions between subunits are mainly composed of hydrophobic and hydrogen bonding interactions. The N-terminus of one subunit projects into another subunit, forming an arm with several hydrogen bonding interactions with the residues in the first β-sheet (β5) of the adjacent C-terminal domain (Figure 4.21B). The N-terminal arm is significant for protein dimerization, in which its mutagenesis as well as the removal of the N-terminal amino acids leads to a non-functional monomer [211, 212]. Investigation on the N-terminal proline indicates that this residue plays a crucial role in dimerization and is conserved in all investigated bleomycin-binding proteins (Figure 4.22). As well, it is also suggested that the dimeric conformation of the bleomycin-binding protein is a significant factor in the sequestration of bleomycin [211, 212].

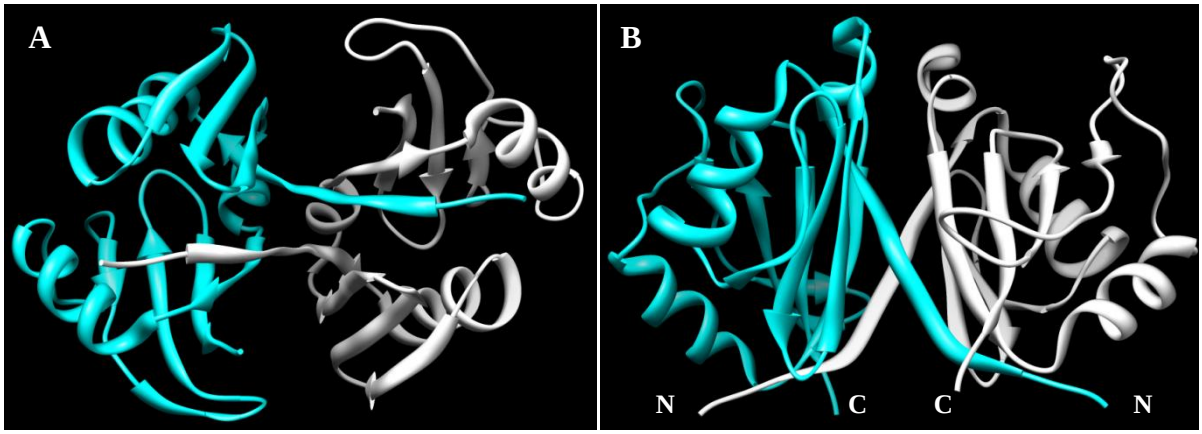


Figure 4.21: The ribbon structure of bleomycin-binding protein from *Streptomyces verticillus* (BMLA, PDB: 1JIF) shows **(A)** two binding pockets within the dimeric interface (cyan and white, side view) and **(B)** the N-terminal arm exchange of each monomer (top view). The 3D structures were generated by using the UCSF Chimera program (University of California, San Francisco) [1].

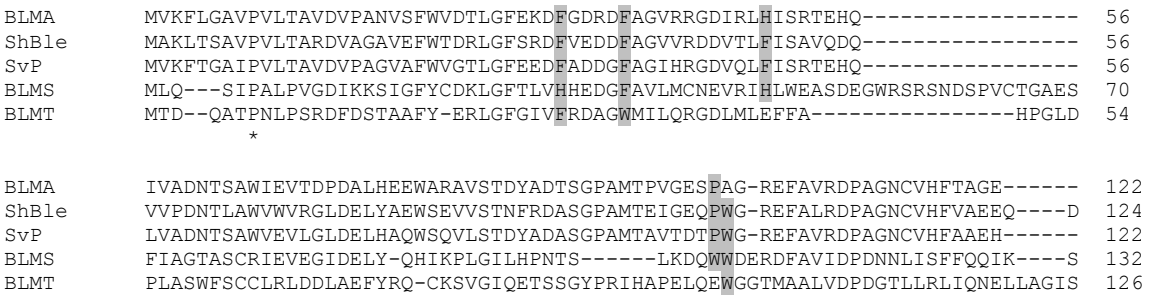


Figure 4.22: Multiple sequence alignment of bleomycin-resistance proteins from various organisms including (organism name follows by National Center for Biotechnology Information (NCBI) accession number and abbreviation of the protein) *Streptomyces verticillus* (1QTO, BLMA), *Streptoalloteichus hindustanus* (P17493, ShBle), *Streptomyces verticillus* (AAB00464, SvP), *Staphylococcus aureus* (AAA26671, BLMS) and *Klebsiella pneumonia* transposon Tn5 (P13081, BLMT). The asterisk indicates the N-terminal proline that is significant for dimerization. The grey highlight presents the predicted amino acids that are involved in interacting with the biothiazole moiety of bleomycin. The alignment was created using CLC Free Workbench (version 3.0.1) with the accurate alignment algorithm (http://www.clcbio.com).

In the binding pocket of the bleomycin-resistance protein (Figure 4.23), the bithiazole moiety of BM lies almost perpendicular to the series of N-terminal β-strand from one subunit and the type-I-β-turn (residues 100–103) of another subunit that acts as a lid to cover the binding pocket from above. The first terminal thiazolium ring of BM stacks with the benzene ring of Phe^{33}, while the second thiazolium ring interacts with aromatic residues (Phe^{38} and His^{49}) and Pro^{101} from the type-I-β-turn, an environment that is similar to the binding of aromatic rings by bases of DNA [178]. The methylene group that extends from the penultimate thiazolium ring of the drug is also stabilized by a hydrophobic interaction with Phe^{38}. These aromatic residues are conserved in bleomycin-resistance proteins from all investigated organisms (Figure 4.22). The terminal γ-aminopropyldimethylsulfonium group in BM that is located at the end of the groove is found to be flexible [187]. This terminal amine contains the positive charges that are necessary for electrostatic interaction with both DNA and the bleomycin-resistance protein [187]. The carboxyl group of the threonine moiety and the hydroxyl group of the methylvalerae moiety of BM interact with the carboxylate side chain of Glu^{105} *via* a water molecule. This Glu^{105} residue is also stabilized by His^{117} (the distance of $O_{\varepsilon 1}$ of Glu^{105} and $N_{\delta 1}$ of His^{117} is 2.92 Å). The hydroxyl and methyl side chains of the threonine moiety of BM are shifted by 47° and 54.5°, respectively. In the metal binding domain of BM, only the pyrimidinyl propionamide moiety interacts with the BLMA. The oxygen from the terminal amide of the pyrimidinyl propionamide moiety of BM is stabilized by the hydroxyl

side chain of Ser^{51}, while its amide nitrogen interacts with the main chain carbonyl group of Arg^{52}. The carbonyl oxygen from the amide of pyrimidinyl propionamide moiety that links to the β-hydroxyl histidine of BM forms a hydrogen bond with the hydroxyl group of Tyr^{86}, while the amino group in the pyrimidine ring of pyrimidinyl propionamide moiety interacts with the main chain carbonyl group of Gly^{113} and Thr^{62}. The multiple sequence alignment suggests that these amino acids from the bleomycin-binding proteins are not well conserved since the stabilization of BM only requires interactions from side chains of charged residues and main chain carbonyl groups (Figure 4.22). In the disaccharide domain, the carbamoyl nitrogen that is attached to the 3-hydroxyl group on α-D-mannose forms hydrogen bonding interactions with a main chain carbonyl oxygen of Asp^{60}. In this sugar molecule, the hydroxyl group that is attached to the C4 position interacts with the guanidine side chain of Arg^{109} and carbonyl oxygen from the amide in the pyrimidinyl propionamide moiety that links to the β-hydroxyl histidine of BM. As well, the hydroxyl group that is attached to C3 in the α-L-gulose interacts with the terminal amide nitrogen and primary amine nitrogen of β-aminoalanine moiety of BM.

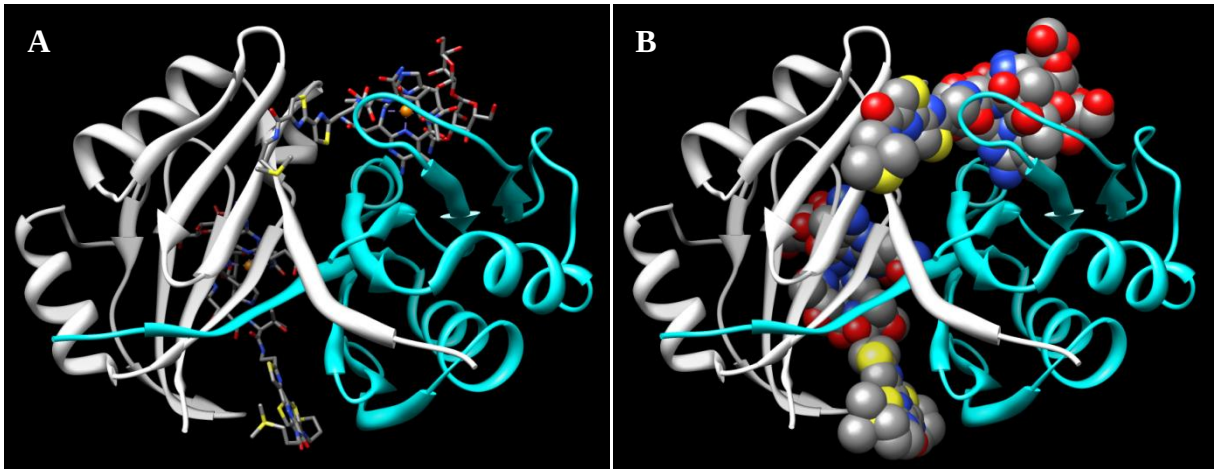


Figure 4.23: The ribbon structure of bleomycin-binding protein from *Streptomyces verticillus* (BMLA) with two molecules of Cu(II)-bleomycin complex in its binding pockets (PDB: 1JIF). Each subunit of BLMA is represented as one color (white and cyan). The binding pockets at the dimeric interface are formed by a series of β-strands from both subunits. The drug molecules are represented as **(A)** stick and **(B)** space fill with Cu^{2+} atom as an orange sphere. The 3D structures were generated by using the UCSF Chimera program (University of California, San Francisco) [1].

The conformations of the two Cu(II)-BM molecules in each binding pocket of BLMA are different from each other (PDB: 1JIF). It was found that the bithiazole moiety of BM is in its *trans* conformation (BM-I) in one binding pocket (pocket I formed by the N-terminus of subunit A and the C-terminus of subunit B) and a mixture of *cis* and *trans* conformations (BM-II) in the other pocket (pocket II formed by the N-terminus of subunit B and the C-terminus of subunit A). The interactions of BM-II and BLMA are similar to those of BM-I in its binding pocket as previously described. However, the major change in the *cis* conformation of BM-II is the terminal γ-aminopropyldimethylsulfonium tail that bends upward and lies above the groove (instead of laying along the groove like in the one in the *trans* conformation). This tail is flexible and does not form any interaction with the protein.

The two subunits of BLMA employ high structural similarity (r.m.s.d. of 0.44 Å for 122 Cα pairs) with significant differences in the type-I-β-turn (residues 100–103) and the type-II-β-turn (residues 43–46). The difference in these two subunits are not clearly observed in the unoccupied BLMA (PDB: 1QTO); however, the superimposed structures of the apo-BLMA and the Cu(II)-BM bound BLMA (PDB: 1QTO and 1JIF, respectively) clearly suggest two asymmetric binding pockets and protein movement upon drug binding (r.m.s.d. of 0.82 Å for 122 Cα pairs). The movement of the type-I-β-turn shifts toward the drug molecule and allows Pro^{101} to interact with the bithiazole moiety of BM by hydrophobic interactions. The shift of Pro^{101} from subunit B in binding pocket I (Cα of Pro^{101} shifts 5.63 Å toward the bithiazole moiety) is larger than the one from subunit A in binding pocket II (Cα of Pro^{101} shifts 3.69 Å). As well, the shift of the side chain of Asp^{45} in the type-II-β-turn from subunit B is larger than the one in subunit A (the shift of Cγ of Asp^{45} is 2.39 Å in subunit B and is 0.53 Å in subunit A). Since this residue shifts away from the N-terminus of another subunit, the type-II-β-turn of subunit B cannot interact with the N-terminal β-strand from subunit A, while the type-II-β-turn of subunit A can still form interactions with the N-terminal β-stand from subunit B. These data suggest that subunit B in BLMA is more flexible than subunit A, and it is more likely that binding pocket II is more flexible than binding pocket I.

The difference in flexibility also causes one binding pocket to have a higher BM binding affinity than another. It was found that the dissociation constant (K_d) of the first BM binding to BLMA is 630 ± 60 nm, while that of the second BM binding is 120 ± 10 nm,

suggesting that the second drug molecule binds more strongly to the protein and is more favourable than the first one [187]. The guanidino group from the side chain of Arg^{47} forms salt links with the carboxylate side chain of Glu^{67} (distances of 2.83 and 2.80 Å) within the binding pocket I and shifts away from the terminal γ-aminopropyldimethylsulfonium of the BM-I, thus reducing the repulsion between their positive charges. In binding pocket II, only one salt link is formed between the side chains of Arg^{47} and Glu^{67} due to the shift of the type-II-β-turn that is located further from the N-terminus of subunit A. This lower degree of interaction causes higher flexibility in binding pocket II. Thus, the BM-II is more likely binding loosely inside the binding pocket, which and is probably the reason why it can occupy both *cis* and *trans* conformations. The movement of Pro^{101} in binding pocket I upon the binding of BM-I is greater than that of binding pocket II, causing tighter interaction of BM-I in the binding pocket I than that of the Bm-II in the binding pocket II. Thus, due to this tighter binding, BM-I can only occupy the *trans* conformation.

The asymmetric pockets in the βαβββ superfamily were suggested earlier in the homodimeric apo-GlxI, where the asymmetric catalytic sites can be observed in the X-ray crystallographic structure and the ^{15}N-^{1}H HSQC NMR experiments [38, 45]. One of these asymmetric pockets employs higher substrate binding affinity and, thus, is chosen by the substrate [38, 45]. The binding of BM to bleomycin-binding protein probably follows a similar fashion. It is possible that the binding pockets in the apo-bleomycin binding protein are slightly asymmetric, although this cannot be clearly observed from the X-ray crystallographic data. The first BM (BM-II) enters the binding pocket of bleomycin-binding protein (binding pocket II) that is more flexible compared to the other one (binding pocket I). This first BM is flexible and is able to form either a *cis* or a *trans* conformation. The binding of this BM in the binding pocket induces a conformational change in the second binding pocket (binding pocket I), thus it becomes narrower and has a higher binding affinity for the second incoming BM molecule (BM-I), suggesting cooperative binding of BM in the bleomycin-binding protein [187].

The stabilizing residues of bleomycin-binding proteins from the organisms that produce BM and its analogues (BLMA, ShBle and SvP) are more conserved than those from organisms that are found to produce other antibiotics (BLMS and BLMT) (Figure 4.22). It was found that ShBle, the bleomycin-resistance protein from the tallysomycin-producing organism, employs a

lower dissociation constant for BM (55 nM) than that of BLMA (280 nM), the bleomycin-resistance protein from the BM-producing organism [194, 213]. The structures of ShBle (PDB: 2ZHP) and BLMA (PDB: 1JIF) are well superimposed (r.m.s.d. of 0.66 Å for 120 Cα pairs) with a major deviation in the type-I-β-turn (residues 100–103 in BLMA and ShBle), in which this β-turn in ShBle locates closer to the binding pocket than that of BLMA. Besides, the multiple sequence alignment (Figure 4.22) as well as the X-ray crystallographic structure of ShBle (PDB: 2ZHP) indicates that ShBle contains more aromatic residues that stabilize the bisthiazole moiety and its methylene group of BM than that of BLMA, thus creating a more hydrophobic environment and providing a higher binding affinity for BM.

BLMT from the transposon Tn5 of *E. coli*, an organism that does not produce BM, has a lower dissociation constant for BM (32 nM) than that of BLMA and ShBle (280 and 55 nM, respectively) [194, 213]. The multiple sequence alignment (Figure 4.22) and the X-ray crystallographic structure of BLMT (PDB: 1MH6) suggests that it contains more stabilizing residues that interact with the disaccharide unit of BM than those from BLMA (PDB: 1JIF) (r.m.s.d. between BLMT and BLMA is 1.09 Å for 74 Cα pairs). Since the binding pockets of BLMT are more packed and most of the stabilizing residues locate in the loops close to the binding cleft, interactions of BLMT and BM are stronger than those of BLMA and ShBle. As well, it has been suggested that the flexibility of the type-I-β-turn loop in BLMT (residues 95–103) that interacts with the bisthiazole moiety of BM is higher than those of BLMA and ShBle, thus allowing it to tightly bind to the drug [213]. This flexibility also decreases substrate specificity so that the protein interacts with different antibiotics [213]. This is an advantage for the non-BM producing organisms such that it can also bind to BM as well as its analogues using the very same protein to prevent cell damage.

It was also found that the mitomycin-binding protein (MRD) from *Streptomyces caespitosus*, the mitomycin producing organism, is able to accept bleomycin in its binding pocket [195]. Both drug binding proteins are from the βαβββ superfamily, thus they share high structural similarity to Glyoxalase I. The apo-MRD (protein without bound drug) from *S. lavendulae* is a homodimer (by X-ray crystallographic analysis (PDB: 1KMZ) and gel permeation chromatography) with a molecular weight of approximately 14 kDa (Figure 4.24A) [214]. Each subunit consists of two domains, the N-terminal domain (residues 3–53) with a β1α1α2β2β3β4 topology and the C-terminal domains (residues 74–130) with a β5α4β6β7β8

arrangement. These two domains are connected by a two-turned α-helix (α3, residues 54–73). Both domains share structural similarity (r.m.s.d. of 1.6 Å for 38 Cα pairs) but low sequence identity and similarity (14% identity and 23% similarity). The major differences in the N-terminal and the C-terminal domains are the short α2-helix that connects α1 and β2 and the β3-sheet that acts as a loop, connecting β2 and β4. Residues 67–69 are not well defined and are not shown in the crystallographic structure. The drug binding pocket is located in a deep cleft, forming by a series of β-sheets (total of 8 β-strands). This drug binding cleft is formed within one subunit (two binding pockets per dimer) similarly to that of the enzyme methylmalonyl-CoA epimerase (PDB: 1JC5) (more details in Chapter 5). The connecting two-turn α-helix is located almost in the middle of the binding cleft, partially blocking the bound drug from outside solvent.

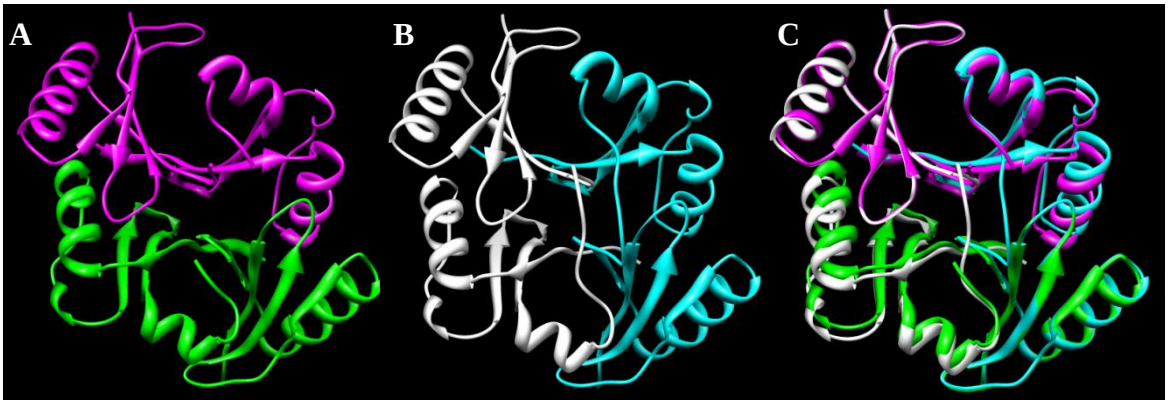


Figure 4.24: The ribbon structures of **(A)** *Streptomyces lavendulae* mitomycin-binding protein (MRD) (PDB: 1KMZ, magenta nd green), **(B)** MRD with bound 1, 2-*cis*-1-hydroxy-2, 7-diaminomitosenes (bound drug not shown, PDB: 1KLL, white and cyan) and **(C)** the superimposed structures of MRD with and without bound drug. These 3D strcutures were generated by UCSF Chimera program (University of California, San Francisco) [1].

Interestingly, the *S. lavendulae* MRD with bound 1, 2-*cis*-1-hydroxy-2, 7-DAMS forms a different dimeric arrangement (PDB: 1KLL) (Figure 4.24B) [214]. Two binding pockets are found at the dimeric interface, where one binding pocket is formed by a 4 β-strand from the N-terminal domain of one subunit and another 4 β-strand from the C-terminal domain of the other subunit. The structures of MRD with and without bound drug (PDB: 1KLL and 1KMZ, respectively) are well superimposed with r.m.s.d. of 0.22 Å for 60 Cα pairs (Figure 4.24C). The major difference between these two structures is the connecting loop (residues 64–71) between the N-terminal and the C-terminal domains. In the MRD structure with bound

ligands, these residues form an extended loop that lies along with another similar loop from adjacent subunit in an antiparallel fashion, forming six hydrogen bonds that hold these two loops together. In the MRD structure without bound ligands, this loop is flexible and forms a turn that connects to the nearby β-strand of the C-terminal domain (β5) of the same subunit. The binding pocket for the compound 1, 2-*cis*-1-hydroxy-2, 7-DAMS locates at the bottom of the cleft. This compound lies flat in the binding pocket and its hydroquinone moiety is stabilized by the π-stacking interaction with the imidazole group from His^{38} and an indole ring from Trp^{108} of MRD (Figure 4.25). It was found that the dissociation constant of 1-hydroxy-2, 7-DAMS is lower than that of MC, suggesting a stronger binding of the former compound to the protein [214]. Due to the difference in dissociation constant, it was suggested that the aziridine ring and methoxide moiety in MC are unnecessary for effective binding. Other interactions of drug and MRD include the hydrogen bonding between the terminal carbonyl oxygen of the drug and the amine side chain of Gln^{110} of the protein. The hydroxyl side chain of Tyr^{112} also interacts with the terminal amine nitrogen and hydroxyl oxygen that attaches to C1 of DAMS (Figure 4.20B). This amino acid is significant for drug binding such that the mutagenesis of Y112F causes the dissociation constant of MC to drop (0.4 mM), thus generating a weaker binding to the protein [214]. Both the hydroxyl side chain of Tyr^{60} and the carboxyl side chain of Asp^{52} interact with the terminal amine and the amine nitrogen that attaches to C2 of DAMS.

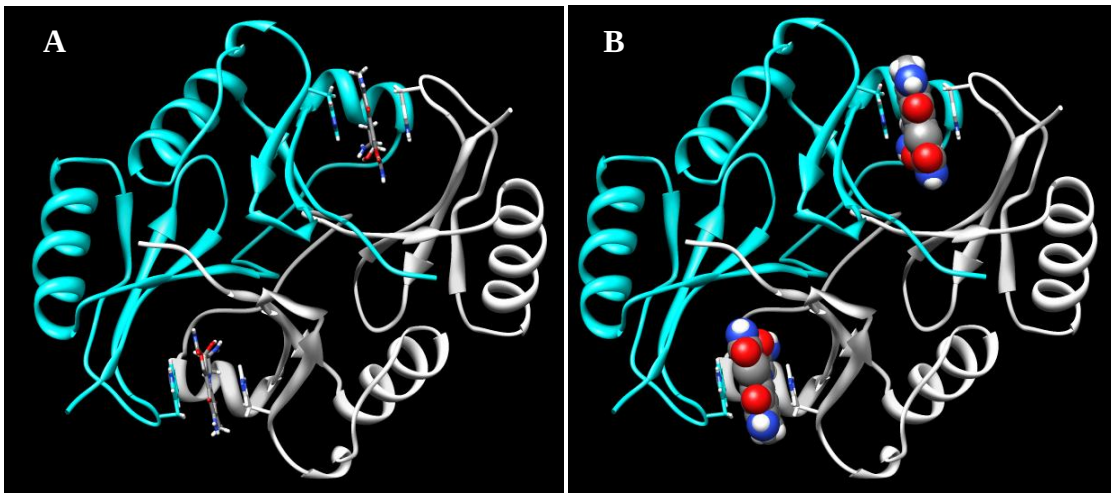


Figure 4.25: The ribbon structures of *Streptomyces lavendulae* mitomycin-binding protein (MRD) with bound 1, 2-*cis*-1-hydroxy-2, 7-diaminomitosenes (PDB: 1KLL) showing the drugs in **(A)** stick and **(B)** space fill, which lie in between aromatic residues (His^{38} and Trp^{108}) of MRD in the binding pockets. Each color represents one subunit. These 3D strcutures were generated by UCSF Chimera program (University of California, San Francisco) [1].

It was also found that the mitomycin-binding protein from *Streptomyces caespitosus*, the mitomycin producing organism, is able to accept bleomycin in its binding pocket [195]. This *Streptomyces caespitosus* MRD shares the exact same amino acid sequence as the *Streptomyces lavendulae* MRD. Instead of binding two molecules of BM like the bleomycin-binding protein, MRD is capable of binding to only one molecule of BM [195]. The X-ray crystallographic structures of Cu(II)-BM bound *S. caespitosus* MRD (PDB: 2A4W) and 1, 2-*cis*-1-hydroxy-2, 7-DAMS bound *S. lavendulae* MRD are well superimposed with an r.m.s.d. of 0.53 Å for 125 Cα pairs. However, instead of laying along the groove and cavity like the drug in the bleomycin-binding protein, only the γ-aminopropyldimethylsulfonium moiety of the BM molecule projects down to the binding pocket of MRD, while the remaining part of the molecule positions itself above the protein (Figure 4.26A). The methylene group of the γ-aminopropyldimethylsulfonium moiety is stabilized between the indole ring of Trp^{108} and the imidazole ring of His^{38} of the protein. The positively charged γ-aminopropyldimethylsulfonium moiety forms electrostatic interactions with the negatively charged side chains of Glu^{40}, Asp^{52} and Asp^{124}. Other subunits (such as nitrogen atom from β-hydroxy histidine moiety, amide nitrogen of threonine moiety and oxygen atom from β-aminoalanine moiety of BM) interact with the protein *via* water molecules.

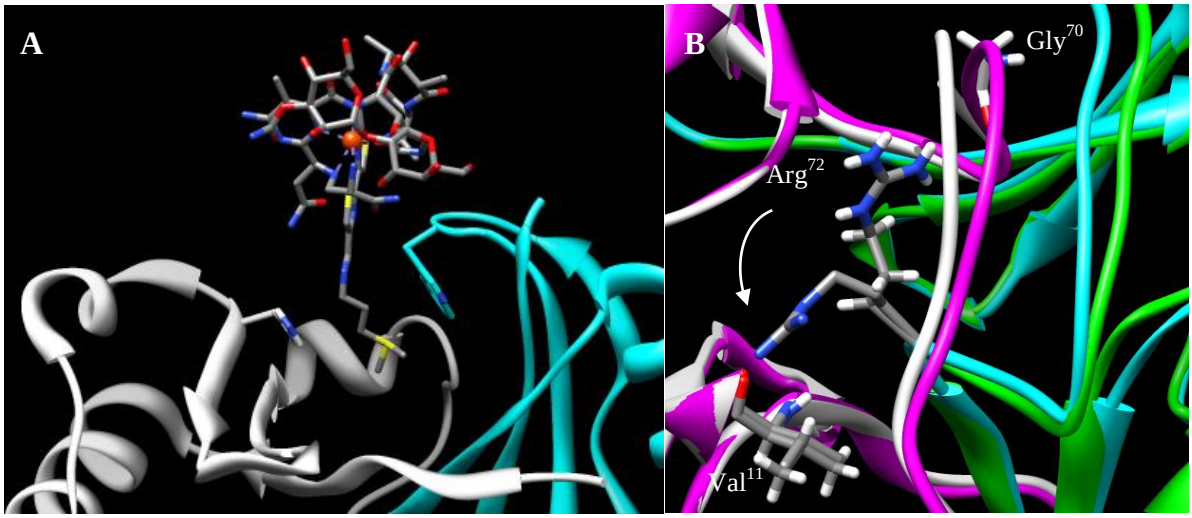


Figure 4.26: (A) The ribbon structure of mitomycin-binding protein from *Streptomyces caespitosus* that binds bleomycin (stick) in its binding pocket (PDB: 2A4W). Only the methylene unit of the bleomycin molecule projects into the binding cleft and is stabilized by Trp^{108} and His^{38} of the protein. **(B)** The superimposed structures of Cu(II)-Bm bound MRD (PDB: 2A4W) and 1, 2-*cis*-1-hydroxy-2, 7-DAMS bound MRD (PDB: 1KLL) show the movement of the side chain of Arg^{72} in Cu(II)-Bm bound MRD to form interactions with the main chain carbonyl oxygen of Val^{11}. These 3D strcutures were generated by UCSF Chimera program (University of California, San Francisco) [1].

The binding of only one BM molecule in MRD produces two asymmetric binding pockets with r.m.s.d. of 0.84 Å for 128 Cα pairs between two subunits (PDB: 2A4W). The superimposed structures of Cu(II)-BM bound MRD (PDB: 2A4W) and 1, 2-*cis*-1-hydroxy-2, 7-DAMS bound MRD (PDB: 1KLL) clearly indicate the major shifts of the extended loops (residues 54–73) that connect the N-terminal and the C-terminal domains and the connecting loop (residues 117–121) between β7 and β8. The major shifts are observed in subunit B, while the Cu(II)-BM is bound to binding pocket I of MRD that is formed by the N-terminal domain of subunit B and the C-terminal domain of subunit A and the unoccupied binding pocket II is formed by the N-terminal domain of subunit A and the C-terminal domain of subunit B. The residue Tyr^{60} in the occupied binding pocket of Cu(II)-BM bound MRD moves toward the core and interacts with His^{71} (the distance of hydroxyl oxygen of Tyr^{60} and side chain $N_{\varepsilon 2}$ of His^{71} is 5 Å) compared to those residues in the structure of the 1, 2-*cis*-1-hydroxy-2, 7-DAMS bound MRD (distance of 5.95 Å for both binding pockets), while these residues in the empty pocket of MRD do not cause significant shifts in their side chains (distance of 5.34 Å). This interaction also causes the rest of the extended loop (residues 54–73) in subunit B to move closer to its connecting loop (residues 117–121) in the same subunit. On the other hand, the extended loop in subunit A shifts away from its connecting loop. In the 1, 2-*cis*-1-hydroxy-2,

7-DAMS bound MRD structure, the side chain amine of Arg^{72} from one subunit is able to interact with the carbonyl oxygen of Gly^{70} from another subunit. The same interaction is applied for Arg^{72} of subunit A and Gly^{70} from subunit B in the Cu(II)-BM bound MRD structure. However, the side chain of Arg^{72} of subunit B swings away from Gly^{70} of subunit A and forms an interaction with the carbonyl oxygen of Val^{11} of subunit A instead (Figure 4.26B). This causes a shift in the C-terminal domain of subunit B (that forms half of the unoccupied binding pocket), resulting in a wider pocket that reduces interaction between the Cu(II)-BM and MRD. Thus, these conformational changes of the MRD lead to only one bound molecule of Cu(II)-BM.

Interestingly, the superimposed structures of *E. coli* GlxI with bound hydroxamate analogue (TSI, unpublished data) and bleomycin-binding protein from *Streptomyces verticillus* (BMLA) with two molecules of Cu(II)-bleomycin complex in its binding pockets (PDB: 1JIF) are overlapped with r.m.s.d. of 1.17 Å for 43 Cα atoms and allow us to compare the location of the GSH moiety and BM in both proteins. The hydroxamate inhibitor of GlxI could fit into the binding pocket of BMLA, which seems to be larger than the catalytic pocket of GlxI due to its natural larger substrate, bleomycin (Figure 4.27A). In turn, the Cu(II)-bleomycin complex could also lie along the groove of GlxI, where the head of bleomycin (β-aminoalanine, pyrimidinyl propionamide, β-hydroxy histidine, α-L-gulose, α-D-mannose, threonine and methylvalerae groups) is located in the catalytic pocket and the terminal bisthiazole and γ-aminopropyldimethylsulfonium groups lie along the active site and exist at one end of the catalytic pocket (Figure 4.27C). This long tail might interfere with the C-terminus of GlxI, which partially covers the catalytic pocket from outer solvent. On the other hand, the superimposed structures of *E. coli* GlxI with bound hydroxamate inhibitor (unpublished data) and *Streptomyces lavendulae* mitomycin-binding protein (MRD) with bound 1, 2-*cis*-1-hydroxy-2, 7-diaminomitosenes (PDB: 1KLL) are well overlapped with r.m.s.d. of 0.84 Å for 48 Cα atoms. The smaller molecule of mitomycin analogue seems to fit better into the catalytic pocket of GlxI than a larger size of BM (Figure 4.27D). However, this MC drug molecule is larger than the hemithioacetal, a substrate of GlxI, due to its aromatic structure and might interfere with the loops that act as lids of the catalytic pocket in GlxI. As well, the glycine group of GSH moiety might also interfere with a two-turned α-helix that locates at the entrance of the binding pocket of MRD (similar structure as methylmalonyl-CoA epimerase

with predicted GSH moiety, Chapter 5) (Figure 4.27B). Nevertheless, these proteins share high structural similarity, which is believed to be the result of evolution within the same superfamily of proteins.

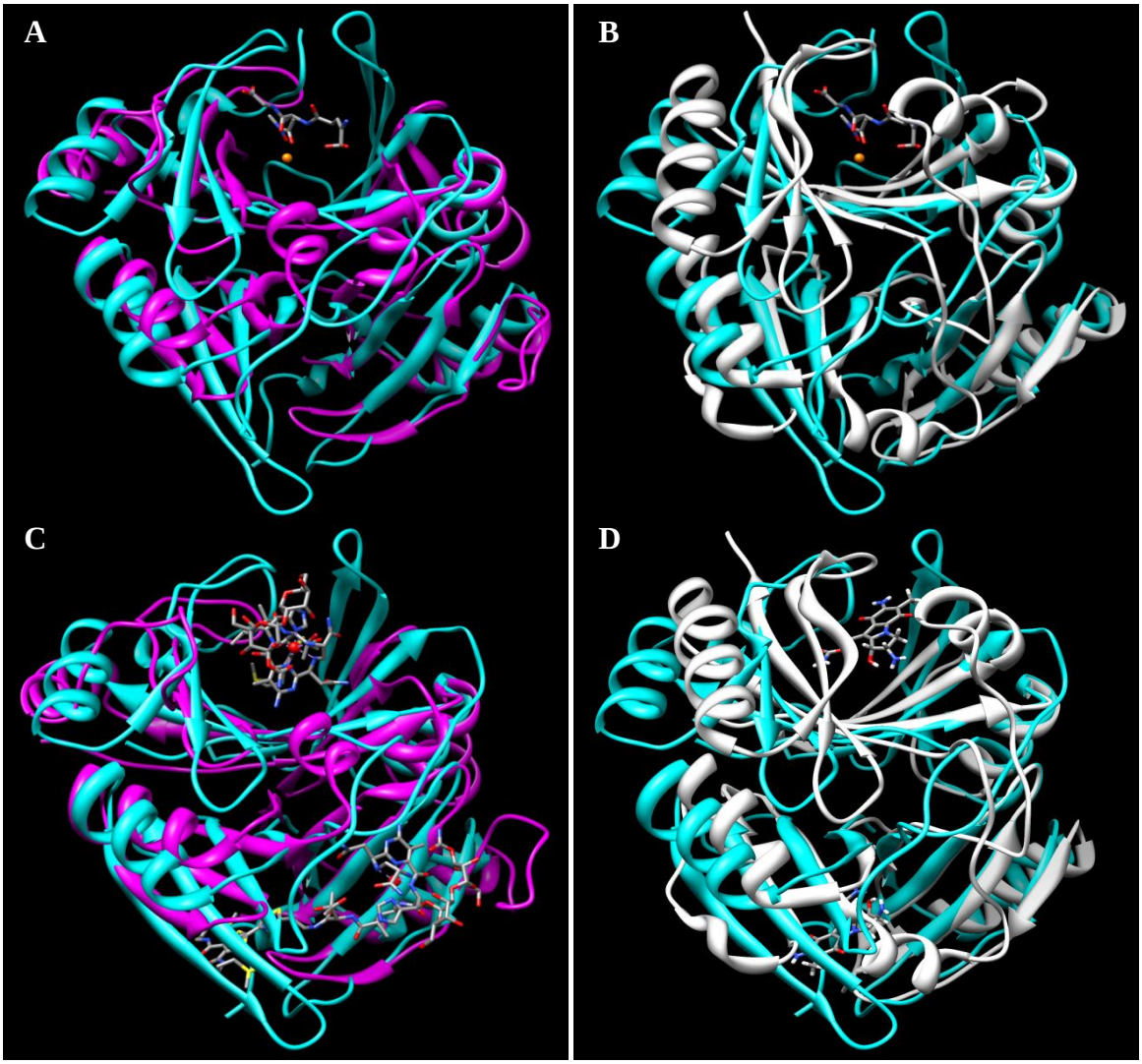


Figure 4.27: The superimposed structures of *E. coli* GlxI (cyan, unpublished data) and bleomycin-binding protein from *Streptomyces verticillus* (BMLA, magenta, PDB: 1JIF) show (**A**) bound hydroxamate analogue (TSI, an inhibitor of GlxI) and (**C**) two molecules of Cu(II)-bleomycin complex in the binding pockets. The metal in GlxI, a Ni^{2+} atom, is shown in orange and the Cu^{2+} atoms in metal-bleomycin complex are shown in red. The superimposed structured of *E. coli* GlxI (cyan, unpublished data) and *Streptomyces lavendulae* mitomycin-binding protein (MRD, white, PDB: 1KLL) show (**B**) bound hydroxamate analogue (TSI, an inhibitor of GlxI) and (**D**) bound 1, 2-*cis*-1-hydroxy-2, 7-diaminomitosenes in the binding pockets. These 3D strcutures were generated by UCSF Chimera program (University of California, San Francisco) [1].

4.5: Conclusions

The mutagenesis of *E. coli* GlxI-E56A was generated by PCR and the protein was purified by anion exchange chromatography and gel permeation chromatography. The secondary and quaternary structure of the mutant that was examined by CD analysis and gel permeation chromatography, respectively, was unchanged compared to that of the wild type enzyme [174]. The mutant showed a decrease in GlxI activity (3.4% of that of wild type enzyme). Its metal activation profile, however, was the same as that of wild type *E. coli* GlxI, where the enzymatic activity was observed with Ni^{2+}/Co^{2+}- reconstitution [19]. The metal titration experiment suggested that the mutant required at least 10 equivalents of metal in order to reach its maximum activity. This metal per enzyme ratio was ten times higher than the metal concentration required by the wild type enzyme to achieve its optimal activity [19]. The fosfomycin resistance activity of this mutant, however, was undetectable by using a DTNB assay. Additionally, the identification of the GS-fosfomycin adduct by ESI-MS after the assay reaction was purified by Sephadex G-10 column was within experimental background.

Due to the high structural similarity of GlxI and FosA, it is possible that GlxI might be a tetradentated progenitor of the tridentated FosA. However, upon evolution, loss of one of the four metal binding ligands as well as genetic mutations provided further functional diversification by opening an additional coordination site to support the particular reaction. The structural comparison between *E. coli* GlxI and *Pseudomonas aeruginosa* FosA indicates that the weaker K^+ and fosfomycin binding sites as well as the lack of significant residues for catalysis (such as Tyr^{39} in FosA) might be a crucial problem for the lack of readily observable fosfomycin conjugating activity with GlxI-E56A. Additional mutagenesis experiments to create K^+ and fosfomycin binding sites as well as the conversion of Phe^{43} in *E. coli* GlxI to Tyr might increase fosfomycin resistance in GlxI.

Other resistance proteins in the $\beta\alpha\beta\beta\beta$ superfamily are bleomycin and mitomycin resistance proteins. Unlike fosfomycin-binding protein, these two proteins are neither enzymes nor metalloproteins. They are capable of accepting two molecules of drug in their relatively hydrophobic binding pockets (two binding pockets per dimer), thus inactivating the drug and resulting in resistance to these drugs by the antibiotic-producing organisms. The difference in flexibility of two subunits in the protein creates two asymmetric binding pockets. The binding of the first drug molecule in the more flexible binding pocket of the protein appears to

introduce conformational changes in the other binding pocket such that this second binding pocket is able to bind tighter to the second incoming drug molecule (cooperative binding pockets). Since bleomycin- and mitomycin-binding proteins share high structural similarity, they are able to functionally replace/substitute one another. This is an advantage for microorganisms such that they can neutralize the effect of several antibiotics or drugs using the very same effective protein (multidrug resistance). It remains to be seen how recombinant DNA techniques might be used to introduce drug-binding activities to GlxI. To our knowledge, this has not been explored by other laboratories.

CHAPTER 5: SIMILARITY WITHIN THE βαβββ SUPERFAMILY: THE INVESTIGATION OF METHYLMALONYL-COENZYME A EPIMERASE ACTIVITY BY GLYOXALASE I

5.1: Introduction

Methylmalonyl-coenzyme A epimerase is a member of the βαβββ superfamily [30, 40]. The proteins in this superfamily including Glyoxalase I (GlxI), methylmalonyl-CoA epimerase (MMCE), extradiol dioxygenase, fosfomycin resistance protein, mitomycin resistance protein and bleomycin resistance protein share low sequence similarity and identity but similar structural arrangements (Figure 5.1 and 5.3). Most proteins in this superfamily form a dimeric domain-domain swapping molecular structure. Some GlxI (yeast and *Plasmodium falciparum* GlxI) and extradiol dioxygenase can also form a larger monomeric four-domain arrangement [39, 145, 151]. Among these proteins, MMCE shows significant structural and functional similarities to GlxI, especially to GlxI from *Clostridium acetobutylicum* (Chapter 2). Based on multiple sequence alignments, a phylogenic tree suggests a closer relationship of CLO GlxI with MMCE than with GlxI that catalyze the same enzymatic reaction (Figure 5.2). These structural similarities lead to a common reaction mechanism for MMCE and GlxI, where the former functions as an epimerase using the thioester methylmalonyl-CoA and the latter accomplishes an isomerization reaction upon the hemithioacetal of 2-oxoaldehydes as substrates (Figure 5.4). The activities of both enzymes also depend on the presence of divalent metals.

The biological function of methylmalonyl-CoA epimerase is to take part in the conversion of the three-carbon propionyl-CoA molecule, formed from degradation of branched amino acids and odd chain fatty acids, into the four-carbon succinyl-CoA molecule (Figure 5.5). The individual steps involved in this conversion process are as follows. First, propionyl-CoA is converted to (*2S*)-methylmalonyl-CoA by the enzyme propionyl-CoA carboxylase in an ATP-dependent reaction. The stereospecific enzyme, methylmalonyl-CoA epimerase (racemase) then converts this (*2S*)-methylmalonyl-CoA to its *R* form, which is subsequently converted to succinyl-CoA by methylmalonyl-CoA mutase, a coenzyme B12 dependent-

enzyme. These last two reactions are reversible. Succinyl-CoA then enters central metabolism *via* the citric acid cycle.

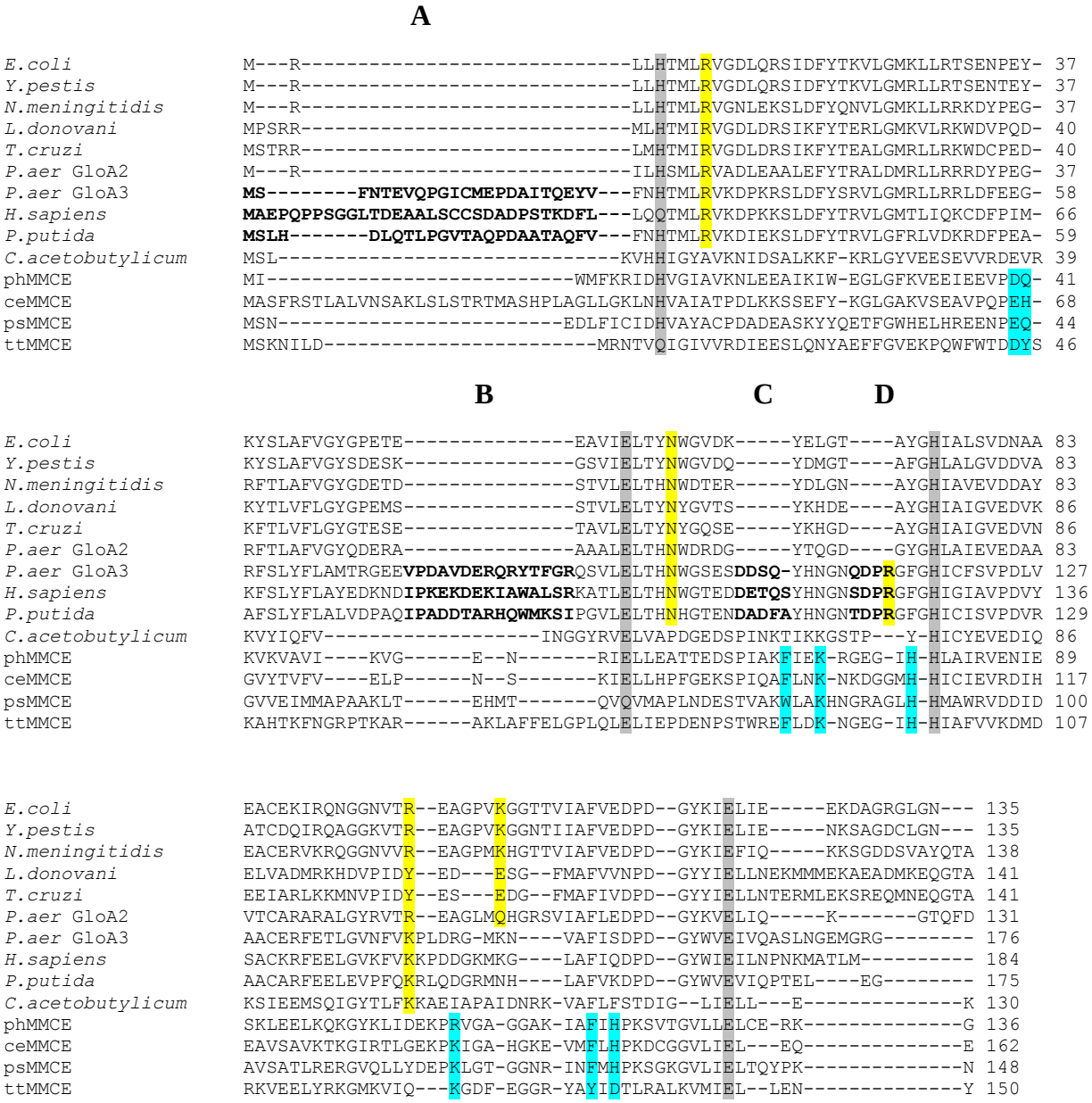


Figure 5.1: Multiple sequence alignments of Glyoxalase I (GlxI) from various organisms (organism name follows by National Center for Biotechnology Information (NCBI) accession number) including *E. coli* (NP_310387), *Y. pestis* (ZP_01887743), *N. meningitidis* (CAA74673), *L. donovani* (AAU87880), *T. cruzi* (XP_818456), *P. aeruginosa* GloA2 (AAG04099), *P. aeruginosa* GloA3 (AAG08496), *H. sapiens* (AAB49495), *P. putida* (AAN69360) and *C. acetobutylicum* (AAK80149) and methylmalonyl-CoA epimerase (MMCE) from *Pyrococcus horikoshii* (AAK52053, phMMCE), *Caenorhabditis elegans* (AAT92095, ceMMCE) and *Propionibacterium shermanii* (AAL57846, psMMCE) as well as putative MMCE from *Thermoanaerobacter tengcongensis* (TTE0360, ttMMCE). The metal binding residues for both enzymes are highlighted in grey. The yellow highlight indicates conserved residues in GlxI for stabilizing its substrate, the hemithioacetal. The cyan highlight

emphasizes putative residues that possibly bind to aromatic and negatively charged groups of methylmalonyl CoA, the substrate of MMCE. The loop regions that only exist in Zn^{2+} activated GlxI (and not in Ni^{2+}/Co^{2+}-activated GlxI) are in bold letter and marked with capital letters (A, B, C and D). The alignment was created using CLC Free Workbench (version 3.0.1) with the accurate alignment algorithm (http://www.clcbio.com).

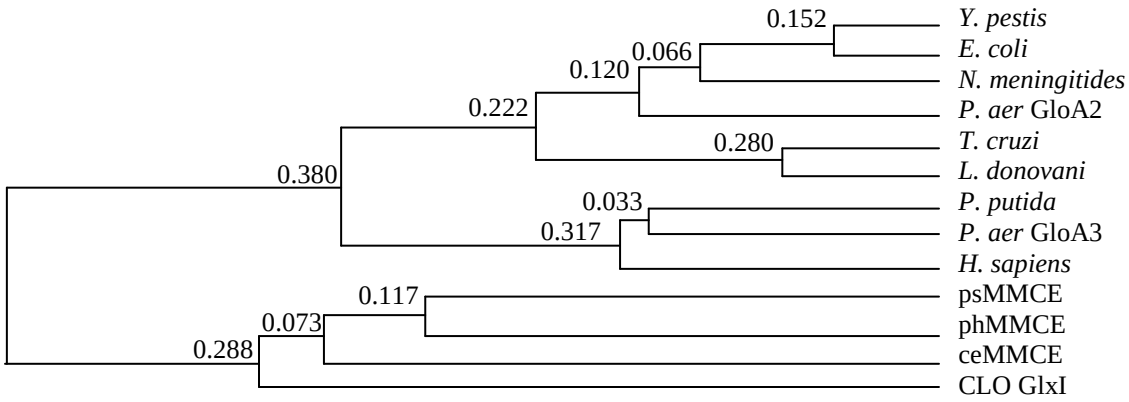


Figure 5.2: The phylogenic tree of GlyoxalaseI and methylmalonyl-CoA epimerase generated from the multiple sequence alignments shown in Figure 5.1. The amino acid sequence of GlxI from *Clostridium acetobutylicum* (CLO GlxI) was found to be closely related to methylmalonyl-CoA epimerase. The phylogeny was created using CLC Free Workbench (version 3.0.1) with UPGMA algorithm (http://www.clcbio.com).

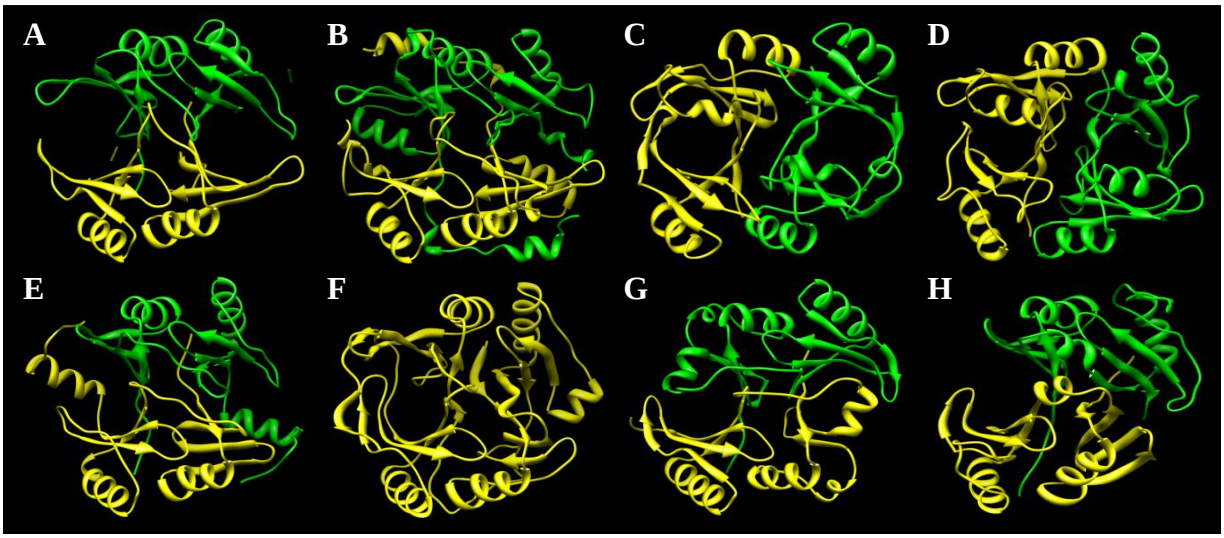


Figure 5.3: The ribbon structures of the proteins in the βαββββ superfamily including **(A)** *E. coli* GlxI (PDB: 1F9Z), **(B)** *H. sapiens* GlxI (PDB: 1QIN), **(C)** *C. acetobutylicum* GlxI (PDB: 2QH0), **(D)** *P. shermanii* methylmalonyl-CoA epimerase (PDB: 1JC5), **(E)** *Pseudomonas aeruginosa* fosfomycin resistance protein (PDB: 1LQO), **(F)** *Pseudomonas* biphenyl-cleaving extradiol dioxygenase (PDB: 1HAN), **(G)** *Streptomyxes caespitosus* mitomycin resistance protein (PDB: 2A4X) and **(H)** *Streptomyces verticillus* bleomycin resistance protein (PDB: 1JIF). Subunit A is labeled as yellow and subunit B is labeled as green. The 3D structures were generated by using the UCSF Chimera program (University of California, San Francisco) [1].

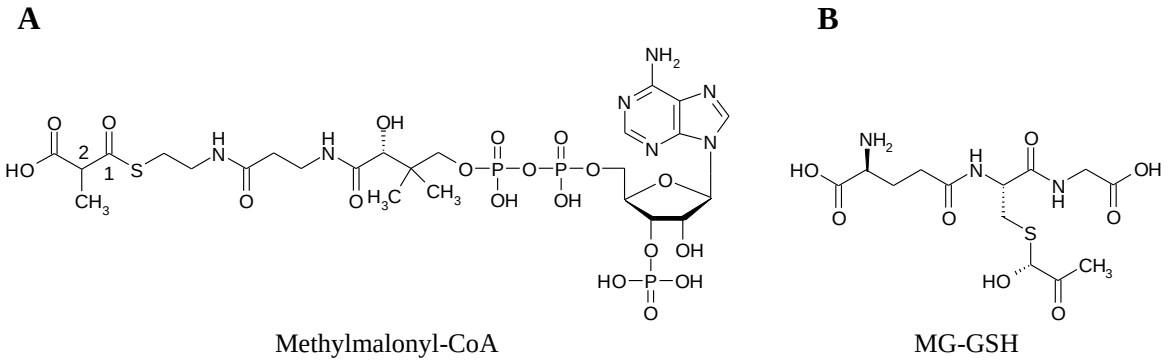


Figure 5.4: (A) The substrate for methylmalonyl-CoA epimerase, methylmalonyl-CoA and **(B)** the substrate for Glyoxalase I, hemithioacetal of MG-GSH. The numbers on methylmalonyl-CoA indicate the significant carbon atoms that involve proton abstraction and transfer in reaction mechanism.

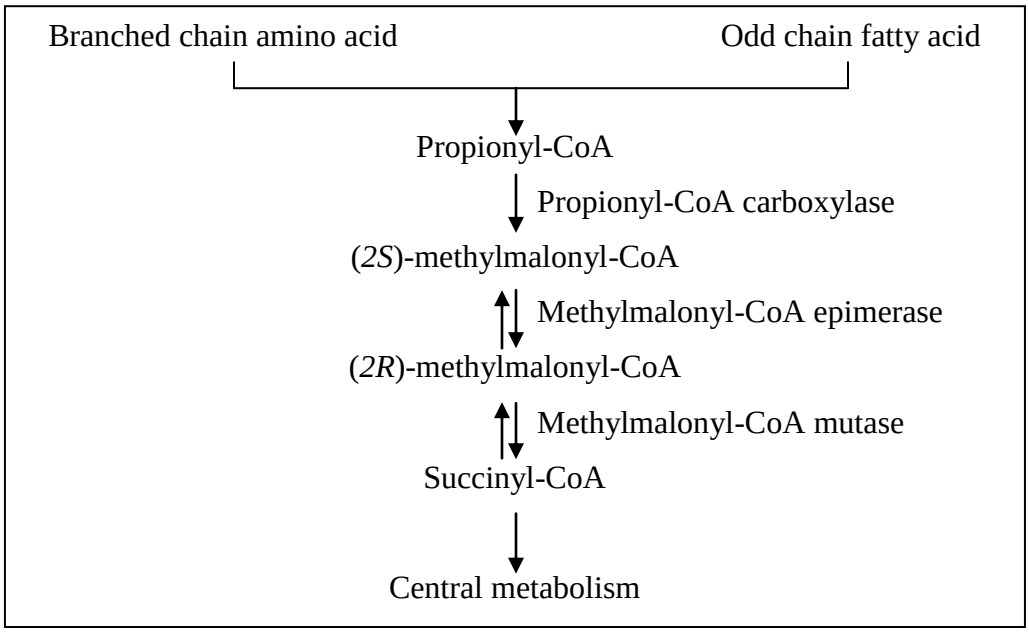


Figure 5.5: The pathway of propionyl-CoA to succinyl-CoA involves three reaction mechanisms with three enzymes, propionyl-CoA carboxylase, methylmalonyl-CoA epimerase and methylmalonyl-CoA mutase. The last two reactions are reversible. Succinyl-CoA then enters central metabolism *via* the citric acid cycle.

Methylmalonyl-CoA epimerase has been characterized from many organisms including humans, animals (rat and sheep), the nematode *Caenorhabditis elegans*, the archaea *Pyrococcus horikoshii* and Gram-positive bacteria (*Streptomyces coelicolor* and *Propionibacterium shermanii*) with *P. shermanii* MMCE being the best characterized [105, 106, 108, 215-219]. MMCE is a metalloenzyme that possesses maximum activity in the presence of Co^{2+} [216, 217]. It has been reported that *P. shermanii* MMCE is activated by several divalent metals including Co^{2+}, Mn^{2+}, Ni^{2+}, Mg^{2+}, Zn^{2+}, Cu^{2+}, Cd^{2+} (the ranking of the

activity of the enzyme being $Co^{2+} > Mn^{2+} > Ni^{2+} > Mg^{2+} \sim Zn^{2+} > Cu^{2+} > Cd^{2+}$) [216]. MMCE from rat is also activated by the presence of Co^{2+}, Co^{3+}, Fe^{2+} and Mn^{2+} (high concentration of Co^{3+}, Fe^{2+} and Mn^{2+} (10 mM) is required to detect the MMCE activity, while only 0.2 μM Co^{2+} is enough for the MMCE assay); however, no activity is observed in the presence of Zn^{2+}, Cu^{2+}, Cu^{1+} and Cd^{2+} [217]. Compared to GlxI that can be divided into two classes according to metal activation; a Zn^{2+}-activation and a Ni^{2+}/Co^{2+}-activation (non Zn^{2+}-activation) class; there is no evidence that MMCE falls into the same category. However, broader investigation on metal specificity of MMCE is certainly required.

MMCE is a homodimeric enzyme with a molecular weight of approximately 30 kDa [105]. The dimeric interactions between subunits are mainly hydrophobic and hydrogen bonding in nature with no disulfide bond formation being observed [105]. Maximum activity of this enzyme is achieved by reconstitution with 1.5 mol Co^{2+} per mol of dimeric enzyme [216]. It was suggested that the metal per enzyme ratio is actually 2, and fully occupied enzyme was not observed because of the short experimental incubation time (one hour) [216]. Higher metal concentrations (10–100 μM Co^{2+}) did not cause significant changes in activity however [216]. Compared to the metal titration profile that exhibited by *E. coli* GlxI, where the metal per dimeric enzyme ratio is approximately 1 [19], it is also possible that MMCE may exhibit optimum activity in the presence of only one metal per dimeric enzyme and therefore with only one active site being functional. Higher metal per enzyme ratios required for activation of MMCE as observed previously might be in error due to the specific process used to prepare the apo-enzyme. Detailed experiments will be needed to test this hypothesis.

The multiple sequence alignments have shown that MMCE contains four conserved metal binding residues (His^{12}, Gln^{65}, His^{91} and Glu^{141} for *P. shermanii* MMCE) similar to those of GlxI (His^{5}, Glu^{56}, His^{74} and Glu^{122} for *E. coli* GlxI) (Figure 5.1). The X-ray crystallographic structure of *P. shermanii* MMCE superimposes very well on CLO GlxI, a Ni^{2+}/Co^{2+}-activated enzyme. Similar to GlxI, the active site geometry of MMCE with active metal forms an octahedral coordination with four metal binding residues and two water molecules. Both MMCE and CLO GlxI have an active site that forms within one subunit (two active sites per dimeric enzyme) (Figure 5.3). This type of domain swapping is characteristic of MMCE and has been used to predict putative MMCE from other proteins in the βαββββ superfamily [30,

40]. Once CLO GlxI was discovered, it was suggested that a common protein structure is not enough to predict its biological function.

The reaction mechanism of MMCE is compatible with that of GlxI, which is proposed to involve a two-base catalysis [40, 105]. First, the two oxygens of C1 and C2 from (*2S*)-methylmalonyl-CoA replace two water molecules from the metal-enzyme complex in the catalytic site (Figure 5.4A). One metal binding ligand (Glu^{141} in *P. shermanii* MMCE) that acts as a base then abstracts the C2 proton from one side of the substrate, leaving the stereochemistry at C2 to invert and form an anionic intermediate that is stabilized by the metal. Then, another metal binding ligand (Gln^{65} in *P. shermanii* MMCE) is believed to donate a proton to C2 from the opposite site, resulting in the *R*-epimer of methylmalonyl-CoA. This proposed reaction mechanism, however, requires more experimental evidence.

Due to the structural and functional similarities between MMCE and GlxI, it was possible that one enzyme might be able to act as a bifunctional enzyme, having both MMCE and GlxI activities. Several GlxI with different metal activation profiles (Zn^{2+}-activation and non Zn^{2+}-activation) and sizes (monomeric or dimeric four-domain GlxI) from various organisms that have already been characterized in our laboratory were examined to determine if they also possess MMCE activity using (*2S*)-methylmalonyl-CoA as a substrate. The structural comparison between MMCE and GlxI suggested significant features that are possibly necessary for substrate selection including the binding space for the substrate within the active site and the stability of the bound substrate. Due to limitations of available X-ray crystallographic structures, three GlxI including human GlxI (PBD: 1QIN and 1QIP), *E. coli* GlxI (PDB: 1F9Z and 1FA8) and Clostridial GlxI (PDB: 2QH0 and 3HDP) were used as the examples for structural comparison with MMCE from *P. shermanii* (PDB: 1JC5). The evolution of the proteins in the $\beta\alpha\beta\beta\beta$ superfamily was also proposed and discussed in detail. This is the first investigation of cross functional activity on the enzymes within this same superfamily.

5.2: Reagents, Materials and Instrumentation

All reagents, materials and equipment used in this chapter are stated below otherwise they are included in the previous chapters. Methylmalonyl-CoA tetralithium salt hydrate

(assay, ≥90%), coenzyme B12 (assay, ≥97%) and *Saccharomyces cerevisiae* Glyoxalase I (grade IV, 1590 units per mg protein) were supplied from Sigma Aldrich (St. Louis, MO). The assay of methylmalonyl-CoA epimerase was performed on a high performance liquid chromatography (HPLC) using a Waters 625 LC system and a 994 programmable photodiode array detector with a μBondapak^{TM} C18 column (3.9 × 300 mm) from Waters Corporation (Chicago, IL).

5.3: Experimental Protocols

5.3.1: Protein Expression, Induction and Purification of Methylmalonyl-CoA Mutase

The bacterial strain *E. coli* BL21 (DE3) CodonPlus RIL containing the plasmid pTA925 expression vector with inserted methylmalonyl-CoA mutase (MCM) coding gene was kindly provided by Dr. Thomas Bobik from Iowa State University (IA). The cells were inoculated in LB (10 mL) containing kanamycin (30 μg/mL) at 37 °C overnight before transferring into 1 L LB broth. The bacterial culture was grown at 30 °C in a thermostated 250 rpm shaker until OD_{600} reached 0.6. Protein production was then induced by the addition of 0.5 mM IPTG followed by continued shaking at 30 °C for 8.5 hours. Cells were harvested by centrifugation at 6,000×*g* for 10 min. The cell pellet was resuspended with minimum wash buffer (20 mM Tris (pH 7.0), 500 mM KCl and 20 mM imidazole) and centrifuged at 6,000×*g* for another 10 min. The cell pellet was collected and flash frozen in liquid nitrogen before storing in a -80 °C freezer.

The protein purification of MCM was performed according to the previous report by Bobik and Rasche, 2003 [107], which is also similar to those of the His-tagged CLO GlxI enzyme purifications (Chapter 2). After eluting from the HisTrap affinity column utilizing a buffer containing 500 mM imidazole, the His-tagged MCM was dialyzed against 2 L of buffer containing 20 mM sodium phosphate buffer (pH 7.0) and 25 mM NaCl overnight at 4 °C. The identity of the purified protein was confirmed by SDS-PAGE and ESI-MS. The protein concentration was determined by the Bradford assay using BSA as a standard [134]. The holo-MCM was prepared by mixing MCM (0.56 mg/mL) with coenzyme B12 (63 μM) in buffer containing 10 mM HEPES (pH 7.0), 25 mM NaCl, 10 mM KCl and 2 mM DTT. The enzyme was stored in the dark at -80 °C.

5.3.2: Enzymatic Assay of Methylmalonyl-CoA Epimerase

The methylmalonyl-CoA epimerase assay was performed as reported by Bobik and Rasche, 2003 [107]. The commercial mixture of (*R*)- and (*S*)-methylmalonyl-CoA was prepared as a 1 mM stock in the assay buffer, which contained 50 mM KPB (pH 7.0), 25 mM NaCl and 2 mM $MgCl_2$. The MCM assay was performed by mixing 2.8 μg holo-MCM and 75 μM methylmalonyl-CoA in the assay buffer followed by incubation at 37 °C for 5 min. The purified enzymes, which had been incubated prior with 5 equivalents of metals overnight at 4 °C, were added to the MCM reaction and incubation was continued at 37 °C for another 5 min. The reaction was then quenched by adding 100 μL of 1 M acetic acid. The assay mixture was separated by a reverse phase HPLC with UV-detection at 260 nm and a μBondapAK^{TM} C18 column using a solvent system as shown in Table 5.1 (solvent A was 100 mM sodium acetate (pH 4.6) in 10% methanol in water and solvent B was 100 mM sodium acetate (pH 4.6) in 90% methanol in water). The methylmalonyl-CoA and succinyl-CoA were eluted with retention times (t_R) of 6 min and 7.3 min, respectively.

Table 5.1: The solvent gradient used for detecting MMCE assay. A = 100 mM sodium acetate (pH 4.6) in 10% methanol in water and B = 100 mM sodium acetate (pH 4.6) in 90% methanol in water.

Time (min)	*Flow rate (mL/min)*	*% A*	*% B*
0	1	100	0
10	1	100	0
22	1	40	60
23	1	0	100
35	1	0	100

5.3.3: Structural Investigation of Methylmalonyl-CoA Epimerase and Glyoxalase I

Structural visualizations and molecular images were generated by the UCSF Chimera package from the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco (supported by NIH P41 RR-01081) [1] and the PyMOL Molecular Graphics System, version 1.2r3pre, Schrödinger, LLC (http://www.pymol.org/).

Molecular docking for prediction of methylmalonyl-CoA binding poses in the active site of *P. shermanii* MMCE (PDB: 1JC5) was performed by Molegro Virtual Docker version MVD 2010.4.1.0; Molegro ApS, Denmark [220]. The substrate (*2R*)-methylmalonyl-CoA was drawn using the program Spartan '08 (version 1.2.2, Wavefunction, Inc, California). The structure was then geometry optimized in Spartan ’08 using the molecular mechanics MMFF94

force field [221-227] as implemented by Wavefunction, Inc. The docking wizard was utilized in Molegro to set up the docking experiment. The default parameters were used and MolDock Score (Grid) was chosen as the scoring function. Grid resolution was 0.3 Å. Ligand evaluation was calculated by internal energy score. Binding site was predicted by a user-defined constraint based on the active site cleft, which had a radius of 19 Å, including the active site groove. The Search Algorithm used was MolDock SE and the number of runs was 10. Docking poses were constrained to the user-defined cavity. After docking, the poses were energy minimized and H-bonds were optimized. Maximum iterations were set at the default of 1500 with a maximum population size default of 50. Energy threshold was the default of 100. Simplex evolution was set at the default of 300 for maximum steps and 1.00 for neighbour distance factor. Ten maximum poses were returned for each of the ten runs. Similar poses were clustered and the program was set to ignore similar poses. No error warnings were detected in any of the docking runs.

5.4: Results and Discussion

5.4.1: Protein Purification of Methylmalonyl-CoA Mutase

Methylmalonyl-CoA epimerase from *E. coli* (NP_417392) was generated and cloned into *E. coli* BL21 (DE3) CodonPlus RIL by Bobik and Rasche, 2003 [107]. The enzyme was engineered to contain an N-terminal His_8-tag for protein purification purposes. This modification changes the molecular weight of the enzyme from 77871.2 Da (714 amino acids) to 79099.5 Da (723 amino acids) and changes the predicted p*I* of 5.05 to 5.31 (Figure 5.6). The His_8-MCM was purified using the facile one-step purification, HisTrap affinity column approach (Figure 5.7). The cell pellet of ~4 g per 1 L culture was harvested and the final 37 mg of purified protein was recovered. The existence of the purified protein was indicated in SDS-PAGE and ESI-MS with no post translational modification being observed (Figure 5.8).

MHHHHHHHHMSNVQEWQQLANKELSRREKTVDSLVHQTAEGIAIKPLYTEADLDNLEVTGTLPGLPPYV
RGPRATMYTAQPWTIRQYAGFSTAKESNAFYRRNLAAGQKGLSVAFDLATHRGYDSDNPRVAGDVGKAGV
AIDTVEDMKVLFDQIPLDKMSVSMTMNGAVLPVLAFYIVAAEEQGVTPDKLTGTIQNDILKEYLCRNTYI
YPPKPSMRIIADIIAWCSGNMPRFNTISISGYHMGEAGANCVQQVAFTLADGIEYIKAAISAGLKIDDFA
PRLSFFFGIGMDLFMNVAMLRAARYLWSEAVSGFGAQDPKSLALRTHCQTSGWSLTEQDPYNNVIRTTIE
ALAATLGGTQSLHTNAFDEALGLPTDFSARIARNTQIIIQEESELCRTVDPLAGSYYIESLTDQIVKQAR
AIIQQIDEAGGMAKAIEAGLPKRMIEEASAREQSLIDQGKRVIVGVNKYKLDHEDETDVLEIDNVMVRNE
QIASLERIRATRDDAAVTAALNALTHAAQHNENLLAAAVNAARVRATLGEISDALEVAFDRYLVPSQCVT
GVIAQSYHQSEKSASEFDAIVAQTEQFLADNGRRPRILIAKMGQDGHDRGAKVIASAYSDLGFDVDLSPM
FSTPEEIARLAVENDVHVVGASSLAAGHKTLIPELVEALKKWGREDICVVAGGVIPPQDYAFLQERGVAA
IYGPGTPMLDSVRDVLNLISQHHD

Figure 5.6: Amino acid sequence of His-tagged methylmalonyl-CoA mutase generated by Bobik and Rasche, 2003 [107]. The bold letters indicate the His_8-tag of the protein for protein purification purpose.

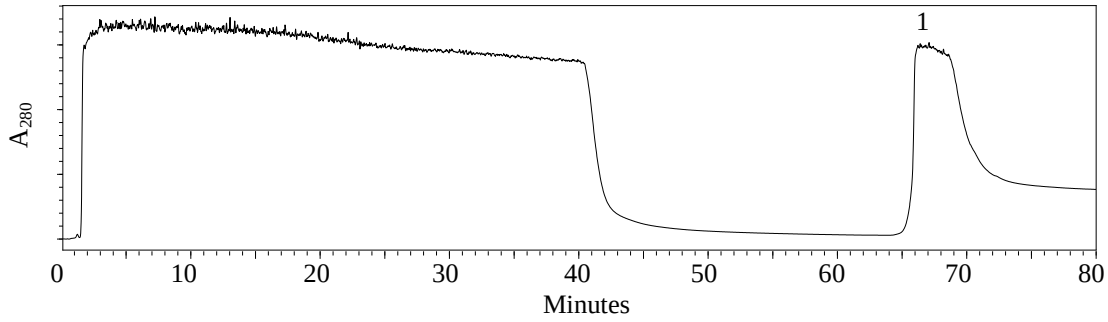


Figure 5.7: The HPLC chromatogram of the methylmalonyl-CoA mutase purification using the HisTrap HP affinity column with a flow rate of 0.5 mL/min and a buffer containing 20–500 mM imidazole. The His-tagged MCM was eluted with the buffer containing 500 mM imidazole and was labeled as a peak number 1 (65–73 min).

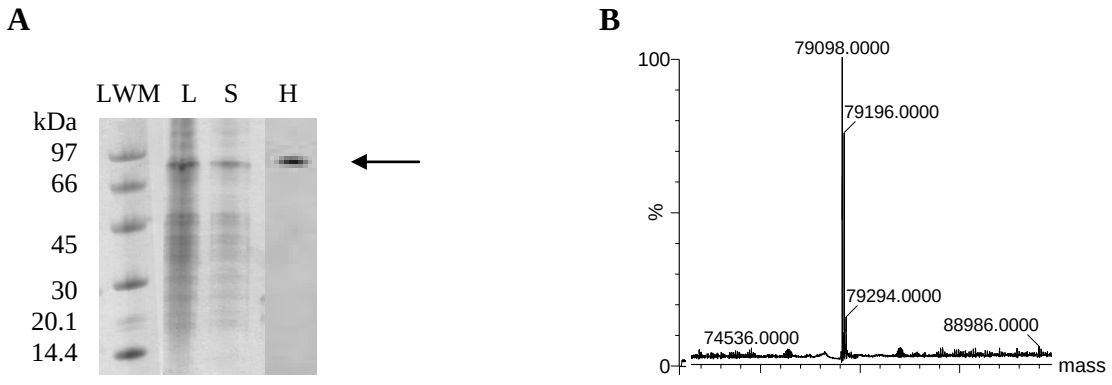


Figure 5.8: (A) The SDS–PAGE of the MCM purification: LMW = low molecular weight marker, L = lysate, S = supernatant after sonication and H = HisTrap eluted fraction. The arrow indicates the purified MCM. **(B)** The electrospray mass spectrum of MCM shows the molecular weight of a single subunit at 79098.0 Da (calculated MW is 79099.5 Da) and its adduct with $2K^+$ and Na^+ atoms at 79196.0 Da.

5.4.2: Methylmalonyl-CoA Epimerase Activity in Glyoxalase I

Structural similarity between MMCE and GlxI suggests a possible common catalytic reaction mechanism for their epimerization and isomerization reactions, respectively. Several GlxI enzymes including *Clostridium acetobutylicum* GlxI (Chapter 2), *Pseudomonas aeruginosa* GlxI (GloA2 and GloA3, Chapter 3) and commercial yeast GlxI were investigated to determine if they might possess epimerase activity. The coupled assay of MMCE and MCM was initiated by the reaction of MCM (2.8 μg) that converts (*2R*)-methylmalonyl-CoA in the commercial mixture (*RS*)-methylmalonyl-CoA (75 μM) to succinyl-CoA, leaving its *S*-form in the reaction mixture. This reaction is completed within 5 min at 37 °C [107]. Since the enzyme is inactive in acidic buffer (pH< 5), acetic acid was then used to quench the reaction [107]. The HPLC chromatogram of the MCM reaction exhibited the expected peaks for methylmalonyl-CoA and succinyl-CoA with t_R of 6 min and 7.3 min, respectively (Figure 5.9). The μBondapak^{TM} C18 column cannot be used to isolate the *S*- and *R*-epimers of methylmalonyl-CoA due to poor resolution; hence, both eluted at approximately the same retention time. Adding MMCE into this reaction mixture would convert (*2S*)-methylmalonyl-CoA to its *R*-form, which then subsequently would be converted to succinyl-CoA by MCM. If the enzyme of interest possessed MMCE enzyme activity, a decrease in the methylmalonyl-CoA peak and an increase in the succinyl-CoA peak would be observed.

It was hypothesized that the Clostridial GlxI (Ni^{2+}/Co^{2+}-activated GlxI), with significant structural similarity to MMCE, might possibly exhibit MMCE activity. Since MMCE was found to be most activated in the presence of Co^{2+} atom, Co^{2+}-reconstituted CLO GlxI was prepared and used in the coupled assay of MMCE and MCM. However, incubation of Co^{2+}-bound CLO GlxI (1 μg) for 5 min in the assay mixture did not show any detectable MMCE activity (Figure 5.9). Neither was there any activity observed when higher protein concentrations (10 μg) and longer incubation times (20 min) were employed. Ni^{2+}-reconstituted CLO GlxI (30 μg) was also examined, since Ni^{2+} is the most active metal that optimizes the Clostridial GlxI activity. However, no MMCE activity was detected. Compared to other MMCE enzymes and their published assay conditions (Table 5.2), the protein concentration and incubation time of CLO GlxI utilized in our studies should have been sufficient to detect MMCE activity (if any). However, none was observed, indicating that CLO GlxI possesses no epimerase activity under our studied experimental conditions.

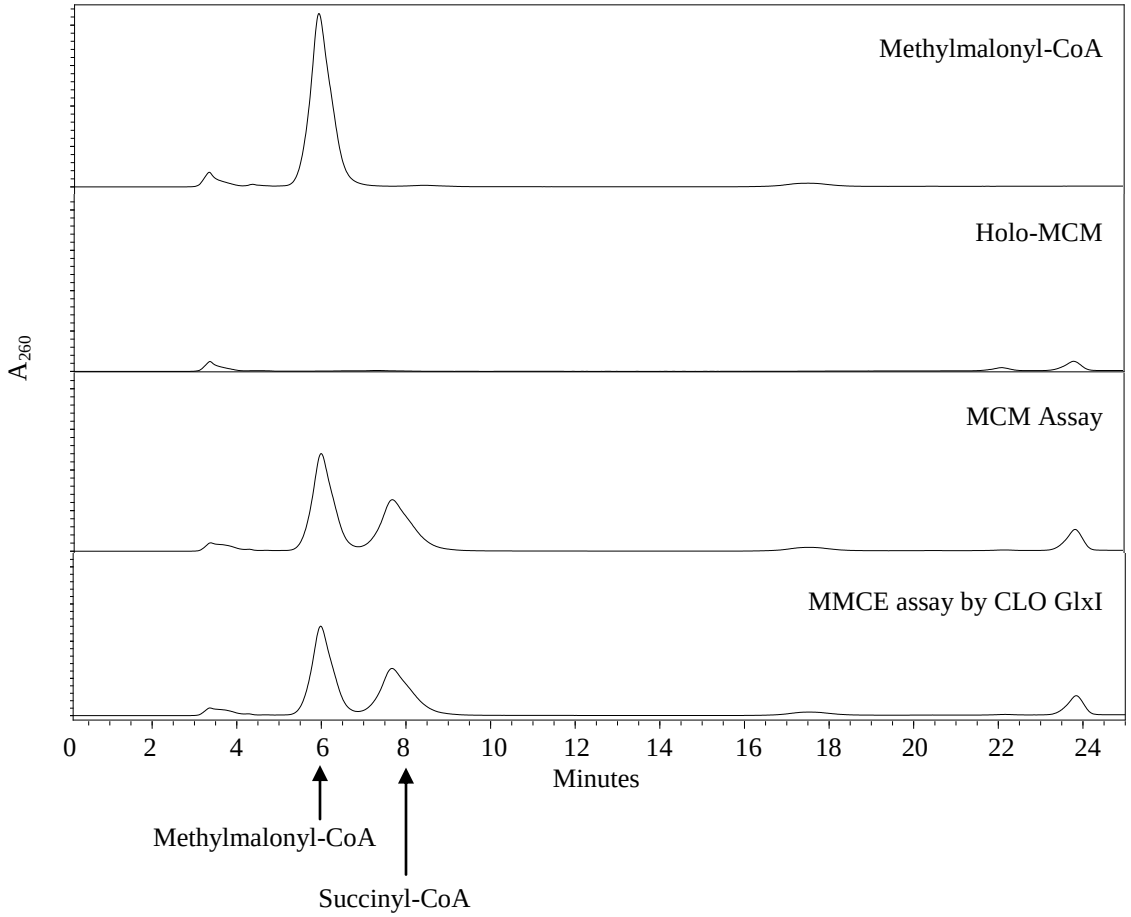


Figure 5.9: The reverse phase HPLC chromatograms of methylmalonyl-CoA epimerase activity of CLO GlxI (1 μg) measured at 260 nm using a μBondapak^{TM} C18 column for separation of substrates and products. The MCM-MMCE coupled assay was performed at 37 °C using commercial (*RS*)-methylmalonyl-CoA as substrate. Methylmalonyl-CoA mutase (MCM) converts (*2R*)-methylmalonyl-CoA to succinyl-CoA, leaving (*2S*)-methylmalonyl-CoA in the assay solution to be converted to its *R*-form by MMCE (if any).

Table 5.2: The reported kinetic parameters for methylmalonyl-CoA epimerase and amount of enzyme used from various organisms.

Source of MMCE	***Enzyme in assay (μg)***	***V_{max} (μmol/min/mg)***	***K_m (μM)***	***k_{cat} (s^{-1})***	***Ref.***
Pyrocuccus horikoshii	0.02	467	79	240	[108]
Propionobacterium shermanii	0.06	270	38	150	[107]
Staphylococcus coelicolor	0.058	140	57	75	[228]
Caenorhabditis elegans	0.007–0.016	191	-	121	[215]
Human	0.084	833	-	625	[106]
Rat liver	0.67	8400	100	4480	[217]

Zn^{2+}-activated GlxI from *P. aeruginosa*, GloA3, was chosen as a studied case to examine if it might exhibit MMCE activity. The metal-reconstituted form of this enzyme was

found to be most activated in the presence of Ni^{2+} in the GlxI assay (140% activity comparing to the activity of the isolated enzyme) as well as with Zn^{2+} (80% activity) [33]. GloA3 that was overexpressed in the *E. coli* system with the presence of $ZnCl_2$ (1 mM) was prepared by adding 5 equivalents of Zn^{2+} to the purified enzyme, followed by incubation overnight at 4 °C. The holo-enzyme (1 μg) was then added into the MCM-MMCE coupled assay. However, no MMCE activity was detected. Neither was activity observed with its Ni^{2+}-incorporated enzyme, which had been prepared in a similar fashion. These observations suggested that there is no MMCE activity exhibited by Zn^{2+}-activated GlxI enzymes that have been reconstituted with Ni^{2+} and Zn^{2+}.

The Ni^{2+}/Co^{2+}-activated GlxI from *P. aeruginosa*, GloA2, was also investigated. Purified GloA2 (1 μg), which was overexpressed in the *E. coli* system in the presence of $NiCl_2$ (1 mM), was incubated with an additional 5 equivalents of $NiCl_2$ overnight at 4 °C. The coupled assay reaction, however, showed no detectable MMCE activity. Neither was activity observed with Co^{2+}-reconstituted GloA2 (0.048–1 μg). These results suggested that the non-Zn^{2+}-activated GlxI class of enzyme, which had been reconstituted with Ni^{2+} or Co^{2+} do not contain epimerase activity.

Since the previous results showed that small Glyoxalase I (homodimer of two-domain GlxI) do not exhibit MMCE activity, it was interesting to also investigate the large GlxI (monomeric four-domain GlxI) such as the GlxI from *Saccharomyces cerevisiae*. This enzyme is a Zn^{2+}-activated enzyme based on the Glx I assay [26]. Commercial yeast GlxI (2–9 ng) was added into the MCM-MMCE coupled assay and incubated at 37 °C for 5 min. However, no MMCE activity was detected, suggesting that both metal activation classes and sizes of GlxI do not possess epimerase activity. Even though both GlxI and MMCE share structural and functional similarity, the GlxI enzymes are very specific to their own substrate.

It has been reported that the activity of MCM is decreased in the presence of divalent metal ions (100 μM) [107]. More than 70% activity is maintained in the presence of Mg^{2+}, Co^{2+} and Fe^{2+}, while Ni^{2+} ion causes a decrease in MCM activity to a level of 48% [107]. Other metals such as Zn^{2+} and Mn^{2+} have been reported to cause a dramatic effect on the MCM assay such that the activity is decreased to only 3% [107]. Our experiments were performed with less than 0.3 μM of metals, so the effect of metal inhibition on MCM in the MCM-MMCE coupled assay was likely negligible.

5.4.3: Structural Comparison of Methylmalonyl-CoA Epimerase and Glyoxalase I

The structural and functional similarity of methylmalonyl-CoA epimerase and Glyoxalase I indicates that they are closely related and have likely evolved from a common ancestor gene. Like CLO GlxI (Chapter 2), homodimeric MMCE from *P. shermanii* packs back-to-back, forming an active site within one subunit (two active sites per dimer) (PDB: 1JC5) (Figure 5.10A). The interactions between a dimer are mainly hydrophobic and hydrogen bonding. One subunit of MMCE consists of two domains, the N-terminus (residues 7–72) and the C-terminus (residues 86–146). These two domains share low sequence similarity and identity (12–15%) but significant structural similarity (42 Cα matches with root mean square deviation (r.m.s.d.) of 1.52 Å). Domain I contains a β1α1β2β3β4 topology (residues 2–73) followed by a three-turn α-helix (residues 74–84) that connects to Domain II with a β5α2β6β7β8 arrangement (residues 85–146). A series of β-sheets curve up to create a deep cleft and provide a floor for the substrate, methylmalonyl-CoA. The catalytic pocket is relatively hydrophobic (Ala^{14}, Ala^{16}, Leu^{36}, Val^{46}, Met^{50}, Ala^{52}, Leu^{57}, Met^{67}, Val^{76}, Ala^{93}, Leu^{114}, Ile^{127}, Phe^{129} and Leu^{139}) and lies deep at the bottom of the cleft, protecting the substrate from the outside environment. There are four conserved metal binding residues (His^{12}, Gln^{65}, His^{91} and Glu^{141} in *P. shermanii* MMCE) located at the active site. The active site geometry with active metal (Co^{2+}) forms an octahedral coordination with four metal binding residues and two water molecules [105]. It is also found that the apo-enzyme and Co^{2+}-bound MMCE possess similar active site geometry [105]. The superimposed structures indicated that Co^{2+} ion in MMCE is only 0.2 Å from the position of the Zn^{2+} ion in human GlxI and 1.3 Å from that of the Co^{2+} ion in *E. coli* GlxI [45, 105].

Despite structural similarity between MMCE and GlxI, their enzymatic reactions are highly specific and cannot be replaced by one another. These suggest crucial structural differences that impact on their catalytic function. Homodimeric GlxI from *H. sapiens* with different 3D domain swapping forms an interaction between Domain I of one subunit and Domain II of another subunit, creating an active site at the dimeric interface. Domain I (residues 31–104) contains a β1α1β2β3α2β4 motif with a long N-terminal arm (residues 1–30) that extends into another subunit. The extra α-helix (α2) between strands 3 and 4 only exists in GlxI having a longer sequence or a Zn^{2+}-activated GlxI (Chapter 3). Domain II (residues 124–183) has slightly different topology of β1α1β2β3β4α2. A long flexible loop (residues

105–123) connects Domain I to Domain II. The structural similarity of Domain I and Domain II creates 42 out of 59 Cα matches with r.m.s.d. of 1.8 Å. The active site of the enzyme forms an octahedral coordination with a Zn^{2+} atom, four metal binding ligands (Gln^{33}, Glu^{99}, His^{126} and Glu^{172}) and two water molecules (PDB: 1QIP) [46]. The catalytic pocket is relatively hydrophobic, which is composed of Cys^{60}, Phe^{62}, Phe^{67}, Leu^{69}, Phe^{71}, Ile^{88}, Leu^{92} from one subunit and Met^{157}, Leu^{160}, Phe^{162}, Leu^{174} from another subunit.

The superimposed structures of MMCE and human GlxI indicate the opposite location of the three-turn α-helix of MMCE (residues 74–84) and a long helix B of GlxI (residues 81–95) (Figure 5.10B). This helix B locates almost in the middle of the active site, creating a wall that completely seals the active site from outside solvent. Thus, it is highly possible that the substrate would enter from another wider side (Figure 5.11). The structure of human GlxI with a bound inhibitor (*S*-(*N*-hydroxy-*N*-*p*-iodophenylcarbamoyl)glutathione) supports this observation, where the side chain carboxylate groups of γ-glutamate and the C-terminal carboxylate group of glycine in GSH moiety are pointing outward and interacting with enzyme residues at the entrance of the active site, while the iodophenyl group prefers a hydrophobic environment within the active pocket (PDB: 1QIN) (Figure 5.10C). As well, several charged residues that are located at the entrance of the active site are conserved for all GlxI and are believed to be important in stabilizing the bound hemithioacetal (except trypanothione-utilizing enzymes such as *T. cruzi* and *L. donovani* GlxI) (Figure 5.1). For the human GlxI, the guanidino group of Arg^{37} forms a salt link with the carboxylate group of the γ-glutamate of GSH moiety, which can also interact with Arg^{122} from the connecting loop of Domain I and Domain II of another subunit that lies within H-bonding distance (3.1 Å). As well, an amide group from the side chain of Asn^{103} can form a hydrogen bond with the carboxylate and the amino groups of this γ-glutamate. The amide group of the Cα from Lys^{156} may also interact with the glycyl residue of GSH moiety. These residues create a relatively hydrophilic path to the active site.

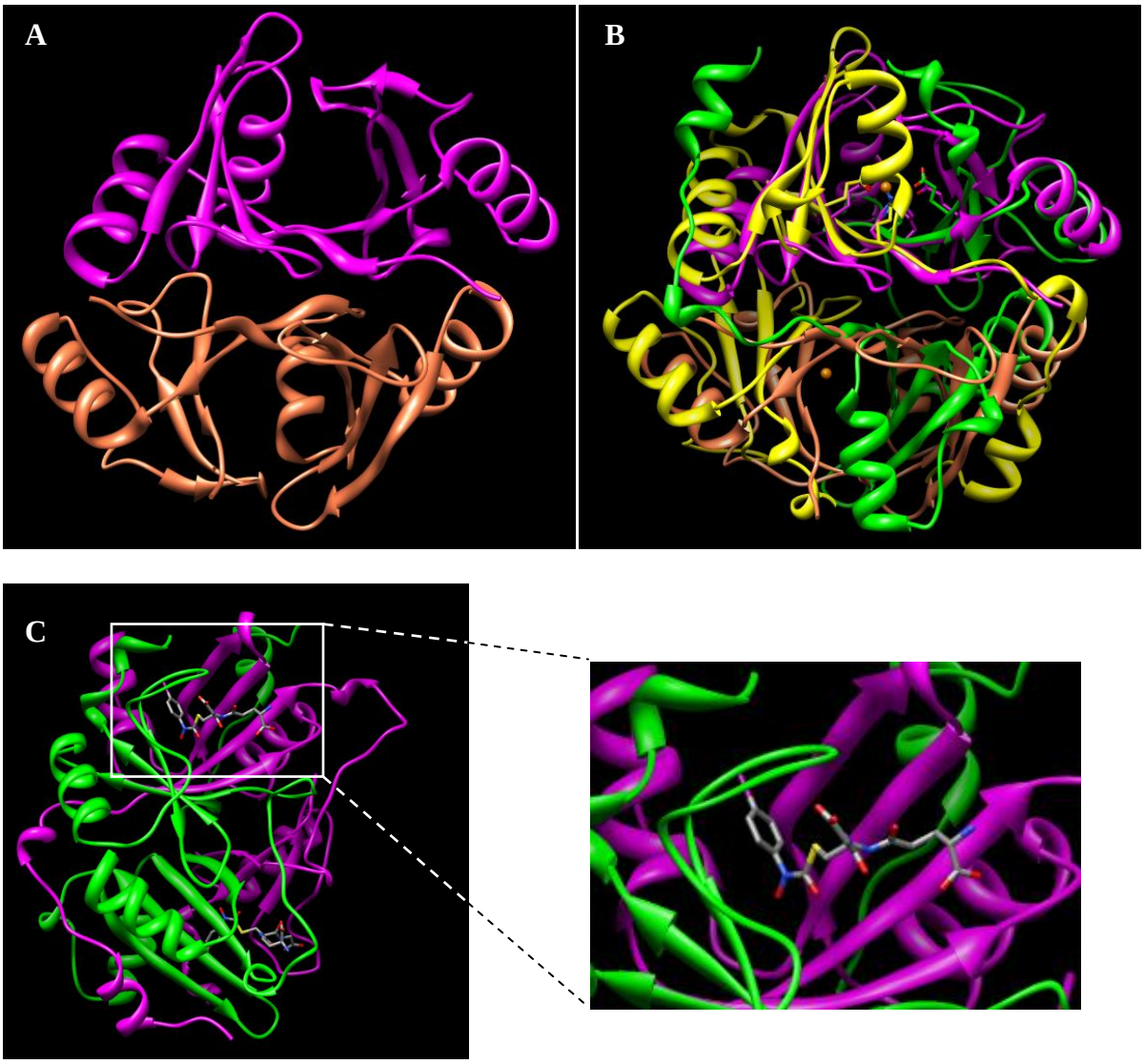


Figure 5.10: (A) The ribbon structure of *P. shermanii* MMCE (magenta and coral, PDB: 1JC5) shows the enzyme dimerization, forming by the back-to-back arrangement of two monomers. **(B)** The superimposed structures of *P. shermanii* MMCE (magenta and coral, PDB: 1JC5) with *H. sapiens* GlxI (green and yellow, PDB: 1QIN) show the metal binding residues (CPK stick) and Zn^{2+} atoms (orange spheres) that are located at the active sites of the GlxI enzyme. **(C)** The ribbon structure of *H. sapiens* GlxI (magenta and green, PDB: 1QIN) shows the binding of the transition state analogue of *S*-(*N*-hydroxy-*N*-*p*-iodophenylcarbamoyl)glutathione in the active site. The 3D structures were generated by using the UCSF Chimera program (University of California, San Francisco) [1].

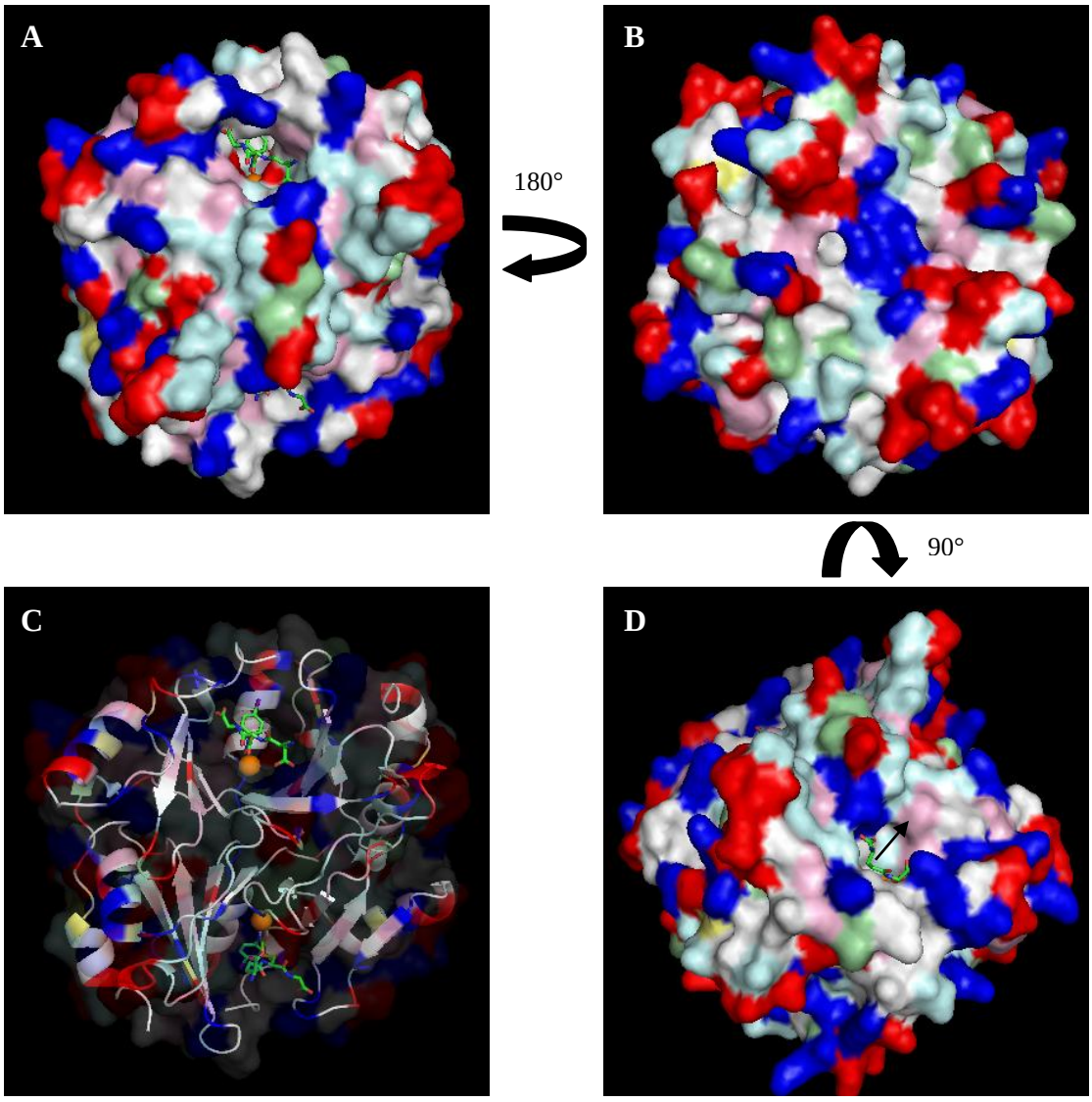


Figure 5.11: The surface illustration of *H. sapiens* GlxI (PDB: 1QIN) showing **(A)** front view with the active site, **(B)** back view, **(C)** ribbon structure with inhibitor *S*-benzyl-glutathione in stick and Zn^{2+} atom in orange (80% surface transparency) and **(D)** top view with the arrow indicating a possible entrance to the active site. Hydrophobic residues (AGILMPV) = white, polar residues (CHNQST) = cyan, positive residues (KR) = blue, negative residues (DE) = red and aromatic residues (FWY) = magenta. The 3D structures were generated by using the PyMOL Molecular Graphics program (http://www.pymol.org/).

Methylmalonyl-CoA epimerase, on the other hand, contains a three-turn α-helix that partially blocks the active site, creating a more narrowed entrance for the substrate (Figure 5.13). Thus, it is possible that the substrate would enter the active site of MMCE from another entrance. If methylmalonyl-CoA enters the active site from the wider entrance (without partial blocking by the three-turn α-helix), several aromatic residues (Trp^{33}, His^{34}, His^{37}, His^{60} and Tyr^{145}) are possibly important for interacting with the adenosine base of the coenzyme A. The

positively charged residue Arg^{95} may interact with the α-hydroxy ketone (–CO-COH–), while Lys^{56} may interact with the phosphate groups of the substrate. However, analysis of the multiple sequence alignment of MMCE from different organisms showed that all these residues are not conserved (Figure 5.1). On the other hand, if the substrate enters MMCE from the more narrowed entrance (with partial blocking by the three-turn α-helix), several aromatic and charged residues (Thr^{75}, Trp^{79}, Lys^{82}, His^{83}, His^{90}, Tyr^{115}, Phe^{129} and His^{131}) are possibly crucial for interacting with the adenosine base and the phosphate groups of methylmalonyl-CoA. As well, charged residues (Asn^{41}, Glu^{43}, Gln^{44}, Lys^{119} and Thr^{122}) might be interacting with the α-hydroxy ketone of the substrate. The multiple sequence alignment analysis indicates that most of these residues are conserved (Figure 5.1). The charged residues are located at a distance from the entrance; hence, they are likely to make the base of the entrance relatively hydrophobic (Figure 5.13). This hypothesis seemed to be supported by the molecular docking computational experiments which calculated likely docked poses of (*2S*)-methylmalonyl-CoA in the X-ray crystallographic structure of *P. shermanii* MMCE, where the coenzyme A portion of the substrate was found to frequently lie along the path created by the three-turn α-helix (Figure 5.12).

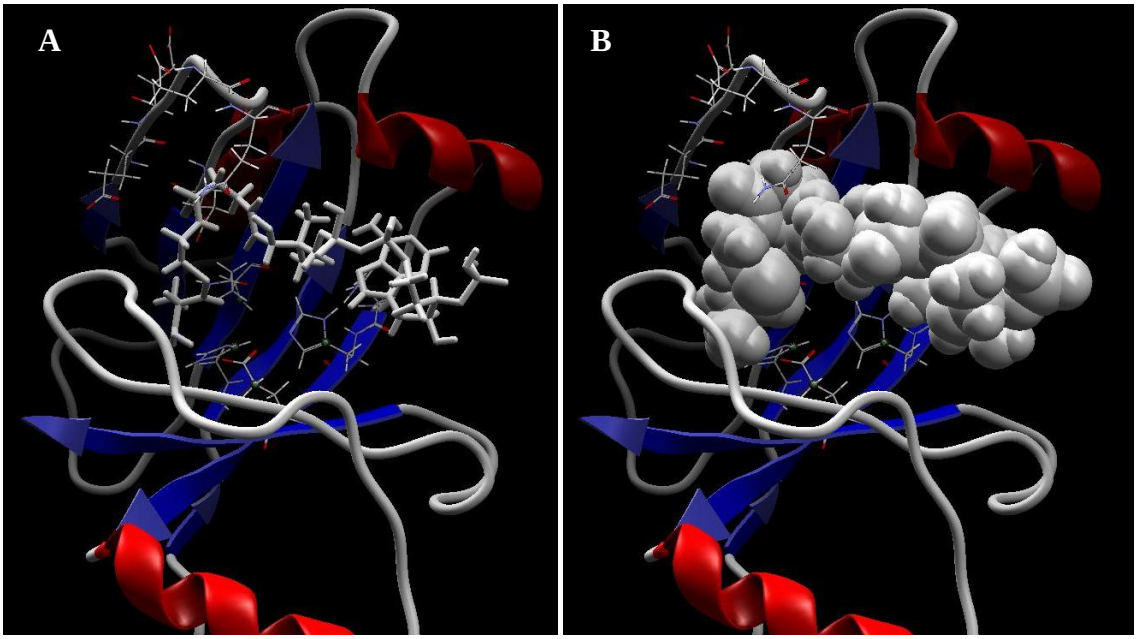


Figure 5.12: The molecular docking of (*2S*)-methylmalonyl-CoA (**(A)** stick and **(B)** spacefilled molecule) into the active site of *P. shermanii* methylmalonyl-CoA epimerase (PDB: 1JC5) using Molegro Virtual Docker program [220].

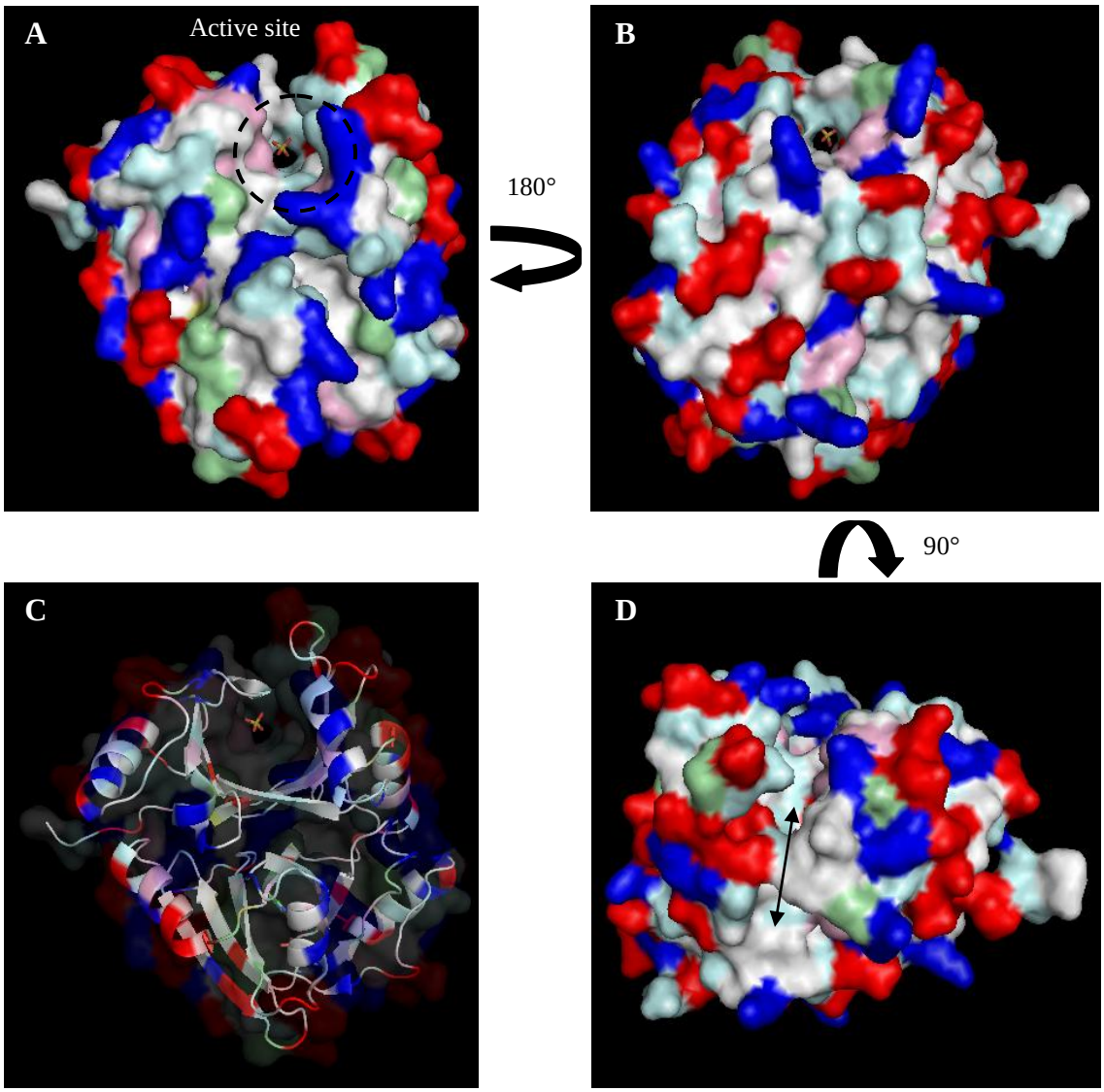


Figure 5.13: The surface illustration of *P. shermanii* MMCE (PDB: 1JC5) showing **(A)** front view with the active site, **(B)** back view, **(C)** ribbon structure with SO_4 in CPK stick (80% surface transparency) and **(D)** top view with the arrow indicating two possible entrances to the active site. Hydrophobic residues (AGILMPV) = white, polar residues (CHNQST) = cyan, positively charged residues (KR) = blue, negatively charged residues (DE) = red and aromatic residues (FWY) = magenta. The 3D structures were generated by using the PyMOL Molecular Graphics program (http://www.pymol.org/).

The residues that might interact with the substrate are necessary for GlxI function and may limit and distinguish GlxI from MMCE. Without reported evidence, MMCE might not be expected to have enzymatic activity with the hemithioacetal used as a substrate for GlxI enzymes. Based on the previous observations discussing possible key residues that interact with substrate, if the hemithioacetal is considered to bind in the active site in a "Y"

conformation, where the thiol moiety of the cysteine residue is the base and the glutamate and glycine residues are the forks, the entrance of the GlxI requires a large space in order to support the size of its substrate (the distance from δ-carbon of glutamate to carboxyl-carbon of glycyl residue of GSH is 8.93 Å). This entrance space is considered to be wider than the one in MMCE. In addition, the overlapped structures of human GlxI with *S*-(*N*-hydroxy-*N*-*p*-iodophenylcarbamoyl)glutathione (PDB: 1QIN) and *P. shermanii* MMCE (PDB: 1JC5) allow the clear picture of collision of the inhibitor inside the MMCE cavity (Figure 5.14). The side chain of Val^{46} projects into the γ-glutamate of GSH and the side chains of Val^{76} and Met^{67} locate too close to the carboxylate group of this γ-glutamate. The benzyl group of the inhibitor collides with the loop that covers the active site of MMCE (residues 119–125). This prediction cannot rule out the possibility that hemithioacetal may enter from another section of MMCE (without blocking from the three-turn α-helix). However, this side is also narrowed and hemithioacetal binding residues are not observed in this entrance, making it highly possible that the enzyme may have no activity with this substrate.

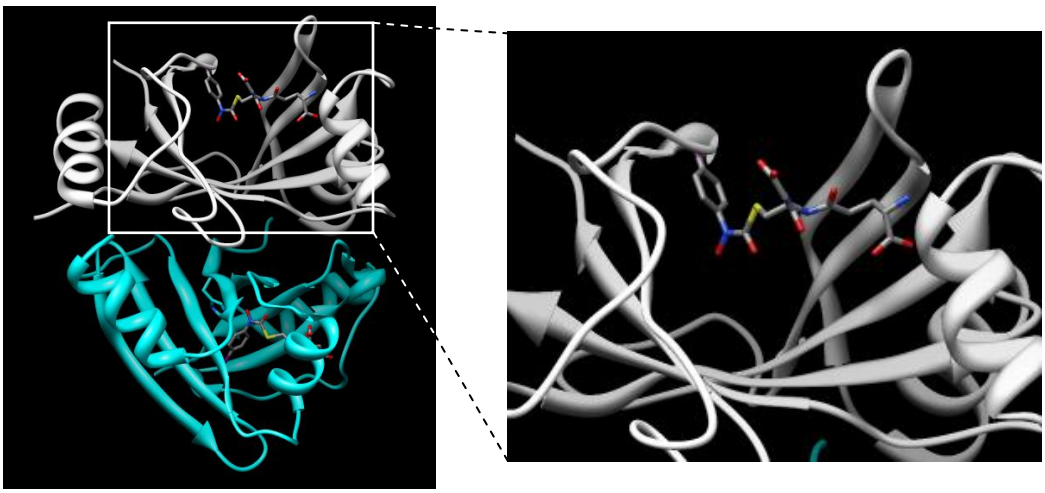


Figure 5.14: The ribbon structure of *P. shermanii* MMCE (PDB: 1JC5) shows a predicted inhibitor of *S*-(*N*-hydroxy-*N*-*p*-iodophenylcarbamoyl)glutathione in the active site. The inhibitor is actually found in a structure deposited for an inhibited *H. sapiens* GlxI (PDB: 1QIN), but by performing structural superimposition with MMCE, the substrate binding sites were predicted. The 3D structures were generated by using the UCSF Chimera program (University of California, San Francisco) [1].

The lack of MMCE activity with Zn^{2+}-activated GlxI, as discussed previously, is possibly due to the absence of methylmalonyl-CoA binding residues. As well, these charged residues are located at the entrance, creating an inappropriate environment for methylmalonyl-

CoA to bind. Another possibility is probably due to the inappropriate interaction of the substrate and the enzyme in the catalytic pocket. As stated previously, the active site of human GlxI forms an octahedral geometry around the active metal (Zn^{2+}-atom); four metal binding residues (Gln^{33}, Glu^{99}, His^{126} and Glu^{172}) and two water molecules surround the metal ion. With the Zn^{2+} atom being a central metal, His^{126} and one water molecule (H_2O_{ax}) locate in the axial positions, while Gln^{33}, Glu^{99}, Glu^{172} and another water molecule (H_2O_{eq}) lie in the equatorial positions. Once the substrate is introduced, two oxygen atoms of the α-hydroxy ketone of the hemithioacetal substrate have been suggested to replace two water molecules, thus maintaining the metal coordination. The distance between the two oxygens in the α-hydroxyl ketone of transition state analogue (*S*-(*N*-hydroxy-*N*-*p*-iodophenylcarbamoyl)glutathione) is 2.69 Å (PDB:1QIN). This distance is close to the distance between H_2O_{ax} and H_2O_{eq} (2.91 Å), which would seem to support the replacement of the two water molecules around the metal ion with the substrate oxygen atoms (PDB: 1QIP). The distance between the two oxygen atoms in the methylmalonyl group is 3.83 Å (PDB: 4REQ). This distance may be too large to fit properly into the active site metal geometry of GlxI and hence it might be less likely to appropriately replace water molecules around the metal ion.

Likewise, the catalytic site of *E. coli* GlxI (non Zn^{2+}-activated enzyme), with a distance between H_2O_{ax} and H_2O_{eq} of 3.06 Å (PDB: 1FA8), might be too small to accept methylmalonyl-CoA. As well, *E. coli* GlxI contains four possible hemithioacetal binding residues including Arg^{9}, Asn^{61}, Arg^{98} and Lys^{104} that might not interact optimally with the methymalonyl-CoA molecule (Figure 5.15). Arg^{9} and Asn^{60} from one subunit stabilize the carboxylate and amino groups of the γ-glutamate of GSH moiety, while Arg^{98} and Lys^{104} stabilize the carboxyl group of its glycyl residue. Even though there are two possible entrances in *E. coli* GlxI due to the absence of helix B, both entering sides are covered with charged residues. Interestingly, only one side of *E. coli* GlxI contains conserved residues for hemithioacetal stabilization, which is in the same entering side as it is for human GlxI (the side that has no hindrance by helix B) and is possibly the preferred side for the substrate to bind (Figure 5.16A). The other side is more narrow due to the presence of a loop (residues 46–54), which connects β3 and β4 and that bends up to partially cover the active site. In fact, our unpublished *E. coli* GlxI X-ray structure with a bound hydroxamate analogue shows that the

last nine amino acids in the C-terminus (residues 127–135) block this side of the active cleft upon binding of the inhibitor (Figure 5.16B). Even though these amino acids are not detected in the enzyme without bound inhibitor, the last two amino acids (Gly^{134} and Gln^{135}) are found to locate distantly from the active site (PDB: 1F9Z). Therefore, it is possible that this loop flips back to cover one side of the active site once the substrate enters to protect it from the outside environment (more detailed discussions in Chapter 6).

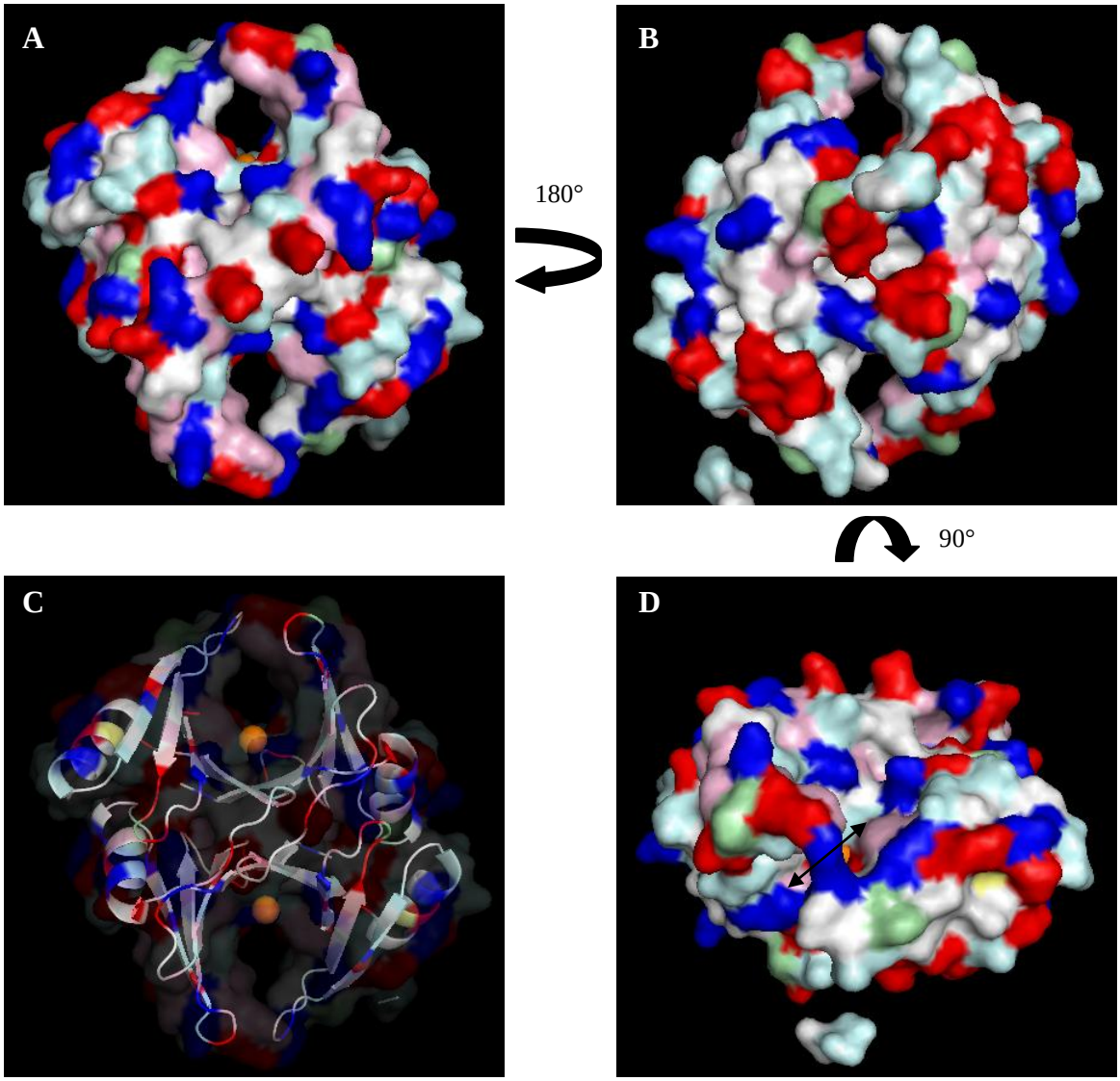


Figure 5.15: The surface illustration of *E. coli* GlxI (PDB: 1F9Z) showing **(A)** front view with the active site, **(B)** back view, **(C)** ribbon structure with Ni^{2+} atom in orange (80% surface transparency) and **(D)** top view with the arrow indicating two possible entrances to the active site. Hydrophobic residues (AGILMPV) = white, polar residues (CHNQST) = cyan, positive residues (KR) = blue, negative residues (DE) = red and aromatic residues

(FWY) = magenta. The 3D structures were generated by using the PyMOL Molecular Graphics program (http://www.pymol.org/).

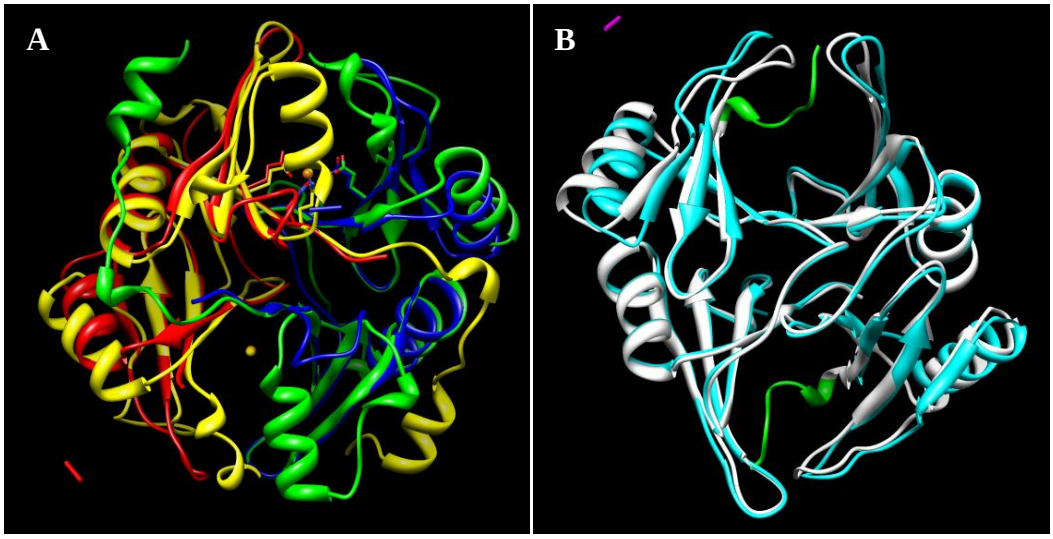


Figure 5.16: (A) The superimposed structures of *E. coli* GlxI (blue and red, PDB: 1F9Z) with *H. sapiens* GlxI (green and yellow, PDB: 1QIN) shows the metal binding residues (CPK stick) and bound metals (orange sphere for Zn^{2+} and yellow sphere for Ni^{2+}) that are located at the active site of the enzymes. **(B)** The superimposed structures of *E. coli* GlxI with a bound hydroxamate (TSI) inhibitor (white, unpublished structure) and *E. coli* GlxI without bound inhibitor (cyan, PDB: 1F9Z) shows the last nine amino acids at the C-terminus of the enzyme with bound inhibitor, which are indicated in green. The last two amino acids at the C-terminus of *E. coli* GlxI without bound inhibitor are indicated in magenta. The 3D structures were generated by using the UCSF Chimera program (University of California, San Francisco) [1] .

CLO GlxI, the first enzyme in the Glyoxalase family that has close structural similarity to MMCE but has GlxI activity, possesses no activity with methylmalonyl-CoA. CLO GlxI contains a loop (residues 107–110) that acts as a lid, covering the active site from above. This loop projects down to one side of the active site, creating a wall as well as forcing the substrate to enter from another available entrance (Figure 5.17). This entrance contains a short two-turn α-helix (residues 61–71) that resembles the three-turn α-helix in MMCE (Figure 5.18A). This entrance is in the same substrate entering side as those of human GlxI and *E. coli* GlxI (Figure 5.18B and 5.18C, respectively). Compared to the three-turn α-helix found in MMCE, the two-turn α-helix present in the structure of CLO GlxI partially blocks the active site and creates a more narrowed path for the substrate to bind. Despite their structural similarities, CLO GlxI contains only two conserved residues for stabilizing methylmaloyl-CoA (Lys^{69} and Phe^{114}). As well, the catalytic space that is closed up by two loops (residues 37–39 and residues 107–110) is possibly another crucial factor for the absence of MMCE activity. These loops create a

smaller catalytic space than the one in MMCE, which is created by loop residues 42–45 and residues 119–125. Besides, CLO GlxI possesses a distance between H_2O_{ax} and H_2O_{eq} of 2.84 Å (PDB: 3HDP) in the metal-enzyme complex, which may be too small for methylmalonyl-CoA to effectively duplicate when binding. Hence, due to lack of residues that actually bind the substrate and of sufficient space in the active sites, CLO GlxI does not exhibit MMCE activity.

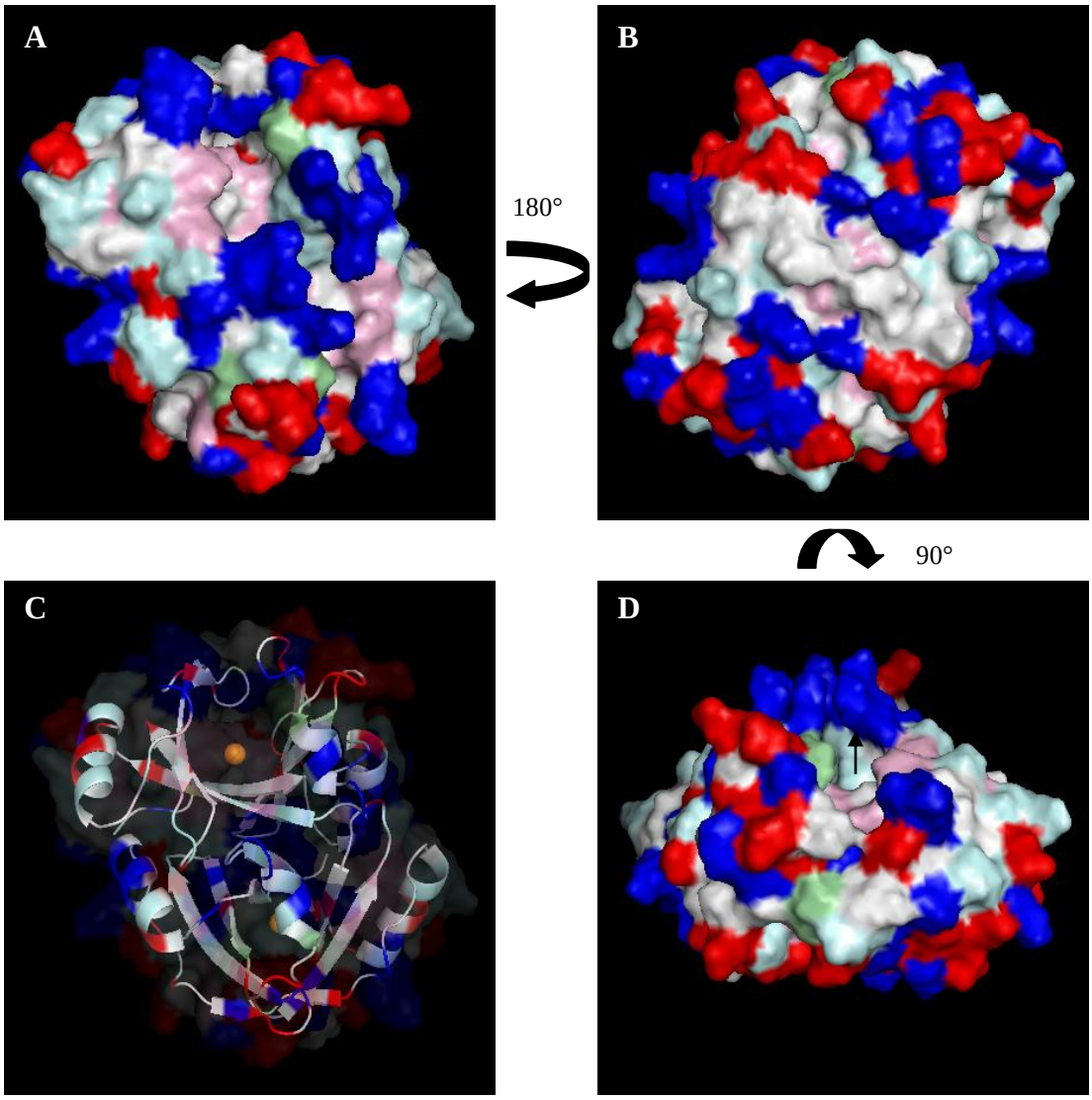


Figure 5.17: The surface illustration of *C. acetobutylicum* GlxI (PDB: 2QH0) showing **(A)** front view with the active site, **(B)** back view, **(C)** ribbon structure with Zn^{2+} atom in orange (80% surface transparency) and **(D)** top view with the arrow indicating a possible entrance to the active site. Hydrophobic residues (AGILMPV) = white, polar residues (CHNQST) = cyan, positive residues (KR) = blue, negative residues (DE) = red and aromatic residues (FWY) = magenta. The 3D structures were generated by using the PyMOL Molecular Graphics program (http://www.pymol.org/).

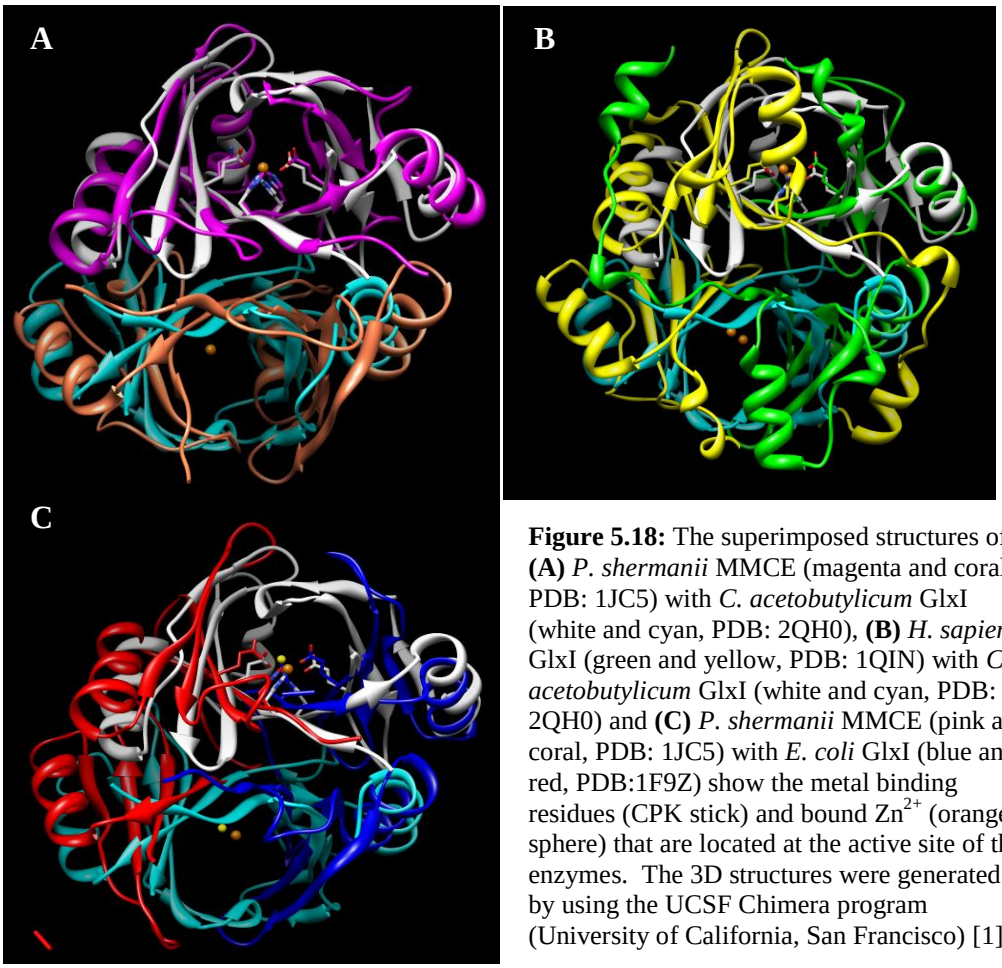


Figure 5.18: The superimposed structures of **(A)** *P. shermanii* MMCE (magenta and coral, PDB: 1JC5) with *C. acetobutylicum* GlxI (white and cyan, PDB: 2QH0), **(B)** *H. sapiens* GlxI (green and yellow, PDB: 1QIN) with *C. acetobutylicum* GlxI (white and cyan, PDB: 2QH0) and **(C)** *P. shermanii* MMCE (pink and coral, PDB: 1JC5) with *E. coli* GlxI (blue and red, PDB:1F9Z) show the metal binding residues (CPK stick) and bound Zn^{2+} (orange sphere) that are located at the active site of the enzymes. The 3D structures were generated by using the UCSF Chimera program (University of California, San Francisco) [1].

CLO GlxI with a lower GlxI activity only contains one conserved hemithioacetal binding residue Lys^{101} and a non-conserved Arg^{39} that is located at the entrance of the active site (Figure 5.1). Structural comparison with human GlxI indicates that CLO GlxI can barely accept the inhibitor, *S*-(*N*-hydroxy-*N*-*p*-iodophenylcarbamoyl)glutathione, in its active site (Figure 5.19). The loop (residues 37–39) that acts as a lid bends down to cover the catalytic site of CLO GlxI. This loop contains Arg^{39} that may interact with the carboxylate group of γ-glutamate of the GSH. As well, Lys101 is within H-bonding distance to the carboxyl group of glycine residue. The benzyl group of the inhibitor, however, would appear to collide with another loop (residue 107–110). The catalytic space within the enzyme is limited and more closed than that of other GlxI. The entrance for the substrate is narrowed by the two-turn α-helix that partially blocks the active site. Due to these factors, CLO GlxI might be expected to have a lower enzyme activity compared to other GlxI (Chapter 2).

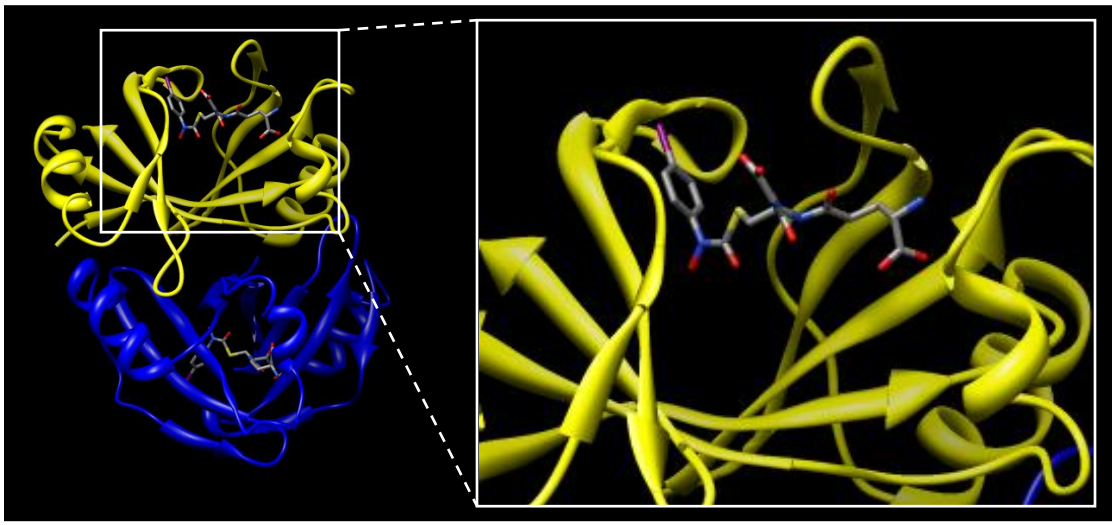


Figure 5.19: The ribbon structure of *C. acetobutylicum* GlxI (PDB: 2QH0) with the predicted inhibitor *S*-(*N*-hydroxy-*N*-*p*-iodophenylcarbamoyl)glutathione in the active site. The inhibitor is actually found in a structure deposited for an inhibited *H. sapiens* GlxI (PDB: 1QIN), but by performing structural superimposition with the *C. acetobutylicum* enzyme, the substrate binding sites were predicted. The 3D structures were generated by using the UCSF Chimera program (University of California, San Francisco) [1].

So far, there are several features that can be used to predict and distinguish GlxI and MMCE using their amino acid sequences and structures. This has not really been proposed previously in the literature to our knowledge. One is to consider conserved substrate stabilizing residues by performing multiple sequence alignment analysis. GlxI has positively charged residues that bind the hemithioacetal substrate and MMCE has several key aromatic and charged residues that likely bind the methylmalonyl-CoA substrate. These residues appear to be conserved in all organisms in which these enzyme activities have been identified. Along with these binding residues, the entrance of the active site requires some consideration. If the entrance to the active site is relatively hydrophobic and narrower, the enzyme is more suitable to accept methylmalonyl-CoA. GlxI employs a wider entrance to support the structure of its substrate. As well, the orientation of the metal-enzyme complex in the active site might also be useful in predicting enzyme function. Despite these features the experimental evidence of enzymatic activity is required to identify the actual function. It would be interesting, for example, to determine if methylmalonyl-CoA can bind to GlxI, even though it is not turned over by the enzyme, and also if the hemithioacetal (or an analogue such as an inhibitor of GlxI) can bind to MMCE. These experiments could possibly be done by isothermal titration calorimetry experiments.

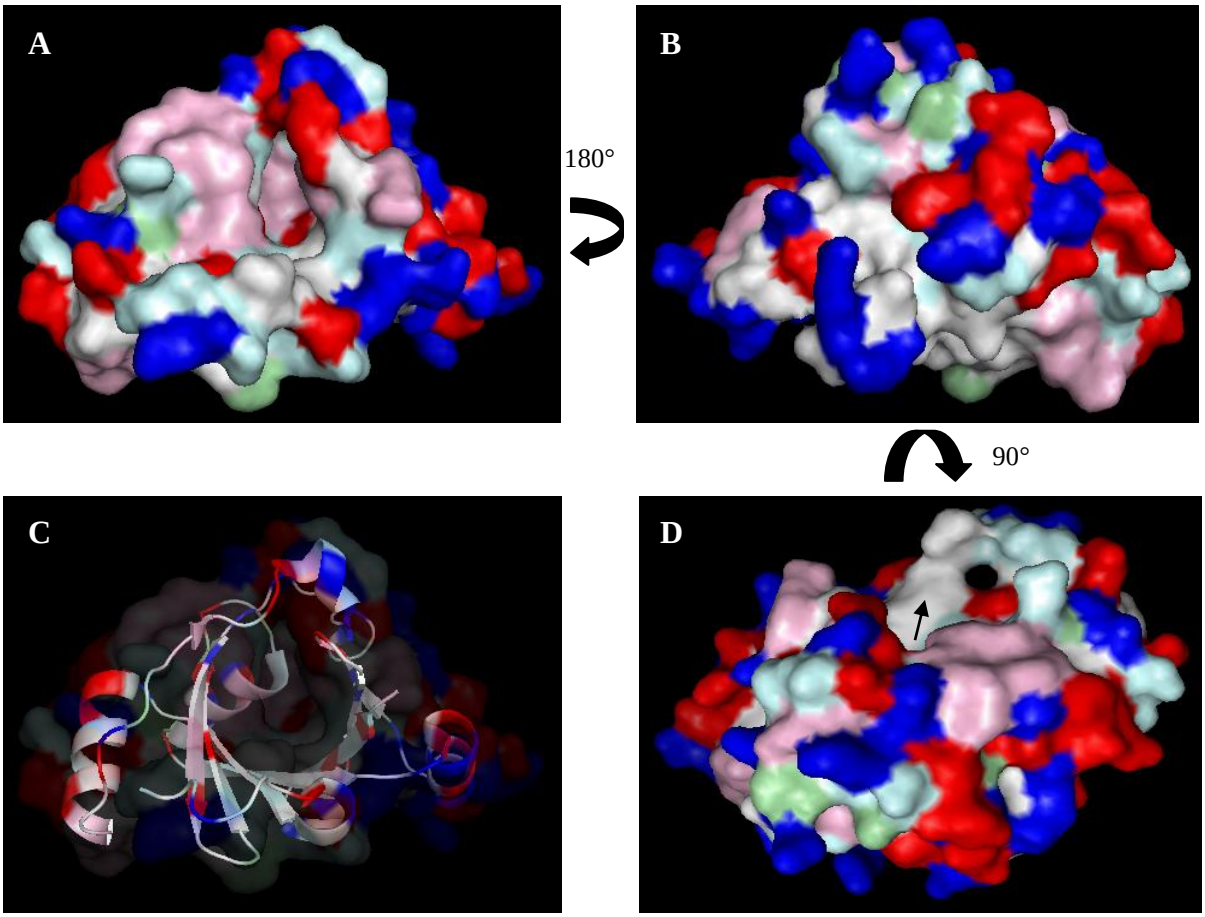


Figure 5.20: The surface illustration of putative MMCE from *T. tengcongensis* (PDB: 3GM5) showing **(A)** front view with the active site, **(B)** back view, **(C)** ribbon structure (80% surface transparency) and **(D)** top view with the arrow indicating a possible entrance to the active site. Hydrophobic residues (AGILMPV) = white, polar residues (CHNQST) = cyan, positive residues (KR) = blue, negative residues (DE) = red and aromatic residues (FWY) = magenta. The 3D structures were generated by using the PyMOL Molecular Graphics program (http://www.pymol.org/).

In addition, the X-ray crystallographic structure of a putative MMCE from *Thermoanaerobacter tengcongensis* (PDB: 3GM5) has been investigated (Figure 5.20) [229]. Based on its structure, this protein was hypothesized to be a MMCE; however, the researchers did not examine the protein for enzymatic activity [229]. The catalytic site of the protein in this Gram-positive bacterium is more packed than the ones in *P. shermanii* MMCE due to the presence of two longer loops (residues 42–60 and residues 121–126) that act as lids to the active site. As well, a three-turn α-helix that locates almost in the middle of the active site completely seals the active site from the outside environment. This wall might force the substrate to enter from another available entrance, which is relatively hydrophobic.

Interestingly, this entrance is opposite from that of *P. shermanii* MMCE, where the substrate is hypothesized to bind at the side containing the three-turn α-helix. The putative MMCE enzyme from *T. tengcongensis* contains five substrate binding residues that are conserved in all studied MMCE enzymes and none of the residues that are conserved in GlxI enzymes (Figure 5.1). Therefore, it is highly possible that this enzyme may have MMCE activity. However, its activity may be low due to restriction of catalytic and substrate binding spaces. These predictions will need to be experimentally evaluated in the future.

5.4.4: Suggested Evolution of the βαββββ Superfamily

Several proteins such as Glyoxalase I, methylmalonyl-CoA epimerase, fosfomycin resistance protein, extradiol dioxygenase, bleomycin resistance protein and mitomycin C resistance protein share a βαββββ motif, suggesting divergent evolution within this superfamily [40, 41]. It was hypothesized that the dimeric members of the superfamily share a common one-module ancestor, in which upon gene duplication and fusion as well as mutagenesis, these proteins possess low sequence homology but structural similarity (Figure 5.21) [40, 41]. These enzymes may use the evolutionarily conserved common fold as a platform but bind a specific metal ion, vary the ligands and change the active site geometry, thus achieving diverse function (Table 5.3). A proposed early event is the gene modification of a single motif ancestor to give a C2-symmetric metallo-dimer that links $(\beta\alpha\beta\beta\beta)_1$ and $(\beta\alpha\beta\beta\beta)_2$ motifs through metal binding ligands. These metal-bound motifs then form a more stable two-domain pseudosymmetric metallo-monomer as a result of gene duplication and fusion such as may be found in the Zn monomeric *P. putida* GlxI, which was reported to be somewhat catalytically active [30]. This active monomer is considered a precursor of a back-to-back dimeric formation in MMCE, which suggests that this protein may have evolved earlier than other homodimeric proteins in this family. Other dimeric formations may have formed at the subunit interface to give an edge-to-edge-fusion of $(\beta\alpha\beta\beta\beta)_1$ of one subunit with $(\beta\alpha\beta\beta\beta)_2$ of another subunit and may be the result of gene modification and arrangement termed 3D-domain swapping [40, 41]. These small dimeric two-subunit modules are believed to be progenitors of the larger monomeric four-domain proteins such as the large GlxI from yeast and *Plasmodium falciparum* as well as several extradiol dioxygenases [41].

Among these proteins, MMCE and GlxI show a very close relationship. Both are metalloenzymes with four metal binding residues, forming an octahedral geometry when bound to their activating metal ions [44-46, 105]. The symmetry of the metal binding ligands is inherent from the symmetry in the paired βαββββ motifs. These enzymes may have preceded the tridendate enzymes (three metal binding ligands) such as extradiol dioxygenase and FosA that catalyze more complex bisubstrate reactions. Loss of one of the four metal binding ligands provides further functional diversification by opening an additional coordination site to support the particular reaction. Other proteins in this structural class such as the bleomycin resistance proteins and the mitomycin resistance proteins are neither enzymes nor metalloproteins. The metal binding ability of the progenitor of the bleomycin resistance proteins may have been omitted in favor of a more hydrophobic cavity to accommodate the antibiotic binding.

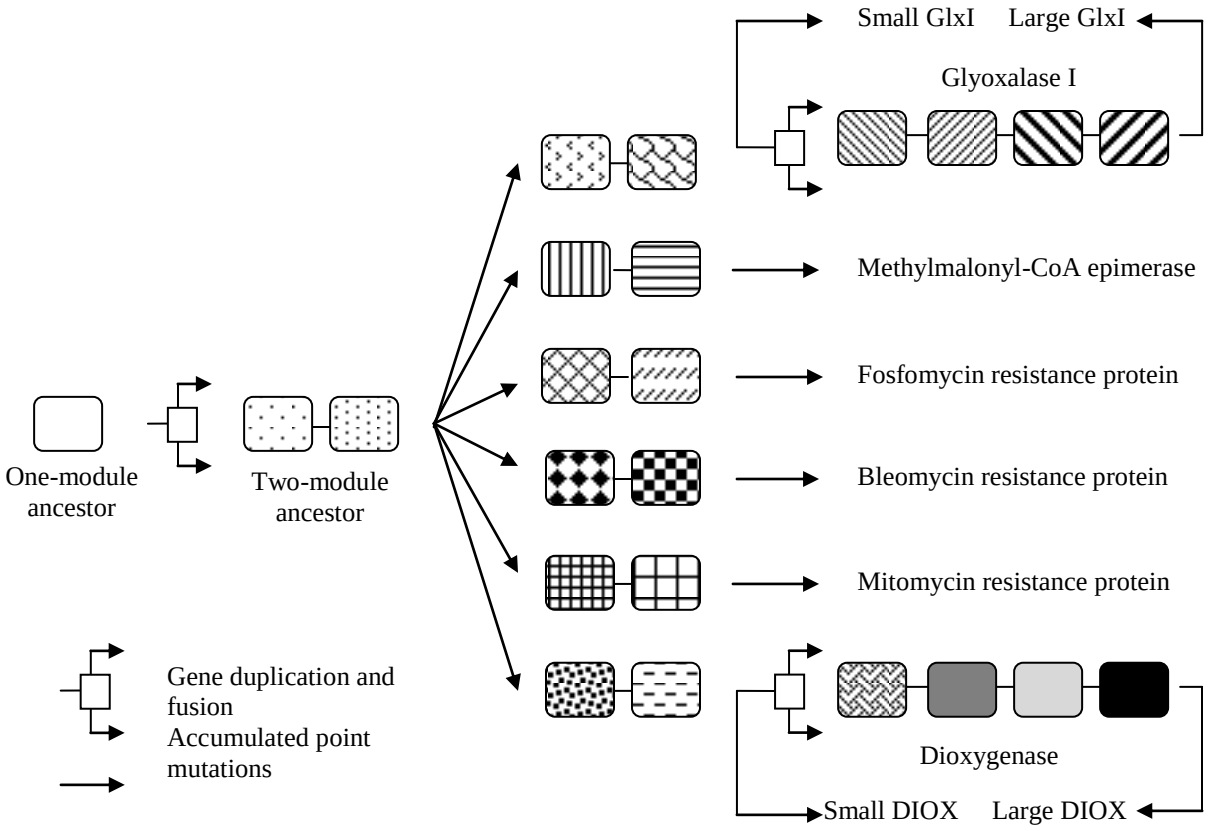


Figure 5.21: The hypothesis of evolution within the βαββββ superfamily based on the report of Bergdoll *et al*., 1998 [41].

Table 5.3: The summary of metal activations, reaction types and structures of proteins in the βαβββ superfamily

Protein	***Active metal***	***Reaction type***	***Structure***	***Ref.***
Glyoxalase I	Zn^{2+}, Ni^{2+}/Co^{2+}	Isomerization	1F9Z, 1QIN	[45, 46]
Methylmalonyl-CoA epimerase	Co^{2+}	Epimerization	1JC4, 1JC5	[105]
Extradiol dioxygenase	Fe^{2+}, Mn^{2+}	C-C bonding cleavage	1HAN	[145]
Fosfomycin resistance protein	Mn^{2+}, Mg^{2+}	Epoxide opening	1LQO	[169]
Bleomycin resistance protein	None	Drug binding	1JIF	[187]
Mitomycin resistance protein	None	Drug binding	2A4X	[195]

5.5: Conclusions

Even though methylmalonyl-CoA epimerase and Glyoxalase I possess structural similarity, their catalytic functions are specific and cannot be replaced by one another. No epimerase activity was detected with either small (dimeric four-domain) GlxI for both the Zn^{2+}-activated (i.e. *P. aeruginosa* GloA3) and Ni^{2+}/Co^{2+}-activated (non Zn^{2+}-activated GlxI i.e. *P.aeruginosa* GloA2 and CLO GlxI) enzymes. Neither was MMCE activity found in the large (monomeric four-domain) GlxI (i.e. yeast GlxI). These results suggest that an alternate MMCE activity does not occur with any metal classes or sizes of GlxI.

The structural investigation suggests several main factors that might affect specificity of the substrate including the entrance to the active site, the space within the metal binding cleft and the interactions of the enzyme with bound substrate. Three GlxI from humans, *E. coli* and *C. acetobutylicum* were considered to explain the impossibility of cross activity with MMCE. The superimposed structures of these GlxI with MMCE indicate a potential entrance of the substrate to the catalytic pocket. Interestingly, methylmalonyl-CoA, the substrate of MMCE, is likely to enter the MMCE enzyme at a more narrowed entrance, created by partially blocking of three-turn α-helix, instead of by the alternate wider side. This helix contains several residues that are important for substrate stabilization. It was found that all conserved substrate binding residues only exist on this side of the enzyme. The entrance to the catalytic pocket is relatively hydrophobic. On the other hand, GlxI requires charged residues that can stabilize its substrate, the hemithioacetal formed between methylglyoxal and the tripeptide glutathione. These residues are mostly positively charged and located at the entrance to the active site. The lack of MMCE activity in GlxI is possibly due to the absence of binding residues for methylmalonyl-CoA. Besides, the orientation of the metal-enzyme complex in

GlxI likely only supports the replacement of two water molecules with the hemithioacetal (and not methylmalonyl-CoA). The structural investigation also predicts that MMCE may not exhibit GlxI activity. The superimposed structures of human GlxI with bound inhibitor, *S*-benzyl glutathione, and MMCE indicate that the hemithioacetal requires a wider binding site, while MMCE with a narrower path is likely not able to bind this substrate. Besides, MMCE lacks binding residues for the hemithioacetal. However, in order to confirm these observations, an investigation of possible GlxI activity by MMCE would be required.

Evolution within the βαβββ superfamily has been hypothesized to have involved gene duplication and fusion events as well as domain swapping events, resulting in the different functional proteins derived from a common precursor. The formation of a symmetrical protein with four metal binding ligands is a crucial step in the evolution process for this structural class. It has been proposed that these metalloenzymes such as GlxI and MMCE evolved earlier from the two-module ancestor as a result of symmetric dimeric interaction of the two βαβββ motifs [41]. The enzymes having only three metal binding protein ligands, such as extradiol dioxygenase and FosA, may have evolved from the tetradentated progenitor to accomplish more complicated reactions and therefore providing biological functional variety. Drug resistance proteins such as bleomycin resistance proteins and mitomycin resistance proteins that contain no metal binding ligands are believed to have evolved by modification of the precursor gene along with various mutations to create a hydrophobic cavity, a suitable environment for sequestering the somewhat hydrophobic antibiotics.

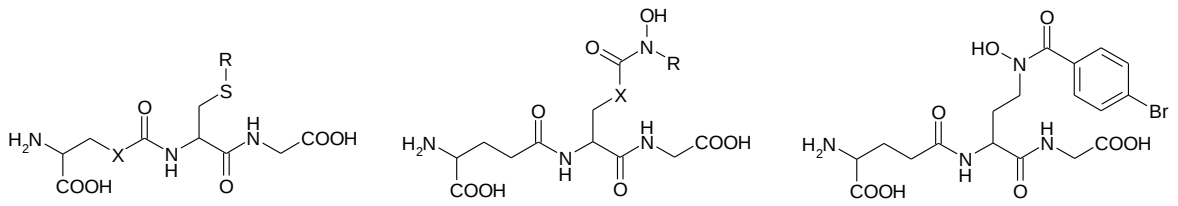
CHAPTER 6: INHIBITION STUDIES AND PROPOSED MECHANISM OF GLYOXALASE I

6.1: Introduction

Methylglyoxal (MG) is an intracellular cytotoxic compound that can initiate cell damage at the level of RNA, DNA and protein, which eventually is believed to lead to cell apoptosis [7-9]. Many organisms thus provide several defense mechanisms to survive under such stressful conditions [13, 14]. The Glyoxalase series of reactions is an area of current interest in our laboratory. These reactions appear to be involved in the detoxification of MG by converting it to the non-toxic molecule D-lactate by action of the enzymes Glyoxalase I (GlxI) and Glyoxalase II (GlxII). Some of the organisms that have been shown experimentally to employ the GlxI system are pathogenic to plants, animals and even humans. The selective inhibition of the pathogen Glyoxalase systems might lead to the discovery of valuable therapeutic agents for curing several bacterial and parasitic infections as well as to the discovery of potential antitumor drugs [230, 231].

The design of inhibitors of GlxI and GlxII has been widely investigated and developed in terms of transport across the membrane, cell targeting and metabolic stability. The inhibition studies of small compounds such as tropolone derivatives, as well as transition state analogues that resemble the suggested enediol(ate) enzyme intermediate produced in the chemical mechanism of GlxI suggest that the hydrophobicity of the inhibitor directly affects its binding ability to the catalytic pocket of the enzyme [232-235]. It has been previously determined that the hydrophobicity of a GlxI inhibitor is reversely related to its inhibitory constant (K_i), inhibition values (IC_{50}) and binding affinity [232]. For example, *S*-octylglutathione (**6-1**, $IC_{50} = 0.52 \pm 0.01$ mM) was found to have a higher binding affinity and lower IC_{50} towards *E. coli* GlxI than that of *S*-methylglutathione ($IC_{50} = 2.97 \pm 0.45$ mM) [174] (Figure 6.1). However, these GSH moieties were found to be unstable in the presence of γ-glutamyl-transpeptidase (γ-GT) *in vivo*, such that the γ-Glu-Cys bond was rapidly cleaved [236]. Further investigations on the modification of *S-p*-bromobenzyl glutathione (**6-2**) by replacement of the γ-carbon of glutamate to a urea-based isosteric analogue (**6-3**) indicated that the hydrolytic breakdown by γ-GT was reduced [237]. Another disadvantage of these *S*-linked moieties is the lack of selectivity toward target cells such as tumor cells [115, 231, 237].

Therefore, analogues that are considered to have some semblance to the reactive intermediate or the transition state of the enzyme, such as hydroxamate moieties, have been explored as potential antitumor drugs. These inhibitors have been termed enediol(ate) analogues and provide tight binding with GlxI due to the chelation of the hydroxamate moieties to a divalent metal of the enzyme, thus resembling the interactions of the metal-intermediate complex [238]. These *S*-linked hydroxamate moieties can function as competitive inhibitors of GlxI but have also been found in some cases to be substrates of GlxII [114, 115, 239]. Due to hydrolysis by GlxII, the half-life of these compounds in biological media containing GlxII is reduced. It was found that tumor cells possess low expression of GlxII; hence, the hydrolysis of compound **6-4** by GlxII is decreased, providing good selectivity to the tumor cells compared to normal cells [237]. Attempts to develop improved stability of these inhibitors toward hydrolysis by GlxII, the carbon-based hydroxamate linkage were discovered to prolong the lifetime of the inhibitor [235]. The replacement of the sulfur atom by carbon in a "transition-state" analogue such as compound **6-5** ($K_i = 1.9 \pm 0.2$ μM with yeast GlxI) was found to provide resistance against GlxII hydrolysis [238]. The charged carboxylate groups of γ-glutamate and glycyl residues at physiological pH have also been found to hinder cell penetration [237, 240]. This problem can be overcome by esterification of these charged groups, which was proven to facilitate compound access across the cell membrane [240]. Recently, the most effective competitive inhibitor at the nanomolar level was discovered, which is compound **6-6** ($K_i = 6.17 \pm 1.64$ nM with yeast GlxI) [241]. The reversal of atom order in hydroxamate, as for compound **6-7** ($K_i = 123 \pm 5.36$ nM with yeast GlxI) also seemed to help to avoid rapid hydrolysis of γ-GT without use of a urea-isostere functional group [241]. This compound, however, exhibits a higher K_i than hydroxamate compound **6-6**, suggesting that appropriate conformation of inhibitor is necessary for the binding ability to the active site [241].



6-1: X = CH_2, R = $(CH_2)_7CH_3$
6-2: X = CH_2, R = $CH_2(C_6H_4)Br$
6-3: X = NH, R = $CH_2(C_6H_4)Br$
6-4: X = S, R = CH_3
6-5: X = CH_2, R = CH_3
6-6: X = CH_2, R = $CH_2(C_6H_4)Br$
6-7

Figure 6.1: The designed inhibitors for GlxI that developed from the structure of GSH (compounds **6-1**–**6-3**) [241] and the suggested reactive intermediate/transition state structure of the likely enedio(late) (compounds **6-4**–**6-7**) [232].

In an attempt to investigate synthetic inhibitors of Zn^{2+}-activated and non Zn^{2+}-activated (being selectively Ni^{2+}/Co^{2+}-activated) GlxI, two transition state analogues, *S*-(*N*-hydroxy-*N*-methylcarbamoyl)glutathione and *S*-(*N*-hydroxy-*N*-phenylcarbamoyl)glutathione were synthesized in this thesis by following the previous reports of Hamilton and Creighton, 1992 and Murthy *et al*., 1994 [114, 115] (Figure 6.2). The enzymatic assay with these inhibitors was performed with *Clostridium acetobutylicum* GlxI (CLO GlxI), *Saccharomyces cerevisiae* GlxI, and *Pseudomonas aeruginosa* GloA2 and GloA3. A comparison of the inhibitory constants of these inhibitors with Ni^{2+}/Co^{2+}-activated GlxI (CLO GlxI and *P. aeruginosa* GloA2) and Zn^{2+}-activated enzymes (yeast GlxI and *P. aeruginosa* GloA3) will be discussed in terms of the hydrophobic interactions occurring in the inhibitor-enzyme complex. As well, the binding affinity of the inhibitor is considered in relation to structural considerations. The reaction mechanism of the human GlxI that has been proposed and has been based on X-ray crystallographic structures with bound benzyl-glutathione, *S-p*-nitrobenzyloxycarbonyl-glutathione and *S*-(*N*-hydroxy-*N*-*p*-iodophenylcarbamoyl)glutathione (PDB: 1FRO, 1QIP and 1QIN, respectively) as well as theoretical calculations is discussed herein as an example of the likely mechanism for the Zn^{2+}-activated GlxI enzymes [44, 46, 48, 49]. The reaction mechanism for Ni^{2+}/Co^{2+}-activated enzymes is proposed, using *E. coli* GlxI as an example, and is compared to the mechanism of the Zn^{2+}-activated enzymes.

Even though several experiments have indicated that only one out of the two asymmetric active sites is functional, these active sites are formed similarly by the same sets of amino acid sequences (homodimer) [19, 45]. It will be interesting to investigate the factors that contribute to this asymmetric arrangement and to determine how the substrate chooses to enter and be converted to product in one functional active site. A better understanding of the

enzymatic system would provide valuable information on inhibitor design and thus support drug development targeting pathogenic bacteria.

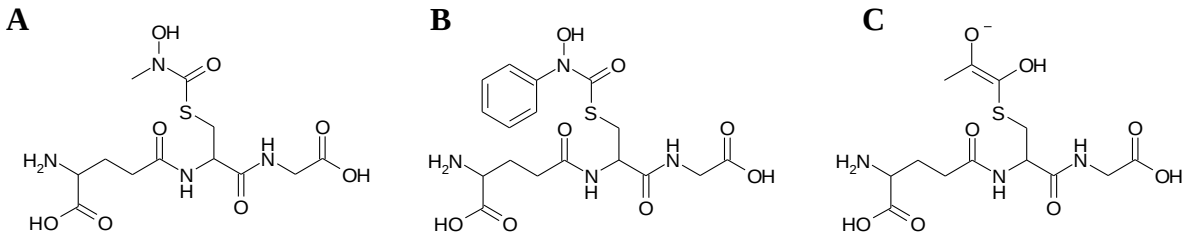


Figure 6.2: The structure of **(A)** *S*-(*N*-hydroxy-*N*-methylcarbamoyl)glutathione and **(B)** *S*-(*N*-hydroxy-*N*-phenylcarbamoyl)glutathione in comparison with **(C)** the hypothetical enediolate intermediate believed to occur in the reaction mechanism of Glyoxalase I reaction.

6.2: Reagents, Materials and Instrumentation

All reagents, materials and equipment used in the experiments in this chapter are listed below otherwise they are included in the previous chapters. Organic chemicals for GlxI inhibitor synthesis such as 4-chlorophenyl chloroformate, *N*-methylhydroxylamine HCl and *N*-phenylhydroxylamine were received from Sigma Chemical Company (St. Louis, MO). Vanillin was obtained from J.T. Baker Chemical Co. (Phillipsburg, NJ), while ninhydrin was obtained from Peirce Chemical Company (Rockford, IL). The vanillin staining solution for TLC analysis was prepared by dissolving vanillin (2.6 g) in the mixture of ethanol (90 mL), acetic acid (2 mL) and sulfuric acid (3.4 mL). The ninhydrin staining solution was prepared by dissolving 0.2% (w/v) ninhydrin in ethanol.

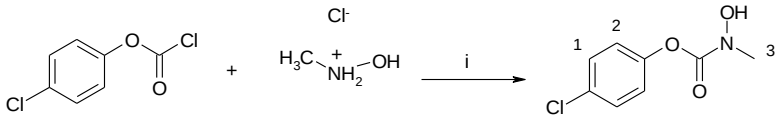
The purification of the GlxI inhibitor was performed using a high performance liquid chromatography (HPLC) system with a Waters μBandapak C18 reverse phase radial compression column (25 × 100 nm) on a Waters 625 LC system with a 994 programmable photodiode array detector from Waters Corporation (Chicago, IL). The purified GlxI inhibitors were identified using electrospray mass spectrometry (ESI-MS) on a Micromass Q-TOF Ultima^{TM} Global (Waterloo Chemical Analysis Facility, University of Waterloo, Waterloo, ON) and nuclear magnetic resonance (NMR) on a Bruker 300 MHz spectrometer with deuterated solvents manufactured by Cambridge Isotopes Laboratories (Andover, MA).

6.3: Experimental Protocols

6.3.1: Synthesis of Glyoxalase I Inhibitors

Two inhibitors, *S*-(*N*-hydroxy-*N*-methylcarbamoyl)glutathione and *S*-(*N*-hydroxy-*N*-phenylcarbamoyl)glutathione were synthesized similarly. The general protocol was initiated by the reaction of 4-chlorophenyl chloroformate and hydroxylamine derivatives to form *N*-hydroxycarbamate 4-chlorophenyl ester. This active ester then subsequently reacted with glutathione, forming a final product containing the *N*-hydroxycarbamoyl moiety. The difference between these inhibitors is the *N*-substitution, *N*-methyl and *N*-phenyl functional groups, which affects the GlxI activity according to the binding affinity in the hydrophobic pocket of the enzyme.

Synthesis of N-hydroxy-N-methylcarbamate 4-chlorophenyl ester (6-3)



6-8 **6-9** **6-10**

Figure 6.3: The synthesis of *N*-hydroxy-*N*-methylcarbamate 4-chlorophenyl ester. (i) K_2CO_3 in diethyl ether and water.

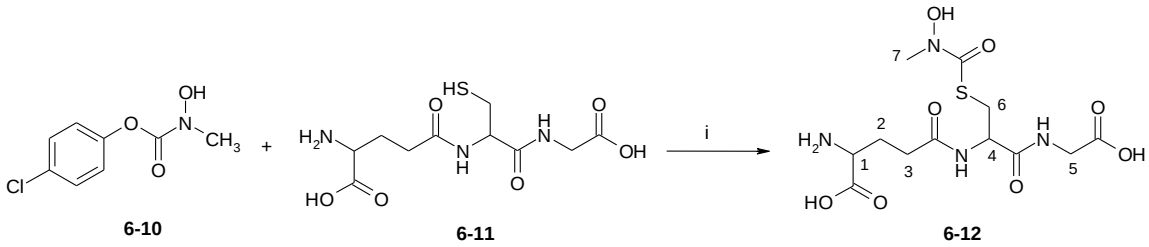
The synthesis of *N*-hydroxy-*N*-methylcarbamate 4-chlorophenyl ester (**6-10**) was performed according to the work of Hamilton and Creighton, 1992 [114]. A solution of 4-chlorophenyl chloroformate (**6-8**, 4.8 g, 25 mmol) in diethyl ether (20 mL) was added drop wise to an ice-cold mixture of *N*-methylhydroxylamine HCl (**6-9**, 2.4 g, 29 mmol) and K_2CO_3 (4.0 g, 29 mmol) in diethyl ether (80 mL) and water (1 mL) over 30 minutes. The reaction was stirred for 1 hour at room temperature. The reaction progress was monitored by silica gel TLC using diethyl ether:hexane (1:1) as a running solvent system (**6-8**, R_f = 0.23; **6-10**, R_f = 0.74). The ether layer was removed and washed twice with water (20 mL) before evaporating to dryness. The product *N*-hydroxy-*N*-methylcarbamate 4-chlorophenyl ester (**6-10**) was crystallized in a minimum volume of methanol and water. The white needle-like crystals of **6-10** were washed with water several times before removing solvent using a lyophilizer (96% yield). The melting point of **6-10** was 118–119 °C (120–121 °C from literature [114]). The

spectroscopic characteristics (^{1}H NMR and ESI-MS) of this compound were in full agreement with those previously reported [114].

^{1}H NMR (300 MHz, $CDCl_3$): δ 7.33–7.05 (m, 4H, H1, H2), 3.34 (s, 3H, H3)

ESI-MS (pos): m/z = 202.0 (MH^+, 100%); 240.0 ($M+K^+$, 70%); 425.0 ($MH+K^+$+8Na^+, 40%); calculated (M) = 201.6 (M)

Synthesis of S-(N-hydroxy-N-methylcarbamoyl)glutathione (6-5)



6-10 **6-11** **6-12**

Figure 6.4: The synthesis of *S*-(*N*-hydroxy-*N*-methylcarbamoyl)glutathione. (i) MeOH and NaOH

N-Hydroxy-*N*-methylcarbamate 4-chlorophenyl ester (**6-10**, 0.20 g, 1 mmol) and glutathione (**6-11**, 0.31 g, 1 mmol) were dissolved in the degassed solvent mixture of methanol (4 mL) and 1 N NaOH (2 mL). The reaction was allowed to stir under argon at room temperature for one week. The reaction progress was monitored by silica gel TLC using *n*-propanol:acetic acid:water (10:1:5) as a solvent system (**6-10**, R_f = 0.95; **6-11**, R_f = 0.34; **6-12**, R_f = 0.44; GSSG, R_f = 0.11). The acid HCl (1 M) was added to the reaction mixture until the pH reached 3.5 to quench the reaction. The mixture was extracted three times with diethyl ether (10 mL). The aqueous layer was concentrated under reduced pressure and purified by reverse-phase HPLC using a Waters μBandapak C18 compression column (25 × 100 mm) and a solvent gradient as shown in Table 6.1. The purified product, *S*-(*N*-hydroxy-*N*-methylcarbamoyl)glutathione (**6-12**), was eluted with a retention time (t_R) of 35 min (35% yield). The spectroscopic characteristics (^{1}H NMR and ESI-MS) of this compound were in full agreement with those previously reported [114].

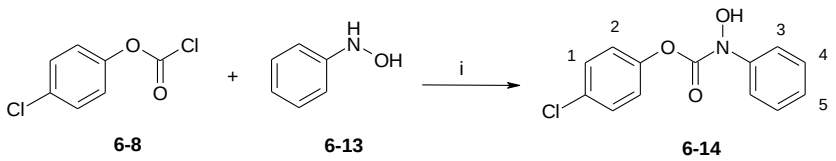
Table 6.1: Solvent gradient used to purify S-(*N*-hydroxy-*N*-methylcarbamoyl)glutathione.

Time (min)	*Flow rate (mL/min)*	*0.25% Acetic acid in water*
0	3	100
10	3	100
11	1	100
100	1	100

^{1}H NMR (300 MHz, D_2O): δ 4.48 (m, 1H, H4), 3.81 (s, 2H, H5), 3.65 (t, J = 6.1 Hz, 1H, H1), 3.28–3.13 (m, 1H, H6a), 3.11 (s, 3H, H7), 3.03–2.76 (m, 1H, H6b), 2.38 (m, 2H, H3), 2.00 (m, 2H, H2)

ESI-MS (pos): m/z = 381.1 (MH^+, 100%); calculated (M) = 380.4

Synthesis of N-hydroxy-N-phenylcarbamate 4-chlorophenyl Ester (6-7)



6-8 **6-13** **6-14**

Figure 6.5: The synthesis of *N*-hydroxy-*N*-phenylcarbamate 4-chlorophenyl ester. (i) K_2CO_3 in diethyl ether and water.

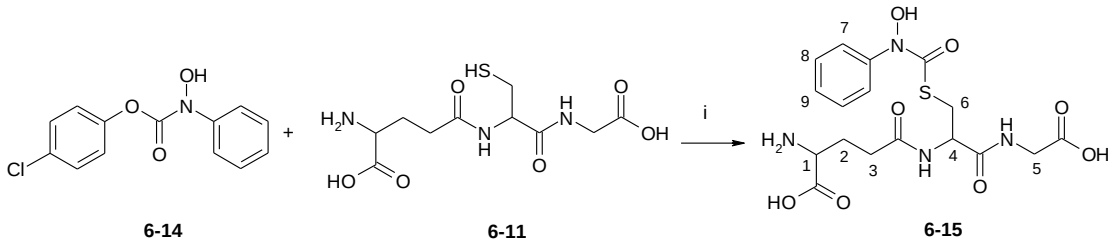
The synthesis of *N*-hydroxy-*N*-phenylcarbamate 4-chlorophenyl ester followed the work of Murthy *et al*. (1994) [115]. Yellow crystals of *N*-phenylhydroxylamine (**6-13**, 2.02 g, 18.5 mmol) and finely powdered K_2CO_3 (1.28 g, 9.3 mmol) were dissolved in a mixture of diethyl ether (20 mL) and water (0.5 mL) and placed on ice. To the ice-cold *N*-phenylhydroxylamine/K_2CO_3 solution, a solution of 4-chlorophenyl chloroformate (**6-8**, 3.5 g, 18 mmol) in diethyl ether (20 mL) was added drop wise over 20 minutes. The reaction was allowed to stir for 3 hours at room temperature. The reaction progress was monitored by silica gel TLC using diethyl ether:hexane (1:1) as a solvent system with UV light or vanillin detection (**6-8**, R_f = 0.88; **6-13**, R_f = 0.36; **6-14**, R_f = 0.41). The ether layer was removed and the reaction was washed twice with water (20 mL) before evaporating the solvent to dryness. The crude product was crystallized in a minimum volume of methanol and water. After recrystallization, the crystals were washed with water several times. The remaining water was removed by lyophilization and the purified product, **6-14**, was obtained as white needle-like

crystals (98% yield). The melting point of **6-14** was 114–115 °C (116–117 °C from literature [115]). The spectroscopic characteristics (^{1}H NMR and ESI-MS) of this compound were in full agreement with those previously reported [115].

^{1}H NMR (300 MHz, $CDCl_3$): δ 7.66–7.56 (d, J = 7.3 Hz, 2H, H3), 7.47–7.34 (m, 4H, H1, H4), 7.27–7.15 (m, 3H, H2, H5)

ESI-MS (pos): m/z = 286.0 ($M+Na^+$, 100%); 549.1 ($2M+Na^+$, 25%); calculated (M) = 263.7

Synthesis of S-(N-hydroxy-N-phenylcarbamoyl)glutathione (6-8)



6-14 **6-11** **6-15**

Figure 6.6: The synthesis of *S*-(*N*-hydroxy-*N*-phenylcarbamoyl)glutathione. (i) MeOH and NaOH.

Glutathione (**6-11**, 307 mg, 1 mmol) was dissolved in a degassed solution of 1 N NaOH (2 mL) and brought to nearly a cloudy point by adding ethanol (2 mL). To the glutathione solution, the solution of *N*-hydroxy-*N*-phenylcarbamate 4-chlorophenyl ester (**6-14**, 264 mg, 1 mmol) in ethanol (2 mL) was added drop wise over 5 min. The reaction was stirred at room temperature under argon for one week. The reaction progress was monitored by silica gel TLC using *n*-propanol:acetic acid:water (10:1:5) as a solvent system with UV light and ninhydrin detection (**6-11**, R_f = 0.45; GSSG, R_f = 0.11; **6-14**, R_f = 0.97; **6-15**, R_f = 0.57). The reaction was then quenched by addition of HCl (1 M) to the final pH of ~3.5. The solvent was then removed *in vacuo*. The crude product was suspended in water (10 mL) and extracted with diethyl ether (3 × 10 mL) to remove 4-chlorophenol and unreacted *N*-hydroxy-*N*-phenylcarbamate 4-chlorophenyl ester. Diethyl ether was back extracted with 10 mL water before combining the aqueous phases and drying to minimum volume. The crude product was subjected to reverse-phase HPLC using a Waters μBandapak C18 compression column (25 × 100 mm) with 0.25% acetic acid in water and methanol solvent system (Table 6.2). The

purified product, *S*-(*N*-hydroxy-*N*-phenylcarbamoyl)glutathione (**6-15**), was eluted at t_R 29 min with a 11% yield. The spectroscopic characteristics (^{1}H NMR and ESI-MS) of this compound were in full agreement with those previously reported [115].

Table 6.2: Solvent gradient used to purify *S*-(*N*-hydroxy-*N*-phenylcarbamoyl)glutathione.

Time (min)	*Flow rate (mL/min)*	*0.25%Acetic acid in water*	*%MeOH*
0	5	100	0
15	5	100	0
40	5	0	100
55	5	0	100

^{1}H NMR (300 MHz, D_2O): δ 7.44–6.94 (m, 5H, H7, H8, H9), 4.56 (m, 1H, H4), 3.82 (s, 2H, H5), 3.65 (t, J = 6.6 Hz, 1H, H1), 3.35–3.27 (q, J = 4.9, 14.8 Hz, 1H, H6b), 3.11–3.01 (q, J = 8.2, 14.8 Hz, 1H, H6a), 2.38 (m, 2H, H3), 2.02 (m, 2H, H2)

ESI-MS (pos): m/z = 443.1 (MH^+, 100%); calculated (M) = 442.5

6.3.2: Enzymatic Assay for Inhibition Studies

Protein expression, induction and purification of *C. acetobutylicum* GlxI (CLO GlxI) as well as *Pseudomonas aeruginosa* GloA2 and GloA3 were performed as previously stated in Chapter 2 and 3, respectively. CLO GlxI was prepared in its apo-form and was incubated with 5 equivalents of $NiCl_2$ overnight at 4 °C. GloA2 and GloA3 were expressed in the *E. coli* system with additional 1 mM $NiCl_2$ and $ZnCl_2$, respectively. No additional metals were added to the purified enzymes.

The enzymatic assay was performed as previously described (Chapter 2) using enzyme in KPB (pH 6.6) and the hemithioacetal substrate, MG-GSH (0.04–1 mM), at room temperature (200 μL assay). The inhibitory constant (K_i) of *S*-(*N*-hydroxy-*N*-methylcarbamoyl)glutathione (80–200 μM) and *S*-(*N*-hydroxy-*N*-phenylcarbamoyl)glutathione (5–40 μM) was determined using Ni^{2+}-reconstituted CLO GlxI (0.47 μg) and commercial yeast GlxI (4.25 ng). The K_i of *S*-(*N*-hydroxy-*N*-phenylcarbamoyl)glutathione (5–40 μM) in GloA2 (25 ng) and GloA3 (5 ng) was investigated under similar conditions. The enzymatic reaction was performed in triplicate on a 96-well plate using the plate reader. The initial rate data were fitted, assuming Michaelis-Menten kinetics with competitive, noncompetitive and

uncompetitive inhibition on a GraphPad Prism version 5.00 for Windows (GraphPad Software, San Diego California USA, www.graphpad.com).

6.3.3: Structural Investigation of Inhibitor-Bound Glyoxalase I

The purified CLO GlxI (3 mg/ml) in buffer containing 50 mM HEPES (pH 7.0), 200 mM KCl, 10 mM Met and 10% (v/v) glycerol and the inhibitor *S*-(*N*-hydroxy-*N*-phenylcarbamoyl)glutathione were sent to the New York Structural Genomics Research Consortium (NYSGXRC) for the X-ray crystallographic investigation of the inhibitor-bound enzyme. As well, the purified *P. aeruginosa* GloA2 (4.8 mg/mL) in buffer containing 20 mM Tris (pH 7.5), 150 mM NaCl and 10% (v/v) glycerol and the purified *P. aeruginosa* GloA3 (5.7 mg/mL) in buffer containing 20 mM Tris (pH 7.5), 150 mM NaCl, 10 mM DTT and 10% (v/v) glycerol with the same inhibitor were sent to Dr. Charles Bond, University of Western Australia (WA, Australia).

Structural visualizations and molecular images were generated by the UCSF Chimera package from the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco (supported by NIH P41 RR-01081) [1] and the PyMOL Molecular Graphics System, version 1.2r3pre, Schrödinger, LLC (http://www.pymol.org/).

6.4: Results and Discussion

6.4.1: Synthesis of Glyoxalase I Inhibitors

In the synthesis of *N*-hydroxycarbamate 4-chlorophenyl esters (both *N*-hydroxy-*N*-methylcarbamate 4-chlorophenyl ester (**6-10**) and *N*-hydroxy-*N*-phenylcarbamate 4-chlorophenyl ester (**6-14**)), the starting materials (4-chlorophenyl chloroformate (**6-8**), K_2CO_3, *N*-phenylhydroxylamine (**6-13**) and *N*-methylhydroxylamine (**6-9**)) did not dissolve appreciably in diethyl ether, the major component of the solvent mixture in this reaction. Despite this, TLC results showed that there was no unreacted compound **6-8** left in the mixture, suggesting a completed reaction. The isolation of the product by crystallization from a mixture of methanol and water was successful with a high recovery yield (> 90%). The structures of these esters were confirmed by ^{1}H NMR (Spectra A7.1 and A7.3, Appendix 7) and ESI-MS

(Figure A6.1 and A6.3, Appendix 6) and characteristics were found to be in complete agreement with previous spectroscopic characterization of these compounds [114].

These *N*-hydroxycarbamate 4-chlorophenyl esters were reacted with glutathione to form derivatives of *N*-hydroxycarbamoyl-glutathione. The solvent mixture for this reaction required degassing to avoid oxidation of GSH (GSSG). Even though the reaction was incubated under argon at room temperature for a week, unreacted GSH was still observed. Product isolation by use of reverse phase HPLC (C18) suggested that GSH and S-(*N*-hydroxy-*N*-methylcarbamoyl)glutathione (**6-12**) eluted with a similar retention time, such that unresolved separation was detected as a broad peak when high flow rates were employed (3 mL/min) (Figure 6.7a). By applying a solvent gradient with a slower flow rate, HPLC chromatograms exhibited a better separation of peaks leading to improved isolation of GSH and compound **6-12** (Figure 6.7b–d). As well, utilization of degassed buffer seemed to improve the recovery yield of the product, in which only a trace amount of unreacted GSH was observed after incubation for a week (Figure 6.7e). The compound **6-12** was confirmed by the analyses of ^{1}H NMR (Spectrum A7.2, Appendix 7) and ESI-MS (Figure A6.2, Appendix 6) experiments and characteristics were found to be in complete agreement with previous spectroscopic characterization of these compounds [114].

The purification of S-(*N*-hydroxy-*N*-phenylcarbamoyl)glutathione (**6-15**) was easier than that of the compound **6-12**. The aryl group of compound **6-15** possessed better interactions with the C18 column than the methyl group of compound **6-12**. The elution times of GSH and compound **6-15** were well separated (t_R of GSH = 10 min and t_R of compound **6-15** = 29 min), so isolation using reverse phase HPLC (C18) with a high flow rate (5 mL/min) was acceptable (Figure 6.8). The compound **6-15** was confirmed by ^{1}H NMR (Spectrum A7.4, Appendix 7) and ESI-MS (Figure A6.4, Appendix 6) and characteristics were found to be in complete agreement with previous spectroscopic characterization of these compounds [115].

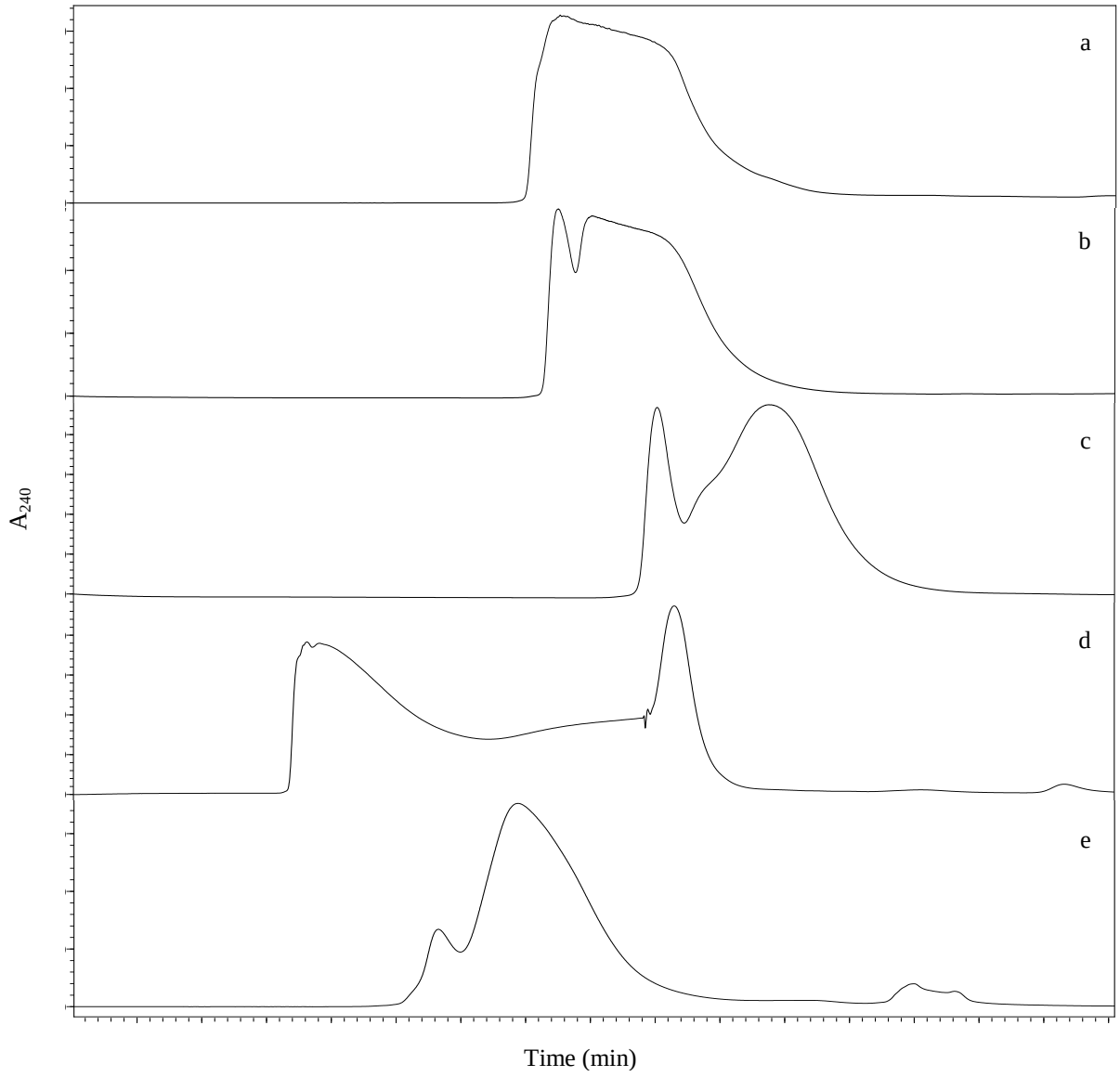


Figure 6.7: The purification of S-(*N*-hydroxy-*N*-methylcarbamoyl)glutathione (**6-12**) using reverse phase HPLC (C18) was monitored at 240 nm with a solvent system as shown in Table 6.1. The HPLC chromatograms of the reaction, which was permitted to react over 7 days without degassing, show a broad peak of unreacted GSH and compound **6-12** when applying a flow rate of (a) 3 mL/min, (b) 1 mL/min, (c) 0.5 mL/min and (d) a gradient of 3 mL/min for 15 min and 0.5 mL/min for 80 min. These HPLC chromatograms were compared with the reaction in the degassed solvent mixture using a gradient flow rate of 3 mL/min for 10 min and 1 mL/min for 80 min (e).

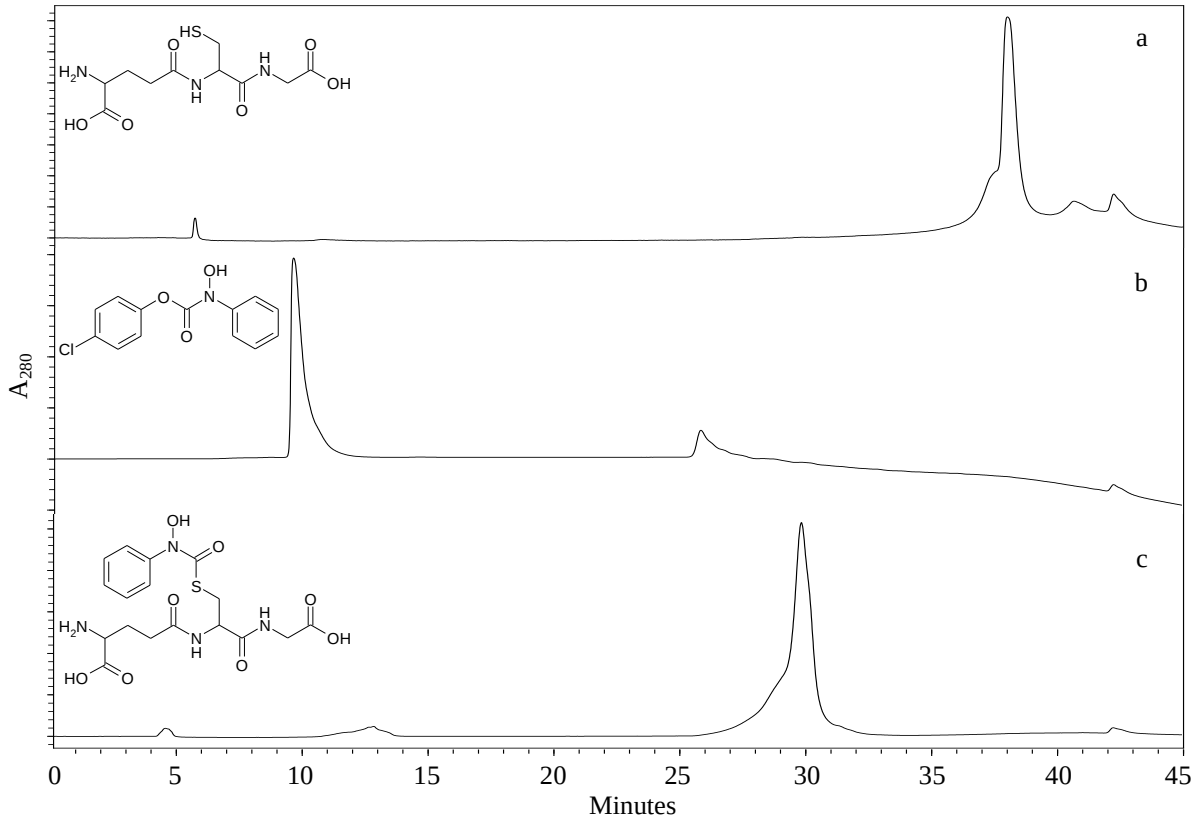


Figure 6.8: The purification of S-(*N*-hydroxy-*N*-phenylcarbamoyl)glutathione (**6-15**) using reverse phase HPLC (C18) was monitored at 280 nm with a solvent system as shown in Table 6.2. The HPLC chromatograms of the starting materials including (a) GSH (**6-11**) and (b) *N*-hydroxy-*N*-phenylcarbamate 4-chlorophenyl ester (**6-14**) were compared with that of (c) S-(*N*-hydroxy-*N*-phenylcarbamoyl)glutathione (**6-15**).

6.4.2: Effect of Inhibitor on Glyoxalase I Kinetics

The inhibition of GlxI from yeast, *Clostridium acetobutylicum* (CLO GlxI) and *Pseudomonas aeruginosa* (GloA2 and GloA3) by *S*-(*N*-hydroxy-*N*-methylcarbamoyl)glutathione (**6-12**) and *S*-(*N*-hydroxy-*N*-phenylcarbamoyl)glutathione (**6-15**) showed that they are competitive inhibitors (Figure 6.9). The inhibition studies on the commercial yeast GlxI, a Zn^{2+}-activated enzyme, showed that the K_i of compound **6-15** (19 ± 1 μM) was lower than compound **6-12** (171 ± 19 μM), suggesting that compound **6-15** possesses higher binding affinity than that of compound **6-12** (Table 6.3). It appears that the phenyl group of *S*-link hydroxamate inhibitor fits into the hydrophobic pocket of the enzyme, thus forming stronger interactions than those of the methyl group of compound **6-12**. The inhibition on CLO GlxI (Ni^{2+}/Co^{2+}-activated enzyme) also followed the similar trend, where inhibition of the compound **6-15** is stronger than that of compound **6-12** (K_i of **6-12** was 200 ± 39 μM and

K_i of **6-15** was 24 ± 6 μM). It is also found that both inhibitors exhibited a lower K_i on Zn^{2+}-activated GlxI, hence possessing higher binding affinity to the Zn^{2+}-activated enzyme than the Ni^{2+}/Co^{2+}-activated enzyme (Table 6.3). These results suggested that the residues in the extra loops and segments (loops B, C and D, Chapter 3) that only exist in the enzyme with the longer amino acid sequence (Zn^{2+}-activated GlxI) might participate in the formation of a more prominent hydrophobic pocket. These observations were also confirmed by inhibition produced by compound **6-15** on GloA2 (Ni^{2+}/Co^{2+}-activated GlxI) and GloA3 (Zn^{2+}-activated GlxI) from *P. aeruginosa* (K_i of GloA3 was 2.2 ± 0.1 μM and K_i of GloA2 was 23.7 ± 1.4 μM). Interestingly, the large GlxI (yeast GlxI) has higher K_i than the small GlxI (GloA3), even though both are Zn^{2+}-activated enzymes. It is possible that the hydrophobicity of the residues that forms the hydrophobic pocket in *P. aeruginosa* GloA3 might be higher than those of yeast GlxI. However, a molecular structure for the enzymes would be needed to substantiate this suggestion.

Table 6.3: The summary of the enzymatic assay of the commercial yeast GlxI, *Clostridium acetobutylicum* GlxI and *Pseudomonas aeruginosa GlxI*, GloA2 and GloA3, in the presence of the competitive inhibitors *S*-(*N*-hydroxy-*N*-methylcarbamoyl)glutathione (**6-12**) and *S*-(*N*-hydroxy-*N*-phenylcarbamoyl)glutathione (**6-15**). The assays were performed in KPB (pH 6.6) at room temperature with 0.04–1 mM substrate (MG-GSH).

	6-12			***6-15***		
Sample	K_m	K_i	K_m/K_i	K_m	K_i	K_m/K_i
Yeast GlxI	308 ± 29 μM	171 ± 19 μM	1.8	247 ± 12 μM	19 ± 1 μM	13.0
Ni^{2+}-CLO GlxI	276 ± 38 μM	200 ± 39 μM	1.4	334 ± 69 μM	24 ± 6 μM	13.9
Ni^{2+}-GloA2	ND	ND	ND	59.5 ± 4.2 μM	23.7 ± 1.4 μM	2.5
Zn^{2+}-GloA3	ND	ND	ND	171.6 ± 11.8 μM	2.2 ± 0.1 μM	78

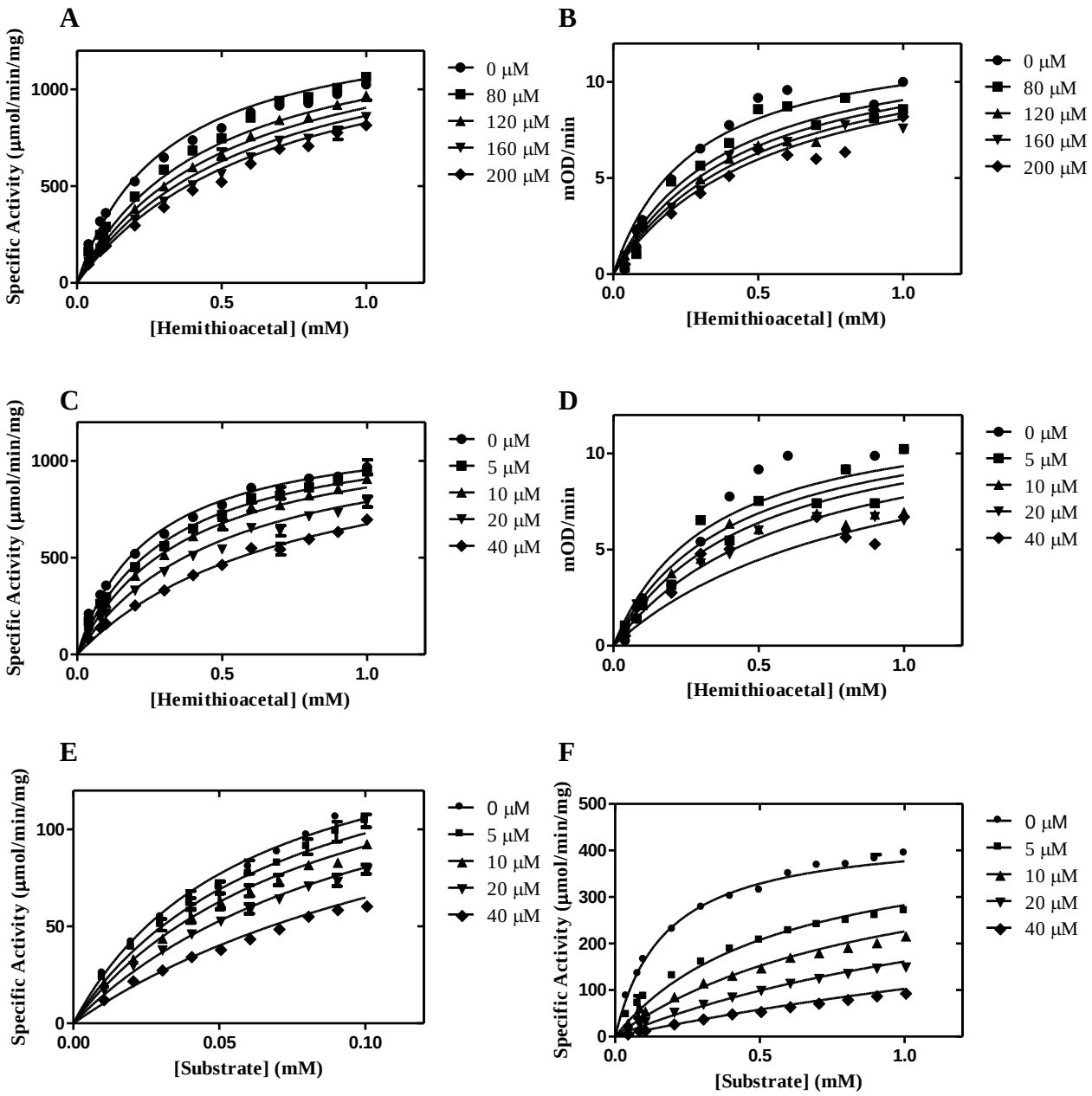


Figure 6.9: The inhibition study using *S*-(*N*-hydroxy-*N*-methylcarbamoyl)glutathione (**A-B**) and *S*-(*N*-hydroxy-*N*-phenylcarbamoyl)glutathione (**C-F**) in the commercial yeast GlxI (**A** and **C**), *Clostridium acetobutylicum* GlxI (**B** and **D**) and *Pseudomonas aeruginosa GlxI*, GloA2 (**E**) and GloA3 (**F**). The assays were performed using 0.01–1 mM substrate (MG-GSH) in 50 mM KPB (pH 6.6) at room temperature.

6.4.3: Structural Investigation of Inhibitor Bound Glyoxalase I

The previous kinetic data on the inhibitors of GlxI indicated that the inhibitory constant decreases with increasing hydrophobicity of the *N*-substituents. The increase in binding affinity by occupying the hydrophobic pocket of the enzyme was found to optimize the interaction between the enzyme and the bound ligands [232]. Therefore, in order to explain the

effect of hydrophobicity of the inhibitor and its interaction with the enzyme, especially the glutathione binding site, it is necessary for us to understand the reaction mechanism. The catalytic pocket of the Zn^{2+}-activated GlxI was examined using the *H. sapiens* GlxI as a model due to its available X-ray crystallographic structure, while the structure of the *E. coli* GlxI was investigated as a representative of the Ni^{2+}/Co^{2+}-activated GlxI enzyme class.

Several inhibitors have been co-crystallized with human GlxI including benzyl-glutathione (B-GSH, PDB: 1FRO), *S*-*p*-nitrobenzyloxycarbonyl-glutathione (NBC-GSH, PDB: 1QIP) and *S*-(*N*-hydroxy-*N*-*p*-iodophenylcarbamoyl)glutathione (HIPC-GSH, PDB: 1QIN) (Figure 6.10). The interactions of these inhibitors with the enzyme suggested that B-GSH might be acting as a substrate-like inhibitor, while HIPC-GSH is a transition state or a reactive intermediate analogue type and NBC-GSH is a product-like compound [44, 46]. The superimposed structures of the human GlxI with these different bound inhibitors suggested particular enzyme movements during catalysis and the proposed reaction mechanism. With the overall structure being described in Chapter 5, it was found that once the inhibitor enters the active sites of human GlxI and two metal-bound water molecules are replaced by the oxygens from the substrate, the loop (residues 152–159) that is located between β7 and β8 moves toward the active site (Figure 6.11A). This loop acts as a lid that covers the catalytic pocket, thus, shielding the substrate from outer solvent. The loop movement also causes a limited catalytic space that influences a change in torsion angle around the bond between the amide nitrogen and the Cα atom of the cysteine residue of the GSH moiety, thus supporting the interactions of the GSH moiety and the enzyme [44, 46]. The loop movement of the enzyme also allows the carboxyl group of the glycine residue in the GSH moiety to form a H-bond with the main chain nitrogen of Met^{157} and Lys^{156} (Figure 6.11B). The Met^{157} residue seems to play a crucial role in substrate binding ability and a mutation at this position (M157A) results in significant loss of activity (exhibiting only 20% of wild-type activity) [242]. Besides, it was suggested that the interaction of the glycine residue of the GSH moiety with the enzyme is a driving force for the dislocation of Glu^{172}, one of the metal binding residues, from the active metal in a catalytic pocket [44, 46]. Other residues that appear to be involved in the interactions with the substrate include the electrostatic interaction of the guanidinium group of Arg^{37} and Arg^{122} with the carboxylate group of the γ-glutamate of GSH as well as a possible H-bond between the amide group from the side chain of Asn^{103} with the carboxylate and the

amino groups of this γ-glutamate. These interactions, however, appear to be only slightly altered upon the binding of the inhibitor. Besides, it has been suggested that the binding of the γ-glutamate of the substrate with the enzyme occurs first, after which the cysteine residue of the substrate swings into the active site. The conformational change of the substrate replaces two metal-bound water molecules and causes the loop that acts as a lid of the enzyme to move toward the active site. The glycine residue of the substrate then forms interactions with the residues in this loop, thus stabilizing the intermediate of the catalytic reaction. This loop is suggested to swing open once the product is formed to allow it to diffuse from the active site.

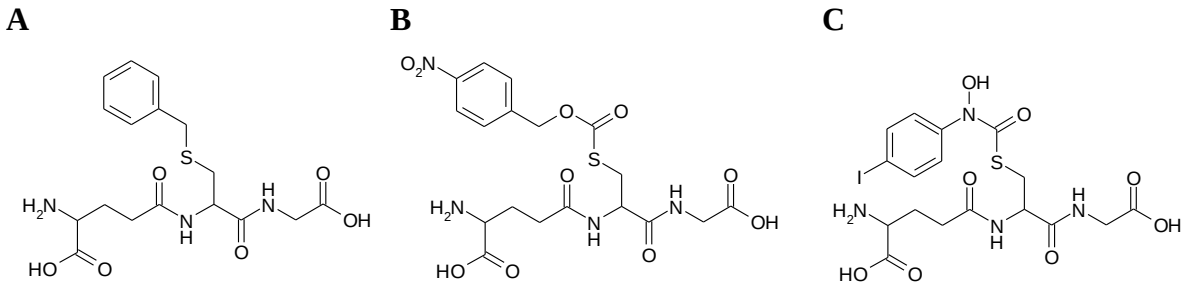


Figure 6.10: The structures of **(A)** benzyl-glutathione (B-GSH), **(B)** *S*-*p*-nitrobenzyloxycarbonyl-glutathione (NBC-GSH) and **(C)** *S*-(*N*-hydroxy-*N*-*p*-iodophenylcarbamoyl)glutathione (HIPC-GSH).

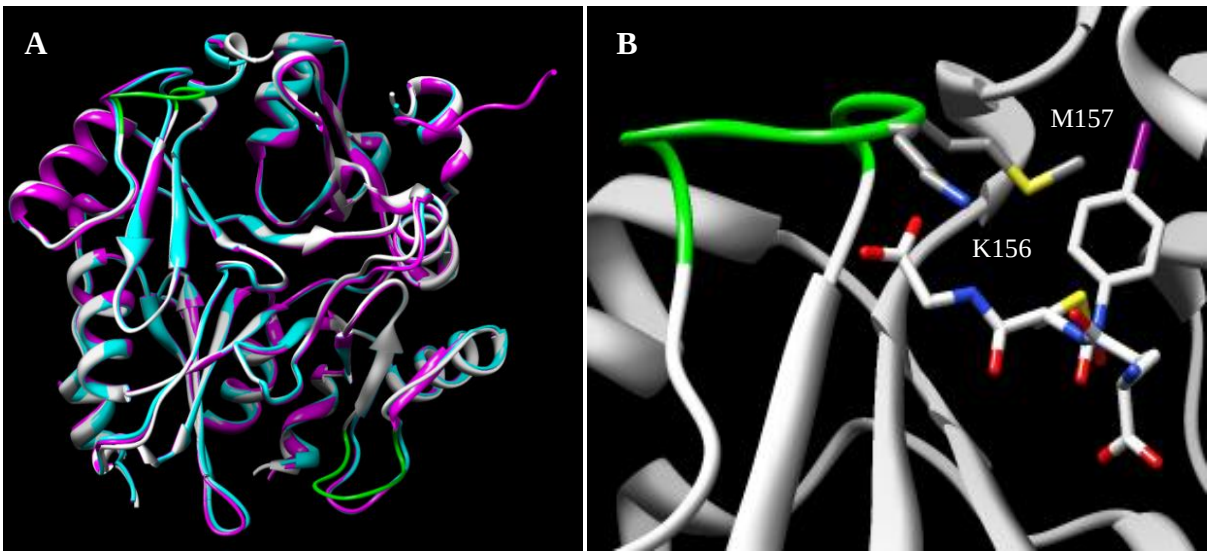


Figure 6.11: The superimposed structures of human GlxI with bound benzyl-glutathione (B-GSH, cyan), *S*-*p*-nitrobenzyloxycarbonyl-glutathione (NBC-GSH, magenta) and *S*-(*N*-hydroxy-*N*-*p*-iodophenylcarbamoyl)glutathione (HIPC-GSH, white) (PDB: 1FRO, 1QIP and 1QIN, respectively) showing **(A)** the loop movement (residues 152–159, green) and **(B)** the interactions of Met^{157} and Lys^{156} with glycine residue of HIPC-GSH. The 3D structures were generated by using the UCSF Chimera program (University of California, San Francisco) [1].

E. coli GlxI, an example of a Ni^{2+}/Co^{2+}-activated GlxI, is a homodimeric enzyme. With the overall structure described in Chapter 4, the active enzyme forms an octahedral geometry in its active site with the activating metal, the ligands surrounding the metal being composed of four metal binding residues (His^{5}, Glu^{56}, His^{74} and Glu^{122}) and two water molecules. The replacements of two metal-bound water molecules with the substrate are similar to those of human GlxI. The superimposed structures of Ni^{2+}-bound *E. coli* GlxI (PDB: 1F9Z) and the enzyme with bound transition state analogue of hydroxamate, TSI (compound **5**, unpublished structure) show high structural similarity with the root mean square deviation (r.m.s.d.) of 0.69 Å (103 Cα pairs). In addition to the loop movement that was observed in human GlxI upon binding of inhibitors, the movements of two loops and the C-terminal fragment are detected in *E. coli* GlxI (Figure 6.12A). The loop that is located between β2 and β3 (residues 34–39) and the loop that is located between β6 and β7 (residues 99–110) move toward the active site, covering the catalytic pocket from above. The movement of the loop residues 99–110 causes a significant change in the position of Arg^{98} and Lys^{104}, where their Cα move 1.386 Å and 4.896 Å, respectively, away from their original positions in the enzyme structure without bound inhibitor. Thus, their side chain amines are able to form H-bonds with the carboxyl group of the glycyl residue in the GSH moiety (Figure 6.12B). The residues Arg^{9} and Asn^{60}, which interact with the carboxylate and the amino groups of γ-glutamate of the GSH moiety, are only slightly altered. The ten C-terminal residues (residues 126–135) that are undetected in the crystal structure of the enzyme without bound hydroxamate analogue are found to locate at one side of the catalytic pocket upon inhibitor binding, suggesting that the C-terminus becomes less flexible upon inhibitor binding, although no direct interaction of this C-terminus and the inhibitor was found. However, interactions among protein residues within the C-terminus (such as Glu126/O–Ala129/N and Asp128/O–Arg131/N) as well as the interactions with the loop residues 99–110 (such as Thr107/N–Leu133/O, Thr107/$O_{\gamma 1}$–Gly134/O and Thr108/$O_{\gamma 1}$–Gly132/O) are detected. These interactions create a network of interactions that likely maintains the hydrophobic environment within the catalytic pocket and possibly causes the structure of the C-terminus to become more defined. Besides, this C-terminus as well as a short loop connecting β3 and β4 (residues 46–54) partially blocks the bound inhibitor from the outside environment. The C-terminus also provides a more hydrophobic environment to the

active site, due to the presence of Leu^{133}, which in turn increases binding affinity of the inhibitor.

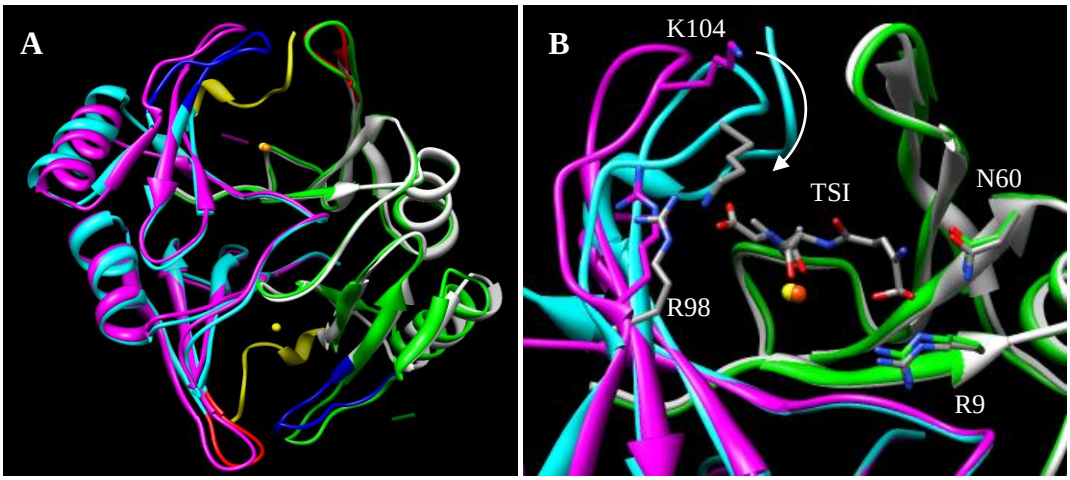


Figure 6.12: (A) The superimposed structures of *E. coli* GlxI (magenta and green, PDB: 1F9Z) and *E. coli* GlxI with bound hydroxamate inhibitor, TSI (cyan and white, unpublished structure) showing the movements of two loops (residues 34–39 (red) and 99–110 (blue)) and the C-terminus (yellow) after binding to the inhibitor. The yellow spheres represent Ni^{2+} atoms in both active sites of the enzyme without bound inhibitor, while the orange sphere represents a Ni^{2+} atom in one active site of the inhibitor bound enzyme. **(B)** The interactions of the GSH moiety with the enzyme in the catalytic pocket including residues Arg^{98} and Lys^{104} that interact with glycine and Arg^{9} and Asn^{61}that form H-bonds with γ-glutamate of inhibitor. The arrow indicates the movement of the residue Lys^{104} upon binding of the hydroxamate analogue. The coordinates are based on our laboratories' unpublished data of these complexes. The 3D structures were generated by using the UCSF Chimera program (University of California, San Francisco) [1].

Interestingly, it was found that the flexibility of the loops in one active site of the Ni^{2+}-bound *E. coli* GlxI (without bound hydroxamate inhibitor) is quite different from the other, such that the r.m.s.d. of the C_α atoms between the two subunits is 0.54 Å with the major difference being observed in two loops regions (residues Gln^{91}–Gly^{94} and Val^{103}–Thr^{108}) (PDB: 1F9Z) [45]. The B-factor variation in these loops (40–70 Å^2) is higher than the average value for the whole protein (19.5 Å^2), especially the B-factors of the residues 91–94 in one subunit that are significant higher than those in the other [45]. These data support our observations on the flexibility of the enzyme, especially the loop residues 99–110 that exhibit a crucial movement of Lys^{104} upon binding to the substrate. In addition, the difference in flexibility of each subunit might lead to the explanation of only one out of two active sites being functional [19]. Analysis of the superimposed structures of Ni^{2+}-bound *E. coli* GlxI without bound inhibitor (PDB: 1F9Z) and its structure with bound hydroxamate inhibitor (unpublished

structure) indicated that one active site employs a larger movement of the loop residues 99–110 and a smaller movement of the loop residues 34–39 (active site A) compared to those of the other active site (active site B) (Figure 6.13). The difference in degree of movement may affect the network connections within the enzyme itself as well as the interactions with the GSH moiety. The shift of the C_α of Lys^{104} in the enzyme without bound inhibitor to its C_α in the enzyme with bound TSI in the active site A (4.896 Å) is larger than the one in the active site B (4.258 Å). The shift of C_α of Arg^{98} from its original position, however, is only slightly different in both active sites (1.386 Å in the active site A and 1.473 Å in the active site B), while its side chain amine possesses a larger movement in the active site A (2.709 Å) comparing to that of the active site B (1.276 Å). The different flexibility of these residues in each active site is possibly crucial for interactions with the glycine residue in the GSH moiety and the loss of these enzyme-substrate interactions would be expected to cause a decrease in activity as what had been observed previously in the human GlxI (the mutation of Met^{157} to Ala in human GlxI, the residue that corresponded to Lys^{104} in *E. coli* GlxI, causes significant loss in activity [242]). As well, the substrate might choose to enter the site with higher flexibility of the loop that acts as a lid to cover the active site since it is possibly easier and quicker to form a conformational change. Nevertheless, the different flexibility of the loops in these two subunits might be a result of the packing of the enzyme in the crystal lattice.

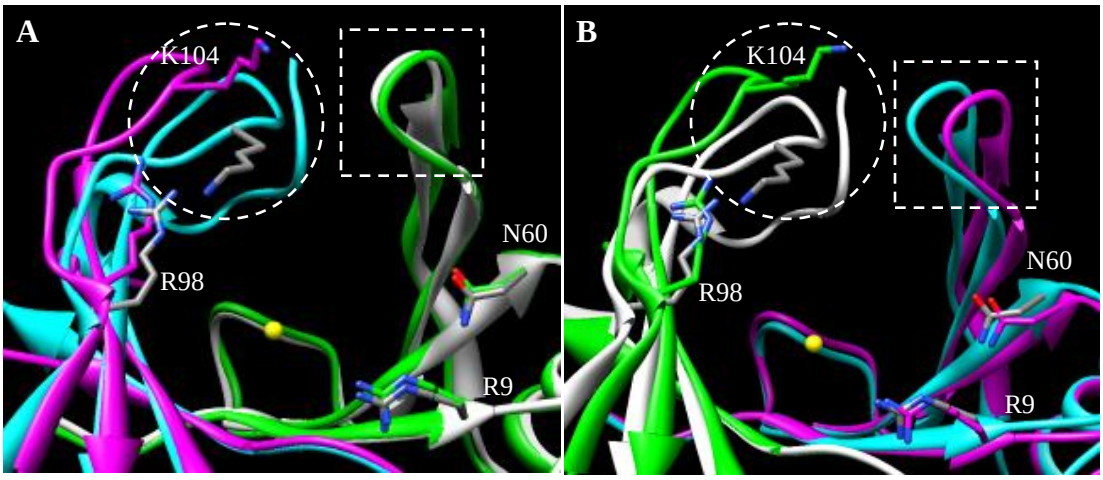


Figure 6.13: The superimposed structures of *E. coli* GlxI (magenta and green, PDB: 1F9Z) and *E. coli* GlxI with bound hydroxamate inhibitor, TSI (cyan and white, unpublished structure) showing **(A)** the active site A that contains a large movement of the loop residues 99–110 and small movement of the loop residues 34–39 and **(B)** the active site B that contains small movement of the loop residues 99–110 and the larger movement of the loop residues 34–39. The residues involved in GSH moiety binding are represented as sticks (Arg^9, Asn^{60}, Arg^{98} and Lys^{104}) and the bound Ni^{2+} atom is represented as a yellow sphere. The loop containing residues 99–110 is

represented as a circle, while the loop containing residues 34–39 is represented as a square. The coordinates are based on our laboratories unpublished data of these complexes. The 3D structures were generated by using the UCSF Chimera program (University of California, San Francisco) [1].

Compared to a domain-swapped dimer of a two-domain structure of a small GlxI (i.e. human and *E. coli* GlxI), the *Saccharomyces cerevisiae* enzyme is a large GlxI, forming a monomeric four-domain arrangement. Since no X-ray crystallographic structure of this large GlxI is available, the multiple sequence alignment of yeast GlxI with others from various organisms was examined (Figure 6.14). Yeast GlxI consists of an N-terminal domain (residues 1–173) that possesses high sequence similarity (64%) and identity (44%) to its C-terminal domain (residues 184–326). These two domains are connected by an extended loop (residues 174–183). The overall structure of yeast GlxI was suggested to be a result of gene duplication and fusion from a dimeric progenitor [151]. This type of protein folding is also found in the malaria parasite, *Plasmodium falciparum,* GlxI [39] and the *Pseudomonas* biphenyl-cleaving extradiol dioxygenase, which is also a member of the βαββββ superfamily of proteins [145]. Three-dimensional structure prediction and mutagenesis studies have led to the suggestion that the yeast enzyme possesses two active sites within a monomer, with each active site being created by half of the N-terminal domain and another half from the C-terminal domain [151]. One catalytic pocket contains four predicted metal binding residues including His^{25}, Glu^{89}, His^{269} and Glu^{318}, while another site is likely formed by His^{117}, Glu^{163}, His^{185} and Glu^{242} [151]. Among these, Glu^{163} and Glu^{318} resemble Glu^{172} in human GlxI, the crucial metal binding ligand that has been suggested to act as a general base in the catalytic mechanism [151]. Analysis of the mutagenesis results targeting the E163Q and E318Q mutations has suggested that one functional active site (formed by His^{25}, Glu^{89}, His^{269} and Glu^{318}) contributes approximately 75% of the total catalytic activity, while the other active site (formed by His^{117}, Glu^{163}, His^{185} and Glu^{242}) only supplies about 25% (k_{cat} of E163Q was 990 ± 50 s^{-1} and k_{cat} of E318Q was 280 ± 20 s^{-1} comparing to the k_{cat} of the wild type enzyme, which has been measured to be 1700 ± 100 s^{-1}) [151]. Besides, each active site appears to prefer different bound metals [151]. The residues that are crucial for substrate binding by interacting with the glycine residue of the GSH moiety are Arg^{302} and Arg^{147}, which resemble Lys^{156} in human GlxI, and Met^{303} and Gln^{148} that correspond to Met^{157} in the human enzyme. Other residues that are predicted to be involved in the interactions with the γ-glutamate residue of GSH

moiety are Arg^{29}, Asn^{93} and Arg^{113} from the N-terminal domain and Arg^{189}, Asn^{246} and Gln^{265} from the C-terminal domain.

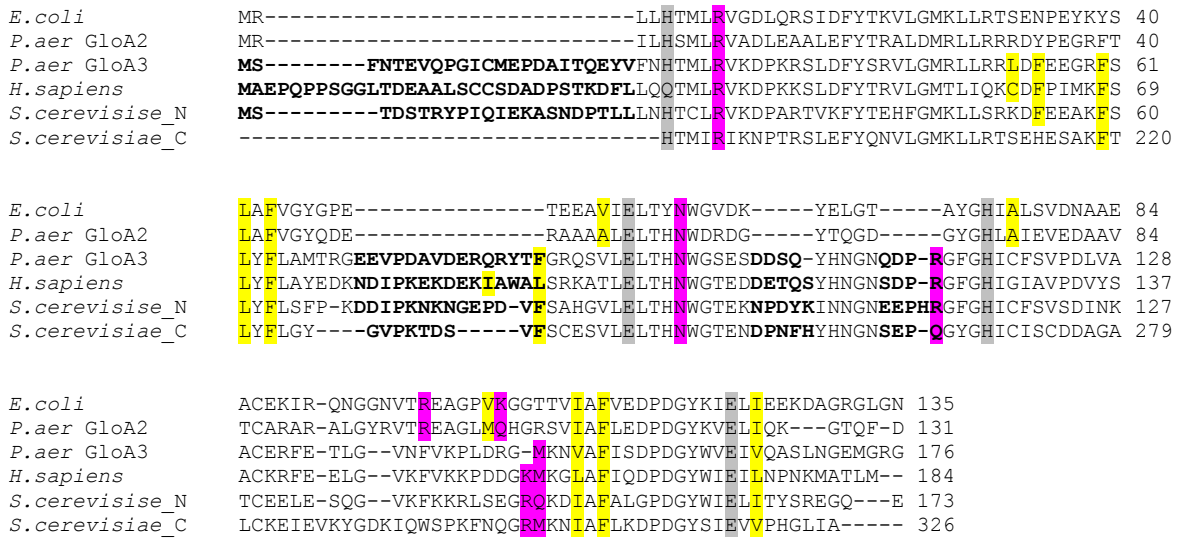


Figure 6.14: The multiple sequence alignment of Glyoxalase I from various organisms (organism name is followed by the National Center for Biotechnology Information (NCBI) accession number) including *E. coli* GlxI (NP_310387), *P. aeruginosa* GloA2 (AAG04099), *P. aeruginosa* GloA3 (AAG08496), *H. sapiens* GlxI (AAB49495) and *S. cerevisiae* GlxI (CAA67622). The amino acid sequence of *S. cerevisiae* GlxI was divided into the N-terminus (residues 1–173) and the C-terminus (residues 184–326). The metal binding residues are marked with a grey highlight and bold letters indicate loop regions that only exist in the Zn^{2+}-activated GlxI. The magenta highlight indicates the residues that are likely involved in substrate binding and the yellow highlight emphasizes the predicted residues that might form the hydrophobic pocket. The alignment was created using CLC Free Workbench (version 3.0.1) with the accurate alignment algorithm (http://www.clcbio.com).

The information obtained from analysis of the X-ray crystallographic structures, enzymatic reactions and ^{15}N-^{1}H HSQC NMR experiments of *E. coli* GlxI suggested a slight asymmetry between the two active sites, in which one active site may possibly possess higher binding affinity for metal and substrate than the other [19, 38, 45]. Yeast GlxI with two asymmetric active sites is a suitable example of that the symmetry between the two active sites is not necessary for the catalytic reaction, since alterations can occur during evolution of the gene such as gene duplication and fusion processes. It was previously found that the apo-form of *E. coli* GlxI contains two asymmetric active sites with one having a higher binding affinity for metal [45]. Addition of metal to the first active site occurs quickly and induces a subtle protein conformational change [38, 45]. The binding of metal to the second unoccupied active site requires additional time, and both active sites of metal-reconstituted enzyme regain more symmetry [38, 45]. However, the flexibility (based on analysis of the B-factor for the proteins)

of the loops that act as lids to cover the active sites containing bound metals is different with one loop being more flexible than the other, indicating that one active site might be more suitable to bind substrate and might have a higher substrate affinity [38, 45]. Once the substrate enters the active site, this interaction likely induces a conformational change in the substrate and in the enzyme, which enhances the enzyme-substrate interaction and maintains the hydrophobic environment within the catalytic site [38, 47]. The conformational change in one active site might induce conformational changes in the other active site, such that this second site may not be suitable to accept the substrate or has developed a lower affinity for substrate. This hypothesis was supported by metal titration experiments, analysis of which have suggested that only one active site is functional, and analysis of mutagenesis experiments in each active site of yeast GlxI, which suggest that one active site is more active than the other [151]. Interestingly, protein symmetry seems to be regained once both active sites are occupied by bound inhibitors [38, 46]. It is possible that the inhibitor binds to one active site first, then eventually the inhibitor can also enter another unoccupied active site (after longer period of time) even though this site possesses a lower binding affinity to the inhibitor. This second bound inhibitor might then induce conformational changes in the enzyme, creating inhibitor-enzyme interactions that result in a more symmetric enzyme structure. The enzyme with bound inhibitors in both active sites was observed because it binds more tightly to the active sites than the substrate, which turns over quickly and so only one active site is functional in the catalytic activity.

In spite of the conclusion that numerous GlxI possess similar interactions with the glycine and the γ-glutamate residues of the GSH moiety, the hydrophobic pockets of these enzymes are likely different. In human GlxI, its hydrophobic pocket is formed by several amino acids such as Cys^{60}, Phe^{62}, Phe^{67}, Leu^{69}, Phe^{71}, Ile^{88}, Leu^{92} from one subunit and Met^{157}, Leu^{160}, Phe^{162}, Leu^{174} from the other subunit (Figure 6.15A). Using this information as a template, a total of 10 hydrophobic residues including Leu^{53}, Phe^{55}, Phe^{60}, Leu^{62}, Phe^{64}, Phe^{85}, Met^{148}, Val^{151}, Phe^{153} and Val^{165} are predicted to compose the hydrophobic pocket of the *P. aeruginosa* GloA3, another Zn^{2+}-activated enzyme. Yeast GlxI forms two hydrophobic pockets within a monomer, each with only 8–9 predicted residues including Phe^{54}, Phe^{59}, Leu^{61}, Phe^{63}, Phe^{82}, Ile^{151}, Phe^{153} and Ile^{165} from the N-terminal domain and Phe^{220}, Leu^{222}, Phe^{224}, Phe^{236}, Met^{303}, Ile^{306}, Phe^{308} and Val^{320} from the C-terminal domain. On the other hand,

the catalytic pocket of *E. coli* GlxI, a Ni^{2+}/Co^{2+}-activated enzyme, consists of 8 hydrophobic residues including Leu^{41}, Phe^{43} and Val^{54} from one subunit and Ala^{76}, Val^{103}, Ile^{110}, Phe^{112} and Ile^{124} from the other subunit (Figure 6.15B). Likewise, the hydrophobic pocket of *P. aeruginosa* GloA2, another Ni^{2+}/Co^{2+}-activating enzyme, is also formed by 8 predicted hydrophobic residues including Leu^{41}, Phe^{43}, Ala^{54}, Ala^{76}, Met^{103}, Ile^{110}, Phe^{112} and Ile^{124}. Analysis of these results appear to indicate that the hydrophobicity of the catalytic pocket of the Zn^{2+}-activated GlxI is greater than that of the Ni^{2+}/Co^{2+}-activated enzymes. Therefore, the binding affinity of inhibitors to the Zn^{2+}-activated GlxI is greater and is the reason for the lower K_i found for these inhibitors with the yeast GlxI and the *P. aeruginosa* GloA3 compared to the affinities to the CLO GlxI and *P. aeruginosa* GloA2 (Table 6.3). It is also possible that other residues that form the hydrophobic pocket cannot be identified by multiple sequence alignments and only the analysis of X-ray crystallographic structure would prove useful in obtaining a better conclusion for the hydrophobicity difference between the two classes of GlxI. Once the X-ray crystallographic structures of *P. aeruginosa* GloA2 and GloA3 are successfully obtained and analyzed, a better appreciation of the nature of the hydrophobic pocket will be obtained.

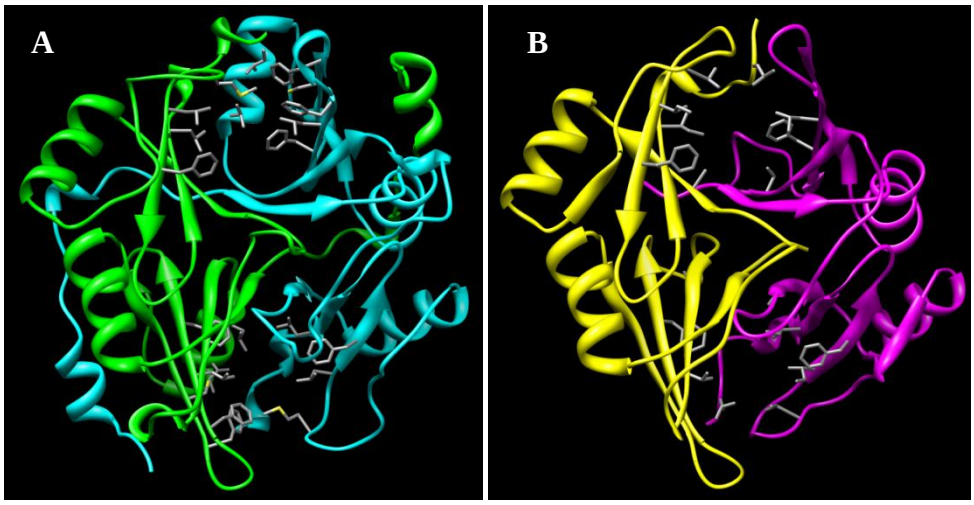


Figure 6.15: The hydrophobic pockets of **(A)** human GlxI consisting of 11 hydrophobic residues in grey (PDB: 1QIN) and **(B)** *E. coli* GlxI consisting of 9 hydrophobic residues in grey (unpublished data). One color represents one subunit for both human (green and cyan) and *E. coli* (yellow and magenta) GlxI. The 3D structures were generated by using the UCSF Chimera program (University of California, San Francisco) [1].

6.4.4: Proposed Reaction Mechanism of Glyoxalase I

Earlier, when the structures of human GlxI with several bound inhibitors were solved, its reaction mechanism was analyzed and proposed. The enzyme with product analogue, NBC-GSH, forms an octahedral coordinated active site with four metal binding ligands (Gln^{33}, Glu^{99}, His^{126} and Glu^{172}) and two water molecules around the Zn^{2+} atom (PDB: 1QIP). However, the structure with bound transition state analogue, HIPC-GSH, shows a trigonal bipyramidal coordination with active metal, inhibitor and only three metal binding ligands including Gln^{33}, Glu^{99} and His^{126} (PDB: 1QIN). The replacement of the metal-bound water molecules by two oxygens from the inhibitor was observed with the distance of metal and metal binding ligands being measured (two oxygens from this inhibitor are 2.04 and 2.131 Å away from the Zn^{2+} atom) (Table 6.4). Besides, no water was observed around the active metal. Another metal binding residue, Glu^{172}, was located distantly from the divalent metal (3.247 Å) in constrast to its position in the enzyme with bound product-like inhibitor (1.991 Å). The Glu^{172} residue is in an ideal position to act as a base for a proton abstraction and transfer as indicated by the distance between $Glu^{172}/O_{\varepsilon 2}$ with the C1 and N atoms of HIPC-GSH inhibitor (3.219 and 3.228 Å, respectively). These distances are smaller than those in the enzyme with bound NBC-GSH (4.474 and 4.836 Å for C1 and O atoms, respectively), suggesting that Glu^{172} only interacts with the intermediate analogue, HIPC-GSH. The removal of Glu^{172} from the active metal and the replacement of metal-bound water molecules by the transition state analogue seem to favor their K_a values and catalytic direction (the K_a value of the carbamoyl functional group of the inhibitor is decreased, while the K_a of Glu^{172} is increased) [48]. The difference in K_a values of the inhibitor and Glu^{172} creates a suitable energy balance for abstraction of a proton from the inhibitor (or the substrate in the normal course of the reaction) [48].

Even though the removal of Glu^{172} from the active metal upon binding of the substrate had been supported by various experiments [46], another theoretical calculation by Himo and Siegbahn, 2001 suggested that displacement of Glu^{172} from the metal may not be necessary [48]. This residue might simply be dislocated because of the presence of the iodophenyl functional group of the HIPC-GSH that does not fit well into the active site, thus possibly creating a steric interaction with the Glu^{172} [46, 48]. The reaction mechanism was proposed to occur as follows (Figure 6.16). The first step of the proton abstraction from C1 of the *S*-hemithioacetal substrate by Glu^{172} (without displacing from the Zn^{2+}-atom) forms an unstable

enolate transition state. The carboxyl group of Glu^{172} then rotates and delivers its proton to the O2 of this transition state, yielding the *cis*-enediol intermediate. Glu^{172} then abstracts a second proton from O1 and delivers this to C2 to form the product, *S*-D-lactoylglutathione. The catalytic reaction mechanism of the *R*-form of the hemithioacetal was proposed to involve another base, Glu^{99}, that is located in the inner coordination sphere of the active metal [48]. The carboxylate of Glu^{99}, which is *trans* to Glu^{172} in the enzyme-metal complex, is in an appropriate position to abstract the proton from the bound *R*-diastereomer of the substrate [46, 48]. It was proposed that the first proton abstraction from C1 and delivery to O2 of the substrate was performed by Glu^{99}. Another base, Glu^{172}, then abstracts the second proton from O1 and delivers this to the C2 position of the intermediate, thus forming a product that diffuses from the active site.

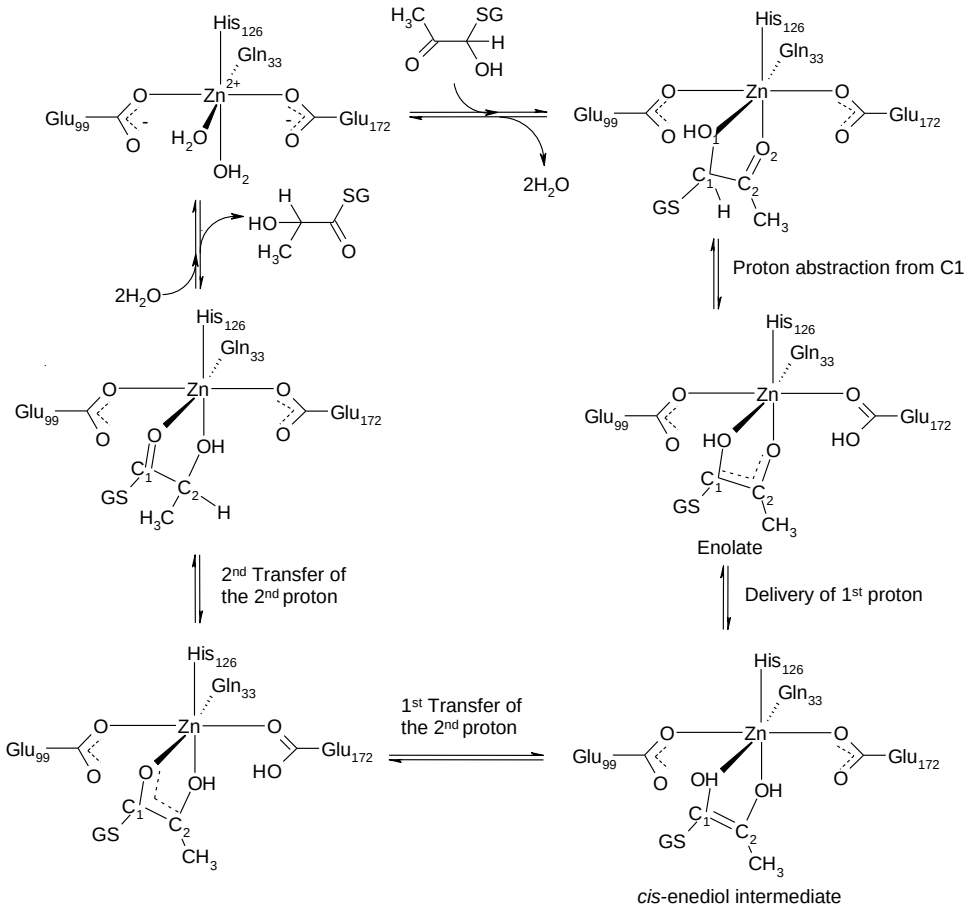


Figure 6.16: The proposed reaction mechanism of human GlxI using Glu^{172} as proton transferred base predicts that there is no disconnect of Glu^{172} from metal [48].

Table 6.4: The distances between active metal and metal binding ligands of human GlxI with bound *S*-*p*-nitrobenzyloxycarbonyl-glutathione (NBC-GSH, product-like inhibitor, PDB: 1QIP) and *S*-(*N*-hydroxy-*N*-*p*-iodophenylcarbamoyl)glutathione (HIPC-GSH, transition state analogue, PDB: 1QIN).

	Distance (Å)	
Type of Interaction	***NBC-GSH***	***HIPC-GSH***
Interactions of metal binding ligands and Zn		
Gln33/$O_{\varepsilon 1}$ – Zn	2.034	2.055
Glu99/$O_{\varepsilon 1}$ – Zn	2.006	1.910
His126/$N_{\varepsilon 2}$ – Zn	2.027	2.049
Glu172/$O_{\varepsilon 1}$ – Zn	1.991	3.247
Interaction of Glu172 and inhibitor		
Glu172/$O_{\varepsilon 2}$ – C1	4.474	3.219
Glu172/$O_{\varepsilon 2}$ – N	4.836	3.228
Interaction of inhibitor and Zn		
Inh/$O_{\zeta 1}$ – Zn	4.8	2.131
Inh/$O_{\phi 2}$ – Zn	-	2.040

The reaction mechanism of *E. coli* GlxI is proposed to be similar to that of the human enzyme. The active site of Ni^{2+}-bound *E. coli* GlxI forms an octahedral geometry with activating metal, four metal binding residues (His^{5}, Glu^{56}, His^{74} and Glu^{122}) and two water molecules (PDB: 1F9Z). Upon binding to the inhibitor, analysis of the X-ray absorption spectroscopy (XAS) data indicated the loss of water molecules and one carboxylate ligand, causing the formation of a trigonal bipyramidal geometry of the active site [47, 243]. The replacement of two metal-bound water molecules by the hydroxamate analogue, TSI, was also suggested by analysis of the X-ray crystallographic data, in which the distances of the two oxygens of the inhibitor and Ni^{2+} atom are 2.016 Å (distance of Ni^{2+} and O-C1) and 2.177 Å (Ni^{2+} and O-C2) (Table 6.5). No water molecule was found within the H-bond distance around the Ni^{2+} atom. The catalytic base, Glu^{122}, that resembles Glu^{172} in human GlxI, locates in a suitable position to abstract a proton from the C1 position (0.292 Å) of the substrate. This residue does not seem to be displaced from the inner coordination of the active metal since the distance between the $Glu^{122}/O_{\varepsilon 2}$ and Ni^{2+} atom is 2.694 Å. This observation appears to support the previous suggestion that Glu^{172} in the human GlxI only disconnects from its active metal due to steric interactions of the phenyl group of the HIPC-GSH inhibitor. The reaction mechanism is proposed such that the catalytic base, Glu^{122}, abstracts the first proton from C1 of the substrate and delivers it to O2 of the hemithioacetal substrate, forming a *cis*-enediol intermediate. Glu^{122} then abstracts the second proton from O1 and donates this to C2 to form

the product, *S*-D-lactoylglutathione. In the R-form of the substrate, Glu^{56} is possibly used as a catalytic base to abstract the proton from C1 of the substrate instead of Glu^{122}.

Table 6.5: The distances between active metal and metal binding ligands of *E. coli* GlxI with bound hydroxamate (TSI) inhibitor (unpublished data)

	Distance (Å)	
Type of Interaction	***No inhibitor***	***TSI***
Interactions of metal binding ligands and Ni		
His5/$N_{\varepsilon 2}$ – Ni	2.032	2.384
Glu56/$O_{\varepsilon 2}$ – Ni	2.163	2.075
His74/$N_{\varepsilon 2}$ – Ni	2.151	2.396
Glu122/$O_{\varepsilon 2}$ – Ni	2.093	2.694
Interaction of Glu122 and inhibitor		
Glu122/$O_{\varepsilon 1}$ – C1	-	3.292
Glu122/$O_{\varepsilon 1}$ – N	-	3.206
Interaction of inhibitor and Ni		
Inh/$O_{\varepsilon 2}$ – Ni	-	2.016
Inh/$O_{\zeta 2}$ – Ni	-	2.177

6.5: Conclusions

In order to investigate the reaction mechanism of GlxI and to explore inhibitor design against this enzyme, two *S*-substituted carbamoyl inhibitors, *S*-(*N*-hydroxy-*N*-methylcarbamoyl)glutathione and *S*-(*N*-hydroxy-*N*-phenylcarbamoyl)glutathione, were successfully synthesized based on previous methodologies [114, 115]. Inhibition studies were performed with both Zn^{2+}-activated and Ni^{2+}/Co^{2+}-activated GlxI using the enzymes from *C. acetobutylicum* (CLO GlxI), *S. cerevisiae* (yeast GlxI) and *P. aeruginosa* (GloA2 and GloA3) as models. It was found that the hydrophobicity of the inhibitor that fits into a catalytic pocket of a particular enzyme directly affects its binding affinity and inhibitory constant (K_i). As well, the hydrophobic environment within the catalytic pocket optimizes the interactions between the active metal and the bound ligands from the glutathione derivatives. Since the catalytic pocket of Zn^{2+}-activated GlxI (*P. aeruginosa* GloA3 and yeast GlxI) is found to be more hydrophobic than that of Ni^{2+}/Co^{2+}-activated enzyme (*P. aeruginosa* GloA2 and CLO GlxI), their K_i values were determined to be lower.

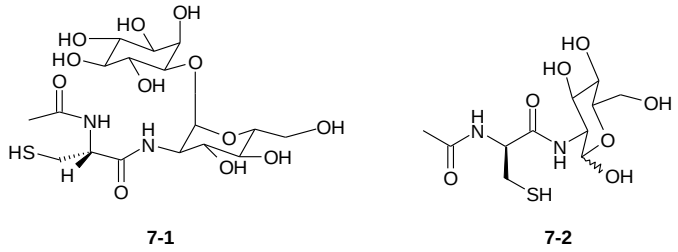
Analysis of the X-ray crystallographic structures of the human and the *E. coli* GlxI suggest that the enzyme alters its loop structure upon binding of the substrate, and thereby

promotes the enzyme-substrate interaction and enzymatic reaction mechanism. The active site of the active enzyme without substrate or inhibitor possesses an octahedral geometry with four metal binding ligands (His^{5}, Glu^{56}, His^{74} and Glu^{122} for *E. coli* GlxI and Gln^{33}, Glu^{99}, His^{126} and Glu^{172} for human GlxI) and two water molecules. Its structure is in an opened conformation, which allows the substrate to enter into its active site. The γ-glutamyl residue of GSH moiety forms interactions with the enzyme, while its cysteine residue swings toward the active site. A conformational change of the substrate allows it to replace two metal-bound water molecules from the metal center. It also allows the loop to slide toward the active site, thus generating a closed conformation of the enzyme. The catalytic base (Glu^{122} in *E. coli* GlxI and Glu^{172} in human GlxI) then abstracts the first proton from C1 of the *S*- hemithioacetal substrate and delivers it to O2, forming a *cis*-enediol intermediate, which is stabilized by the interactions between the glycine residue of the GSH moiety with the enzyme and, of course, the oxygens and the metal center. This same base is then proposed to abstract the second proton from O1 and deliver it to C2, thus generating the product, *S*-D-lactoylglutathione that dissociates from the active metal. This situation causes conformational changes in the enzyme, in which the loops that act as lids to cover the active site swing open to allow the product to diffuse from the catalytic pocket.

CHAPTER 7: ISOLATION OF MYCOTHIOL FROM NATURAL SOURCES AND SYNTHESIS OF DES-*MYO*-INOSITOL MYCOTHIOL

7.1: Introduction

Cysteine is a unique amino acid, consisting of a thiol moiety (–SH) that is involved in many biological applications such as protein and peptide synthesis and intracellular redox maintenance [244, 245]. Cysteine can easily undergo autoxidation according to the scheme, $2RSH + O_2 = RSSR + H_2O_2$, causing two significant problems within the cell. The first is that the cell requires energy to reduce the disulfide product back into its free thiol form. The second is that the co-product of the reaction, the peroxide, is highly toxic and can initiate cell damage by oxidizing functional groups in the cell resulting in interfering with RNA, DNA and protein biosynthesis, which eventually can lead to cell apoptosis [7-9]. Therefore, to avoid these problems, many organisms store or produce Cys in other "thiol" forms such as glutathione (GSH), mycothiol (MSH), and coenzyme A (CoA) that possess slower rates of autoxidation [83, 96, 97, 246]. Most eukaryotes and Gram-negative bacteria produce glutathione, while mycothiol (1-D-*myo*-inosityl-2-deoxy-2-(*N*-acetyl-L-cysteinyl)amino-α-D-glucopyranoside, **7-1**, Figure 7.1) is a major thiol produced by actinomycetes and mycobacteria [83, 96, 97]. MSH was found to be a cofactor in several enzymatic processes that protect against oxygen toxicity [83, 96, 97]. MSH has a 30-fold slower auto-oxidation in the presence of copper ion than that of cysteine and a 7-fold slower rate than that of GSH [80]. Thus, MSH may enable actinomycetes to better cope with oxidative stress.



7-1 **7-2**

Figure 7.1: The structure of mycothiol (MSH, **7-1**) and its truncated analogue, des-*myo*-inositol mycothiol (tMSH, **7-2**).

Since the reported chemical synthesis of mycothiol involves several steps of protecting and deprotecting of selective hydroxyl groups on the sugar and inositol ring, making the

reaction highly difficult and costly [247, 248], it was considered more convenient to isolate MSH from bacterial cells. The isolation of mycothiol from natural sources was previously reported by using reverse phase high performance liquid chromatography (HPLC) or immobilized affinity column and/or tagged reagent approaches [80, 99, 249-251]. The general steps include harvesting of the cell pellet from a bacterial source, disrupting the bacterial cell wall with collection of the cell lysate, and then these steps are followed by adjusting buffer conditions to be suitable and appropriate for further purification steps. The isolation involves thiol oxidation and reverse-phase HPLC based purification of the disulfide product [252]. This method, however, is not efficient since the product is reported to always coelute with other contaminants [252]. By adjusting the buffer pH and precipitating out some contaminants from the cell lysate, a C18 reverse-phase HPLC column (25 mm x 250 mm) utilizing 0.1% trifluoroacetic acid (TFA) in Milli-Q purified water with a flow rate of 10 mL/min is able to isolate the free thiol with a retention time of 8 min and its disulfide (MSSM) having a retention time of 12 min [253]. HPLC on an Asahipak ODP-50 column, which has alkyl chains bonded to porous vinyl alcohol polymer beads, at pH 2.9 was reported as an alternative approach to the isolation of MSH [99]. Spectrophotometric monitoring of the eluate was performed at 220 nm. Another reported column that has been utilized to purify MSH is a thiol-selective column, thiopropyl Sepharose 6B (activated with thiopyridyl) [253]. With this column, the bound MSH can be eluted off the column with DTT in ammonium bicarbonate buffer (pH 8.0) [253]. A commonly used avidin column can also be employed for purification by biotinylating MSH with the thiol-specific reagent, 3-(*N*-maleimidopropionyl)biocytin, before loading onto the avidin column [250]. The biotinylated MSH can be eluted from the column using glycine or with a high concentration of biotin [250].

Our research of interest is to investigate the enzymatic function of the putative mycoGlyoxalase system including mycoGlyoxalase I (mcGlxI) and mycoGlyoxalase II (mcGlxII) from the Gram-positive bacterium, *Streptomyces coelicolor*. This bacterial species was selected because it was previously reported to produce MSH as a major source of thiol [83]. Unlike other mycobacteria, *S. coelicolor* is non-pathogenic and its genome sequence is known, which makes it easier to handle and a suitable choice for DNA and protein investigation. We hypothesized that the reaction mechanism of this system would be similar to that of the well-studied Glyoxalase system, but instead of using GSH as cofactor, the mcGlx

system would employ MSH. The reaction would likely be initiated by the conversion of the hemithioacetal, the product of a non-enzymatic reaction between methylglyoxal (MG) and its cofactor mycothiol, to *S*-D-lactoylmycothiol by mcGlxI. The second enzyme, mcGlxII, then would likely hydrolyze this compound to D-lactate and regenerate MSH.

Unlike GSH, which is commercially available, MSH is not commercially available and it was necessary for us to isolate it from a natural source. *Streptomyces jumonjinesis* was chosen as the natural source of MSH. Even though *S. jumonjinesis* does not produce the highest level of MSH compared to some other actinomycetes (with *Mycobacterium smegmatis* being the highest MSH producer), it produces the highest MSH amounts among the non-pathogenic and relatively easily handled strains. In an attempt to isolate MSH from bacterial cells, the disulfide tag 2-*S*-(2'-thiopyridyl)-6-hydroxynaphthyldisulfide was synthesized [251]. This compound was used to tag MSH in the cell lysate by forming a disulfide bond, thus, generating tagged MSH. The crude product was purified using reverse phase HPLC and the tag was cleaved off using tris(2-carboxyethyl)phosphine (TCEP), a powerful reducing agent that reduces disulfide bonds.

Since air oxidation, undesired adducts and low yield of the thiol were found to contribute to the problem of isolation of MSH from natural sources, the truncated MSH (des-*myo*-inosital mycothiol, **7-2**, Figure 7.1), that can function as MSH for some actinomycetes enzymes, was synthesized. Several attempts to synthesize tMSH have been reported previously [116, 253-260]. It was also found that tMSH was able to act as a substrate in the reaction of MSH disulfide reductase and had a K_m of 463 ± 43 μM and k_{cat} of 70.05 ± 5.88 s^{-1} compared to MSH, which had a K_m of 113 ± 10 μM and k_{cat} of 68.76 ± 2.27 s^{-1} [116]. From these data, we hypothesized (and hoped) that tMSH would likely also be a substrate for the mcGlxI catalyzed reaction, thus solving the problem of MSH limitation and extending further investigation on the reactions involving MSH. The synthesis of truncated mycothiol covered several strategies including repetition of the synthetic method outlined by Patel and Blanchard's Method, a synthetic approach utilizing a cyclic thiolactone intermediate, a resin-bound active ester reaction and finally a peptide-like synthetic strategy (Figure 7.2). Each approach was investigated to optimize the synthetic conditions to increase yield of tMSH, which was further used for the investigation of the putative mcGlxI and mcGlxII in Chapter 8 and 9.

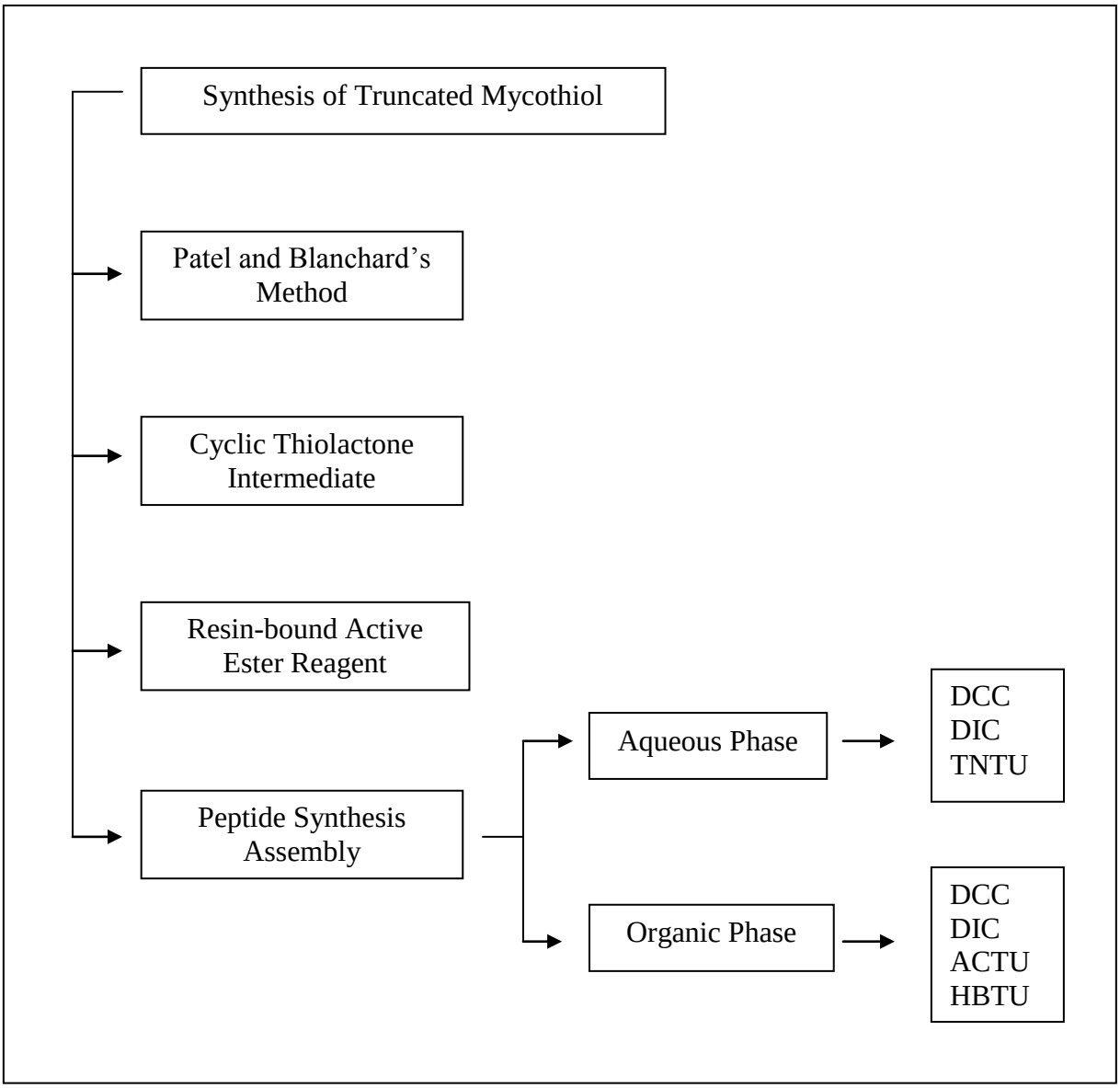


Figure 7.2: The summary of the synthesis of des-*myo*-inositol mycothiol (tMSH, **7-2**) using several strategies including the synthesis outlined in the Patel and Blanchard's Method, the synthesis *via* a cyclic thiolactone intermediate, the resin-bound active ester reaction and the peptide-like synthetic strategy [116, 253, 261].

7.2: Reagents, Materials and Instrumentation

Reagents and Materials

All reagents, materials and equipment used for the experiments described in this chapter are stated below otherwise they are included in the previous chapters. The chemicals, 6-hydroxy-2-naphthyldisulfide, 2-thiopyridyl disulfide, sodium borohydride ($NaBH_4$), disulfide 2,2'-dithiobis(5-nitropyridine) (DTNP), glucosamine hydrochloride (GlcN-HCl), *N*,*N*′-diisopropylcarbodiimide (DIC), *N*,*N*′-dicyclohexylcarbodiimide (DCC), dithiothreitol (DTT), 2-(5-norbornene-2,3-dicarboximido)-1,1,3,3-tetramethyluronium tetrafluoroborate (TNTU), thioanisole, anisole and Ellman's reagent (5, 5'-dithiobis-(2-nitrobenzoic acid) or DTNB) were purchased from Sigma Chemical Company (St. Louis, MO). *S*-Trityl-*N*-acetyl-L-cysteine (Ac-Cys(Trt)-OH) and *N*-(9-fluorenylmethoxycarbonyl)-*S*-acetamidylmethyl-cysteine-pentafluorophenyl ester (Fmoc-Cys(Acm)-OPfP) were received from Bachem Bioscience Inc. (King of Prussia, PA). 1-Hydroxybenzotriazole-6-sulfonamidomethyl polystyrene (PS-HOBt(HL), 0.9 mmol/g) and 2-chloro-1,1,3,3-tetramethyluronium hexachloroantimonate (ACTU) were received from Biotage (Uppsala, Sweden). 1-Hydroxybenzotriazole (HOBt) and *O*-(benzotriazol-1-yl)-*N,N,N′,N′*-tetramethyluronium hexafluorophosphate (HBTU) were obtained from Novabiochem (EMD Chemicals, Darmstadt, Germany). Imidazole and tris(2-carboxyethyl)phosphine (TCEP) were purchased from BioShop Canada Inc. (Burlington, ON). Iodine was received from BDH Chemicals Ltd (Carle Place, NY).

Salts were obtained from EMD Pharmaceuticals (Durham, NC) including sodium bicarbonate ($NaHCO_3$) and potassium phosphate, dibasic (K_2CO_3). Potassium carbonate (K_2CO_3) was received from BDH Chemicals Ltd (Carle Place, NY).

Acid-base reagents acquired from Fisher Scientific (Nepean, ON) were as follows: perchloric acid ($HClO_4$), glacial acetic acid (CH_3CO_2H) and ethylenediamine tetraacetic acid (EDTA). Trifluoroacetic acid (TFA), piperidine, pyridine and sodium hydroxide (NaOH) were received from Caledon Laboratories Ltd. (Georgetown, ON), while diisopropylethylamine (DIEA) was purchased from Sigma Chemical Company (St. Louis, MO).

All solvents such as acetonitrile (ACN), methanol (MeOH), ethanol (EtOH) dimethylformamide (DMF), chloroform ($CHCl_3$) and ethyl acetate (EtOAc) were HPLC grade from Caledon Laboratories Ltd. (Georgetown, ON). The Aldrich Sure-Seal anhydrous solvents including anhydrous ether, anhydrous dioxane and anhydrous DMF were purchased from

Sigma Chemical Company (St. Louis, MO). Water used in all experiments was Milli-Q water (18 MΩ-cm conductivity).

Equipment

Streptomyces jumonjinesis cells were grown by using an incubator shaker (series 25) from New Brunswick Scientific Co., Inc. (Edison, NJ) and were harvested using a Avanti® J-E Centrifuge from Beckman Coulter, Inc. (Palo Alto, CA). The cell wall disruption of crude cell extract was performed using a Sonicator™ cell disruptor (model W225) with converter model #2 and a standard tapered microtip with the output control set at 5 (Heat Systems-Ultrasonics, Inc., Plainview, NY).

The purification of organic synthetic products utilizing high performance liquid chromatography (HPLC) was performed with a Waters 625 LC system and a 994 programmable photodiode array detector using a Waters μBondapak C18 reverse phase radial compression column (25 × 100 nm) from Waters Corporation (Chicago, IL). The isolation using a RediSep C18 reverse phase chromatography column (43 g) (Teledyne Isco Inc., Lincoln, NE) was performed on a BioFlash system from Biotage (Uppsala, Sweden). Aluminum-backed silica gel plates for thin layer chromatography (TLC) were purchased from EMD Pharmaceuticals (Durham, NC). The visualization for TLC plates was performed using UV light and/or ninhydrin staining solution that was prepared by dissolving 0.2% (w/v) ninhydrin (Peirce Chemical Company, Rockford, IL) in ethanol.

The identification of organic material was performed using an electrospray ionization mass spectrometer (ESI-MS), a Micromass Q-TOF Ultima™ Global instrument, supplied by the Waterloo Chemical Analysis Facility, University of Waterloo (Waterloo, ON). The nuclear magnetic resonance (NMR) data were obtained using Bruker 300 MHz or 500 MHz spectrometers with deuterated solvents manufactured by Cambridge Isotopes Laboratories (Andover, MA).

7.3: Experimental Protocols

7.3.1: Isolation of Mycothiol from *Streptomyces jumonjinesis*

The protocols were adapted from Steenkamp and Vogt, 2004 with some changes as follows [251]. The reduction of 6-hydroxy-2-naphthyldisulfide (**7-3**) was adjusted to pH 9.0, which is the maximum activity for $NaBH_4$. All buffers in this experiment were degassed before use and the reaction was performed under argon to eliminate reformation of homodisulfide. The reaction between 2-thio-6-hydroxynaphthalene (**7-4**) and 2-thiopyridyl disulfide (**7-5**) was stirred for 1.5 hours instead of 20 min and precipitate was removed by filtration instead of centrifugation.

Synthesis of 2-S-(2'-thiopyridyl)-6-hydroxynaphthyldisulfide (7-5)

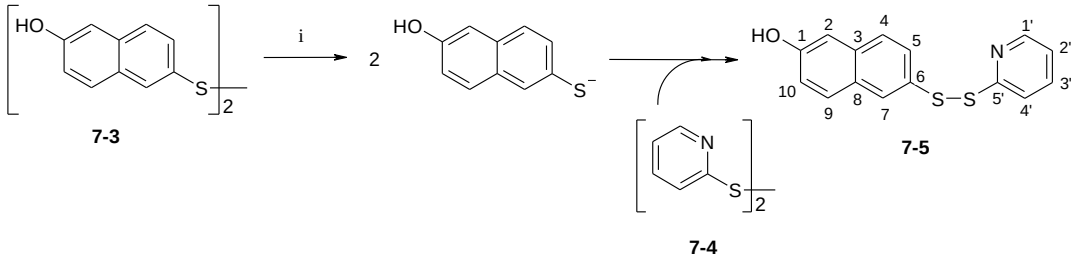


Figure 7.3: The synthesis of 2-*S*-(2'-thiopyridyl)-6-hydroxynaphthyldisulfide. (i) $NaBH_4$.

The starting material, 6-hydroxy-2-naphthyldisulfide (**7-3**, 0.198 g, 0.56 mmol), was dissolved in a degassed solution of (1:1) acetonitrile and 0.1 M $NaHCO_3$ solution (pH 9.0) (100 mL). The reducing agent, $NaBH_4$, (0.6 g, 20 mmol, 28 equiv.) was added to the solution, which was then flushed three times with argon and left stirring at room temperature for 2 hours in a sealed vessel. The reaction was quenched by adding glacial acetic acid (2 mL), and the pH of buffer was adjusted to 4.6 by either dipotassium phosphate or acetic acid to initiate the formation of the naphthalene thiol. The reaction was then permitted to continue to degas for 5 min.

In the mean time, 2-thiopyridyl disulfide (**7-4**, 0.474 g, 2.06 mmol) was dissolved in a degassed solution of a 1:1 mixture of acetonitrile and water (50 mL). The solution of **7-3** was then added dropwise to the solution of **7-4** over 20 min (~6 mL/min). The mixture was allowed to stir at room temperature for 1.5 hours then diluted with 4 °C water (150 mL) to precipitate the crude mixture of product, 2-*S*-(2'-thiopyridyl)-6-hydroxynaphthyldisulfide (**7-5**).

This precipitate was collected by filtration and washed with MeOH several times. The remaining solvent was removed by rotary evaporation. The dry precipitate was then redissolved in 50% (v/v) acetonitrile/water. The product was purified by HPLC using a Waters μBandapak C18 reverse phase radial compression column (25 × 100 mm) with a solvent system as shown in Table 7.1 (solvent A was 1% TFA in water and solvent B was acetonitrile). The purified product (**7-5**) was eluted with a retention time (t_R) of 42 min and a 20% recovery yield was obtained.

Table 7.1: Solvent gradient used to purify 2-*S*-(2'-thiopyridyl)-6-hydroxynaphthyldisulfide. A = 1% TFA in water and B = acetonitrile.

Time (min)	***Flow rate (mL/min)***	***%A***	***%B***
0	4	100	0
15	4	100	0
20	4	50	50
60	4	0	100
75	4	0	100

Since the purification using the reverse phase HPLC required multiple sample loadings, which led to sample loss during filtrations and injections, an alternative method for isolation of **7-5** was used and it was based on silica gel chromatography. After the dry precipitate was recovered from methanol evaporation, it was dissolved in a minimal amount of 95:5 chloroform:MeOH. The sample was then filtered through a 0.2 μm PTFE filter. The filtrate was loaded onto the silica gel column using a 95:5 chloroform:MeOH solvent system. The purified product was collected by gravity (flow rate ~3.5 mL/min), and the product **7-5** was monitored by silica gel TLC using chloroform:MeOH (95:5) mixture as a running solvent (**7-5**, R_f = 0.32; **7-4**, R_f = 0.62; **7-3**, R_f = 0.18). The solvent was removed by rotary evaporation and the remaining dry yellow residue gave approximately 31% yield of purified **7-5**, identical in spectroscopic properties as those reported previously [251].

^{1}H NMR (300 MHz, MeOH-d_4): δ 8.4 (d, J = 5.2 Hz, 1H, H1'), 7.9 (d, J = 1.4 Hz, 1H, H9), 7.8 (m, 1H, H3'), 7.7 (m, 1H, H4'), 7.6 (m, 2H, H4, H7), 7.5 (dd, J = 2.0, 2.0 Hz, 1H, H2), 7.2 (m, 1H, H2′), 7.0 (m, 2H, H5, H10).

^{13}C NMR (300 MHz, MeOH-d_4): δ 156.1 (C5'), 156.1 (C1), 148.9 (C1'), 137.8 (C3), 134.5 (C7), 129.1 (C9), 128.8 (C8), 128.3 (C3'), 127.6 (C6), 127.4 (C5), 126.1 (C4), 121.2 (C2'), 120.0 (C4'), 118.9 (C10), 108.6 (C2)

ESI-MS (pos): m/z $(MH)^+$ = 286, calculated (M) = 286.03

Growth of Streptomyces jumonjinesis Cells

S. jumonjinesis cells were received from the American Tissue Culture Collection (ATCC, number 298640). The stock cells were grown in 5 mL International Streptomyces Project (ISP) medium (5 g tryptone and 3 g yeast extract in 1 L distilled water and adjusted pH to 7.0–7.2) in a 220 rpm shaker at 26 °C for 2 days. The culture was then transferred to ISP medium (1 L) and continued shaking at the same speed and temperature for 5 days. The cells were harvested by centrifugation at 29,030×*g* for 20 min and washed with a minimum volume of buffer containing 50 mM Tris (pH 8.0) and 150 mM KCl. The cells were then centrifuged at the same speed and time before being flash frozen in liquid nitrogen and stored at -80 °C.

Purification of Mycothiol from Streptomyces jumonjinesis

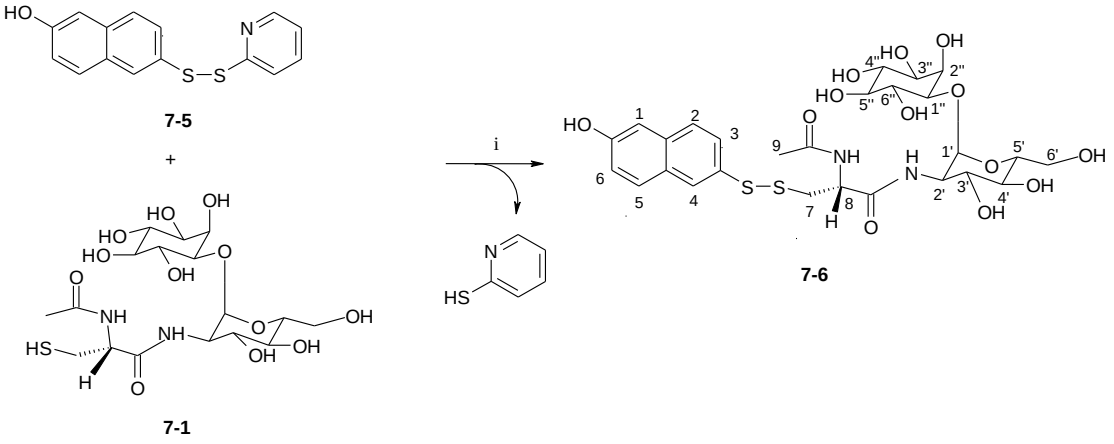


Figure 7.4: The isolation of mycothiol using a tag of 2-*S*-(2'-thiopyridyl)-6-hydroxynaphthyldisulfide. (i) K_2CO_3.

S. jumonjinesis cells (7.1 g) were thawed in 30 mL lysis buffer (0.25 M perchloric acid, 2 mM EDTA and 40% (v/v) acetonitrile/water). The cell wall was disrupted by sonication for 30 × 10 sec with cooling. The cell lysate was then centrifuged at 48,384×*g* for 15 min. The

supernatant was collected, and the pH of the culture was adjusted to ~4.8 with an addition of K_2CO_3, followed by degassing for 10 min. The solution was added dropwise over 20 min to 3.65 mL of degassed 4.4 mM (saturated) solution of **7-5** in 60% (v/v) acetonitrile/water (0.016 mmol). The reaction was stirred under argon at room temperature for 4.5 hours before diluting with 168 mL water (5 equiv.). The mixture was filtered through a Waters classic Sep-pak (C18) that was previously conditioned with methanol and water, respectively. The column was then rinsed with water and the crude product was eluted with a minimum amount of acetonitrile:water (1:1). The solvent was removed by rotary evaporation. The remaining dry product was redissolved in 50% (v/v) acetonitrile in water containing 0.1% TFA. The crude product was purified by utilizing a RediSep C18 column on a BioFlash system (43 g, 20–100% acetonitrile in water containing 0.1% TFA, with 20% increments of acetonitrile). The fractions containing purified product were monitored at 220 and 280 nm. Further purification was performed by HPLC using a Waters μBandapak C18 reverse phase radial compression column (25 × 100 mm) with a solvent system as shown in Table 7.2 (solvent A was 0.1% TFA in water and solvent B was acetonitrile). The purified product, *S*-2-(mycothiolyl)-6-hydroxynaphthyldisulfide (**7-6)**, was eluted at t_R 38 min with 3.1 mg recovery yield per 7.1 g cell pellet. To improve the recovery yield of the product, every 10% increment of acetonitrile from 20–100% acetonitrile in water containing 0.1% TFA solvent system of the RediSep C18 chromatography was performed. The pure product of **7-6** was eluted with 30% acetonitrile with a recovery yield of 10.8 mg per 6.3 g cell pellet. The isolated material had spectroscopic properties identical to those previously reported [251].

Table 7.2: Solvent gradient used to purify *S*-2-(mycothiolyl)-6-hydroxynaphthyldisulfide. A = 1% TFA in water and B = acetonitrile.

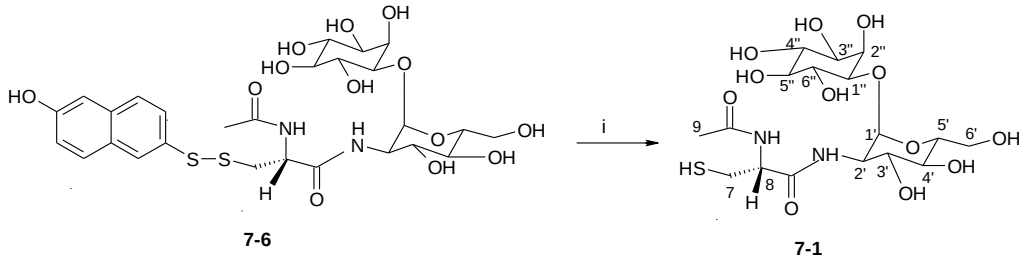
Time (min)	***Flow rate (mL/min)***	***%A***	***%B***
0	4	100	0
15	4	100	0
55	4	0	100
70	4	0	100
75	4	100	0
90	4	100	0

^{1}H NMR (500 MHz, D_2O): δ 7.9 (s, 1H, H4), 7.6–7.7 (m, 2H, H2, H5), 7.5 (d, J = 8.8 Hz, 1H, H1), 7.2 (m, 2H, H3, H6), 5.0 (s, 1H, H1'), 4.5 (m, 1H, H8), 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, J = 9.1 Hz, 1H, H7a), 2.9 (m, 1H, H7b), 2.1 (s, 3H, H9)

ESI-MS (neg): m/z 659.1 $(tagged\ MSH)^-$; calculated 660.7 (tagged MSH)

Cleavage of Tagged Mycothiol



7-6 **7-1**

Figure 7.5: The cleavage of tagged mycothiol. (i) TCEP

The approximated yield of MSH (**7-1**) before the cleavage of the 6-hydroxy-2-naphthylsulphydryl tag group was determined by mixing 10 μL of MSH in water, 143 μL of 7 mM DTT in 25 mM K_2HPO_4 (pH 8.0) and 847 μL water. The absorbance at 298 nm was measured and the concentration of MSH was calculated using the Beer-Lambert Law ($A= \varepsilon_{298}bc$, where A is absorbance at 298 nm, ε is extinction coefficient at 298 nm (18.5 $mM^{-1}cm^{-1}$ [251]), b is path length (1 cm) and c is concentration of MSH in molar unit.

For the cleavage of the 6-hydroxy-2-naphthylsulphydryl group from MSH, 100 μL of 50 mM TCEP in 50 mM Tris buffer (pH 8.0) was mixed with 900 μL MSH in water (required at least 1:1 ratio of TCEP:tagged MSH). The mixture was purified by reverse phase HPLC using a Waters μBandapak C18 reverse phase radial compression column (25 × 100 mm) with a solvent system as shown in Table 7.3 (solvent A was 0.1% TFA in water and solvent B was acetonitrile). However, the TCEP adduct was observed after this purification as indicated by ^{1}H NMR and ESI-MS. The TCEP resin is an alternative choice for future purification.

^{1}H NMR (500 MHz, D_2O): δ 5.1 (s, 1H, H1'), 4.5 (m, 1H, H8), 4.1 (s, 1H, H2″), 3.9–3.7 (m, 7H, H2', H3', H5', H6' (2H), H4", H6"), 3.6–3.4 (m, 4H, H4', H1", H3", H5"), 3.2 (t, J = 9.2 Hz, 1H, H7a), 2.94 (m, 1H, H7b), 2.1 (s, 3H, H9) with contamination of TCEP adduct.

ESI-MS (neg): m/z 249.0 (100%) $(TCEP)^-$, 265.0 (50%) $(TCEP\text{-}O)^-$, 485.1 (20%) $(MSH)^-$, 499.1 (30%) $(2TCEP)^-$, 515.1 (25%) $(2TCEP\text{-}O)^-$; calculated 486.5 (MSH), 250.2 (TCEP)

Table 7.3: Solvent gradient used to purify mycothiol. A = 1% TFA in water and B = acetonitrile.

Time (min)	*Flow rate (mL/min)*	*%A*	*%B*
0	4	100	0
10	4	100	0
40	4	75	25

Only trace amounts of MSH as well as the TCEP oxide and the naphthylthiol were recovered. It was reported previously that *S. jumonjinesis* produces 6.7 μmol/g residual dry weight (or 3.2 mg per g cell) [83]. Our experiment used 6–7 g of *S. jumonjinesis* cell, which should produce maximal 19–23 mg MSH. However, the product loss during purification steps limited the yield of purified MSH. Other bacterial strains such as *Mycobacterium smegmatis* that can produce MSH as high as 19 μmol/g residual dry weight have been used as alternative sources for MSH isolation [83]. Nevertheless, this bacterial strain is a mild pathogen, so we decided to not pursue any experiments utilizing this bacterium but to focus on the synthetic aspects initially to determine if there was indeed a mycoGlyoxalase.

7.3.2: Synthesis of Des-*myo*-Inositol Mycothiol

Truncated Mycothiol Synthesis by Patel and Blanchard's Method

The addition of L-glucosamine onto N-(9-fluorenylmethoxycarbonyl)-S-acetamidylmethyl-cysteine-pentafluorophenyl ester

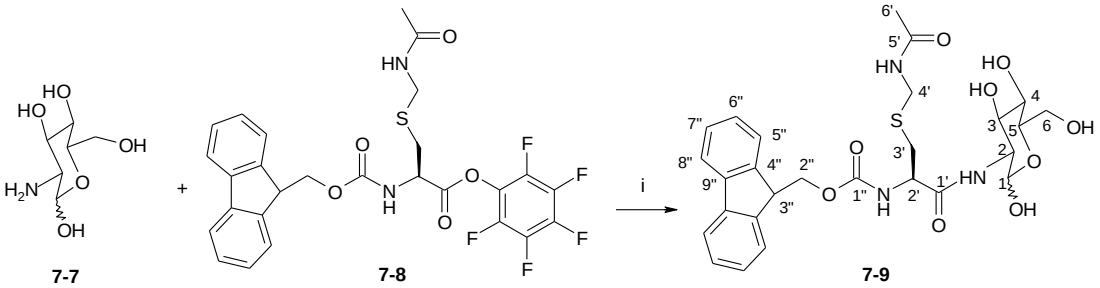


Figure 7.6: The synthesis of *N*-α-Fmoc-*S*-acetamidomethyl-L-cysteinyl-2-amino-2-deoxy-*α*-D-glucopyranoside (Fmoc-Cys(Acm)-GlcN). (i) HOBt, DMF.

This protocol is developed from the synthesis of des-*myo*-inositol mycothiol by Patel and Blanchard, 1998 [116]. The glucosamine-HCl (3.72 g, 17.22 mmol) was neutralized in 1 M NaOH (17 mL, 17 mmol) to obtain the free base of α-D-glucosamine (**7-7**). The solution was allowed to stir at room temperature for 5 minutes before being flash frozen in liquid nitrogen and dried on a lyophilizer overnight. The starting material, Fmoc-Cys(Acm)-OPfp (**7-8**, 2 g, 3.44 mmol) was dissolved in Aldrich Sure-Seal anhydrous DMF (10 mL) before adding HOBt (0.93 g, 6.89 mmol, 1.8 equiv.). The mixture was then transferred to the glucosamine (**7-7**, 17 mmol, 5 equiv.) immediately and stirred vigorously. The reaction was allowed to stir at room temperature for 3 hours under argon. The solvent was removed by evaporation using a rotary evaporator attached to a vacuum pump. The remaining material was precipitated with water (100 mL). The product was isolated by filtration and thoroughly rinsed with water to remove excess **7-7**, HOBt, NaCl, pentafluorophenol (side product) and residual DMF. The product was dissolved in methanol before being dried *in vacuo* overnight to remove any residual water. The dry product, Fmoc-Cys(Acm)-GlcN (**7-9**) gave 73% yield and had spectroscopic characteristics identical to those previously reported [116].

^{1}H NMR (300 MHz, MeOH-d_4): δ 7.87 (d, J = 8.0 Hz, 2H, H8″), 7.75 (d, J = 6.4 Hz, 2H, H5″), 7.46–7.37 (t, J = 7.0 Hz, 2H, H7″), 7.36–7.27 (t, J = 7.4 Hz, 2H, H6″), 4.91 (d, J = 2.3 Hz, 1H, H1), 4.47–4.22 (m, 6H, H2′, H4′, H2″, H3″), 3.79–3.61 (m, 6H, H2, H3, H4, H5, H6), 3.01–2.56 (m, 2H, H3′), 1.87 (s, 3H, H6′)

^{13}C NMR (300 MHz, DMSO-d_6): δ 173.4 (C1´), 172.9 (C5´), 158.7 (C1″), 146.3 (C4″), 143.3 (C9″), 130.3 (C7″), 129.7 (C8″), 127.8 (C6″), 122.6 (C5″), 93.2 (C1), 74.5 (C2), 73.3 (C3), 73.0 (C6), 68.4 (C2″), 63.4 (C5), 56.8 (C4), 56.6 (C2′), 49.0 (C3″), 35.6 (C3′), 25.0 (C6′), C4′ is expected to appear at ~δ 40.8, which would be hidden by the DMSO-d_6 solvent peak.

ESI-MS (pos): m/z = 576.2 $(MH)^+$, calculated (M) = 575.2

*The removal of the 9-fluorenylmethoxycarbonyl group producing S-acetamidomethyl-L-cysteinyl-2-amino-2-deoxy-α,β-D-glucopyranoside (**7-10**)*

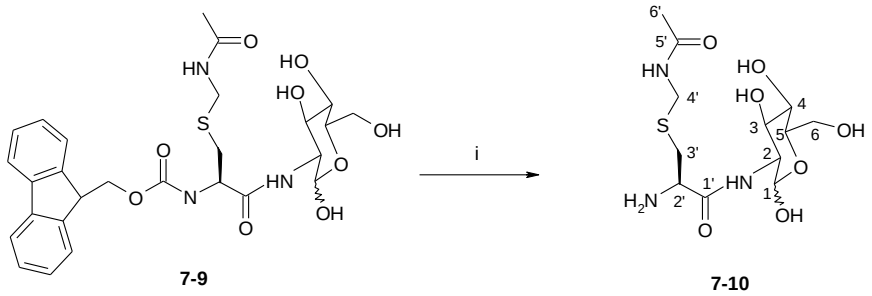


Figure 7.7: The synthesis of *S*-acetamidomethyl-L-cysteinyl-2-amino-2-deoxy-*α,β*-D-glucopyranoside (NH_2-Cys(Acm)-GlcN). (i) 5% piperidine in DMF.

The dry Fmoc-Cys(Acm)-GlcN (**7-9**, 1.45 g, 2.53 mmol) was dissolved in 5 mL of DMF solution containing 5% piperidine and allowed to stir for 30 minutes. The solvent was removed by using a rotary evaporator attached to a vacuum pump. The crude product of NH_2-Cys(Acm)-GlcN (**7-10**) was dissolved in water (100 mL). The precipitate containing cleaved dibenzofulvene (side product) was removed by filtration (filter paper #42), and the filtrate was dried *in vacuo* overnight. The recovered product of **7-10** gave 99% yield. The spectroscopic characteristics (^{1}H NMR and ESI-MS) of this compound were in full agreement with those previously reported [116].

^{1}H NMR (300 MHz, D_2O): δ 5.40 (d, J = 4.6 Hz, 1H, H1 for α-amomer), 5.16 (d, J = 3.3 Hz, 1H, H1 for β-anomer), 4.39–4.19 (m, 3H, H4′/ H2′), 3.91–3.27 (m, 6H, H2, H3, H4, H5, H6), 3.05–2.63 (m, 2H, H3′), 1.87 (s, 3H, H6′)

ESI-MS (pos): m/z = 354.1 $(MH)^+$, calculated (M) = 353.1

*Acetylation of the amino group of S-acetamidomethyl-L-cysteinyl-2-amino-2-deoxy-α,β-D-glucopyranoside to produce compound **7-11***

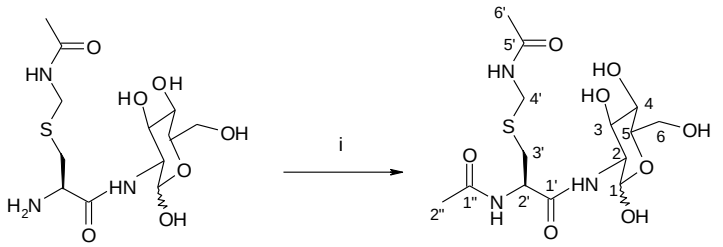


Figure 7.8: The synthesis of *N*-acetyl-*S*-acetamidomethyl-L-cysteinyl-2-amino-2-deoxy-α,β-D-glucopyranoside (Ac-Cys(Acm)-GlcN). (i) Ac_2O, K_2CO_3.

The NH_2-Cys(Acm)-GlcN (**7-10**, 0.884 g, 2.5 mmol) produced from the previous reaction was dissolved in water (6.3 mL) before adding potassium carbonate (0.69 g, 5.01 mmol, 2 equiv.) with stirring until all K_2CO_3 was dissolved. To the solution, acetic anhydride (260 μL, 0.281 g, 1.1 equiv.) was added dropwise over 30 min. The mixture was allowed to stir at room temperature for 45 minutes before being dried *in vacuo*. The dry sample was redissolved in ethanol (30 mL) and precipitated residual K_2CO_3 was removed by filtration. The filtrate was dried to become an oil-like crude product, compound **7-11**, which was further purified by HPLC using a Waters μBandapak C18 reverse phase radial compression column (25 × 100 mm) and a solvent system as shown in Table 7.4 (solvent A was water and solvent B was methanol). The isolated product was dried on the lyophilizer overnight. The white precipitate of pure **7-11** was eluted at t_R 11–12 min with ~20% yield. The spectroscopic characteristics (^{1}H NMR and ESI-MS) of this compound were in full agreement with those previously reported [116].

Table 7.4: Solvent gradient used to purify *N*-acetyl-*S*-acetamidomethyl-L-cysteinyl-2-amino-2-deoxy-α,β-D-glucopyranoside (solvent A = water and solvent B = methanol).

Time (min)	*Flow rate (mL/min)*	*%A*	*%B*
0	5	100	0
17	5	100	0
22	5	0	100
42	5	0	100
47	5	100	0
57	5	100	0

^{1}H NMR (300 MHz, D_2O): δ 5.37 (d, J = 4.3 Hz, 1H, H1 for α-amomer), 5.10 (d, J = 3.0 Hz, 1H, H1 for β-anomer), 4.37–4.09 (m, 3H, H4′/ H2′), 3.93–3.23 (m, 6H, H2, H3, H4, H5, H6), 3.09–2.70 (m, 2H, H3′), 1.94 (s, 3H, H6′), 1.82 (s, 3H, H2″)

ESI-MS (pos): m/z = 434.1 $(MK)^+$, calculated (M) = 395.4

*The removal of acetamidomethyl group by thallium to produce disulfide **7-12***

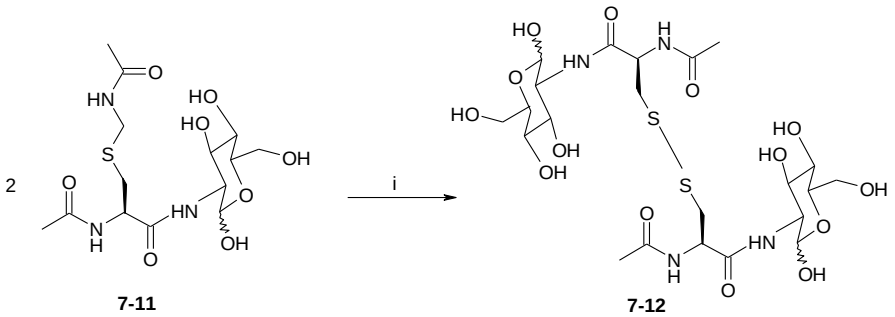


Figure 7.9: The synthesis of *N*-acetyl-L-cysteinyl-2-amino-2-deoxy-*α,β*-D-glucopyranoside disulfide (GlcN-AcCys-AcCys-GlcN). (i) $Tl(CF_3CO_2)_3$, anisole and TFA.

Ac-Cys(Acm)-GlcN (**7-11**, 12.9 mg, 32.66 μmol) was dissolved in 7 μL of anisole solution (7.06 mg, 65.32 μmol) in trifluoroacetic acid at 4 °C under nitrogen. Thallium(III) trifluoroacetate (22.8 mg, 41.56 μmol) was added to the solution and the mixture was allowed to stir for 1 hour under a blanket of nitrogen to avoid air oxidation. The solvent was dried *in vacuo*. The crude product was precipitated with anhydrous ether (50 mL) and washed three times with fresh anhydrous ether. The sample was dried *in vacuo* and further purified by HPLC using a Waters μBandapak C18 reverse phase radial compression column (25 × 100 mm) and a solvent system as in Table 7.5 (solvent A was water and solvent B was methanol). The ^{1}H NMR and ESI-MS experiments suggested no formation of the product GlcN-AcCys-AcCy-GlcN (**7-12**).

Since previous experiments on the removal of the acetamidomethyl group were unsuccessful, an alternative cleavage reagent, iodine, was used. This protocol is developed from the work of Kamber and coworkers [262]. Ac-Cys(Acm)-GlcN (**7-11**, 29 mg, 0.073 mmol) was dissolved in DMF (5.6 mL). To the solution, 0.1 N solution of iodine (4.4 mL, 0.438 mmol, 6 equiv.) was added, and the reaction was allowed to stir at room temperature for

1 min. The crude product in the aqueous layer was isolated from the mixture of water (10 mL) and chloroform (10 mL). The reaction was washed several times with chloroform until the chloroform layer became colorless. The aqueous layer was dried *in vacuo*. The crude product was further purified by HPLC using a Waters μBandapak C18 reverse phase radial compression column (25 × 100 mm) and a solvent system as in Table 7.5 (solvent A was water and solvent B was methanol). The ^{1}H NMR and ESI-MS experiments suggested no formation of compound **7-12**.

Table 7.5: Solvent gradient used to purify *N*-acetyl-L-cysteinyl-2-amino-2-deoxy-α,β-D-glucopyranoside disulfide (solvent A = water and solvent B = methanol).

Time (min)	*Flow rate (mL/min)*	*%A*	*%B*
0	5	100	0
14	5	100	0
17	5	0	100
30	5	0	100
35	5	100	0

The removal of acetamidomethyl group by disulfide 2,2'-dithiobis(5-nitropyridine)

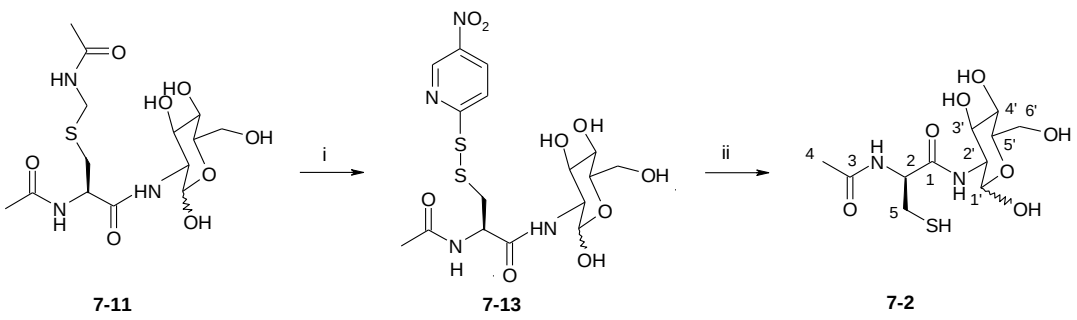


Figure 7.10: The synthesis of des-*myo*-inositol mycothiol *via* the formation of Ac-Cys(5-Npys)-GlcN intermediate. (i) DTNP, thioanisole and TFA. (ii) DTT and TFA

Thallium (III) salts and iodine can generate harsh reaction conditions. Disulfide 2,2'-dithiobis(5-nitropyridine) (DTNP), on the other hand, has been reported to be a milder reagent that has been used for the deprotection of cysteine and selenocysteine side-chain protecting groups [263]. Ac-Cys(Acm)-GlcN (**7-11**, 74 mg, 0.187 mmol) was dissolved in cocktails containing DTNP (0.87 g, 2.805 mmol, 15 equiv.) in 2% thioanisole in TFA (0.36 mL in 17.78 mL TFA, 16 equiv.) and allowed to stir at room temperature for 2 hours. TFA was then removed by rotary evaporation before adding a mixture of 1:2 water:diethyl ether. The

aqueous layer was separated and the crude product was purified by HPLC using a Waters μBandapak C18 reverse phase radial compression column (25 × 100 mm) and a solvent system as in Table 7.6 (solvent A was 0.1% TFA in water and solvent B was acetonitrile). ^{1}H NMR and ESI-MS results neither showed the formation of AcCys(5-Npys)-GlcN (**7-13**) nor tMSH (**7-2**).

The second trial of this experiment was performed using 2 equivalents of DTNP with excess thioanisole. Ac-Cys(Acm)-GlcN (**7-11**, 5 mg, 10.46 μmol) solution was prepared by dissolution of compound **7-11** in 2.94 mL TFA, and DTNP (11.7 mg, 35.4 μmol) solution was prepared by dissolution in 10 mL TFA. Ac-Cys(Acm)-GlcN solution (400 μL) was mixed with the DTNP solution (400 μL) and TFA (2 mL) before adding a mixture of thioanisole (80 μL) and DTNP solution (400 μL). The reaction was allowed to stir at room temperature for 2 hours. TFA was removed by rotary evaporation. A mixture of water:diethyl ether (1:2) was added to the remaining sample. The reaction was mixed by vortex. The aqueous layer was separated and the crude product was purified as stated above. Trace amounts of the product **7-2** were found in the HPLC eluted peak with t_R 9 min, while compound **7-13** were found in the eluted peak with t_R 12 min. ^{1}H NMR confirmed the existence of **7-2** as it was identical to spectroscopic properties of previously synthesized **7-2** [116]. However, due to the very low amounts of compound **7-2** recovered from the purification, no ESI-MS experiment was performed on this sample. Thus, larger scale purification to recover higher yield of the product and the comfirmation on the molecular mass of compound **7-2** would be necessary before pursuing any further experiments.

Table 7.6: Solvent gradient used to purify Cys(5-Npys) intermediate (solvent A = 0.1% TFA in water and solvent B = acetonitrile).

Time (min)	*Flow rate (mL/min)*	*%A*	*%B*
0	5	100	0
15	5	100	0
75	5	40	60
80	5	0	100
90	5	0	100
95	5	100	0
105	5	100	0

^{1}H NMR (300 MHz, D_2O): δ 5.21–5.12 (d, 1H, H1'), 4.35–4.05 (m, 1H, H2), 3.93–3.35 (m, 6H, H2', H3', H4', H5', H6'), 2.62–2.42 (m, 2H, H5), 1.91 (s, 3H, H4)

Truncated Mycothiol Synthesis via Cyclic Thiolactone Intermediate

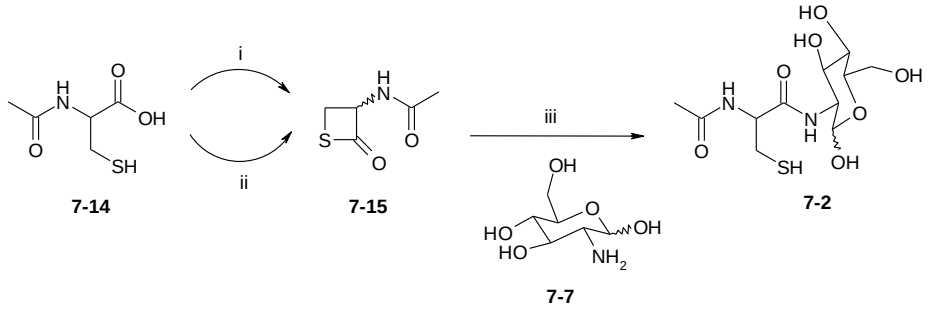


Figure 7.11: The synthesis of des-*myo*-inositol mycothiol *via* cyclic thiolactone intermediate. (i) Ac_2O and pyridine. (ii) DCC in dioxane. (iii) NaOH, water and chloroform.

The synthesis of truncated mycothiol *via* cyclic thiolactone intermediate was adapted from the work of Ramirez and coworkers [261]. *N*-Acetyl-(*L,D*)-cysteine (**7-14**, 6.50 g, 3.75 mmol) was dissolved in a mixture of dry pyridine (40 mL) and acetic anhydride (20 mL). The mixture was allowed to stir at room temperature overnight. The solvent was removed *in vacuo*. Chloroform (300 mL) was added into the remaining residual before washing with 1 M HCl (100 mL), water (100 mL) and brine (100 mL). The solution was dried over anhydrous magnesium sulphate. The remaining residue was washed with petroleum ether (3 x 30 mL). The cyclic *N*-acetyl cysteine thiolactone (**7-15**) was re-crystallized from ethanol. However, the product required heat in order to dissolve in ethanol and no crystal was formed upon cooling as reported in the literature and therefore the synthesis was halted.

An alternative protocol to synthesize **7-15** was the reaction of *N*-acetyl-(*L,D*)-cysteine and dicyclohexyl carbodiimide (DCC) in dioxane. This protocol was adapted from Patent # 3952092 by William Bowen, 1976. *N*-Acetyl-(*L,D*)-cysteine (**7-14**, 398 mg, 2.45 mmol) was dissolved in a solution of DCC (477 mg, 2.31 mmols, 0.94 equiv.) in Aldrich Sure-Seal anhydrous dioxane (6 mL). The reaction was allowed to stir at room temperature for 16 hours before being filtered. The filtrate was then evaporated by rotary evaporation. The remaining solid (290 mg, 2.0 mmols) was dissolved in chloroform (1 mL) and mixed slowly (dropwise) with a solution of the free base glucosamine (**7-7**) that was prepared previously by dissolving solid D-glucosamine-HCl (430 mg, 2.4 mmols) in 1 M NaOH (1.9 mL) and water (2.1 mL). The reaction was left to shake overnight. The aqueous phase of the reaction was isolated and washed with chloroform (2 x 4 mL). The solvent was then evaporated *in vacuo*. The purification was performed by HPLC using a Waters μBandapak C18 reverse phase radial

compression column (25 x 100 mm) and a solvent system as in Table 7.5 (solvent A was water and solvent B was methanol). However, the HPLC chromatogram showed unreactive compound **7-14**, suggesting that the reaction was unsuccessful.

Truncated Mycothiol Synthesis Using Resin-bound Active Ester Reagent

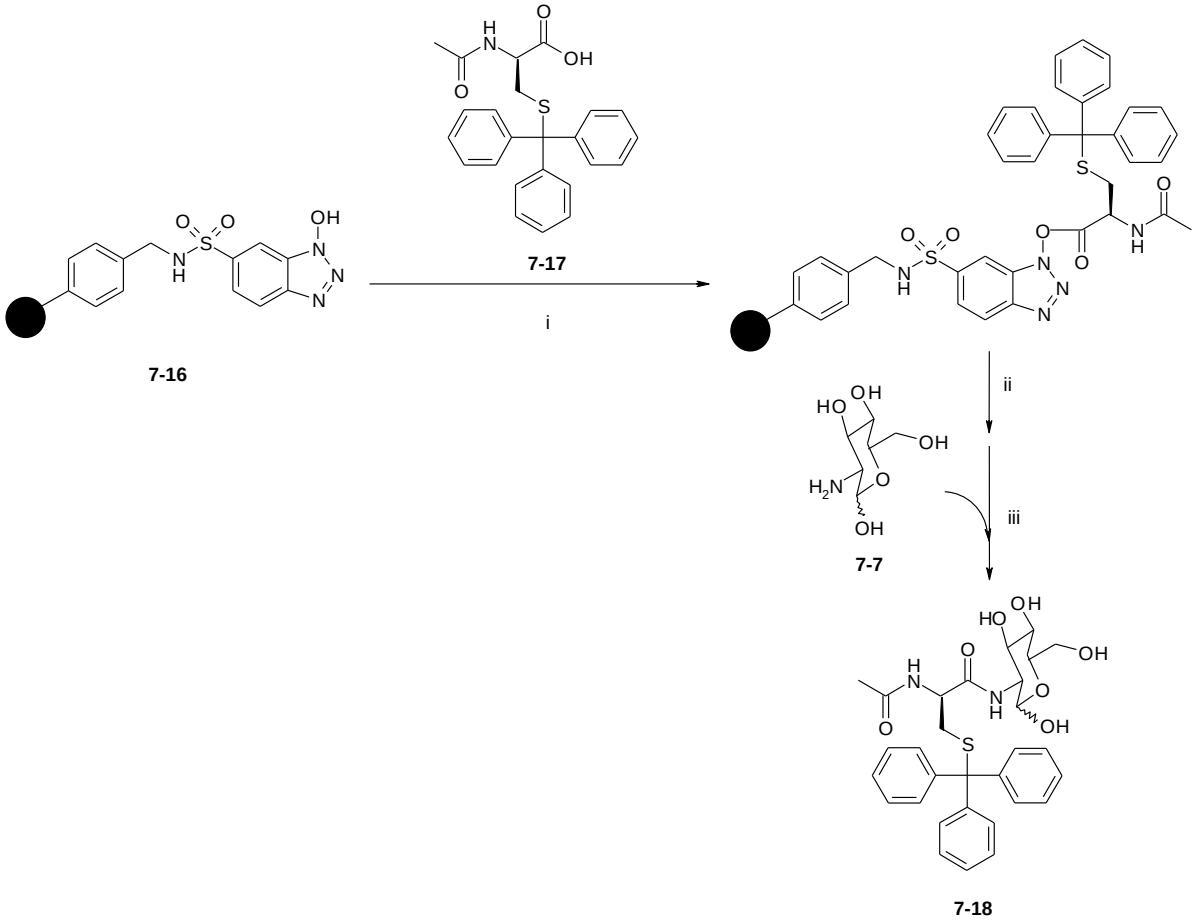


Figure 7.12: The synthesis of *N*-acetyl-*S*-trityl-L-cysteinyl-2-amino-2-deoxy-α,β-D-glucopyranoside (Ac-Cys(Trt)-GlcN). (i) Activating reagent (i.e. ACTU and DIC), DIEA in DMF. (ii) Filtration. (iii) DIEA.

This protocol was adapted from the Synthesis and Purification Catalog (2008-2009) from Biotage (Uppsala, Sweden) with the advantage of having a convenient purification step (filtration and aqueous partitioning instead of chromatography). Ac-Cys(Trt)-OH (**7-17**, 100 mg, 0.24 mmol) was dissolved in DMF (1 mL) and DIEA (0.11 mL, 0.63 mmol, 3 equiv.). To the solution, the resin PS-HOBt(HL) (**7-16**, 220 mg, 0.2 mmol) was added and allowed to stir at room temperature for 2 min before adding the solution of ACTU (99 mg, 0.2 mmol, 1 equiv.) in DMF (1 mL) dropwise. The activating reaction was stirred at room temperature for

1.5 hours. The resin was allowed to settle to the bottom of the reaction vessel and a solution phase was pulled off with pipette. The resin was then washed with dry DMF (6 x 3 mL) before adding to a solution of limited GlcN (**7-7**, 32 mg, 0.147 mmol) in DMF (2 mL) and DIEA (0.1 mL). The coupling reaction was allowed to stir overnight at room temperature. The resin was filtered out and washed with DMF (6 x 3 mL). The filtrate was dried *in vacuo*. The crude product was further purified by silica chromatography (Biotage Flash) with MeOH:CH_2Cl_2 (20:80) since the product mixture was not as pure as was hoped. However, ^{1}H-NMR and ESI-MS did not support the formation of the desired product, **7-18**.

A second trial was performed on larger scale by increasing the activation time from 1.5 hour to overnight. The separation of resin and filtrate was performed by filtration. However, neither ^{1}H NMR nor ESI-MS analyses provided evidence for the existence of **7-18**. The experiment was repeated with increased activating time (overnight) and coupling time (3 days). ^{1}H NMR and ESI-MS experiments still did not provide evidence for the formation of **7-18**.

The protocol was repeated using DIC as activating reagent. Ac-Cys(Trt)-OH (**7-17**, 202 mg, 0.5 mmol) was dissolved in dry DMF (3 mL) and purified by distillation DIEA (0.175 mL, 1 mmol, 2 equiv.). To the solution, the resin PS-HOBt(HL) (**7-16**, 0.55 g, 0.5 mmol) and DIC (0.08 mL, 0.5 mmol, 1 equiv.) were added and allowed to stir at room temperature for 2.5 hours under argon. The consumption of **7-17** was measured by TLC (5% (v/v) methanol and 0.1% (v/v) TFA in dichloromethane with ninhydrin detection). The mixture was then filtered and washed with DMF (6 x 3 mL). The solution of GlcN (**7-7**, 75.4 mg, 0.35 mmol) in DMF (4 mL) and DIEA (0.2 mL) was added to the resin and stirred at 45–50 °C for 4 days. The resin was filtered out and washed with methanol (5 x 5 mL). The filtrate was dried *in vacuo*. ^{1}H-NMR analysis did not support the formation of the desired product, **7-18**. The results are summarized in Table 7.7.

Table 7.7: The summary of des-*myo*-inositol mycothiol synthesis using resin-bound active ester reagent

Reagent	***Activating Time***	***Coupling Time***	***Temperature***	***Yield***
ACTU, PS-HOBt(HL)	1.5 hours	Overnight	RT	NR
	Overnight	3 days	RT	NR
DIC, PS-HOBt(HL)	2.5 hours	4 days	45–50 °C	NR

Truncated Mycothiol Synthesis by Peptide Synthesis Assembly

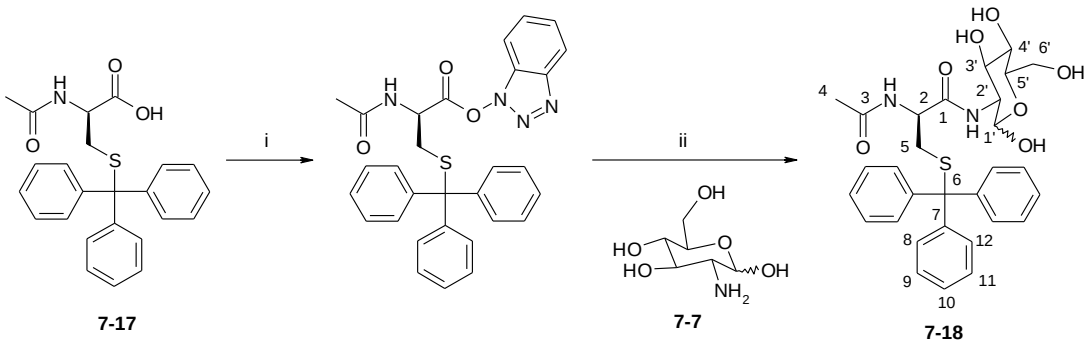


Figure 7.13: The synthesis of *N*-acetyl-*S*-trityl-L-cysteinyl-2-amino-2-deoxy-α,β-D-glucopyranoside (Ac-Cys(Trt)-GlcN). (i) Activating reagent, HOBt, DIEA and DMF. (ii) Imidazole/HCl (pH 7.0) in 50% acetonitrile.

This experiment was performed in two solvent systems; organic (DMF) and aqueous (imidazole/ACN) phases. In both phases, several activating reagents (DCC, DIC, ACTU, HBTU and TNTU), activating time and coupling time were investigated in order to optimize efficiency and product yield. The protocol was adapted from the previous report by Unson *et al*., 1998 [253]. The following is an example of the synthesis in an organic phase using DCC as activating reagent. *S*-Trityl-*N*-acetyl-L-cysteine (**7-18**, 405 mg, 1.0 mmol) and HOBt (0.148 mg, 1.1 mmol, 1.1 equiv.) were dissolved in anhydrous DMF (1 mL) with DIEA (0.26 mL, 1.5 mmol, 1.5 equiv.) under argon before adding dropwise to the solution of DCC (206 mg, 1.0 mmol, 1 equiv.) in anhydrous DMF (0.5 mL). The activating reaction was allowed to stir at room temperature for 15 min and then transferred to a solution of glucosamine-HCl (**7-7**, 129.4 mg, 0.6 mmol, 0.6 equiv.) in anhydrous DMF (10 mL) and DIEA (0.14 mL, 0.84 mmol, 0.84 equiv.). The coupling reaction was stirred at room temperature for 2 hours. The reaction was then quenched by acidifying with TFA (pH 2.0) and concentrated under reduced pressure overnight. The crude product was dissolved in methanol and purified by utilizing either a RediSep C18 chromatographic column (43 g, 20–100% ACN/H_2O + 0.1% TFA) or reverse-phase HPLC on a Waters μBandapak C18 reverse phase radial compression column (25 × 100 mm) using a linear gradient of 40–100% ACN: H_2O + 0.1% TFA with a flow rate of 5 mL/min. ^{1}H NMR and ESI-MS experiments supported the suggestion that purified product, **7-18** was obtained. The synthesis conditions and results are summarized in Table 7.8. The highest recovery yield (40%) was achieved from the reaction using TNTU as activating reagent

with 2 hours activating time and 3 hours coupling time in aqueous phase. The control experiment was performed on sugar, *S*-trityl-*N*-acetyl-L-cysteine and HOBt with the same parameters and conditions on a reverse-phase HPLC in order to compare the retention time with the sample.

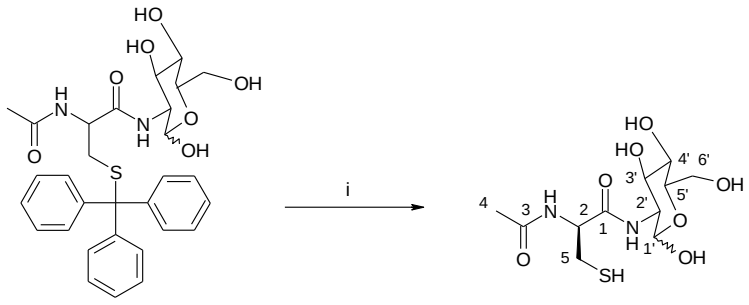
Table 7.8: The summary of des-*myo*-inositol mycothiol synthesis by Peptide Synthesis Assembly

[a]*Reagent*	[b]*Activation Time*	[d]*Coupling Time*	[c]*Solvent System*	*Yield (%)*
DCC, HOBt	15 min	2 hours	DMF	<3
	15 min	3 hours	DMF	Trace
	2 hours	Overnight	DMF	Trace
	3 hours	Overnight	Imidazole/ACN	Trace
DIC, HOBt	15 min	3 hours	DMF	Trace
	3 hours	Overnight	Imidazole/ACN	Trace
ACTU, HOBt	15 min	3 hours	DMF	NR
HBTU	15 min	3 hours	DMF	15
	15 min	1 hour	Imidazole/ACN	NR
TNTU	15 min	1 hour	Imidazole/ACN	25
	2 hours	3 hours	Imidazole/ACN	40

[a] The activating reagents that required HOBt to form HOBt active ester were DCC, DIC and ACTU. If HBTU and TNTU were used, no HOBt was needed. The last two were self-attached (analogue) HOBt active ester.
[b] The activation time was changed in order to maximize active ester intermediate.
[c] Glucosamine was dissolved in 1 N NaOH and freeze dried in lyophilizer overnight prior to increase solubility in DMF. For aqueous phase, GlcN-HCl was dissolved in 0.25 M imidazole HCl buffer (pH 7, 50% ACN).
[d] Various coupling times were applied for the same purpose as activation time.

^{1}H NMR (300 MHz, MeOH-d_4): δ 7.44–7.14 (m, 15H, H8, H9, H10, H11, H12), 5.05 (d, J = 3.3 Hz, 1H, H1′ for α-amomer), 5.02 (d, J = 3.0 Hz, 1H, H1′ for β-anomer), 4.46–4.19 (m, 1H, H2), 3.83–3.18 (m, 6H, H2', H3', H4', H5', H6'), 2.62–2.42 (m, 2H, H5), 1.91 (s, 3H, H4)

ESI-MS (pos): m/z = 567 (MH^+, 85%), 243 (Trt^+, 100%), calculated (M) = 566.68



7-18 **7-2**

Figure 7.14: The synthesis of *N*-acetyl-L-cysteinyl-2-amino-2-deoxy-*α,β*-D-glucopyranoside (Ac-Cys-GlcN). (i) DTT and TFA

The protocol was adapted from the previous report by Photaki *et al*., 1970 [264]. Ac-Cys(Trt)-GlcN (**7-18**, 45 mg, 0.08 mmol) and DTT (6 mg, 0.04 mmol, 0.5 equiv.) were dissolved in TFA (1 mL, 0.07 M) to give a bright yellow solution. The reaction mixture was stirred at room temperature for 30 min then diluted with H_2O (1 mL) and extracted with ethyl acetate (10 x 1 mL). The aqueous phase was freeze dried under reduced pressure overnight to give a white solid of compound **7-2** (21 mg, 82%). The free thiol was quantified by Ellman's reagent (DTNB) [265]. Purified tMSH was titrated with excess DTNB in 50 mM phosphate buffer (pH 8.0). The reaction mixture was incubated in the dark at room temperature for 20 min before measuring the absorbance at 412 nm. The concentration of free thiol was calculated using the Beer-Lambert Law ($A= \varepsilon_{412}BC$, where A is absorption at 412 nm, B is path length (1 cm), C is concentration of free thiol (M) and ε is extinction coefficient (ε_{412} = 14150 $M^{-1}cm^{-1}$ [140]) of free thiol measured at 412 nm). The recovery yield of purified tMSH (**7-2**) was 54% yield (12 mg). The spectroscopic characteristics (^{1}H NMR and ESI-MS) of this compound were in full agreement with those previously reported [116].

^{1}H NMR (300 MHz, D_2O): δ 5.07 (d, J = 2.9 Hz, 1H, H1', only one observed anomer), 4.54–4.52 (m, 1H, H2), 3.84–3.28 (m, 6H, H2', H3', H4', H5', H6'), 2.93–2.73 (m, 2H, H5), 1.94 (s, 3H, H4)

ESI-MS (pos): m/z = 649.2 ($MSSMH^+$, 20%), 325.1 (MH^+, 100%), calculated (M) = 324.1

7.4: Results and Discussion

7.4.1: Isolation of Natural Mycothiol

The reduction of 6-hydroxy-2-naphthyldisulfide (**7-3)** by $NaBH_4$ to form the thiolate of 2-thio-6-hydroxynaphthalene was indicated by the color change (cloudy white to clear yellow). This thiolate anion has a pKa ~6.4, so by adjusting the pH of the buffer to 4.6, the thiolate anion ($–S^-$) would be in its protonated form (–SH). This acidic buffer also quenched the activity of $NaBH_4$, since its maximum activity is at pH 9.0 and it reacts rapidly in acidic aqueous solutions to produce hydrogen and borate. The disappearance of the thiolate anion and the formation of 2-*S*-(2'-thiopyridyl)-6-hydroxynaphthyldisulfide (**7-5**) by reacting the thiol with 2-thiopyridyl disulfide (**7-4)** was indicated by the color change from yellow to colorless. It was found that the crude product could not dissolve appreciably in 50% (v/v) acetonitrile/water mixture and the undissolved precipitate of the crude product was approximately 42% (or 83 mg) of the starting material. This precipitate could, however, be dissolved in methanol. The ESI-MS and ^{1}H NMR analyses suggested that this precipitate was a mixture of the starting material (**7-4)** and the product (**7-5)** (~1:1 ratio). The purification of the crude product by reverse phase HPLC (C18) showed three major peaks with t_R at 38, 42 and 48 min (Figure 7.15). The analyses by ^{1}H NMR experiments indicated that the peak with t_R 38 min was compound **7-4**, the peak with t_R 42 min was compound **7-5** and the peak with t_R 48 min was compound **7-3**. The order of the eluted compounds, where t_R of compound **7-4** (38 min) < compound **7-5** (42 min) <compound **7-3** (48 min), corresponded to the number of aromatic groups (2, 3 and 4 aromatic groups for compounds **7-4**, **7-5** and **7-3**, respectively) in the molecule, suggesting their different hydrophobic interactions with the C18 column. However, the ^{1}H NMR and ESI-MS experiments also suggested the coexistence of compound **7-5** with unreacted starting material **7-4** in the peak with t_R 42 min. The contamination of compound **7-4** might react with mycothiol, leading to the formation of MSSM (Figure 7.16). In order to eliminate the impurity and recover a higher quantity of the product, the purification using gravity silica gel chromatography (with a 10/30 glass column) possessed some advantages over HPLC (with a 25 × 100 mm compression column) including the ability to handle larger scale purification, higher recovery yields and purer product. It was found that the purification using the silica gel chromatographic approach recovered 31% of compound **7-5** compared to 20% yield with reverse phase HPLC, in which the major product loss was due to

multiple injections. The ^{1}H and ^{13}C NMR (Spectra A7.5 and A7.6, Appendix 7) and ESI-MS (Figure A6.5, Appendix 6) of this purification confirmed the purified compound **7-5** or a disulfide tag.

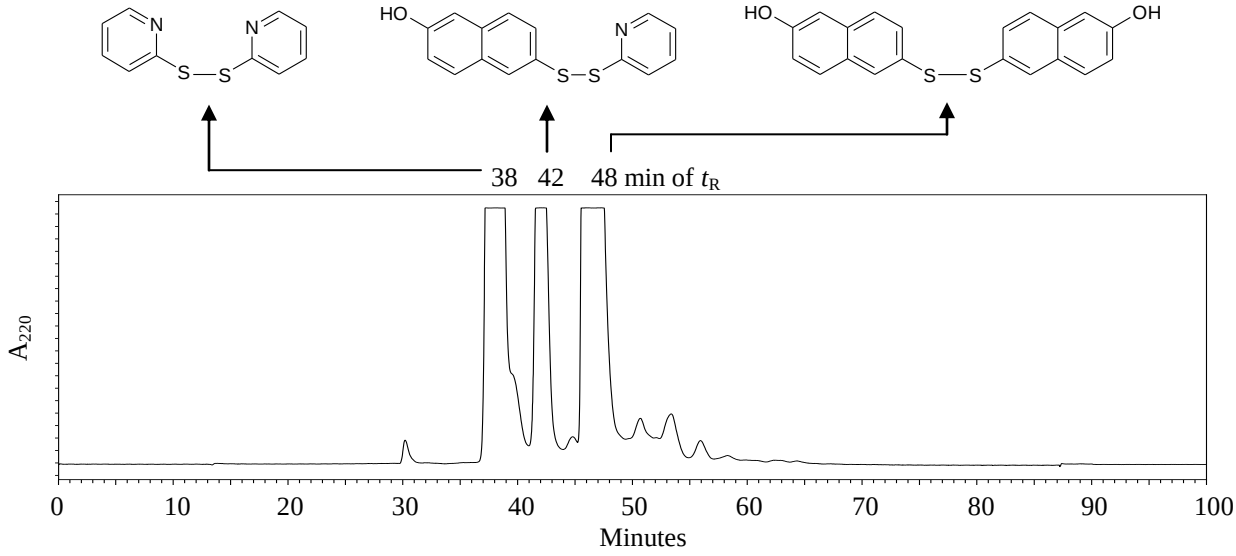


Figure 7.15: The purification of 2-*S*-(2'-thiopyridyl)-6-hydroxynaphthyldisulfide (**7-5**) using reverse phase HPLC (C18). The HPLC chromatogram was monitored at 220 nm with a solvent system shown in Table 7.1. The eluted peak with t_R of 42 min was the desired product **7-5**, while the eluted peaks with t_R 38 and 48 min were 2-thiopyridyl disulfide (**7-4**) and 6-hydroxy-2-naphthyldisulfide (**7-3**), respectively.

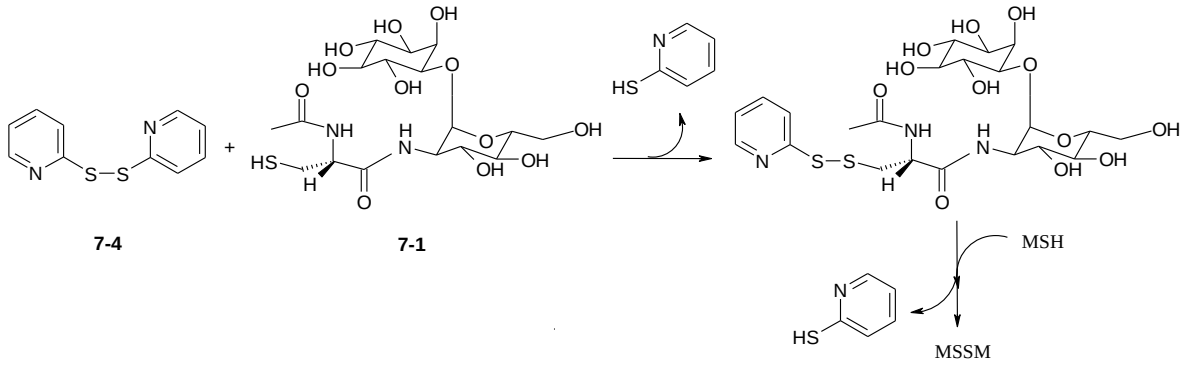


Figure 7.16: The reaction of 2-thiopyridyl disulfide and mycothiol initiates the formation of mycothiol disulfide (MSSM).

The cell lysate from *S. jumonjinesis* after sonication was a cloudy brownish yellow. The color changed to light yellow after the addition of disulfide tag (**7-5**). This step was followed by filtering through a Waters classic Sep-pak (C18) cartridge. Further purification

using the RediSep C18 column gave four eluted peaks with the first fraction being eluted with 40% acetonitrile in 0.1% TFA-water, while the second and the third fractions eluted with the 60% acetonitrile and the forth fractions eluted with 80% acetonitrile (Figure 7.17). These fractions were further purified by reverse phase HPLC (C18) (Figure 7.18). The ESI-MS analysis indicated that fraction 1 from the purification using the RediSep C18 column contained *S*-2-(mycothiolyl)-6-hydroxynaphthyldisulfide or tagged MSH (**7-6**), which showed as the double peaks with t_R ~38–39 min in the HPLC chromatogram. However, this fraction also contained other contaminants. By comparing the HPLC chromatograms, it was suggested that fraction 2 contained compound **7-4** with t_R 42 min, fraction 3 contained compound **7-5** with t_R 46 min and fraction 4 contained compound **7-3** with t_R 49 min (Figure 7.18).

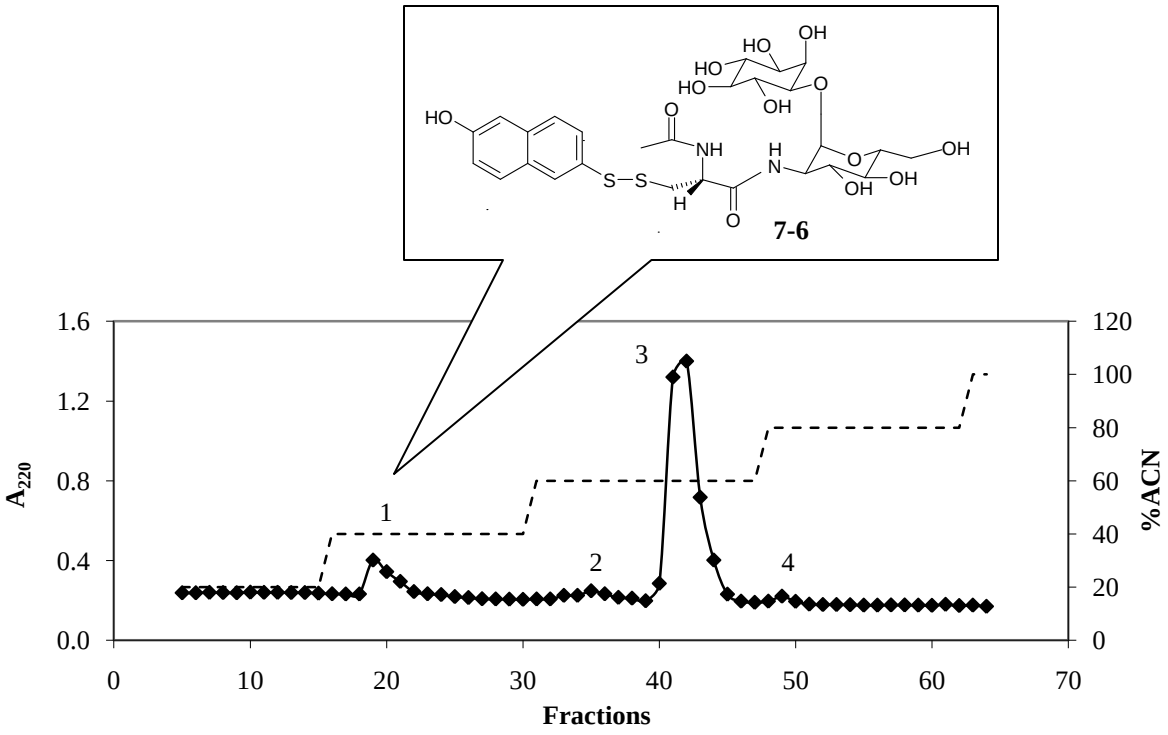


Figure 7.17: The purification of *S*-2-(mycothiolyl)-6-hydroxynaphthyldisulfide (**7-6**) using a RediSep C18 column with acetonitrile in 0.1% TFA-water (20% increment of acetonitrile). The chromatogram was monitored at 220 nm. The number indicated fractions that eluted with different levels of acetonitrile.

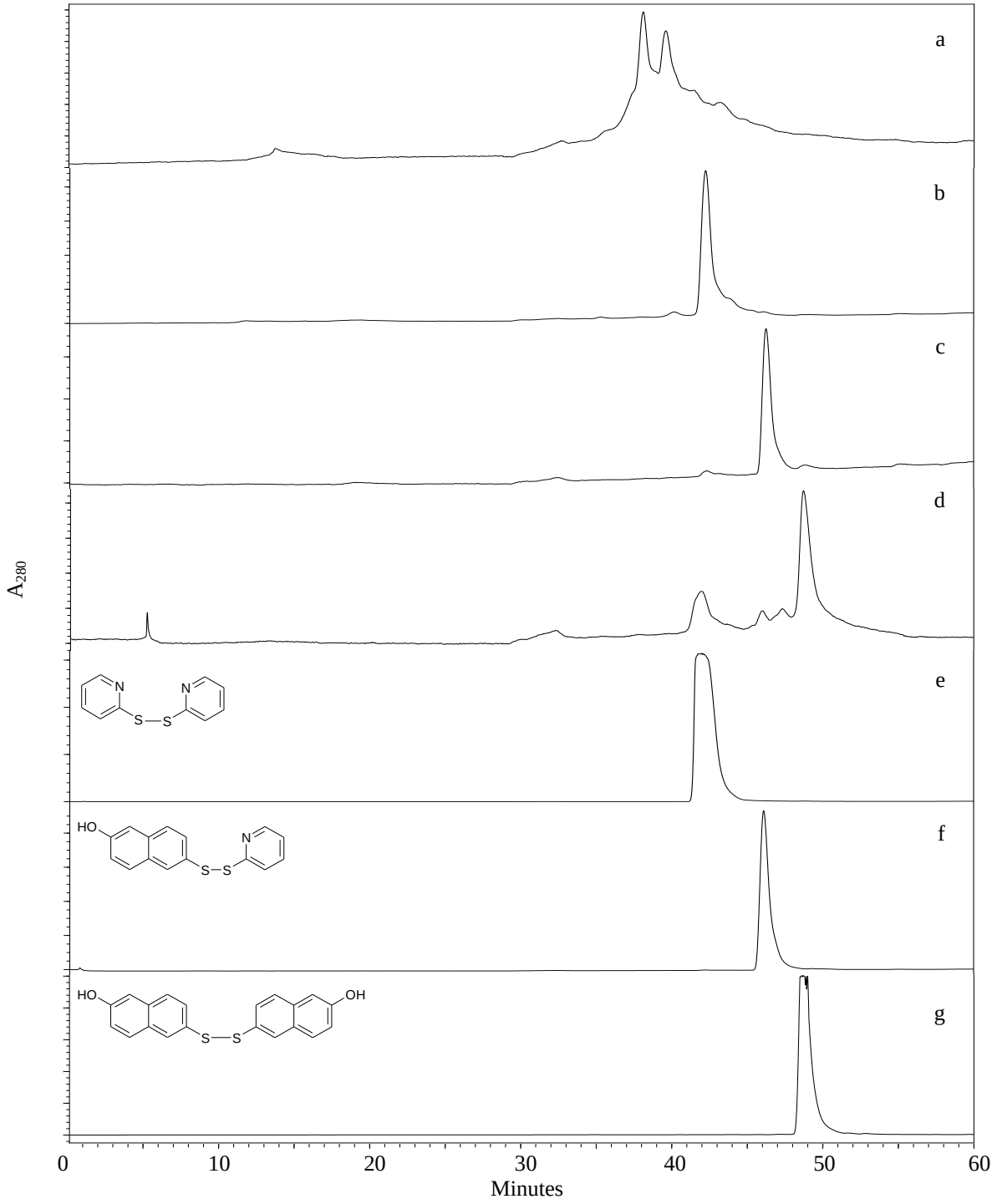


Figure 7.18: The purification of *S*-2-(mycothiolyl)-6-hydroxynaphthyldisulfide (**7-6**) using reverse phase HPLC (C18) and monitoring at 280 nm with a solvent system listed in Table 7.2. The HPLC chromatogram of (a) the fraction 1 (eluted with 40% ACN in 0.1% TFA-water) from the purification using the RediSep C18 column contained compound **7-6** with other contaminations. Other fractions including (b) fraction 2, (c) fraction 3 and (d) fraction 4 that eluted with 60% (fractions 2 and 3) and 80% (fraction 4) acetonitrile in 0.1% TFA-water were eluted at the same retention time as starting materials of (e) 2-thiopyridyl disulfide (**7-4**), (f) 2-*S*-(2'-thiopyridyl)-6-hydroxynaphthyldisulfide (**7-5**) and (g) 6-hydroxy-2-naphthyldisulfide (**7-3**).

To improve recovery yield and avoid product loss from purification using the reverse phase HPLC procedure, a solvent system with each step increasing by 10% (v/v) acetonitrile increment in 0.1% TFA-water (instead of every 20% acetonitrile) was applied to the BioFlash system using the RediSep C18 column to isolate the tagged MSH (**7-6**) (Figure 7.19). Analysis of the reverse phase HPLC chromatogram suggested that the eluted fraction from the RediSep C18 column with 30% ACN in 0.1% TFA-water (fraction 3) contained purified compound **7-6** (Figure 7.20). The result was confirmed by analysis of the ESI-MS (Figure A6.6, Appendix 6) and the ^{1}H NMR (Spectrum A7.7, Appendix 7) experiments. Fraction 2 (eluted with the solvent mixture of 20–30% acetonitrile in 0.1% TFA-water) also contained tagged MSH with other contaminants. Therefore, the compound **7-6** could be isolated using the RediSep C18 column on the BioFlash system utilizing 0–100% acetonitrile in 0.1% TFA-water with 10% acetonitrile increment, where the purified product **7-6** was eluted with 30% acetonitrile. The reverse phase HPLC protocol and other spectroscopic properties were then used to identify the purified compound **7-6**, which eluted with t_R of 38 min as shown in the HPLC chromatogram.

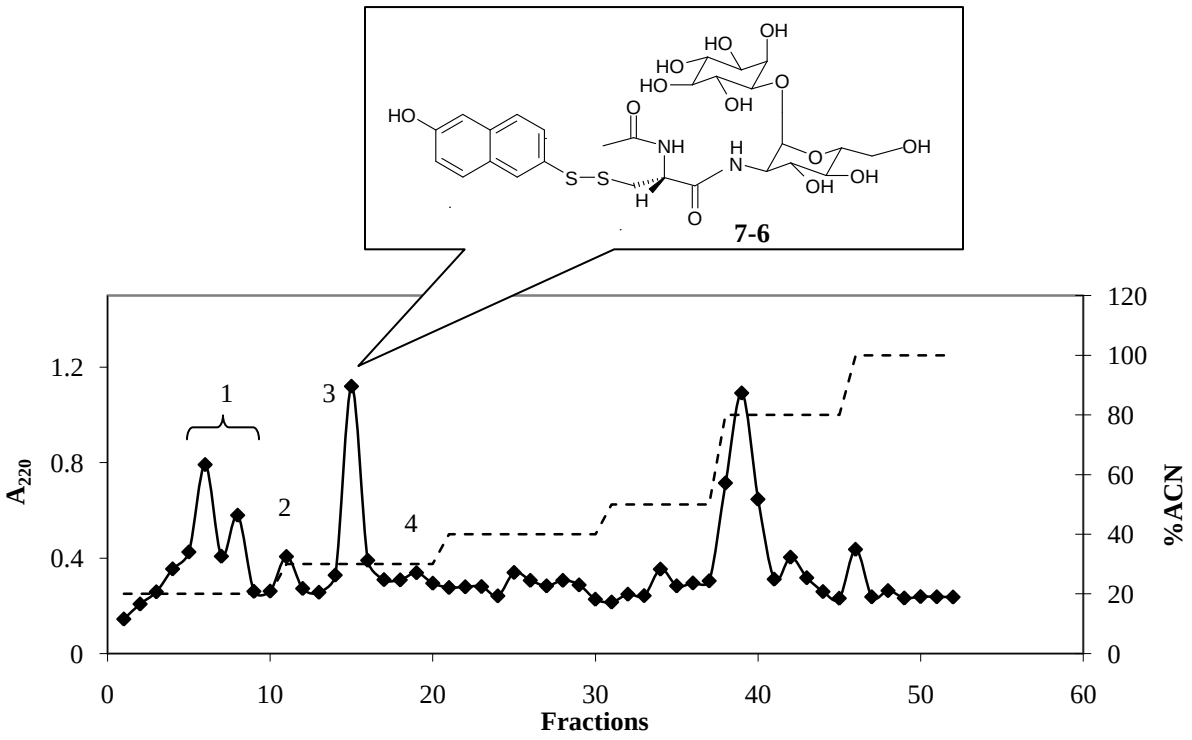


Figure 7.19: The purification of *S*-2-(mycothiolyl)-6-hydroxynaphthyldisulfide (**7-6**) using a RediSep C18 column with acetonitrile in 0.1% TFA-water (10% increment of acetonitrile). The chromatogram was monitored at 220 nm. The number indicated fractions that eluted with different levels of acetonitrile.

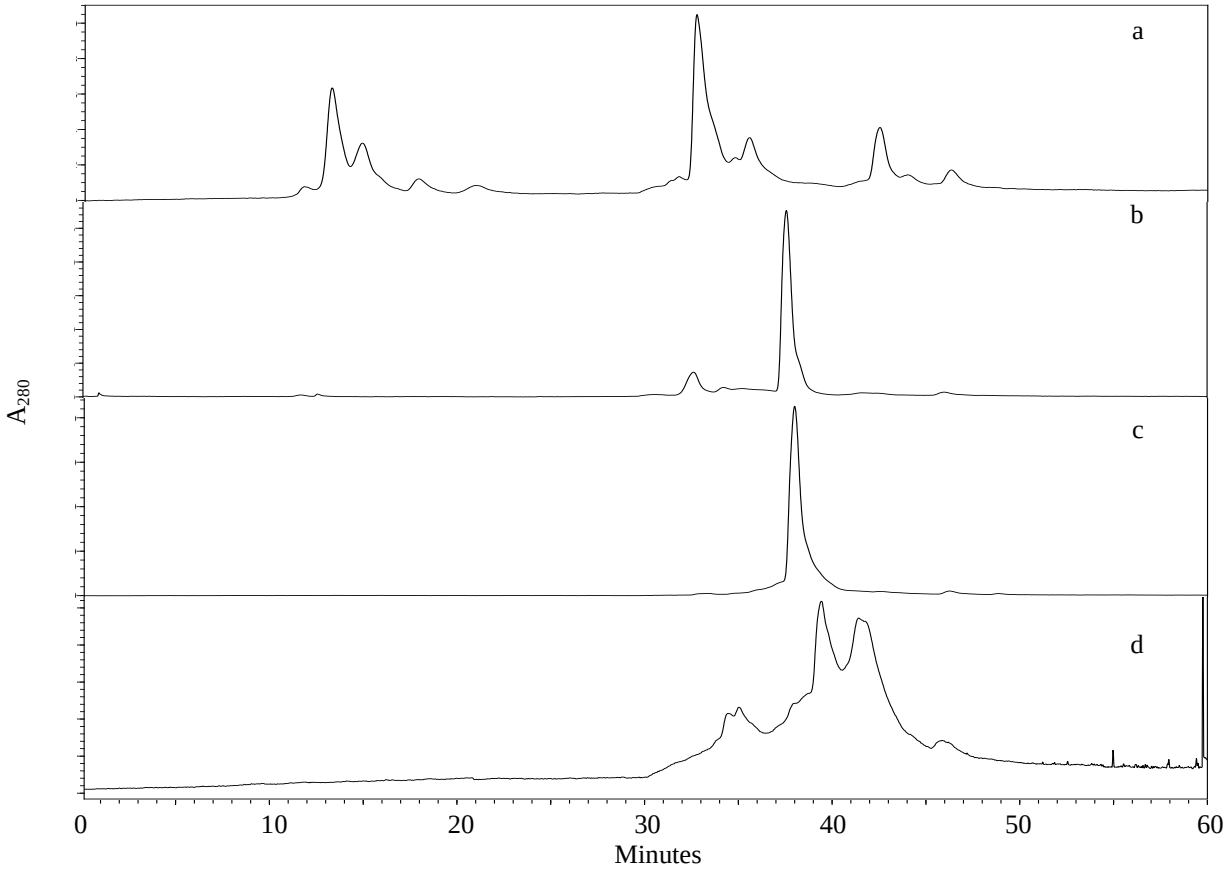


Figure 7.20: The purification of *S*-2-(mycothiolyl)-6-hydroxynaphthyldisulfide (**7-6**) using the reverse phase HPLC (C18) protocol with monitoring at 280 nm with a solvent system as indicated in Table 7.2. The HPLC chromatograms showed impurities in (a) the fraction 1 (eluted with 20% acetonitrile in 0.1% TFA-water), (b) fraction 2 (eluted with 20–30% acetonitrile in 0.1% TFA-water), (c) fraction 3 (eluted with 30% acetonitrile in 0.1% TFA-water) and (d) fraction 4 (eluted with 30% acetonitrile in 0.1% TFA-water) from the purification using the RediSep C18 column with 0–100% acetonitrile in 0.1% TFA-water (10% acetonitrile increment). Only fraction 3 contained purified compound **7-6**.

The tagged MSH (**7-6**) was then cleaved by using tris(2-carboxyethyl)phosphine (TCEP), which is a stable rapid disulfide reducing reagent that is resistant to air oxidation [251]. The reaction of TCEP with the disulfide compound yielded TCEP oxide as well as reduced thiol product, in our case, reduced MSH (Figure 7.21). The isolation of MSH (**7-1**) was performed by using the reverse phase HPLC (C18) protocol, which gave a peak with t_R 13–14 min (Figure 7.22). The ESI-MS and ^{1}H NMR analysis showed the existence of MSH as well as unreacted TCEP and TCEP oxide (Figure A6.7, Appendix 6 and Spectrum A7.7, Appendix 7). The co-elution of MSH with TCEP and TCEP oxide might correspond to poor isolation using the C18 column with a high flow rate (4 mL/min). Lower flow rates with

longer elution times might be useful in isolating compound **7-1** from the reducing reagent. Nevertheless, the results from ESI-MS and ^{1}H NMR experiments corresponded to the presence of the disulfide form of mycothiol, a small amount for reference purposes was kindly supplied by Dr. Gerald Newton from University of California (SanDiego, La Jolla, CA) (Figure A6.8, Appendix 6 and Spectrum A7.8, Appendix 7). However, due to the low amount of purified tagged MSH, the isolation of MSH from *S. jumonjinesis* cells was put on hold.

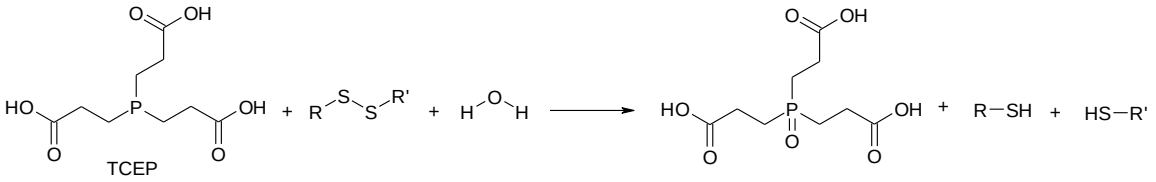


Figure 7.21: The reduction of the disulfide bond by Tris(2-carboxyethyl)phosphine (TCEP).

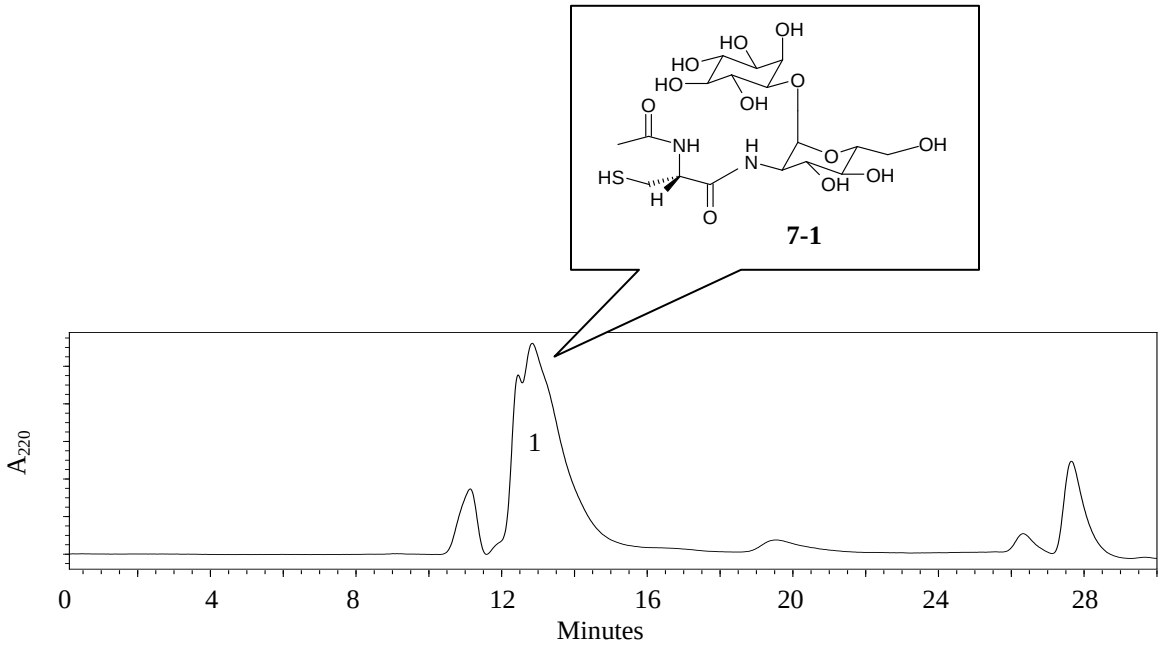


Figure 7.22: The purification of mycothiol (**7-1**) after the cleavage of hydroxynaphthyldisulfide by TCEP using reverse phase HPLC (C18) and a solvent system in Table 7.3. The chromatogram was monitored at 220 nm. Fraction 1 was found to contain MSH but contaminated with TCEP.

7.4.2: Synthesis of Des-*myo*-Inositol Mycothiol

Truncated Mycothiol Synthesis by Patel and Blanchard's Method

The first step of this method was the coupling reaction of glucosamine (**7-7**) to Fmoc-Cys(Acm)-OPfp (**7-8**) using HOBt in DMF. The coupling reagent HOBt acts as a nucleophile toward Fmoc-Cys(Acm)-OPfp, resulting in the formation of an active ester Fmoc-Cys(Acm)-

OBt that was reactive toward the nucleophilic glucosamine (Figure 7.23). The product, Fmoc-Cys(Acm)-GlcN (**7-9**), was soluble in aqueous solvent and gave a yellowish fine powder, while an insoluble white precipitate of a side product, pentafluorophenol, could be easily filtered out. The ^{1}H and ^{13}C NMR (Spectra A7.9 and A7.10, Appendix 7) and ESI-MS (Figure A6.9, Appendix 6) experiments indicated purified product **7-9**. If water was presented in the reaction, it could act as a nucleophile in place of glucosamine and form an acid side product, Fmoc-Cys(Acm)-OH (Figure 7.23).

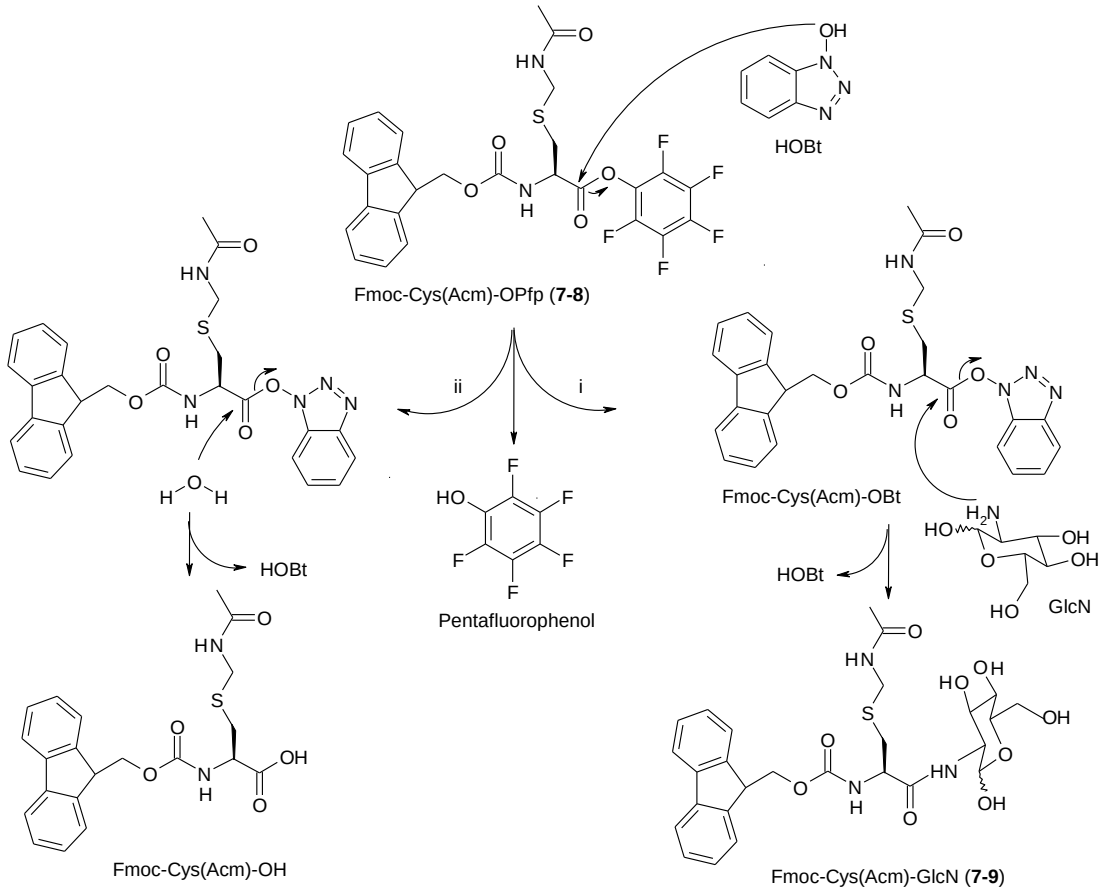


Figure 7.23: The proposed reaction for the synthesis of *N*-α-Fmoc-*S*-acetamidomethyl-L-cysteinyl-2-amino-2-deoxy-α-D-glucopyranoside (Fmoc-Cys(Acm)-GlcN, **7-9**), where glucosamine acts as a nucleophile to form an active ester Fmoc-Cys(Acm)-OBt (i). When water is present in the reaction, the side product Fmoc-Cys(Acm)-OH was present (ii).

The next step was to remove the N-terminal protecting group, Fmoc, using piperidine in DMF. The reaction generated a glassy mixture of a water soluble product, NH_2-Cys(Acm)-GlcN (**7-10**) and a white precipitate of cleaved dibenzofulvene-piperidine adduct, which

formed by the reaction of the exocyclic carbon of dibenzofulvene with nitrogrn from the piperidine (Figure 7.24). The desired product was easily isolated by filtration. The ^{1}H NMR (Spectra A7.11, Appendix 7) and ESI-MS (Figure A6.10, Appendix 6) indicated the purified product **7-10**.

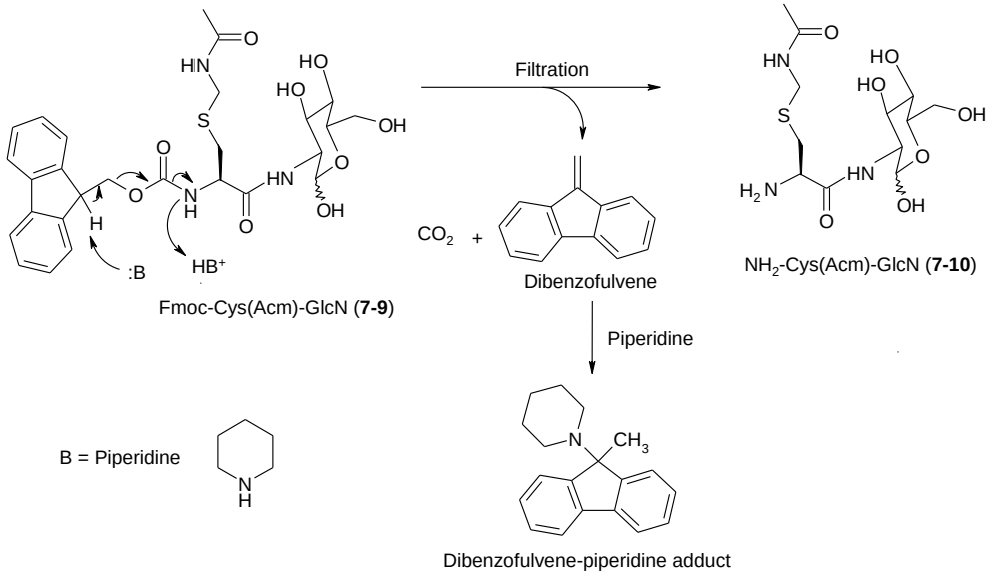


Figure 7.24: The proposed reaction mechanism of the synthesis of *S*-acetamidomethyl-L-cysteinyl-2-amino-2-deoxy-α,β-D-glucopyranoside (NH_2-Cys(Acm)-GlcN, **7-10**).

The acetylation at the N-terminus of the compound **7-10** was performed in an acidic solution at a selected pH. Slow addition of acetic anhydride was used to accomplish the acetylation, preferentially on the amino group (pKa of NH_2-Cys ~9 and pKa of hydroxide on sugar ~15). If acetic anhydride is added too quickly, acetylation of sugar hydroxyls might occur as the exothermic reaction and might result in a less specific acetylation. The ^{1}H NMR analysis of the crude product, Ac-Cys(Acm)-GlcN (**7-11**), showed contamination due to small traces of piperidine and ethanol. The reverse phase HPLC (C18) protocol was used to isolate the compound **7-11**. The HPLC chromatogram showed three eluted peaks with t_R 9, 11 and 12 min (Figure 7.25). The ^{1}H NMR data suggested that the fractions with t_R 11 and 12 min contained stereoisomers of compound **7-11** (Spectrum A7.12, Appendix 7). These results were also confirmed by analysis of the ESI-MS (Figure A6.11, Appendix 6).

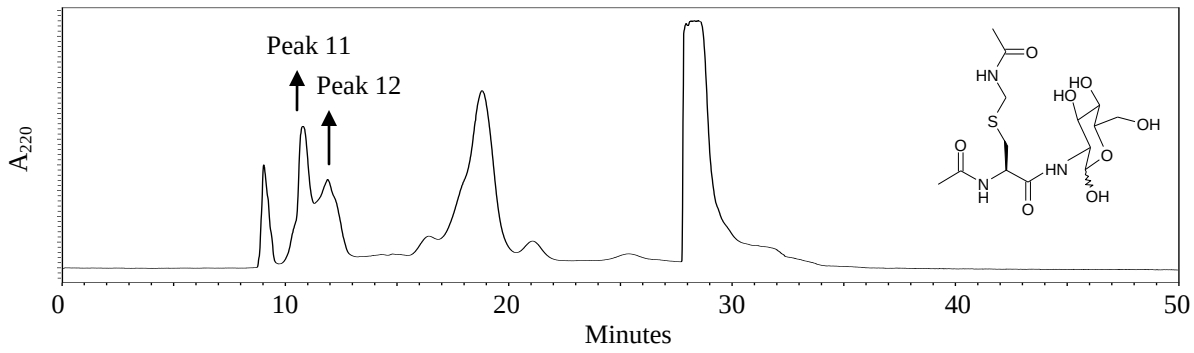


Figure 7.25: The purification of *N*-acetyl-*S*-acetamidomethyl-L-cysteinyl-2-amino-2-deoxy-α,β-D-glucopyranoside (Ac-Cys(Acm)-GlcN, **7-11**) using reverse phase HPLC. The chromatogram was monitored at 220 nm with solvent system as shown in Table 7.4.

Heavy metals such as Hg^{2+}, Ag^{+}, and Tl^{3+} were used as reported previously to remove 2-(4-methoxytrityl)sulfenyloxymethylbenzoyl (Mob) and acetamidomethyl (Acm) protecting groups from Cys and Sec sidechains in peptide synthesis [266-268]. Thallium compound $(CF_3COO)_3Tl$ in TFA acts as a soft acid that is reactive toward sulfur in Ac-Cys(Acm)-GlcN (**7-11**), forming a thallium-sulfur complex and an iminium ion that can be removed by a scavenger such as anisole (Figure 7.26). Excess compound **7-11** then removes metal and forms a disulfide compound GlcN-AcCys-AcCys-GlcN (**7-12**). Despite their successful reported use as a deprotecting reagent, the disadvantages of using heavy metals includes their toxicity, handling difficulty, formation of difficult to redissolve precipitates and the production of strong metal-sulfur interactions, which lead to incomplete removal of metal and, thus, results in product loss. Since the disulfide compound **7-12** was expected to possess high solubility in water, the HPLC chromatogram of this reaction indicated two fractions with t_R 7 and 9 min that might contain the desired product (Figure 7.27). The ^{1}H NMR analysis suggested that the fraction with t_R 9 min might contain compound **7-12**, but the detection was very low due to product loss during the purification. However, the ESI-MS result did not support this conclusion, since neither free thiol compound **7-2** with m/z 324 nor its disulfide compound **7-12** with m/z 646 were detected.

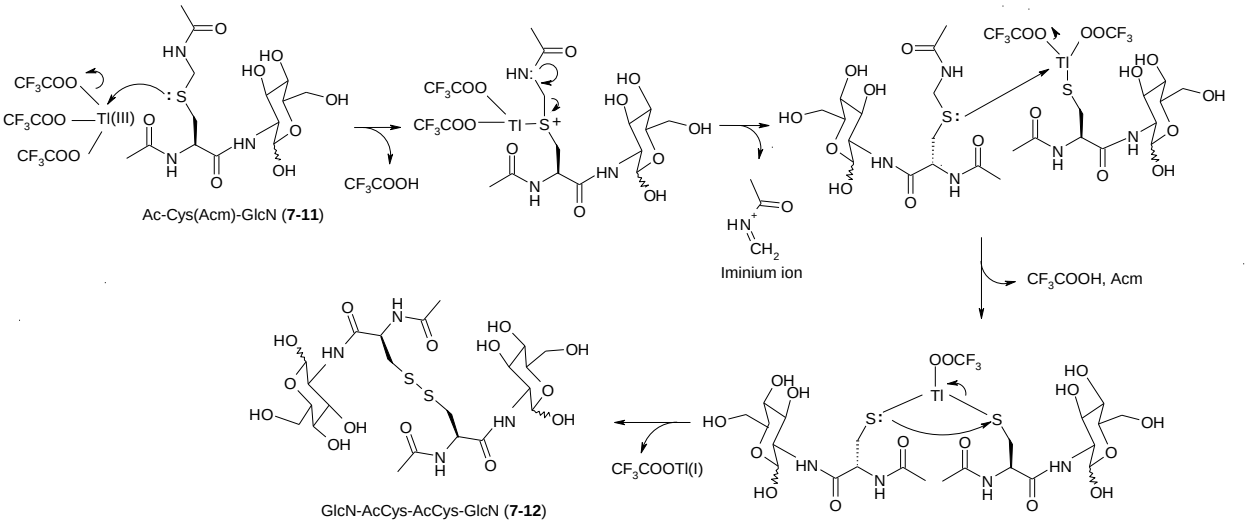


Figure 7.26: The proposed reaction mechanism of *N*-acetyl-L-cysteinyl-2-amino-2-deoxy-α,β-D-glucopyranoside disulfide (GlcN-AcCys-AcCys-GlcN, **7-12**) by oxidation in the presence of $(CF_3COO)_3Tl^{III}$. Mechanism is as suggested by Fujii *et al*., 1987 [266].

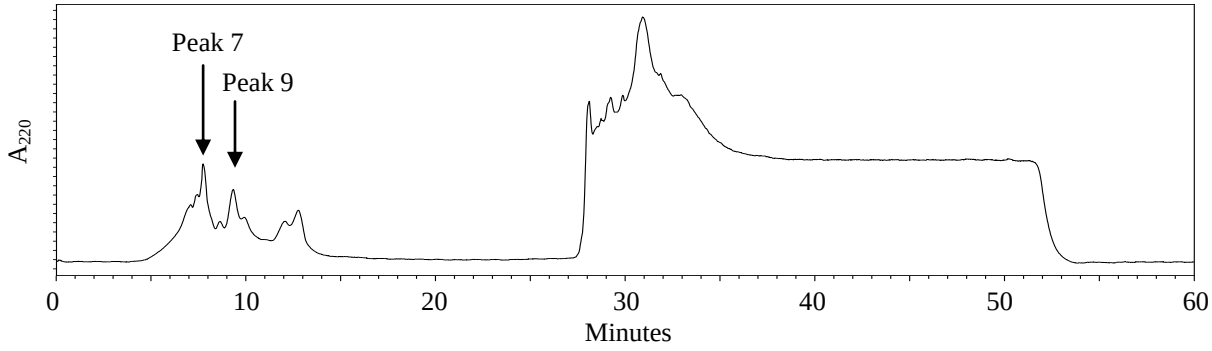


Figure 7.27: The purification using reverse phase HPLC (C18) of the attempted deprotection of *N*-acetyl-L-cysteinyl-2-amino-2-deoxy-α,β-D-glucopyranoside disulfide (GlcN-AcCys-AcCys-GlcN, **7-12**) with thallium (III) trifluoroacetate as a deprotecting reagent. The HPLC chromatogram was monitored at 220 nm using the solvent system in Table 7.5.

Due to the toxicity of thallium, iodine was introduced as an alternative choice of deprotection reagent. A divalent sulfur atom that acts as a donor in Ac-Cys(Acm)-GlcN (**7-11**) is reactive toward an acceptor iodine, forming an iodosulfonium ion (Figure 7.29). This ion can perform self-cleavage to form a sulfenyl iodide complex and an iminium ion, which is removed by a scavenger such as water. The electrophilic sulfur atom in sulfenyl iodides can react with another sulfenyl iodide to form a disulfide product (**7-12**) with iodine (I_2) as side product. As well, it can also react with a nucleophilic sulfur atom in compound **7-11** to form compound **7-12** and iminium ion. Even though I_2 might have the advantage of being less toxic

than heavy metals, its application is limited to a unique condition on each individual reaction and a side product of iodinated adducts are commonly found [262, 263]. Choice of solvent and protecting group also altered the efficiency of the cleavage reaction. It was reported previously that the trityl protecting group in Cys(Trt) is more reactive than that of Cys(Acm) in most solvents (i.e. methanol, AcOH, dioxane, etc.) except for DMF [262]. It was also predicted that water, acting as a scavenger, is a driving force for this reaction [262]. However, in our experiments the use of water as the reaction solvent was not successful. However, the reaction was maximized by performing the experiment in DMF with a one minute reaction time, even though the half life of a similar reaction was reported to be as short as 2–3 sec [262]. Besides, upon the addition of the water-chloroform mixture, the reaction became hot and smoke was observed. The HPLC chromatogram showed two fractions with t_R 10 and 12 min that eluted with water containing 0.1% TFA (Figure 7.28). Analysis of the ^{1}H NMR of the fraction with t_R of 10 min indicated a low quantity of the sugar group, while the fraction with t_R of 12 min showed low levels of contamination along with the compound **7-12**. Neither was detected in the ESI-MS analysis for either fraction however.

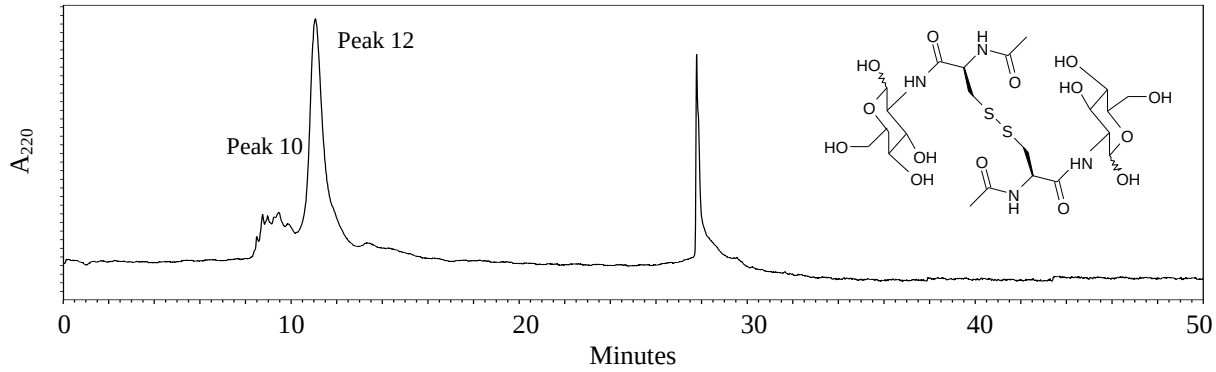


Figure 7.28: The purification using the reverse phase HPLC (C18) protocol of the reaction of *N*-acetyl-L-cysteinyl-2-amino-2-deoxy-α,β-D-glucopyranoside (GlcN-AcCys-AcCys-GlcN, **7-12**) using iodine as deprotecting reagent. The HPLC chromatogram was monitored at 220 nm using the solvent system in Table 7.5.

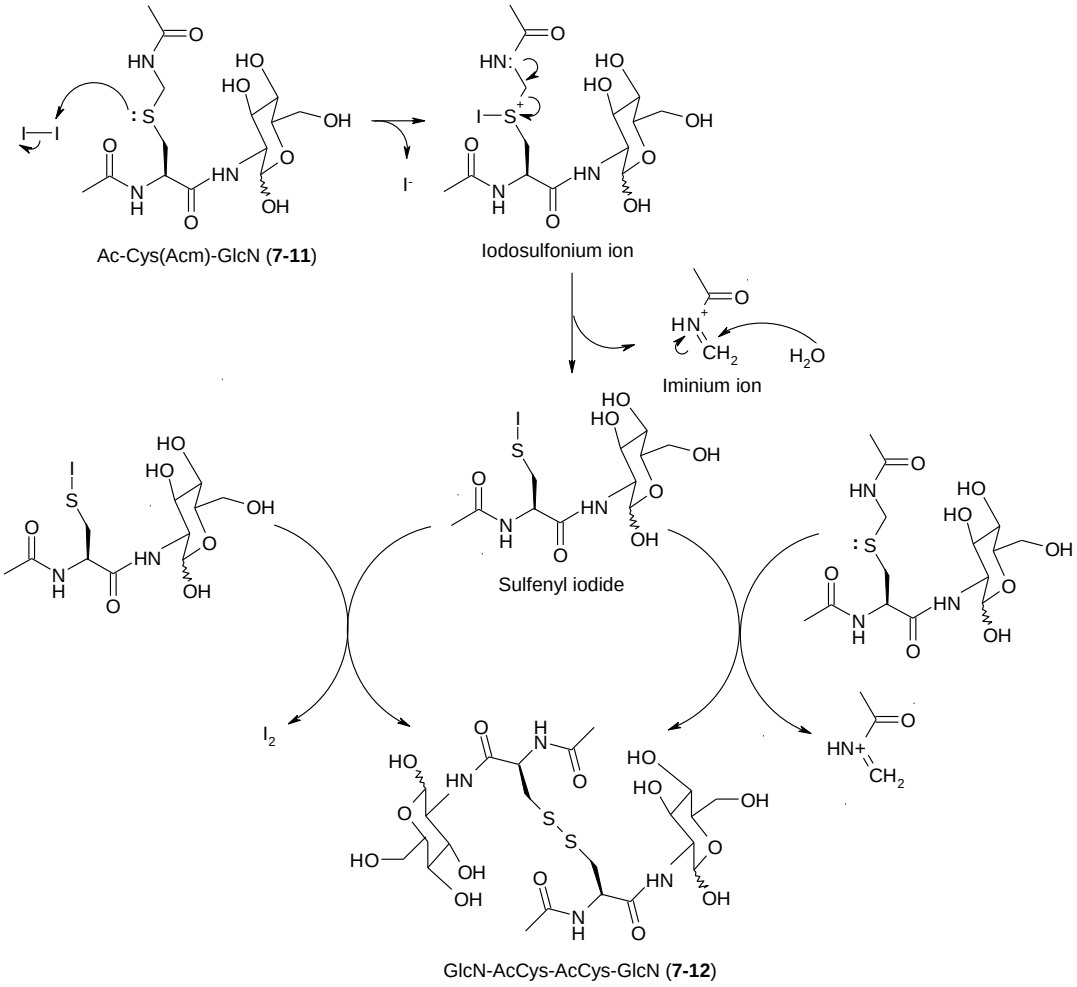


Figure 7.29: The proposed reaction mechanism of *N*-acetyl-L-cysteinyl-2-amino-2-deoxy-α,β-D-glucopyranoside (GlcN-AcCys-AcCys-GlcN, **7-12**) synthesis using iodine as deprotecting reagent to remove the acetamidomethyl group.

The next deprotection reagent used in our experiment was disulfide 2,2'-dithiobis(5-nitropyridine) (DTNP). The reaction of highly electrophilic DTNP with thioanisole in TFA is involved in many applications of protecting group removal including removal of the Acm group, which is difficult to remove and the removal requires more than 15 equivalents of DTNP in order to obtain a 62% recovery yield of deprotected Cys peptide [263]. In our experiments that were concerned with the deprotection of the Acm functionality from compound **7-11**, the reaction was initiated by activating the deprotection reagent, DTNP (Figure 7.30). TFA activates DTNP by protonating the nitrogen atom of the pyridine ring, thus creating an electron sink and generating a more electrophilic disulfide bond on the DTNP. Thioanisole that acts as a catalyst reacts with DTNP, forming a reactive trivalent sulfonium

cation that is more reactive than DTNP and the co-product 2-thio-5nitro-pyridine. The less nucleophilic sulfur atom of Ac-Cys(Acm)-GlcN (**7-11**) then attacks this trivalent cation, releasing thioanisole and forming Ac-Cys(5-Npys)-GlcN (**7-13**) with the addition of a scavenger such as water.

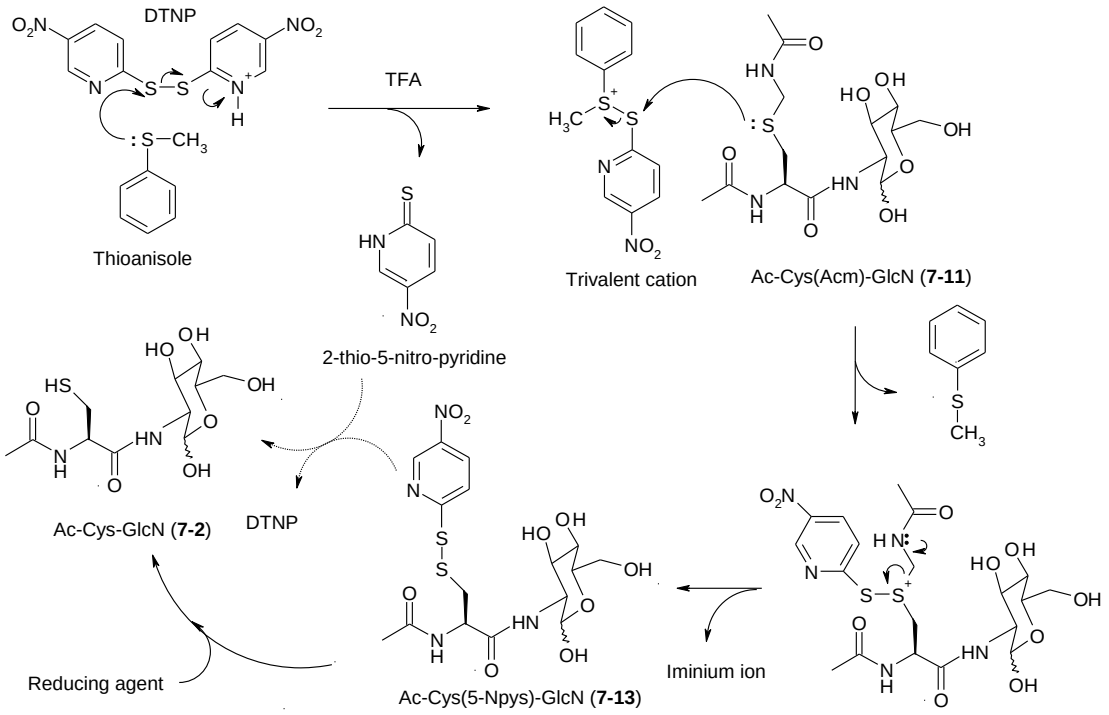


Figure 7.30: The proposed reaction mechanism for the synthesis of des-*myo*-inositol mycothiol (Ac-Cys-GlcN, **7-2**) by the cleavage of the acetamidomethyl group using DTNP as deprotecting reagent.

The purification using reverse phase HPLC protocol (Figure 7.31) and ^{1}H NMR analysis suggested that the peaks with t_R 9 min might contain compound **7-2**. However, ESI-MS analysis did not support this observation since no molecular mass of the compound **7-2** was detected. The eluted peak with t_R 12 min might contain compound **7-13**, which was also detected by the ^{1}H NMR experiment. However, the detection was low and its mass could not be found by ESI-MS analysis. These data suggested that two forms of deprotected Cys existed; the 5-Npys protected derivative compound **7-13** and the free thiol compound **7-2**. It was found earlier that the ratio of free seleno-peptide to the protected peptide is in reversed proportion to the quantity of DTNP. The ratio of free seleno-peptide to 5-Npys protected Sec was reoprted to be 82/18 when 0.5 equivalents of DTNP was applied and the ratio increased to

56/44 with 5 equivalents of DTNP [263]. A possible explanation for the formation of compound **7-2** without the addition of reducing agent was the attack of free 2-thio-5-nitro-pyridine that was released during the deprotection to the disulfide (–S-S–), and interaction with compound **7-13**, forming the free thiol and regenerating DTNP (Figure 7.30). This free thiol eluted earlier with higher concentration of the aqueous solvent than the 5-Npys protected derivative. The retention time (9 min) of compound **7-2** corresponed to what had been observed earlier in the deprotection of the Acm group of compound **7-11** by thallium. The product yield was low due to the disadvantages of HPLC purification, which included product loss due to filtration and multiple injections.

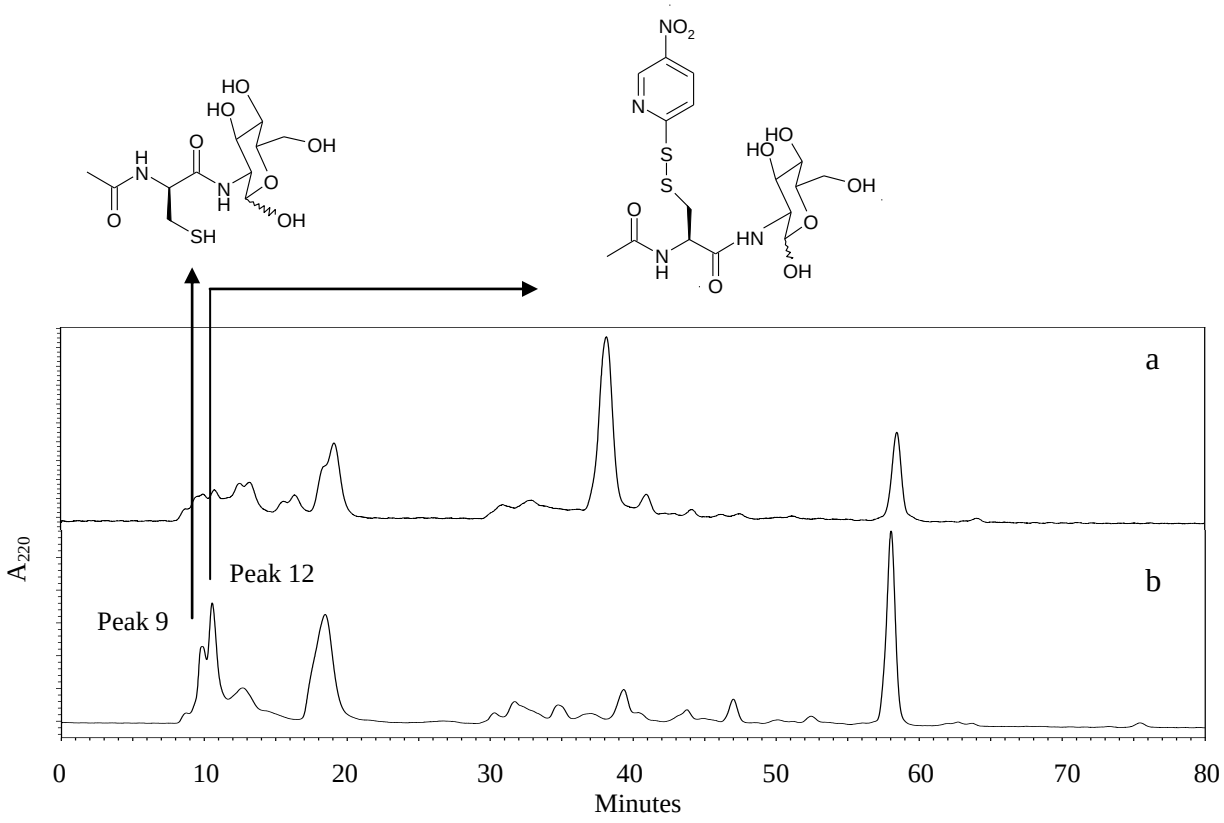


Figure 7.31: The removal of the acetamidomethyl group from *N*-acetyl-*S*-acetamidomethyl-L-cysteinyl-2-amino-2-deoxy-α,β-D-glucopyranoside (Ac-Cys(Acm)-GlcN, **7-11**) using 2 equivalents (a) and 15 equivalents (b) of DTNP in TFA. The reverse phase HPLC (C18) chromatogram was measured at 220 nm using the solvent system in Table 7.6. The peaks with t_R 19 min contained **7-11**, the peak with t_R 38 might contain **7-2** with contaminants and the peak with t_R 58 min might contain **7-13**.

Truncated Mycothiol Synthesis via a Cyclic Thiolactone Intermediate

The synthesis of des-*myo*-inositol mycothiol (AcCys-GlcN, **7-2**) *via* a cyclic thiolactone intermediate was adapted from the synthesis of glycol-*S*-nitrosothiol [261]. This is a simple

two-step protocol that required less time and chemicals (Figure 7.32). It was previously reported that the reaction of *N*-acetyl penicillamine with acetic anhydride and pyridine initiated the cyclic 3-acetamido-4,4-dimethylthietan-2-one *N*-acetyl penicillamine (51% yield), followed by the coupling reaction with aminated sugar to form glycol-*S*-nitrosothiol (30% yield) [261]. In an attempt to synthesize compound **7-2**, *N*-acetyl cysteine (*N*-AcCys) (**7-14**) was chosen as a starting material with either DCC in acid solution or acetic anhydride in basic solution to form a cyclic *N*-acetyl cysteine thiolactone (**7-15**).

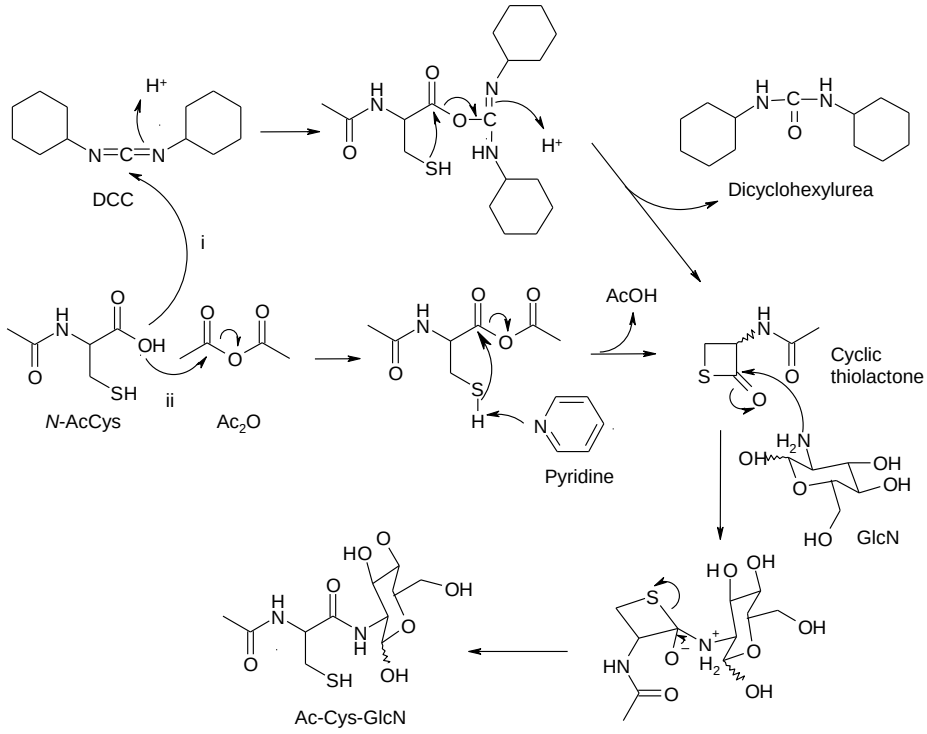


Figure 7.32: The proposed mechanism of des-*myo*-inositol mycothiol *via* a cyclic thiolactone intermediate. (i) DCC in acetic solution. (ii) Acetic anhydride and pyridine in basic solution.

The first synthesis of cyclic *N*-acetyl cysteine thiolactone (**7-15**) by Ac_2O and pyridine was initiated by the reaction of the hydroxyl oxygen from compound **7-14** that attaches to the acyl carbon of acetic anhydride, forming the mixed anhydride *N*-AcCys-*C*-OAc. The sulfur from the thiol group of Cys is then deprotonated by pyridine and self attaches with its C-terminal carboxyl carbon, thus generating the cyclic thiolactone and AcOH as the co-product. The reaction, however, was unsuccessful since no crystals were formed during the cooling

process in ethanol. The failure might be due to excess solvent such that it might take a longer time to from crystals of compound **7-15**.

The second approach was to use DCC, which activates the carboxylic acid, thus, creating the active cyclic thiolactone, which would be reactive toward the amine group of the glucosamine. However, dicyclohexylurea would be a co-product. This reaction gave unclear results since there was no supporting evidence by ESI-MS or ^{1}H NMR analyses. The reaction solution of the possible cyclic thiolate was directly coupling with the free base of D-glucosamine. Analyses of the HPLC chromatograms for the purification of des-*myo*-inositol mycothiol showed the overlapping peaks of the reaction mixture and compound **7-14** (data not shown), indicating that the reaction was unsuccessful.

In order to improve this protocol in the future, the isolation of the cyclic thiolate and its unambiguous identification by MS and NMR analyses should be performed before continuing with the coupling step. The cyclic thiolate should be identified by CI-MS (and not ESI-MS) due to its chemical structure that was difficult to protonate.

Truncated Mycothiol Synthesis Using Resin-bound Active Ester Reagent

This method used the starting material of Ac-Cys(Trt)-OH (**7-17**), which might contain the D, L-racemic mixtures of the *N*-acetyl cysteine. Thus, racemization might occur with the acetyl protecting group in this synthesis, which is acceptable since we primarily aimed for testing the method by using this approach. The reaction was performed in DMF in order to avoid the decomposition of the PS-HOBt active ester from reacting with water. Since the sugar possessed low solubility in organic solvents, this might be the reason for the failure of the reaction. Analyses of the ^{1}H NMR of the filtrate from the activation step using ACTU as the activating reagent indicated the presence of starting material Ac-Cys(Trt)-OH (**7-17**), which suggested that compound **7-17** might not attach to the resin. A brown wax was recovered after the coupling step with D-glucosamine. Analysis of the ^{1}H-NMR only indicated the presence of unreacted sugar (**7-7**) with no resonance which would indicate presence of the product Ac-Cys(Trt)-GlcN (**7-18**). Neither was detected based on analyses of the ESI-MS results. By increasing the activation time from 1.5 hours to overnight, analyses of the ^{1}H NMR of the crude product only indicated the presence of the unreacted sugar (**7-7)**. The ESI-MS analysis showed low levels of a mass corresponding to the cleaved Trt group (m/z 243, Trt^{+})

but no product mass was observed. These results indicated that only trace amounts of compound **7-17** might have been attached to the resin, thus the coupling reaction was likely inefficient. The reaction might require longer activating and coupling times.

For the experiment using DIC as an activating reagent, the coupling reaction was observed to have failed even though the solubility of the D-glucosamine in DMF was increased at the higher temperature (45–50 °C) and longer coupling time (4 days) used in this experiment. In order to improve the protocol using this approach, higher temperatures and even longer activating and coupling times might be necessary. As well, the filtrate after the activation and coupling reactions should be examined for unreacted compounds **7-17** and **7-7** such that the suitable time and temperature for this reaction can be optimized.

Truncated Mycothiol Synthesis by Peptide Synthesis Assembly

There were several factors concerning the synthesis of tMSH by peptide synthesis assembly that affected the formation of product and recovery yield. These factors likely included the activating reagent, the activation and coupling times, the purification protocols and the limited solubility of D-glucosamine in the reaction solvents. The synthesis was initiated by the reaction of starting material, Ac-Cys(Trt)-OH (**7-17**), with various activating reagents including DCC, DIC, ACTU, HBTU and TNTU (Figure 7.33). The base used in the experiment (DIEA) supported the deprotonation of the carboxyl group of compound **7-17**, which then formed a reactive ester with the activating reagent. HOBt then reacted with this activated carboxyl group, resulting in the formation of an activated ester of compound **7-17**, which then coupled with the free base form of D-glucosamine to form the product, Ac-Cys(Trt)-GlcN **(7-18)**.

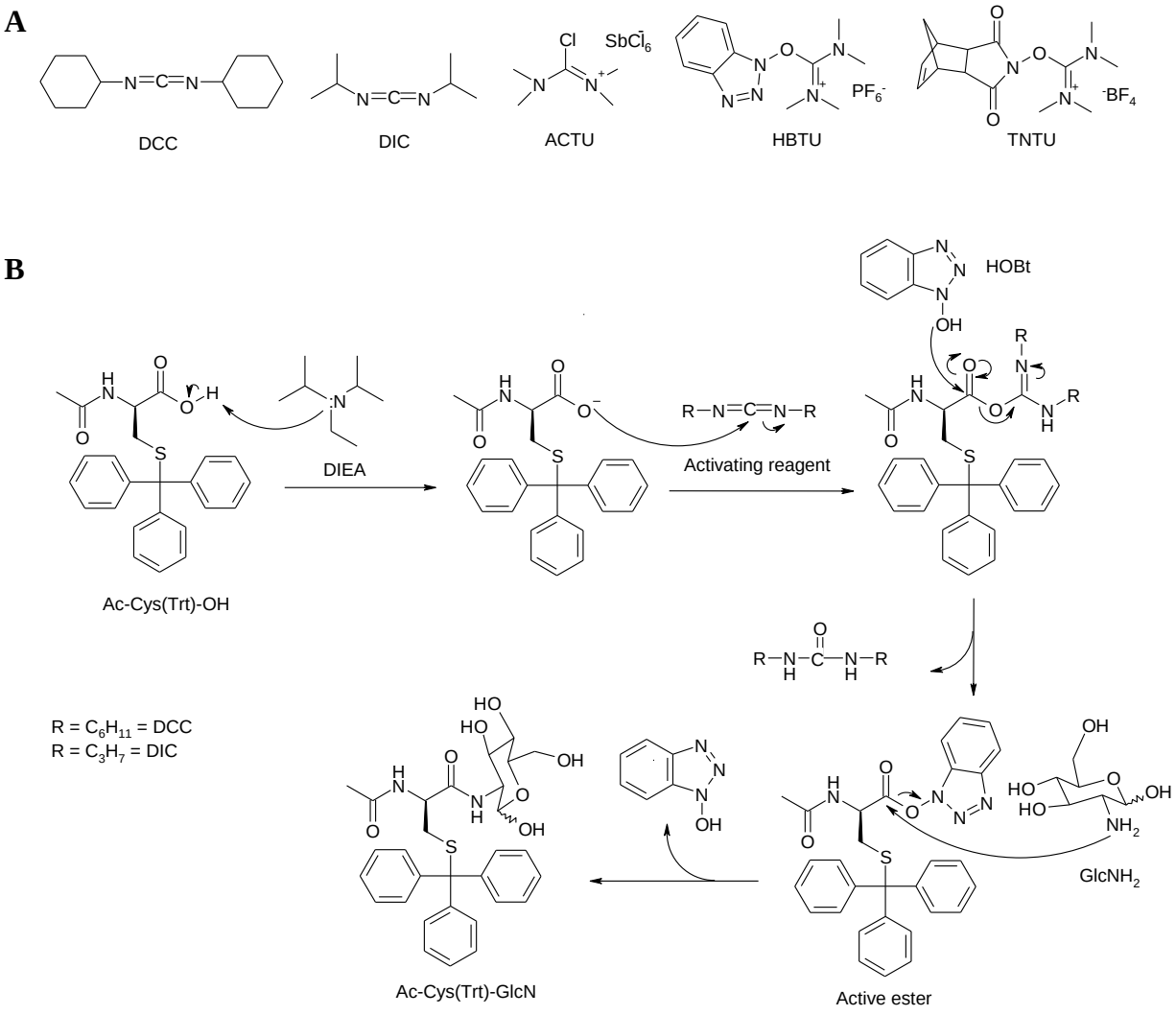


Figure 7.33: (A) Various activating reagents used in tMSH synthesis by the peptide synthesis assembly approach and **(B)** the example of the reaction mechanism using DCC and DIC as activating reagents. DCC, *N*,*N*′-dicyclohexylcarbodiimide; DIC, *N*,*N*′-diisopropylcarbodiimide; ACTU, 2-chloro-1,1,3,3-tetramethyluronium hexachloroantimonate; HBTU, *O*-(benzotriazol-1-yl)-*N,N,N′,N′*-tetramethyluronium hexafluorophosphate; TNTU, 2-(5-norbornene-2,3-dicarboximido)-1,1,3,3-tetramethyluronium tetrafluoroborate.

Purification of compound **7-18** (Ac-Cys(Trt)-GlcN) from the crude mixture was a major concern using this protocol. Due to the polarity of the product that required purification by reverse phase chromatography, two methods were utilized to purify compound **7-18** from the crude mixture: reverse-phase C18 HPLC and reverse-phase Redisep C18-based chromatography (43 g) using a BioFlash system. The advantages of the HPLC method included the ability to detect the product (**7-18**), which would allow ready collection at its known retention time (t_R~13 min) and the solvent gradient (40–100% acetonitrile in water

containing 0.1% TFA) could be easily set using the HPLC system. However, only limited amounts of product could be purified each HPLC run time, typically only a few milligrams per injection to avoid overloading the HPLC column. The Redisep C18 utilizing the BioFlash system had the advantage of being able to purify larger amounts (gram scale) at once; however, it lacked a UV detection system. The product could be monitored by TLC and visualized by ninhydrin staining or detected by absorption at 220 nm after fractions were collected. Thus, the prefered method which required less repetition and time as well as being capable of large scale purification was the Redisep C18 chromatography method utilizing a gradient of 20–100% acetonitrile in water containing 0.1% TFA (10% acetonitrile increment). The pure product **7-18** was collected at a gradient of ~40% acetonitrile, which was detected as a single peak in the reverse phase HPLC chromatogram.

The solubility of GlcN-HCl was also a major issue during the synthesis. Since the HOBt-ester was very reactive towards nucleophiles and was quite unstable in aqueous media, it was expected to react with water in addition to the amine, causing a loss of product formation and unreacted compound **7-17**. Thus, performing the coupling step in DMF should prevent the loss of the HOBt active ester but forces a compromise with the poor solubility of D-glucosamine in the organic solvent (DMF) even upon addition of base (DIEA) and the application of high temperature (45–50 °C). Neutralization of the amine hydrochloride by 1 N NaOH and freeze-drying to give the solid free-based glucosamine **7-7** was performed in an attempt to increase the solubility. An alternative method was to change the solvent system in the coupling reaction to 1:1 of 0.25 M imidazole HCl buffer (pH 7) in acetonitrile.

The reaction using DCC/HOBt as an activating reagent gave a white precipitate that formed upon the addition of these reagents to the reaction mixture in approximately 2–5 min. This precipitation indicated the formation of insoluble co-product *N,N'*-dicyclohexylurea (DCU), which could be conveniently removed by filtration. Longer activation times were evaluated (from 15 min up to 3 hours) due to the lower reactivity of the carbodiimides (DCC and DIC) in relation to the uronium activation reagents (ACTU, HBTU and TNTU). However, analysis of the ^{1}H NMR indicated the presence of impurities even after HPLC purification, and thus required repurification. Purification on a larger scale (RediSep C18 column using a gradient of 40–100% acetonitrile in water containing 0.1% TFA) was not able to obtain a significant amount of product and yet still required further purification by reverse phase HPLC.

It was also found that a longer activation time did not significantly affect the recovery yield of the compound **7-18** from the crude mixture as the same purification issues were encountered.

The alternative carbodiimide reagent, DIC, was used in order to avoid purification issues associated with DCU. The corresponding urea by-product, *N,N'*-diisopropyl urea (DIU), is soluble in the reaction solvent, thus eliminating product loss due to the filtration step as in the DCC activating reaction. However, similar purification issues were encountered and the trace amount of product recovered from the reverse phase HPLC required further purification.

The peptide coupling reaction utilizing the reagent ACTU/HOBt showed no product formation. The HPLC analysis of the crude product showed no peak at t_R 13 min corresponding to the product **7-18** and a large peak at t_R 17 min corresponding to compound **7-17**, suggesting that either the activated carboxylic acid reverted back to the acid or did not become activated at all. This reagent might not be reactive enough for this coupling reaction or might too sensitive to moisture.

The coupling reagent HBTU was evaluated in both aqueous and organic solvents. Its side product, tetramethyl urea (TMU), was solubilized in both water and organic solvent. However, only undertaking the coupling reaction in DMF gave the desired product **7-18**. HOBt ester of this reagent might be too reactive toward water such that compound **7-17** was reformed, causing loss of product. Impurity of this reaction was still found with the purification by Redisep C18 chromatography with a gradient of 40–100% acetonitrile in water containing 0.1% TFA.

The reaction utilizing the coupling reagent TNTU gave the most efficient results with a 25% yield. The active ester derivative is a norbornene-active ester, which was more stable in aqueous media and so was more suitable for acylation of GlcN in a mixed aqueous/organic medium. The reaction with TNTU indicated that the pure product **7-18** was isolated. The elution gradient of the RediSep C18 system was started at 20% acetonitrile in water containing 0.1% TFA to eliminate unreacted sugar, while the pure product was eluted with 40% acetonitrile. The HPLC chromatogram showed only one peak of the pure product with t_R 13 min (Figure 7.34). This reagent was successfully employed on the synthesis of tMSH previously [253]. The ^{1}H NMR (Spectrum 7.13, Appendix 7) and ESI-MS (Figure A6.12, Appendix 6) analyses suggested purified product **7-18**.

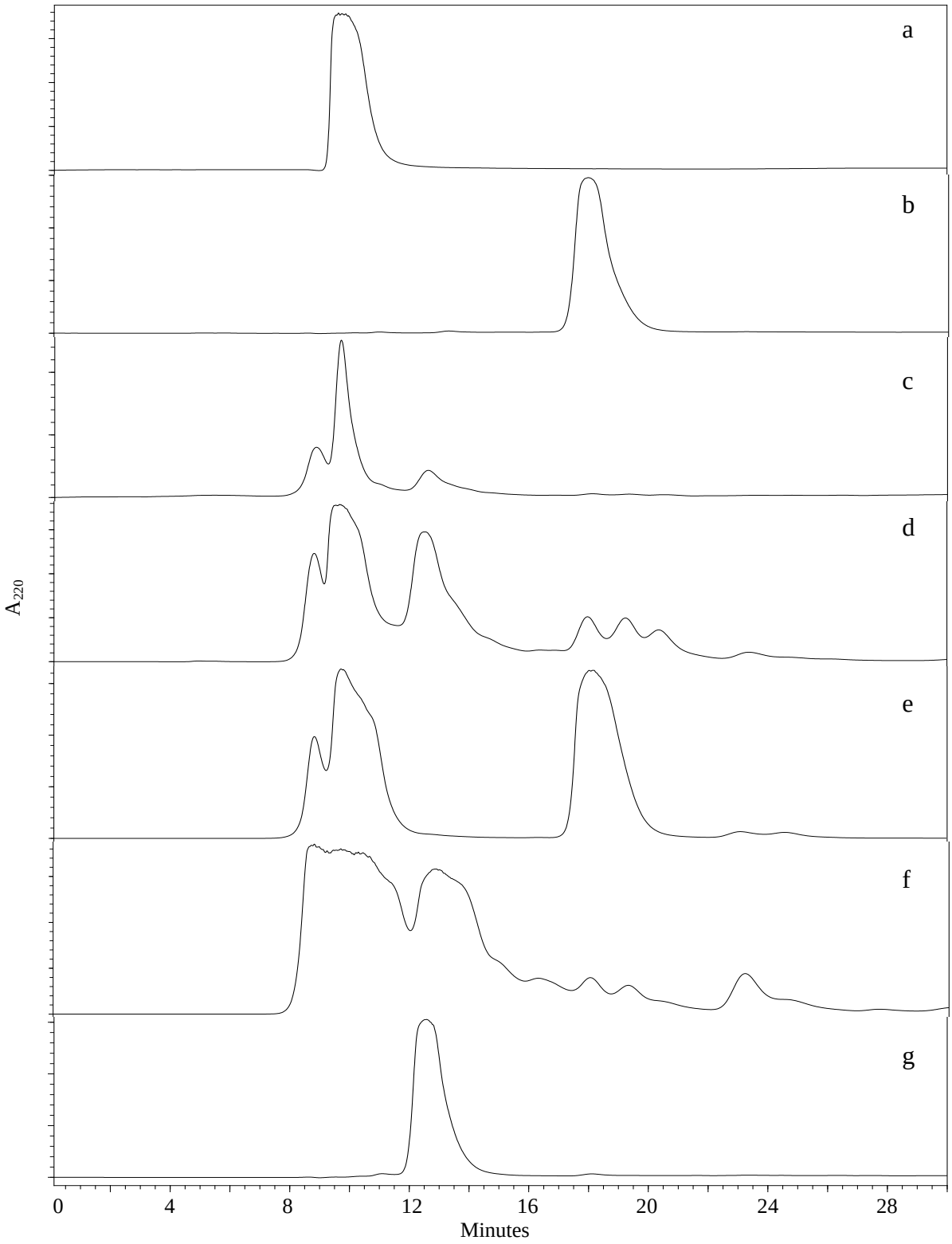


Figure 7.34: The reverse phase HPLC (C18) chromatograms of (a) HOBt and (b) starting material *N*-acetyl-*S*-trityl-L-cysteinyl-2-amino-2-deoxy-α,β-D-glucopyranoside (Ac-Cys(Trt)-GlcN, **7-18**) in comparison with the crude product from the reactions of *N*-acetyl-*S*-trityl-L-cysteinyl-2-amino-2-deoxy-α,β-D-glucopyranoside (Ac-Cys(Trt)-GlcN, **7-18**) synthesis with activating reagent (c) DCC, (d) DIC, (e) ACTU, (f) HBTU and (g) TNTU.

The reactions with DCC, DIC, ACTU and HBTU were performed in DMF with 15 min activation time and 3 hours coupling time. The reaction using TNTU was performed in imidazole/ACN with 15 min activation time and 1 hour coupling time. The compound **7-18** was in the peak with retention time at 13 min.

The cleavage of the Trt protecting group in the compound **7-18** was easily performed using TFA and DTT. The free thiol compound **7-2** was quantitated by the colorimetric assay using 5, 5'-dithiobis-(2-nitrobenzoic acid) (DTNB). The basic solution is required since DTNB reacts with thiolate anion, exchanges thiol-disulfide and generates 2-nitro-5-thiobenzoic acid (NTB), which can be monitored at 412 nm. By using the Beer-Lambert's Law, the concentration of tMSH was estimated. The compound **7-2** was confirmed by analysis of the ESI-MS (Figure A6.13, Appendix 6) and ^{1}H NMR (Spectrum 7.14, Appendix 7). Neither ESI-MS nor ^{1}H NMR experiements suggested the remaining free DTT in the final product.

7.5: Conclusion

The isolation of mycothiol (**7-1**) from *S. jumonjinesis* cells using a tagged disulfide of 2-*S*-(2'-thiopyridyl)-6-hydroxynaphthyldisulfide (**7-5**) only gave trace amounts of isolated product with contaminating reducing reagent (TCEP). An alternative reducing agent is immobilized TCEP, which should be conveniently removed upon filtration. However, due to low production of MSH in *S. jumonjinesis* and product loss during purification, this experiment was halted.

Numerous approaches to synthesize the truncated mycothiol, des-*myo*-inositol mycothiol (**7-2**), were attempted including synthesis by Patel and Blanchard's Method, *via* a cyclic thiolactone intermediate, using the resin-bound active ester reaction and using a peptide synthesis assembly (Table 7.9). The synthesis by Patel and Blanchard's Method involved several steps of protecting and deprotecting processes, which eventually produced only trace amounts of product being recovered. The most difficult aspect was to remove the acetamidomethyl (Acm) protecting group. Several deprotection reagents were chosen including thallium(III), iodine and disulfide 2,2'-dithiobis(5-nitropyridine) (DTNP). DTNP was the most efficient reagent for the removal of the Acm group; however, only trace amounts of the product were recovered. The synthesis of compound **7-2** *via* the cyclic thiolactone intermediate was unsuccessful. It was suspected that the conditions were not able to prepare

the highly strained thiolactone in any useful quantity in our hands in spite of what was stated in the literature. Neither was the product successfully synthesized from the synthetic approach using the resin-bound active ester reaction with two activating reagents (ACTU and DIC). The failure in coupling reaction with glucosamine might be the solubility of the sugar in DMF. The successful synthesis of the product was achieved by the peptide synthesis assembly method using TNTU as activating reagent in an aqueous phase. The recovery yield was increased from 25% to 40% when applying longer activating and coupling times. The cleavage of the Trt protecting group was easier than that of removal of the Acm protecting group. Acid hydrolysis and DTT was used to remove the Trt protecting group from compound **7-18**, yielding the free thiol compound **7-2** that would be used in the enzymatic assay of putative mycoGlyoxalase I and mycoGlyoxalase II (Chapter 8 and 9).

Table 7.9: The summary for the synthesis of des-*myo*-inositol mycothiol

Patel and Blanchard's Method[a]	Cyclic Thiolactone Intermediate	Resin-bound Active Ester Reaction	Peptide Synthesis Assembly[b]
Fmoc-Cys(Acm)-OPfP ↓ 73% Fmoc-Cys(Acm)-GlcN ↓ 99% (72%)[c] NH_2-Cys(Acm)-GlcN ↓ 20% (14%)[c] Ac-Cys(Acm)-GlcN ↓ Trace Ac-Cys-GlcN	*N*-AcCys ↓ *N*-AcCys thiolactone ↓ NR Ac-Cys-GlcN	Ac-Cys(Trt)-OH ↓ PS-OBt-Ac-Cys(Trt)-OH ↓ NR Ac-Cys-GlcN	Ac-Cys(Trt)-OH ↓ 25% Ac-Cys(Trt)-GlcN ↓ 82% (21%)[c] Ac-Cys-GlcN

[a] The last step of the synthesis used DTNP to remove acetamidomethyl group.
[b] The last step of the synthesis used TNTU in aqueous phase.
[c] Relative % yield from the first step reaction.

CHAPTER 8: PUTATIVE MYCOGLYOXALASE I FROM *STREPTOMYCES COELICOLOR*

8.1: Introduction

Gram-positive *Mycobacterium tuberculosis* is the causative agent of human tuberculosis (TB), which is a major health problem that is responsible for the deaths of approximately 1.6 million people annually [104, 269]. Due to developments in TB control including better diagnostics, antibiotics and vaccination, TB has been decreasing in numbers since the 19th century [104, 269]. However, this mycobacterial infection has emerged again with multidrug resistance in association with the spread of HIV/AIDS, causing one third of the world's population to be infected [104, 269]. Individuals with pulmonary TB infection produce airborne droplets that can spread to others through talking, coughing and sneezing. An infected person would rely on his/her self-immune response to fight the pathogen's replication. However, once the immune system cannot suppress the bacterial replication, the primary infection of TB would then become active. Healthy individuals normally develop inactive TB and can only be carriers or have the latent form of the infection. However, infected people with a weak immune system, such as HIV patients, have a greater chance to develop life-threatening symptoms.

Several antibiotics such as isoniazid, rifampicin, ethambutol and pyrazinamide are recommended for TB treatment [269-271]. These drugs are normally effective toward different targeted proteins in mycobacterial cells but are ineffective toward host cells. Rifampicin, for example, is an inhibitor of bacterial DNA-dependent RNA polymerase that prevents RNA transcription and subsequently protein translation, thus slowing down the replication rate of the *Mycobacterium* [272, 273]. This drug, however, creates complicated drug-drug interactions in patients with co-infections of TB and HIV as well as being ineffective toward some *Mycobacterium* infections due to the rapid development of bacterial drug resistance [272, 273]. Thus, an investigation on the combination of rifampicin with other antibiotics is necessary to avoid drug-drug interactions due to bacterial multidrug resistance. As well, reduction of treatment time is one of several main factors in the future of drug development.

While emphasis on chemical synthesis for drug improvement against tuberculosis has been increasing, other groups are interested in the structural genomics of *M. tuberculosis*, which include some proteins that play significant roles within the *Mycobacterium* such as urease, malate synthase and fumarase [274]. Protein structural investigations are important contributions to rational structure-based drug design. This approach is of interest to our laboratory, which relates to our previous studies on Glyoxalase enzymes. Our interest in the putative Glyoxalase (Glx) system in *M. tuberculosis* is described in this chapter, which focuses on a closely related organism, *Streptomyces coelicolor*. These microorganisms, instead of producing glutathione (GSH), mainly contain a low molecular weight mycothiol (MSH), in which 12 μmol MSH and <0.02 μmol GSH in 1 g residual dry weight in *M. tuberculosis* were found [83]. Since the Glyoxalase system using GSH as a cofactor has already been widely investigated, it is worthwhile to examine the parallel pathway of the putative mycothiol-dependent Glyoxalase or mycoGlyoxalase (mcGlx) system.

The putative enzymes in the mcGlx system, mycoGlyoxalase I (mcGlxI) and mycoGlyoxalase II (mcGlxII) are investigated from another Gram-positive bacterium, *Streptomyces coelicolor*. This Gram-positive bacterium possesses the genes for mycothiol biosynthesis (including *mshA, mshB, mshC* and *mshD*, more details in Chapter 1) as well as mycothiol catabolic enzymes (such as maleyl pyruvate isomerase) similar to mycobacteria [275], thus assuming that it possesses a similar mcGlx system. *S. coelicolor* cells produce higher levels of MSH than GSH (found 2.8 μmol MSH and <0.02 μmol GSH in 1 g residual dry weight) [83]. This bacterial strain was selected because it is non-pathogenic to humans and, thus, is easy to handle. Besides, its whole genome sequence has already been reported [276]. Streptomycetes are also known as a main source for antibiotic production such as the antibiotics cerulenin and rifampicin [277]. The self-defense mechanisms to avoid toxicity of these antibiotics produced within the bacterium involve MSH conjugation (thiol reactive agent) and elimination as mercapturic acid derivatives as well as regeneration of MSH, similar to the GSH detoxifying mechanisms (more details in Chapter 1) [98]. Mutagenesis studies on MSH-deficient *M. smegmatis* cells also indicated that these mutants are hypersentitive to various antibiotics and other toxic agents [278], suggesting that targeting MSH biosynthetic and metabolic pathways can allow design of novel antimycobacterial agents for effective drug

improvements against mycobacteria. Besides, targeting MSH metabolism would help limit toxicity of antituberculosis drugs because these enzymes are not found in the human host.

Sequence searching analysis for a putative mcGlxI from *S. coelicolor* on the StrepDB-The Streptomyces Annotation Server (http://strepdb.streptomyces.org.uk) revealed two genes with annotations of a putative dioxygenase (SCO1970) and a putative lyase (SCO2237) that exhibit high sequence similarity to the GlxI from other organisms. Analysis of the multiple sequence alignment of the putative dioxygenase (PDO) and the putative lyase (PLA) with other GlxI from various sources including Gram-negative bacteria (*E. coli*, *Y. pestis*, *N. meningitides*, *P. aeruginosa* and *P. putida*), protozoa (*L. donovani* and *T. cruzi*) and humans (*H. sapiens*) suggests that both proteins might function as GlxI and, thus, exhibit GlxI activity (Figure 8.1). The sequence identity and similarity between PDO and PLA are low (7% sequence identity and 16% similarity). However, these ranges of sequence identity and similarity are similar to those of the putative enzymes with other GlxI (3–9% identity and 12–19% similarity for PDO and 3–8% identity and 12–17% similarity for PLA). Since PDO employs four predicted conserved metal binding residues (His^{9}, Glu^{59}, His^{85} and Glu^{142}) and PLA only possesses one predicted metal binding residue (Glu^{6}, Asp^{58}, Gln^{76} and Glu^{128}, only Glu^{128} is conserved), PDO is more likely than PLA to exhibit GlxI activity. Additionally, since it was suggested previously that the metal activation classes (Zn^{2+}-activation and non-Zn^{2+}-activation) can be predicted by the length of the amino acid sequences (more details in Chapter 3), it is possible to suggest that PDO and PLA with the lack of the extended N-terminal arm and other three extra regions (helix B, loop C and loop D, Figure 8.1) might fall into the non Zn^{2+}-activated class and might be Ni^{2+}/Co^{2+}-activated.

To investigate the mcGlxI reaction, the putative enzymes (PDO and PLA) were cloned from a *S. coelicolor* plasmid into the pET-28b expression vector using PCR and various restriction enzymes. The vectors containing *pdo* and *pla* genes were overexpressed in *E. coli*. A two-step purification scheme using HisTrap HP and Benzamidine FF affinity columns was employed to isolate PDO and PLA from the endogenous *E. coli* proteins. The proteins' stabilities and their initial structures were examined using circular dichroism, fluorescence spectroscopy, ^{1}H NMR and gel permeation chromatography. The substrate formation, the hemithioacetal from the non-enzymatic reaction of methylglyoxal (MG) and truncated mycothiol (tMSH or des-*myo*-inositol mycothiol, details for synthesis in Chapter 7), for the

mcGlxI reaction was examined in terms of its equilibration time and dissociation constant in comparison with those of the reported GlxI reaction [85, 86, 141, 279-281]. The metal characterizations of the putative enzymes (metal analysis, metal activation and metal titration), pH profile, substrate specificity and kinetics were then investigated based on these preliminary data of the hemithioacetal formation. As well, the isolation and identification of the prospective product of the mcGlxI reaction, *S*-D-lactoyl-des-*myo*-inositol mycothiol, supported and confirmed the enzymatic activity of the putative enzymes. Additionally, due to the close structural relation between GlxI and methylmalonyl-CoA epimerase (MMCE, more details in Chapter 5), the MMCE activities were examined in these putative enzymes.

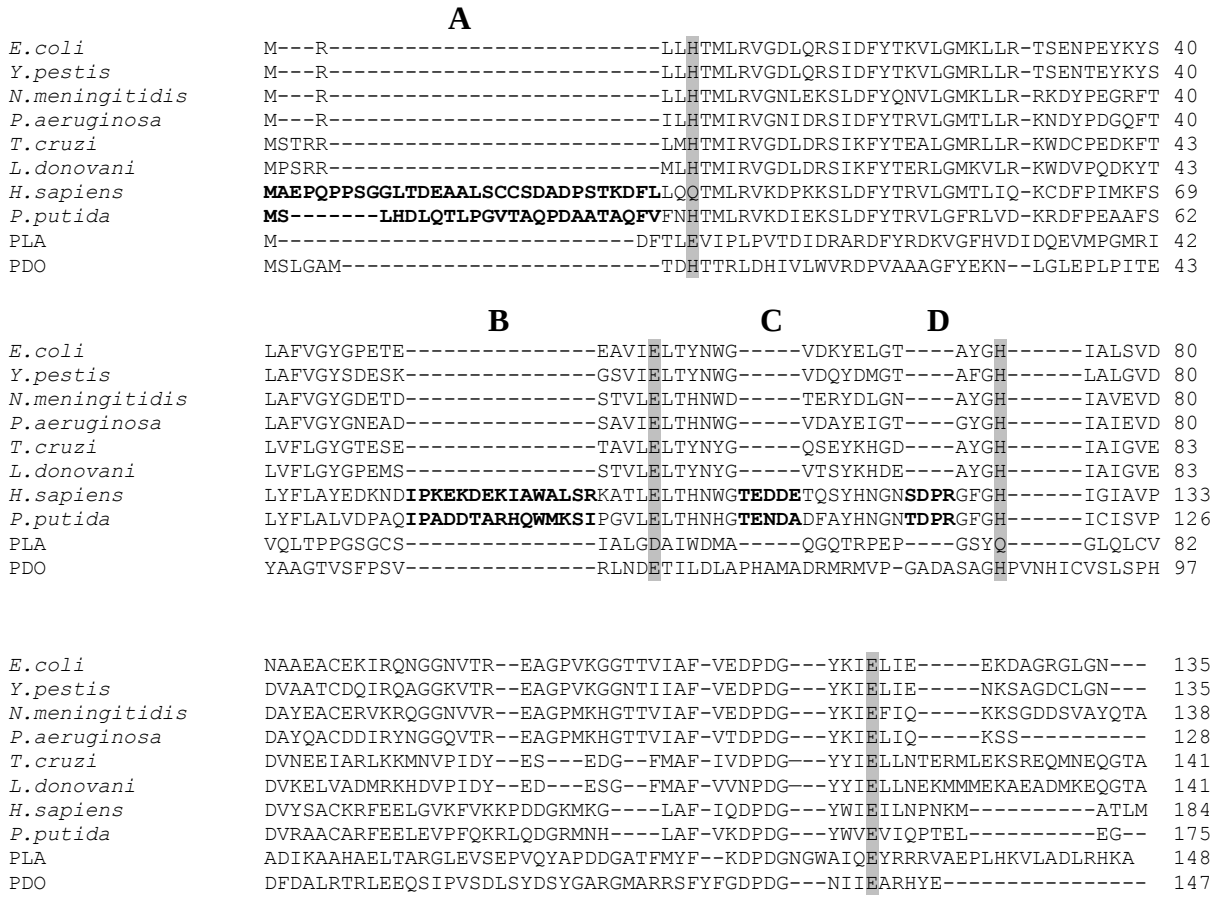


Figure 8.1: The multiple sequence alignment of the putative dioxygenase (PDO, CAC42744) and the putative lyase (PLA, CAC37263) from *Streptomyces coelicolor* with Glyoxalase I from other organisms (organism name follows by National Center for Biotechnology Information (NCBI) accession number) including *E. coli* (NP_310387), *Y. pestis* (ZP_01887743), *N. meningitides* (CAA74673), *P. aeruginosa* (AAG06912), *L. donovani* (AAU87880), *T. cruzi* (XP_818456), *H. sapiens* (AAB49495) and *P. putida* (AAN69360). The metal binding residues are highlighted in grey. The loop regions in the Zn^{2+}-activated GlxI are indicated in bold letters and

marked with bold capital letters (A, B, C and D). The alignment was created using CLC Free Workbench (version 3.0.1) with the accurate alignment algorithm (http://www.clcbio.com).

8.2: Reagents, Materials and Instrumentation

All reagents, materials and equipment used in this chapter are listed below otherwise they are included in the previous chapters. The designed primers for DNA cloning were received from Sigma-Genosys Canada (Oakville, ON). The purification column for gel permeation chromatography, Superose^{TM} 6 10/300 GL column, was obtained from GE Healthcare (Piscataway, NJ). Protein standards for gel permeation chromatography including γ-globulin, bovine serum albumin (BSA), ovalbumin, carbonic anhydrase, myoglobulin and vitamin B12 (molecular weights of 158, 66, 44, 29, 17 and 1.35 kDa, respectively) were obtained from Bio-Rad laboratories (Hercules, CA).

8.3: Experimental Protocols

8.3.1: DNA Cloning and Manipulation

All DNA manipulations and purifications were performed according to the protocols outlined by Sambrook and Russell [135]. The cloning of *pdo* and *pla* genes into the pET-28b(+) expression vector to generate a protein with the N-terminal His_6-tag followed by a thrombin protease cleavage site utilizing PCR and the restriction enzymes, *Nde*I and *BamH*I, was performed similarly to the cloning of *clo glxI* (Chapter 2) and is summarized in Appendix 1 (Figure A1.3). The forward and reverse primers were designed as in Table 8.1. PCR reagents and plasmids were prepared as in Table 8.2, and thermal cycles were programmed as in Table 8.3. The circular DNA was heat shock transformed into the competent *E. coli* DH5α cells followed by growth on agar plates containing antibiotic kanamycin (30 μg/mL) overnight at 37 ºC. A single colony was picked and inoculated overnight with LB (5 mL) containing Kan (30 μg/mL) at 37 ºC in a 220 rpm shaker. The plasmid was purified using a QIAprep Spin Miniprep Kit and sent for sequencing when appropriate. The pET-28b-*pdo* and pET-28b-*pla* plasmids were heat shock transformed into the competent *E. coli* BL21 (DE3) cells for protein expression purposes.

Table 8.1: The forward and reverse primers for DNA cloning of putative dioxygenase and putative lyase

Gene	*Primer name*	*Sequence of primer (5′ to 3′)*
pdo	PDONdeI.For[†]	CCGAAGCTT**CATATG**AGCCTGGGAGCC[£]
	PDOBamHI.Rev[†]	GGCGAATTCGGATCCTACTCGTAGTGCCGG[¥]
pla	PLANdeI.For[€]	CCGAAGCTT**CATATG**GACTTCACGCTCG[£]
	PLABamHI.Rev[€]	GGCGAATTCGGATCCTAGGCCTTGTGCCGG[¥]

[†] The melting temperatures (T_m) for PDONdeI.For and PDOBamHI.Rev primer are 64 and 67 ºC, respectively.
[€] The T_m for PLANdeI.For and PLABamHI.Rev primers are 63 and 68 ºC, respectively.
[£] Bold letters indicate *Nde*I restriction sites and the underlined letters are coding for the N-terminus of the protein.
[¥] The grey highlights indicate *BamH*I restriction sites and the underlined letters are coding for the C-terminus of the protein.

Table 8.2: PCR recipe for DNA amplification of putative dioxygenase and putative lyase

Plasmids/Reagents	*Stock Concentration*	*Amount Used (μL)*	*Final Concentration*
DNA	5.4 ng/μL	4.0	20 ng
Primer.For	10 μM	3.0	600 nM
Primer.Rev	10 μM	3.0	600 nM
dNTP	10 mM	1.0	0.2 mM
PWO buffer	10×	5.0	
Mg^{2+}	25 mM	*a*	*b*
DMSO	100 %	2.5	5 %
ddH_2O		*c*	
PWO polymerase		0.5	
Total		50.0	

a is varied as 1, 2, or 3 μL (corresponded to *b* values, which are 0.5 mM, 1.0 mM and 1.5 mM, respectively).
b is amount of Mg^{2+} used to maximize PCR products.
c is amount of distilled water (ddH_2O) used to fulfill the final PCR volume of 50 μL.

Table 8.3: PCR programming set for DNA cloning of putative dioxygenase and putative lyase

Steps	*Temperature (ºC)*	*Time*	
Initial denaturation[†]	95	5 min	
Hot start[‡]	85		
Denaturation	95	30 sec	Repeat 30 cycles
Annealing	60	1 min	
Elongation	72	45 sec	
Final extension	72	2 min	
Hold	4		

[†] Initial denaturing temperature is used to ensure DNA template denaturation.
[‡] Hot start is used to maintain the temperature at 85 ºC during addition of PWO DNA polymerase to minimize self annealing of primers.

8.3.2: Protein Expression, Induction and Purification

Various conditions including growth and induction temperatures as well as amount of IPTG were investigated in order to optimize protein expression (Table 8.4). The *E. coli* BL21 (DE3) containing the desired plasmid was inoculated in LB (5 mL) containing Kan (30 μg/mL LB) overnight at 37 ºC in an incubator shaker (shaken at 220 rpm). The overnight culture (1 mL) was transferred into 50 mL LB containing Kan (30 μg/mL LB) and was shaken until OD_{600} reached 0.6 at the particular temperature. The proteins were then induced by the addition of IPTG followed by continued shaking at the same speed. An aliquot (1 mL) was taken every hour for 4 hours. The cell pellets from these aliquots were collected by centrifugation at 13,000×*g* for 1 minute before performing SDS-PAGE experiments. The rest of the culture was centrifuged at 14,031×*g* for 10 minutes and pellets were resuspended in a minimum volume of Tris buffer (50 mM Tris (pH 8.0) and 150 mM KCl). The suspension was disrupted by sonication (15 × 10 sec) before centrifugation at 14,031×*g* for another 10 minutes. The proteins in the cell pellet and the supernatant were visualized by Coomassie-staining SDS-PAGE gel.

Table 8.4: The conditions for optimization of protein expression including growth and induction temperatures as well as concentration of IPTG

Growth Temperature (°C)	*Induction Temperature (°C)*	*IPTG (mM)*
25	25	0.5
25	25	1.0
37	25	0.5
37	25	1.0
37	37	0.5
37	37	1.0

The protein induction and expression were performed according to the previous investigations on growth and induction temperatures, amount of IPTG and expression time (optimized conditions were growth at 25 ºC and expression with 0.5 mM IPTG at 25 °C for four hours). The protein purification was performed using a HisTrap HP affinity column (1 mL) and HiTrap Benzamidine FF affinity column (1 mL). Metal analysis by ICP-MS and apo-enzyme preparation was performed as previously described in Chapter 2. The existence of the desired protein was confirmed and the molecular weight was estimated by the analysis of SDS-PAGE and ESI-MS experiments.

8.3.3: Protein Stability and Initial Structural Investigation

The secondary structural investigation of the proteins was determined by circular dichroism (CD) as described previously in Chapter 2. The quaternary structures of the proteins were examined by gel permeation chromatography using a Superose6 10/300 GL column and a buffer containing 50 mM Tris (pH 8.0) and 150 mM KCl with a flow rate of 0.5 mL/min. The molecular weight was estimated using the Bio-Rad standard proteins including γ-globulin (158 kDa), bovine serum albumin (BSA) (66 kDa), ovalbumin (44 kDa), carbonic anhydrase (29 kDa), myoglobulin (17 kDa) and vitamin B_{12} (1.35 kDa) (Appendix 3).

The stability of the proteins and initial structural investigations under specified conditions was examined by fluorescence spectroscopy and ^{1}H NMR. The fluorescence spectra were scanned using a low volume quartz cuvette (45 μL), a step size of 1 nm, an integration of 1 sec, an incoming slit of 1 nm and an outgoing slit of 5 nm with a fixed excitation at 289 nm and a range of emission wavelength between 300–500 nm. ^{1}H NMR (600 MHz) was performed using purified proteins in 50 mM Tris (pH 8.0), 150 mM KCl and 10% (v/v) glycerol. Additionally, the determination of denaturation temperature was also investigated by differential scanning calorimetric (DSC) as described previously in Chapter 2.

8.3.4: Glyoxalase I Enzymatic Assay

Determination of Hemithioacetal Equilibrium Time

In order to investigate the equilibrium time for the formation of hemithioacetal (GlxI substrate), the enzymatic assay as well as ^{1}H NMR were performed under various incubation times of MG-tMSH. MG-tMSH (using 16 mM MG and 0.6 mM tMSH or 0.5 mM MG-tMSH initially assuming that the dissociation constant (K_d) would be similar to that of MG-GSH, which is 3.1 mM [85, 86, 141, 279-281]) was incubated for various time periods over 0–3 hours in 50 mM KPB (pH 6.6) at room temperature. The assay was then performed with Ni^{2+}-reconstituted PDO (4.5 μg in 300 μL assay) on a 96-well UV plate using a SpectraMax spectrophotometer. The formation of mcGlxI product (*S*-D-lactoyl-des-*myo*-inositol mycothiol) was indicated by an increase in absorption at 240 nm assuming that this product has the same absorption as *S*-D-lactoylglutathione (the product of GlxI reaction using GSH as cofactor). Additionally, the hemithioacetal of MG-tMSH (7 mM, K_d = 3.1 mM) in D_2O was

incubated at room temperature for different time periods, and the equilibrium time was analyzed using ^{1}H NMR.

Analysis of the formation of the hemithioacetal of MG-GSH was performed in parallel as a control. The enzymatic assay was performed using the commercial yeast GlxI and MG-GSH (0.5 mM) in 50 mM KPB (pH 6.6) at room temperature. The ^{1}H NMR was also performed on MG-GSH (7 mM) in D_2O.

Determination of the Hemithioacetal Dissociation Constant

The dissociation constant of MG-tMSH has been previously derived for thiols reacting with MG [85, 86, 141, 279-281] and can be calculated from the equation

$$\Delta A_{240} = C[MG]_t/(K_d + [MG]_t)$$

where ΔA_{240} is the net increase in absorbance at 240 nm after the reaction reaches its equilibrium and is read against a blank containing thiol, C is the total concentration of thiol multiplied by the collection of extinction coefficients of MG, tMSH and MG-tMSH (($\varepsilon_{MG\text{-}tMSH} - \varepsilon_{MG} - \varepsilon_{tMSH}$)$[tMSH]_t$), $[MG]_t$ is the total concentration of MG and K_d is dissociation constant (see detailed derivation in Appendix 5). The dissociation constant can be determined from the plot of the change in A_{240} as a function of the concentration of MG.

The determination of the dissociation constant of MG-tMSH was performed using various concentrations of MG (0.5–100 mM), which was incubated with a fixed concentration of tMSH (0.2 mM) in 50 mM KPB (pH 6.6) for 30 min at room temperature (1 mL total). The increase in absorbance at 240 nm after the reaction reached its equilibrium (30 min, data from the previous experiment on hemithioacetal equilibrium time) was detected using a SpectraMax spectrophotometer. The investigation on the dissociation constant of MG-GSH under the same conditions was performed in parallel as a control. However, in this case MG-GSH was incubated for 15 minutes to allow the reaction to reach its equilibrium.

Maximum Absorbance of Glyoxalase I Product

The optimum absorbance for the monitoring of the mcGlxI reaction was considered using Ni^{2+}-reconstituted PDO (4.5 μg in 300 μL assay) and MG-tMSH (0.5 mM) in 50 mM KPB (pH 6.6). The substrate, MG-tMSH (K_d of 3.3 mM, data from previous experiment) was incubated for 30 minutes before mixing with the enzyme. The reactions were incubated at

different time periods and the absorbance between 230–300 nm was scanned. The maximum absorbance to detect the formation of the product of the mcGlxI reaction (*S*-D-lactoyl-des-*myo*-inositol mycothiol) was determined.

Control experiments were performed using the commercial yeast GlxI (0.15 μg in 300 μL assay) with MG-GSH (0.1 mM, K_d of 3.1 mM) in 50 mM KPB (pH 6.6). MG-GSH was incubated for 15 minutes at room temperature before performing the enzymatic assay. Since the optimum absorbance for the product of the GlxI reaction, *S*-D-lactoylglutathione, was already reported [85, 279, 281], this experiment was repeated to confirm the accuracy of our protocols.

Detection of Glyoxalase I Product

The identification of the product from the mcGlxI reaction, *S*-D-lactoyl-des-*myo*-inositol mycothiol, was performed to confirm the enzymatic function of the putative enzymes. Apo-PDO (25 μg in 600 μL assay) was incubated with $NiCl_2$ (5 equivalents) overnight at 4 °C. The assay was performed using the substrate MG-tMSH (5 mM, K_d of 3.3 mM) in 50 mM KPB (pH 6.6) that had been incubated at room temperature for 30 minutes prior to assaying. The reaction was allowed to remain at room temperature for one hour. The product of the reaction was isolated by HPLC using a Waters μBandapak C18 reverse phase radial compression column (25 × 100 nm) and a flow rate of 1 mL/min with 1% increment of acetonitrile in water containing 0.1% TFA over a 100 minute interval. The product was monitored at 280 nm and was identified by ^{1}H NMR and ESI-MS.

8.3.5: Specificity of mycoGlyoxalase I Activity

The specificity of the mcGlxI reaction was investigated in terms of substrates, metals and enzymes. The substrates were prepared as MG, GSH, tMSH, hemithioacetal of MG-GSH (0.5 mM, 15 minute incubation with K_d of 3.1 mM) and hemithioacetal of MG-tMSH (0.5 mM, 30 minute incubation with K_d of 3.3 mM). The Ni^{2+}-substrate complexes were also prepared in parallel as controls by adding 5 equivalents of metal (relative to concentration of PDO in the enzymatic assay) to the substrate in order to examine the effect of metal on the reaction without addition of the enzyme. The enzymes (1.5 μg in 200 μL assay) being examined were native PDO, Ni^{2+}-reconstituted PDO, apo-PDO, denatured PDO (boiled for 10 minutes) with

the substrate prepared as describe above. The assays using the commercial yeast GlxI (0.05 μg in 200 μL assay) were performed in parallel. The activity was monitored at 240 nm using a 96-well plate on a plate reader.

8.3.6: Metal Characterization

Metal characterizations including metal analysis, metal activation, metal titration and pH profile were performed as previously described in Chapter 2. Metal activation was investigated using apo-PDO (3.1 μg in 200 μL assay) and apo-PLA (3.1 μg in 200 μL assay) in 50 mM Tris (pH 8.0) and 150 mM KCl that were incubated with various divalent metals (5 equivalents of $ZnCl_2$, $NiCl_2$, $CoCl_2$, $CuCl_2$, $MnCl_2$, $MgCl_2$, $CaCl_2$ and $CdCl_2$) overnight at 4 °C prior to performing the assay with MG-tMSH (0.5 mM, K_d of 3.3 mM with 30 minute incubation) in 50 mM KPB (pH 6.6) at room temperature. For metal titration, various concentrations of Ni^{2+} and Cu^{2+} were incubated with apo-PDO (3.1 μg in 200 μL assay) overnight at 4 °C before performing the assay under the same conditions as previously described. The investigation of the optimal pH was performed on the Ni^{2+}-reconstituted PDO (5 equivalents of Ni^{2+}, 3.1 μg enzyme in 200 μL assay) and MG-tMSH (0.5 mM, K_d of 3.3 mM with 30 minute incubation) in buffer at various pH (pH 5.8–8). The activities of all experiments were measured at 240 nm using a 96-well plate on the plate reader.

8.3.7: Kinetic Studies of Glyoxalase I Function

Enzyme kinetics were evaluated using Ni^{2+}-reconstituted PDO (30 ng in 1 mL assay) with MG-tMSH (0.04–1 mM, K_d of 3.3 mM with 30 minute incubation) in 50 mM KPB (pH 6.6) at room temperature, assuming that the concentration of tMSH is required to be maintained at 0.1 mM in order to avoid enzyme inhibition (similar to that of GSH). The activity was measured as an increase in absorbance at 240 nm (ε_{240} = 2860 $M^{-1}cm^{-1}$) on a SpectraMax spectrophotometer. The initial rate was fitted by the Michaelis-Menten equation with least squares fit parameters using GraphPad Prism software version 5.00 (GraphPad Software, Inc., La Jolla, CA).

8.3.8: Methylmalonyl-CoA Epimerase Enzymatic Assay

The assay was performed as described previously in Chapter 5. Apo-PDO and apo-PLA in 20 mM Tris (pH 7.0) and 150 mM NaCl were incubated with 5 equivalents of Ni^{2+}, Co^{2+}, Cu^{2+} and Zn^{2+} overnight at 4 °C to a final concentration of 1 mg/mL. The coupled assay of methylmalonyl-CoA mutase (supplied by Dr. Thomas Bobik from the Iowa State University, IA) and metal-reconstituted PDO and PLA (1 μg in 500 μL assay) was performed using the commercial (*RS*)-methylmalonyl-CoA as a substrate on a reverse phase HPLC instrument with a μBondapak^{TM} C18 column (3.9 × 300 mm) (Waters, Milford, MA).

8.4: Results and Discussion

8.4.1: Amino Acid Sequence Analysis

The results from sequence searching for the putative mycoGlyoxalase I from *Streptomyces coelicolor* A3(2) genome (NC_003888) suggested two related sequences, a putative dioxygenase (PDO) and a putative lyase (PLA). The nucleic acid coding for PDO (CAC42744) is located in the region between 2109895 and 2110338 on the minus strand. This putative protein contains 444 nucleic acids that codes for 147 amino acids (Figure 8.2). The PCR product for DNA cloning of *pdo* gene into the pET-28b(+) expression vector was optimized by the addition of 0.5 mM $MgCl_2$ and 5% (v/v) DMSO (Figure 8.3). This cloning procedure generated a plasmid coding for PDO with the N-terminal His_6-tag followed by a thrombin protease cleavage site. This His-tag was used for protein purification purposes. The DNA cloning and the cleavage by thrombin protease to remove the His-tag after purification would add the extra three amino acids at the N-terminus (Gly-Ser-His) of PDO as well as replace the first Val residue with Met. These modifications changed the predicted molecular weight of PDO from 16256.2 Da to 16569.5 Da and the predicted p*I* from 5.25 to 5.38.

```
gtgagcctgggagccatgacggaccacacgacacgtctcgaccacatcgtcctctgggtg
 V  S  L  G  A  M  T  D  H  T  T  R  L  D  H  I  V  L  W  V
cgcgacccggttgccgcggccggcttctacgagaagaacctgggcctcgagcccctgcgg
 R  D  P  V  A  A  A  G  F  Y  E  K  N  L  G  L  E  P  L  R
atcaccgagtacgccgcggggaccgtgagtttcccctccgtgcgcctcaacgacgagacc
 I  T  E  Y  A  A  G  T  V  S  F  P  S  V  R  L  N  D  E  T
atcctcgacctcgcaccgcacgccatggcggaccgcatgcgcatggtccccggcgccgac
 I  L  D  L  A  P  H  A  M  A  D  R  M  R  M  V  P  G  A  D
gccagtgccgggcatcccgtcaaccacatctgcgtgtccctgtccccgcacgacttcgac
 A  S  A  G  H  P  V  N  H  I  C  V  S  L  S  P  H  D  F  D
gccctgcgcactcgcctggaggagcagtccatcccggtctcggacctctcgtacgactcc
 A  L  R  T  R  L  E  E  Q  S  I  P  V  S  D  L  S  Y  D  S
tacggcgcccgcggcatggccaggcgcagcttctacttcggcgacccggacggcaacatc
 Y  G  A  R  G  M  A  R  R  S  F  Y  F  G  D  P  D  G  N  I
atcgaggcccggcactacgagtag
 I  E  A  R  H  Y  E  -

Amino acid composition:
16A   12R   4N    13D   1C    1Q    8E    10G   7H    7I
13L   1K    5M    5F    9P    11S   7T    1W    6Y    10V
```

Figure 8.2: The DNA sequence (444 nucleic acids), amino acid sequence (147 residues) and amino acid composition of the putative dioxygenase from *Streptomyces coelicolor*.

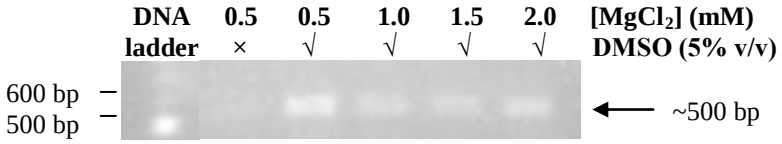


Figure 8.3: The agarose gel (1%) illustrates the PCR product of the PDO amplification using various concentrations of $MgCl_2$ with/without the presence of DMSO.

The gene coding for PLA (CAC37263) is located on the region between 2406519 and 2406965 on the plus strand. This protein consists of 148 amino acids that are coded from 447 nucleic acids (Figure 8.4). The cloning of *pla* gene into the pET-28b(+) expression vector was similar to that of the *pdo* cloning. The PCR product of *pla* amplification was optimized by the addition of 0.5 mM $MgCl_2$ and 5% (v/v) DMSO (Figure 8.5). This DNA cloning generated the His_6-tag and the thrombin protease cleavage site at the N-terminus of PLA. After thrombin protease cleavage, its predicted molecular weight was changed from 16468.6 Da to 16749.9 Da and the predicted p*I* from 4.86 to 4.98.

```
atggacttcacgctcgaagtcatcccgctccccgtcaccgacatcgaccgcgcccgggac
 M  D  F  T  L  E  V  I  P  L  P  V  T  D  I  D  R  A  R  D
ttctaccgcgacaaggtcggcttccacgtcgacatcgaccaggaggtgatgccgggcatg
 F  Y  R  D  K  V  G  F  H  V  D  I  D  Q  E  V  M  P  G  M
cggatcgtccagctgacgcctccgggttccggctgttcaatcgccctcggcgacgccatc
 R  I  V  Q  L  T  P  P  G  S  G  C  S  I  A  L  G  D  A  I
tgggacatggcgcaggggcagacccggccggaaccgggctcgtaccagggcctccagctc
 W  D  M  A  Q  G  Q  T  R  P  E  P  G  S  Y  Q  G  L  Q  L
tgcgtcgccgacatcaaggcggcccacgcggagctgaccgcccgtggtctggaggtctcg
 C  V  A  D  I  K  A  A  H  A  E  L  T  A  R  G  L  E  V  S
gagccggtgcagtacgccccggacgacggcgccaccttcatgtacttcaaggacccggac
 E  P  V  Q  Y  A  P  D  D  G  A  T  F  M  Y  F  K  D  P  D
ggcaacggctgggcgatccaggagtaccggcgccgggtggcggagccgctgcacaaggtg
 G  N  G  W  A  I  Q  E  Y  R  R  R  V  A  E  P  L  H  K  V
ctggccgacctccggcacaaggcctag
 L  A  D  L  R  H  K  A  -

Amino acid composition:
15A   10R   1N    15D   2C    8Q    8E    12G   4H    8I
11L   5K    5M    5F    11P   4S    6T    2W    5Y    11V
```

Figure 8.4: The DNA sequence (447 nucleic acids), amino acid sequence (148 residues) and amino acid composition of the putative lyase from *Streptomyces coelicolor*.

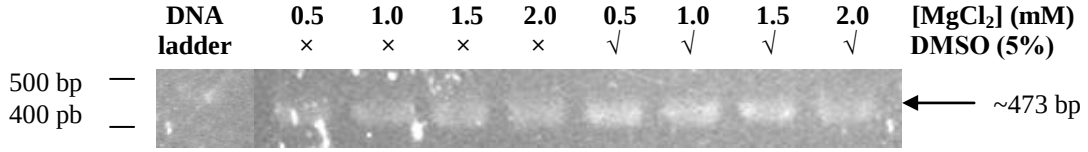


Figure 8.5: The agarose gel (1%) illustrates the PCR product of the putative lyase amplification using various concentrations of $MgCl_2$ with/without DMSO.

8.4.2: Protein Characterization

The optimal conditions for induction and expression of PDO were investigated in terms of growth and expression temperatures, amount of inducer (IPTG) and expression time. Analysis of the preliminary data suggested that optimal overexpression could be achieved under growth at 25 °C and expression at 37 °C with induction by 0.5 mM IPTG for 4 hours (Figure 8.6a-c). However, after cell disruption by sonication, more soluble proteins were observed under expression at 25 °C than at 37 °C (Figure 8.6d). Thus, the overexpression of PDO was performed at 25 °C.

Cell pellets of ~2.5 g were harvested from 1 L culture. The purification using a HisTrap HP affinity column could be used to isolate approximately 17.6 mg of His-tagged proteins (~22 mg/mL) from the cell lysate (Figure 8.7). The cleavage of these His-tagged proteins by thrombin protease was performed overnight at 4 °C. The following purification by

the combined HisTrap HP and HiTrap Benzamidine FF affinity columns recovered approximately 10 mg of cleaved protein (Figure 8.8). The purification of the purified PDO was monitored by SDS-PAGE experiment (Figure 8.9A) and its identity was confirmed by ESI-MS analysis (Figure 8.9B).

The molecular weight of purified PDO was determined using SDS-PAGE, ESI-MS and gel permeation chromatography. The SDS-PAGE estimated the molecular mass of a single subunit of PDO at ~20 kDa, while ESI-MS provided a more precise molecular mass of 16565.0 Da (calculated MW is 16569.5 Da). No post-translational modification of the purified protein was observed. Analysis of the results of a gel permeation chromatographic separation using a Superose6 10/300 GL column with a flow rate of 0.5 mL/min suggested a molecular weight of the protein under native conditions of approximately 27 kDa (Figure 8.10), which indicated a dimeric structure for PDO in 50 mM Tris (pH 8.0) and 150 mM KCl.

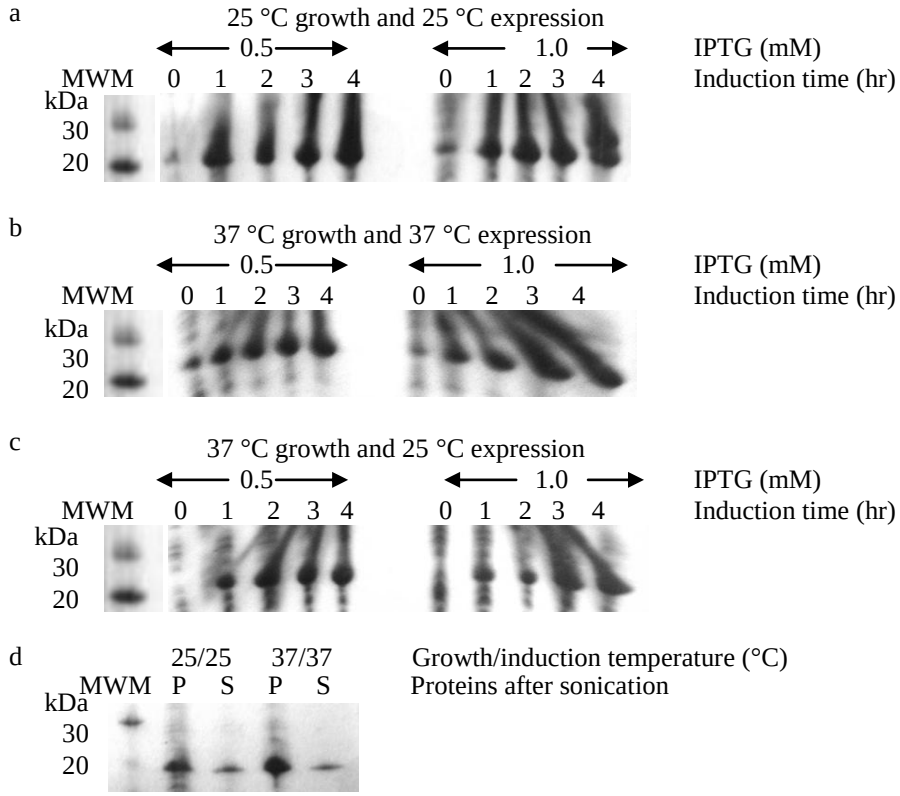


Figure 8.6: The SDS–PAGE of the optimized conditions for induction and expression tests of PDO including (a) 25 °C growth and induction, (b) 37 °C growth and induction and (c) 37 °C growth and 25 °C induction. Amount of IPTG and induction time were varied. (d) The SDS–PAGE of expressed proteins under various growth and induction temperatures with addition of 0.5 mM IPTG shows proteins in pellet (P) and supernatant (S) after sonication.

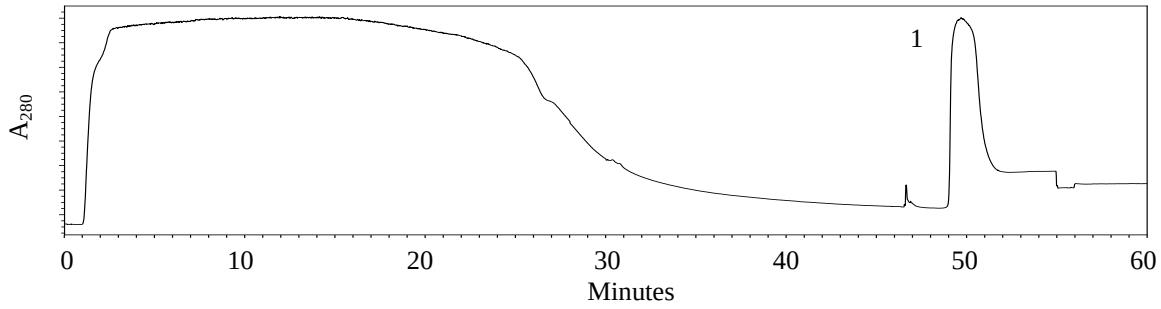


Figure 8.7: The HPLC chromatogram of PDO purification using a HisTrap HP affinity column with a flow rate of 0.5 mL/min. The lysate was loaded into the column and washed with 50 mM Tris (pH 8.0), 500 mM KCl and 20 mM imidazole (0–45 min). His-tagged proteins were then eluted using the same buffer but with 500 mM imidazole (peak 1). The proteins in the eluted fraction were monitored by SDS–PAGE.

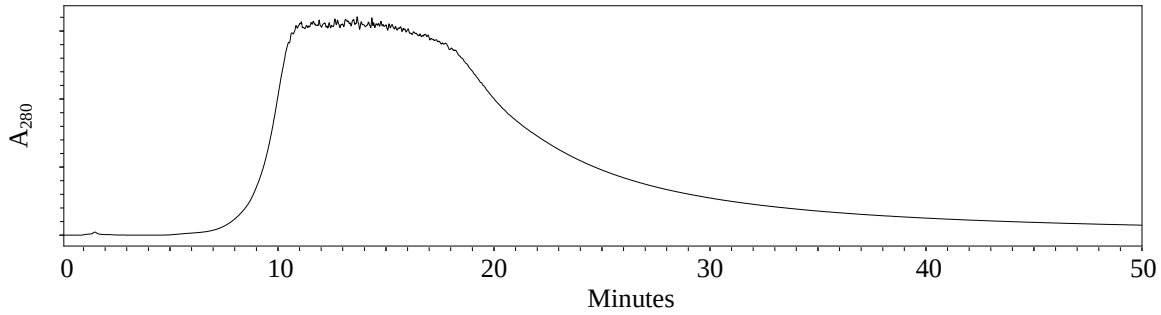


Figure 8.8: The HPLC chromatogram of PDO purification after treatment with thrombin protease using a HiTrap Benzamidine FF and HisTrap HP affinity columns with a flow rate of 0.5 mL/min and a buffer containing 50 mM Tris (pH 8.0), 500 mM KCl and 20 mM imidazole. The purified PDO (without His-tag) was eluted as shown as a broad peak over 7–30 min and was monitored by SDS–PAGE.

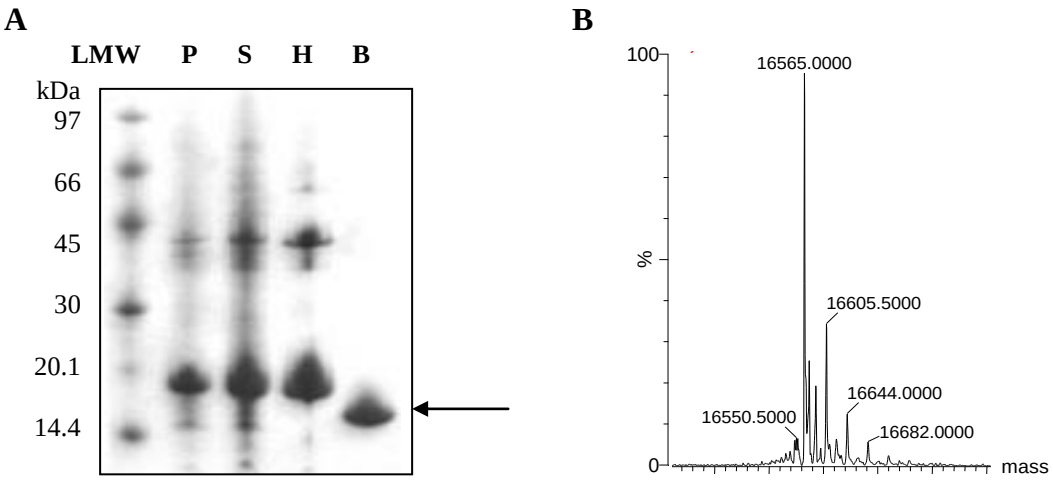


Figure 8.9: (A) The SDS–PAGE of the PDO purification: LMW = low molecular weight marker, P = pellet after sonication, S = supernatant after sonication, H = HisTrap eluted fraction and B = HisTrap and Benzamidine eluted fraction after thrombin protease treatment. The arrow indicates the presence of PDO. **(B)** The electrospray mass spectrum of PDO exhibits a molecular mass of 16565 Da (calculated MW is 16569.5 Da).

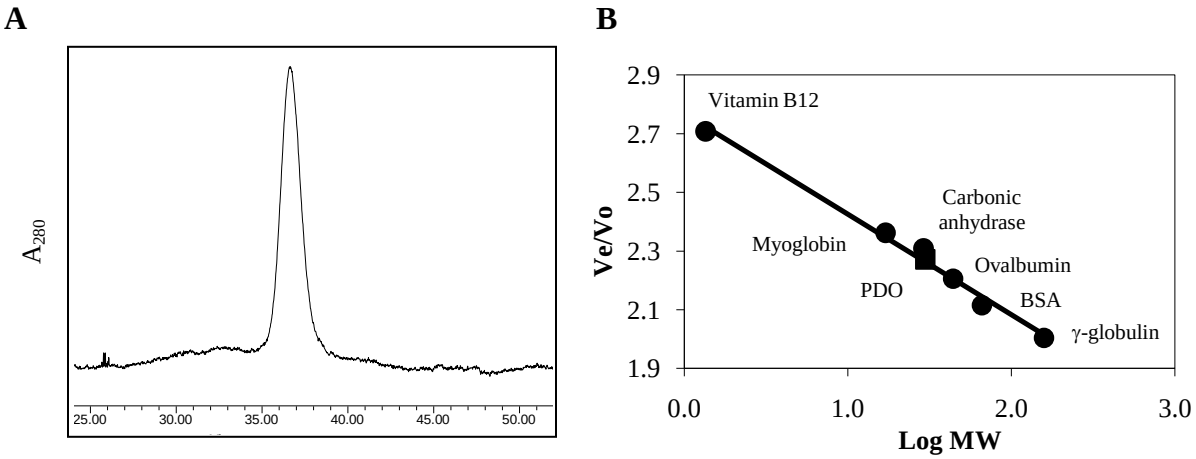


Figure 8.10: (A) The gel permeation chromatogram of PDO in 50 mM Tris (pH 8.0) and 150 mM KCl using a Superose6 10/300 GL column with a flow rate of 0.5 mL/min. The peak indicates the eluted fraction of PDO, which was identified by SDS–PAGE. **(B)** The gel permeation chromatographic profile of PDO (■) that fits into a plot of the Bio-Rad protein standards (●) including γ-globulin (158 kDa), BSA (66 kDa), ovalbumin (44 kDa), carbonic anhydrase (29 kDa), myoglobulin (17 kDa) and vitamin B12 (1.35 kDa).

Similarly, the overexpression of PLA was optimized using growth conditions of 25 °C and induction by 0.5 mM IPTG at the same temperature for a total of four hours (Figure 8.11a-d). A cell pellet of approximately 2.9 g was harvested from 1 L cell culture. The purification using the HisTrap HP affinity column isolated approximately 32.5 mg of His-tagged proteins (~13 mg/mL) from the cell lysate (Figure 8.12). After cleavage by thrombin protease and purification by the combined HisTrap HP and HiTrap Benzamidine FF affinity columns, the purified PLA was recovered at approximately 15 mg (Figure 8.13). The purity of PLA during the purification steps was visualized by SDS-PAGE analysis (Figure 8.14A). The molecular weight of PLA was estimated to be approximately 20 kDa by SDS-PAGE, while ESI-MS provided a more precise molecular mass of 16748.0 Da (calculated MW of 16749.9 Da, Figure 8.14B). No post-translational modification was observed based on the mass analysis. The mass of PLA under native conditions was estimated by analysis of the gel permeation chromatography data using a Superose6 10/300 GL column with a flow rate of 0.5 mL/min to be approximately 33 kDa (Figure 8.15), suggesting the purified PLA forms a dimer in 50 mM Tris (pH 8.0) and 150 mM KCl.

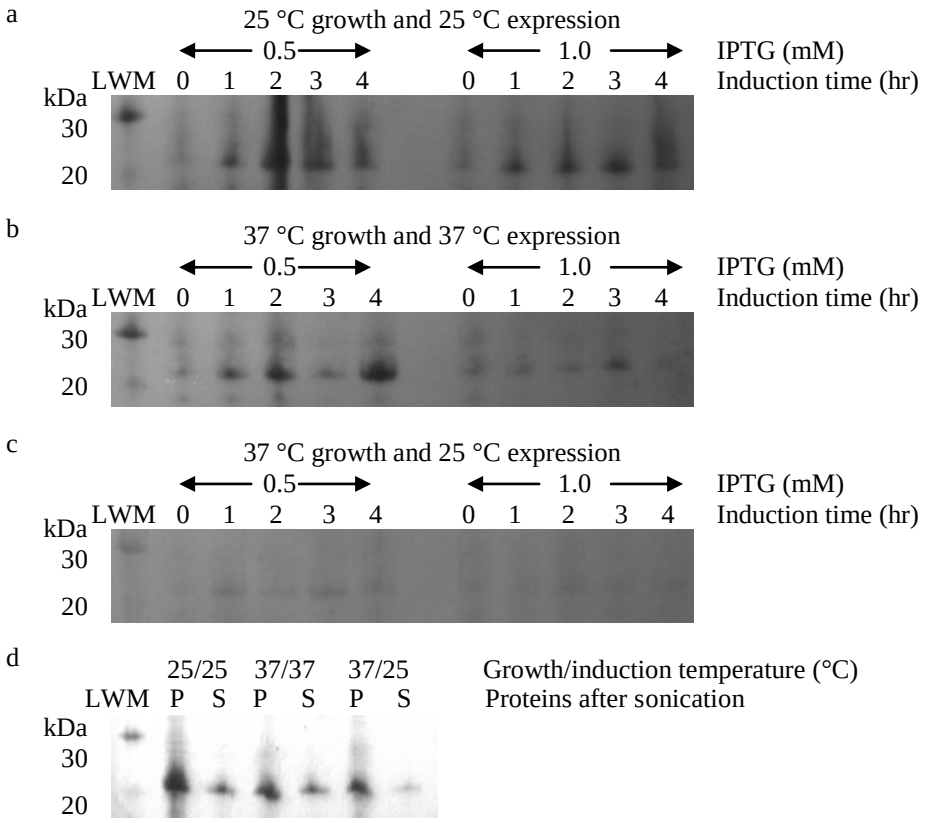


Figure 8.11: The SDS–PAGE of the optimized conditions for induction and expression tests of PLA purification including (a) 25 °C growth and induction, (b) 37 °C growth and induction and (c) 37 °C growth and 25 °C induction. Amount of IPTG and induction time were varied. (d) The SDS–PAGE of expressed proteins under various growth and induction temperatures with addition of 0.5 mM IPTG shows proteins in pellet (P) and supernatant (S) after sonication.

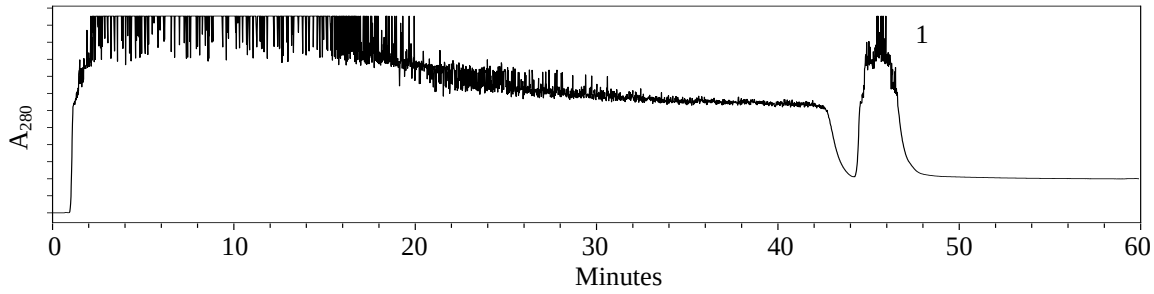


Figure 8.12: The HPLC chromatogram of PLA purification using a HisTrap HP affinity column with a flow rate of 0.5 mL/min. The lysate was loaded into the column and washed with 50 mM Tris (pH 8.0), 500 mM KCl and 20 mM imidazole (0–44 min). His-tagged proteins were then eluted using the same buffer but with 500 mM imidazole (peak 1). The proteins in the eluted fraction were monitored by SDS–PAGE.

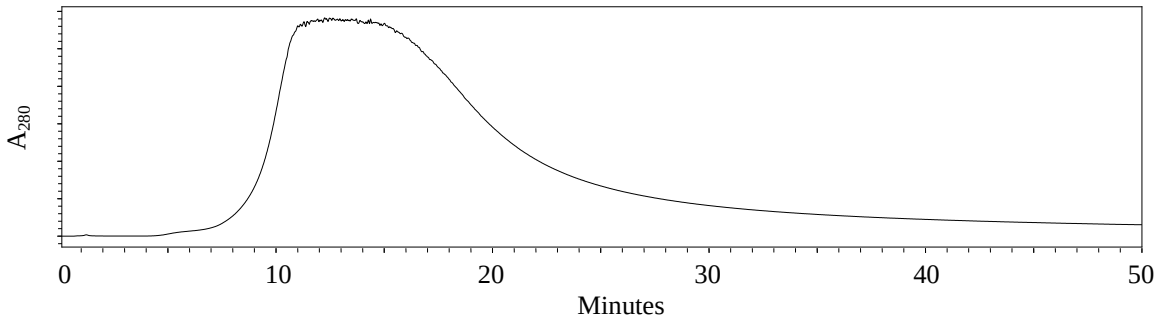


Figure 8.13: The HPLC chromatogram of PLA purification after treatment with thrombin protease using a HiTrap Benzamidine FF and HisTrap HP affinity columns with a flow rate of 0.5 mL/min and a buffer containing 50 mM Tris (pH 8.0), 500 mM KCl and 20 mM imidazole. The purified PLA (without His-tag) was eluted as shown as a broad peak over 7–30 min and was monitored by SDS–PAGE.

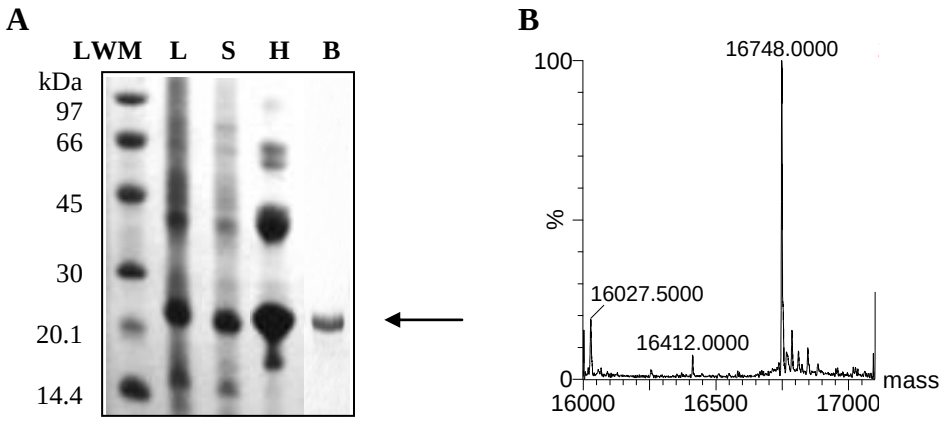


Figure 8.14: (A) The SDS–PAGE result of PLA purification: LMW = low molecular weight marker, L = lysate, S = supernatant after sonication, H = HisTrap fraction and B = HisTrap and Benzamidine fraction after thrombin protease treatment. The arrow indicates the presence of PLA. **(B)** The electrospray mass spectrum of PLA with a molecular mass of 16748.0 Da (calculated MW is 16749.9 Da).

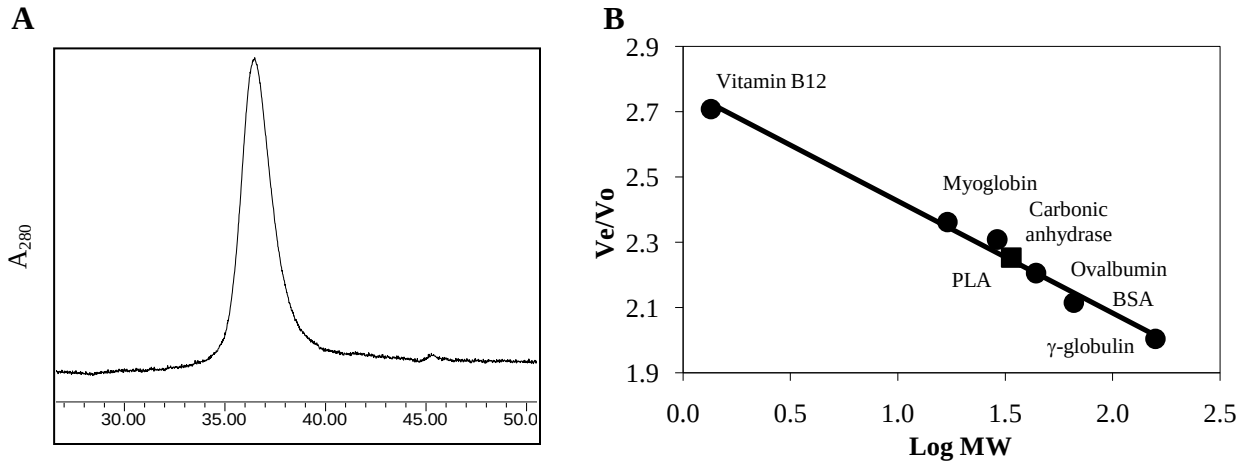


Figure 8.15: (A) The gel permeation chromatogram of PLA using a Superose6 10/300 GL column with 50 mM Tris (pH 8.0) and 150 mM KCl and a flow rate of 0.5 mL/min. The peak indicates the eluted fraction of PLA,

which was identified by SDS–PAGE. **(B)** The gel permeation chromatographic profile of PLA (■) that fits into a plot of the Bio-Rad protein standards (●) including γ-globulin (158 kDa), BSA (66 kDa), ovalbumin (44 kDa), carbonic anhydrase (29 kDa), myoglobulin (17 kDa) and vitamin B12 (1.35 kDa).

8.4.3: Protein Stability and Initial Structural Investigation

Putative Dioxygenase

The protein stability under various conditions was investigated using differential scanning calorimetry (DSC), circular dichroism (CD), fluorescence spectroscopy and ^{1}H NMR. Analysis of the DSC data on PDO in 50 mM MOPS (pH 7.0) and 10% (v/v) glycerol indicated that denaturation was reversible and occurred at a midpoint (T_m) of 57.8 °C (Figure 8.16A). An investigation of the secondary structure of the protein as determined by analysis of CD data performed on the protein suggested that the native PDO at protein concentrations lower than 5.6 mg/mL (168 μM) is stable in buffer containing 50 mM KPB (pH 7.0) and 200 mM KCl. The CD spectra of PDO at low protein concentrations (< 5.6 mg/mL) possessed a negative maximum at 222 nm, a small shoulder at 211 nm and a positive band at wavelengths less than 209 nm (Figure 8.16B). The shape of this CD spectrum suggested the secondary structure of the protein consists predominantly of β-sheet. The K2D secondary structural prediction [152, 153] estimated the secondary structural contents to be 3% α-helix, 49% β-sheet and 47% random coil. As the protein concentration increased (>168 μM), the negative band at 222 nm shifted to 227 nm and the small shoulder at 211 nm became unnoticeable. The changes in shape and shift of the CD spectra might suggest the possible formation of oligomers or a low quality of CD signal, the later would significantly affect the samples with high protein concentrations (>168 μM) at wavelengths below 210 nm due to high tension (HT) voltage. The increase in the positive absorbance at wavelength less than 209 nm was also observed as the protein concentration was decreased. The purified PDO at low protein concentration probably contains looser structure, which becomes more exposed to solvent, thus creating more hydrogen bonds with solvent and causing the increase in these positive bands. The structure of the oligomeric state, on the other hand, would be more packed, thus avoiding interactions with solvent and causing these bands to disappear. The change in positive absorbance, however, does not affect the predicted secondary structural contents.

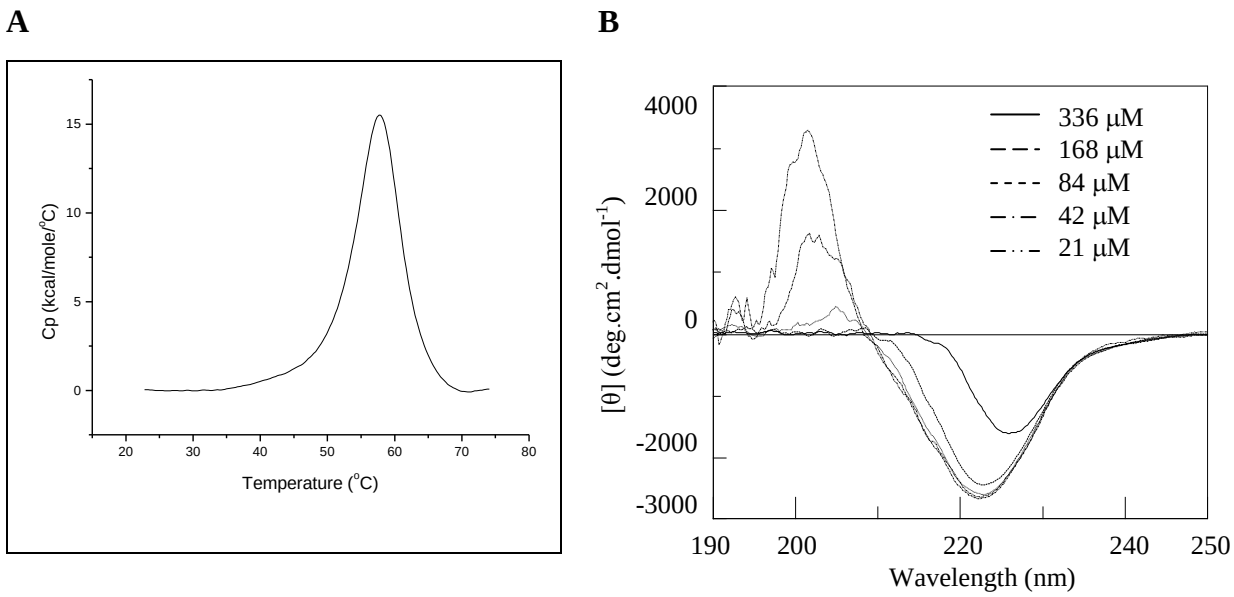


Figure 8.16: (A) The plot of heat capacity (Cp) and temperature (t) showing the peak of the temperature for transition midpoint (T_m) of PDO in 50 mM MOPS (pH 7.0) and 10% (v/v) glycerol. The T_m was estimated to be 57.8 °C. **(B)** The CD spectra of PDO (336, 168, 84, 42 and 21 μM) in buffer containing 50 mM KPB (pH 7.0) and 200 mM KCl scanning between 190–250 nm. The CD analysis indicated that the protein was stable when the concentration was less than 168.08 μM (5.6 mg/mL).

Interestingly, a change in buffer pH only affected the positive maxima, while the negatively bands were unaltered (Figure 8.17A). The increased positive absorptions were observed with protein in relatively acidic buffers. A possible explanation may be that the structure of the protein becomes looser, thus allowing more interactions with solvent. Similarly, the effect of buffers only caused a change in the positive CD absorptions (Figure 8.17B). Only PDO in KPB and Tris buffer possessed the positive bands, while others (HEPES, MOPS and MOPSO) did not. It is possible that the protein is more stable in KPB and Tris buffer since other buffers might have a greater propensity to create protein oligomers for some reason. Despite these observable changes in the CD spectra upon changes in pH and buffer, salt (KCl) and glycerol did not seem to alter the secondary structure of PDO (Figure 8.18A-B).

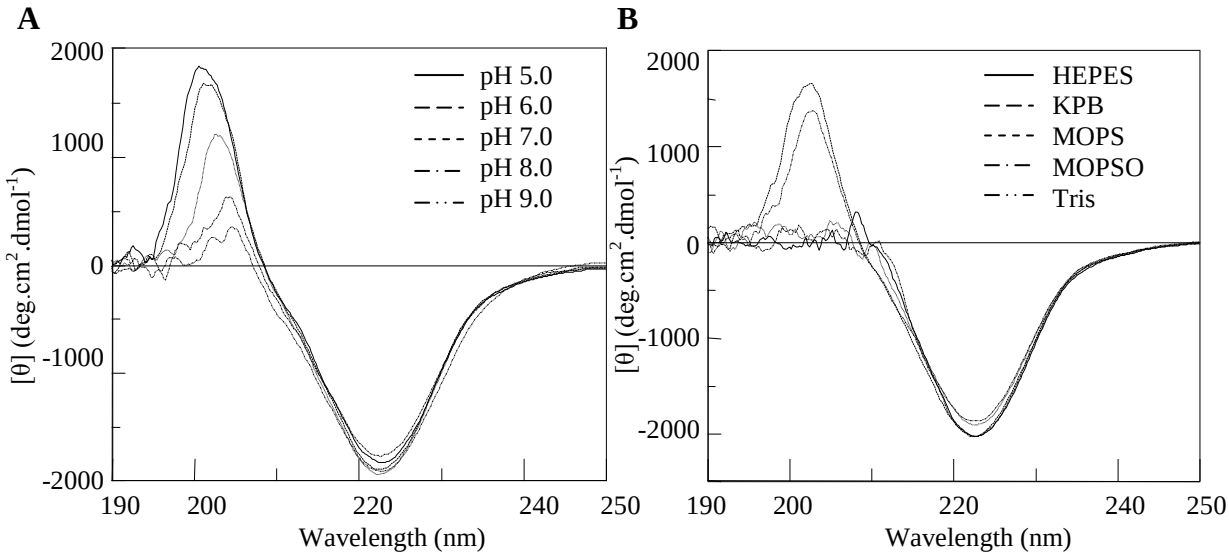


Figure 8.17: The CD spectra of PDO (2.8 mg/mL) in **(A)** Tris buffer (pH of 5.0, 6.0, 7.0, 8.0 and 9.0) and **(B)** in different buffers (HEPES, KPB, MOPS, MOPSO and Tris, all with pH of 7.0) scanning between 190–250 nm. The CD analysis indicated that the secondary structure of PDO was not significantly altered in buffer pH of 5–9 and type of buffers.

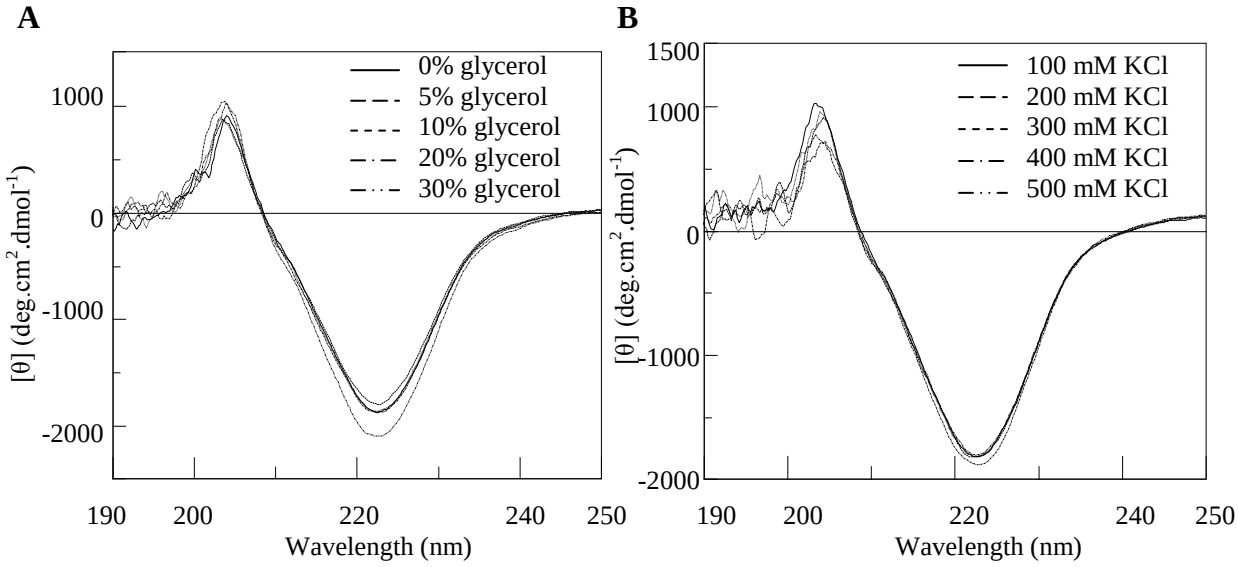


Figure 8.18: The CD spectra of PDO (2.8 mg/mL) in **(A)** Tris buffer (pH 7.0) containing various glycerol concentrations (0, 5, 10, 20 and 30% (v/v) glycerol) and **(B)** in Tris buffer (pH 7.0) containing different salt concentration (100, 200, 300, 400 and 500 mM KCl) scanning between 190–250 nm. The CD analysis indicated that the secondary structure of the protein was not significantly altered in buffer with 0–30% (v/v) glycerol and salt concentration of 100–500 mM.

The fluorescence emission scanned from 300 to 500 nm with excitation wavelength at 289 nm was measured to detect the signal of the Trp residue in PDO (only one Trp is present in this protein) in attempts to investigate protein folding and stability under various conditions.

Normally, the folded protein that is excited at 289 nm emits the fluorescence, which belongs to aromatic residues such as Trp, Tyr and Phe with Trp being a major fluorescence emission source at the wavelength between 300 and 350 nm depending on polarity of the protein environment. The free acid form of Trp in water exhibits a fluorescence emission wavelength at 348 nm with a fixed excitation wavelength at 280 nm. However, this emission wavelength can shift to 333 nm when Trp is buried within the folded structure of the protein [282]. Thus, the fluorescence emission wavelength can shift to shorter wavelength and its intensity increases as the function of decreased polarity of the solvent surrounding of Trp residue. It was also reported that Trp residues that are buried in the hydrophobic core of proteins could shift by 10 to 20 nm comparing to Trp residues at the surface of the protein [282]. Additionally, Trp residues are considered as buried residues in the non-polar environment (or hydrophobic core of the protein) if λ_{max} is less than ~330 nm [282]. On the other hand, if λ_{max} shifts to greater wavelength, Trp residues likely reside in a polar environment [282].

The fluorescence spectra of PDO exhibited a maximum emission wavelength at 335 nm and the intensity of this peak was elevated as protein concentration was increased (Figure 8.19A). This suggested that more Trp residues were detected, which corresponded to the increase in protein concentration. The detection of this emission wavelength (335 nm) also suggested that the Trp residue in PDO is possibly located in the interior of the protein. However, once the protein concentration was reduced to less than 5.3 μM (< 0.2 mg/mL), the maximum emission wavelength shifted from 335 to 380 nm (red shift). This shift possibly suggests that the tertiary structure of PDO becomes more unstable such that the Trp residue is exposed to the outside environment (Figure 8.19B). Thus, the protein at concentrations lower than 5.3 μM might be unstable. However, this migt be an artifact from buffers. Thus, combining the fluorescence analysis with the previous CD investigation on protein concentration, a suitable concentration to maintain protein integrity was between 5 μM–168 μM (0.2–5.6 mg/mL). On the other hand, the fluorescence emission spectra of PDO under various buffer pH conditions (pH of 5–9), salt concentrations (100–500 mM), glycerol concentrations (0–30% v/v) and types of buffer (Tris, HEPES, KPB, MOPS and MOPSO) suggested that these factors did not significantly interfere with the tertiary structure such that the maximum emission wavelength of PDO under these conditions was unaltered (Figure 8.20 and 8.21).

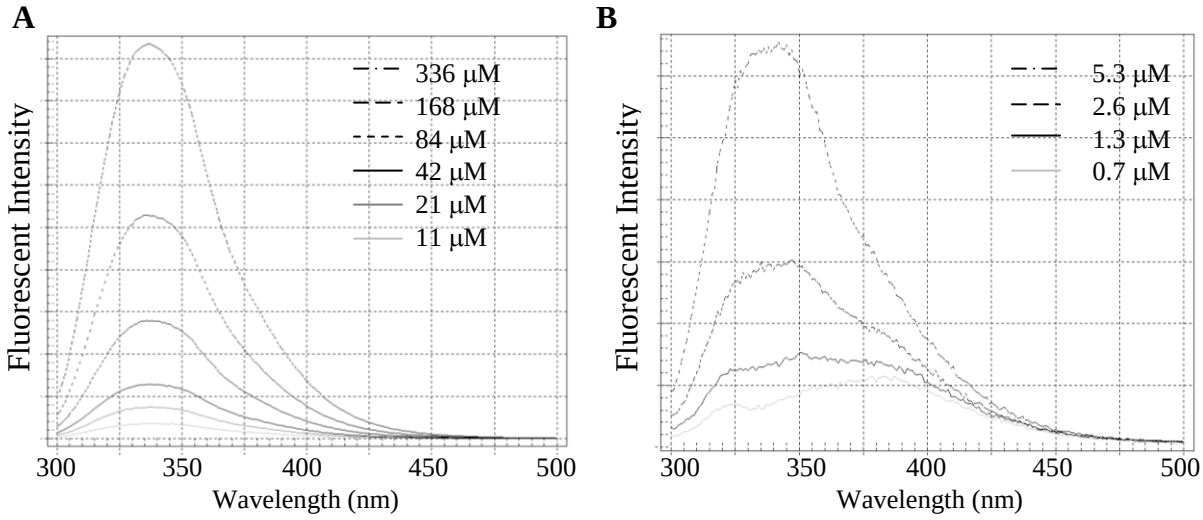


Figure 8.19: The fluorescence emission spectra of PDO under **(A)** high protein concentrations (336, 168, 84, 42, 21 and 11 µM) and **(B)** low protein concentrations (5.3, 2.6, 1.3 and 0.7 µM) in 50 mM KPB (pH 7.0) and 200 mM KCl scanning between 300–500 nm with a fixed excitation wavelength at 289 nm.

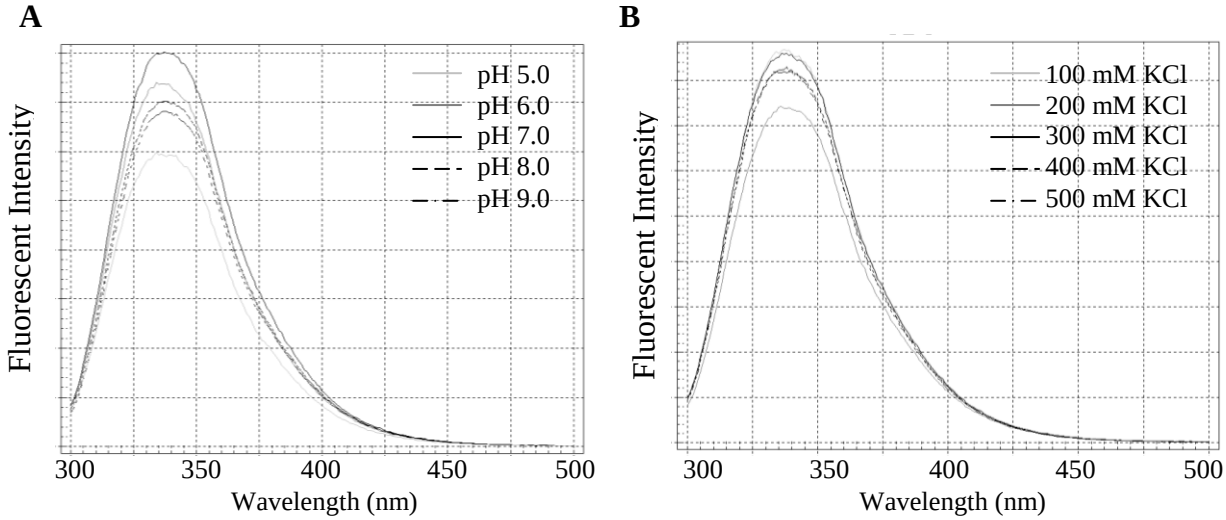


Figure 8.20: The fluorescence emission spectra of PDO (2.8 mg/mL) in **(A)** Tris buffer pH of 5.0, 6.0, 7.0, 8.0 and 9.0 and **(B)** in Tris buffer (pH 7.0) containing different salt concentrations (100, 200, 300, 400 and 500 mM KCl) scanning between 300–500 nm with a fixed excitation wavelength at 289 nm.

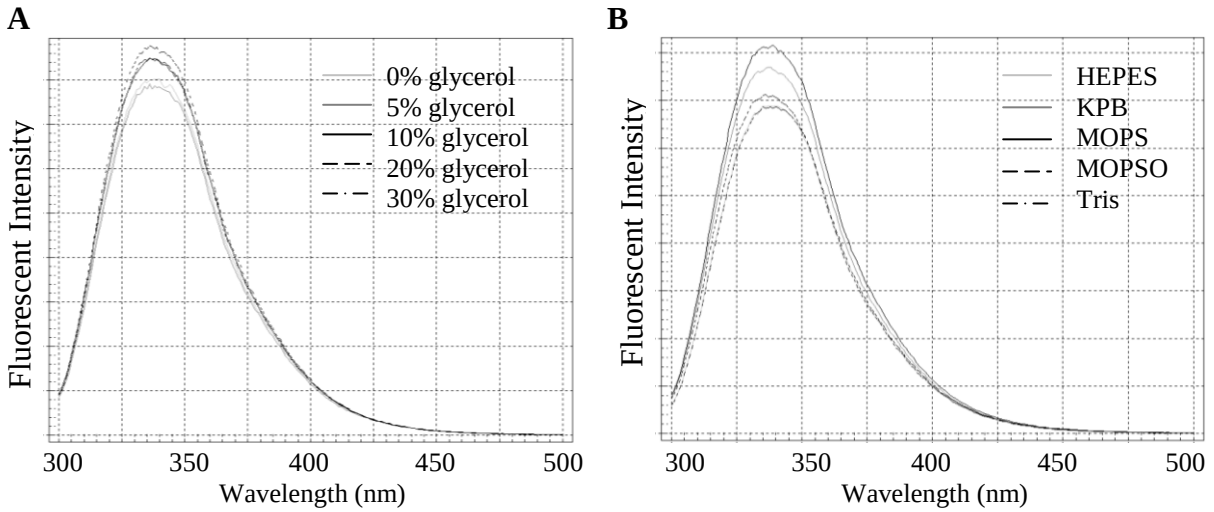


Figure 8.21: The fluorescence emission spectra of PDO (2.8 mg/mL) in **(A)** Tris buffer (pH 7.0) containing various glycerol concentrations (0, 5, 10, 20 and 30% (v/v) glycerol) and **(B)** in different buffers (HEPES, KPB, MOPS, MOPSO and Tris, all with pH of 7.0) scanning between 300–500 nm with a fixed excitation wavelength at 289 nm.

The ^{1}H NMR (600 MHz) of PDO in buffer containing 50 mM Tris (pH 8.0), 150 mM KCl and 10% (v/v) glycerol can be characterized mainly in two areas including a low field region between 9.5–6.5 ppm and a high field region between 2.3–0.5 ppm. The resonances in the low-field region represent backbone amide hydrogen (–NH–) and the resonances at high-field region indicate methyl and methylene groups (–CH_3 and –CH_2–, respectively) from the side chains of some of the amino acids of the protein. The protein in its native state (properly folded) normally possesses each hydrogen in its unique environment (with its unique characteristic resonance), thus exhibiting resonances as multiple resonances in these regions. On the other hand, if the protein is in a denatured or unfolded state, its amide, methyl and methylene groups will be in the same environment and its NMR signal is expected to be represented as a single broad resonance. However, the report on the proton NMR of denatured lysozyme (pH 3.8 at 82 °C) and SH3 domain (3.5 M guanidine-hydrochloride at pH 7.2 and at pH 2.0) in D_2O suggested that the resonances in the low-field and high-field regions still exhibits a residual heterogeneity of chemical shift but contain a considerable loss of dispersion in chemical shifts relative to those in the spectrum of the native protein [283]. Since multiple resonances in both low- and high-field regions were observed in the ^{1}H NMR spectrum of PDO, we predicted that this protein likely possesses a properly folded structure under the studied conditions (Figure 8.22).

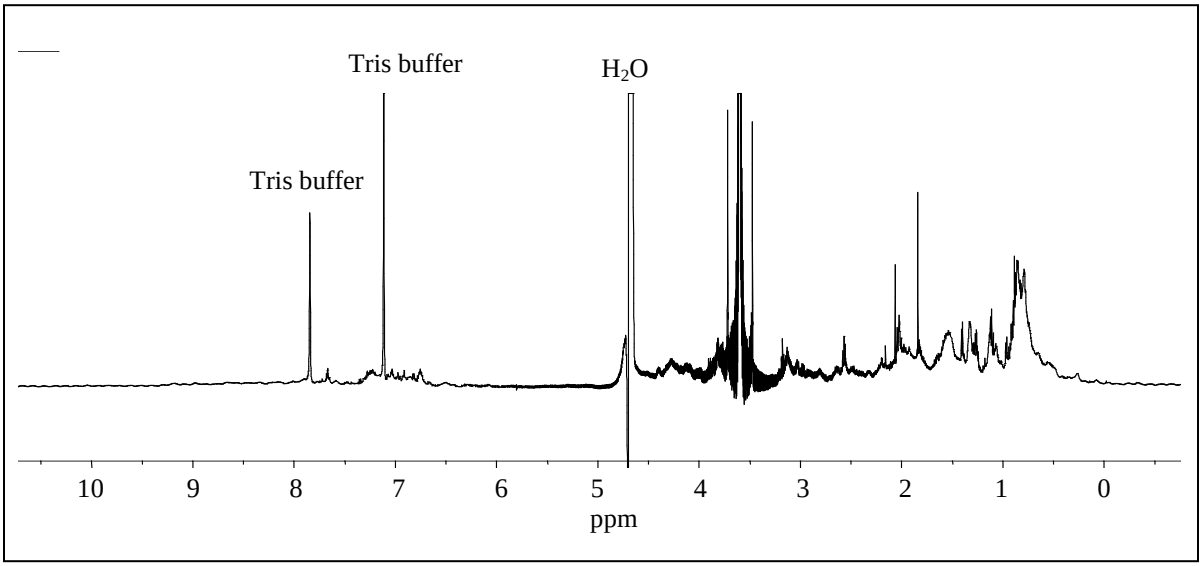


Figure 8.22: The ^{1}H NMR spectrum of PDO in Tris buffer (50 mM Tris (pH 8.0), 150 mM KCl and 10% glycerol) exhibiting evidence for a properly folded protein.

Putative Lyase

The temperature for transition midpoint (T_m) of PLA in 50 mM MOPS (pH 7.0) and 10% glycerol estimated by DSC analysis was determined to be 58.5 °C (Figure 8.23A). The secondary structure investigation of PLA by CD was similar to that of PDO. The CD spectra of PLA under low protein concentrations (< 108 μM or 3.6 mg/mL) exhibited a negative maximum at 222 nm and a shoulder at 208 nm, suggesting the presence of a protein with a predominant β-sheet structure (Figure 8.23B). This protein seemed to contain more β-sheet structure than PDO since it contained higher negative absorbance at 208 nm. The K2D secondary structural prediction [152, 153] supported this hypothesis, where approximately 5% α-helix, 47% β-sheet and 48% random coil structures were estimated to be present. Positive CD absorbance at wavelengths lower than the 205 nm absorbance for PLA under low protein concentrations was also detected, which suggested interactions of the exposed protein with solvent. Once the protein concentration was elevated (> 108 μM), the negative maximum at 222 nm shifted to 225 nm, while the shoulder at 208 nm and the positive CD absorbance at wavelengths lower than 205 nm disappeared. These changes in shape and shift in the CD spectra suggested the possibility that the PLA protein takes on an oligomeric structure at high protein concentrations. However, these might be caused by low quality of CD signal, which

would significantly affect the samples with high protein concentrations (> 3.6 mg/mL) at wavelengths below 210 nm due to high tension (HT) voltage.

As well, analysis of the CD spectra as a function of the buffer pH indicated that all protein samples in buffer with different pH possessed similar negative absorbances (Figure 8.24A). However, the increase in positive CD bands suggested that the protein is more stable in more acidic buffer conditions (pH 5–7). Even though these positive bands disappeared in the CD spectra of the protein under basic buffer conditions (pH 8–9), the secondary structural contents predicted by K2D secondary structural prediction [152, 153] were unaltered. Since the predicted p*I* of PLA is 4.89, a suitable buffer pH to maintain protein stability would be between pH 6 and 7. Similarly, the protein in KPB and Tris buffer exhibited CD spectra with negative maxima at 222 nm, a shoulder at 208 nm and positive absorbances at wavelengths lower than 205 nm (Figure 8.24B). The protein in HEPES, MOPS and MOPSO however only possessed the negative maximum at 222 nm, while the shoulder at 208 nm and the positive band were absent. Analyses of these data suggested that the protein has a greater tendency to form oligomers in these buffers, and KPB and Tris might be more suitable buffers to store the protein. However, this might be an artifact from buffers. On the other hand, there was no significant change in the CD spectra of PLA in buffer containing different concentrations of salt and/or glycerol (Figure 8.25A-B). Analyses of these data suggested that these two factors did not crucially affect protein stability.

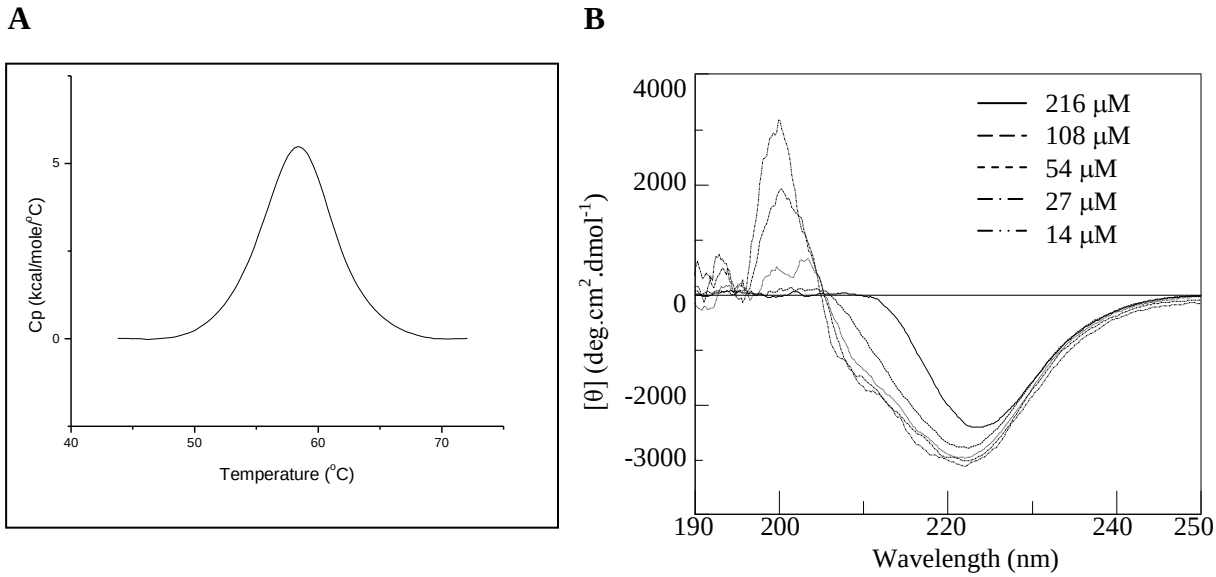


Figure 8.23: (A) The plot of heat capacity (Cp) and temperature (t) showing the temperature for transition midpoint (T_m) of PLA in 50 mM MOPS (pH 7.0) and 10% (v/v) glycerol. The T_m was estimated to be 58.5 °C.

(B) The CD spectra of PLA (216, 108, 54, 27 and 14 μM) in 50 mM KPB (pH 7.0) and 200 mM KCl scanning between 190–250 nm. The CD analysis indicated that the protein was stable when the protein concentration was less than 108 μM (3.6 mg/mL).

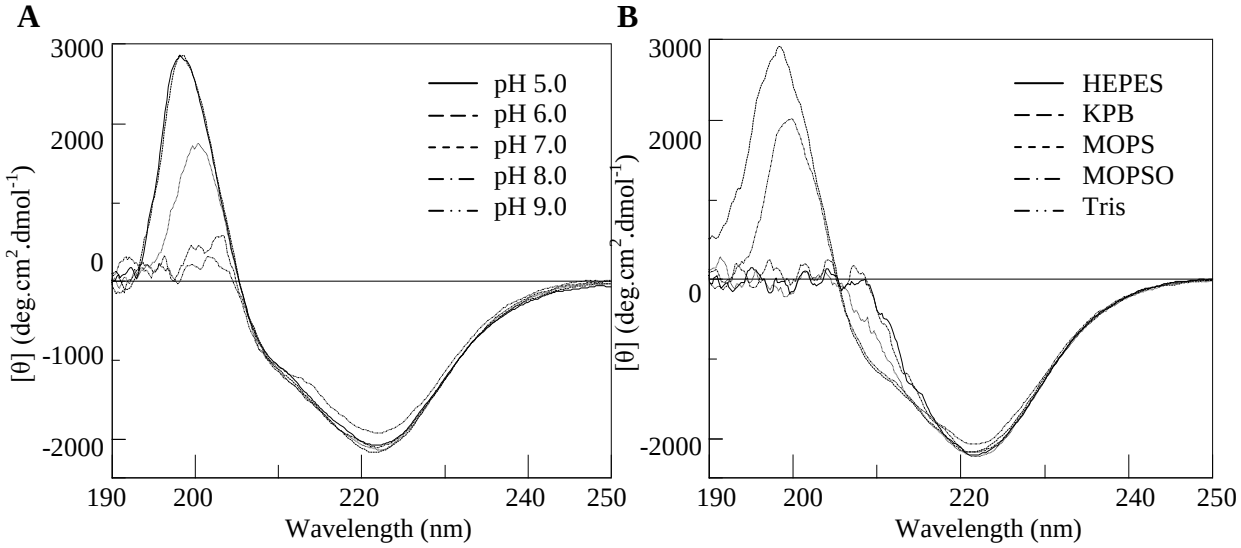


Figure 8.24: The CD spectra of PLA (1.8 mg/mL) in **(A)** Tris buffer pH of 5.0, 6.0, 7.0, 8.0 and 9.0 and **(B)** in different buffers (50 mM HEPES, KPB, MOPS, MOPSO and Tris, all at pH of 7.0) scanning between 190–250 nm. The CD analysis indicated that the secondary structure of the protein was not significantly altered in either buffer pH of 5–9 or type of buffers.

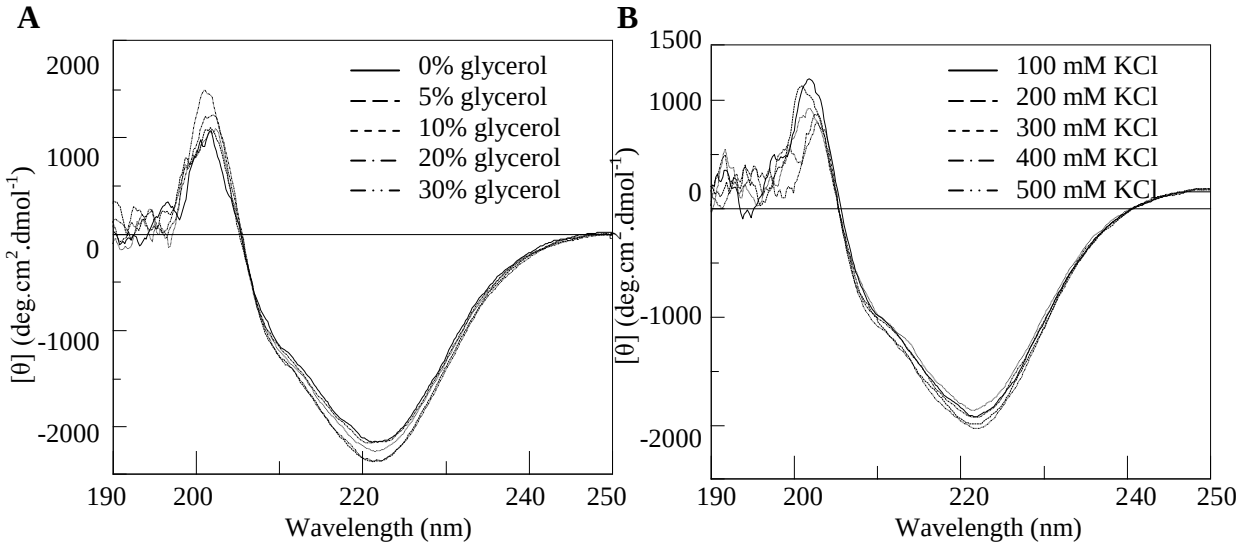


Figure 8.25: The CD spectra of PLA (1.8 mg/mL) in **(A)** Tris buffer (pH 7.0) containing various glycerol concentrations (0, 5, 10, 20 and 30% (v/v) glycerol) and **(B)** in Tris buffer (pH 7.0) containing different salt concentrations (100, 200, 300, 400 and 500 mM KCl) scanning between 190–250 nm. The CD analysis indicated that the secondary structure of the protein was not significantly altered in either buffer with 0–30% (v/v) glycerol or salt concentration of 100–500 mM.

The results with fluorescence spectroscopy of the PLA protein were also similar to those of the PDO protein. Analysis of the maximum fluorescence emission of approximately 335 nm with a fixed excitation wavelength at 289 nm using a protein concentration anywhere between 216–1.7 μM suggested that Trp residues (2 Trp in PLA) may locate in the interior of the protein (Figure 8.26). However, the observed shift of the maximum emission peak from 335 to 345 nm (red shift) when the protein concentration was lowered to values less than 0.8 μM (< 0.03 mg/mL) suggested that the structure of PLA becomes more open to the external medium such that the Trp residues are likely more exposed to the environment (Figure 8.26B). However, the change in shape and shift of emission spectra might be an artifact from buffers. Analyses of these fluorescence data as well as the secondary structure investigated by CD experiments suggested that the protein is stable in buffer between 0.8–108 μM (5.6–0.03 mg/mL). Similarly, a small shift in maximum emission wavelength from 330 to 345 nm as the buffer pH was increased from 5 to 9 indicated that buffer pH might slightly affect the structure of PLA in this pH region (Figure 8.27A). Thus, these data support the CD analysis, which suggested that the protein is stable between pH 6–7. Other factors including salt (100–500 mM KCl) and glycerol (0–30% v/v) as well as choice of buffers (Tris, HEPES, KPB, MOPS and MOPSO) had no major affect on the structure of PLA (Figure 8.27B and 8.28A-B).

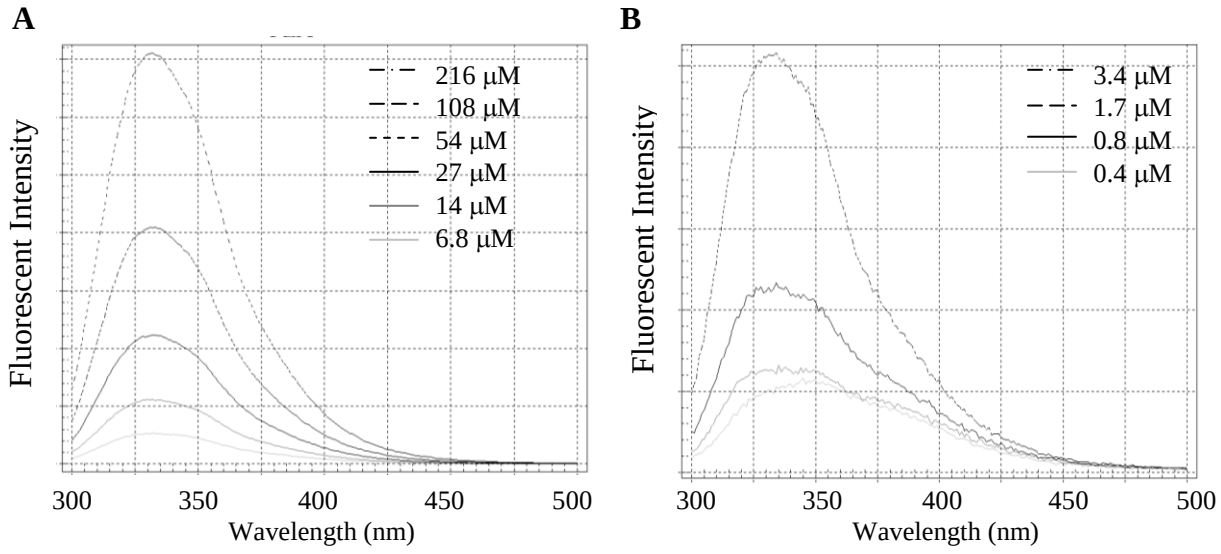


Figure 8.26: The emission spectra of PLA under **(A)** high protein concentrations (216, 108, 54, 27, 14 and 6.8 μM) and **(B)** low protein concentrations (3.4, 1.7, 0.8 and 0.4 μM) in 50 mM KPB (pH 7.0) and 200 mM KCl scanning between wavelengths of 300–500 nm with the fixed excitation wavelength of 289 nm.

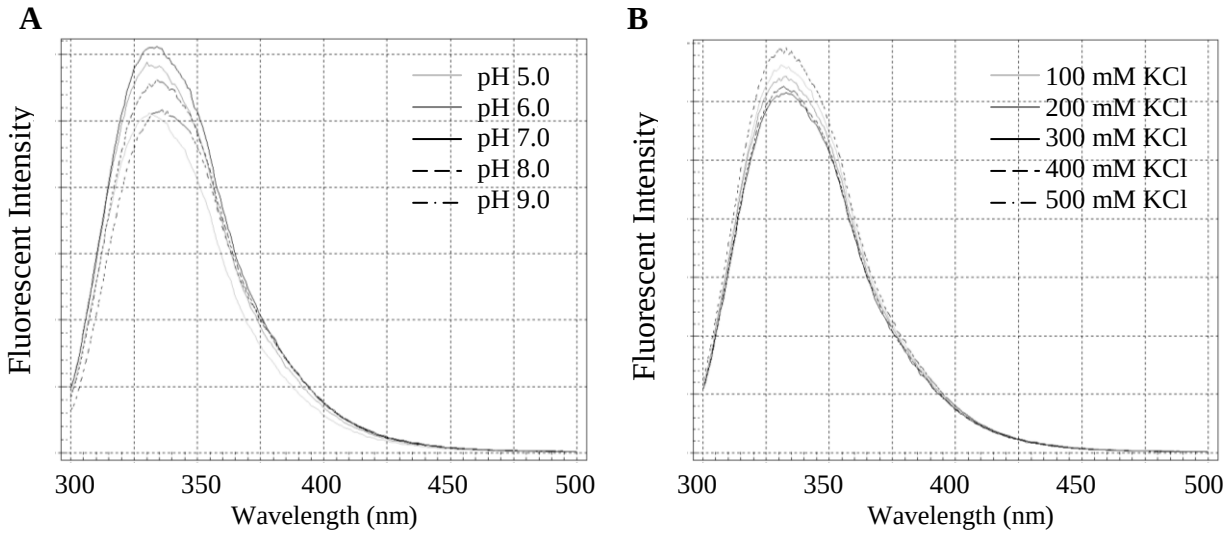


Figure 8.27: The emission spectra of PLA (1.8 mg/mL) in **(A)** Tris buffer pH of 5.0, 6.0, 7.0, 8.0 and 9.0 and **(B)** in Tris buffer (pH 7.0) containing different salt concentrations (100, 200, 300, 400 and 500 mM KCl) scanning between wavelengths of 300–500 nm with a fixed excitation wavelength of 289 nm.

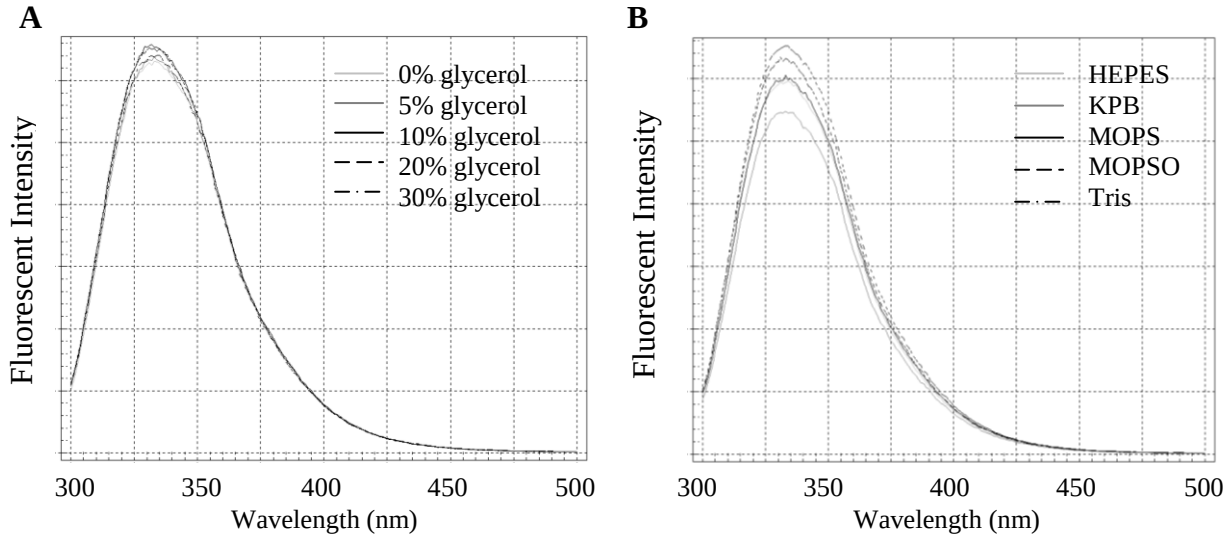


Figure 8.28: The emission spectra of PLA (1.8 mg/mL) in **(A)** Tris buffer (pH 7.0) containing various glycerol concentrations (0, 5, 10, 20 and 30% (v/v) glycerol) and **(B)** in different buffers (HEPES, KPB, MOPS, MOPSO and Tris, all with pH of 7.0) scanning between wavelengths of 300–500 nm with a fixed excitation wavelength of 289 nm.

As described previously, the ^{1}H NMR (700 MHz) of PLA in 50 mM MOPS (pH 7.0) and 10% (v/v) glycerol showed multiple resonances at both low field (6.5–9.5 ppm) and high field (2.3–0.5 ppm), suggesting that this protein possesses a properly folded structure under the studied conditions (Figure 8.29).

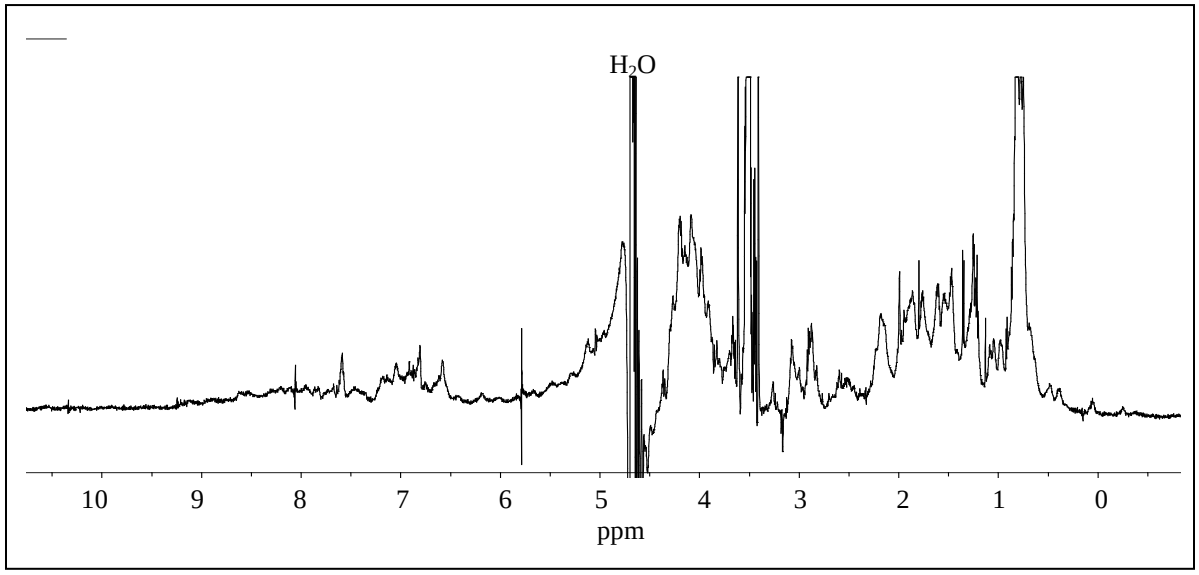


Figure 8.29: The ^{1}H NMR spectrum of PLA in 50 mM MOPS (pH 7.0) and 10% glycerol.

8.4.4: Enzymatic Activity of MycoGlyoxalase I

Equilibrium Time for Hemithioacetal Formation

The time for the formation of the hemithioacetal substrate to form and reach equilibrium when methylglyoxal (MG) and the truncated MSH are added together was investigated using commercial MG and the synthesized truncated MSH (tMSH, detailed synthesis in Chapter 7). The non-enzymatic product of these two substrates is the hemithioacetal of MG-tMSH, the actual substrate expected for the putative mcGlxI. The hemithioacetal (0.5 mM) was incubated at various time periods before performing the assay using Ni^{2+}-reconstituted PDO (apo-enzyme with additional 5 equivalents of metal). We assumed that the detection of the product for the mcGlxI reaction (240 nm) and the dissociation constant of MG-tMSH (3.1 mM) would be similar to that of MG-GSH due to the presence of the common thioester moiety in both products. The activity profiles accomplished in this work suggested that the equilibrium time for formation of the MG-GSH hemithioacetal was 15 minutes, which corresponds well to what has been reported previously (Figure 8.30A) [85, 86, 141, 279-281]. However, a longer incubation time for the equilibrium formation of the MG-tMSH hemithioacetal was observed (30 minutes) (Figure 8.30B), suggesting that the formation of the hemithioacetal between these two substrates reaches equilibrium more slowly. These

results likely correspond to the lower nucleophilicity of MSH compared to that of GSH [93-98].

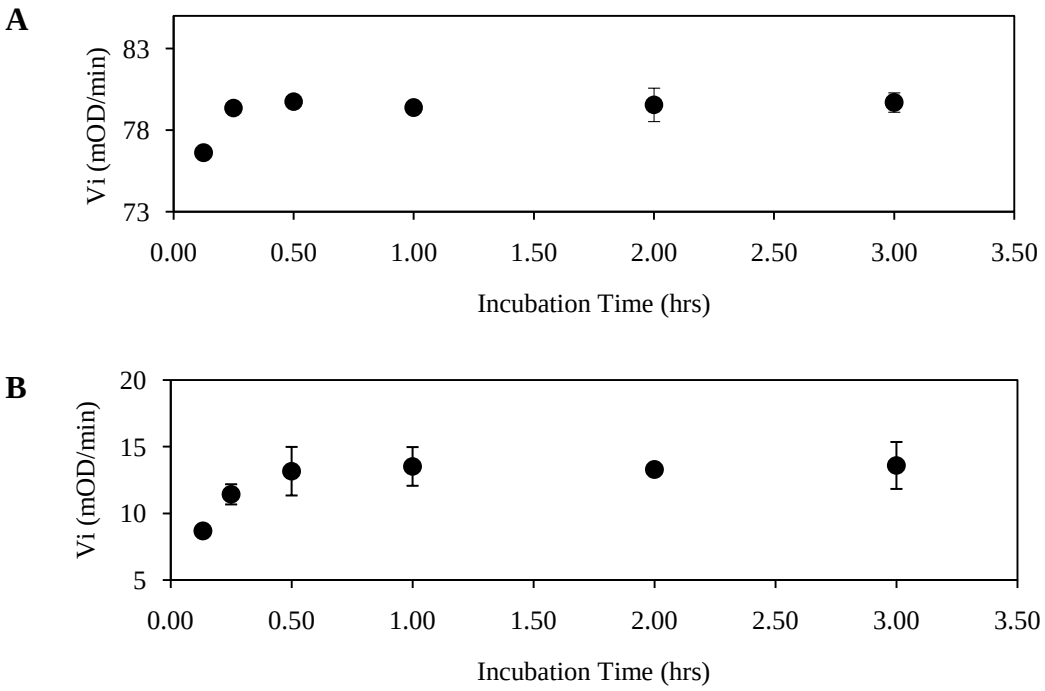


Figure 8.30: The enzymatic assay using various incubation times of **(A)** MG-GSH (0.5 mM) followed by reaction with yeast GlxI (0.15 μg in 300 μL assay) and **(B)** MG-tMSH (0.5 mM) with Ni^{2+}-reconstituted PDO (4.5 μg in 300 μL assay) in 50 mM KPB (pH 6.6) at room temperature (25 °C).

Another approach to support the previous investigation on equilibration time was utilization of ^1H NMR. The substrates, MG and GSH, were incubated in D_2O for various periods of time before performing proton NMR. Our experimental NMR spectra were compared with the reported NMR spectrum by Rae *et al*., 1994 [284], which suggested the diastereomeric mixture of interconverting hemithioacetals of MG-GSH. The appearance of NMR resonances for some of the hemiacetal hydrogens (such as H7 and H8, Table 8.5) as well as the changes in the chemical shifts of the proton resonances of their surrounding hydrogens (H4 and H5) between the starting thiol and the hemithioacetal product would additionally confirm the formation of the MG-GSH hemithioacetal (Figure 8.31). However, the resonance of one of the hemithioacetal's hydrogens, H7 (with estimated chemical shift at 4.39 ppm), is obscured under the water and MG resonances. The estimated chemical shift of the purified MG by ChemBioDraw Ultra 12.0 software (http://www.cambridgesoft.com) is at 2.17 ppm for

methyl hydrogens; however, our experimental proton NMR of MG suggested the same hydrogens at 2.09 ppm as well as other resonances (δ 1.93, 3.19, 5.13) that might correspond to polymerization of MG and/or the existence of other monomeric MG in different forms. It was reported that synthetic MG in carbon tetrachloride showed resonances of methyl protons at δ 2.2, while aldehydic protons of MG in water move to the δ 5–6 region as a result of hydration of aldehyde group [285, 286]. As well, the complex pattern in the methyl region and the appearance of a peak at δ 2.7 suggested the involvement of the ketonic carbonyl group in aldehyde hydration [285, 286]. The reported proton NMR spectrum of the hemithioacetal of MG and GSH also indicated the resonance of methylglyoxal monohydrate $–CH_3$ at δ 2.31 and $–CH(OH)_2$ at δ 5.32 as well as protons of unhydrolyzed methylglyoxal dimethyl acetal at δ 3.44, 3.56, 2.26 [284]. As well, this spectrum reported the clearly observed diastereomer of the hemithioacetal at H5 [284]. In our experiment, the change in chemical shift of H5 (δ 2.94–2.80 in GSH and δ 2.92–2.80 in the hemithioacetal) as well as additional splitting at δ 3.09–2.92 (as a result of diastereomeric formation) suggested the formation of the hemithioacetal. Another key hemithioacetal resonance, the three hydrogens, H8, was observed at δ 2.26. However, its integration was over-estimated due to the overlap with MG resonances. The change in chemical shift of H4 in GSH (4.49 ppm) to H4* in the hemithioacetal (5.51 ppm) were also clearly observed. These changes, however, occurred quickly after a few minutes of mixing MG and GSH. The lag in time might, as well, be taken into consideration (few minutes) due to the delay before starting the measurement. Nevertheless, no significant change in NMR signal was observed after 15 min incubation, thus supporting the previous investigation using the literature enzymatic assay.

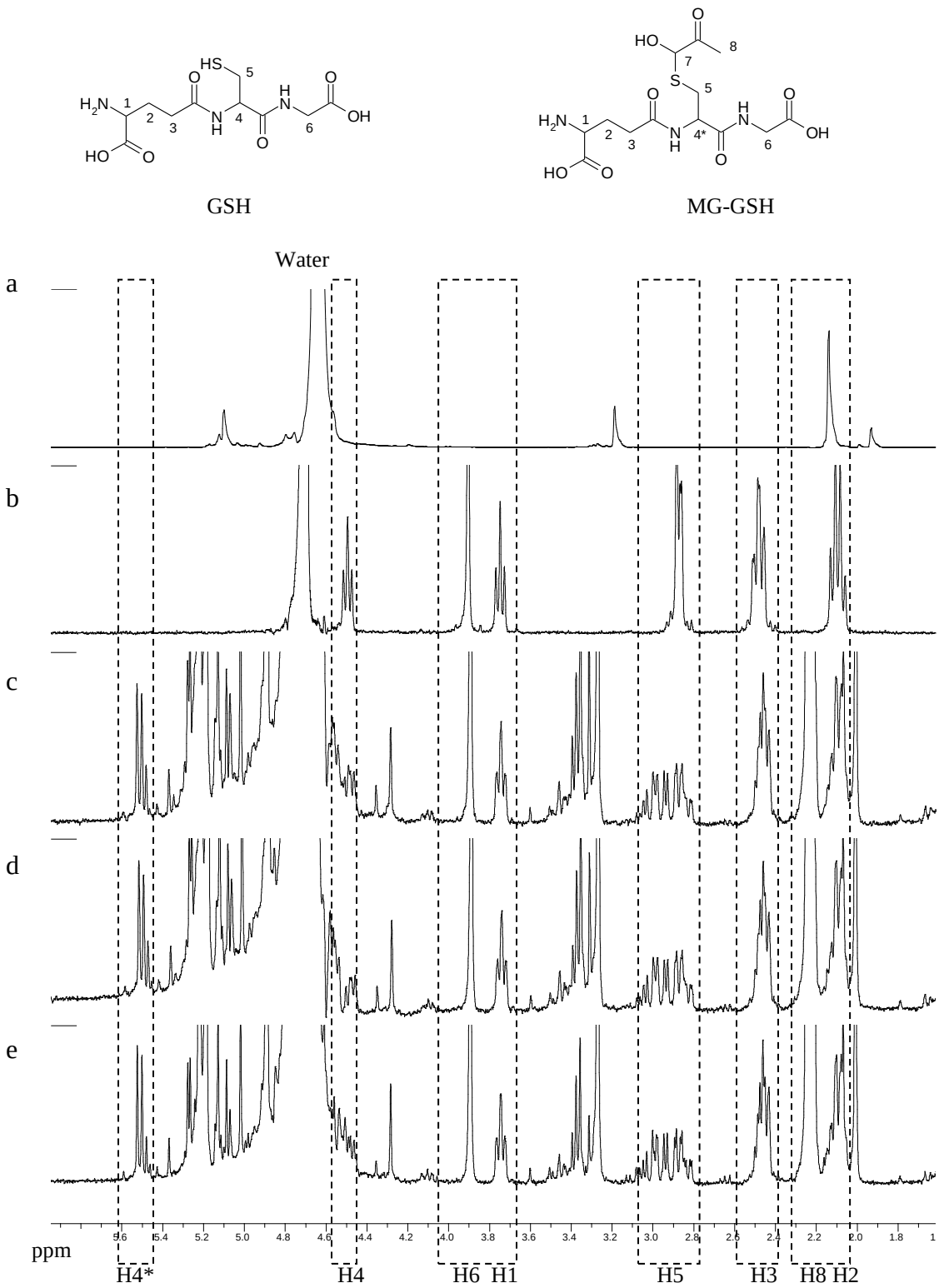
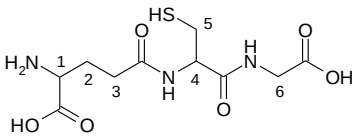
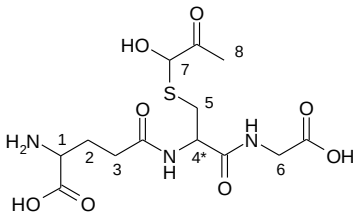


Figure 8.31: The ^{1}H NMR spectra for various incubation times for the formation of the hemithioacetal of MG-GSH; (a) MG, (b) GSH, (c) MG-GSH incubated for 8 min, (d) MG-GSH incubated for 15 min and (e) MG-GSH incubated for 30 min.

Table 8.5: The summary for chemical shift (CS) and integration (int.) of NMR signal for MG-GSH equilibration time compared to the resonances for GSH.





	GSH				***MG-GSH***			
Peak	***†Predicted CS (ppm)***	***Actual CS (ppm)***	***Predicted int.***	***Actual int.***	***†Predicted CS (ppm)***	***Actual Cs (ppm)***	***Predict Int.***	***Actual int.***
H1	4.06	3.75	1	1.01	2.89	3.75	1	0.86
H2	2.23	2.09	2	2.25	1.75–1.96	2.18–2.05	2	2.79
H3	2.60	2.49	2	2.16	2.35	2.47	2	2.03
H4	4.01	4.49	1	1.03	5.62	5.51	1	1.24
H5	2.94	2.88	2	2.12	3.21	3.07–2.80	2	1.79
H6	4.03	3.91	2	2.19	4.13	3.91	2	2.02
H7	-	-			4.39	-	1	-
H8	-	-			2.19	2.26	3	4.42

†Chemical shifts are estimated by ChemBioDraw Ultra 12.0 (http://www.cambridgesoft.com) software.

Similarly to the detection of the MG-GSH hemithioacetal adduct, the NMR signals for the formation of MG-tMSH as a function of time indicated the presence of the additional hemithioacetal hydrogen, H4 and the methyl group H5, as well as changes in chemical shift of H2 (Figure 8.32). However, the hemithioacetal hydrogen H4 was obscured by the water and MG peaks, while H5 overlapped with MG peaks (Table 8.6). Thus, their appearances were unclear. The chemical shifts of others were unaltered; however, some signals overlapped with MG peaks, causing the overestimated integration (H3, H4, H5, H1′-H6′). Only the change in chemical shift of H2, which could be observed as a series of unclear splitting after 2 min of incubation, could readily be employed to indicate the formation of MG-tMSH. This pattern became more defined after 30 min incubation, thus being in agreement with the previous investigation by use of the enzymatic reaction.

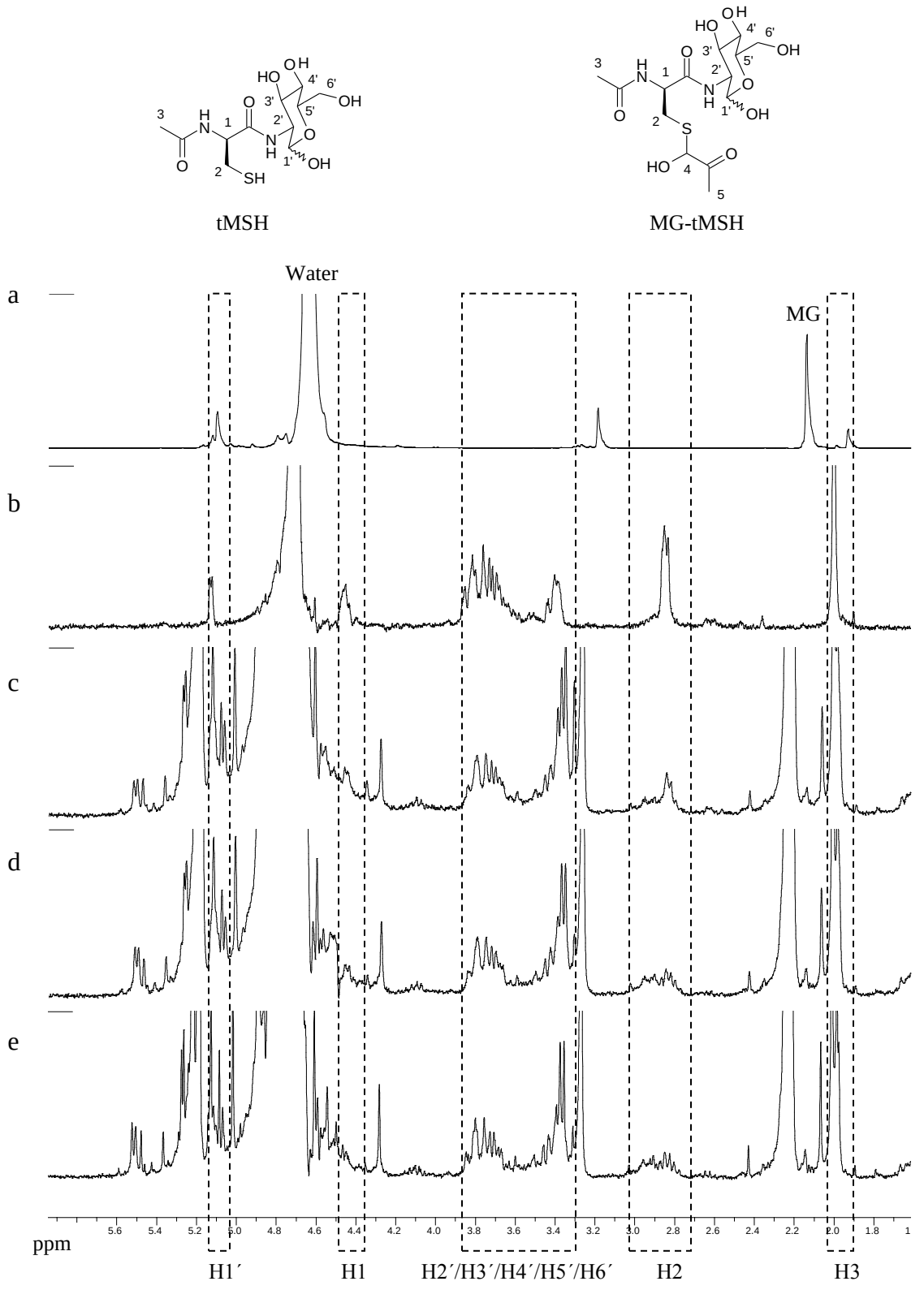
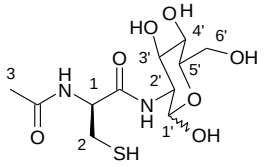
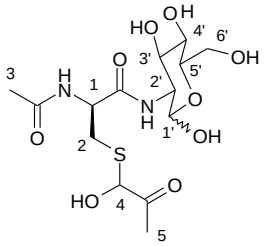


Figure 8.32: The ^{1}H NMR spectra of various incubation times for the formation of the hemithioacetal of MG-tMSH; (a) MG, (b) tMSH, (c) MG-tMSH incubated for 2 min, (d) MG-tMSH incubated for 15 min and (e) MG-tMSH incubated for 30 min.

Table 8.6: The summary for chemical shift (CS) and integration (int.) of NMR signal for MG-GSH equilibration time comparing to the signal of GSH (N/A = not available because the signal was obscured by water and MG peaks).





Peak	tMSH				MG-tMSH			
	†Predicted CS (ppm)	Actual CS (ppm)	Predicted int.	Actual int.	†Predicted CS (ppm)	Actual Cs (ppm)	Predict Int.	Actual int.
H1	4.39	4.50–4.43	1	0.93	4.37	5.51	1	1.37
H2	2.73–2.96	2.89–2.80	2	2.07	3.21	3.05–2.76	2	2.00
H3	2.06	2.00	3	3.00	2.06	2.04–1.96	3	9.37
H4	-	-	-	-	4.39	N/A	1	N/A
H5	-	-	-	-	2.19	2.35–2.24	3	8.73
H1'	4.88	5.14	1	0.61	4.88	N/A	1	N/A
H2'-H6'	3.52–4.16	3.89–3.35	6	7.15	3.52–4.16	3.89–3.23	6	22.55

†Chemical shifts are estimated by ChemBioDraw Ultra 12.0 (http://www.cambridgesoft.com) software.

Dissociation Constant of the Hemithioacetal

The dissociation constant (K_d) of the hemithioacetal was determined due to its significance in the calculation of the exact concentration of the substrate used in the GlxI kinetic assays. After allowing the hemithioacetal to reach its equilibrium (15 min), the dissociation constant could be measured using different concentrations of MG with a fixed concentration GSH. Even though K_d of MG-GSH was already reported [85, 86, 141, 279-281], the experiment was repeated to confirm the accuracy of our experimental protocols.

The dissociation profile of MG-GSH suggested its K_d of approx. 3.19 ± 0.29 mM (Figure 8.33A), which was in excellent agreement to the literature value (K_d of 3.1 mM) [85, 86]. The K_d of the hemithioacetal from the reaction of MG and tMSH was determined similarly using the equilibrium time that was previously investigated (30 min). The MG-tMSH dissociation profile suggested that the K_d of MG-tMSH was approx. 3.33 ± 0.41 mM (Figure 8.33B).

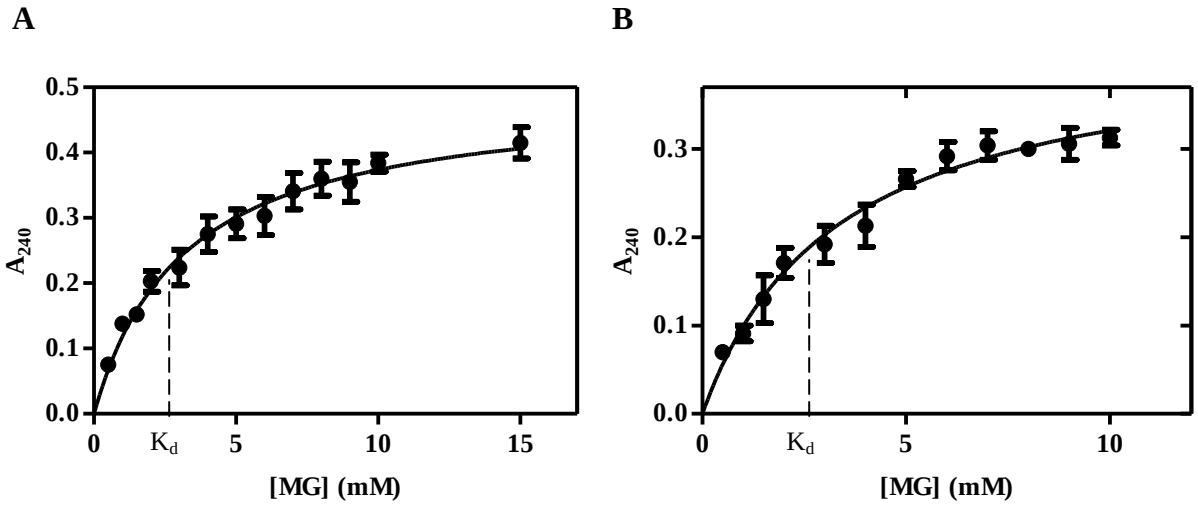


Figure 8.33: The determination of the dissociation constant (K_d) of **(A)** MG-GSH and **(B)** MG-tMSH. The experiments were performed in 50 mM KPB (pH 6.6) at room temperature.

Maximum Absorbance of Glyoxalase I Product

It has been previously reported that the increase in UV absorbance at 240 nm observed during the reaction of GlxI is due to the formation of the product, *S*-D-lactolyglutathione, from the GlxI reaction using MG-GSH as a substrate [85, 279, 281]. Our experiment on the monitoring of the GlxI reaction using the commercial yeast GlxI and MG-GSH suggested that the absorbance between 230–260 nm increases as a function of reaction time (Figure 8.34). Thus, the absorbance at 240 nm was chosen for general GlxI reaction kinetics studies since this wavelength could offer a more selective monitoring of the product from background contributions, which was especially the case when shorter wavelengths (below 230–240 nm) were used as interference from protein and peptide thiol that usually would occur.

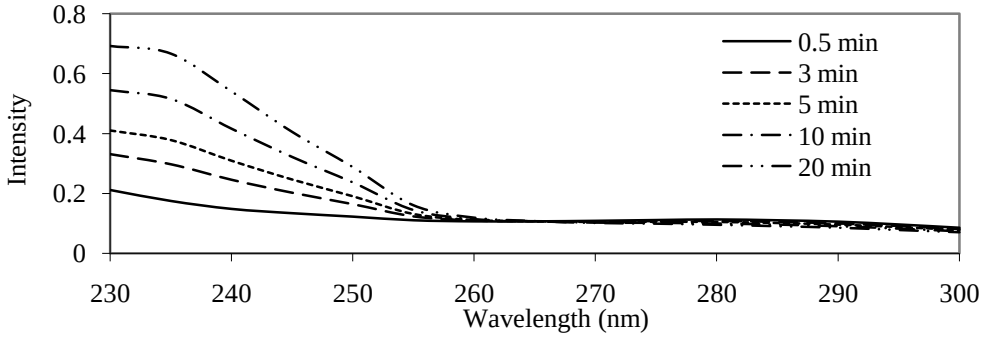


Figure 8.34: The wavelength scan for the reaction of yeast GlxI (0.15 μg in 300 μL assay) with MG-GSH (0.1 mM) in 50 mM KPB (pH 6.6) under different reaction times (0.5–20 min) showing the increase in signal between 230–260 nm due to product thioester formation.

The monitoring of the reaction kinetics of the mcGlxI reaction was similarly investigated. The wavelength scan of the mcGlxI reaction using Ni^{2+}-reconstituted PDO with MG-tMSH indicated that there was only a slight change in intensity when the chosen wavelength to monitor the reaction was greater than 260 nm. Since tMSH is likely not the real substrate for mcGlxI, the difference between the wavelength scans for various time periods were small (data not shown). However, the monitored profiles at various wavelengths suggested that the detection of the product, *S*-D-lactoyl-des-*myo*-inositol mycothiol, at 240 nm offered the most clearly defined signal (Figure 8.35).

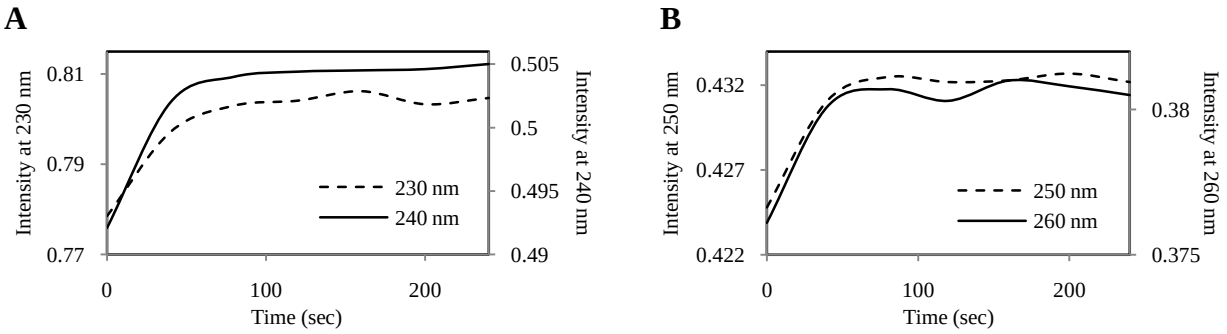


Figure 8.35: The UV detection at specific wavelengths including **(A)** 230 and 240 nm and **(B)** 250 and 260 nm shows the product formation from the reaction of Ni^{2+}-reconstituted PDO (4.5 μg in 300 μL assay) using MG-tMSH (0.5 mM) in 50 mM KPB (pH 6.6) as substrate.

Identification of S-D-lactoyl-des-myo-inositol mycothiol (mycoGlyoxalase I product)

The product *S*-D-lactoyl-des-*myo*-inositol mycothiol, produced from the mcGlxI catalyzed reaction using Ni^{2+}-reconstituted PDO with MG-tMSH, was isolated using reverse phase C18 HPLC (Figure 8.36). After comparing the HPLC chromatograms of the control experiments (MG, tMSH and the enzyme under the same concentrations as in the enzymatic reaction), the HPLC chromatogram of the assay reaction suggested that the product was possibly eluted as a peak having a retention time (t_R) of 54 min. The ^{1}H NMR and ESI-MS analyses on this isolated material was in agreement with the expected product therefore confirming that this peak contained the desired product with possible contamination of the substrates, MG and tMSH (Figure 8.37 and 8.38). Thus, this experiment confirmed that PDO serves as the targeted mcGlxI enzyme.

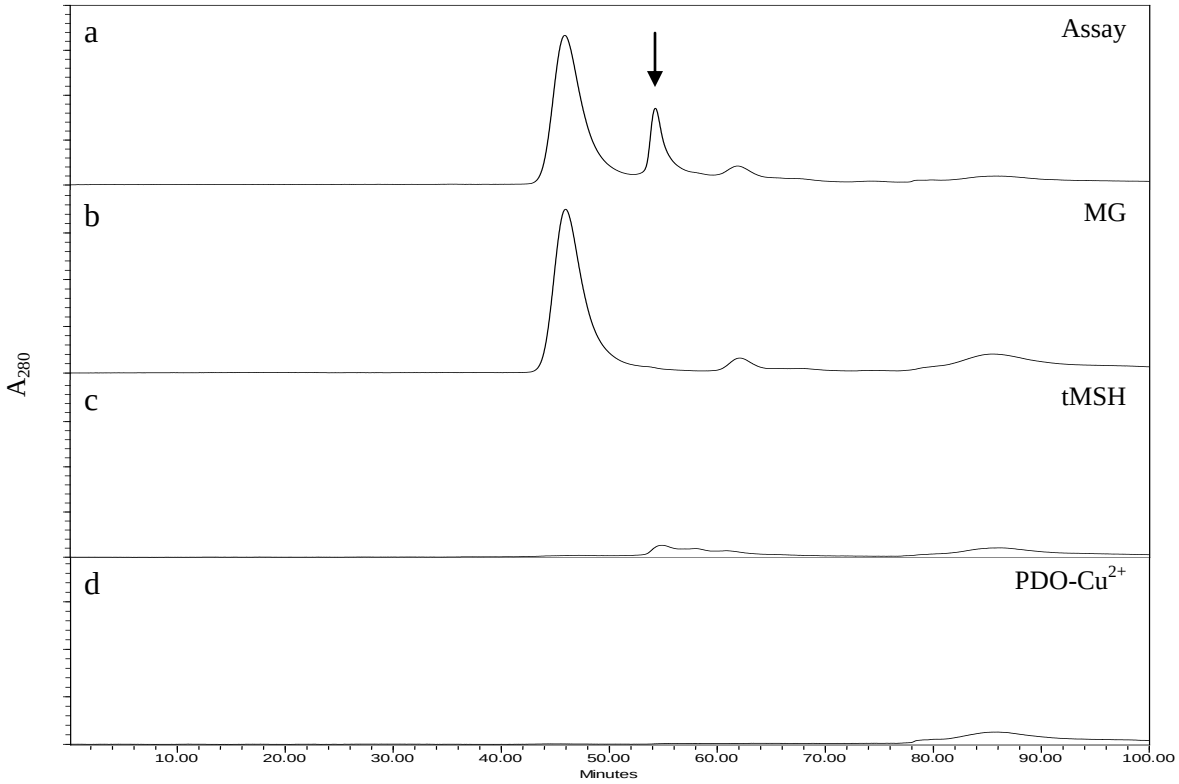


Figure 8.36: The reverse phase C18 HPLC chromatograms of (a) the assay reaction of Ni^{2+}-reconstituted PDO (25 μg) and MG-tMSH (5 mM) compared to those of (b) MG, (c) tMSH and (d) Ni^{2+}-reconstituted PDO that employed the same concentrations as in the reaction. The HPLC chromatograms were monitored at 280 nm and the product was eluted from the column using 1% increment of acetonitrile in water containing 0.1% TFA over 100 min interval with a flow rate of 1 mL/min. The eluted peak with t_R of 54 min contained the desired product (indicated by arrow) as confirmed by ESI-MS and ^{1}H NMR.

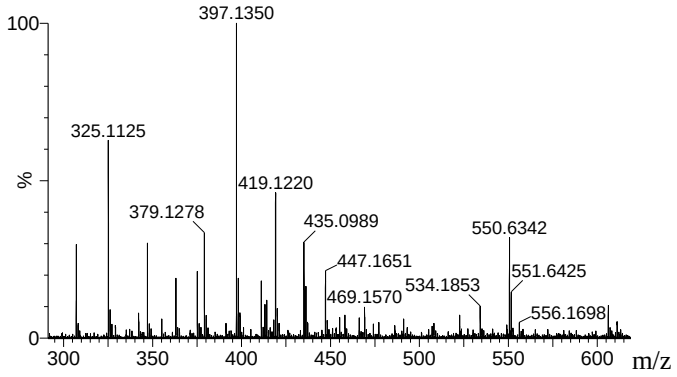


Figure 8.37: The electrospray mass spectrum of the purified *S*-D-lactoyl-des-*myo*-inositol mycothiol exhibits a molecular weight of 397.1 Da (calculated MW of 396.4 Da). Other signals include m/z 325 ($tMSH^+$, 60%), 379.1 ($tMSHK^+$, 30%), 419.1 (Na^+ adduct of *S*-D-lactoyl-des-*myo*-inositol mycothiol Na^+, 50%) and 435.1 (K^+ adduct of *S*-D-lactoyl-des-*myo*-inositol mycothiol, 30%) indicated contamination of the product, tMSH, and the salt adduct of the product. The signals at approximately 550 Da were from background.

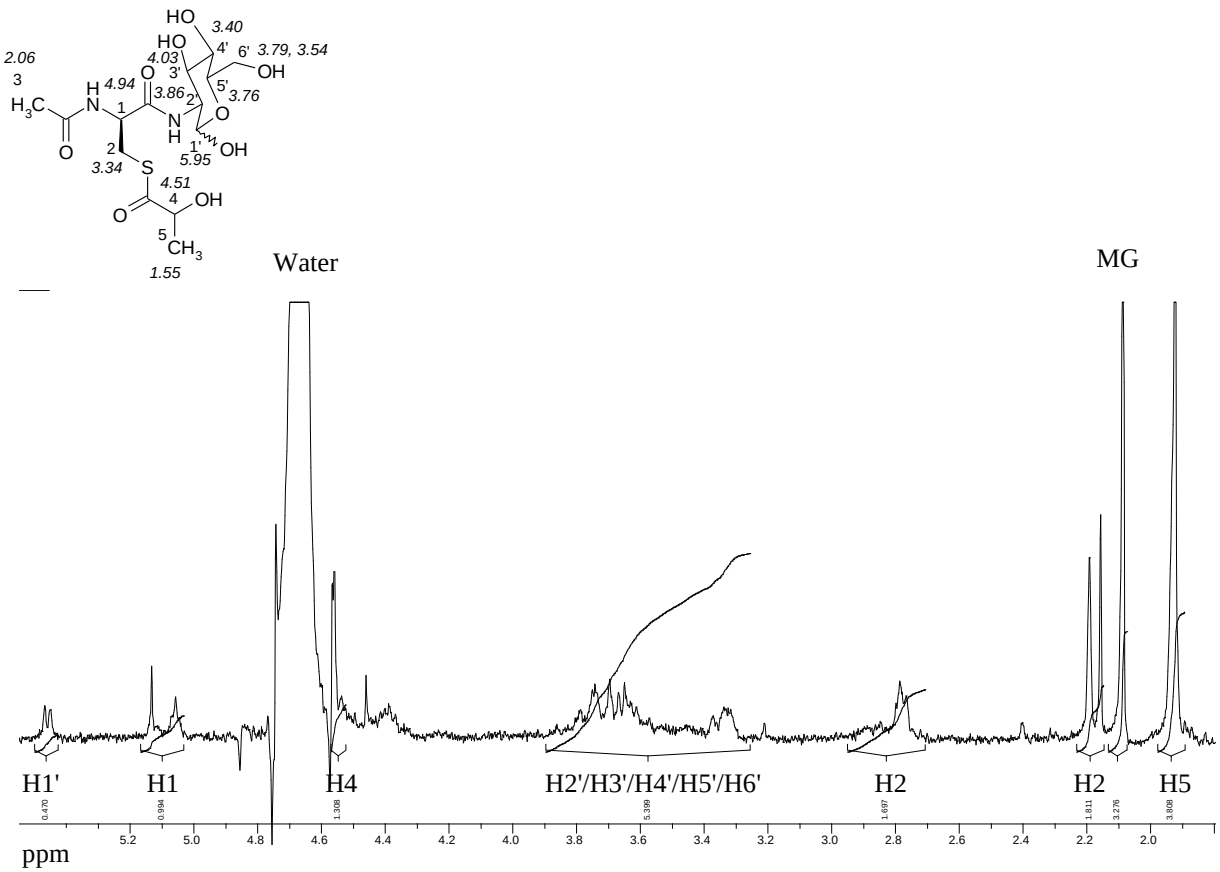


Figure 8.38: ^{1}H NMR spectrum of the purified *S*-D-lactoyl-des-*myo*-inositol mycothiol. The predicted chemical shifts (in the structure of the compound) are estimated by ChemBioDraw Ultra 12.0 (http://www.cambridgesoft.com) software.

Substrate Specificity

The substrate specificity of the GlxI and mcGlxI reactions was investigated in relation to metal-incorporation and types of enzymes (PDO and commercial yeast GlxI). The UV detection of the substrates (without addition of the enzyme) including MG, GSH, tMSH, hemithioacetal of MG-GSH and hemithioacetal of MG-tMSH did not show any significant increase in signal at 240 nm without the corresponding enzyme. Neither was observed in these substrates with an addition of Ni^{2+} atoms (without addition of the enzyme), suggesting that substrates with and without additional Ni^{2+} do not interfere with the detection of the enzyme reaction, if any. No activity was observed in the enzymatic reactions of MG, GSH, tMSH, MG-GSH with Ni^{2+}-reconstituted PDO. Neither were any reactions observed to occur in the reactions of apo-PDO and denatured PDO with MG-tMSH. However, trace activity was obtained in the reaction of native PDO with MG-tMSH, while high activity was observed in

the reaction of Ni^{2+}-reconstituted PDO with the same substrate. These preliminary data suggested that the incorporated metal into PDO is significant for the mcGlxI reaction and the "as isolated" PDO is not completely in its apo-form (confirmed by ICP-MS). This isolated enzyme might accept some metals from the organism's growth environment during the expression and purification processes undertaken in the laboratory. Additionally, the reaction catalyzed by the PDO enzyme is specific to the hemithioacetal substrate, MG-tMSH, while the substrate of the GSH-dependent GlxI reaction, MG-GSH, did not act as a substrate (monitoring at 240 nm and at a concentration of 0.5 mM) for PDO. Similarly, the enzymatic reaction catalyzed by yeast GlxI is specific to its substrate, MG-GSH, in which it exhibited no activity with MG-tMSH (at a concentration of MG-tMSH of 0.5 mM).

Metal Characterization

The ICP-MS analysis suggested that the purified PDO (0.088 mg/mL, 5.28 μM) binds to copper, cobalt, nickel and zinc ions (Table 8.7, more details in Appendix 4). The metal per enzyme ratios of nickel- and zinc-bound PDO were approximately 0.16–0.17, while those of cobalt and copper-bound enzyme were approx. 0.02. Thus, the ratios of Ni^{2+} and Zn^{2+} contents were approximately ten times higher than those of Co^{2+} and Cu^{2+}, which might suggest that PDO has higher metal binding affinity for Ni^{2+} and Zn^{2+} than Co^{2+} and Cu^{2+} ions. However, it was also possible that these metals are the results of bioavailability in the expression system under experimental environment. The concentrations of these incorporated metals in the "as isolated" protein were low and the enzyme appeared almost completely in a de-metallated form. ICP-MS analysis on this as isolated form of the PDO (without metal activation) indicated that the enzyme contained no iron, a common metal found in dioxygenase enzymes, but it did contain ions such as zinc and nickel. Both Zn^{2+} and Ni^{2+}, as previously mentioned, are usually associated with GlxI enzymes. These results suggest that PDO may bind either of these metals in intact *S. coelicolor*, although of course these metals could have been picked up by the protein in steps associated with the proteins isolation.

The metal analysis of as isolated PLA (0.076 mg/mL or 4.54 μM) exhibited similar results, where copper, nickel and zinc were detected bound to various copies of the protein (Table 8.7, more details in Appendix 4). The metal per dimeric enzyme ratios of nickel- and zinc-bound PLA were approximately 0.11–0.16, while the ratio of copper bound enzyme was

0.08 (Table 8.7). The concentrations of these metals were low, which indicated that the "as isolated" native PLA is isolated in predominantly its apo-form.

Table 8.7: The calculation for the amount of metals contained in PDO and PLA using ICP-MS analysis

Proteins	*Metals*	*MW (g/mol)*	*1st [Metal] (mg/L)*	*2nd [Metal] (mg/L)*	*Ave. [metal] (mg/L)*	*[Metal] (μM)*	*[Metal]/[Protein] ratio*
PDO	Cobalt	58.93	0.006	0.006	0.006	2.04	0.019
	Copper	63.55	0.005	0.005	0.005	1.57	0.015
	Nickel	58.69	0.053	0.051	0.052	17.72	0.168
	Zinc	65.39	0.057	0.056	0.057	17.28	0.164
PLA	Copper	63.55	0.024	0.023	0.024	0.37	0.082
	Nickel	58.69	0.03	0.028	0.029	0.49	0.109
	Zinc	65.39	0.047	0.046	0.047	0.71	0.157

The metal activation profile for PDO suggested that the enzyme was activated in the presence of Ni^{2+}, Cu^{2+}, Cd^{2+} and Ca^{2+} (the activity of PDO with $Ni^{2+} > Cu^{2+} > Cd^{2+} > Ca^{2+}$) (Figure 8.39A). The enzyme with incorporated Ni^{2+} exhibited the highest activity, while that of Cu^{2+}-reconstitution gave approx. 60% compared to that of the Ni^{2+}-reconstitution. The activities with Cd^{2+} and Ca^{2+}-were low but detectable (approx. 20%), while other metals hardly activated the enzyme. Thus, PDO could be a Glyoxalase enzyme in the non-Zn^{2+}-activation class. These results support our hypothesis, which predicts that PDO would fit better, by sequence comparisons, into the non-Zn^{2+}-activation class due to its shorter amino acid sequence (Figure 8.1). However, instead of being Ni^{2+}/Co^{2+} activated like other GlxI in this class, PDO is activated with Ni^{2+} and Cu^{2+} atoms. The unusual metal activation by Cu^{2+} in PDO may just be anomalous and not related to any physiological role of this metal in the organism, but it is interesting to speculate that it may well be an important metal for this enzyme in the organism since Gram-positive bacterium is a soil-organism. Streptomycetes are normally found in soil, which contains detectable amounts of copper [287]. Interestingly, due to the bioavailability of the copper ions, analysis of the *S. coelicolor* genome has revealed that the organism has genes coding for several putative copper binding proteins such as lipo-cupredoxin [288]. Thus, it appears reasonable that this enzyme might utilize Cu^{2+} in the cell. But this requires further work to confirm or dispute.

It had been found earlier that isolated mammalian GlxI exhibited high activity in the presence of Mg^{2+}, and it was thought to be a magnesium activated enzyme *in vivo* [289-292]. However, the naturally bound metal was later discovered to be zinc, which was incorporated at a ratio of 1 mole metal per mole enzyme subunit [293]. The magnesium atoms were believed to be recovered from protein preparation processes and had no correlation with enzymatic activity [293]. The analysis by electron paramagnetic resonance **(**EPR) of magnesium bound human GlxI, however, suggested an octahedral metal coordination with four metal binding protein residues and two water molecules [31]. Yeast GlxI was also found to be partially activated by the presence of Mg^{2+} and Ca^{2+} [34, 151, 293]. It was possible that these metals could reactivate the enzyme but play a major role in protein stability. As well, X-ray crystallographic structure of *E. coli* GlxI with bound Cd^{2+} (PDB: 1FA7) suggested that the enzyme could be activated with this metal and its metal coordination forms an octahedral geometry [45]. However, the activity was low compared to the enzyme with Ni^{2+} and Co^{2+}, which might be related to the size of the metal that fits into the active site of the enzyme (more details in Chapter 3). Thus, the low activity observed in PDO with Ca^{2+} and Cd^{2+} might behave in a similar fashion.

Metal titration studies of PDO with Ni^{2+} and Cu^{2+} indicated that 1 mole of metal per mole dimeric enzyme could optimize enzymatic activity (Figure 8.39B), suggesting a tight binding of the metal as well as one functional active site (possibly two active sites per dimeric enzyme). These data are consistent with previous reports on other GlxI (such as *E. coli* GlxI and *Plasmodium falciparum* GlxI), which also exhibit a metal per dimeric enzyme ratio of approximately one [19, 39]. Additionally, the pH profile of PDO suggested that the enzyme could be optimally active in a buffer pH of 7.0 (Figure 8.39C).

The mcGlxI activity of PLA, on the other hand, was undetectable under studied conditions. The enzyme exhibited no activity with MG-tMSH nor MG-GSH (data not shown). Neither activity was observed in the presence of different divalent metals ($ZnCl_2$, $NiCl_2$, $CoCl_2$, $CuCl_2$, $MnCl_2$, $MgCl_2$, $CaCl_2$ and $CdCl_2$), suggesting that PLA does not function as a metalloenzyme Glyoxalase I.

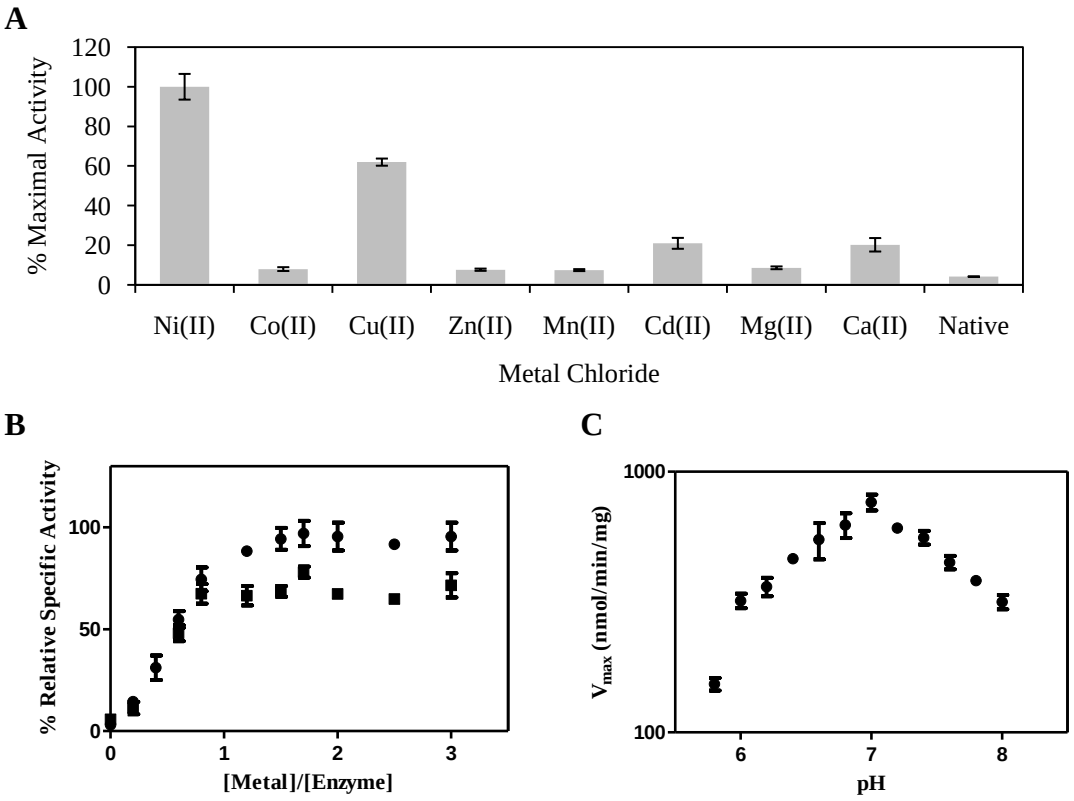


Figure 8.39: The diagram shows **(A)** metal activation profile **(B)** the metal titration with Ni^{2+} (●) and Co^{2+} (■) and **(C)** the pH profile of PDO (3.125 μg in 200 μL assay) with the substrate MG-tMSH (0.5 mM, K_d = 3.3 mM) that was incubated for 30 minutes prior to performing the enzyme assay in 50 mM KPB (pH 6.6) at room temperature (25 °C).

Kinetic Studies

The kinetics of the enzymatic reaction of Ni^{2+}-reconstituted PDO with MG-tMSH yielded a calculated V_{max} of 5.8 ± 1.0 μmol/min/mg, K_m of 1.25 ± 0.13 mM and k_{cat} of 3.2 s^{-1}. The K_m is relatively high compared to other GlxI [17, 19, 32-34, 50, 151, 294, 295], which is probably due to the structure of MG-tMSH being somewhat different from the chemical structure of the real thiol found in Streptomyces, mycothiol, which contains an additional inositol. Thus, further experiments on the kinetic studies of PDO with the real substrate, MG-GSH, are required.

8.4.5: Enzymatic Activity of Methylmalonyl-CoA Epimerase

Due to the close structural relationship between methylmalonyl-CoA epimerase (MMCE) and GlxI, an investigation of possible MMCE activity by PDO and PLA with selected incorporated metals (Ni^{2+}, Co^{2+}, Cu^{2+}, and Zn^{2+}) was performed. Ni^{2+}, Co^{2+} and Zn^{2+}

are general cofactors found in GlxI enzymes, while Co^{2+} is normally found in MMCE enzymes [216, 217]. Cu^{2+} was also chosen for our studies due to its high activation of PDO in the mcGlxI reaction. Besides, since PDO exhibits relatively low mcGlxI activity, it appeared possible that it might be a bifunctional enzyme that possesses activity as both a GlxI and a MMCE enzyme. However, no MMCE activity was observed with PDO under various conditions (data not shown), suggesting that the enzyme only functions as a mcGlxI and not as a MMCE. PLA with no observable GlxI activity exhibited no activity in the MMCE reaction, suggesting that this enzyme is neither a GlxI nor a MMCE.

To support our previous hypothesis on the similar MSH-dependent Glyoxalase system in *Streptomyces coelicolor* and *Mycobacterium tuberculosis*, the database search for putative mycoGlyoxalase I in *Mycobacterium tuberculosis* H37Rv genome (http://www.tbdb.org/) by using blastp with amino acid sequence of PDO suggested the gene Rv0274 (582 nucleic acids and 193 amino acids) with ~6% sequence similarity and ~15% sequence identity with GlxI from other organisms. This putative protein contains all four conserved metal binding residues and might behave as mycoGlyoxalase I, in which its metal characteristic and enzymatic activity should be further examined to confirm its biological function. Thus, our research takes advantage of working on the non-pathogenic bacterial strain, *S. coelicolor*, which possesses a similar Glyoxalase system.

Figure 8.40: The multiple sequence alignment of putative mycoGlyoxalase I from *Mycobacterium tuberculosis* (Rv0274) with Glyoxalase I from various organisms (organism name follows by National Center for Biotechnology Information (NCBI) accession number) including *E. coli* (NP_310387), *N. meningitidis* (CAA74673), *P. aeruginosa* (AAG06912), *L. donovani* (AAU87880), *T. cruzi* (XP_818456), *H. sapiens* (AAB49495), *P. putida* (AAN69360) and mycoGlyoxalase I from *S. coelicolor* (PDO). The metal binding residues are marked with a grey highlight. The alignment was created using CLC Free Workbench (version 3.0.1) with the accurate alignment algorithm (http://www.clcbio.com).

8.5: Conclusion

The putative mcGlxIs, PDO and PLA, were cloned from *S. coelicolor* DNA into the pET-28b(+) expression vector using PCR and restriction enzymes. The proteins were then expressed in *E. coli* and purified using HisTrap HP and Benzamidine FF affinity columns under standard protocols. Protein stability was investigated by CD analysis, fluorescence spectroscopy and ^{1}H NMR. Analysis of the results of these biophysical approaches suggested that the proteins were stable under certain protein concentrations in buffer of pH of 6–7. These proteins had a high propensity to form oligomers in HEPES, MOPS and MOPSO, while KPB and Tris were better buffers to store these proteins. Salt and glycerol had no major effect on protein stability. Additionally, the secondary structural investigation by CD suggested that

both PDO and PLA contain predominant β-sheets, while gel permeation chromatography analysis indicated that the native forms (quaternary structures) of the proteins are dimers. Further investigation by X-ray crystallographic analysis would be required to support our preliminary data as well as extend our knowledge and understanding of these protein structures.

The investigation on the enzymatic reaction suggested that the formation of the hemithioacetal substrate, MG-tMSH, reached equilibrium in 30 minutes and the dissociation constant for this equilibrium was determined to be approximately 3.3 mM. The optimal detection wavelength for the formation of the product, *S*-D-lactoyl-des-*myo*-inositol mycothiol, from the mcGlxI reaction was at 240 nm. The existence of this product was confirmed by ^{1}H NMR and ESI-MS methods.

Studies of the enzymatic reaction of PDO found that the enzyme is activated with Ni^{2+}, Cu^{2+}, Cd^{2+} and Ca^{2+}. Optimal catalytic activity was observed with Ni^{2+} and Cu^{2+}. The metal titration with 1 mole of metal per mole dimeric enzyme ratio suggested tight binding of the metal and the existence of one possible functional active site. The optimum pH for mcGlxI activity was determined to be at pH 7.0. The kinetics of Ni^{2+}-reconstituted PDO with MG-tMSH was also estimated. However, further investigation on kinetics with the real substrate, MG-MSH, is also required.

On the other hand, no GlxI activity was detected in the reaction using PLA and MG-tMSH in the presences of divalent metals ($ZnCl_2$, $NiCl_2$, $CoCl_2$, $CuCl_2$, $MnCl_2$, $MgCl_2$, $CaCl_2$ and $CdCl_2$). No reaction was observed in the enzyme assay using a different substrate (MG-GSH). These data suggest that PLA does not function as a GlxI enzyme.

Additionally, the MMCE activity investigation using metal-reconstituted PDO and PLA showed no activity with Ni^{2+}, Co^{2+}, Cu^{2+}, and Zn^{2+} under the studied conditions, suggesting that PDO and PLA do not function as an MMCE enzyme.

CHAPTER 9: PUTATIVE MYCOGLYOXALASE II FROM *STREPTOMYCES COELICOLOR*

9.1: Introduction

The first enzyme in the Glyoxalase system, Glyoxalase I, converts the hemithioacetal, the non-enzymatic product formed between methylglyoxal (MG) and a thiol cofactor such as glutathione (GSH), to *S*-D-lactolyglutathione (SDL). This compound is then hydrolyzed by the second enzyme of the system, Glyoxalase II (GlxII), to D-lactate and regenerates the cofactor, GSH, in the presence of water. GlxII is a metalloenzyme with binuclear active sites per monomer that can be activated by various metals depending on the particular source organism [63, 67, 72-74]. This property has been compared to the related metallo-β-lactamase group of enzymes where mononuclear metal ion-containing active sites are observed [71, 72]. Unlike GlxI, GlxII does not appear to possess two distinct metal activation classes [63, 67, 72-74]. Even though there are two metal binding sites in GlxII, an investigation on the *Salmonella typhimurium* enzyme has suggested that only one metal atom is catalytically active and the second metal might not be necessary for catalysis [63]. These results are further supported by the discovery of the properties of the GlxII from *Salmonella enteric* serovar Typhimurium (YcbL), where this GlxII possesses only one metal binding site with optimal activity being observed in the presence of iron [68] and the human enzyme also being reported as active in its mono-metallated Zn^{2+} form [76].

Studies of GlxII from almost all organisms have been reported [47-62]. However, none appear to have been characterized from any Gram-positive bacteria, especially from the bacteria that produce higher levels of mycothiol (MSH) than glutathione (GSH). This chapter, thus, is devoted to the study and discovery of a putative mycothiol-dependent Glyoxalase II or mycoGlyoxalase II (mcGlxII) from the Gram-positive bacterium, *Streptomyces coelicolor*, which is continued from the previous chapter (Chapter 8) that concerns a putative mycothiol-dependent Glyoxalase I or mycoGlyoxalase I (mcGlxI) from the same organism. The sequence searching analysis for the putative mcGlxII from *S. coelicolor* on the StrepDB-The Streptomyces Annotation Server (http://strepdb.streptomyces.org.uk) revealed a highly related protein to GlxII and was annotated as a putative hydrolase (PHL, SCO3572). The analysis of a multiple sequence alignment of PHL with GlxII from other organisms including *Escherichia*

coli (AAC73317), *Homo sapiens* (CAA62483), *Arabidopsis thaliana* cytoplasmic (NP_187696), *Arabidopsis thaliana* mitochondrial precursor (NP_565999), *Plasmodium falciparum* pre-protein (AAQ05976), isomer of *Saccharomyces cerevisiae* GlxII including GLOA4 and GLOA2 (CAA58065 and CAA71335, respectively), suggested that PHL contains all conserved metal binding residues (His^{63}, His^{65}, Asp^{67}, His^{68}, His^{118}, Asp^{137} and His^{180}) and the THxHxDH active site motif, a characteristic of the metallo-β-lactamase superfamily (Figure 9.1). This putative protein, however, does not contain conserved substrate binding residues (Thr^{144}, Val^{146}, Pro^{182}, Gln^{242} and His^{249} in PHL) unlike GlxII from other organisms, which mostly consist of hydrophilic amino acids. A comparison with the substrate binding residues found in human GlxII (Lys^{143}, Tyr^{145}, Tyr^{175}, Arg^{249} and Lys^{252}), the hydrophobic residues found in PHL might suggest a role in binding the cosubstrate mycothiol rather than glutathione, thus accounting for the need for a different set of protein residues. The sequence identity and similarity of PHL with other GlxII are in the range of 4–10% and 10–17%, respectively. From the multiple sequence alignment, we hypothesize that PHL could function as mcGlxII, using *S*-D-lactoylmycothiol as substrate. As mention earlier (Chapter 1 and Chapter 6), inhibitors targeting pathogen GlxII over human GlxII might also be effective from a medicinal chemistry perspective. Therefore, it may be worthwhile to investigate the putative mcGlxII to extend our knowledge and provide an alternative pathway for structure-based drug design.

In order to investigate the mcGlxII reaction, the putative mcGlxII enzyme (PHL) was cloned from the *S. coelicolor* genome into the pET expression vectors utilizing the polymerase chain reaction followed by application of various restriction enzymes. The protein was overproduced in *E. coli* cells and the resulting protein was purified using different methodologies to optimize protein recovery. As well, the initial structural investigation of PHL was performed using circular dichroism and gel permeation chromatography. Due to the lack of suitable quantities of MSH, the kinetic studies and enzymatic assay of mcGlxII have yet to undertaken.

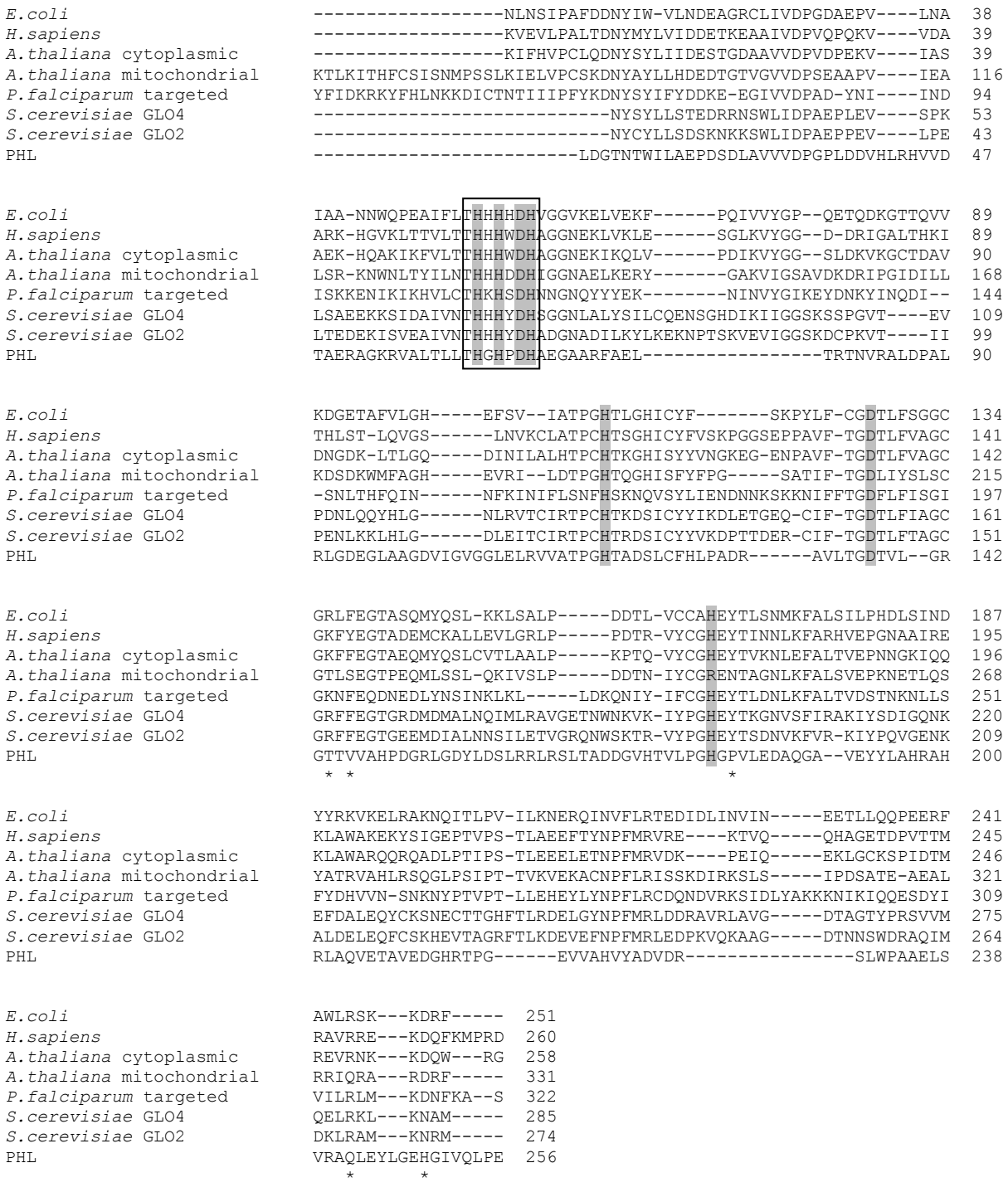


Figure 9.1: The multiple sequence alignment of putative hydrolase (PHL, CAB45552) from *S. coelicolor* compared with GlxII from other organisms (organism name follows by National Center for Biotechnology Information (NCBI) accession number) including *E. coli* (AAC73317), *H. sapiens* (CAA62483), *A. thaliana* cytoplasmic (NP_187696), *A. thaliana* mitochondrial precursor (NP_565999), *P. falciparum* (AAQ05976), *S. cerevisiae* GLOA4 (CAA58065) and *S. cerevisiae* GLOA2 (CAA71335). The metal binding residues are highlighted in grey and the box indicates the THxHxDH active site motif. The substrate binding residues are marked with an asterisk. The alignment was created using CLC Free Workbench (version 3.0.1) with the accurate alignment algorithm (www.clcbio.com).

9.2: Reagents, Materials and Instrumentation

All reagents, materials and equipment used in the experiments in this chapter are listed below otherwise they are included in the previous chapters. The designed primers for DNA cloning were received from Sigma (Sigma-Genosys Canada, Oakville, ON). The enzyme AcTev™ protease was purchased from Invitrogen Corporation (Carlsbad, CA).

9.3: Experimental Protocols

9.3.1: DNA Cloning and Protein Purification

The cloning of the putative hydrolase-encoding gene (*phl*) was performed using three different procedures according to strategies that were required for purification purpose. The first approach was to generate DNA that coded for the N-terminal His_6-Tag followed by a thrombin protease cleavage site and then the protein of interest. This method provided the most convenient purification pathway using HisTrap HP and HiTrap Benzamidine FF affinity columns as previously described for the putative dioxygenase (PDO) and putative lyase (PLA) enzymes in Chapter 8. However, PHL was found to contain an internal thrombin protease cleavage site (VLTGDT), thus an alternative method to purify PHL was developed by altering the thrombin protease cleavage sequence to a Tev protease cleavage site that has a different recognition sequence (ENLYFQS). The protein overproduced from the plasmid generated by this method, however, was recovered in inclusion bodies, suggesting that proper protein folding might be interfered by the additional Tev protease cleavage site. In order to avoid this matter, the third attempt was to generate a non His-tagged PHL encoding DNA. All three methods on DNA cloning and protein purification of PHL are described in detail as follows.

PHL with the N-terminal His_6-tag and thrombin protease cleavage site

The *phl* gene was cloned into the pET-28b(+) expression vector using the polymerase chain reaction (PCR) and restriction enzymes, *Nde*I and *Eco*RI, to generate a plasmid coding for PHL with the N-terminal His_6-tag followed by a thrombin protease cleavage site. The forward and reverse primers were designed as in Table 9.1. The PCR reagents and thermal cycles were prepared as previously described (Table 8.2 and 8.3, Chapter 8) with the exception of annealing temperature (63 °C) and elongation time (1 min). The circular plasmid was

amplified by heat shock transformation into a competent *E. coli* DH5α bacterial cell line and grown on agar plates containing kanamycin (30 μg/mL) overnight at 37 ºC. A single colony was picked and inoculated overnight in LB (5 mL) containing Kan (30 μg/mL) at 37 ºC in a shaker at 220 rpm. The plasmid was purified using a QIAprep Spin Miniprep Kit and sent for sequencing when appropriate.

Table 9.1: The forward and reverse primers designed for DNA cloning of *phl* gene from *Streptomyces coelicolor* into the pET-28b(+) expression vector.

Primer name	***Sequence of primer (5' to 3')***
PHLNdeI.For[Ψ]	CCGAAGCTT**CATATG**AACGTTCTCGCGCCC[£]
PHLEcoRI.Rev[Ψ]	GGCGAATTCGGATCCTACTCGGGCAGCTGGACG[Ц]

[Ψ] T_m for PHLNdeI.For primer and PHLEcoRI.Rev primer are 66 and 71 ºC, respectively.
[£] The bold letters indicate *Nde*I cutting sites and the underlined letters code for the N-terminus of the protein.
[Ц] The grey highlights indicate *EcoR*I cutting sites and the underlined letters code for the C-terminus of the protein.

The pET-28b-thrombin-*phl* plasmid was heat shock transformed into *E. coli* BL21 (DE3) for protein overproduction under the studied conditions (growth and expression at 25 °C and induction in the presence of 0.5 mM ITPG for four hours). However, after cell wall disruption by sonication, the desired protein (PHL) was recovered in the inclusion body fraction. The plasmid was then transformed into *E. coli* BL21 (DE3) pLysS. This cell line carries a low level of gene expression for T7 lysozyme, which binds to T7 RNA polymerase that can inhibit transcription. Thus, a low level of DNA can be produced and a slow production of protein can be achieved. This system also supports a slower process of protein folding, which in turn leads to a more accurate folding. The optimized conditions for protein expression were investigated in terms of induction and growth temperatures, amount of inducer (IPTG) and the presence of $ZnCl_2$ (Table 9.2). The addition of $ZnCl_2$ might support proper protein folding, since zinc is a frequently found metal that is incorporated into GlxII by most organisms although completely folded protein likely exists before the metalation process occurs [63, 67, 72-74].

Table 9.2: The induction tests of His-tagged PHL for optimum expression level in *E. coli* strains BL21 (DE3) and BL21 (DE3) pLysS including varied concentrations of IPTG, $ZnCl_2$ and growth and induction temperatures. The numbers in the box represent the concentration of $ZnCl_2$.

IPTG (mM)[d] ***Growth/ Induction Temp. (ºC)***	***0.1***	***0.2***	***0.5***	***1.0***	
37/15	0	0	0	0	***E. coli BL21 (DE3)***
37/25	0	0	0	0, 1	
37/37	0	0	0	0	
25/25	1	1, 2, 5	1, 2, 5	1, 2, 5	
15/15	0	-	0	0	***E. coli BL31 (DE3) pLysS***
25/25	0	-	0, 1	0, 1	
37/37	0	-	0	0	
37/15	0	-	0	0	
37/25	0	-	0	0	
	[$ZnCl_2$] (μM)				

The protein production was performed under optimized conditions as investigated above. *E. coli* BL21 (DE3) pLysS containing the pET-28b-thrombin-*phl* plasmid was inoculated in LB (10 mL) containing Kan (30 μg/mL) overnight at 37 °C. The culture was then transferred into 1 L LB containing $ZnCl_2$ (1 μM), Kan (30 μg/mL) and chloramphenicol (34 μg/mL) and shaken in a shaker (220 rpm) at 25 °C until OD_{600} reached 0.6. Proteins were then induced with IPTG (0.5 mM) at 25 ºC for 4 hours. The cell pellet was harvested by centrifugation at 6,000×*g* for 10 min before storing at -80 °C.

The purification of PHL was similar to those undertaken for the proteins PDO and PLA (Chapter 8) using HisTrap HP (1 mL) and HiTrap Benzamidine FF (1 mL) affinity columns. However, after cleavage of the N-terminal His_6-tag by thrombin protease (overnight at 4 °C), no protein with a molecular weight of PHL (approximately 29 kDa) was detected by the SDS-PAGE experiment. At first, it was suspected that PHL was overdigested by thrombin protease, thus creating nonspecific cleavages that caused the disappearance of PHL. A test for a suitable digestion time for thrombin protease was then performed using His-tagged PHL (100 μg protein) with incubation with thrombin protease (1 μL/100 μg protein) for various time periods (0.5, 1, 2, 3, 4, 5, 6, 7, 8, 9 and 14 hours). The protein was visualized by SDS-PAGE and it was actually found that PHL possesses an internal thrombin cleavage site. Therefore, this protocol was discontinued.

PHL with the N-terminal His-tag and Tev protease cleavage site

The purpose of this method was to replace the thrombin protease cleavage site with a Tev protease cleavage site in the pET-28b(+) expression vector using PCR (two-stage PCR method similarly to the DNA cloning of mutagenesis of *Pseudomonas aeruginosa* GloA3, Chapter 3) (Figure A1.4, Appendix 1). The first PCR used the pET-28b(+) expression vector with T7.For and Tev.Rev primers to generate the plasmid with a *Xba*I digestion site and a Tev protease cleavage site. The second PCR used the pET-28b-thrombin-*phl* plasmid (from previous DNA cloning) with Tev.For and T7.Rev primers to generate a plasmid containing an N-terminal Tev protease cleavage site followed by the *phl* gene and an *Eco*RI restriction site. Both PCR products were purified using a 1% agarose gel and a QIAquick Gel Extraction Kit. The third PCR combined the previous two PCR products using T7.For and T7.Rev primers to generate a plasmid coding for PHL with the N-terminal His_6-tag followed by the Tev protease cleavage site. The primers were design as in Table 9.4 and the PCR reagents were prepared as in Table 9.3. The thermal cycles were set similarly to Table 8.3 in Chapter 8. The third PCR product was then cloned back into the pET-28b(+) expression vector using *Xba*I and *Eco*RI restriction enzymes and transformed into the competent *E. coli* DH5α. The plasmid was purified using a QIAprep Spin Miniprep Kit and sent for sequencing when appropriate.

Table 9.3: The recipe for the two-stage PCR protocol used to clone *phl* with the Tev protease cleavage site into the pET-28b(+) expression vector.

Plasmids/Reagents	***Stock Concentration***	***Amount Used (μL)***	***Final Concentration***
DNA template[£]	1 ng/μL	1	1 ng
Primer.For[§]	20 μM	1	0.4 μM
Primer.Rev[§]	20 μM	1	0.4 μM
dNTP	10 mM	1	0.2 mM
pfu Buffer	10×	5	
DMSO*	100 %	2.5	5 %
ddH_2O*		40	
pfu polymerase		1	
Total		50.0	

[£] DNA template for the first PCR reaction was pET-28b(+) vector, for the second PCR reaction was pET-28b-thrombin-*phl* and for the third PCR reaction was the purified products from the first and the second PCR reactions.

[§] The primers for the first PCR reaction were T7.For and Tev.Rev, for the second PCR reaction were Tev.For and T7.Rev and for the third reaction were T7.For and T7.Rev.

* DMSO was only added to the second and the third PRC reactions, in which the amount of ddH_2O was changed to 37.5 and 36.5 μL, respectively.

Table 9.4: The forward and reverse primers for the two-stage PCR protocol used to generate the plasmid that coded for PHL with the N-terminal His-tag and Tev protease cleavage site.

Primer name	*Sequence of primer (5' to 3')*
T7.For*	TAATACGACTCACTATAGG
T7.Rev*	TATGCTAGTTATTGCTCAG
Tev.For[£]	AGCAGCGGCGAGAACCTGTACTTCCAGAGCCATATG
Tev.Rev[£]	CATATGGCTCTGGAAGTACAGGTTCTCGCCGCTGCT

* T_m for T7.For and T7.Rev primers are 45 ºC.
[£] T_m for Tev.For and Tev.Rev primers are 69 ºC.

The plasmid was transformed into *E. coli* BL21 (DE3) CodonPlus RP and the cells were grown on an agar plate containing Kan (30 μg/mL) and chloramphenicol (34 μg/mL). This *E. coli* cell line has been commercially developed in order to express proteins from genes that consist of high GC content (especially from Gram-positive bacteria). Among the codons that encoded for amino acids, the arginine codon (AGA and AGG) and the proline codon (CCC) are rarely found in *E. coli*, thus leading to the delay in RNA translation, codon substitution and misincorporation. *E. coli* BL21 (DE3) CodonPlus RP contains a ColE1-compatible plasmid coding for extra *argU* and *proL* tRNA genes; therefore, allowing DNA to translate and providing better protein expression compared to a normal cell strain.

An investigation for optimal protein expression was performed as previously described with various induction temperatures (15, 20, 25 and 30 °C), amount of IPTG (0.1 and 0.5 mM) and the presence of $ZnCl_2$ (1 μM). Under the optimal conditions, the cells were inoculated in TB (5 mL) overnight at 37 °C before transferring into 1 L TB containing $ZnCl_2$ (1 μM), Kan (30 μg/mL) and chloramphenicol (34 μg/mL). The culture was grown at 25 °C in the 220 rpm shaker until an OD_{600} of 0.6 was reached. Protein production was induced with IPTG (0.5 mM) at 15 °C for 4 hours. The cell pellet was harvested by centrifugation at 6,000×*g* for 10 min and stored at -80 °C.

The protocols for protein purification were similar to those of PHL with the thrombin cleavage site as previously described. However, PHL was recovered in inclusion bodies after disrupting cells by sonication. In an attempt to recover the protein, urea (6 M) was added to the inclusion bodies and stirred for an hour at 4 °C. The mixture was then centrifuged at 48,384×*g* for 20 min and the supernatant was loaded onto a HisTrap HP affinity column (1 mL) utilizing a buffer containing 20 mM Tris (pH 7.0), 500 mM KCl, 20 mM imdazole, 6 M

urea and 10% (v/v) glycerol. The denatured His-tagged proteins were eluted with the same buffer but containing 500 mM imidazole. Removal of urea and imidazole were performed by dialysis against buffer containing 20 mM Tris (pH 7.0), 500 mM KCl and 10% glycerol (dialysis buffer) overnight at 4 °C. Precipitation was still observed after dialysis, so the dialyzed sample was centrifuged at 48,384×*g* for 20 min before loading onto another HisTrap affinity column utilizing a similar buffer as previously described but without the presence of 6 M urea. The eluted fraction was dialyzed against dialysis buffer overnight at 4 °C. The protein concentration was determined by the Bradford assay using BSA as a standard. DTT (5 mM) and Tev protease was added into the dialyzed sample in a 100:1 ratio before incubating overnight at 4 °C. The mixture was then loaded onto the HisTrap HP affinity column to isolate PHL from the Tev protease. His-tagged PHL, however, was not cleaved (as indicated by SDS-PAGE experiment). Thus, a suitable buffer condition for the active Tev protease should be further examined. The purification protocol is summarized in Figure A2.4 (Appendix 2).

Non His-tagged PHL

This method involved the cloning of *phl* into the pET-22b(+) expression vector using PCR and the restriction enzymes, *Nde*I and *Eco*RI, to generate a gene coding for non His-tagged PHL. The PCR was performed similarly to the cloning of *phl* into the pET-28b(+) expression vector, where pET-28b-thrombin-*phl* as well as PHLNdeI.for and PHLEcoRI.rev (Table 9.1) were used as DNA template and primers, respectively. The cloning protocol is summarized in Figure A1.5 (Appendix 1). The plasmid was purified using a QIAprep Spin Miniprep Kit and sent for sequencing when appropriate.

The circular DNA was heat shock transformed into *E. coli* BL21 (DE3) and BL21 (DE3) CodonPlus RP, which were then subsequently grown on agar plates containing antibiotic(s) (Amp (50 μg/mL) for strain BL21 (DE3) and Amp (50 μg/mL) and chloramphenicol (30 μg/mL) for strain BL21 (DE3) CodonPlus RP) overnight at 37 ºC. Single colonies from both strains of *E. coli* cells containing the desired plasmids were picked and inoculated overnight with LB (5 mL) containing their corresponding antibiotic(s) in a 37 ºC incubator (shaken at 220 rpm). The culture (1 mL) was transferred into 50 mL LB and was shaken until an OD_{600} of 0.6 at 37 ºC was reached. To optimize the protein production, the induction test was performed by varying growth and induction temperatures. The protein was

induced at either 25 or 37 °C by adding 1 mM IPTG and continued shaking at the same speed. An aliquot (1 mL) was taken every hour for 4 hours. The cell pellets from these aliquots were collected by centrifugation at 13,000×*g* for 1 minute before performing SDS-PAGE.

Under optimized conditions, pET-22b-*phl* in *E. coli* BL21 (DE3) CodonPlus RP was inoculated overnight in TB (50 mL) in a 37 ºC shaker (shaken at 220 rpm). The culture was then transferred into 1 L TB, then allowed to grow at 25 ºC with shaking (220 rpm) for 4 hours (or until OD_{600} reached 0.6) with the presence of $ZnCl_2$ (1 mM). The proteins were overproduced by addition of 0.5 mM IPTG followed by continued shaking at the same speed and temperature for 4 hours. The cell pellet was harvested by centrifugation at 6,000×*g* for 10 min before being flash frozen in liquid nitrogen and being stored at -80 ºC.

The protein purification of PHL was performed similarly to the conditions reported for the purification of the *E. coli* GlxII enzyme by O'Young *et al*., 2007 [67] with some changes as follows. The frozen cell pellet was resuspended in 50 mM Tris (pH 8.0) and 10% glycerol with the addition of 1 mM PMSF. Cell wall was disrupted by sonication before centrifugation at 48,384×*g* for 20 min. The supernatant was loaded onto a Bio-Scale Mini UNOsphere Q cartridge (1 mL) and eluted with a gradient of 0–1 M KCl in 50 mM Tris (pH 8.0) and 10% glycerol with 0.8 mL/min flow rate over 100 min interval. The fractions containing PHL were identified by SDS-PAGE analysis (corresponding to 20–40% KCl) before dialysis against buffer containing 50 mM Tris (pH 8.0) and 10% glycerol overnight at 4 °C. Solid $(NH_4)_2SO_4$ was then slowly added to the dialyzed sample to a final concentration of 1 M before filtering through a 0.2 μm filter. The solution was then loaded onto a Phenyl FF Hi Sub column (1 mL) with a gradient of 1–0 M $(NH_4)_2SO_4$ in 50 mM Tris (pH 8.0) and 10% glycerol with a flow rate of 0.5 mL/min over 15 min. The fractions containing PHL were identified by SDS-PAGE analysis (corresponding to 30–0 M $(NH_4)_2SO_4$) and were then pooled and dialyzed against buffer containing 50 mM Tris (pH 8.0), 150 mM KCl and 10% glycerol overnight at 4 °C. The dialyzed sample was filtered through a 0.2 μm filter and loaded onto a Superose6 10/300 GL column using an eluting buffer of 50 mM Tris (pH 8.0), 150 mM KCl and 10% glycerol with a flow rate of 0.5 mL/min. The fraction containing PHL was visualized on SDS-PAGE gel. The protocols were summarized in Appendix 2 (Figure A2.3).

However, overproduced protein was sometimes found in inclusion bodies, so the purification was developed as follows. After sonication and centrifugation, urea (6 M) was

added to the inclusion bodies and stirred for an hour at 4 °C. The protein concentration was estimated by measuring the absorbance of the solution at 280 nm. The sample was then diluted with dialysis buffer containing 20 mM Tris (pH 7.5) and 10% (v/v) glycerol until OD_{280} reached approximately 0.1 unit before dialyzing overnight at 4 °C. The dialyzed sample was loaded onto a UNOsphere Q cartridge (1 mL) and eluted using a gradient of 0–1 M KCl in 20 mM Tris (pH 7.5) and 10% glycerol with a flow rate of 1 mL/min over a 100 min interval. The fractions containing PHL protein were identified by SDS-PAGE analysis (corresponding to 18–47% KCl) and were dialyzed against buffer containing 20 mM Tris (pH 7.5) and 10% glycerol overnight at 4 °C. The dialyzed sample was further purified by gel permeation chromatography using a Superdex75 HR 10/30 column with a flow rate of 0.5 mL/min. The protocols were summarized in Appendix 2 (Figure A2.4). The purified protein was visualized by SDS-PAGE analysis. Protein concentration was performed by the Bradford assay using BSA as a standard [134]. The molecular weight of PHL was confirmed by ESI-MS experiment.

9.3.2: Initial Structural Investigation

The secondary structural investigation of the purified protein in 20 mM Tris (pH 7.5), 150 mM KCl and 10% glycerol was performed utilizing circular dichroism (CD) as previously described (Chapter 2). The gel permeation chromatography using a Superdex75 HR 10/30 column and buffer containing 20 mM Tris (pH 7.5), 150 mM KCl, 10 mM DTT and 10% (v/v) glycerol with a flow rate of 0.5 mL/min was also performed to investigate its quaternary structure.

9.4: Results and Discussion

9.4.1: DNA Cloning and Protein Characterization

The sequence searching for a putative mycoGlyoxalase in the *S. coelicolor* genome revealed three genes with annotations of putative dioxygenase (PDO), putative lyase (PLA) and putative hydrolase (PHL). The multiple sequence alignment suggested that the first two putative proteins (PDO and PLA, Chapter 8) are matched with GlxI from other organisms due to their amino acid sequence similarity and identity as well as conserved metal binding

residues. The putative hydrolase (effective query of SCO3572 and accession number of CAB45552.1), however, might likely be a GlxII since it contains all conserved metal binding residues and the THxHxDH active site motif. The *phl* gene is located on the minus strand from position 3951085 to 3951855 and consists of 771 nucleic acids coding for 256 amino acids (Figure 9.2). The protein has a predicted molecular weight of 27194.4 Da and a predicted p*I* of 5.11 (ProtParam tool from the ExPASy Proteomics Server, http://ca.expasy.org/tools/protparam.html).

```
  gtgaacgttctcgcgcccaacgcctccgcgatgaccctggacggcacgaacacctggatc
   V  N  V  L  A  P  N  A  S  A  M  T  L  D  G  T  N  T  W  I
  ctcgccgagcccgactccgacctggccgtcgtcgtcgacccgggcccgctggacgacgtc
   L  A  E  P  D  S  D  L  A  V  V  V  D  P  G  P  L  D  D  V
  cacctgcggcacgtcgtcgacacggccgagcgggccggcaagcgggtcgcgctgacgctg
   H  L  R  H  V  V  D  T  A  E  R  A  G  K  R  V  A  L  T  L
  ctgacgcacggtcacccggaccacgcggagggcgccgcgcgcttcgccgagctgacccgc
   L  T  H  G  H  P  D  H  A  E  G  A  A  R  F  A  E  L  T  R
  acgaacgtacgggcactggacccggcgctgcggctgggcgacgagggcctggcggccggt
   T  N  V  R  A  L  D  P  A  L  R  L  G  D  E  G  L  A  A  G
  gacgtgatcggggtcggcggcctggagctgagggtcgtggcgaccccggggcacacggcg
   D  V  I  G  V  G  G  L  E  L  R  V  V  A  T  P  G  H  T  A
  gactccctgtgcttccacctcccggccgaccgggccgtcctgacgggcgacaccgtcctg
   D  S  L  C  F  H  L  P  A  D  R  A  V  L  T  G  D  T  V  L
  ggccgcggcacgacggtggtggcccacccggacggccgtctcggcgactacctggactcg
   G  R  G  T  T  V  V  A  H  P  D  G  R  L  G  D  Y  L  D  S
  ctgcgcaggctcaggtcgctcacggccgacgacggggtgcacacggtgcttccggggcac
   L  R  R  L  R  S  L  T  A  D  D  G  V  H  T  V  L  P  G  H
  gggcccgtcctcgaggacgcccagggtgccgtggagtactacctcgcccatcgggcccac
   G  P  V  L  E  D  A  Q  G  A  V  E  Y  Y  L  A  H  R  A  H
  cggctcgcccaggtcgagacggcggtggaggacggccaccgcacgccgggcgaggtcgtc
   R  L  A  Q  V  E  T  A  V  E  D  G  H  R  T  P  G  E  V  V
  gcccacgtgtacgcggacgtcgaccgctccctgtggccggcggcggagctgtcggtgcgg
   A  H  V  Y  A  D  V  D  R  S  L  W  P  A  A  E  L  S  V  R
  gcgcagctggagtacctgggcgaacacgggatcgtccagctgcccgagtag
   A  Q  L  E  Y  L  G  E  H  G  I  V  Q  L  P  E  -

Amino acid composition:
33A   19R   4N    23D   1C    4Q    15E   25G   15H   3I
34L   1K    2M    2F    14P   7S    18T   2W    5Y    29V
```

Figure 9.2: The DNA sequence (771 nucleic acids), amino acid sequences (256 residues) and amino acid composition of a putative hydrolase from *Streptomyces coelicolor*. An internal thrombin protease cleavage site is shown as underlined letters.

PHL with N-terminal His-tag and Thrombin Protease Cleavage Site

The cloning of *phl* into the pET-28b(+) expression vector using PCR and the restriction enzymes, *Nde*I and *Eco*RI, was successfully accomplished. The DNA amplification of the PCR product (~ 800 pb) was optimized upon addition of DMSO (5%) and $MgCl_2$ (0.5 mM) (Figure 9.3). The presence of the *phl* gene on the plasmid was confirmed by DNA sequencing.

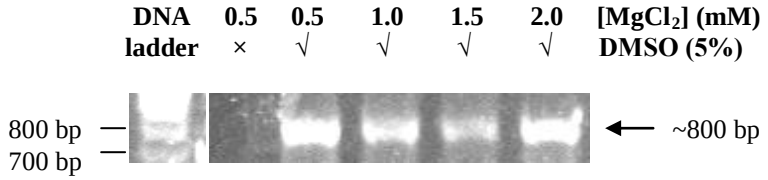


Figure 9.3: The agarose gel illustrates the PCR product of the PHL amplification using various concentrations of $MgCl_2$ with/without DMSO.

The induction tests were performed in order to optimize protein expression by varying growth and induction temperatures, amount of IPTG and $ZnCl_2$. The expression of PHL in *E. coli* BL21 (DE3) was optimized using 37/25 °C growth/induction temperatures with 1 mM IPTG (Figure 9.4). However, more soluble protein was obtained with 25/25 °C growth/induction temperatures as well as with the addition of 1 μM $ZnCl_2$ (Figure 9.5). To increase solubility of the protein, the plasmid containing the *phl* gene was heat shock transformed into a high control *E. coli* strain BL21 (DE3) pLysS, which supports proper protein folding and thus prevents protein aggregation. The induction tests on this *E. coli* cell line suggested that the PHL overexpression could be optimized under the growth and induction conditions of 25 °C with the addition of 0.5 mM IPTG and 1 μM $ZnCl_2$ over four hours induction (Figure 9.6 and 9.7)

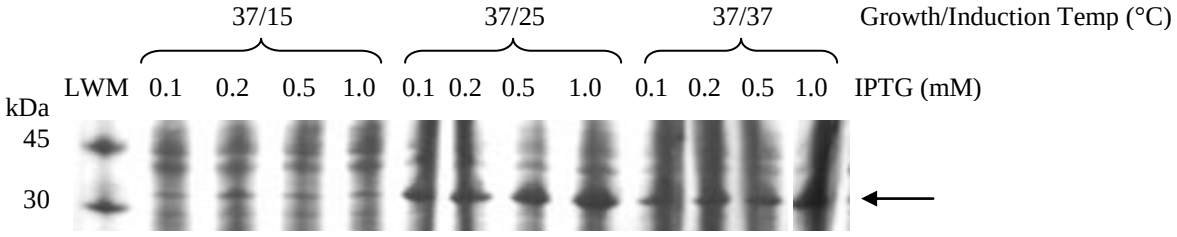


Figure 9.4: The SDS–PAGE of the experiments for PHL overexpression in *E. coli* BL21 (DE3) under different growth/induction temperatures and amount of IPTG. The arrow indicates the presence of His-tagged PHL.

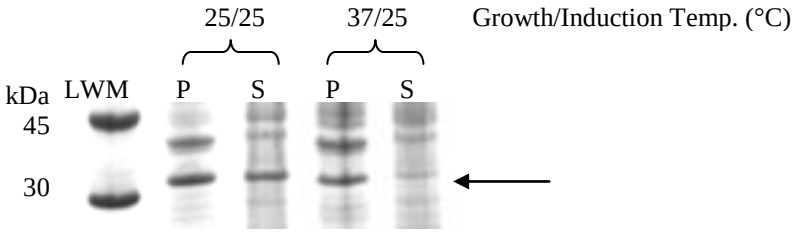


Figure 9.5: The SDS–PAGE for solubility tests of PHL that overexpressed in *E. coli* BL21 (DE3) with additional 1 mM IPTG and 1 μM $ZnCl_2$ under different growth/induction temperatures: S = proteins in the supernatant and P = proteins in pellet after sonication. The arrow indicates the presence of His-tagged PHL.

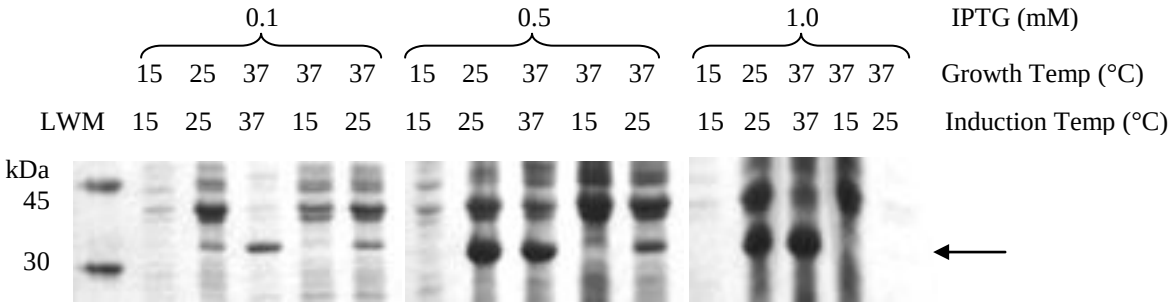


Figure 9.6: The SDS–PAGE of overexpressed PHL in *E. coli* BL21 (DE3) pLysS under different growth/induction temperatures and concentrations of IPTG. The arrow indicates the presence of His-tagged PHL.

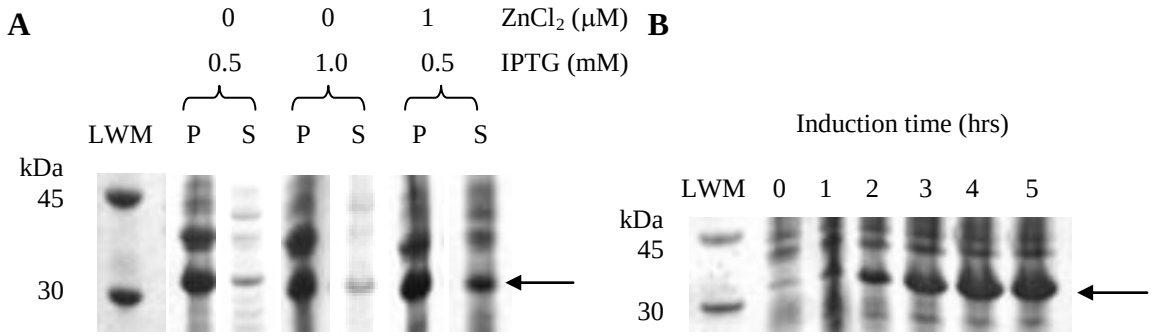


Figure 9.7: (A) The SDS–PAGE of solubility tests for production of PHL in *E. coli* BL21 (DE3) pLysS that overexpressed under 25/25 °C growth/induction temperatures and various concentrations of IPTG and $ZnCl_2$: S = proteins in the supernatant and P = proteins in pellet after sonication. **(B)** The SDS–PAGE for the induction time studies for PHL production in *E. coli* BL21 (DE3) pLysS under 25/25 °C growth/induction temperatures with the addition of 0.5 mM IPTG and 1 μM $ZnCl_2$. The arrow indicates the presence of His-tagged PHL.

The cell pellet of approximately 3.5 g was harvested from a 1 L cell culture of His-tagged PHL overproduced in *E. coli* BL21 (DE3) pLysS. The amount of protein recovered from the purification using HisTrap affinity column chromatography (Figure 9.8) was approximately 6.0 mg/mL (~24 mg total protein), which was higher than the one expressed in

E. coli BL21 (DE3) (3.4 mg/mL or ~15 mg total proteins). However, after treatment with thrombin protease and purification using the combined HiTrap Benzamidine FF and HisTrap HP affinity columns (Figure 9.9), no PHL with a predicted molecular mass of approximately 29 kDa was observed by SDS-PAGE analysis (Figure 9.10A). Thus, non-specific thrombin cleavage and the presence of a possible internal thrombin protease cleavage site in PHL were examined. It was found that PHL was digested completely after incubating with thrombin protease for 7 hours at 4 °C (Figure 9.10B). More importantly, an internal thrombin digestion site in PHL (139-VLGRGT-144) was observed and confirmed by the predicted amino acid sequence of PHL. In this recognition site, the bond between Arg^{142} and Gly^{143} is cleaved by thrombin protease, leading to the reduced molecular masses of approximately 12.4 and 16.9 kDa as observed by SDS-PAGE experiment. Thus, this method was discontinued.

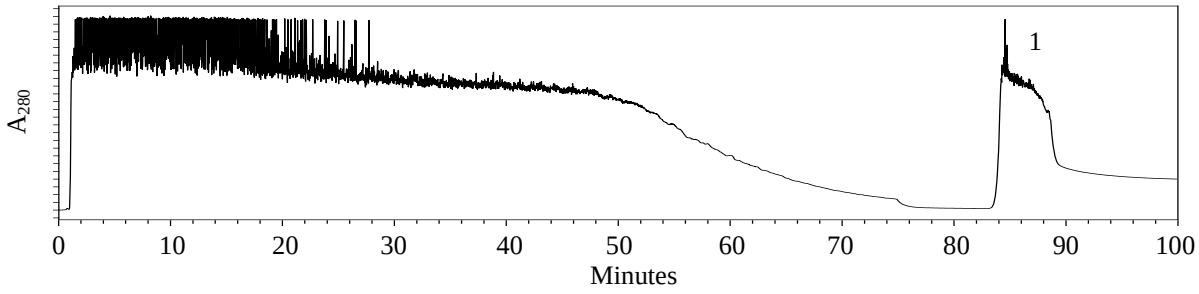


Figure 9.8: The HPLC chromatogram of PHL purification using a HisTrap HP affinity column measured at 280 nm. His-tagged proteins were collected using a buffer containing 500 mM imidazole and a flow rate of 0.5 mL/min (peak 1).

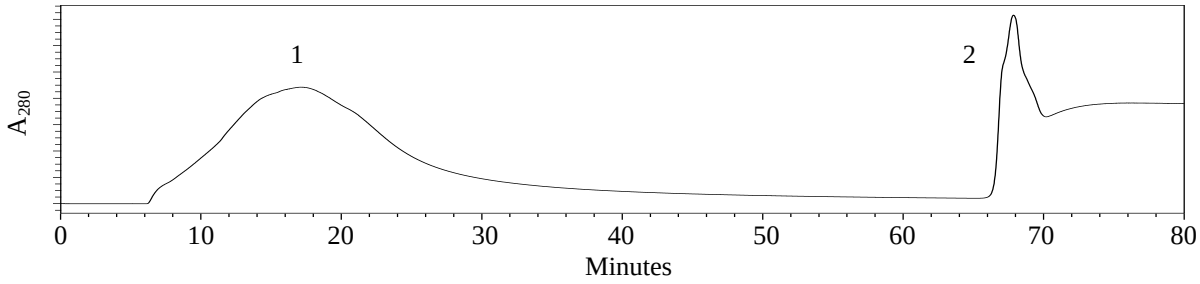


Figure 9.9: The HPLC chromatogram of cleaved PHL obtained by purification using the HiTrap Benzamidine FF and HisTrap HP affinity columns measured at 280 nm. The purified PHL (without His-tag) was collected with buffer containing 20 mM imidazole (peak 1), while non-cleaved His-tagged proteins were eluted with buffer containing 500 mM imidazole (peak 2).

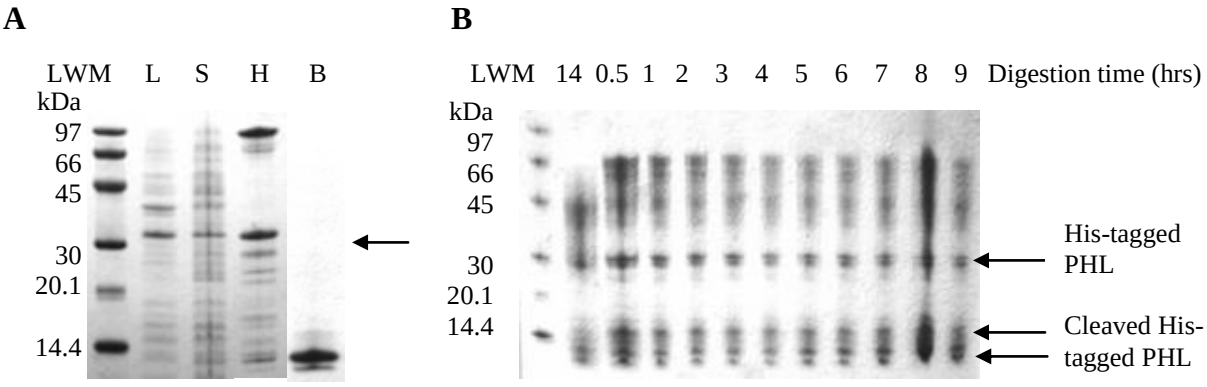


Figure 9.10: (A) The SDS–PAGE of the PHL purification procedure: LWM = low molecular weight marker, L = lysate, S = supernatant after sonication, H = HisTrap eluted fraction and B = HisTrap and Benzamidine flow-through fraction. The arrow indicates the present of His-tagged PHL. **(B)** The SDS–PAGE for the thrombin protease cleavage times (0.5–14 hours).

PHL with the N-terminal His-tag and Tev Protease Cleavage Site

This DNA cloning protocol involved deletion of the thrombin protease cleavage site and insertion of a Tev protease cleavage site using PCR and restriction enzymes. The first PCR (~100 bp) was to replace the thrombin protease cleavage site with a Tev protease encoding gene sequence and generation of the *Xba*I restriction site (Figure 9.11). The second PCR (~900 pb) was to clone the *phl* gene with a *EcoR*I restriction site. The third PCR (~1000 pb) step was to combine the first and the second PCR products and then clone back into the pET-28b(+) expression vector (Figure 9.11). The DNA amplification of the PCR product was achieved upon addition of DMSO (5%) and $MgCl_2$ (0.5 mM). The existence of the *phl* gene on the plasmid was confirmed by DNA sequencing.

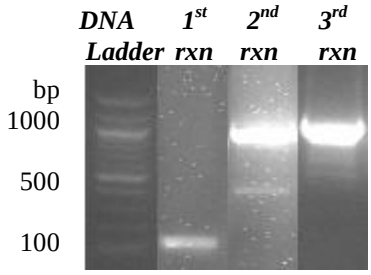


Figure 9.11: The 10% agarose gel illustrates the PCR products of the DNA amplification of the gene coding for PHL with the N-terminal His-tag and Tev protease cleavage site.

The analyses of the results of the induction tests in terms of induction temperatures, amount of inducer (IPTG) and the addition of $ZnCl_2$ suggested that the overexpression of PHL

with the N-terminal His-tag and Tev protease cleavage site could be optimized by growth at 25 °C in the presence of 1 mM $ZnCl_2$ (Figure 9.12). Then, proteins were expressed and induced at 15 °C with 0.5 mM IPTG. Low temperature conditions would allow for slower expression, thus providing more time for properly folding of the protein. As well, addition of Zn^{2+} that was found to bind to most GlxII might help the protein production levels of PHL.

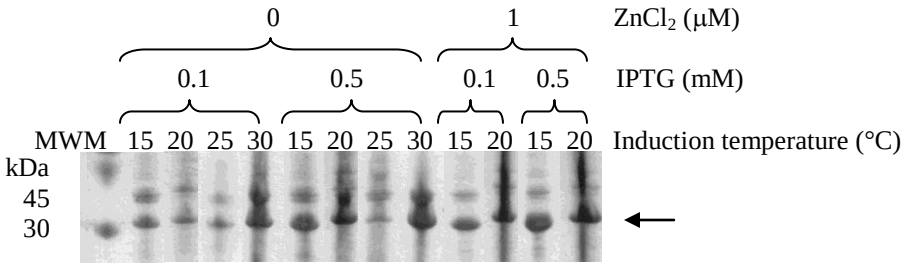


Figure 9.12: The SDS–PAGE of the overexpression of PHL that was grown at 37 °C overnight, then induced with IPTG (0.1 or 0.5 mM) with/without the presence of $ZnCl_2$ (1 μM) at various induction temperatures (15, 20, 25, 30 °C). The arrow indicates the presence of His-tagged PHL.

The protein under these expression conditions was recovered in inclusion bodies, thus its purification using a HisTrap HP affinity column was performed under denaturing conditions (Figure 9.13). The first purification step was to separate His-tagged proteins from *E. coli* proteins, and these proteins were collected utilizing 500 mM imidazole and 6 M urea. The denaturant was removed from the sample by dialysis overnight at 4 °C. Another purification using the HisTrap HP affinity column was repeated under native conditions (without urea). His-tagged proteins were then cleaved by an addition of Tev protease (also possesses His-tag itself). Thus, the Tev protease and the cleaved His-tag could be conveniently removed by the purification using another HisTrap HP affinity column under native conditions. However, no PHL in buffer with low imidazole (20 mM, flow-through fraction) with expected molecular mass (~30 kDa) was collected as observed by SDS-PAGE experiment after cleavage with the Tev protease. The desired protein was, on the other hand, eluted from the column in the presence of high imidazole (500 mM), suggesting that the His-tag was still attached to PHL and the cleavage was unsuccessful. It was possible that the Tev protease was not optimally active under the studied conditions (20 mM Tris (pH 7.0), 500 mM KCl, 5 mM DTT and 10% (v/v) glycerol). Thus, optimal cleavage conditions should be further investigated.

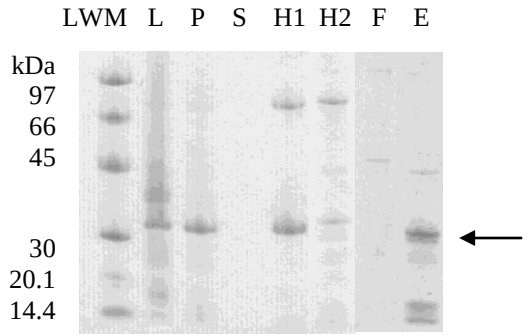


Figure 9.13: The SDS–PAGE of the PHL purification scheme using Tev protease cleavage and the HisTrap HP affinity column: LWM = low molecular weight marker, L = lysate, P = pellet after sonication, S = supernatant after sonication, H1 = denatured fraction from 1st HisTrap column, H2 = native fraction from 2nd HisTrap column, F = flow-through fraction from 3rd HisTrap column and E = eluted fraction after 3rd HisTrap column. The arrow indicates the presence of PHL.

Non His-tagged PHL

The DNA cloning steps for isolation of the *phl* gene onto the pET-22b(+) expression vector was performed similarly to that of the cloning into pET-28b(+) vector. The induction test for PHL expression indicated that PHL could not be effectively produced in *E. coli* BL21 (DE3) (Figure 9.14A). Instead, it could be overexpressed in *E. coli* BL21 (DE3) CodonPlus RP with the optimum conditions of 25 °C growth/induction temperatures and 4 hours induction time, which gave ~4.5 g cell pellet per 1 L of LB (Figure 9.14B). The isolation of PHL from *E. coli* proteins was begun by anion exchange chromatography using a UNOsphere Q cartridge (Figure 9.15). Further purification was accomplished by ammonium sulfate precipitation, in which most *E. coli* proteins were precipitated in the presence of high concentration of salt. PHL was also precipitated in 1.7 M $(NH_4)_2SO_4$, thus the salt concentration was decreased to 1 M. Using a Phenyl Sepharose chromatographic step, PHL was eluted in buffer containing low salt (Figure 9.16). After this purification step, a small amount of pure protein with a molecular weight of 30 kDa was observed by SDS-PAGE analysis (Figure 9.17).

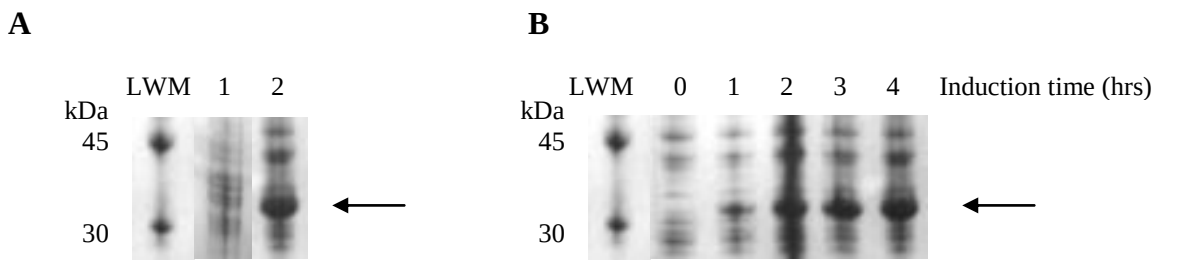


Figure 9.14: (A) The SDS–PAGE of PHL production from (1) *E. coli* strain BL21 (DE3) and (2) BL21 (DE3) CodonPlus RP under 25/25 °C growth/induction temperatures with 1 mM IPTG. **(B)** The SDS–PAGE of the induction tests of PHL in *E. coli* strain BL21 (DE3) CodonPlus RP under 25/25 °C growth/induction temperatures with 1 mM IPTG. The arrow indicates the band representing the presence of PHL.

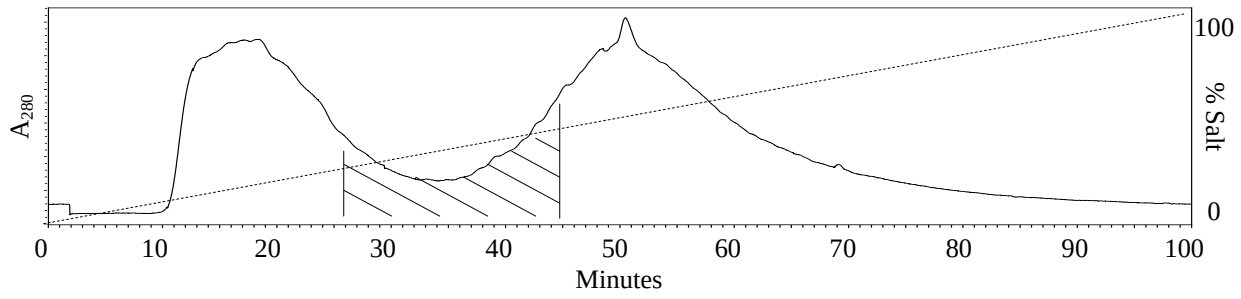


Figure 9.15: The chromatogram of PHL purification using a Bio-Scale Mini UNOsphere Q cartridge and a gradient buffer of 0–1 M KCl in 50 mM Tris (pH 8.0) and 10% glycerol over 100 min with 0.8 mL/min flow rate. The shaded area contains PHL protein as indicated by SDS–PAGE analysis.

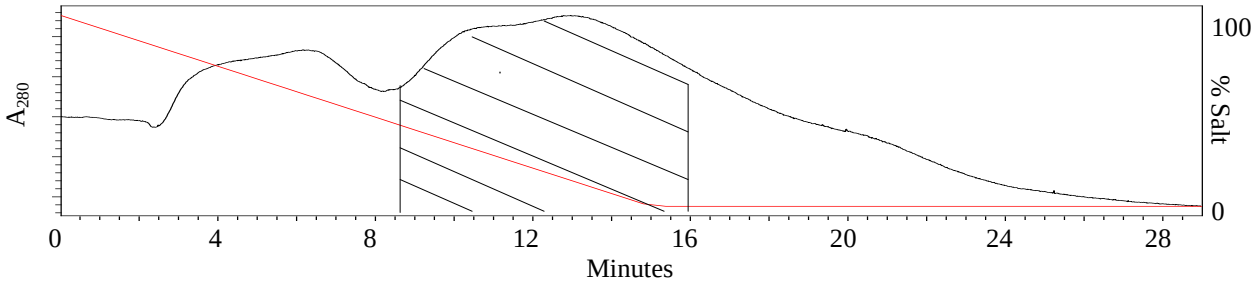


Figure 9.16: The chromatogram of the PHL purification using a Phenyl FF Hi Sub column and a gradient buffer of 0–1 M $(NH_4)_2SO_4$ in 50 mM Tris (pH 8.0) buffer over 15 min with a 0.5 mL/min flow rate. The shaded area contained the PHL protein as indicated by SDS–PAGE analysis.

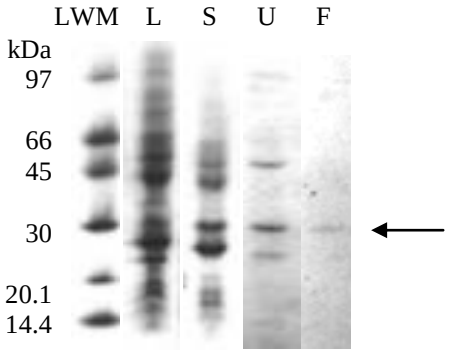


Figure 9.17: The SDS–PAGE of the PHL purification: LWM = low molecular weight marker, L = lysate, S = supernatant after sonication, U = UNOsphere Q fraction and F = phenyl FF fraction. The arrow indicates the presence of PHL.

PHL was recovered from inclusion bodies in buffer containing 6 M urea, which was then followed by dialysis in order to remove the denaturant. The proteins were then loaded onto the UNOsphere Q cartridge (Figure 9.18). After this step, the protein was quite pure as indicated by SDS-PAGE analysis, thus further purification by ammonium sulfate precipitation and by Phenyl Sepharose chromatography could be skipped. The gel permeation chromatography step using the Superdex75 HR 10/30 column under native conditions was employed for the final purification of this protein (Figure 9.19). The molecular mass estimated by SDS-PAGE experiment was approximately 30 kDa (Figure 9.20A), while ESI-MS provided a more precise molecular mass of 27193.0 Da (calculated MW is 27194.4 Da) (Figure 9.20B). No post-translational modifications were observed.

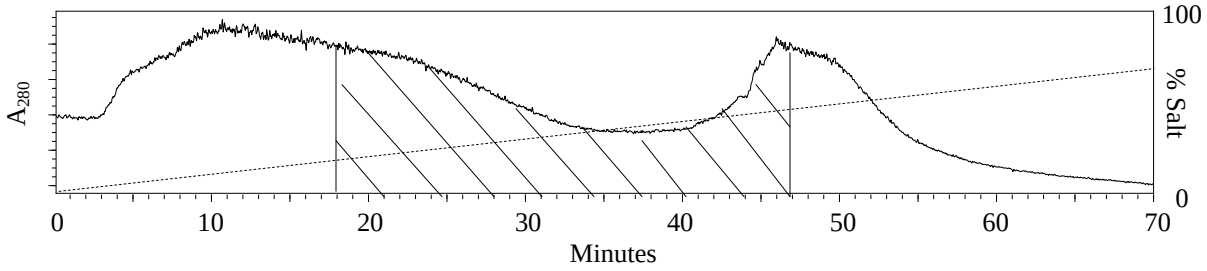


Figure 9.18: The HPLC chromatogram of the PHL purification using an anion exchange chromatographic step using the UNOsphere Q cartridge and a gradient buffer of 0–1 M KCl in 20 mM Tris (pH 7.5), 10 mM DTT and 10% (v/v) glycerol with a flow rate of 1 mL/min. The shaded area contained PHL, which was identified by SDS–PAGE analysis.

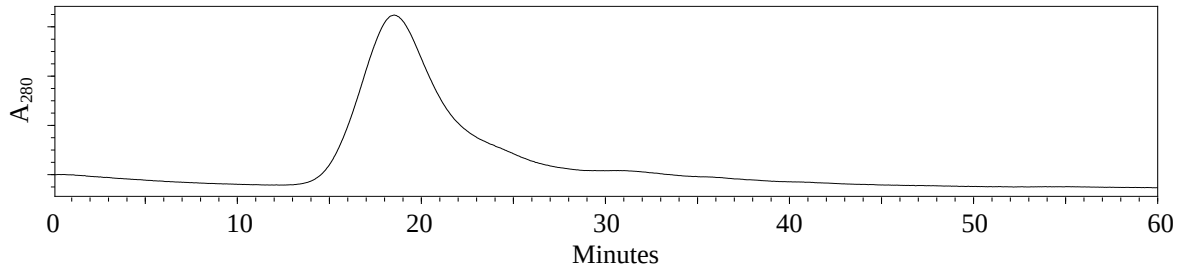


Figure 9.19: The HPLC chromatogram of the PHL purification by a gel permeation chromatographic step using a Superdex75 HR 10/30 column and a buffer containing 20 mM Tris (pH 7.5), 150 mM KCl, 10 mM DTT and 10% (v/v) glycerol with a flow rate of 0.5 mL/min. The peak containing PHL was identified by SDS–PAGE analysis.

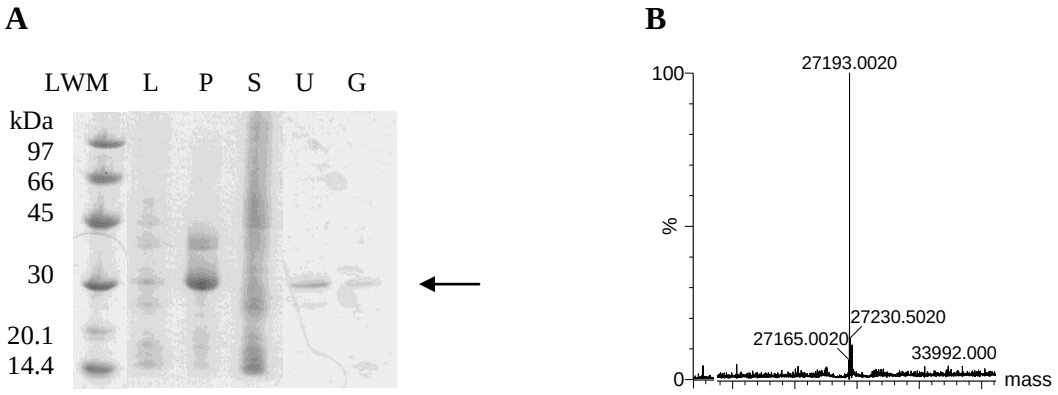


Figure 9.20: (A) SDS–PAGE of PHL purification: LWM = low molecular weight marker, L = lysate, P = pellet after sonication, S = supernatant after sonication, U = UNOsphere Q fraction and G = gel permeation fraction. The arrow indicates the presence of PHL. **(B)** The electrospray mass spectrum of PHL shows the molecular weight of a single subunit at 27193.0 Da (calculated MW is 27194.4 Da).

9.4.2: Structural Investigation

A secondary structural investigation of the purified PHL (0.3 mg/mL) in 20 mM Tris (pH 7.5), 150 mM KCl and 10% glycerol using CD analysis suggested that the protein possessed a negative maximum at 208 nm and a shadow at 220 nm, which is characteristic of a predominant β-sheet containing protein (Figure 9.21). The shape and negative bands of this CD spectrum were similar to that of human GlxII [296]. The web-based K2D secondary structural prediction protocol [152, 153] estimated the contents of PHL to be 17% α-helix, 27% β-sheet and 56% random coil. This prediction corresponded quite well to the available X-ray crystallographic structure of human GlxII (PDB: 1QH3), where the N-terminal domain employs a series of β-sheet (βββαβαββ and βββββαβ motifs) and the C-terminal domain consists of five α-helices.

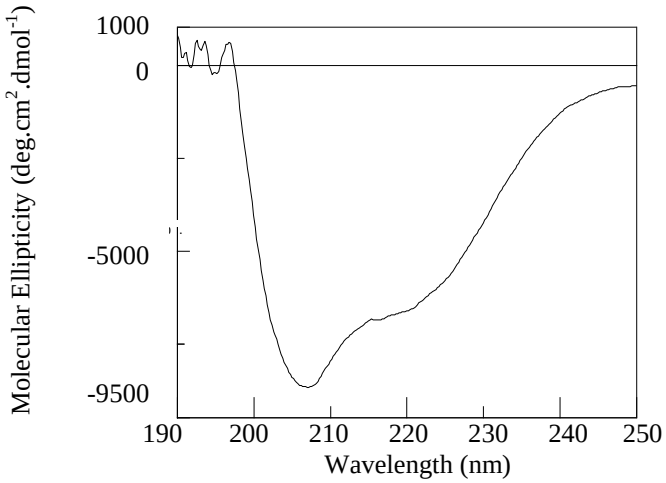


Figure 9.21: The CD spectrum of PHL (0.3 mg/mL) in 20 mM Tris (pH 7.5), 150 mM KCl and 10% glycerol scanning between 200–250 nm.

Additionally, the quaternary structure of the PHL protein in buffer containing 20 mM Tris (pH 7.5), 150 mM KCl, 10 mM DTT and 10% (v/v) glycerol was investigated by gel permeation chromatography using a Superdex75 HR 10/30 column with a flow rate of 0.5 mL/min. The molecular weight of PHL estimated from this experiment was approximately 30 kDa, similar to the calculated molecular mass of 27.2 kDa (Figure 9.22). This result suggested that the purified PHL is a monomeric protein under the studied conditions.

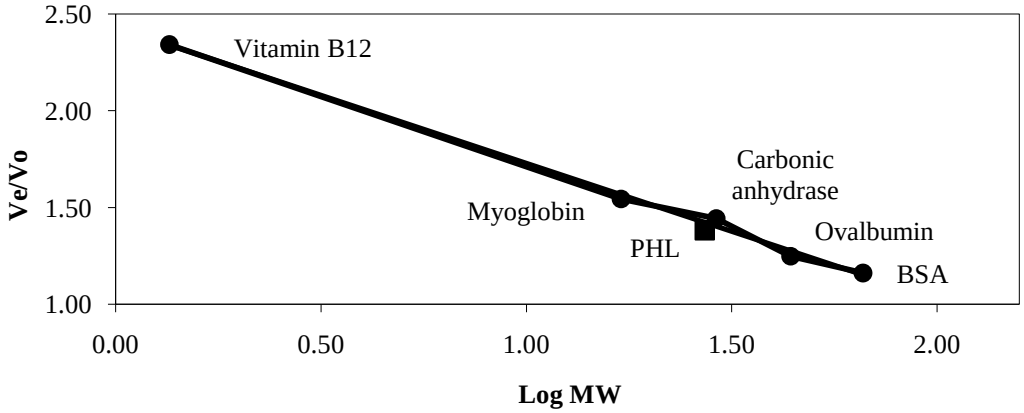


Figure 9.22: The gel permeation chromatographic profile of PHL (■) (0.3 mg/mL) that fits into the Bio-Rad protein standard (●) plot using a Superdex75 HR 10/30 column and a buffer containing 20 mM Tris (pH 7.5), 150 mM KCl, 10 mM DTT and 10% (v/v) glycerol with a flow rate of 0.5 mL/min. The protein standard consists of BSA (66 kDa), ovalbumin (44 kDa), carbonic anhydrase (29 kDa), myoglobulin (17 kDa) and vitamin B12 (1.35 kDa).

9.5: Conclusions

The putative GlxII was searched for through the genome of *Streptomyces coelicolor* and the highly related protein was identified in the database as a "putative hydrolase". The multiple sequence alignment of PHL with GlxII from other organisms suggested that this protein could function as a GlxII since it contains all the expected metal binding residues (Figure 8.1). Several methodologies for DNA cloning of *phl* from *S. coelicolor* into the pET expression vectors and their protein purifications were investigated in order to optimize protein expression and recovery. The purified protein with expected molecular mass was predicted to be a monomer containing a predominant β-sheet, which corresponded to other previously characterized GlxII.

Further investigation on GlxII activity in terms of substrate and metal specificities should be performed to determine the mcGlxII activity of this protein. The enzymatic assay using DTNB could be carried out using PHL and the substrate, *S*-D-lactolymycothiol. This substrate would need to be obtained using the enzymatic reactions of mcGlxI (putative dioxygenase from Chapter 8) with MG and MSH. Then, the coupled assay of this mcGlxI reaction with PHL could indicate if this protein functions as GlxII. As well, the assay reaction of PHL with the commercial substrate, *S*-D-lactolyglutathione, should also be performed to investigate the substrate specificity of this putative enzyme. The assay reaction using the commercial yeast GlxII with *S*-D-lactolygluathione and *S*-D-lactolymycothiol should be performed in parallel as a control as well as to investigate the substrate specificity of this enzyme. These experimental data would extend our knowledge and understanding of the substrate binding affinity and substrate binding site based on the structure of the enzymes in the same family that exhibit different, but similar, reactions.

It would also be interesting to examine the activity of PHL in the presence of different metals. It was previously reported that GlxI contains two metal activation classes; Zn^{2+}-activated class and a non-Zn^{2+}-activated class (being Ni^{2+}/Co^{2+}-activation), while there is no evidence of distinct metal activation classes for the GlxII enzymes [67]. In attempts to investigate the metal dependence of GlxII, the previous report by O'Young *et al*. suggested that the *E. coli* GlxII should be chosen since the GlxI from this organism was found to be Ni^{2+}/Co^{2+}-activated [67]. Thus, it was worthwhile to investigate its GlxII from the same bacterial source to determine if this enzyme could also be activated with Ni^{2+}. However, a

study of *E. coli* GlxII in the presence of various metals (Mn^{2+}, Ni^{2+}, Zn^{2+} and Co^{2+}) found the enzyme to be inactivated in the presence of Ni^{2+}, thus there is no distinct metal activation class in GlxII from this organism [67]. Based on these experiments, the metal activation profile of PHL could be performed similarly to investigate the metal specificity of mcGlxII in comparison with that of *E. coli* GlxII.

Additionally, structural investigations on PHL could be performed by X-ray crystallographic analysis. Based on the multiple sequence alignment, it could be predicted that PHL might possess a similar 3D structure as other GlxII. This structural investigation would be significant to elucidate the enzyme mechanism as well as to extend structural knowledge that might be critical to the development of novel enzyme inhibitors of this new group of GlxII enzymes in the Streptomyces and, more importantly, possibly the Mycobacteria.

CHAPTER 10: SUMMARY AND FUTURE WORK

10.1: Summary of Accomplishments

For almost a century the Glx system, isolated from a variety of organisms, has been under investigation [17-35]. Several bacterial Glx enzymes from human and animal pathogens have been studied. One alternative method to combat these bacteria may be to suppress the Glx pathway by inhibiting GlxI and/or GlxII, which eventually could lead to cell death due to the increased levels of cytotoxic MG. Since both the infectious agents and host cells possess the Glx pathway, it would be necessary to be able to differentiate their enzymes (GlxI and GlxII), which would then support the potential for rational structure-based drug design as well as lead to the discovery of an effective drug that can specifically target bacterial enzymes. On a more fundamental note, the main reason for our work is to increase our fundamental understanding of the structure-function relationship of these enzymes, especially of Glyoxalase I, particularly from Gram-positive and Gram-negative systems. Additionally, we are also interested in the evolution within the βαβββ superfamily including Glyoxalase I, fosfomycin resistance protein (FosA), methylmalonyl-CoA epimerase (MMCE), extradiol dioxygenase (DIOX), mitomycin C resistance protein (MRP) and bleomycin resistance protein (BRP). Several of these are drug resistance proteins, which create crucial problems in the treatment of infectious diseases. These resistance proteins are suspected to have evolved from proteins with different functions in the bacterial cells, thus an investigation of these proteins within the same superfamily could provide additional insight into their relatedness. As well, this investigation would extend our fundamental understanding of drug resistance proteins, which eventually could lead to improvements in drug design.

Even though GlxI has been widely investigated from various biological sources, no detailed report, to our knowledge, has appeared on the enzymes from a Gram-positive bacterium. We discovered that the GlxI from *Clostridium acetobutylicum* (CLO GlxI), which possesses a shorter amino acid sequence, is inactive in the presence of Zn^{2+} (non-Zn^{2+}-activation class) and is similar in metal-activation profile to the GlxI from the Gram-negative *Escherichia coli* [19] (Chapter 2). The enzyme is indeed activated in the presence of Ni^{2+} and Co^{2+} atoms. As expected, its active site geometry with a bound Zn^{2+} atom forms a trigonal bipyramidal metal-coordination with four metal binding protein residues and one water

molecule around the metal center, while its Ni^{2+}-bound counterpart forms a six-coordinated octahedral geometry with the same four metal binding protein residues but with two water molecules. Its catalytic pocket, however, is formed within a monomer (two active sites per dimer). This is the first discovery for a "small" GlxI to possess such a unique dimeric arrangement, where other small GlxI form two active sites at the dimeric interface.

This new discovery in the structural characterization of CLO GlxI led to the investigation on the cross functional activity of the enzymes within the same βαβββ superfamily (Chapter 5). The dimeric arrangement of CLO GlxI is similar to that of methylmalonyl-CoA epimerase (MMCE), a metalloenzyme that converts (2*S*)-methylmalonyl-CoA to its 2*R*-epimer [105-108]. MMCE forms a six-coordinated octahedral geometry with four metal binding protein residues and two water molecules around the active metal [105]. Its activity can be optimized in the presence of Co^{2+} ion [216, 217], similar to the activating metals of GlxI in the Ni^{2+}/Co^{2+}-activation class [19, 32]. Due to its metal-dependency profile and its overall structural similarity, we hypothesized that GlxI might behave as a bifunctional enzyme that could turnover both hemithioacetal (GlxI substrate) and (2*S*)-methylmalonyl-CoA (MMCE substrate). However, several GlxI from different biological sources including yeast, Gram-positive *Clostridium acetobutylicum* and Gram-negative *Pseudomonas aeruginosa* (GloA2 and GloA3) as well as two putative GlxI enzymes from *Streptomyces coelicolor* (annotated as putative dioxygenase and putative lyase) exhibit no MMCE activity, suggesting that GlxI and its putative enzymes cannot accomplish MMCE reaction regardless of metal activation class nor size of protein. No study in the Glyoxalase field, to our knowledge, has ever reported analysis of alternate enzymatic activities for the Glyoxalase I proteins.

From a similar perspective, we were also interested in investigating the relatedness in terms of evolution of bacterial drug resistance mechanisms within the βαβββ superfamily, especially between GlxI and the fosfomycin resistance protein (FosA) (Chapter 4). Previously, it was suggested that FosA might have evolved from the already existing gene in the same bacterial source, and the closest evolution would be the proteins in the same superfamily. FosA is a metalloenzyme that opens the epoxide of the antibiotic fosfomycin using GSH as a cosubstrate [113, 168]. FosA employs two metal binding sites, one for a divalent metal such as Mn^{2+} and another one for the monovalent metal K^{+} [110, 111, 169, 170, 175]. Both metals are significant for enzymatic activity [110, 111, 169, 170, 175]. The enzyme's metal coordination

in the Mn^{2+} binding site (relevant to the metal binding site in GlxI) is formed by three metal binding protein residues (instead of four as in GlxI) [169]. Thus, a single point mutation in *E. coli* GlxI (E56A) was generated to approximately mimic the metal binding site in FosA. However, no FosA activity was observed by this GlxI mutein. Analysis of the multiple sequence alignment as well as structural comparison suggested that the substrate binding sites of FosA and GlxI might be sufficiently different and, thus, responsible for the lack of observed enzymatic activity. However, we felt it was important to determine if *any* FosA activity might be found for the GlxI mutein. This was not the case.

As mentioned previously, the metal activation profile of GlxI can be quite reliably predicted by the length of the amino acid sequence of the protein. We hypothesized that the extra amino acids that only exist in the Zn^{2+}-activated enzyme, including a long α-helix (helix B) and two small loops (loops C and D) might be responsible for different metal activation profiles (Chapter 3). Deletional mutagenesis of the Zn^{2+}-activated GloA3 from *P. aeruginosa* was used to generate a structural mimic of the Ni^{2+}/Co^{2+}-activated GloA2 from the same bacterial organism. The enzymatic activity of the GloA3 muteins as well as structural comparison of human GlxI (Zn^{2+}-activation) and *E. coli* GlxI (Ni^{2+}/Co^{2+}-activation) suggested that the long α-helix (15 amino acids) that locates almost in the middle of the catalytic pocket is significant for metal specificity of GlxI. The effect of this loop on metal specificity may be due to several factors such as the hydrophobic environment, limited catalytic space and interaction of the residue(s) in this helix with the metal binding residues. On the other hand, the two small loops (4 and 5 amino acids) that lie close to one of the metal binding residues appear to be crucial for the magnitude of the enzymatic activity, but do not have strong effects on metal specificity.

We also attempted to obtain additional insight into the reaction mechanism of the Ni^{2+}/Co^{2+}-bound GlxI as well as investigate the structural flexibility upon binding of inhibitors (Chapter 6). Two transition state analogues including *S*-(*N*-hydroxy-*N*-methylcarbamoyl)glutathione and *S*-(*N*-hydroxy-*N*-phenylcarbamoyl)glutathione were synthesized based on previous reports [114, 115]. It was found that *S*-(*N*-hydroxy-*N*-phenylcarbamoyl)glutathione is a stronger competitive inhibitor than *S*-(*N*-hydroxy-*N*-methylcarbamoyl)glutathione because it contains a phenyl group which, we hypothesize, fits better into the relatively hydrophobic pocket of the enzyme. An X-ray crystallographic

investigation of CLO GlxI, GloA2 and GloA3 with *S*-(*N*-hydroxy-*N*-phenylcarbamoyl)glutathione is currently being undertaken in collaboration with Professor Charles Bond, University of Western Australia (WA, Australia). In the meantime, analysis of the determined but as yet to be published structure of *E. coli* GlxI with a bound hydroxamate analogue suggested an enzyme movement upon binding of the inhibitor. Using this 3D structure as well as other experimental data, we proposed a reaction mechanism of the Ni^{2+}/Co^{2+}-activated GlxI in comparison to what had been reported for the Zn^{2+}-activated enzyme [44, 46, 48, 49].

Our interest was then extended to the Glx systems that might utilize a different type of thiol cofactor, mycothiol (MSH). Some mycobacteria are human pathogens, one of which, *Mycobacterium tuberculosis*, is the causative agent of tuberculosis. These microorganisms produce high levels of MSH instead of GSH. Thus, our purpose was to investigate the putative mycoGlyoxalase (mcGlx) system using MSH as a cofactor in comparison to the GSH-dependent Glx system. Due to difficulty and low recovery yield in the isolation of MSH from natural sources, numerous approaches to synthesize the truncated mycothiol (tMSH) or des-*myo*-inositol mycothiol were carried out (Chapter 7) since it had been previously reported that the disulfide form of this tMSH could be used as a substrate in the reaction of a mycobacterial MSH disulfide reductase [116]. Among several methodologies, tMSH was successfully synthesized utilizing a peptide synthesis assembly approach. Using this synthesized cofactor, the formation of the mcGlxI substrate, the hemithioacetal from the non-enzymatic reaction of MG and tMSH, was investigated in terms of dissociation constant (K_d) and equilibrium time (Chapter 8) in comparison to those characteristics of the hemithioacetal produced by reaction of MG and GSH.

We then examined the putative enzymes in the mcGlx system, mycoGlyoxalase I (mcGlxI) and mycoGlyoxalase II (mcGlxII) (Chapter 8 and 9). Sequence searching and data mining of the genome of *Streptomyces coelicolor*, a non-pathogenic Gram-positive bacterium that produces higher level of MSH than GSH, revealed two putative mcGlxI annotated as a putative dioxygenase (PDO) and a putative lyase (PLA) and one putative mcGlxII that was annotated as a putative hydrolase (PHL). These genes were cloned and the gene products were overexpressed, purified and studied. The enzymatic reactions of mcGlxI on PDO and PLA were examined in terms of assay detection wavelength, substrate specificity and metal

activation profile. However, only PDO exhibited mcGlxI activity with the hemithioacetal product of MG and tMSH, while PLA was found to be inactive in these reactions. The product of the mcGlxI reaction using PDO was isolated and identified to confirm the existence of the enzymatic reaction. This putative mcGlxI, PDO, can be activated in the presence of Ni^{2+} and Cu^{2+} atoms, instead of Ni^{2+} and Co^{2+} as in other GlxI in the non-Zn^{2+}-activation class. The different activating metals might be the result of the bioavailability of these metals for the organism in its natural soil habitat, although this is purely conjectural at present. The putative mcGlxII, PHL, was found to be difficult to purify due to its solubility, thus the enzymatic assay has yet to be pursued on this protein.

10.2: Future Work

Several future experiments are suggested in order to more fully understand the Glx system from various organisms. As mentioned earlier, the *C. acetobutylicum* GlxI exhibits a lower than expected activity with the hemithioacetal of MG and GSH (Chapter 2). Follow-up studies of the assay using different types of thiol cofactor might be able to increase its activity. For example, it was reported that *C. perfringens* produces higher levels of H_2S than GSH (2.2 μmol H_2S and 0.24 μmol GSH/g residual dry weight) [83], thus this thiol could be a possible candidate as a cofactor in the GlxI reaction. As well, the X-ray crystallographic analysis of the inhibitor-bound CLO GlxI is necessary to support our hypothesis on limited catalytic space in this enzyme and a slight interference of a helix loop at the entrance of the catalytic pocket (Chapter 5). In addition, we suspect that the Cys residue that locates close to the active site (Cys^{79}) might possibly form a metal-enzyme complex, resulting in a distorted metal-coordination as well as producing a block at the entrance of the active site. A single point mutagenesis study of this residue could address these suggestions. It is also interesting to investigate another putative GlxI from *C. acetobutylicum* (NP_346890) in terms of enzymatic activity as well as structural analysis. The multiple sequence alignment of this putative enzyme with other GlxI indicates that this enzyme possesses four conserved metal binding residues and has a shorter amino acid sequence. Thus, it might behave as another Ni^{2+}-activated GlxI in Clostridium. It is also worthwhile to perform additional structural

investigations, since analysis of the multiple sequence alignment suggested that this putative enzyme might possess a similar dimeric arrangement as the first CLO GlxI.

To follow-up on the investigation of the deletional mutagenesis of the Zn^{2+}-activated GlxI from *P. aeruginosa* (Chapter 3), the X-ray crystallographic analysis of GloA2 and GloA3 is required as one part of the experiments to explain the effect of the extra amino acids that only exist in the Zn^{2+}-activated enzyme on its metal specificity. As well, we suspect that a hydrophobic residue (Phe^{82} in GloA3) on the long α-helix (helix B) in the Zn^{2+}-activated GlxI might play an important role in metal specificity. Thus, experiments generating single mutations at this residue might lead to a possible explanation on the effect of this helix.

We are also interested in the evolutionary relatedness between GlxI and FosA. Due to the different substrate binding residues between the *E. coli* GlxI and *P. aeruginosa* FosA, no FosA activity in the generated GlxI-E56A was observed (Chapter 4). FosA might possibly have evolved from the same ancestor as GlxI but has accumulated more point mutations, thus only mutagenesis in the metal binding residue would be insufficient to alter its enzymatic reaction. More mutagenesis on the fosfomycin binding residues in GlxI to mimic that of FosA might be able to change its function.

It would also be important to continue our studies on the design and synthesis of inhibitors of GlxI (Chapter 6). The proposed reaction mechanism as well as enzyme movement upon binding of inhibitor to Ni^{2+}/Co^{2+}-activated GlxI requires additional structures from other experimental data such as X-ray crystallographic analysis.

The enzymatic reaction of mcGlxI was successfully performed using the substrate of MG-tMSH with the annotated dioxygenase (Chapter 8). However, the kinetic studies of this enzyme with the real substrate, MG-MSH, will also be required. Since little was recovered from the isolation of MSH from bacterial sources, a large-scale purification is needed (Chapter 7). As well, the enzymatic activity of the putative mcGlxII could be easily performed using the utilization of Ellman's reagent or DTNB as previously reported [67, 74] (Chapter 9) once sufficient quantities of the required thioester, *S*-lactoylmycothiol, are obtained.

REFERENCES

1. Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C. and Ferrin, T.E. 2004. UCSF Chimera- a visualization system for exploratory research and analysis. *J Comput Chem*. **25**(13): p. 1605-12.
2. Booth, I.R., Ferguson, G.P., Miller, S., Li, C., Gunasekera, B. and Kinghorn, S. 2003. Bacterial production of methylglyoxal: a survival strategy or death by misadventure? *Biochem Soc Trans*. **31**(Pt 6): p. 1406-8.
3. Desai, K.M., Wang, H., Banigesh, A., Dhar, A., Liu, J., Untereiner, A. and Wu, L. 2010. Oxidative stress and aging: is methylglyoxal the hidden enemy? *Can J Physiol Pharmacol*. **88**(3): p. 273-84.
4. Kalapos, M.P. 1999. Methylglyoxal in living organisms: chemistry, biochemistry, toxicology and biological implications. *Toxicol Lett*. **110**(3): p. 145-75.
5. Kalapos, M.P. 2008. Methylglyoxal and glucose metabolism: a historical perspective and future avenues for research. *Drug Metabol Drug Interact*. **23**(1-2): p. 69-91.
6. Totemeyer, S., Booth, N.A., Nichols, W.W., Dunbar, B. and Booth, I.R. 1998. From famine to feast: the role of methylglyoxal production in *Escherichia coli*. *Mol Microbiol*. **27**(3): p. 553-62.
7. Krautwald, M. and Munch, G. 2011. Advanced glycation end products as biomarkers and gerontotoxins- A basis to explore methylglyoxal-lowering agents for Alzheimer's disease? *Exp Gerontol*. **45**(10): p. 744-51.
8. Thornalley, P.J. 2007. Dietary AGEs and ALEs and risk to human health by their interaction with the receptor for advanced glycation endproducts (RAGE)- an introduction. *Mol Nutr Food Res*. **51**(9): p. 1107-10.
9. Thornalley, P.J. 2008. Protein and nucleotide damage by glyoxal and methylglyoxal in physiological systems- role in ageing and disease. *Drug Metabol Drug Interact*. **23**(1-2): p. 125-50.
10. Gaens, K.H., van Der Kallen, C.J., van Greevenbroek, M.M., Feskens, E.J., Stehouwer, C.D. and Schalkwijk, C.G. 2008. Receptor for advanced glycation end product polymorphisms and type 2 diabetes: the CODAM study. *Ann N Y Acad Sci*. **1126**: p. 162-5.
11. Wang, S.H., Sun, Z.L., Guo, Y.J., Yuan, Y. and Yang, B.Q. 2009. Diabetes impairs hippocampal function *via* advanced glycation end product mediated new neuron generation in animals with diabetes-related depression. *Toxicol Sci*. **111**(1): p. 72-9.
12. Kuhla, B., Luth, H.J., Haferburg, D., Boeck, K., Arendt, T. and Munch, G. 2005. Methylglyoxal, glyoxal, and their detoxification in Alzheimer's disease. *Ann N Y Acad Sci*. **1043**: p. 211-6.
13. Thornalley, P.J. 2003. Glyoxalase I--structure, function and a critical role in the enzymatic defence against glycation. *Biochem Soc Trans*. **31**(Pt 6): p. 1343-8.
14. Thornalley, P.J. 2003. The enzymatic defence against glycation in health, disease and therapeutics: a symposium to examine the concept. *Biochem Soc Trans*. **31**(Pt 6): p. 1341-2.
15. Sukdeo, N. and Honek, J. 2008. Microbial glyoxalase enzymes: metalloenzymes controlling cellular levels of methylglyoxal. *Drug Metabol Drug Interact*. **23**(1-2): p. 29-50.

16. MacLean, M.J., Ness, L.S., Ferguson, G.P. and Booth, I.R. 1998. The role of glyoxalase I in the detoxification of methylglyoxal and in the activation of the KefB K^+ efflux system in *Escherichia coli*. *Mol Microbiol*. **27**(3): p. 563-71.
17. Akoachere, M., Iozef, R., Rahlfs, S., Deponte, M., Mannervik, B., Creighton, D.J., Schirmer, H. and Becker, K. 2005. Characterization of the glyoxalases of the malarial parasite *Plasmodium falciparum* and comparison with their human counterparts. *Biol Chem*. **386**(1): p. 41-52.
18. Ariza, A., Vickers, T.J., Greig, N., Armour, K.A., Dixon, M.J., Eggleston, I.M., Fairlamb, A.H. and Bond, C.S. 2006. Specificity of the trypanothione-dependent *Leishmania major* glyoxalase I: structure and biochemical comparison with the human enzyme. *Mol Microbiol*. **59**(4): p. 1239-48.
19. Clugston, S.L., Barnard, J.F., Kinach, R., Miedema, D., Ruman, R., Daub, E. and Honek, J.F. 1998. Overproduction and characterization of a dimeric non-zinc glyoxalase I from *Escherichia coli*: evidence for optimal activation by nickel ions. *Biochemistry*. **37**(24): p. 8754-63.
20. Clugston, S.L., Daub, E. and Honek, J.F. 1998. Identification of glyoxalase I sequences in *Brassica oleracea* and *Sporobolus stapfianus*: evidence for gene duplication events. *J Mol Evol*. **47**(2): p. 230-4.
21. Clugston, S.L., Daub, E., Kinach, R., Miedema, D., Barnard, J.F. and Honek, J.F. 1997. Isolation and sequencing of a gene coding for glyoxalase I activity from *Salmonella typhimurium* and comparison with other glyoxalase I sequences. *Gene*. **186**(1): p. 103-11.
22. Deswal, R. and Sopory, S.K. 1991. Purification and partial characterization of glyoxalase I from a higher plant *Brassica juncea*. *FEBS Lett*. **282**(2): p. 277-80.
23. Deswal, R. and Sopory, S.K. 1998. Biochemical and immunochemical characterization of *Brassica juncea* glyoxalase I. *Phytochemistry*. **49**(8): p. 2245-53.
24. Garcia-Iniguez, L., Powers, L., Chance, B., Sellin, S., Mannervik, B. and Mildvan, A.S. 1984. X-ray absorption studies of the Zn^{2+} site of glyoxalase I. *Biochemistry*. **23**(4): p. 685-9.
25. Greig, N., Wyllie, S., Vickers, T.J. and Fairlamb, A.H. 2006. Trypanothione-dependent glyoxalase I in *Trypanosoma cruzi*. *Biochem J*. **400**(2): p. 217-23.
26. Marmstal, E., Aronsson, A.C. and Mannervik, B. 1979. Comparison of glyoxalase I purified from yeast (*Saccharomyces cerevisiae*) with the enzyme from mammalian sources. *Biochem J*. **183**(1): p. 23-30.
27. Martins, A.M., Cordeiro, C.A. and Ponces Freire, A.M. 2001. *In situ* analysis of methylglyoxal metabolism in *Saccharomyces cerevisiae*. *FEBS Lett*. **499**(1-2): p. 41-4.
28. Rhee, H., Murata, K. and Kimura, A. 1986. Purification and characterization of glyoxalase I from *Pseudomonas putida*. *Biochem Biophys Res Commun*. **141**(3): p. 993-9.
29. Ridderstrom, M. and Mannervik, B. 1996. Optimized heterologous expression of the human zinc enzyme glyoxalase I. *Biochem J*. **314 (Pt 2)**: p. 463-7.
30. Saint-Jean, A.P., Phillips, K.R., Creighton, D.J. and Stone, M.J. 1998. Active monomeric and dimeric forms of *Pseudomonas putida* glyoxalase I: evidence for 3D domain swapping. *Biochemistry*. **37**(29): p. 10345-53.

31. Sellin, S., Eriksson, L.E., Aronsson, A.C. and Mannervik, B. 1983. Octahedral metal coordination in the active site of glyoxalase I as evidenced by the properties of Co(II)-glyoxalase I. *J Biol Chem*. **258**(4): p. 2091-3.
32. Sukdeo, N., Clugston, S.L., Daub, E. and Honek, J.F. 2004. Distinct classes of glyoxalase I: metal specificity of the *Yersinia pestis, Pseudomonas aeruginosa* and *Neisseria meningitidis* enzymes. *Biochem J*. **384**(Pt 1): p. 111-7.
33. Sukdeo, N. and Honek, J.F. 2007. *Pseudomonas aeruginosa* contains multiple glyoxalase I-encoding genes from both metal activation classes. *Biochim Biophys Acta*. **1774**(6): p. 756-63.
34. Takatsume, Y., Izawa, S. and Inoue, Y. 2004. Identification of thermostable glyoxalase I in the fission yeast *Schizosaccharomyces pombe*. *Arch Microbiol*. **181**(5): p. 371-7.
35. Vickers, T.J., Greig, N. and Fairlamb, A.H. 2004. A trypanothione-dependent glyoxalase I with a prokaryotic ancestry in *Leishmania major*. *Proc Natl Acad Sci USA*. **101**(36): p. 13186-91.
36. Martins, A.M., Mendes, P., Cordeiro, C. and Freire, A.P. 2001. *In situ* kinetic analysis of glyoxalase I and glyoxalase II in *Saccharomyces cerevisiae*. *Eur J Biochem*. **268**(14): p. 3930-6.
37. Clugston, S.L., Yajima, R. and Honek, J.F. 2004. Investigation of metal binding and activation of *Escherichia coli* glyoxalase I: kinetic, thermodynamic and mutagenesis studies. *Biochem J*. **377**(Pt 2): p. 309-16.
38. Su, Z., Sukdeo, N. and Honek, J.F. 2008. 15N-1H HSQC NMR evidence for distinct specificity of two active sites in *Escherichia coli* glyoxalase I. *Biochemistry*. **47**(50): p. 13232-41.
39. Iozef, R., Rahlfs, S., Chang, T., Schirmer, H. and Becker, K. 2003. Glyoxalase I of the malarial parasite *Plasmodium falciparum*: evidence for subunit fusion. *FEBS Lett*. **554**(3): p. 284-8.
40. Armstrong, R.N. 2000. Mechanistic diversity in a metalloenzyme superfamily. *Biochemistry*. **39**(45): p. 13625-32.
41. Bergdoll, M., Eltis, L.D., Cameron, A.D., Dumas, P. and Bolin, J.T. 1998. All in the family: structural and evolutionary relationships among three modular proteins with diverse functions and variable assembly. *Protein Sci*. **7**(8): p. 1661-70.
42. Feierberg, I., Cameron, A.D. and Aqvist, J. 1999. Energetics of the proposed rate-determining step of the glyoxalase I reaction. *FEBS Lett*. **453**(1-2): p. 90-4.
43. Deponte, M., Sturm, N., Mittler, S., Harner, M., Mack, H. and Becker, K. 2007. Allosteric coupling of two different functional active sites in monomeric *Plasmodium falciparum* glyoxalase I. *J Biol Chem*. **282**(39): p. 28419-30.
44. Cameron, A.D., Olin, B., Ridderstrom, M., Mannervik, B. and Jones, T.A. 1997. Crystal structure of human glyoxalase I- evidence for gene duplication and 3D domain swapping. *Embo J*. **16**(12): p. 3386-95.
45. He, M.M., Clugston, S.L., Honek, J.F. and Matthews, B.W. 2000. Determination of the structure of *Escherichia coli* glyoxalase I suggests a structural basis for differential metal activation. *Biochemistry*. **39**(30): p. 8719-27.
46. Cameron, A.D., Ridderstrom, M., Olin, B., Kavarana, M.J., Creighton, D.J. and Mannervik, B. 1999. Reaction mechanism of glyoxalase I explored by an X-ray crystallographic analysis of the human enzyme in complex with a transition state analogue. *Biochemistry*. **38**(41): p. 13480-90.

47. Davidson, G., Clugston, S.L., Honek, J.F. and Maroney, M.J. 2001. An XAS investigation of product and inhibitor complexes of Ni-containing GlxI from *Escherichia coli*: mechanistic implications. *Biochemistry*. **40**(15): p. 4569-82.
48. Himo, F. and Siegbahn, P.E. 2001. Catalytic mechanism of glyoxalase I: a theoretical study. *J Am Chem Soc*. **123**(42): p. 10280-9.
49. Richter, U. and Krauss, M. 2001. Active site structure and mechanism of human glyoxalase I- an ab initio theoretical study. *J Am Chem Soc*. **123**(29): p. 6973-82.
50. Ridderstrom, M., Cameron, A.D., Jones, T.A. and Mannervik, B. 1998. Involvement of an active-site Zn^{2+} ligand in the catalytic mechanism of human glyoxalase I. *J Biol Chem*. **273**(34): p. 21623-8.
51. Principato, G.B., Rosi, G., Talesa, V., Bocchini, V. and Giovannini, E. 1984. Purification of *S*-2-hydroxyacylglutathione hydrolase (Glyoxalase II) from calf brain. *Biochem Int*. **9**(3): p. 351-9.
52. Principato, G.B., Rosi, G., Talesa, V., Giovannini, E. and Uotila, L. 1987. Purification and characterization of two forms of glyoxalase II from the liver and brain of Wistar rats. *Biochim Biophys Acta*. **911**(3): p. 349-55.
53. Talesa, V., Principato, G.B., Norton, S.J., Contenti, S., Mangiabene, C. and Rosi, G. 1990. Isolation of glyoxalase II from bovine liver mitochondria. *Biochem Int*. **20**(1): p. 53-8.
54. Talesa, V., Uotila, L., Koivusalo, M., Principato, G., Giovannini, E. and Rosi, G. 1988. Demonstration of glyoxalase II in rat liver mitochondria. Partial purification and occurrence in multiple forms. *Biochim Biophys Acta*. **955**(1): p. 103-10.
55. Talesa, V., Uotila, L., Koivusalo, M., Principato, G., Giovannini, E. and Rosi, G. 1989. Isolation of glyoxalase II from two different compartments of rat liver mitochondria. Kinetic and immunochemical characterization of the enzymes. *Biochim Biophys Acta*. **993**(1): p. 7-11.
56. Bito, A., Haider, M., Briza, P., Strasser, P. and Breitenbach, M. 1999. Heterologous expression, purification, and kinetic comparison of the cytoplasmic and mitochondrial glyoxalase II enzymes, Glo2p and Glo4p, from *Saccharomyces cerevisiae. Protein Expr Purif*. **17**(3): p. 456-64.
57. Bito, A., Haider, M., Hadler, I. and Breitenbach, M. 1997. Identification and phenotypic analysis of two glyoxalase II encoding genes from *Saccharomyces cerevisiae*, GLO2 and GLO4, and intracellular localization of the corresponding proteins. *J Biol Chem*. **272**(34): p. 21509-19.
58. Cho, M.Y., Bae, C.D., Park, J.B. and Lee, T.H. 1998. Purification and cloning of glyoxalase II from rat liver. *Exp Mol Med*. **30**(1): p. 53-7.
59. Norton, S.J., Talesa, V., Yuan, W.J. and Principato, G.B. 1990. Glyoxalase I and glyoxalase II from *Aloe vera*: purification, characterization and comparison with animal glyoxalases. *Biochem Int*. **22**(3): p. 411-8.
60. Talesa, V., Rosi, G., Contenti, S., Mangiabene, C., Lupattelli, M., Norton, S.J., Giovannini, E. and Principato, G.B. 1990. Presence of glyoxalase II in mitochondria from spinach leaves: comparison with the enzyme from cytosol. *Biochem Int*. **22**(6): p. 1115-20.
61. Allen, R.E., Lo, T.W. and Thornalley, P.J. 1993. Purification and characterisation of glyoxalase II from human red blood cells. *Eur J Biochem*. **213**(3): p. 1261-7.

62. Ball, J.C. and Vander Jagt, D.L. 1979. Purification of *S*-2-hydroxyacylglutathione hydrolase (glyoxalase II) from rat erythrocytes. *Anal Biochem*. **98**(2): p. 472-7.
63. Campos-Bermudez, V.A., Leite, N.R., Krog, R., Costa-Filho, A.J., Soncini, F.C., Oliva, G. and Vila, A.J. 2007. Biochemical and structural characterization of *Salmonella typhimurium* glyoxalase II: new insights into metal ion selectivity. *Biochemistry*. **46**(39): p. 11069-79.
64. Kizil, G., Wilks, K., Wells, D. and Ala'Aldeen, D.A. 2000. Detection and characterisation of the genes encoding glyoxalase I and II from *Neisseria meningitidis*. *J Med Microbiol*. **49**(7): p. 669-73.
65. Murata, K. and Kimura, A. 1986. [Cloning of metabolic regulators and regulation of cell proliferation]. *Tanpakushitsu Kakusan Koso*. **31**(11): p. 1010-21.
66. Oray, B. and Norton, S.J. 1980. Purification and characterization of mouse liver glyoxalase II. *Biochim Biophys Acta*. **611**(1): p. 168-73.
67. O'Young, J., Sukdeo, N. and Honek, J.F. 2007. *Escherichia coli* glyoxalase II is a binuclear zinc-dependent metalloenzyme. *Arch Biochem Biophys*. **459**(1): p. 20-6.
68. Stamp, A.L., Owen, P., El Omari, K., Nichols, C.E., Lockyer, M., Lamb, H.K., Charles, I.G., Hawkins, A.R. and Stammers, D.K. 2010. Structural and functional characterization of *Salmonella enterica* serovar *Typhimurium* YcbL: an unusual Type II glyoxalase. *Protein Sci.* **19**(10): p. 1897-905.
69. Talesa, V., Rosi, G., Bistoni, F., Marconi, P., Norton, S.J. and Principato, G.B. 1990. Presence of a plant-like glyoxalase II in *Candida albicans*. *Biochem Int*. **21**(3): p. 397-403.
70. Uotila, L. 1973. Purification and characterization of *S*-2-hydroxyacylglutathione hydrolase (glyoxalase II) from human liver. *Biochemistry*. **12**(20): p. 3944-51.
71. Daiyasu, H., Osaka, K., Ishino, Y. and Toh, H. 2001. Expansion of the zinc metallo-hydrolase family of the beta-lactamase fold. *FEBS Lett*. **503**(1): p. 1-6.
72. Wenzel, N.F., Carenbauer, A.L., Pfiester, M.P., Schilling, O., Meyer-Klaucke, W., Makaroff, C.A. and Crowder, M.W. 2004. The binding of iron and zinc to glyoxalase II occurs exclusively as di-metal centers and is unique within the metallo-beta-lactamase family. *J Biol Inorg Chem*. **9**(4): p. 429-38.
73. Cameron, A.D., Ridderstrom, M., Olin, B. and Mannervik, B. 1999. Crystal structure of human glyoxalase II and its complex with a glutathione thiolester substrate analogue. *Structure*. **7**(9): p. 1067-78.
74. Marasinghe, G.P., Sander, I.M., Bennett, B., Periyannan, G., Yang, K.W., Makaroff, C.A. and Crowder, M.W. 2005. Structural studies on a mitochondrial glyoxalase II. *J Biol Chem*. **280**(49): p. 40668-75.
75. Schilling, O., Wenzel, N., Naylor, M., Vogel, A., Crowder, M., Makaroff, C. and Meyer-Klaucke, W. 2003. Flexible metal binding of the metallo-beta-lactamase domain: glyoxalase II incorporates iron, manganese, and zinc *in vivo*. *Biochemistry*. **42**(40): p. 11777-86.
76. Limphong, P., McKinney, R.M., Adams, N.E., Bennett, B., Makaroff, C.A., Gunasekera, T. and Crowder, M.W. 2009. Human glyoxalase II contains an Fe(II)Zn(II) center but is active as a mononuclear Zn(II) enzyme. *Biochemistry*. **48**(23): p. 5426-34.

77. Ridderstrom, M., Jemth, P., Cameron, A.D. and Mannervik, B. 2000. The active-site residue Tyr-175 in human glyoxalase II contributes to binding of glutathione derivatives. *Biochim Biophys Acta*. **1481**(2): p. 344-8.
78. Misra, K., Banerjee, A.B., Ray, S. and Ray, M. 1995. Glyoxalase III from *Escherichia coli*: a single novel enzyme for the conversion of methylglyoxal into D-lactate without reduced glutathione. *Biochem J*. **305** (Pt 3): p. 999-1003.
79. Benov, L., Sequeira, F. and Beema, A.F. 2004. Role of rpoS in the regulation of glyoxalase III in *Escherichia coli*. *Acta Biochim Pol*. **51**(3): p. 857-60.
80. Newton, G.L., Bewley, C.A., Dwyer, T.J., Horn, R., Aharonowitz, Y., Cohen, G., Davies, J., Faulkner, D.J. and Fahey, R.C. 1995. The structure of U17 isolated from *Streptomyces clavuligerus* and its properties as an antioxidant thiol. *Eur J Biochem*. **230**(2): p. 821-5.
81. Fahey, R.C., Brown, W.C., Adams, W.B. and Worsham, M.B. 1978. Occurrence of glutathione in bacteria. *J Bacteriol*. **133**(3): p. 1126-9.
82. Meister, A. and Anderson, M.E. 1983. Glutathione. *Annu Rev Biochem*. **52**: p. 711-60.
83. Newton, G.L., Arnold, K., Price, M.S., Sherrill, C., Delcardayre, S.B., Aharonowitz, Y., Cohen, G., Davies, J., Fahey, R.C. and Davis, C. 1996. Distribution of thiols in microorganisms: mycothiol is a major thiol in most actinomycetes. *J Bacteriol*. **178**(7): p. 1990-5.
84. Pastore, A., Federici, G., Bertini, E. and Piemonte, F. 2003. Analysis of glutathione: implication in redox and detoxification. *Clin Chim Acta*. **333**(1): p. 19-39.
85. Vander Jagt, D.L., Han, L.P. and Lehman, C.H. 1972. Kinetic evaluation of substrate specificity in the glyoxalase-I-catalyzed disproportionation of -ketoaldehydes. *Biochemistry*. **11**(20): p. 3735-40.
86. Vince, R., Daluge, S. and Wadd, W.B. 1971. Studies on the inhibition of glyoxalase I by *S*-substituted glutathiones. *J Med Chem*. **14**(5): p. 402-4.
87. Kosower, N.S. and Kosower, K.E. 1978. The glutathione status of cells. *Int Rev Cytol*. **54**: p. 109-60.
88. Richard, J.P. 1991. Kinetic parameters for the elimination reaction catalyzed by triosephosphate isomerase and an estimation of the reaction's physiological significance. *Biochemistry*. **30**(18): p. 4581-5.
89. Richard, J.P. 1993. Mechanism for the formation of methylglyoxal from triosephosphates. *Biochem Soc Trans*. **21**(2): p. 549-53.
90. Romao, P.R., Tovar, J., Fonseca, S.G., Moraes, R.H., Cruz, A.K., Hothersall, J.S., Noronha-Dutra, A.A., Ferreira, S.H. and Cunha, F.Q. 2006. Glutathione and the redox control system trypanothione/trypanothione reductase are involved in the protection of *Leishmania spp*. against nitrosothiol-induced cytotoxicity. *Braz J Med Biol Res*. **39**(3): p. 355-63.
91. Oza, S.L., Shaw, M.P., Wyllie, S. and Fairlamb, A.H. 2005. Trypanothione biosynthesis in *Leishmania major*. *Mol Biochem Parasitol*. **139**(1): p. 107-16.
92. Thomson, L., Denicola, A. and Radi, R. 2003. The trypanothione-thiol system in *Trypanosoma cruzi* as a key antioxidant mechanism against peroxynitrite-mediated cytotoxicity. *Arch Biochem Biophys*. **412**(1): p. 55-64.
93. Hand, C.E., Auzanneau, F.I. and Honek, J.F. 2006. Conformational analyses of mycothiol, a critical intracellular glycothiol in Mycobacteria. *Carbohydr Res*. **341**(9): p. 1164-73.

94. Hand, C.E. and Honek, J.F. 2005. Biological chemistry of naturally occurring thiols of microbial and marine origin. *J Nat Prod*. **68**(2): p. 293-308.
95. Newton, G.L., Av-Gay, Y. and Fahey, R.C. 2000. A novel mycothiol-dependent detoxification pathway in mycobacteria involving mycothiol *S*-conjugate amidase. *Biochemistry*. **39**(35): p. 10739-46.
96. Newton, G.L., Buchmeier, N. and Fahey, R.C. 2008. Biosynthesis and functions of mycothiol, the unique protective thiol of Actinobacteria. *Microbiol Mol Biol Rev*. **72**(3): p. 471-94.
97. Newton, G.L. and Fahey, R.C. 2002. Mycothiol biochemistry. *Arch Microbiol*. **178**(6): p. 388-94.
98. Rawat, M. and Av-Gay, Y. 2007. Mycothiol-dependent proteins in actinomycetes. *FEMS Microbiol Rev*. **31**(3): p. 278-92.
99. Misset-Smits, M., van Ophem, P.W., Sakuda, S. and Duine, J.A. 1997. Mycothiol, 1-*O*-(2'-[*N*-acetyl-L-cysteinyl]amido-*2'*-deoxy-alpha-D-glucopyranosyl)-D-*myo*-inositol, is the factor of NAD/factor-dependent formaldehyde dehydrogenase. *FEBS Lett*. **409**(2): p. 221-2.
100. Feng, J., Che, Y., Milse, J., Yin, Y.J., Liu, L., Ruckert, C., Shen, X.H., Qi, S.W., Kalinowski, J. and Liu, S.J. 2006. The gene ncgl2918 encodes a novel maleylpyruvate isomerase that needs mycothiol as cofactor and links mycothiol biosynthesis and gentisate assimilation in *Corynebacterium glutamicum*. *J Biol Chem*. **281**(16): p. 10778-85.
101. Buchmeier, N.A., Newton, G.L., Koledin, T. and Fahey, R.C. 2003. Association of mycothiol with protection of *Mycobacterium tuberculosis* from toxic oxidants and antibiotics. *Mol Microbiol*. **47**(6): p. 1723-32.
102. Newton, G.L., Unson, M.D., Anderberg, S.J., Aguilera, J.A., Oh, N.N., delCardayre, S.B., Av-Gay, Y. and Fahey, R.C. 1999. Characterization of *Mycobacterium smegmatis* mutants defective in 1-D-*myo*-inosityl-2-amino-2-deoxy-alpha-D-glucopyranoside and mycothiol biosynthesis. *Biochem Biophys Res Commun*. **255**(2): p. 239-44.
103. Ordonez, E., Van Belle, K., Roos, G., De Galan, S., Letek, M., Gil, J.A., Wyns, L., Mateos, L.M. and Messens, J. 2009. Arsenate reductase, mycothiol, and mycoredoxin concert thiol/disulfide exchange. *J Biol Chem*. **284**(22): p. 15107-16.
104. Halloran, M.E. 1994. *Mycobacterium tuberculosis*: just desserts for an ungrateful guest. *Trends Ecol Evol*. **9**(2): p. 72-5.
105. McCarthy, A.A., Baker, H.M., Shewry, S.C., Patchett, M.L. and Baker, E.N. 2001. Crystal structure of methylmalonyl-coenzyme A epimerase from *P. shermanii*: a novel enzymatic function on an ancient metal binding scaffold. *Structure*. **9**(7): p. 637-46.
106. Bobik, T.A. and Rasche, M.E. 2001. Identification of the human methylmalonyl-CoA racemase gene based on the analysis of prokaryotic gene arrangements. Implications for decoding the human genome. *J Biol Chem*. **276**(40): p. 37194-8.
107. Bobik, T.A. and Rasche, M.E. 2003. HPLC assay for methylmalonyl-CoA epimerase. *Anal Bioanal Chem*. **375**(3): p. 344-9.
108. Bobik, T.A. and Rasche, M.E. 2004. Purification and partial characterization of the *Pyrococcus horikoshii* methylmalonyl-CoA epimerase. *Appl Microbiol Biotechnol*. **63**(6): p. 682-5.

109. Beharry, Z. and Palzkill, T. 2005. Functional analysis of active site residues of the fosfomycin resistance enzyme FosA from *Pseudomonas aeruginosa*. *J Biol Chem*. **280**(18): p. 17786-91.
110. Bernat, B.A., Laughlin, L.T. and Armstrong, R.N. 1997. Fosfomycin resistance protein (FosA) is a manganese metalloglutathione transferase related to glyoxalase I and the extradiol dioxygenases. *Biochemistry*. **36**(11): p. 3050-5.
111. Bernat, B.A., Laughlin, L.T. and Armstrong, R.N. 1999. Elucidation of a monovalent cation dependence and characterization of the divalent cation binding site of the fosfomycin resistance protein (FosA). *Biochemistry*. **38**(23): p. 7462-9.
112. Brown, D.W., Schaab, M.R., Birmingham, W.R. and Armstrong, R.N. 2009. Evolution of the antibiotic resistance protein, FosA, is linked to a catalytically promiscuous progenitor. *Biochemistry*. **48**(9): p. 1847-9.
113. Rigsby, R.E., Brown, D.W., Dawson, E., Lybrand, T.P. and Armstrong, R.N. 2007. A model for glutathione binding and activation in the fosfomycin resistance protein, FosA. *Arch Biochem Biophys*. **464**(2): p. 277-83.
114. Hamilton, D.S. and Creighton, D.J. 1992. Inhibition of glyoxalase I by the enediol mimic *S*-(*N*-hydroxy-*N*-methylcarbamoyl)glutathione. The possible basis of a tumor-selective anticancer strategy. *J Biol Chem*. **267**(35): p. 24933-6.
115. Murthy, N.S., Bakeris, T., Kavarana, M.J., Hamilton, D.S., Lan, Y. and Creighton, D.J. 1994. *S*-(*N*-aryl-*N*-hydroxycarbamoyl)glutathione derivatives are tight-binding inhibitors of glyoxalase I and slow substrates for glyoxalase II. *J Med Chem*. **37**(14): p. 2161-6.
116. Patel, M.P. and Blanchard, J.S. 1998. Synthesis of des-*myo*-inositol mycothiol and demonstration of a mycobacterial specific reductase activity. *J Am Chem Soc*. **120**: p. 11538-9.
117. Hatheway, C.L. 1990. Toxigenic clostridia. *Clin Microbiol Rev*. **3**(1): p. 66-98.
118. Bayer, E.A., Lamed, R., White, B.A. and Flint, H.J. 2008. From cellulosomes to cellulosomics. *Chem Rec*. **8**(6): p. 364-77.
119. Demain, A.L., Newcomb, M. and Wu, J.H. 2005. Cellulase, clostridia, and ethanol. *Microbiol Mol Biol Rev*. **69**(1): p. 124-54.
120. Mitchell, W.J. 1998. Physiology of carbohydrate to solvent conversion by clostridia. *Adv Microb Physiol*. **39**: p. 31-130.
121. Ting, P.T. and Freiman, A. 2004. The story of *Clostridium botulinum*: from food poisoning to Botox. *Clin Med*. **4**(3): p. 258-61.
122. Aktories, K. 1997. Bacterial toxins that target Rho proteins. *J Clin Invest*. **99**(5): p. 827-9.
123. Bermudes, D., Zheng, L.M. and King, I.C. 2002. Live bacteria as anticancer agents and tumor-selective protein delivery vectors. *Curr Opin Drug Discov Devel*. **5**(2): p. 194-9.
124. Fenicia, L. and Anniballi, F. 2009. Infant botulism. *Ann Ist Super Sanita*. **45**(2): p. 134-46.
125. Poulain, B., Lonchamp, E., Jover, E., Popoff, M.R. and Molgo, J. 2009. [Mechanisms of action of botulinum toxins and neurotoxins]. *Ann Dermatol Venereol*. **136** (Suppl 4): p. S73-6.
126. Popoff, M.R. and Bouvet, P. 2009. Clostridial toxins. *Future Microbiol*. **4**(8): p. 1021-64.

127. Turton, K., Chaddock, J.A. and Acharya, K.R. 2002. Botulinum and tetanus neurotoxins: structure, function and therapeutic utility. *Trends Biochem Sci*. **27**(11): p. 552-8.
128. Goonetilleke, A. and Harris, J.B. 2004. Clostridial neurotoxins. *J Neurol Neurosurg Psychiatry*. **75** (Suppl 3): p. iii35-9.
129. Dixon, B. 1996. Chaim Weizmann: from fermentation chemist to president of Israel. *Anaerobe*. **2**(4): p. 195-6.
130. Nolling, J., Breton, G., Omelchenko, M.V., Makarova, K.S., Zeng, Q., Gibson, R., Lee, H.M., Dubois, J., Qiu, D., Hitti, J., Wolf, Y.I., Tatusov, R.L., Sabathe, F., Doucette-Stamm, L., Soucaille, P., Daly, M.J., Bennett, G.N., Koonin, E.V. and Smith, D.R. 2001. Genome sequence and comparative analysis of the solvent-producing bacterium *Clostridium acetobutylicum*. *J Bacteriol*. **183**(16): p. 4823-38.
131. Gasteiger, E., Jung, E. and Bairoch, A. 2001. SWISS-PROT: connecting biomolecular knowledge *via* a protein database. *Curr Issues Mol Biol*. **3**(3): p. 47-55.
132. Krissinel, E. and Henrick, K. 2007. Inference of macromolecular assemblies from crystalline state. *J Mol Biol*. **372**(3): p. 774-97.
133. Holm, L., Kaariainen, S., Rosenstrom, P. and Schenkel, A. 2008. Searching protein structure databases with DaliLite v.3. *Bioinformatics*. **24**(23): p. 2780-1.
134. Bradford, M.M. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem*. **72**: p. 248-54.
135. Sambrook, J. and Russell, D. 2001. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
136. Huet-Minkowski, M., Gasson, J.C. and Bourgeois, S. 1981. Induction of glucocorticoid-resistant variants in a murine thymoma line by antitumor drugs. *Cancer Res*. **41**(11 Pt 1): p. 4540-6.
137. McCall, K.A. and Fierke, C.A. 2000. Colorimetric and fluorimetric assays to quantitate micromolar concentrations of transition metals. *Anal Biochem*. **284**(2): p. 307-15.
138. Pfister, T., Jones, K.W. and Wimmer, E. 2000. A cysteine-rich motif in poliovirus protein 2C(ATPase) is involved in RNA replication and binds zinc *in vitro*. *J Virol*. **74**(1): p. 334-43.
139. Bergmeyer, H., Grassl, M. and Walter, H. 1983. Methods of Enzymatic Analysis, 3rd ed. (Bergmeyer, H.U., Ed.), Verlag Chemie, Deerfield Beach, FL. **2**: p. 217-8.
140. Riddles, P.W., Blakeley, R.L. and Zerner, B. 1979. Ellman's reagent: 5,5'-dithiobis(2-nitrobenzoic acid)- a reexamination. *Anal Biochem*. **94**(1): p. 75-81.
141. Cliffe, E.E. and Waley, S.G. 1961. The mechanism of the glyoxalase I reaction, and the effect of ophthalmic acid as an inhibitor. *Biochem J*. **79**: p. 475-82.
142. Combet, C., Jambon, M., Deleage, G. and Geourjon, C. 2002. Geno3D: automatic comparative molecular modelling of protein. *Bioinformatics*. **18**(1): p. 213-4.
143. Serre, L., Sailland, A., Sy, D., Boudec, P., Rolland, A., Pebay-Peyroula, E. and Cohen-Addad, C. 1999. Crystal structure of Pseudomonas fluorescens 4-hydroxyphenylpyruvate dioxygenase: an enzyme involved in the tyrosine degradation pathway. *Structure*. **7**(8): p. 977-88.
144. Kita, A., Kita, S., Fujisawa, I., Inaka, K., Ishida, T., Horiike, K., Nozaki, M. and Miki, K. 1999. An archetypical extradiol-cleaving catecholic dioxygenase: the crystal

structure of catechol 2,3-dioxygenase (metapyrocatechase) from *Pseudomonas putida* mt-2. *Structure*. **7**(1): p. 25-34.
145. Han, S., Eltis, L.D., Timmis, K.N., Muchmore, S.W. and Bolin, J.T. 1995. Crystal structure of the biphenyl-cleaving extradiol dioxygenase from a PCB-degrading pseudomonad. *Science*. **270**(5238): p. 976-80.
146. Chen, Z.Y., Brown, R.L., Damann, K.E. and Cleveland, T.E. 2004. Identification of a maize kernel stress-related protein and its effect on aflatoxin accumulation. *Phytopathology*. **94**(9): p. 938-45.
147. Brophy, P.M., Crowley, P. and Barrett, J. 1990. Relative distribution of glutathione transferase, glyoxalase I and glyoxalase II in helminths. *Int J Parasitol*. **20**(2): p. 259-61.
148. Liyanage, H., Kashket, S., Young, M. and Kashket, E.R. 2001. *Clostridium beijerinckii* and *Clostridium difficile* detoxify methylglyoxal by a novel mechanism involving glycerol dehydrogenase. *Appl Environ Microbiol*. **67**(5): p. 2004-10.
149. Kino, K., Kuratsu, S., Noguchi, A., Kokubo, M., Nakazawa, Y., Arai, T., Yagasaki, M. and Kirimura, K. 2007. Novel substrate specificity of glutathione synthesis enzymes from *Streptococcus agalactiae* and *Clostridium acetobutylicum*. *Biochem Biophys Res Commun*. **352**(2): p. 351-9.
150. Birkenmeier, G., Stegemann, C., Hoffmann, R., Gunther, R., Huse, K. and Birkemeyer, C. 2010. Posttranslational modification of human glyoxalase 1 indicates redox-dependent regulation. *PLoS One*. **5**(4): p. e10399.
151. Frickel, E.M., Jemth, P., Widersten, M. and Mannervik, B. 2001. Yeast glyoxalase I is a monomeric enzyme with two active sites. *J Biol Chem*. **276**(3): p. 1845-9.
152. Andrade, M.A., Chacon, P., Merelo, J.J. and Moran, F. 1993. Evaluation of secondary structure of proteins from UV circular dichroism spectra using an unsupervised learning neural network. *Protein Eng*. **6**(4): p. 383-90.
153. Merelo, J.J., Andrade, M.A., Prieto, A. and Morán, F. 1994. Proteinotopic Feature Maps. *Neurocomputing*. **6**: p. 443-54.
154. Nicolle, L.E. 2002. Urinary tract infection: traditional pharmacologic therapies. *Am J Med*. **113** (Suppl 1A): p. 35S-44S.
155. Hidaka, T., Goda, M., Kuzuyama, T., Takei, N., Hidaka, M. and Seto, H. 1995. Cloning and nucleotide sequence of fosfomycin biosynthetic genes of *Streptomyces wedmorensis*. *Mol Gen Genet*. **249**(3): p. 274-80.
156. Higgins, L.J., Yan, F., Liu, P., Liu, H.W. and Drennan, C.L. 2005. Structural insight into antibiotic fosfomycin biosynthesis by a mononuclear iron enzyme. *Nature*. **437**(7060): p. 838-44.
157. Woodyer, R.D., Shao, Z., Thomas, P.M., Kelleher, N.L., Blodgett, J.A., Metcalf, W.W., van der Donk, W.A. and Zhao, H. 2006. Heterologous production of fosfomycin and identification of the minimal biosynthetic gene cluster. *Chem Biol*. **13**(11): p. 1171-82.
158. Metcalf, W.W. and van der Donk, W.A. 2009. Biosynthesis of phosphonic and phosphinic acid natural products. *Annu Rev Biochem*. **78**: p. 65-94.
159. El Zoeiby, A., Sanschagrin, F. and Levesque, R.C. 2003. Structure and function of the Mur enzymes: development of novel inhibitors. *Mol Microbiol*. **47**(1): p. 1-12.
160. Kumar, S., Parvathi, A., Hernandez, R.L., Cadle, K.M. and Varela, M.F. 2009. Identification of a novel UDP-*N*-acetylglucosamine enolpyruvyl transferase (MurA)

from *Vibrio fischeri* that confers high fosfomycin resistance in *Escherichia coli*. *Arch Microbiol*. **191**(5): p. 425-9.
161. Eschenburg, S., Priestman, M. and Schonbrunn, E. 2005. Evidence that the fosfomycin target Cys115 in UDP-*N*-acetylglucosamine enolpyruvyl transferase (MurA) is essential for product release. *J Biol Chem*. **280**(5): p. 3757-63.
162. Skarzynski, T., Kim, D.H., Lees, W.J., Walsh, C.T. and Duncan, K. 1998. Stereochemical course of enzymatic enolpyruvyl transfer and catalytic conformation of the active site revealed by the crystal structure of the fluorinated analogue of the reaction tetrahedral intermediate bound to the active site of the C115A mutant of MurA. *Biochemistry*. **37**(8): p. 2572-7.
163. Hendlin, D., Stapley, E.O., Jackson, M., Wallick, H., Miller, A.K., Wolf, F.J., Miller, T.W., Chaiet, L., Kahan, F.M., Foltz, E.L., Woodruff, H.B., Mata, J.M., Hernandez, S. and Mochales, S. 1969. Phosphonomycin, a new antibiotic produced by strains of streptomyces. *Science*. **166**(901): p. 122-3.
164. Kobayashi, S., Kuzuyama, T. and Seto, H. 2000. Characterization of the fomA and fomB gene products from *Streptomyces wedmorensis*, which confer fosfomycin resistance on *Escherichia coli*. *Antimicrob Agents Chemother*. **44**(3): p. 647-50.
165. Pakhomova, S., Bartlett, S.G., Augustus, A., Kuzuyama, T. and Newcomer, M.E. 2008. Crystal structure of fosfomycin resistance kinase FomA from *Streptomyces wedmorensis*. *J Biol Chem*. **283**(42): p. 28518-26.
166. Garcia, P., Arca, P. and Evaristo Suarez, J. 1995. Product of fosC, a gene from *Pseudomonas syringae*, mediates fosfomycin resistance by using ATP as cosubstrate. *Antimicrob Agents Chemother*. **39**(7): p. 1569-73.
167. Fillgrove, K.L., Pakhomova, S., Newcomer, M.E. and Armstrong, R.N. 2003. Mechanistic diversity of fosfomycin resistance in pathogenic microorganisms. *J Am Chem Soc*. **125**(51): p. 15730-1.
168. Laughlin, L.T., Bernat, B.A. and Armstrong, R.N. 1998. Mechanistic imperative for the evolution of a metalloglutathione transferase of the vicinal oxygen chelate superfamily. *Chem Biol Interact*. **111-112**: p. 41-50.
169. Rife, C.L., Pharris, R.E., Newcomer, M.E. and Armstrong, R.N. 2002. Crystal structure of a genomically encoded fosfomycin resistance protein (FosA) at 1.19 Å resolution by MAD phasing off the L-III edge of Tl(+). *J Am Chem Soc*. **124**(37): p. 11001-3.
170. Bernat, B.A. and Armstrong, R.N. 2001. Elementary steps in the acquisition of Mn^{2+} by the fosfomycin resistance protein (FosA). *Biochemistry*. **40**(42): p. 12712-8.
171. Rigsby, R.E., Rife, C.L., Fillgrove, K.L., Newcomer, M.E. and Armstrong, R.N. 2004. Phosphonoformate: a minimal transition state analogue inhibitor of the fosfomycin resistance protein, FosA. *Biochemistry*. **43**(43): p. 13666-73.
172. Pakhomova, S., Rife, C.L., Armstrong, R.N. and Newcomer, M.E. 2004. Structure of fosfomycin resistance protein FosA from transposon Tn2921. *Protein Sci*. **13**(5): p. 1260-5.
173. Lan, Y., Lu, T., Lovett, P.S. and Creighton, D.J. 1995. Evidence for a (triosephosphate isomerase-like) "catalytic loop" near the active site of glyoxalase I. *J Biol Chem*. **270**(22): p. 12957-60.
174. Stokvis, E., Clugston, S.L., Honek, J.F. and Heck, A.J. 2000. Characterization of glyoxalase I (*E. coli*)-inhibitor interactions by electrospray time-of-flight mass spectrometry and enzyme kinetic analysis. *J Protein Chem*. **19**(5): p. 389-97.

175. Smoukov, S.K., Telser, J., Bernat, B.A., Rife, C.L., Armstrong, R.N. and Hoffman, B.M. 2002. EPR study of substrate binding to the Mn(II) active site of the bacterial antibiotic resistance enzyme FosA: a better way to examine Mn(II). *J Am Chem Soc*. **124**(10): p. 2318-26.
176. Walsby, C.J., Telser, J., Rigsby, R.E., Armstrong, R.N. and Hoffman, B.M. 2005. Enzyme control of small-molecule coordination in FosA as revealed by 31P pulsed ENDOR and ESE-EPR. *J Am Chem Soc*. **127**(23): p. 8310-9.
177. Chen, J., Ghorai, M.K., Kenney, G. and Stubbe, J. 2008. Mechanistic studies on bleomycin-mediated DNA damage: multiple binding modes can result in double-stranded DNA cleavage. *Nucleic Acids Res*. **36**(11): p. 3781-90.
178. Goodwin, K.D., Lewis, M.A., Long, E.C. and Georgiadis, M.M. 2008. Crystal structure of DNA-bound Co(III) bleomycin B2: Insights on intercalation and minor groove binding. *Proc Natl Acad Sci USA*. **105**(13): p. 5052-6.
179. Hecht, S.M. 2000. Bleomycin: new perspectives on the mechanism of action. *J Nat Prod*. **63**(1): p. 158-68.
180. Galm, U., Wang, L., Wendt-Pienkowski, E., Yang, R., Liu, W., Tao, M., Coughlin, J.M. and Shen, B. 2008. *In vivo* manipulation of the bleomycin biosynthetic gene cluster in *Streptomyces verticillus* ATCC15003 revealing new insights into its biosynthetic pathway. *J Biol Chem*. **283**(42): p. 28236-45.
181. Shen, B., Du, L., Sanchez, C., Edwards, D.J., Chen, M. and Murrell, J.M. 2001. The biosynthetic gene cluster for the anticancer drug bleomycin from *Streptomyces verticillus* ATCC15003 as a model for hybrid peptide-polyketide natural product biosynthesis. *J Ind Microbiol Biotechnol*. **27**(6): p. 378-85.
182. Akkerman, M.A., Haasnoot, C.A. and Hilbers, C.W. 1988. Studies of the solution structure of the bleomycin-A2-zinc complex by means of two-dimensional NMR spectroscopy and distance geometry calculations. *Eur J Biochem*. **173**(1): p. 211-25.
183. Otvos, J.D., Antholine, W.E., Wehrli, S. and Petering, D.H. 1996. Metal coordination environment and dynamics in 113cadmium bleomycin: relationship to zinc bleomycin. *Biochemistry*. **35**(5): p. 1458-65.
184. Lehmann, T.E., Ming, L.J., Rosen, M.E. and Que, L. Jr. 1997. NMR studies of the paramagnetic complex Fe(II)-bleomycin. *Biochemistry*. **36**(10): p. 2807-16.
185. Sugiyama, M., Thompson, C.J., Kumagai, T., Suzuki, K., Deblaere, R., Villarroel, R. and Davies, J. 1994. Characterization by molecular cloning of two genes from *Streptomyces verticillus* encoding resistance to bleomycin. *Gene*. **151**(1-2): p. 11-6.
186. Kawano, Y., Kumagai, T., Muta, K., Matoba, Y., Davies, J. and Sugiyama, M. 2000. The 1.5 Å crystal structure of a bleomycin resistance determinant from bleomycin-producing *Streptomyces verticillus*. *J Mol Biol*. **295**(4): p. 915-25.
187. Sugiyama, M., Kumagai, T., Hayashida, M., Maruyama, M. and Matoba, Y. 2002. The 1.6-Å crystal structure of the copper(II)-bound bleomycin complexed with the bleomycin-binding protein from bleomycin-producing Streptomyces verticillus. *J Biol Chem*. **277**(3): p. 2311-20.
188. Oda, K., Matoba, Y., Noda, M., Kumagai, T. and Sugiyama, M. 2010. Catalytic mechanism of bleomycin *N*-acetyltransferase proposed on the basis of its crystal structure. *J Biol Chem*. **285**(2): p. 1446-56.

189. Sugiyama, M., Kumagai, T., Shionoya, M., Kimura, E. and Davies, J.E. 1994. Inactivation of bleomycin by an *N*-acetyltransferase in the bleomycin-producing strain *Streptomyces verticillus*. *FEMS Microbiol Lett*. **121**(1): p. 81-5.
190. Drocourt, D., Calmels, T., Reynes, J.P., Baron, M. and Tiraby, G. 1990. Cassettes of the *Streptoalloteichus hindustanus* ble gene for transformation of lower and higher eukaryotes to phleomycin resistance. *Nucleic Acids Res*. **18**(13): p. 4009.
191. Gatignol, A., Durand, H. and Tiraby, G. 1988. Bleomycin resistance conferred by a drug-binding protein. *FEBS Lett*. **230**(1-2): p. 171-5.
192. Calcutt, M.J. and Schmidt, F.J. 1994. Bleomycin biosynthesis: structure of the resistance genes of the producer organism. *Ann N Y Acad Sci*. **721**: p. 133-7.
193. Galm, U., Hager, M.H., Van Lanen, S.G., Ju, J., Thorson, J.S. and Shen, B. 2005. Antitumor antibiotics: bleomycin, enediynes, and mitomycin. *Chem Rev*. **105**(2): p. 739-58.
194. Dumas, P., Bergdoll, M., Cagnon, C. and Masson, J.M. 1994. Crystal structure and site-directed mutagenesis of a bleomycin resistance protein and their significance for drug sequestering. *Embo J*. **13**(11): p. 2483-92.
195. Danshiitsoodol, N., de Pinho, C.A., Matoba, Y., Kumagai, T. and Sugiyama, M. 2006. The mitomycin C (MMC)-binding protein from MMC-producing microorganisms protects from the lethal effect of bleomycin: crystallographic analysis to elucidate the binding mode of the antibiotic to the protein. *J Mol Biol*. **360**(2): p. 398-408.
196. Kumagai, T., Nakano, T., Maruyama, M., Mochizuki, H. and Sugiyama, M. 1999. Characterization of the bleomycin resistance determinant encoded on the transposon Tn5. *FEBS Lett*. **442**(1): p. 34-8.
197. Sugiyama, M., Kumagai, T., Matsuo, H., Bhuiyan, M.Z., Ueda, K., Mochizuki, H., Nakamura, N. and Davies, J.E. 1995. Overproduction of the bleomycin-binding proteins from bleomycin-producing *Streptomyces verticillus* and a methicillin-resistant *Staphylococcus aureus* in *Escherichia coli* and their immunological characterisation. *FEBS Lett*. **362**(1): p. 80-4.
198. Seow, H.A., Penketh, P.G., Baumann, R.P. and Sartorelli, A.C. 2004. Bioactivation and resistance to mitomycin C. *Methods Enzymol*. **382**: p. 221-33.
199. Tomasz, M. 1995. Mitomycin C: small, fast and deadly (but very selective). *Chem Biol*. **2**(9): p. 575-9.
200. Tomasz, M., Chawla, A.K. and Lipman, R. 1988. Mechanism of monofunctional and bifunctional alkylation of DNA by mitomycin C. *Biochemistry*. **27**(9): p. 3182-7.
201. McClelland, R.A. and Lam, K. 1985. Kinetics and mechanism of the acid hydrolysis of mitomycins. *J Am Chem Soc*. **107**(18): p. 5182-6.
202. Palom, Y., Belcourt, M.F., Kumar, G.S., Arai, H., Kasai, M., Sartorelli, A.C., Rockwell, S. and Tomasz, M. 1998. Formation of a major DNA adduct of the mitomycin metabolite 2,7-diaminomitosene in EMT6 mouse mammary tumor cells treated with mitomycin C. *Oncol Res*. **10**(10): p. 509-21.
203. August, P.R., Flickinger, M.C. and Sherman, D.H. 1994. Cloning and analysis of a locus (mcr) involved in mitomycin C resistance in *Streptomyces lavendulae*. *J Bacteriol*. **176**(14): p. 4448-54.
204. Sheldon, P.J., Johnson, D.A., August, P.R., Liu, H.W. and Sherman, D.H. 1997. Characterization of a mitomycin-binding drug resistance mechanism from the producing organism, *Streptomyces lavendulae*. *J Bacteriol*. **179**(5): p. 1796-804.

205. Belcourt, M.F., Penketh, P.G., Hodnick, W.F., Johnson, D.A., Sherman, D.H., Rockwell, S. and Sartorelli, A.C. 1999. Mitomycin resistance in mammalian cells expressing the bacterial mitomycin C resistance protein MCRA. *Proc Natl Acad Sci USA*. **96**(18): p. 10489-94.
206. Johnson, D., August, P., Shackleton, C., Liu, H. and Sherman, D. 1997. Microbial resistance to mitomycins involves a redox relay mechanism. *J Am Chem Soc*. **119**: p. 2576-7.
207. He, M., Sheldon, P.J. and Sherman, D.H. 2001. Characterization of a quinone reductase activity for the mitomycin C binding protein (MRD): Functional switching from a drug-activating enzyme to a drug-binding protein. *Proc Natl Acad Sci USA*. **98**(3): p. 926-31.
208. Sheldon, P.J., Mao, Y., He, M. and Sherman, D.H. 1999. Mitomycin resistance in *Streptomyces lavendulae* includes a novel drug-binding-protein-dependent export system. *J Bacteriol*. **181**(8): p. 2507-12.
209. Kumagai, T. and Sugiyama, M. 1998. Protection of mammalian cells from the toxicity of bleomycin by expression of a bleomycin-binding protein gene from *Streptomyces verticillus*. *J Biochem*. **124**(4): p. 835-41.
210. Gennimata, D., Davies, J. and Tsiftsoglou, A.S. 1996. Bleomycin resistance in *Staphylococcus aureus* clinical isolates. *J Antimicrob Chemother*. **37**(1): p. 65-75.
211. Kumagai, T., Hibino, R., Kawano, Y. and Sugiyama, M. 1999. Mutation of the N-terminal proline 9 of BLMA from *Streptomyces verticillus* abolishes the binding affinity for bleomycin. *FEBS Lett*. **450**(3): p. 227-30.
212. Nuttall, S.D., Hattarki, M., Guthrie, R.E., Hudson, P.J. and Kortt, A.A. 1999. Utilization of the *Streptoalloteichus hindustanus* resistance determinant ShBle as a protein framework: effect of mutation upon ShBle dimerization and interaction of C-terminal displayed peptide epitopes. *J Protein Chem*. **18**(8): p. 813-21.
213. Maruyama, M., Kumagai, T., Matoba, Y., Hayashida, M., Fujii, T., Hata, Y. and Sugiyama, M. 2001. Crystal structures of the transposon Tn5-carried bleomycin resistance determinant uncomplexed and complexed with bleomycin. *J Biol Chem*. **276**(13): p. 9992-9.
214. Martin, T.W., Dauter, Z., Devedjiev, Y., Sheffield, P., Jelen, F., He, M., Sherman, D.H., Otlewski, J., Derewenda, Z.S. and Derewenda, U. 2002. Molecular basis of mitomycin C resistance in streptomyces: structure and function of the MRD protein. *Structure*. **10**(7): p. 933-42.
215. Kuhnl, J., Bobik, T., Procter, J.B., Burmeister, C., Hoppner, J., Wilde, I., Luersen, K., Torda, A.E., Walter, R.D. and Liebau, E. 2005. Functional analysis of the methylmalonyl-CoA epimerase from *Caenorhabditis elegans*. *Febs J*. **272**(6): p. 1465-77.
216. Leadlay, P.F. 1981. Purification and characterization of methylmalonyl-CoA epimerase from *Propionibacterium shermanii*. *Biochem J*. **197**(2): p. 413-9.
217. Stabler, S.P., Marcell, P.D. and Allen, R.H. 1985. Isolation and characterization of DL-methylmalonyl-coenzyme A racemase from rat liver. *Arch Biochem Biophys*. **241**(1): p. 252-64.
218. Allen, S.H., Kellermeyer, R., Stjernholm, R., Jacobson, B. and Wood, H.G. 1963. The isolation, purification, and properties of methylmalonyl racemase. *J Biol Chem*. **238**: p. 1637-42.

219. Mazumder, R., Sasakawa, T., Kaziro, Y. and Ochoa, S. 1962. Metabolism of propionic acid in animal tissues. IX. Methylmalonyl coenzyme A racemase. *J Biol Chem*. **237**: p. 3065-8.
220. Thomsen, R. and Christensen, M.H. 2006. MolDock: A New Technique for High-Accuracy Molecular Docking. *J Med Chem*. **49**: p. 3315 - 3321.
221. Halgren, T.A. 1996. Merck Molecular Force Field. II. MMFF94 van der Waals and Electrostatic Parameters for Intermolecular Interactions. *J Comp Chem*. **17**: p. 520-552.
222. Halgren, T.A. 1996. Merck Molecular Force Field. III. Molecular Geometries and Vibrational Frequencies. *J Comp Chem.* **17**: p. 553-586.
223. Halgren, T.A. 1996. Merck Molecular Force Field. V. Extension of MMFF94 Using Experimental Data, Additional Computational Data and Empirical Rules. *J Comp Chem*. **17**: p. 616-641.
224. Halgren, T.A. 1996. Merck Molecular Force Field: I. Basis, Form, Scope, Parameterization and Performance of MMFF94. *J Comp Chem*. **17**: p. 490-519.
225. Halgren, T.A. 1999. MMFF VI. MMFF94s Option for Energy Minimization Studies. *J Comp Chem*. **20**: p. 720-729.
226. Halgren, T.A. 1999. MMFF VII. Characterization of MMFF94, MMFF94s, and Other Widely Available Force Fields for Conformational Energies and for Intermolecular-Interaction Energies and Geometries. *J Comp Chem*. **20**: p. 730-748.
227. Halgren, T.A. and Nachbar, R. B. 1996. Merck Molecular Force Field. IV. Conformational Energies and Geometries for MMFF94. *J Comp Chem*. **17**: p. 587-615.
228. Dayem, L.C., Carney, J.R., Santi, D.V., Pfeifer, B.A., Khosla, C. and Kealey, J.T. 2002. Metabolic engineering of a methylmalonyl-CoA mutase-epimerase pathway for complex polyketide biosynthesis in *Escherichia coli*. *Biochemistry*. **41**(16): p. 5193-201.
229. Shi, L., Gao, P., Yan, X.X. and Liang, D.C. 2009. Crystal structure of a putative methylmalonyl-coenzyme A epimerase from *Thermoanaerobacter tengcongensis* at 2.0 Å resolution. *Proteins*. **77**(4): p. 994-9.
230. Sharkey, E.M., O'Neill, H.B., Kavarana, M.J., Wang, H., Creighton, D.J., Sentz, D.L. and Eiseman, J.L. 2000. Pharmacokinetics and antitumor properties in tumor-bearing mice of an enediol analogue inhibitor of glyoxalase I. *Cancer Chemother Pharmacol*. **46**(2): p. 156-66.
231. Hamilton, D.S., Kavarana, M.J., Sharkey, E.M., Eiseman, J.L. and Creighton, D.J. 1999. A new method for rapidly generating inhibitors of glyoxalase I inside tumor cells using *S*-(*N*-aryl-*N*-hydroxycarbamoyl)ethylsulfoxides. *J Med Chem*. **42**(10): p. 1823-7.
232. Kalsi, A., Kavarana, M.J., Lu, T., Whalen, D.L., Hamilton, D.S. and Creighton, D.J. 2000. Role of hydrophobic interactions in binding *S*-(*N*-aryl/alkyl-*N*-hydroxycarbamoyl)glutathiones to the active site of the antitumor target enzyme glyoxalase I. *J Med Chem*. **43**(21): p. 3981-6.
233. Barnard, J.F. and Honek, J.F. 1989. Investigation on glyoxalase I inhibitors. *Biochem Biophys Res Commun*. **165**(1): p. 118-24.
234. Barnard, J.F., Vander Jagt, D.L. and Honek, J.F. 1994. Small molecule probes of glyoxalase I and glyoxalase II. *Biochim Biophys Acta*. **1208**(1): p. 127-35.
235. Ly, H.D., Clugston, S.L., Sampson, P.B. and Honek, J.F. 1998. Syntheses and kinetic evaluation of hydroxamate-based peptide inhibitors of glyoxalase I. *Bioorg Med Chem Lett*. **8**(7): p. 705-10.

236. Vince, R., Wolf, M. and Sanford, C. 1973. Glutaryl-*S*-(p-bromobenzyl)-L-cysteinylglycine. A metabolically stable inhibitor of glyoxalase I. *J Med Chem*. **16**(8): p. 951-3.
237. More, S.S. and Vince, R. 2006. A metabolically stable tight-binding transition-state inhibitor of glyoxalase-I. *Bioorg Med Chem Lett*. **16**(23): p. 6039-42.
238. More, S.S. and Vince, R. 2007. Design, synthesis, and binding studies of bidentate Zn-chelating peptidic inhibitors of glyoxalase-I. *Bioorg Med Chem Lett*. **17**(13): p. 3793-7.
239. Zheng, Z.B. and Creighton, D.J. 2003. Bivalent transition-state analogue inhibitors of human glyoxalase I. *Org Lett*. **5**(25): p. 4855-8.
240. Lo, T.W. and Thornalley, P.J. 1992. Inhibition of proliferation of human leukaemia 60 cells by diethyl esters of glyoxalase inhibitors *in vitro*. *Biochem Pharmacol*. **44**(12): p. 2357-63.
241. More, S.S. and Vince, R. 2009. Inhibition of glyoxalase I: the first low-nanomolar tight-binding inhibitors. *J Med Chem*. **52**(15): p. 4650-6.
242. Ridderstrom, M., Cameron, A.D., Jones, T.A. and Mannervik, B. 1997. Mutagenesis of residue 157 in the active site of human glyoxalase I. *Biochem J*. **328**(Pt 1): p. 231-5.
243. Davidson, G., Clugston, S.L., Honek, J.F. and Maroney, M.J. 2000. XAS investigation of the nickel active site structure in *Escherichia coli* glyoxalase I. *Inorg Chem*. **39**(14): p. 2962-3.
244. Hesse, H., Nikiforova, V., Gakiere, B. and Hoefgen, R. 2004. Molecular analysis and control of cysteine biosynthesis: integration of nitrogen and sulphur metabolism. *J Exp Bot*. **55**(401): p. 1283-92.
245. Jones, D.P. 2010. Redox sensing: orthogonal control in cell cycle and apoptosis signalling. *J Intern Med*. **268**(5): p. 432-48.
246. Spies, H.S. and Steenkamp, D.J. 1994. Thiols of intracellular pathogens. Identification of ovothiol A in *Leishmania donovani* and structural analysis of a novel thiol from *Mycobacterium bovis*. *Eur J Biochem*. **224**(1): p. 203-13.
247. Jardine, M.A., Spies, H.S., Nkambule, C.M., Gammon, D.W. and Steenkamp, D.J. 2002. Synthesis of mycothiol, 1D-1-*O*-(2-[*N*-acetyl-L-cysteinyl]amino-2-deoxy-alpha-D-glucopyranosyl)-*myo*-inositol, principal low molecular mass thiol in the actinomycetes. *Bioorg Med Chem*. **10**(4): p. 875-81.
248. Lee, S. and Rosazza, J.P. 2004. First total synthesis of mycothiol and mycothiol disulfide. *Org Lett*. **6**(3): p. 365-8.
249. Sakuda, S., Zhou, Z.Y. and Yamada, Y. 1994. Structure of a novel disulfide of 2-(*N*-acetylcysteinyl)amido-2-deoxy-alpha-D-glucopyranosyl-myo-inositol produced by Streptomyces sp. *Biosci Biotechnol Biochem*. **58**(7): p. 1347-8.
250. Unson, M.D., Newton, G.L., Arnold, K.F., Davis, C.E. and Fahey, R.C. 1999. Improved methods for immunoassay of mycothiol. *J Clin Microbiol*. **37**(7): p. 2153-7.
251. Steenkamp, D.J. and Vogt, R.N. 2004. Preparation and utilization of a reagent for the isolation and purification of low-molecular-mass thiols. *Anal Biochem*. **325**(1): p. 21-7.
252. Steenkamp, D.J. and Spies, H.S. 1994. Identification of a major low-molecular-mass thiol of the trypanosomatid *Crithidia fasciculata* as ovothiol A. Facile isolation and structural analysis of the bimane derivative. *Eur J Biochem*. **223**(1): p. 43-50.
253. Unson, M.D., Newton, G.L., Davis, C. and Fahey, R.C. 1998. An immunoassay for the detection and quantitative determination of mycothiol. *J Immunol Methods*. **214**(1-2): p. 29-39.

254. Bornemann, C., Jardine, M.A., Spies, H.S. and Steenkamp, D.J. 1997. Biosynthesis of mycothiol: elucidation of the sequence of steps in *Mycobacterium smegmatis*. *Biochem J*. **325**(Pt 3): p. 623-9.
255. Gammon, D.W., Hunter, R., Steenkamp, D.J. and Mudzunga, T.T. 2003. Synthesis of 2-deoxy-2-*C*-alkylglucosides of *myo*-inositol as possible inhibitors of a *N*-deacetylase enzyme in the biosynthesis of mycothiol. *Bioorg Med Chem Lett*. **13**(12): p. 2045-9.
256. Knapp, S., Gonzalez, S., Myers, D.S., Eckman, L.L. and Bewley, C.A. 2002. Shortcut to mycothiol analogues. *Org Lett*. **4**(24): p. 4337-9.
257. Metaferia, B.B., Ray, S., Smith, J.A. and Bewley, C.A. 2007. Design and synthesis of substrate-mimic inhibitors of mycothiol-*S*-conjugate amidase from *Mycobacterium tuberculosis*. *Bioorg Med Chem Lett*. **17**(2): p. 444-7.
258. Nicholas, G.M., Eckman, L.L., Ray, S., Hughes, R.O., Pfefferkorn, J.A., Barluenga, S., Nicolaou, K.C. and Bewley, C.A. 2002. Bromotyrosine-derived natural and synthetic products as inhibitors of mycothiol-*S*-conjugate amidase. *Bioorg Med Chem Lett*. **12**(17): p. 2487-90.
259. Slattegard, R., Gammon, D.W. and Oscarson, S. 2007. Synthesis of fused bicyclic thioglycosides of *N*-acylated glucosamine as analogues of mycothiol. *Carbohydr Res*. **342**(12-13): p. 1943-6.
260. Stewart, M.J., Jothivasan, V.K., Rowan, A.S., Wagg, J. and Hamilton, C.J. 2008. Mycothiol disulfide reductase: solid phase synthesis and evaluation of alternative substrate analogues. *Org Biomol Chem*. **6**(2): p. 385-90.
261. Ramirez, J., Yu, L., Li, J., Braunschweiger, P. and Wang, P. 1996. Glyco-*S*-nitrosothiols, a novel class of no donor compounds. *Bioorganic & Medicinal Chemistry Letters*. **6**: p. 2575-2580.
262. Kamber, B., Hartmann, A., Eisler, K., Riniker, B., Rink, H., Sieber, P. and Rittel, W. 1980. The Synthesis of Cystine Peptides by Iodine Oxidation of *S*-Trityl-cysteine and *S*-Acetamidomethyl-cysteine Peptides. *Helvetica Chimica Acta*. **63**: p. 899-915.
263. Harris, K.M., Flemer Jr., S. and Hondal, R.J. 2007. Studies on deprotection of cysteine and selenocysteine side-chain protecting groups. *J Pept Sci*. **13**(2): p. 81-93.
264. Photaki, I., Taylor-Papadimitriou, J., Sakarellos, C., Mazarakis, P. and Zervas, L. 1970. On cysteine and cystine peptides. V. *S*-trityl and *S*-diphenylmethyl-cysteine and -cysteine peptides. *J Chem Soc Perkin 1*. **19**: p. 2683-7.
265. Ellman, G.L. 1959. Tissue sulfhydryl groups. *Arch Biochem Biophys*. **82**(1): p. 70-7.
266. Fujii, N., Otaka, A., Funakoshi, S., Bessho, K. and Yajima, H. 1987. New procedure for the synthesis of cystine-peptides by oxidation of *S*-substituted cysteine-peptides with thallium(III) trifluoroacetate. *J Chem Soc Chem Commun*. **3**: p. 163-164.
267. Nishimura, O., Kitada, C. and Fujino, M. 1978. A new method for removing the *S*-*p* methoxybenzyl and *S*-*t*-butyl groups of cysteme residues with mercuric trifluoroacetate *Chem. Pharm. Bull*. **26**: p. 1576-1585.
268. Hunter, M.J. and Komives, E.A. 1995. Deprotection of *S*-acetamidomethyl cysteine-containing peptides by silver trifluoromethanesulfonate avoids the oxidation of methionines. *Anal Biochem*. **228**(1): p. 173-7.
269. Koul, A., Arnoult, E., Lounis, N., Guillemont, J. and Andries, K. 2011. The challenge of new drug discovery for tuberculosis. *Nature*. **469**(7331): p. 483-90.
270. Ginsberg, A.M. 2010. Tuberculosis drug development: progress, challenges, and the road ahead. *Tuberculosis (Edinb)*. **90**(3): p. 162-7.

271. Prasad, R. 2010. Multidrug and extensively drug-resistant TB (M/XDR-TB): problems and solutions. *Indian J Tuberc*. **57**(4): p. 180-91.
272. Mikhailovich, V.M., Lapa, S.A., Gryadunov, D.A., Strizhkov, B.N., Sobolev, A.Y., Skotnikova, O.I., Irtuganova, O.A., Moroz, A.M., Litvinov, V.I., Shipina, L.K., Vladimirskii, M.A., Chernousova, L.N., Erokhin, V.V. and Mirzabekov, A.D. 2001. Detection of rifampicin-resistant *Mycobacterium tuberculosis* strains by hybridization and polymerase chain reaction on a specialized TB-microchip. *Bull Exp Biol Med*. **131**(1): p. 94-8.
273. Chugh, S. 1020. Drug resistance in tuberculosis and options available. *J Indian Med Assoc*. **108**(2): p. 106-8.
274. Chim, N., Habel, J.E., Johnston, J.M., Krieger, I., Miallau, L., Sankaranarayanan, R., Morse, R.P., Bruning, J., Swanson, S., Kim, H., Kim, C.Y., Li, H., Bulloch, E.M., Payne, R.J., Manos-Turvey, A., Hung, L.W., Baker, E.N., Lott, J.S., James, M.N., Terwilliger, T.C., Eisenberg, D.S., Sacchettini, J.C. and Goulding, C.W. 2010. The TB Structural Genomics Consortium: A decade of progress. *Tuberculosis (Edinb)*. **91**(2): p. 155-72.
275. Park, J.H., Cha, C.J. and Roe, J.H. 2006. Identification of genes for mycothiol biosynthesis in *Streptomyces coelicolor* A3(2). *J Microbiol*. **44**(1): p. 121-5.
276. Bentley, S.D., Chater, K.F., Cerdeno-Tarraga, 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., Rabbinowitsch, 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. and Hopwood, D.A. 2002. Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). *Nature*. **417**(6885): p. 141-7.
277. Ceylan, O., Okmen, G. and Ugur, A. 2008. Isolation of soil Streptomyces as source antibiotics active against antibiotic-resistant bacteria. *EurAsia J BioSci*. **2**: p. 73-82.
278. Rawat, M., Newton, G.L., Ko, M., Martinez, G.J., Fahey, R.C. and Av-Gay, Y. 2002. Mycothiol-deficient *Mycobacterium smegmatis* mutants are hypersensitive to alkylating agents, free radicals, and antibiotics. *Antimicrob Agents Chemother*. **46**(11): p. 3348-55.
279. Creighton, D.J., Migliorini, M., Pourmotabbed, T. and Guha, M.K. 1988. Optimization of efficiency in the glyoxalase pathway. *Biochemistry*. **27**(19): p. 7376-84.
280. Griffis, C.E., Ong, L.H., Buettner, L. and Creighton, D.J. 1983. Nonstereospecific substrate usage by glyoxalase I. *Biochemistry*. **22**(12): p. 2945-51.
281. Vander Jagt, D.L., Daub, E., Krohn, J.A. and Han, L.P. 1975. Effects of pH and thiols on the kinetics of yeast glyoxalase I. An evaluation of the random pathway mechanism. *Biochemistry*. **14**(16): p. 3669-75.
282. Vivian, J.T. and Callis, P.R. 2001. Mechanisms of tryptophan fluorescence shifts in proteins. *Biophys J*. **80**(5): p. 2093-109.
283. Dobson, C.M., Evans, P.A. and Williamson, K.L. 1984. Proton NMR studies of denatured lysozyme. *FEBS Lett*. **168**(2): p. 331-4.

284. Rae, C., O'Donoghue, S.I., Bubb, W.A. and Kuchel, P.W. 1994. Stereospecificity of substrate usage by glyoxalase 1: nuclear magnetic resonance studies of kinetics and hemithioacetal substrate conformation. *Biochemistry*. **33**(12): p. 3548-59.
285. Nemet, I., Vikic-Topic, D. and Varga-Defterdarovic, L. 2004. Spectroscopic studies of methylglyoxal in water and dimethylsulfoxide. *Bioorg Chem*. **32**(6): p. 560-70.
286. Fodor, G., Mujumdar, R. and Szent-Gyorgyi, A. 1978. Isolation of methylglyoxal from liver. *Proc Natl Acad Sci USA*. **75**(9): p. 4317-9.
287. Volk, W.A., Benjamin, D.C., Kadner, R.J. and Parseus, J.T. 2005. Essentials of medical microbiology (fourth edition), J.B. Lippincott, Philadelphia, Pennsylvania.
288. Worrall, J.A., Machczynski, M.C., Keijser, B.J., di Rocco, G., Ceola, S., Ubbink, M., Vijgenboom, E. and Canters, G.W. 2006. Spectroscopic characterization of a high-potential lipo-cupredoxin found in *Streptomyces coelicolor*. *J Am Chem Soc*. **128**(45): p. 14579-89.
289. Aronsson, A.C. and Mannervik, B. 1977. Characterization of glyoxalase I purified from pig erythrocytes by affinity chromatography. *Biochem J*. **165**(3): p. 503-9.
290. Han, L.P., Schimandle, C.M., Davison, L.M. and Vander Jagt, D.L. 1977. Comparative kinetics of Mg^{2+}-, Mn^{2+}-, Co^{2+}-, and Ni^{2+}-activated glyoxalase I. Evaluation of the role of the metal ion. *Biochemistry*. **16**(25): p. 5478-84.
291. Mannervik, B., Lindstrom, L. and Bartfai, T. 1972. Partial purification and characterization of glyoxalase I from porcine erythrocytes. *Eur J Biochem*. **29**(2): p. 276-81.
292. Uotila, L. and Koivusalo, M. 1975. Purification and properties of glyoxalase I from sheep liver. *Eur J Biochem*. **52**(3): p. 493-503.
293. Aronsson, A.C., Marmstal, E. and Mannervik, B. 1978. Glyoxalase I, a zinc metalloenzyme of mammals and yeast. *Biochem Biophys Res Commun*. **81**(4): p. 1235-40.
294. Deswal, R., Chakaravarty, T.N. and Sopory, S.K. 1993. The glyoxalase system in higher plants: regulation in growth and differentiation. *Biochem Soc Trans*. **21**(2): p. 527-30.
295. Johansen, K.S., Svendsen, I.I. and Rasmussen, S.K. 2000. Purification and cloning of the two domain glyoxalase I from wheat bran. *Plant Sci*. **155**(1): p. 11-20.
296. Dragani, B., Cocco, R., Ridderstrom, M., Stenberg, G., Mannervik, B. and Aceto, A. 1999. Unfolding and refolding of human glyoxalase II and its single-tryptophan mutants. *J Mol Biol*. **291**(2): p. 481-90.

Appendix 1: Flow Chart for DNA Cloning

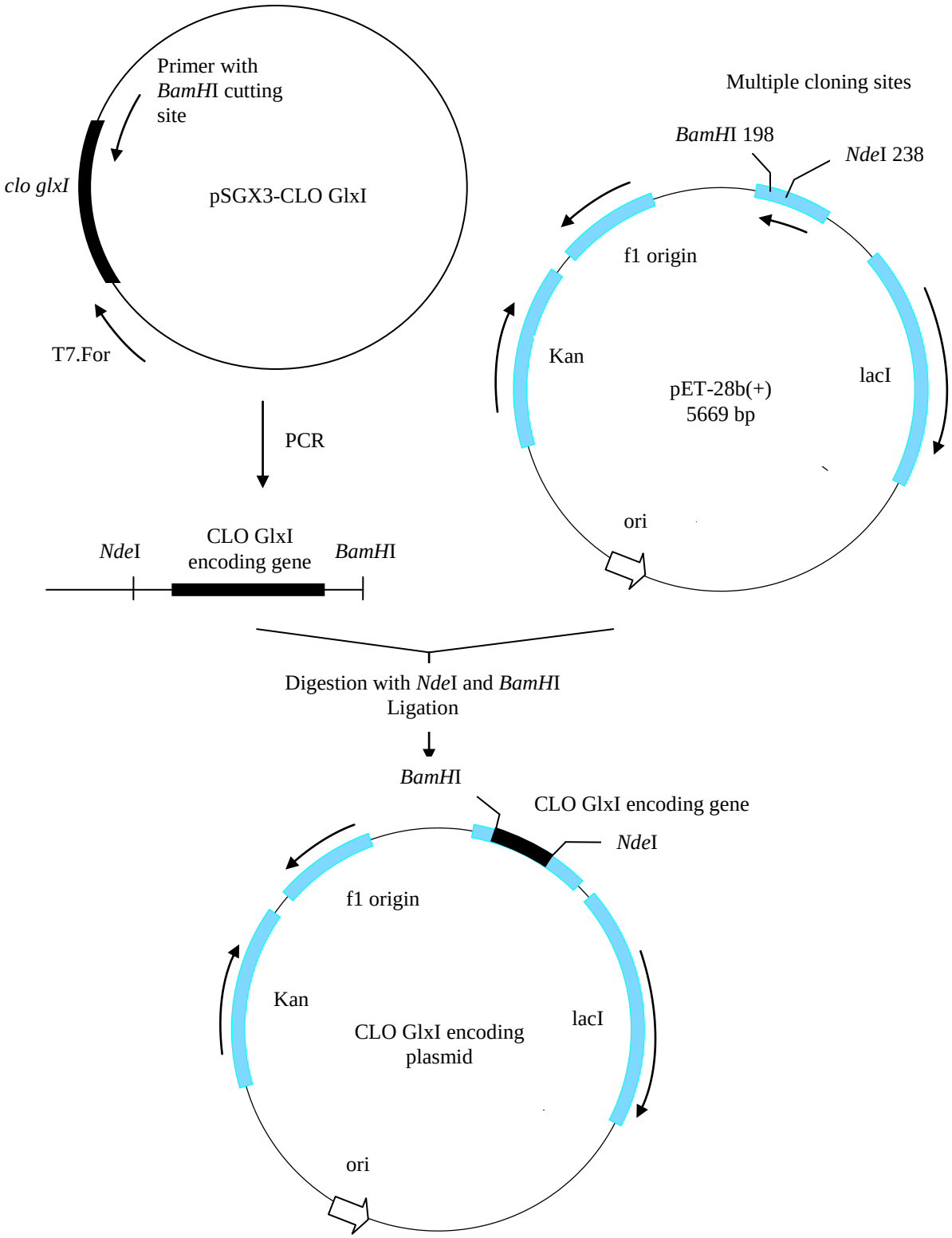


Figure A1.1: DNA cloning of *Clostridium acetobutylicum* GlxI (CLO GlxI) into the pET-28b(+) expression vector using PCR and restriction enzymes (*Nde*I and *BamH*I). The produced protein from this clone possesses the N-terminal His_6-tag followed by a thrombin protease cleavage site and the protein of interest.

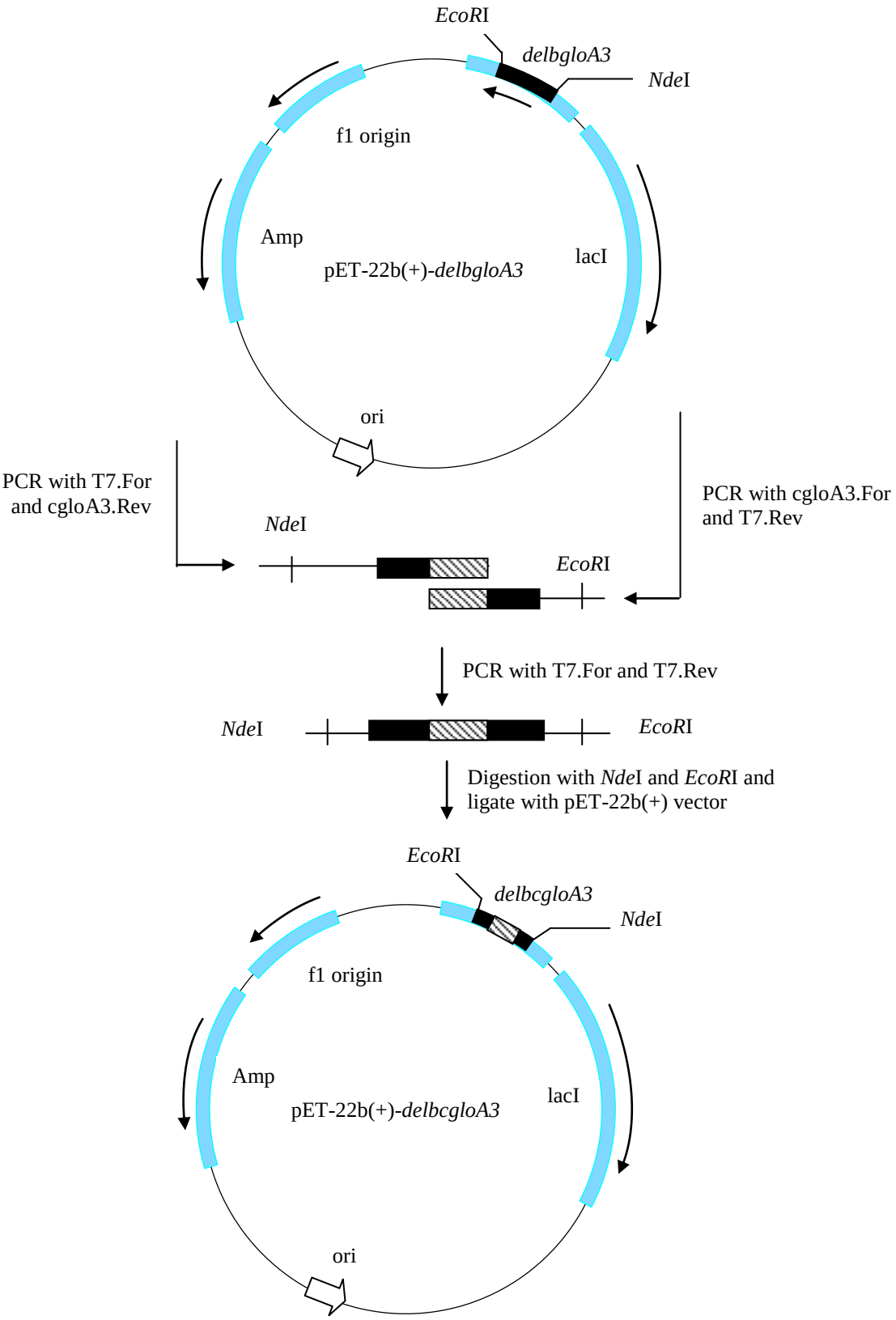


Figure A1.2: DNA cloning of *Pseudomonas aeruginosa delbcgloA3* gene into a pET-22b(+) expression vector using *delbgloA3* plasmid as a template and the utilization of two-stage PCR and restriction enzymes (*Nde*I and *Eco*RI).

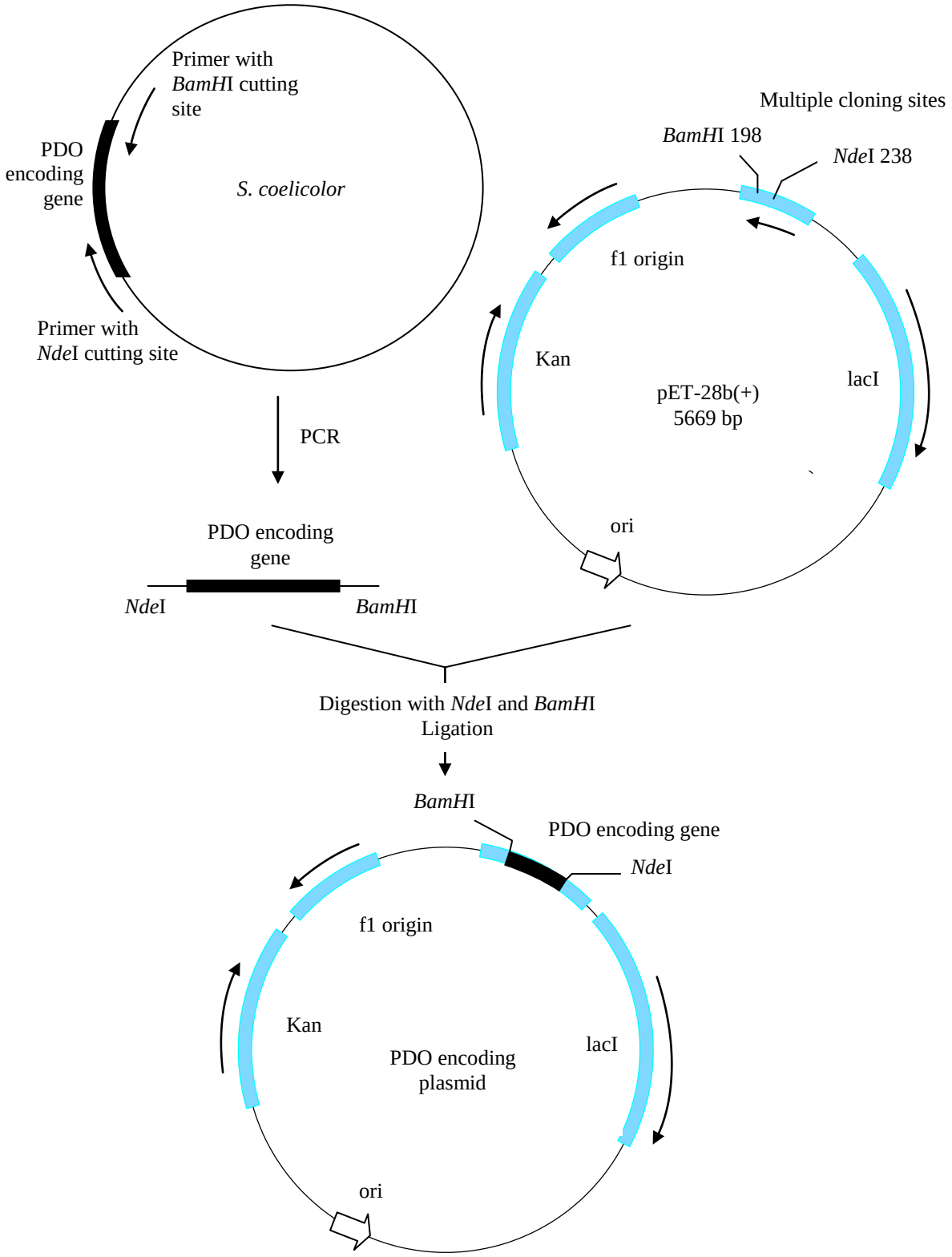


Figure A1.3: DNA cloning of the putative dioxygenase (PDO) from *Streptomyces coelicolor* into the pET-28b(+) expression vector using PCR and restriction enzymes (*Nde*I and *BamH*I). The produced protein from this clone possesses the N-terminal His_6-tag followed by a thrombin protease cleavage site and the protein of interest.

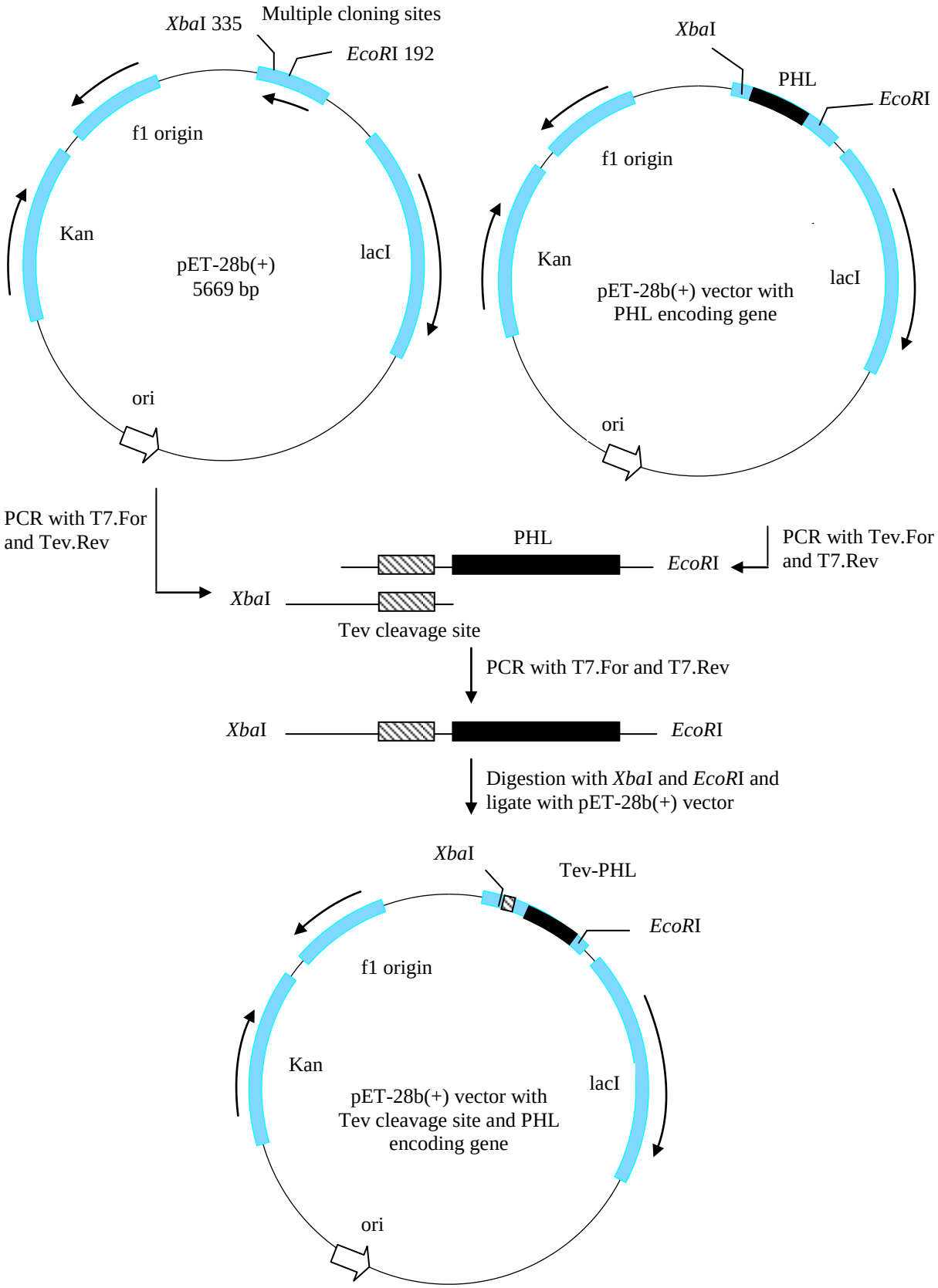


Figure A1.4: DNA cloning of the putative hydrolase (PHL) from *Streptomyces coelicolor* into the pET-28b(+) expression vector with Tev protease recognition site. The produced protein from this clone possesses the N-terminal His_6-tag followed by Tev protease cleavage site and the protein of interest.

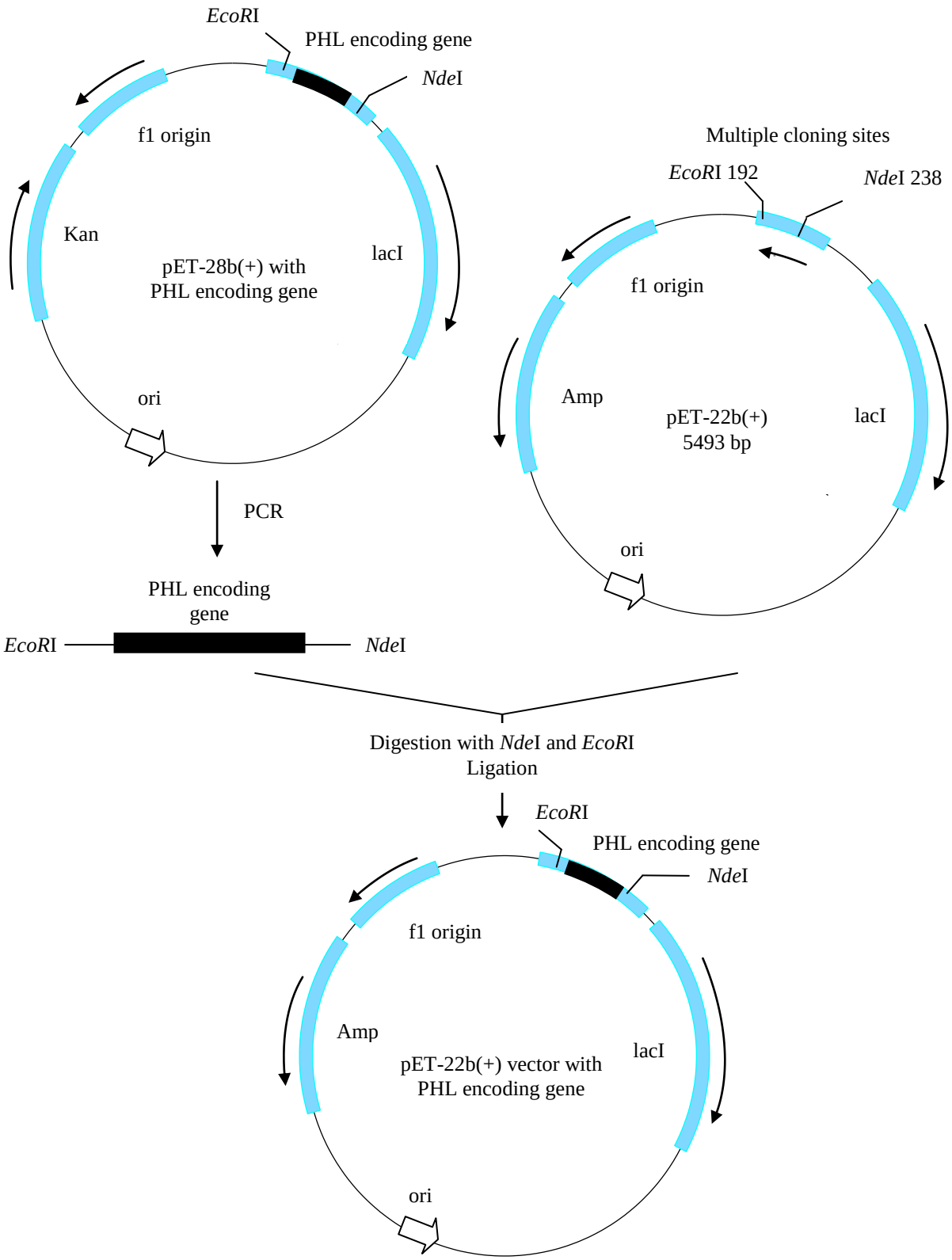


Figure A1.5: DNA cloning of putative hydrolase (PHL) from *Streptomyces coelicolor* into the pET-22b(+) expression vector using PCR and restriction enzymes (*Nde*I and *EcoR*I). The produced protein from this clone does not possess any His-tag.

Appendix 2: Flow Chart for Protein Purification

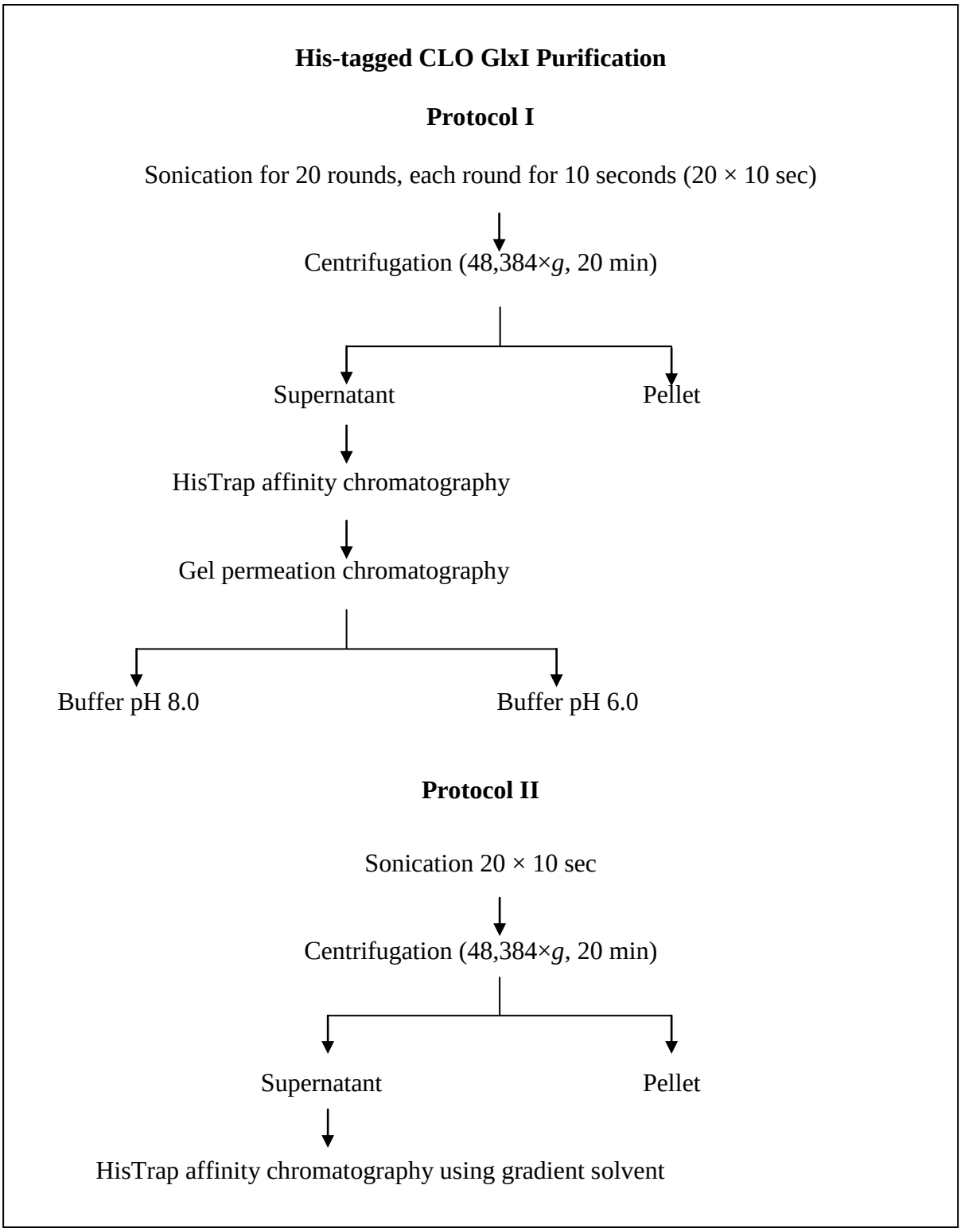


Figure A2.1: The flow chart of protein purification for His-tagged *C. acetobutylicum* GlxI (CLO GlxI). The first protocol includes two steps, HisTrap affinity chromatography and gel permeation chromatography using a

Superdex75 HR 10/30 column with a flow rate of 0.5 mL/min. The second protocol is developed from the first one in an attempt to improve purification efficiency, in which the protocol contains only one step which is the HisTrap affinity chromatography using gradient solvent.

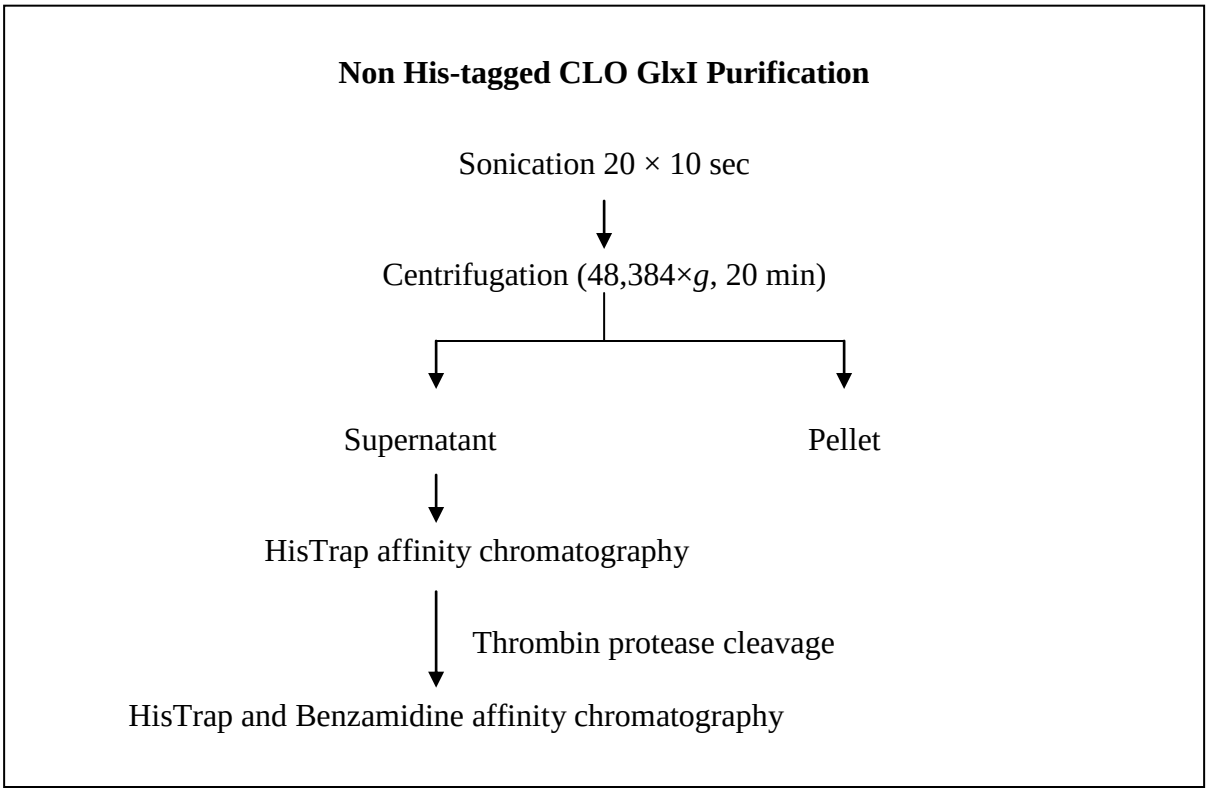


Figure A2.2: The flow chart showing the protein purification of non His-tagged *C. acetobutylicum* GlxI (CLO GlxI). The first step is to purify the His-tagged CLO GlxI by utilizing a HisTrap HP affinity column followed by a cleavage by thrombin protease. The non His-tag CLO GlxI is then purified by using a combined HisTrap HP and Benzamidine FF affinity column approach.

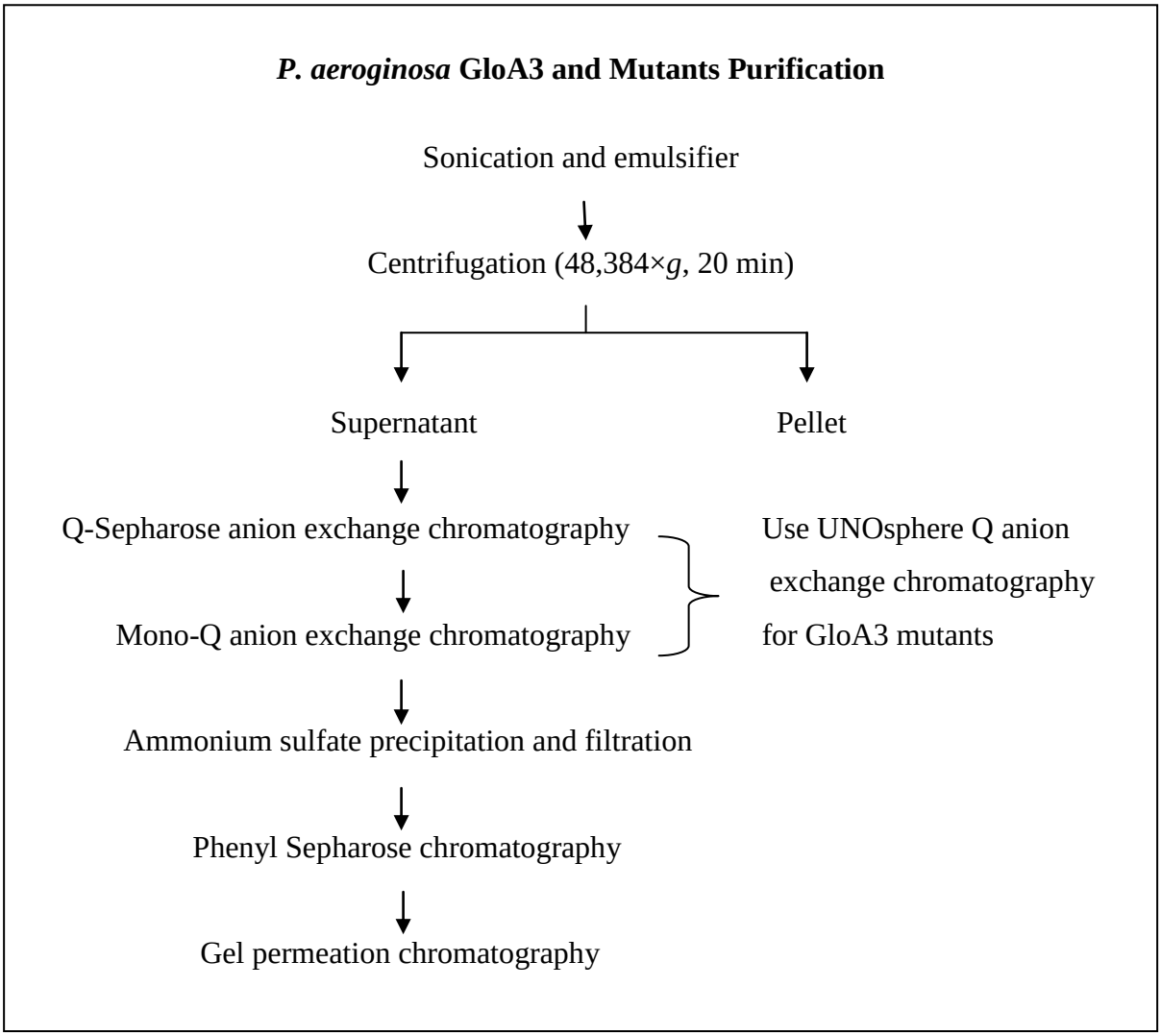


Figure A2.3: The flow chart showing the protein purification of *P. aeroginosa* GloA3 and its mutants, which includes 5 step-purifications utilizing two anion exchange columns (Q-Sepaharose Fast Flow 10/30 column and Mono-Q HR 5/5 column), ammonium sulfate precipitation, phenyl Sepharose chromatography (phenyl HP column, 5 mL) and gel permeation chromatography (Sperdex75 HR 10/30 column). The Q-Sepharose and Mono-Q columns were replaced by use of a UNOsphere Q anion exchange cartridge (1 mL) for GloA3 mutants for higher efficiency of purification.

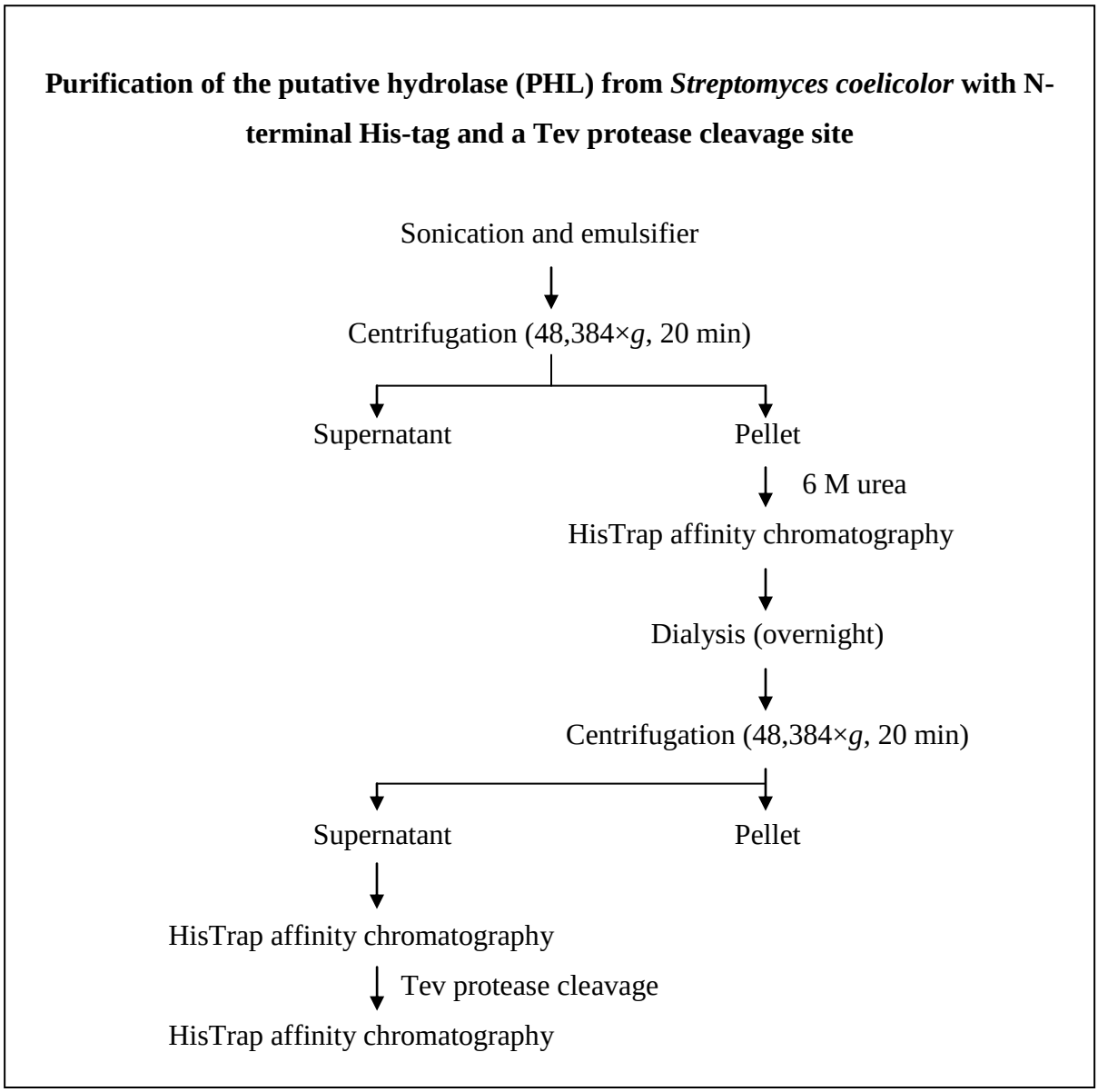


Figure A2.4: The flow chart showing the protein purification under denaturing followed by refolding conditions of the putative hydrolase (PHL) from *Streptomyces coelicolor* with the N-terminal His-tag and Tev protease cleavage site. The first step involved solubilization of the inclusion bodies in buffer containing 6 M urea, following by separation on the HisTrap HP affinity column under denaturing conditions. The denaturant is removed by dialysis overnight at 4 °C. The proteins were then loaded onto the 2nd HisTrap affinity column under native conditions. The His-tagged proteins were cleaved by Tev protease incubation and the desired proteins were recovered from the purification using the 3rd HisTrap HP affinity column.

Appendix 3: Standard Curve for Gel Permeation Chromatography

Gel Permeation Chromatography using Superdex75 HR 10/30 Chromatography

The protein standard plot for gel permeation chromatography using the Sigma standard proteins including blue dextran (2000 kDa), bovine serum albumin or BSA (66 kDa), carbonic anhydrase (29 kDa), cytochrome *c* (12.4 kDa) and aprotinin (1.35 kDa)) was performed on a HPLC system using a Superdex75 HR 10/30 column with a buffer under specific conditions and a flow rate of 0.5 mL/min (Figure A3.1). The protein standard plot presents a relationship between molecular weight of a particular protein and its retention time determined from gel permeation chromatography results. The plot of K_a versus log molecular weight (MW) is used to determine and estimate the molecular weight of an unknown protein (Figure A3.2 and Table A3.1). K_a represents the eluted volumes for each protein and can be calculated as $K_a = (V_e – V_0)/(V_t-V_0)$, where V_0 is the void volume, V_e is the elution volume and V_t is the total bed volume. The void volume could be obtained from the elution volume of blue dextran, which is approximately 8.07 mL and the total bed volume, which is the volume of the column resin, is approx. 24 mL. The *Clostridium acetobutylicum* GlxI (CLO GlxI) in 50 mM Tris (pH 8.0), 200 mM KCl and 10% (v/v) glycerol had an elution volume of 11.4 mL, resulting in a K_a of 0.232 and a molecular mass of 27.5 kDa as estimated from the standard curve.

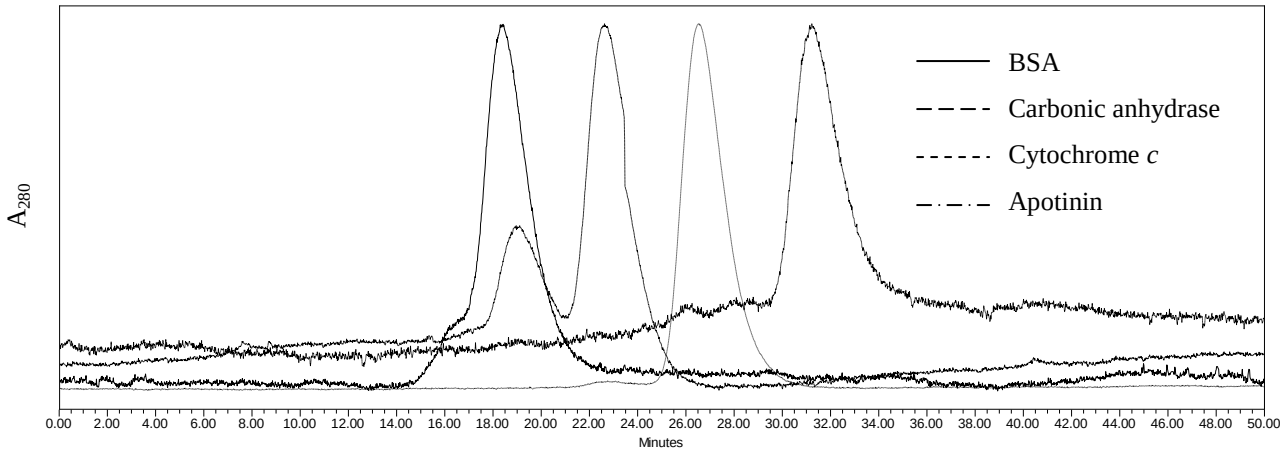


Figure A3.1: The gel permeation chromatogram of the Sigma protein standards including BSA (66 kDa), carbonic anhydrase (29 kDa), cytochrome *c* (12.4 kDa) and aprotinin (1.35 kDa) using a Superdex75 HR 10/30 column and a buffer containing 50 mM Tris (pH 8.0), 200 mM KCl and 10% (v/v) glycerol with a flow rate of 0.5 mL/min. Each protein is eluted from the column with a particular retention time according to its size, which can be used to plot a protein standard curve as shown in Figure A3.2.

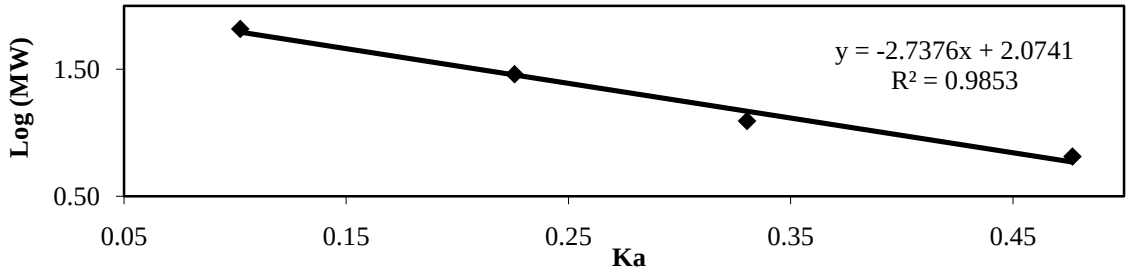


Figure A3.2: The Sigma protein standard curve from gel permeation chromatography using a Superdex75 HR 10/30 column for molecular weight calculation of protein samples. The data for this plot are shown in Table A3.1.

Table A3.1: The table shows the molecular weights and eluted volumes of the standard proteins using a Superdex75 HR 10/30 column, which were then used to plot the standard protein curve and to estimate the molecular weight of the protein sample (in this case, CLO GlxI).

Standard Protein	***MW (kDa)***	***log (MW)***	***V_e-1 (mL)***	***V_e-2 (mL)***	***V_e-3 (mL)***	***Average V_e (mL)***	***K_a***
Blue dextran (V_o)			8.10	8.00	8.10	8.07	
BSA	66.00	1.819544	9.70	9.70	9.70	9.70	0.10232
Carbonic anhydrase	29.00	1.462398	11.70	11.70	11.60	11.67	0.22578
Cytochrome *c*	12.40	1.093422	13.30	13.30	13.40	13.33	0.33040
Aprotinin	6.50	0.812913	15.70	15.70	15.60	15.67	0.47688
CLO GlxI	27.47	1.438821	11.70	11.80	11.80	11.77	0.23206

Similar experiments were performed on a different set of protein standards obtained from Bio-Rad laboratories (Hercules, CA) and included bovine serum albumin (66 kDa), ovalbumin (44 kDa), carbonic anhydrase (29 kDa), myoglobulin (17 kDa) and vitamin B12 (1.35 kDa). The protein standard plot of the gel permeation chromatography using a HPLC system with the Superdex75 HR 10/30 column and having a flow rate of 0.5 mL/min is shown in Figure A3.3. An example calculation for the molecular mass of the mutein from *Pseudomonas aeroginosa*, delbGloA3, in 20 mM Tris (pH 7.5) and 150 mM NaCl is estimated from the protein standard plot as shown in Table A3.2.

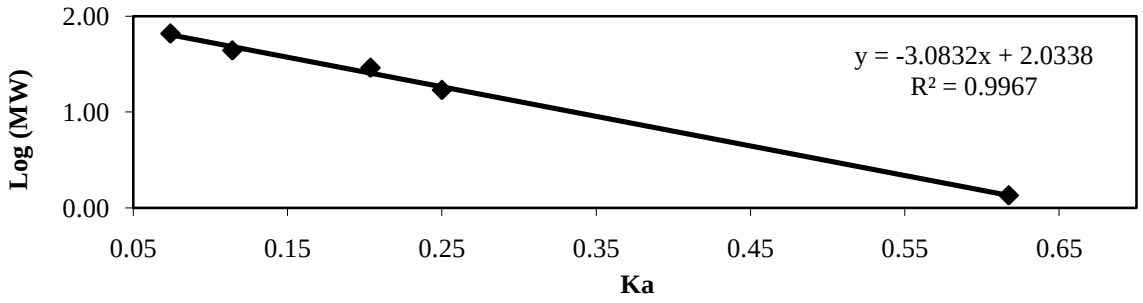


Figure A3.3: A Bio-Rad protein standard curve from gel permeation chromatography results using a Superdex75 HR 10/30 column for molecular weight calculation of protein samples. The data for this plot are shown in Table A3.2.

Table A3.2: The table shows the molecular weights and eluted volumes of the Bio-Rad standard proteins using a Superdex75 HR 10/30 column, which were used to plot the standard protein curve and to estimate the molecular weight of the protein sample (in this case, delbGloA3).

Standard Protein	*MW (kDa)*	*log (MW)*	*V_e-1 (mL)*	*V_e-2 (mL)*	*Average V_e (mL)*	*K_a*
Blue dextran (V_o)			7.8	7.8	7.80	
BSA	66.00	1.82	9	9	9.00	0.07407
Ovalbumin	44.00	1.64	9.6	9.7	9.65	0.11420
Carbonic anhydrase	29.00	1.46	11.1	11.1	11.10	0.20370
Myoglobin	17.00	1.23	11.8	11.9	11.85	0.25000
Vitamin B12	1.35	0.13	17.8	17.8	17.80	0.61728
delbGloA3	37.77	1.58	10.2	10.2	10.20	0.14815

Gel Permeation Chromatography using a Superose6 10/300 GL Column

The molecular weight calculation by the gel permeation chromatography step using a Superose6 10/300 GL column and a buffer under the specified conditions with a flow rate of 0.5 mL/min is similar to those results obtained by using the Superdex75 HR 10/30 column. Due to larger bead size and capacity to separate globular proteins with molecular weight between 5000–5×10^6 Da, the Bio-Rad protein standards including γ-globulin (158 kDa), BSA (66 kDa), ovalbumin (44 kDa), carbonic anhydrase (29 kDa), myoglobulin (17 kDa), and vitamin B_{12} (1.35 kDa) (Figure A3.4) were used to plot the protein standard curve. The void volume from blue detran of 8.1 mL and the total bed volume of 24 mL were determined from this gel permeation chromatography procedure. An example calculation for the putative dioxygenase (PDO) from *Streptomyces coelicolor* in 50 mM Tris (pH 8.0) and 150 mM KCl with the elution volume of 18.35 mL suggests a molecular mass of approx. 31 kDa (Figure A3.5 and Table A3.3).

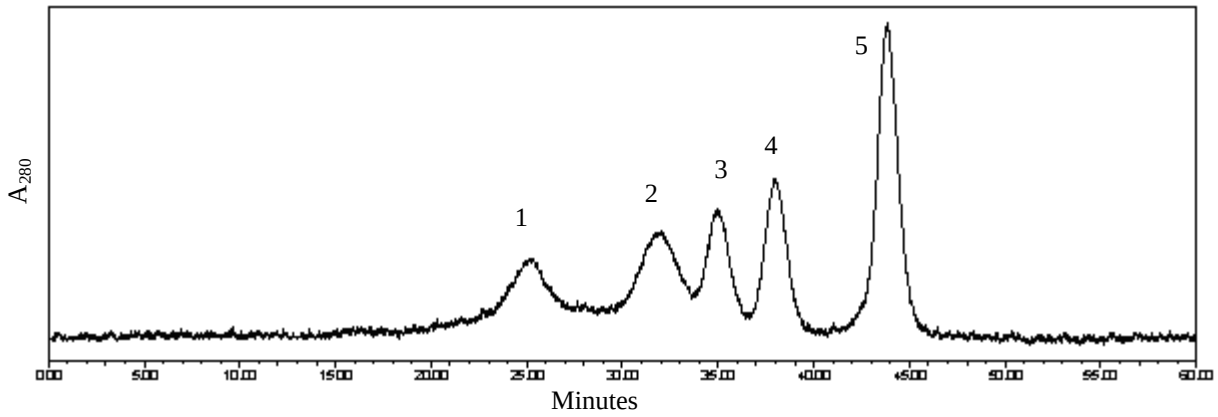


Figure A3.4: The gel permeation chromatogram of the Bio-Rad standard proteins including γ-globulin (158 kDa), BSA (66 kDa), ovalbumin (44 kDa), carbonic anhydrase (29 kDa), myoglobulin (17 kDa), and vitamin B_{12} (1.35 kDa) using a Superose6 10/300 GL column and a buffer containing 50 mM Tris (pH 8.0) and 150 mM KCl with a flow rate of 0.5 mL/min. Each peak represented a particular protein as stated in Table A3.3 and plotted in Figure A3.5.

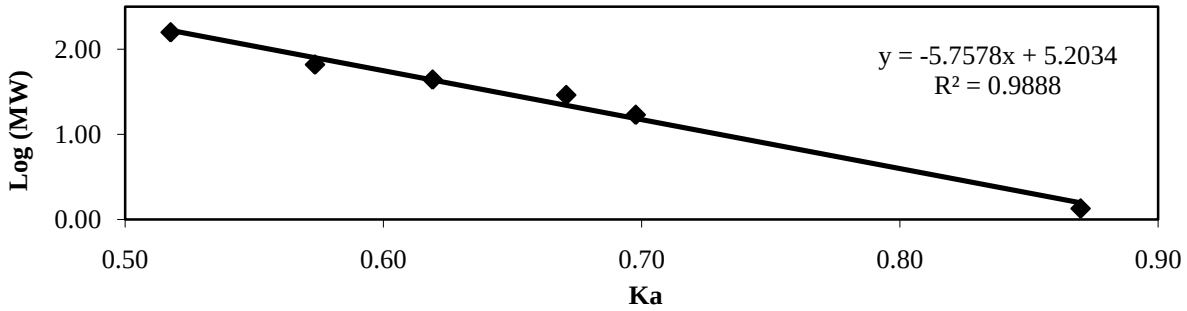


Figure A3.5: The Bio-Rad protein standard curve from gel permeation chromatography using a Superose6 10/300 GL column for molecular weight calculation of protein samples. The data for this plot are shown in Table A3.3.

Table A3.3: The table shows the molecular weights and eluted volumes of the Bio-Rad standard proteins using a Superose6 10/300 GL column, which were used to plot the standard protein curve and to estimate the molecular weight of the protein sample (in this case, putative dioxygenase (PDO) from *Streptomyces coelicolor*).

Standard Protein	***MW (kDa)***	***log (MW)***	***V_e-1 (mL)***	***V_e-2 (mL)***	***V_e-3 (mL)***	***Ave. V_e (mL)***	***K_a***
Blue Dextran (V_o)			8.10	8.10	8.10	8.10	
γ-globulin (1)	158.00	2.198657	16.30	16.20	16.20	16.23	0.51760
BSA (2)	66.00	1.819544	17.10	17.20	17.10	17.13	0.57350
Ovalbumin (3)	44.00	1.643453	17.80	17.90	17.90	17.87	0.61905
Carbonic anhydrase (4)	29.00	1.462398	18.70	18.70	18.70	18.70	0.67081
Myoglobin (5)	17.00	1.230449	19.10	19.20	19.10	19.13	0.69772
Vitamin B12 (6)	1.35	0.130334	21.90	22.00	21.90	21.93	0.87002
PDO	31.02	1.491611	18.35				0.64465

Appendix 4: Metal Analysis by Inductively Coupled Plasma Mass Spectrometry (ICP-MS)

Table A4.1: The table shows metal contents of the purified *Clostricium acetobutylicum* GlxI (CLO GlxI, 2.2 μM) in Chelex-treated water. The experiment was performed in duplicated using Chelex-treated water as reference. The metals that have higher level of detection than those of the reference are highlight in grey.

Metal analysis	*Detection limit (mg/L)*	*Metal detection (mg/L) Trial 1*	*Metal detection (mg/L) Trial 2*	*Metal detection in Chelex-treated water (mg/L)*
Aluminum (Al)	0.01	<0.01	<0.01	<0.01
Antimony (Sb)	0.005	<0.005	<0.005	<0.005
Arsenic (As)	0.001	<0.001	<0.001	<0.001
Barium (Ba)	0.01	<0.01	<0.01	<0.01
Berylium (Be)	0.001	<0.001	<0.001	<0.001
Bismuth (Bi)	0.001	<0.001	<0.001	<0.001
Boron (B)	0.05	<0.05	<0.05	<0.05
Cadmium (Cd)	0.0001	<0.0001	<0.0001	<0.0001
Calcium (Ca)	0.5	<0.5	<0.5	<0.5
Chromium (Cr)	0.001	<0.001	<0.001	<0.001
Cobalt (Co)	0.0005	<0.0005	<0.0005	<0.0005
Copper (Cu)	0.001	<0.001	0.001	<0.001
Iron (Fe)	0.05	<0.05	<0.05	<0.05
Lead (Pb)	0.001	0.002	0.002	<0.001
Magnesium (Mg)	0.5	<0.5	<0.5	<0.5
Manganese (Mn)	0.001	0.001	0.001	<0.001
Molybdenum (Mo)	0.001	<0.001	<0.001	<0.001
Nickel (Ni)	0.002	<0.002	<0.002	<0.002
Phosphorus (P)	0.05	<0.05	<0.05	<0.05
Potassium (K)	1	<1	<1	<1
Selenium (Se)	0.005	<0.005	<0.005	<0.005
Silicon (Si)	0.1	<0.1	<0.1	<0.1
Silver (Ag)	0.0001	<0.0001	<0.0001	<0.0001
Sodium (Na)	0.5	2.8	2.6	2.3
Strontium (Sr)	0.001	<0.001	<0.001	<0.001
Thallium (Tl)	0.0003	<0.0003	<0.0003	<0.0003
Tin (Sn)	0.001	<0.001	<0.001	<0.001
Titanium (Ti)	0.002	<0.002	<0.002	<0.002
Tungsten (W)	0.01	<0.01	<0.01	<0.01
Uranium (U)	0.005	<0.005	<0.005	<0.005
Vanadium (V)	0.001	<0.001	<0.001	<0.001
Zinc (Zn)	0.003	0.006	0.006	<0.003
Zirconium (Zr)	0.004	<0.004	<0.004	<0.004

Table A4.2: The table shows metal contents of Ni^{2+}-reconstituted *Clostridium acetobutylicum* GlxI (CLO GlxI, 1.42 μM) in Chelex-treated buffer containing 50 mM HEPES (pH 7.0), 200 mM KCl, 10 mM Met and 10% glycerol). The experiment was performed in duplicated using the Chelex-treated buffer as reference. The metals that have higher level of detection than those of the reference are highlight in grey.

Metal analysis	*Detection limit (mg/L)*	*Metal detection (mg/L) Trial 1*	*Metal detection (mg/L) Trial 2*	*Metal detection in Chelex-treated buffer (mg/L)*
Aluminum (Al)	0.01	<0.01	<0.01	<0.01
Antimony (Sb)	0.005	<0.005	<0.005	<0.005
Arsenic (As)	0.001	<0.001	<0.001	<0.001
Barium (Ba)	0.01	<0.01	<0.01	<0.01
Berylium (Be)	0.001	<0.001	<0.001	<0.001
Bismuth (Bi)	0.001	<0.001	<0.001	<0.001
Boron (B)	0.05	<0.05	<0.05	<0.05
Cadmium (Cd)	0.0001	<0.0001	<0.0001	<0.0001
Calcium (Ca)	0.5	<0.5	<0.5	<0.5
Chromium (Cr)	0.001	0.002	0.002	0.002
Cobalt (Co)	0.0005	<0.0005	<0.0005	<0.0005
Copper (Cu)	0.001	<0.001	<0.001	<0.001
Iron (Fe)	0.05	<0.05	<0.05	<0.05
Lead (Pb)	0.001	<0.001	<0.001	<0.001
Magnesium (Mg)	0.5	<0.5	<0.5	<0.5
Manganese (Mn)	0.001	0.002	0.002	<0.001
Molybdenum (Mo)	0.001	<0.001	<0.001	<0.001
Nickel (Ni)	0.002	0.179	0.211	<0.002
Phosphorus (P)	0.05	<0.05	<0.05	<0.05
Potassium (K)	1	60	50	60
Selenium (Se)	0.005	<0.005	<0.005	<0.005
Silicon (Si)	0.1	0.5	0.5	0.5
Silver (Ag)	0.0001	<0.0001	<0.0001	<0.0001
Sodium (Na)	0.5	5.8	6.0	5.8
Strontium (Sr)	0.001	<0.001	<0.001	<0.001
Thallium (Tl)	0.0003	<0.0003	<0.0003	<0.0003
Tin (Sn)	0.001	<0.001	0.001	<0.001
Titanium (Ti)	0.002	<0.002	<0.002	<0.002
Tungsten (W)	0.01	<0.01	<0.01	<0.01
Uranium (U)	0.005	<0.005	<0.005	<0.005
Vanadium (V)	0.001	<0.001	<0.001	<0.001
Zinc (Zn)	0.003	0.003	0.004	<0.003
Zirconium (Zr)	0.004	<0.004	<0.004	<0.004

Table A4.3: The table shows metal contents of *Pseudomonas aeroginosa* GloA3 in Chelex-treated buffer containing 20 mM Tris (pH 7.5) and 150 mM NaCl. The protein was grown and expressed in the presence of $ZnCl_2$ (2 mM) in *E. coli* system. The experiment was performed in duplicated using the Chelex-treated buffer as reference. The metals that have higher level of detection than those of the reference are highlight in grey.

Metal analysis	*Detection limit (mg/L)*	*Metal detection (mg/L) Trial 1*	*Metal detection (mg/L) Trial 2*	*Metal detection in Chelex-treated buffer (mg/L)*
Aluminum (Al)	0.1	<0.1	<0.1	<0.1
Antimony (Sb)	0.05	<0.05	<0.05	<0.05
Arsenic (As)	0.01	<0.01	<0.01	<0.01
Barium (Ba)	0.1	<0.1	<0.1	<0.1
Berylium (Be)	0.01	<0.01	<0.01	<0.01
Bismuth (Bi)	0.01	<0.01	<0.01	<0.01
Boron (B)	0.5	<0.5	<0.5	<0.5
Cadmium (Cd)	0.001	<0.001	<0.001	<0.001
Calcium (Ca)	5	<5	<5	<5
Chromium (Cr)	0.01	0.015	0.014	0.015
Cobalt (Co)	0.005	<0.005	<0.005	<0.005
Copper (Cu)	0.01	<0.01	<0.01	<0.01
Iron (Fe)	0.5	<0.5	<0.5	<0.5
Lead (Pb)	0.01	<0.01	<0.01	<0.01
Magnesium (Mg)	5	<5	<5	<5
Manganese (Mn)	0.01	<0.01	<0.01	<0.01
Molybdenum (Mo)	0.01	<0.01	<0.01	<0.01
Nickel (Ni)	0.02	<0.02	<0.02	<0.02
Phosphorus (P)	0.5	1.79	1.85	<0.5
Potassium (K)	10	<10	<10	<10
Selenium (Se)	0.05	<0.05	<0.05	<0.05
Silicon (Si)	10	<10	<10	<10
Silver (Ag)	0.001	<0.001	<0.001	<0.001
Sodium (Na)	5	226	226	213
Strontium (Sr)	0.01	<0.01	<0.01	<0.01
Thallium (Tl)	0.003	<0.003	<0.003	<0.003
Tin (Sn)	0.01	<0.01	0.01	<0.01
Titanium (Ti)	0.02	<0.02	<0.02	<0.02
Tungsten (W)	0.1	<0.1	<0.1	<0.1
Uranium (U)	0.05	<0.05	<0.05	<0.05
Vanadium (V)	0.01	<0.01	<0.01	<0.01
Zinc (Zn)	0.03	0.382	0.380	<0.030
Zirconium (Zr)	0.04	<0.04	<0.04	<0.04

Table A4.4: The table shows metal contents of putative dioxygenase (PDO, 5.28 μM) from *Streptomyces coelicolor* in Chelex-treated water. The experiment was performed in duplicated using Chelex-treated water as reference. The metals that have higher level of detection than those of the reference are highlight in grey.

Metal analysis	*Detection limit (mg/L)*	*Metal detection (mg/L) Trial 1*	*Metal detection (mg/L) Trial 2*	*Metal detection in Chelex-treated water (mg/L)*
Aluminum (Al)	0.01	<0.01	<0.01	<0.01
Antimony (Sb)	0.005	<0.005	<0.005	<0.005
Arsenic (As)	0.001	<0.001	<0.001	<0.001
Barium (Ba)	0.01	<0.01	<0.01	<0.01
Berylium (Be)	0.001	<0.001	<0.001	<0.001
Bismuth (Bi)	0.001	<0.001	<0.001	<0.001
Boron (B)	0.05	<0.05	<0.05	<0.05
Cadmium (Cd)	0.0001	<0.0001	<0.0001	<0.0001
Calcium (Ca)	0.5	<0.5	<0.5	<0.5
Chromium (Cr)	0.001	<0.001	<0.001	<0.001
Cobalt (Co)	0.0005	0.0006	0.0006	<0.0005
Copper (Cu)	0.001	0.005	0.005	<0.001
Iron (Fe)	0.05	<0.05	<0.05	<0.05
Lead (Pb)	0.001	<0.001	<0.001	<0.001
Magnesium (Mg)	0.5	<0.5	<0.5	<0.5
Manganese (Mn)	0.001	<0.001	<0.001	<0.001
Molybdenum (Mo)	0.001	<0.001	<0.001	<0.001
Nickel (Ni)	0.002	0.053	0.051	<0.002
Phosphorus (P)	0.05	<0.05	<0.05	<0.05
Potassium (K)	1	<1	<1	<1
Selenium (Se)	0.005	<0.005	<0.005	<0.005
Silicon (Si)	0.1	<0.1	<0.1	<0.1
Silver (Ag)	0.0001	<0.0001	<0.0001	<0.0001
Sodium (Na)	0.5	<0.5	<0.5	<0.5
Strontium (Sr)	0.001	<0.001	<0.001	<0.001
Thallium (Tl)	0.0003	<0.0003	<0.0003	<0.0003
Tin (Sn)	0.001	<0.001	<0.001	<0.001
Titanium (Ti)	0.002	<0.002	<0.002	<0.002
Tungsten (W)	0.01	<0.01	<0.01	<0.01
Uranium (U)	0.005	<0.005	<0.005	<0.005
Vanadium (V)	0.001	<0.001	<0.001	<0.001
Zinc (Zn)	0.003	0.057	0.056	<0.003
Zirconium (Zr)	0.004	<0.004	<0.004	<0.004

Table A4.5: The table shows metal contents of putative lyase (PLA, 4.54 μM) from *Streptomyces coelicolor* in Chelex-treated water. The experiment was performed in duplicated using Chelex-treated water as reference. The metals that have higher level of detection than those of the reference are highlight in grey.

Metal analysis	*Detection limit (mg/L)*	*Metal detection (mg/L) Trial 1*	*Metal detection (mg/L) Trial 2*	*Metal detection in Chelex-treated water (mg/L)*
Aluminum (Al)	0.01	<0.01	<0.01	<0.01
Antimony (Sb)	0.005	<0.005	<0.005	<0.005
Arsenic (As)	0.001	<0.001	<0.001	<0.001
Barium (Ba)	0.01	<0.01	<0.01	<0.01
Berylium (Be)	0.001	<0.001	<0.001	<0.001
Bismuth (Bi)	0.001	<0.001	<0.001	<0.001
Boron (B)	0.05	<0.05	<0.05	<0.05
Cadmium (Cd)	0.0001	0.0002	0.0002	<0.0001
Calcium (Ca)	0.5	<0.5	<0.5	<0.5
Chromium (Cr)	0.001	<0.001	<0.001	<0.001
Cobalt (Co)	0.0005	<0.0005	<0.0005	<0.0005
Copper (Cu)	0.001	0.024	0.023	<0.001
Iron (Fe)	0.05	<0.05	<0.05	<0.05
Lead (Pb)	0.001	0.003	0.003	<0.001
Magnesium (Mg)	0.5	<0.5	<0.5	<0.5
Manganese (Mn)	0.001	<0.001	<0.001	<0.001
Molybdenum (Mo)	0.001	<0.001	<0.001	<0.001
Nickel (Ni)	0.002	0.03	0.028	<0.002
Phosphorus (P)	0.05	<0.05	<0.05	<0.05
Potassium (K)	1	<1	<1	<1
Selenium (Se)	0.005	<0.005	<0.005	<0.005
Silicon (Si)	0.1	<0.1	<0.1	<0.1
Silver (Ag)	0.0001	0.0006	0.0005	<0.0001
Sodium (Na)	0.5	2.4	2.4	2.1
Strontium (Sr)	0.001	<0.001	<0.001	<0.001
Thallium (Tl)	0.0003	<0.0003	<0.0003	<0.0003
Tin (Sn)	0.001	<0.001	<0.001	<0.001
Titanium (Ti)	0.002	<0.002	<0.002	<0.002
Tungsten (W)	0.01	<0.01	<0.01	<0.01
Uranium (U)	0.005	<0.005	<0.005	<0.005
Vanadium (V)	0.001	<0.001	<0.001	<0.001
Zinc (Zn)	0.003	0.047	0.046	<0.003
Zirconium (Zr)	0.004	<0.004	<0.004	<0.004

Appendix 5: Determination of the Hemithioacetal Dissociation Constant

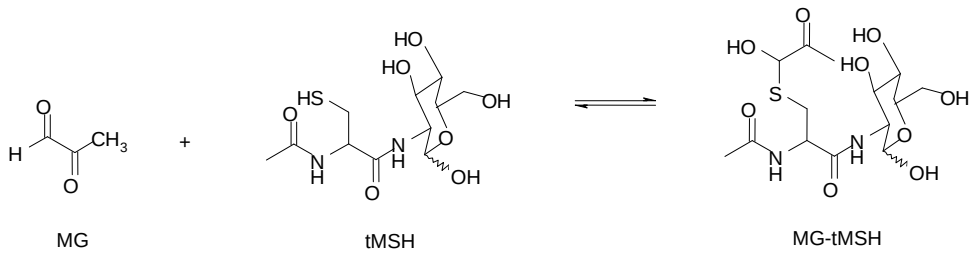


Figure A5.1: The non-enzymatic reaction of methylglyoxal (MG) and des-*myo*-inositol mycothiol (tMSH) to form the hemithioacetal, the Glyoxalase I substrate.

The following derived equations are based on the previous reports for the formation of the hemithioacetal using glutathione as the thiol [85, 86, 141, 279-281]. From Figure A5.1, the dissociation constant of MG and tMSH can be defined as

$$K_d = [MG][tMSH]/[MG\text{-}tMSH] \quad (1),$$

where [MG] is the concentration of the aldehyde methylglyoxal and its hydrate form, [tMSH] is the concentration of free thiol and [MG-tMSH] is the concentration of the hemithioacetal product of the non-enzymatic reaction of MG and tMSH. In this experiment, the concentrations of free thiol and aldehyde are related to their total concentrations as well as the amount used in the hemithioacetal formation that can be written as

$$[MG]_t = [MG] + [MG\text{-}tMSH] \quad (2) \text{ and}$$

$$[tMSH]_t = [tMSH] + [MG\text{-}tMSH] \quad (3)$$

When $[MG] \geq [tMSH]$, $[MG]_t \sim [MG]$ and from the equations (1), (2) and (3), the concentration of hemithioacetal formation can be derived as

$$[MG\text{-}tMSH] = \frac{[MG][tMSH]_t}{K_d + [MG]} \sim \frac{[MG]_t[tMSH]_t}{K_d + [MG]_t} \quad (4)$$

The dissociation of MG-tMSH can be detected by measuring the increase in absorbance at 240 nm using the Beer-Lambert Law, $A = \varepsilon BC$, where A is absorption, ε is extinction coefficient, B is path length and C is concentration. The absorbance at 240 nm of the final absorbance (A_f) of [MG-tMSH], the initial absorbance of $[MG]_t$ (A_i) and reference of $[tMSH]_t$ (A_{ref}) can be defined as

$$A_i = \varepsilon_{MG}[MG]_t \quad (5)$$

$$A_{ref} = \varepsilon_{tMSH}[tMSH]_t \quad (6)$$

$$A_f = \varepsilon_{MG}[MG] + \varepsilon_{tMSH}[tMSH] + \varepsilon_{MG\text{-}tMSH}[MG\text{-}tMSH] \quad (7)$$

The change in absorbance at 240 nm (ΔA_{240}) is the subtraction of the final absorbance with the initial absorbance and its reference such that

$$\Delta A_{240} = A_f - A_i - A_{ref}$$

From the equations (2), (3), (5), (6) and (7), ΔA_{240} can be derived as

$$\Delta A_{240} = (\varepsilon_{MG\text{-}tMSH} - \varepsilon_{MG} - \varepsilon_{tMSH})[MG\text{-}tMSH] \quad (8)$$

By substituting (4) into (8), the dissociation constant can be calculated from the equation

$$\Delta A_{240} = \frac{C[MG]_t}{K_d + [MG]_t}$$

where $C = (\varepsilon_{MG\text{-}tMSH} - \varepsilon_{MG} - \varepsilon_{tMSH})[tMSH]_t$.

Appendix 6: ESI-MS Results

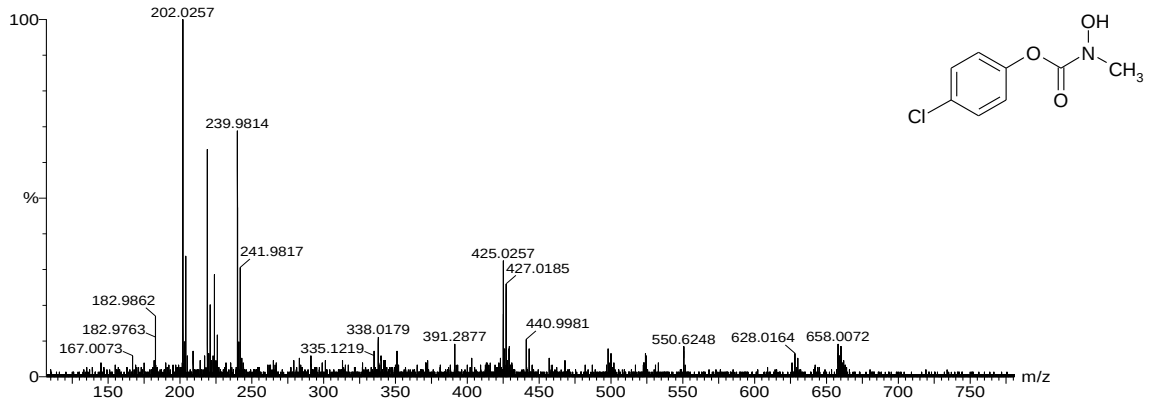


Figure A6.1: The mass spectrum of purified *N*-hydroxy-*N*-methylcarbamate 4-chlorophenyl ester (m/z 202 Da, MH^+). The mass at 240 Da corresponds to *N*-hydroxy-*N*-methylcarbamate 4-chlorophenyl ester with a K^+ atom and the mass at 425 Da indicated the mass of the product with possibly a K^+ and $8Na^+$ atoms or a noncovalent dimer with a Na^+ ion as an alternate possibility. The mass spectrometric results are in complete agreement with the literature data [114].

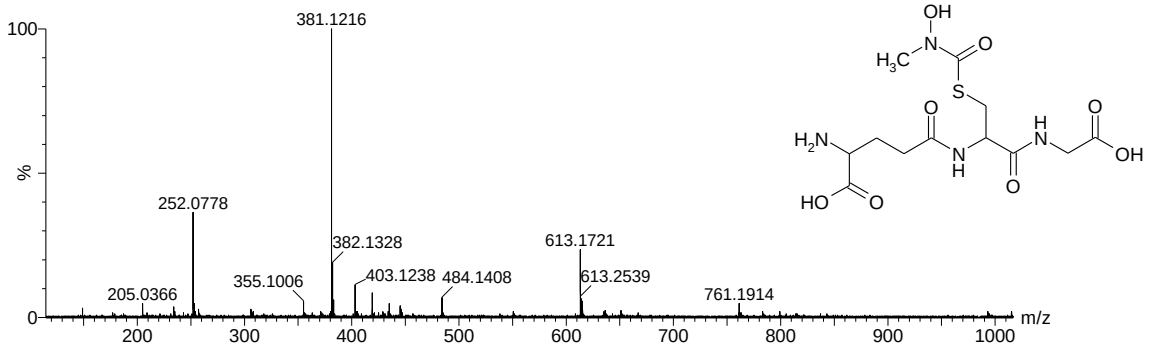


Figure A6.2: The mass spectrum of purified *S*-(*N*-hydroxy-*N*-methylcarbamoyl)glutathione (m/z 381 Da, MH^+). The mass at 252 Da corresponds to *S*-(*N*-hydroxy-*N*-methylcarbamoyl)glutathione without the glutamate moiety $(M\text{-}C(O)CH_2CH_2CH(CO_2H)NH_2+2H)^+$. The mass spectrometric results are in complete agreement with the literature data [114].

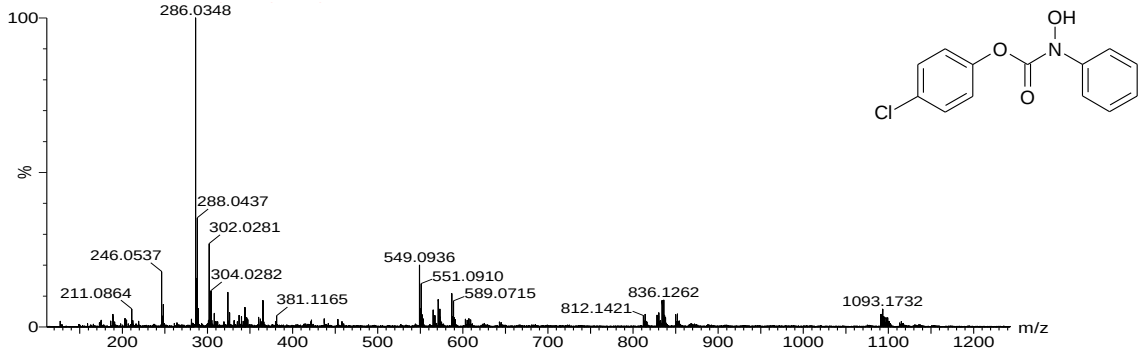


Figure A6.3: The mass spectrum of purified *N*-hydroxy-*N*-phenylcarbamate 4-chlorophenyl ester (predicted m/z 263 Da). The mass at 286 Da corresponds to *N*-hydroxy-*N*-methylcarbamate 4-chlorophenyl ester with Na^+ atom and the mass at 549 Da indicates the mass of a possible noncovalent dimer and Na^+ atom. The mass spectrometric results are in complete agreement with the literature data [115].

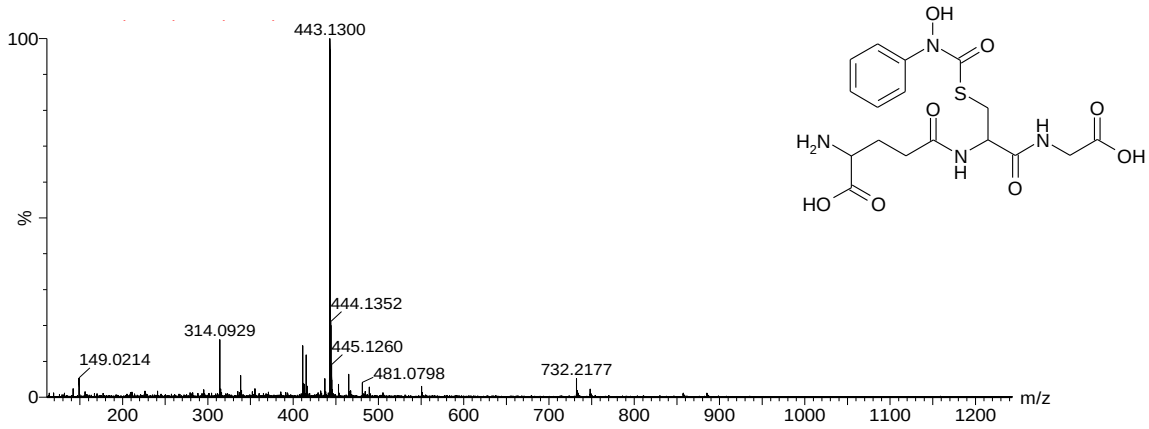


Figure A6.4: The mass spectrum of purified *S*-(*N*-hydroxy-*N*-phenylcarbamoyl)glutathione (m/z 443 Da, MH^+). The mass spectrometric results are in complete agreement with the literature data [115].

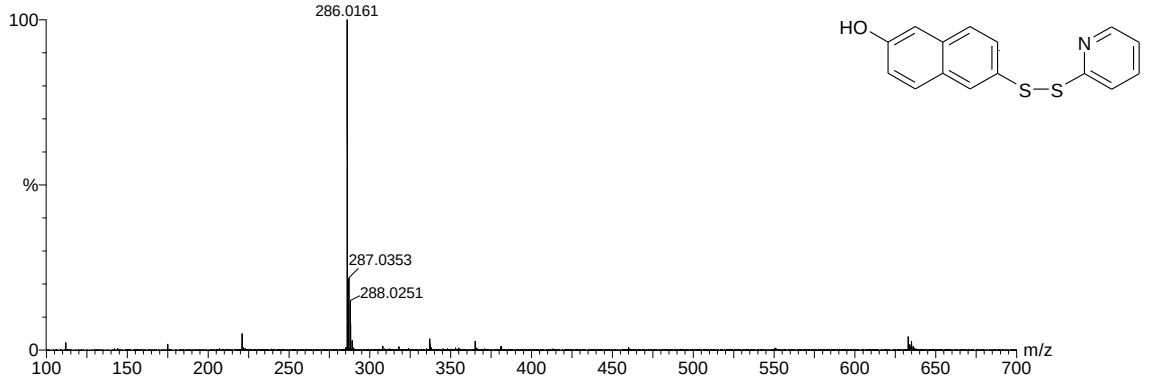


Figure A6.5: The mass spectrum of purified 2-*S*-(2'-thiopyridyl)-6-hydroxynaphthyldisulfide (m/z 286 Da, MH^+). The mass spectrometric results are in complete agreement with the literature data [251].

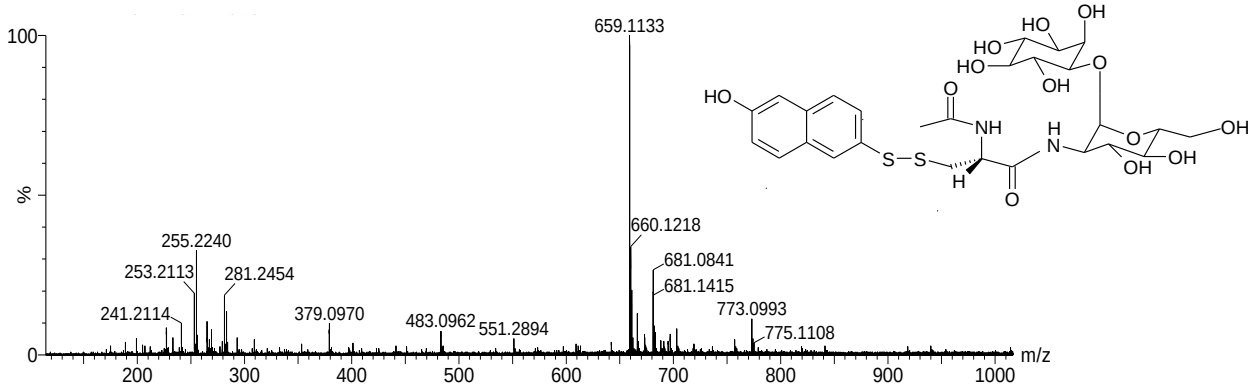


Figure A6.6: The mass spectrum of purified *S*-2-(mycothiolyl)-6-hydroxynaphthyldisulfide (M-H) (m/z 659 Da). Mass at ~ 681 likely corresponds to mass of a gas phase complex of *S*-2-(mycothiolyl)-6-hydroxynaphthyldisulfide with Na^+ atom. The mass spectrometric results are in complete agreement with the literature data [251].

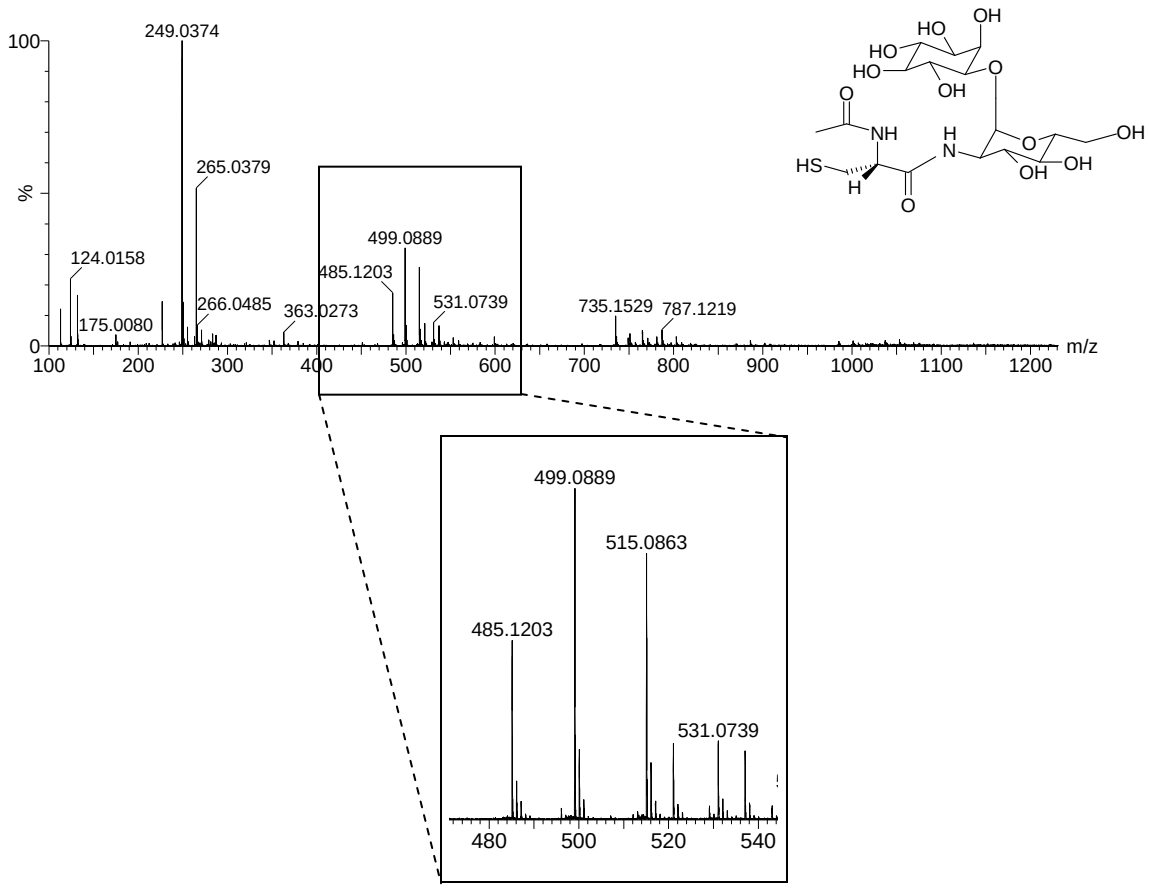


Figure A6.7: The mass spectrum of the isolated natural mycothiol from *Streptomyces jumonjinesis*. The mass at m/z 485 corresponds to MSH $(MSH)^-$. The mass at m/z 249 and 265 corresponds to TCEP and TCEP oxide, respectively. The mass at m/z 499 Da corresponds to that of 2TCEP and the mass at m/z 515 Da was 2TCEP-O.

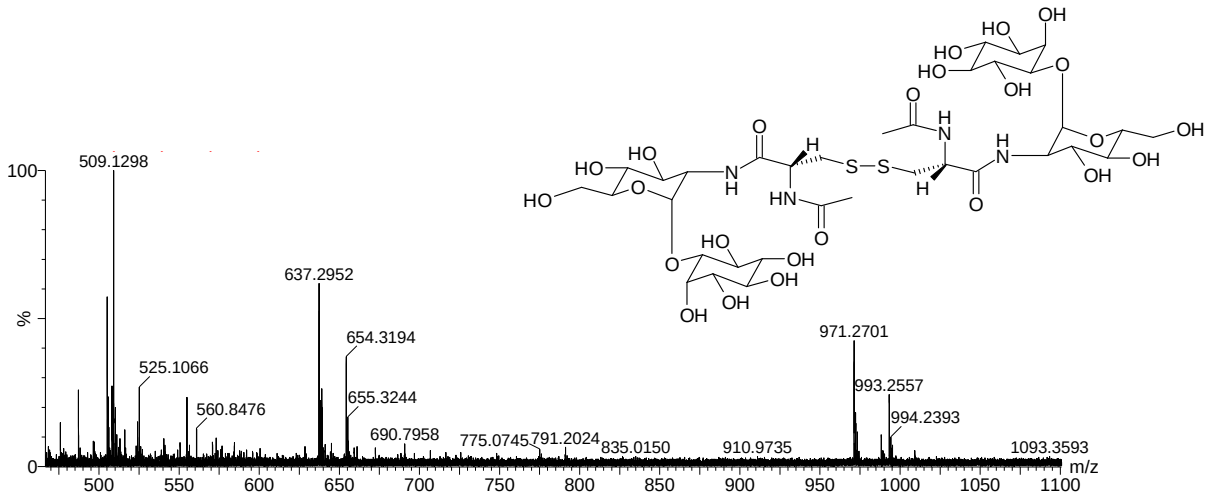


Figure A6.8: The mass spectrum of the disulfide form of mycothiol (supplied by Dr. Gerald Newton from University of California, SanDiego, La Jolla, CA). The mass at m/z 971.3 corresponded to disulfide form of mycothiol $(MSSMH^+)$ (calculated molecular mass is 970.98 Da). The mass at m/z 993.3 and 509.1 corresponds to Na^+ adducts of MSSM and MSH, respectively.

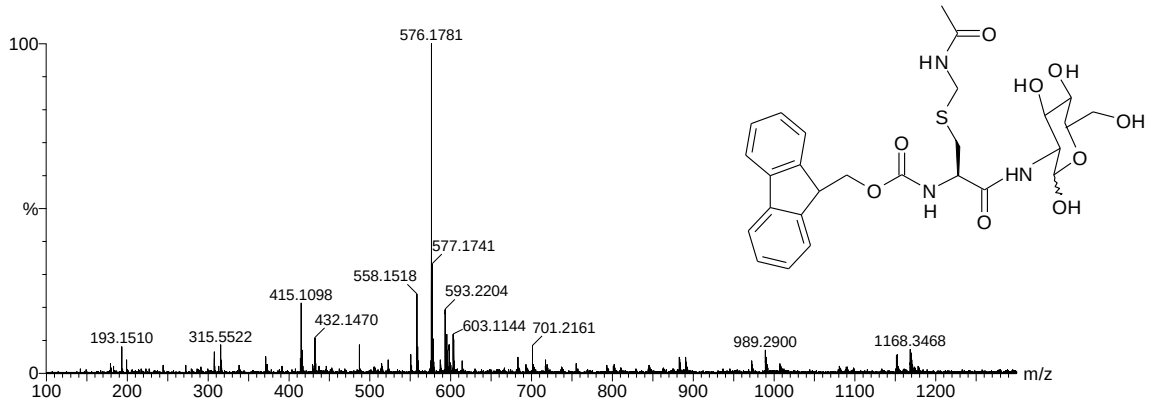


Figure A6.9: The mass spectrum of purified *N*-α-Fmoc-*S*-acetamidomethyl-L-cysteinyl-2-amino-2-deoxy-*α*-D-glucopyranoside (Fmoc-Cys(Acm)-GlcN) (m/z 576 Da, MH^+). The mass spectrometric results are in complete agreement with the literature data [116].

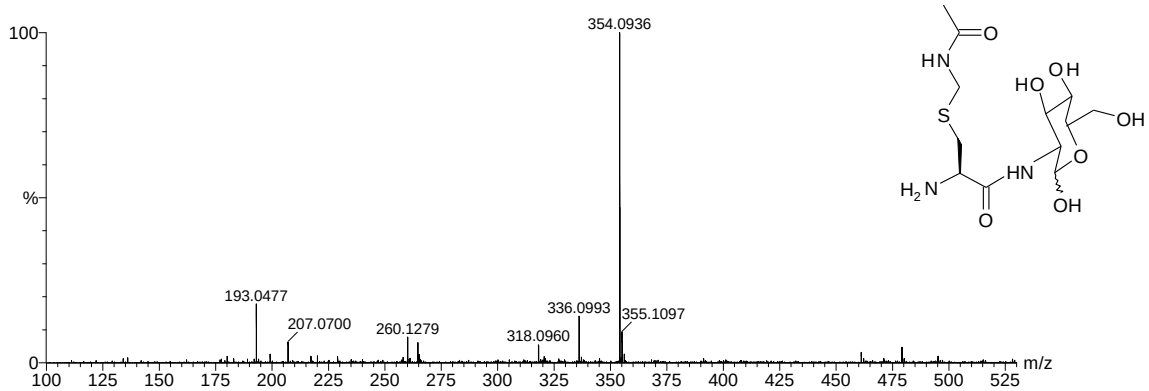


Figure A6.10: The mass spectrum of purified *S*-acetamidomethyl-L-cysteinyl-2-amino-2-deoxy-*α,β*-D-glucopyranoside (NH_2-Cys(Acm)-GlcN) (m/z 354 Da, MH^+). The mass spectrometric results are in complete agreement with the predicted molecular mass of the compound.

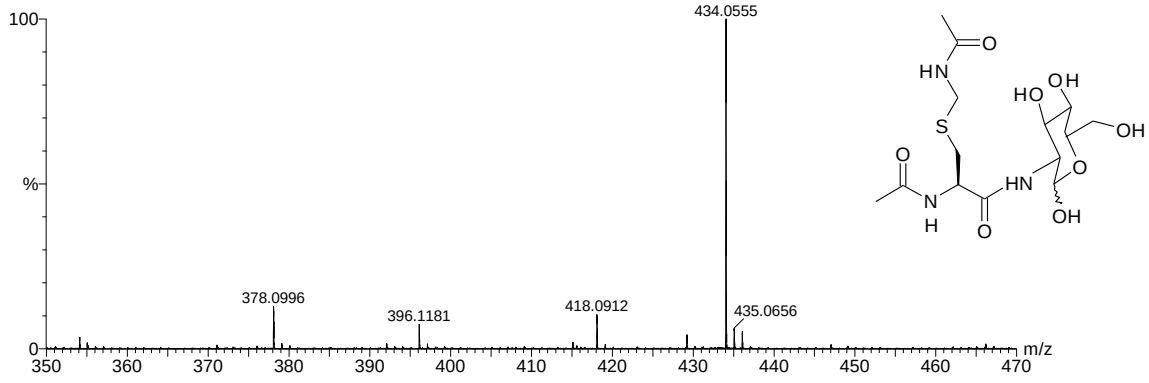


Figure A6.11: The mass spectrum of *N*-acetyl-*S*-acetamidomethyl-L-cysteinyl-2-amino-2-deoxy-*α,β*-D-glucopyranoside (Ac-Cys(Acm)-GlcN) (the predicted m/z 395 Da). The mass at m/z 434 Da indicated the mass of Ac-Cys(Acm)-GlcN with a K^+ atom. The mass spectrometric results are in complete agreement with the literature data [116].

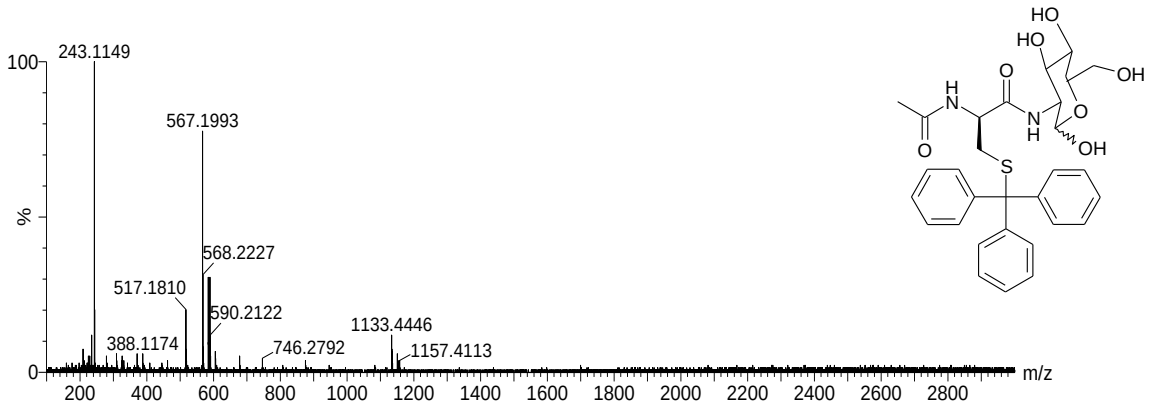


Figure A6.12: The mass spectrum of *N*-acetyl-*S*-trityl-L-cysteinyl-2-amino-2-deoxy-α,β-D-glucopyranoside (Ac-Cys(Trt)-GlcN). The mass at m/z 567 corresponds to Ac-Cys(Trt)-GlcN (MH^+) and m/z 243 indicates the trityl group (Trt). The mass spectrometric results are in complete agreement with the predicted molecular mass of the compound [253].

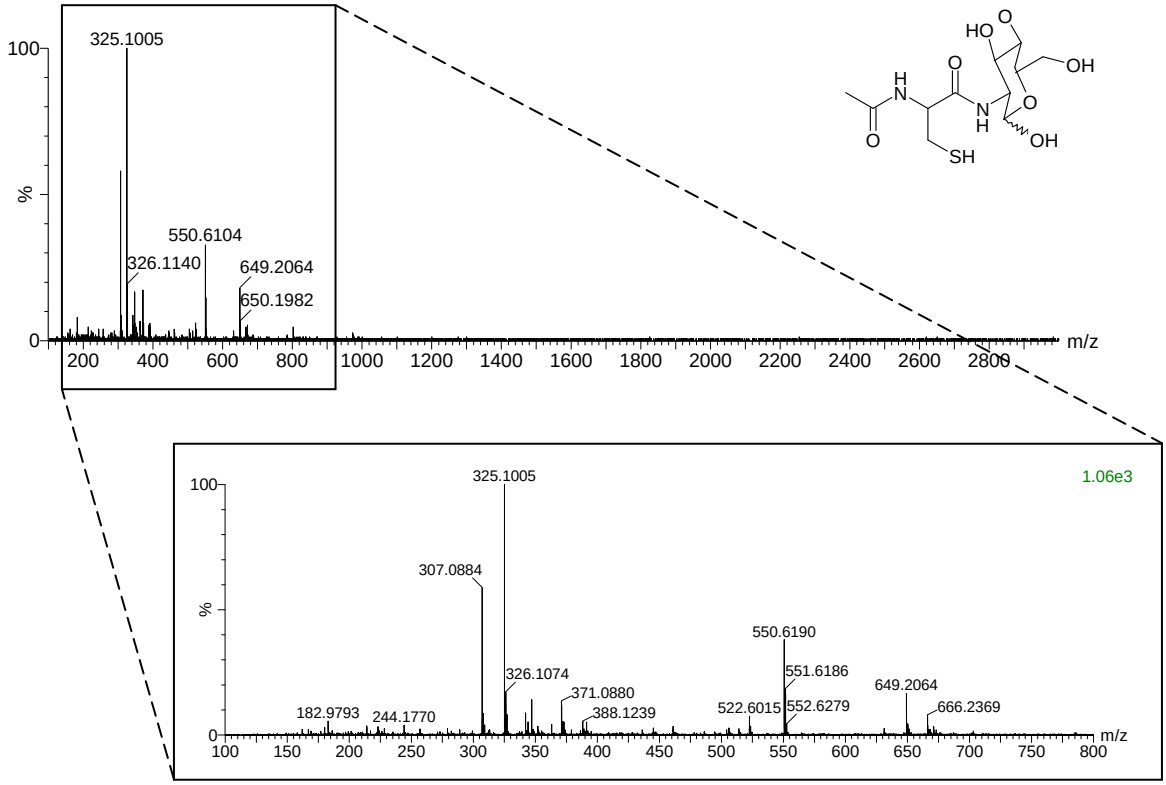
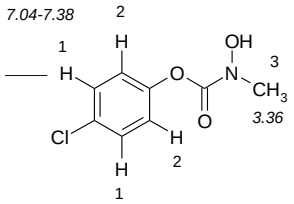


Figure A6.13: The mass spectrum of *N*-acetyl-L-cysteinyl-2-amino-2-deoxy-α,β-D-glucopyranoside (AcCys-GlcN). The mass at m/z 325 corresponds in mass to AcCys-GlcN and the mass at m/z 649 corresponds to GlcN-AcCys-AcCys-GlcN. The mass at m/z 550 is from background. The mass spectrometric results are in complete agreement with the literature data [116].

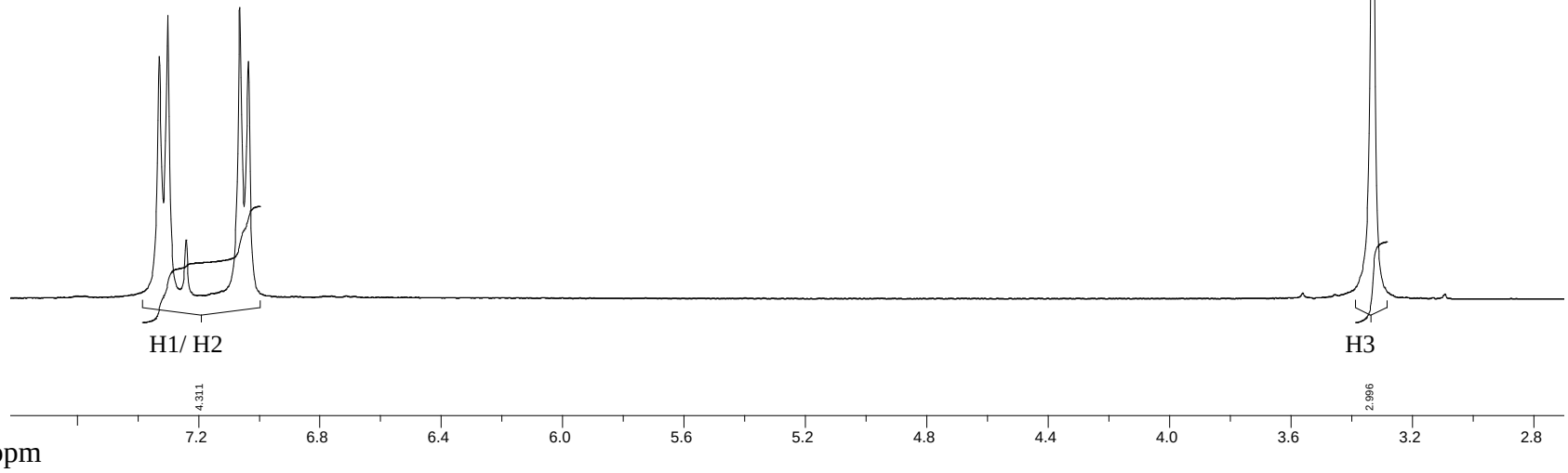
Appendix 7: ^{1}H and ^{13}C NMR Spectra

The figures show the structure of the compounds of interest with numbers to indicate the hydrogens (or carbons) that correspond to the chemical shifts in the NMR spectra as well as the predicted chemical shifts (in the structure of each compound) that are estimated by ChemBioDraw Ultra 12.0 (http://www.cambridgesoft.com) software or by previous literature reports.

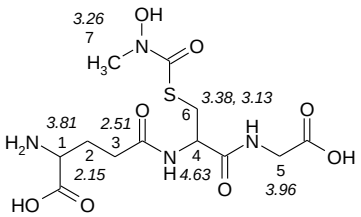
Spectrum A7.1: ^{1}H NMR of *N*-hydroxy-*N*-methylcarbamate 4-chlorophenyl ester in $CDCl_3$



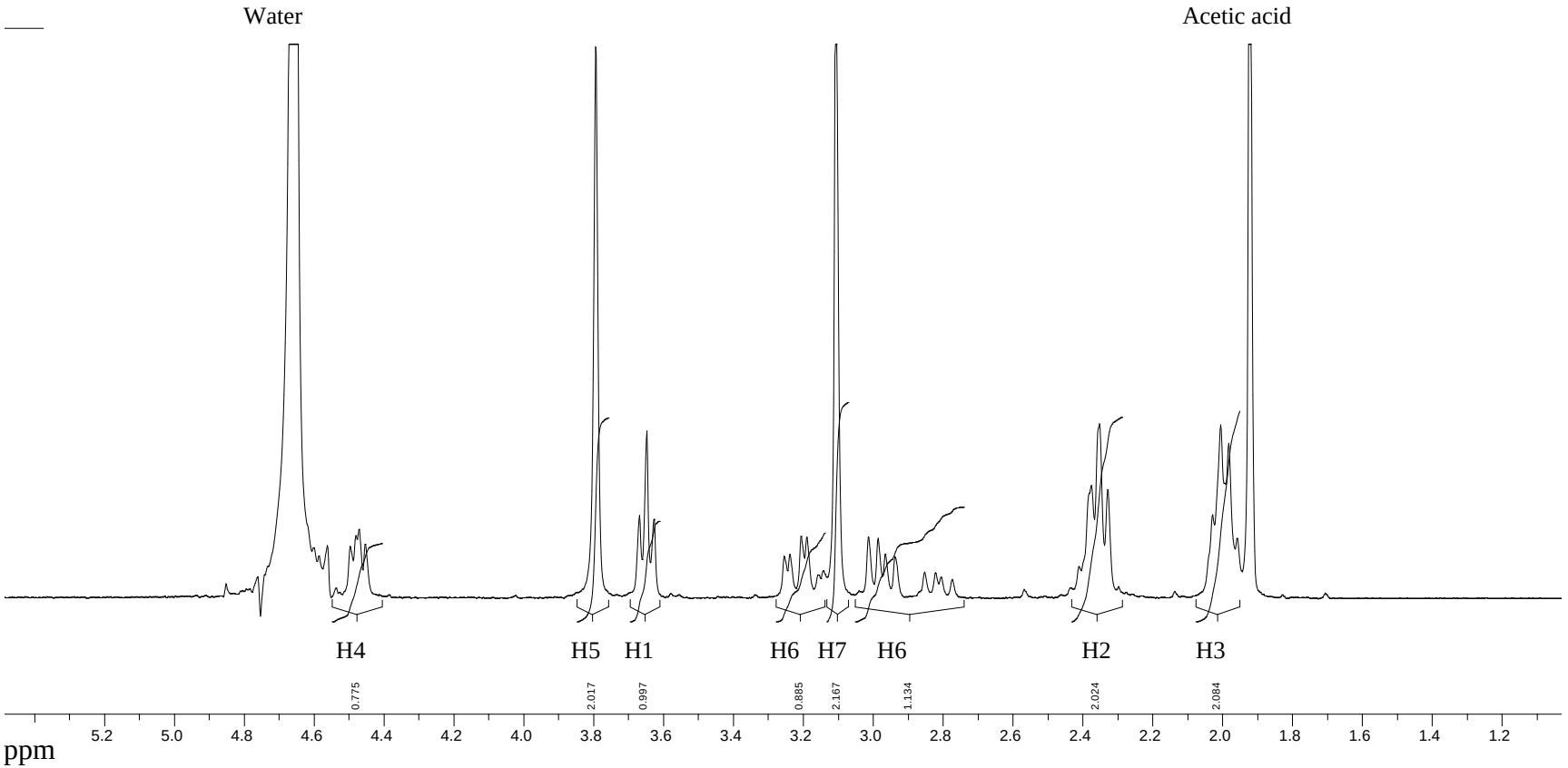
Chemical shifts are shown according to the previous report by Hamilton and Creighton, 1992 [114]



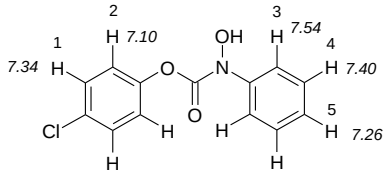
Spectrum A7.2: ^{1}H NMR of *S*-(*N*-hydroxy-*N*-methylcarbamoyl)glutathione in D_2O



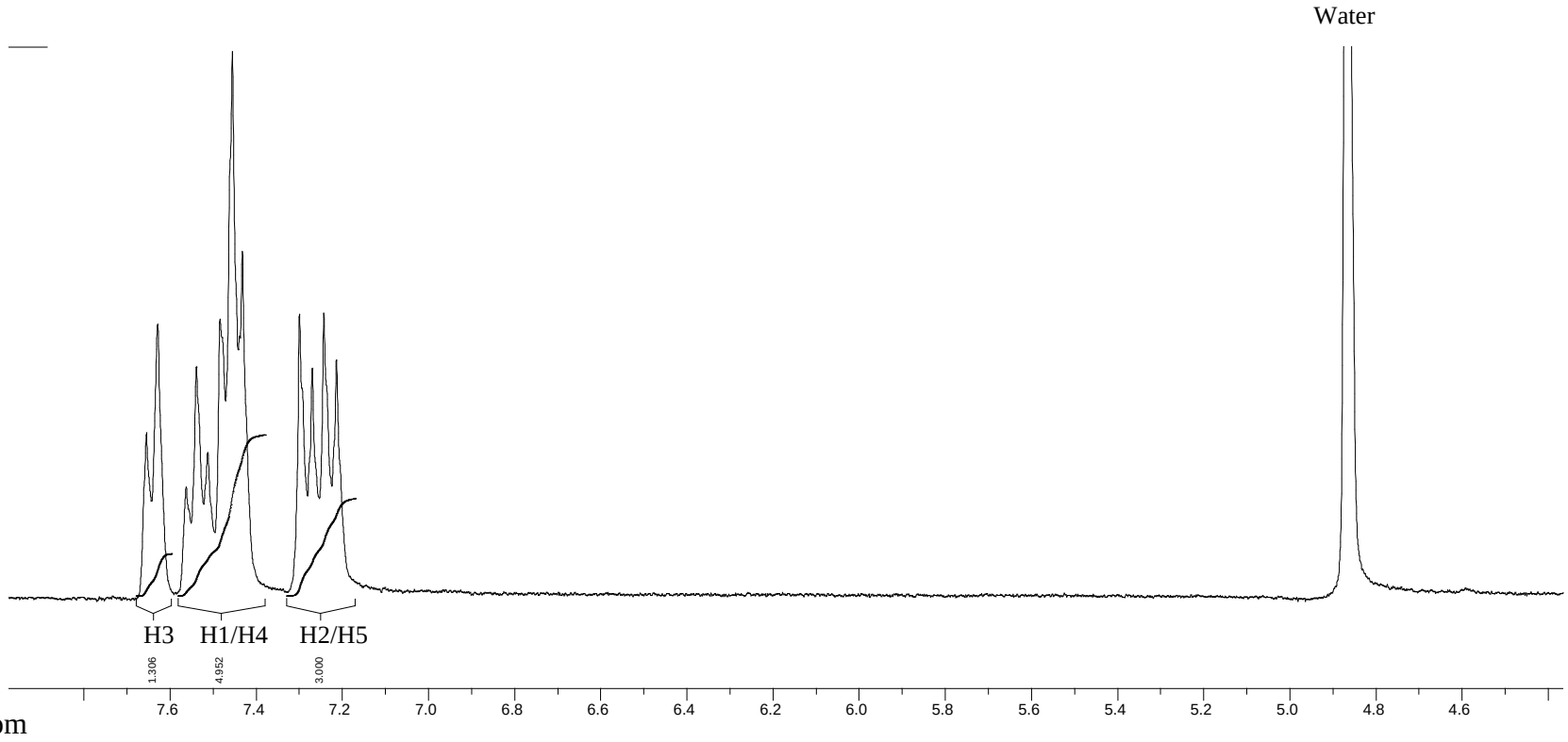
Chemical shifts are shown according to the previous report by Hamilton and Creighton, 1992 [114]



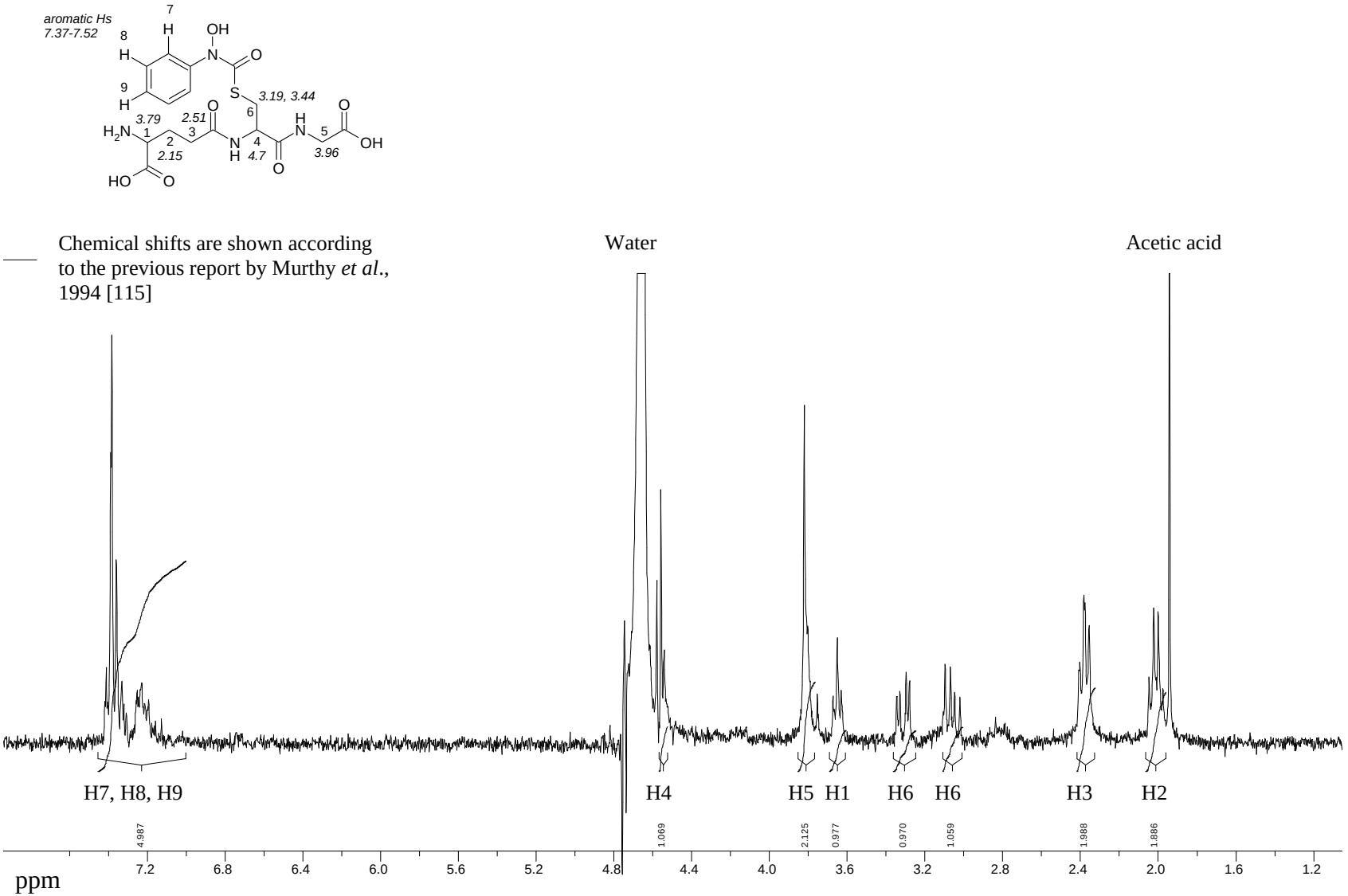
Spectrum A7.3: ^{1}H NMR of *N*-hydroxy-*N*-phenylcarbamate 4-chlorophenyl ester in $CDCl_3$



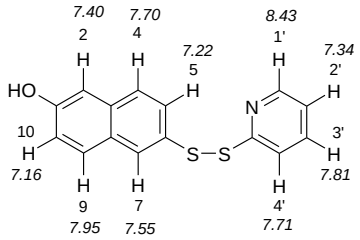
Chemical shifts are shown according to the previous report by Murthy *et al*., 1994 [115]



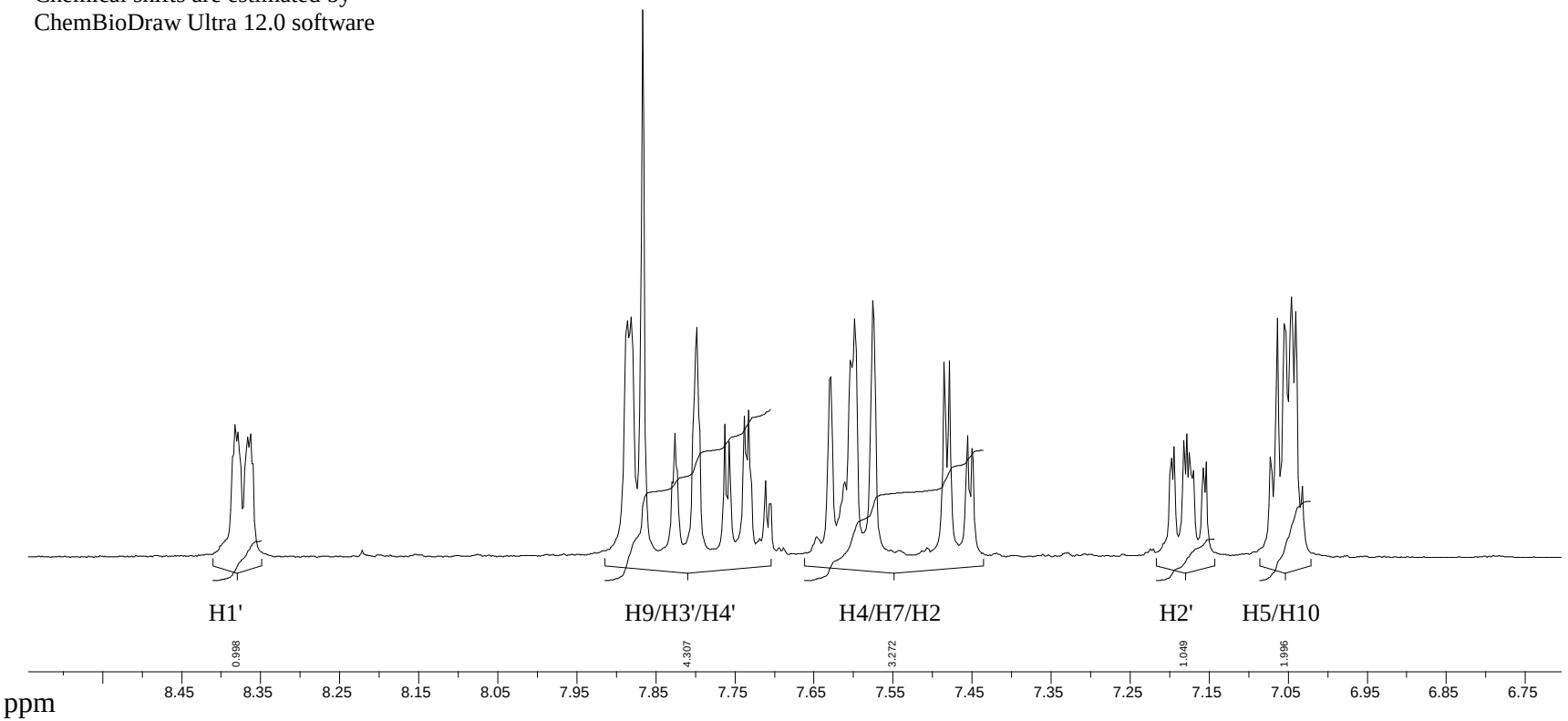
Spectrum A7.4: ^{1}H NMR of *S*-(*N*-hydroxy-*N*-phenylcarbamoyl)glutathione in D_2O



Spectrum A7.5: ^{1}H NMR of 2-*S*-(2'-thiopyridyl)-6-hydroxynaphthyldisulfide in CD_3OD

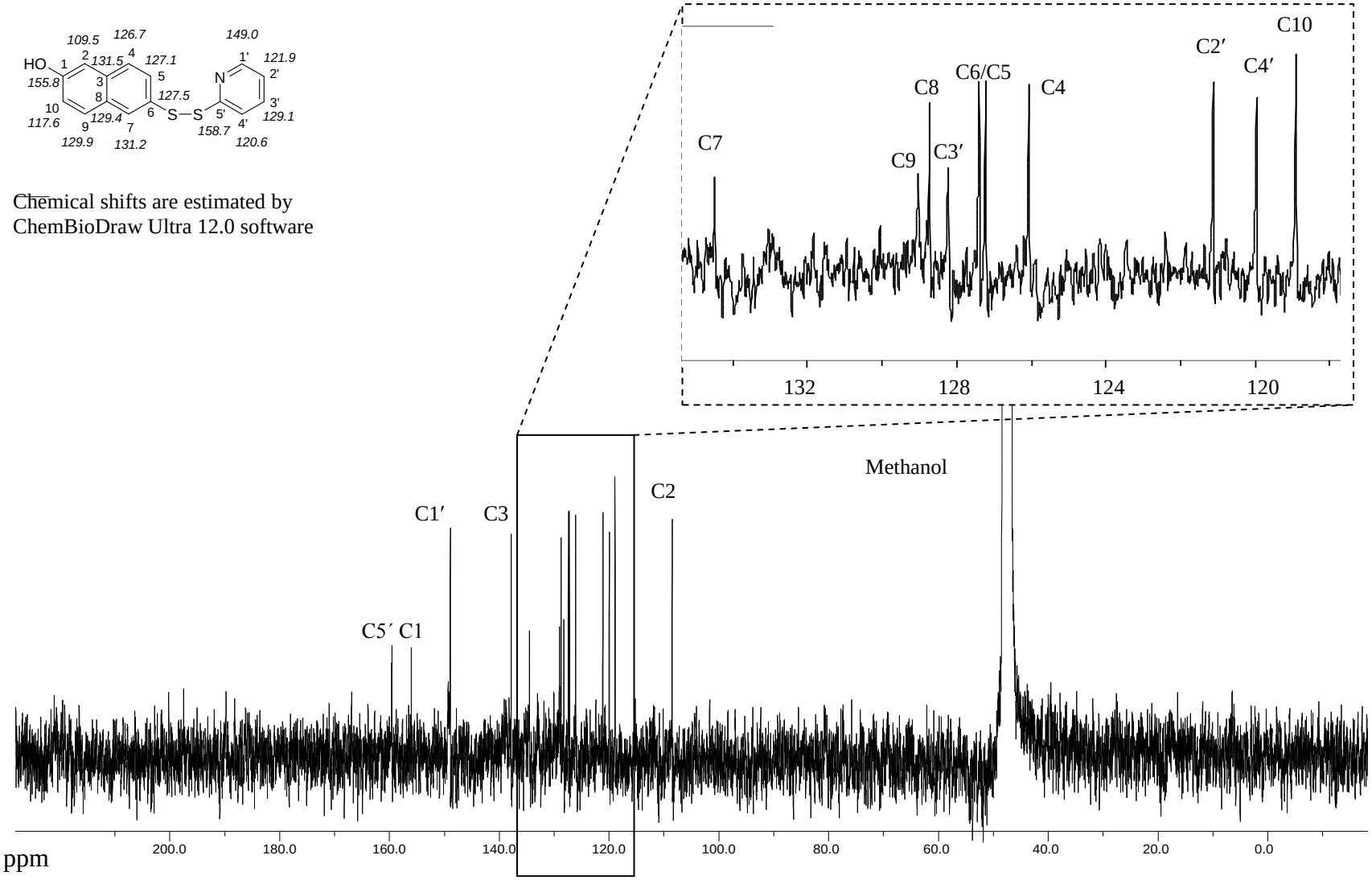


Chemical shifts are estimated by ChemBioDraw Ultra 12.0 software

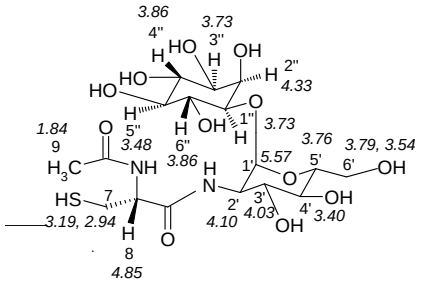


Spectrum A7.6: ^{13}C NMR of 2-*S*-(2'-thiopyridyl)-6-hydroxynaphthyldisulfide in CD_3OD

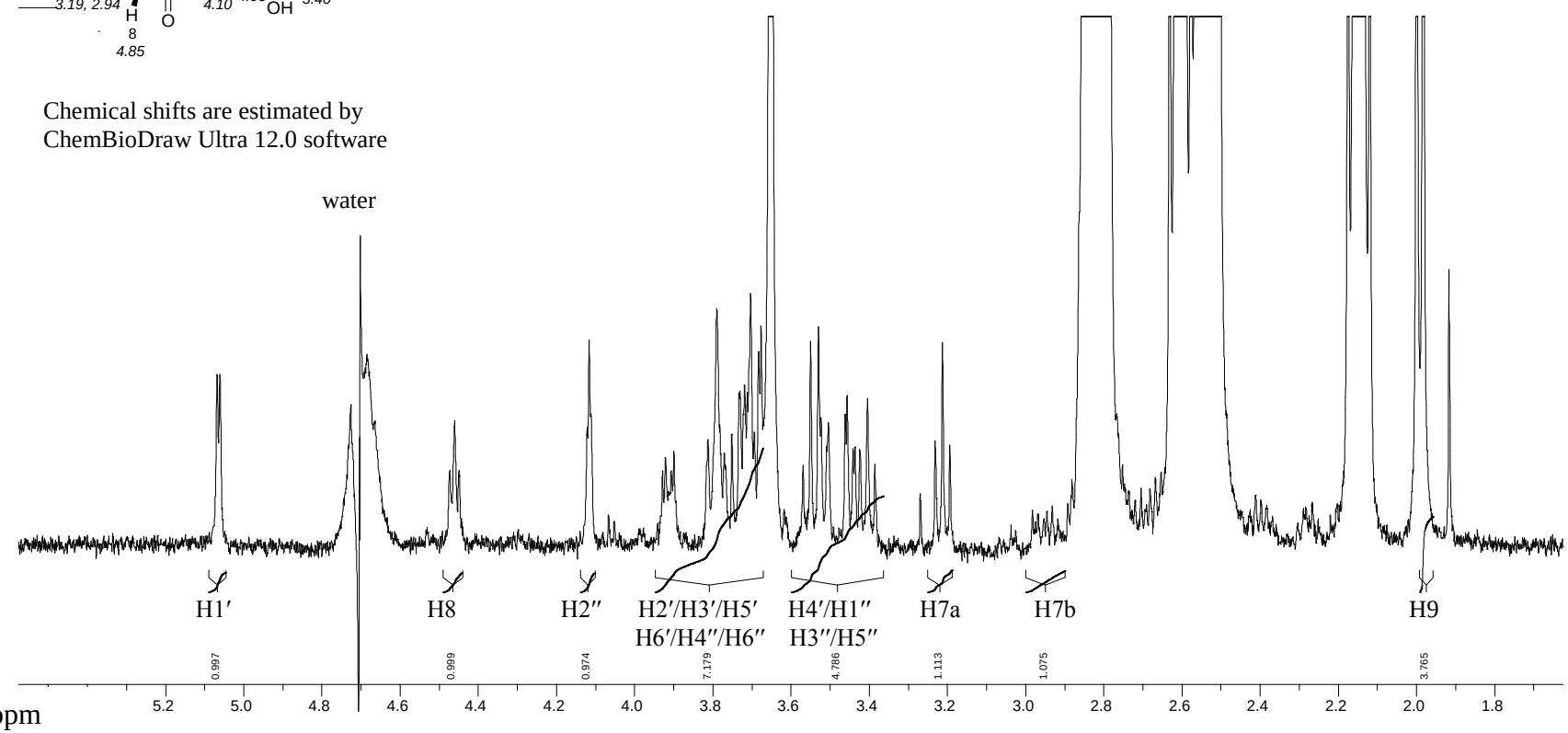
Chemical shifts are estimated by ChemBioDraw Ultra 12.0 software



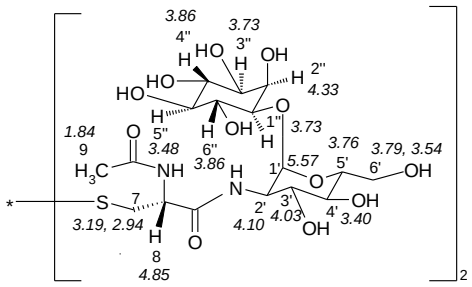
Spectrum A7.7: ^{1}H NMR of mycothiol in CD_3OD



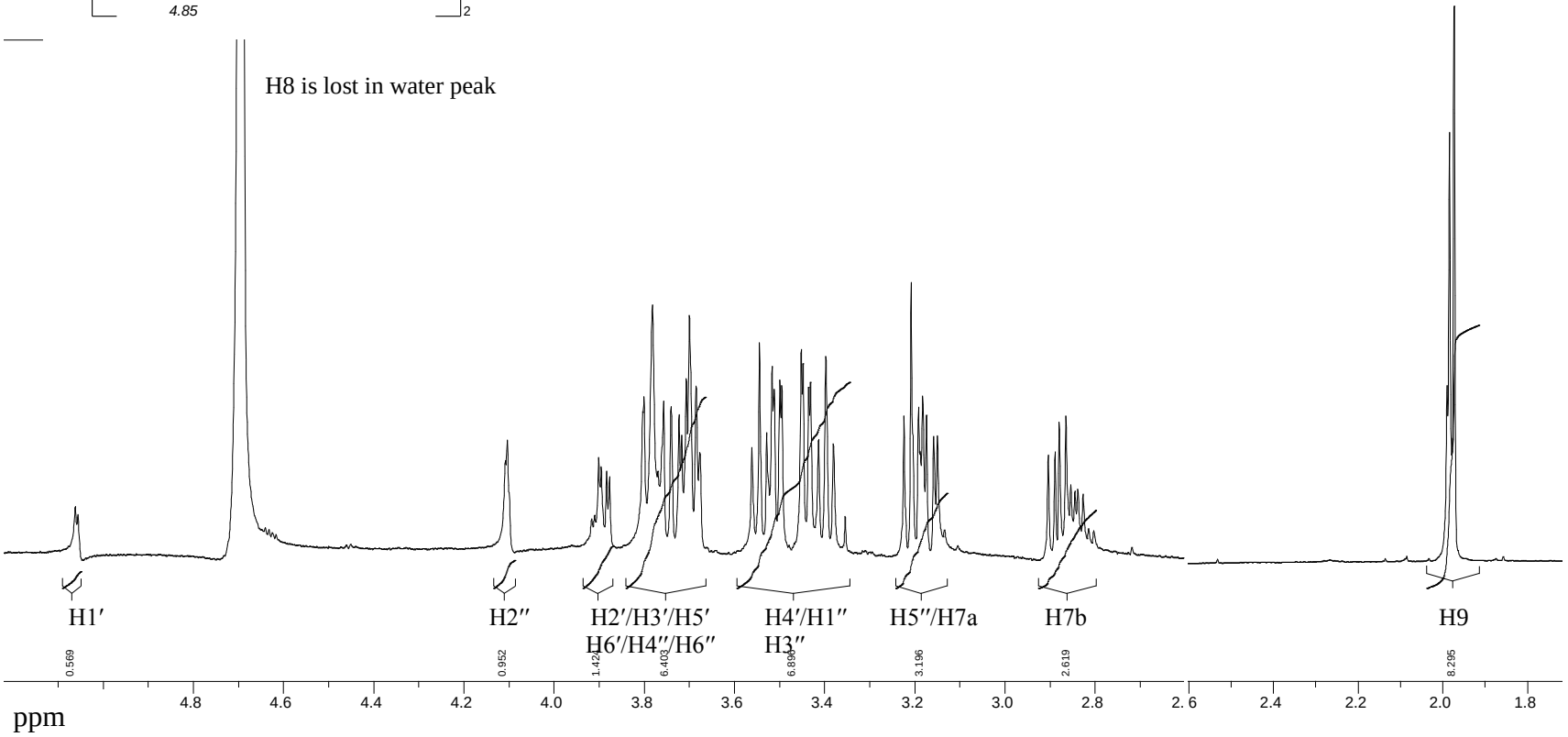
Chemical shifts are estimated by
ChemBioDraw Ultra 12.0 software



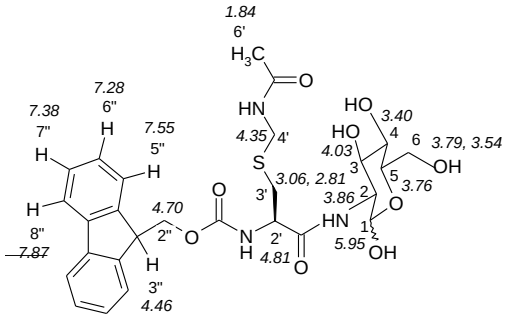
Spectrum A7.8: ^{1}H NMR of disulfide mycothiol in D_2O



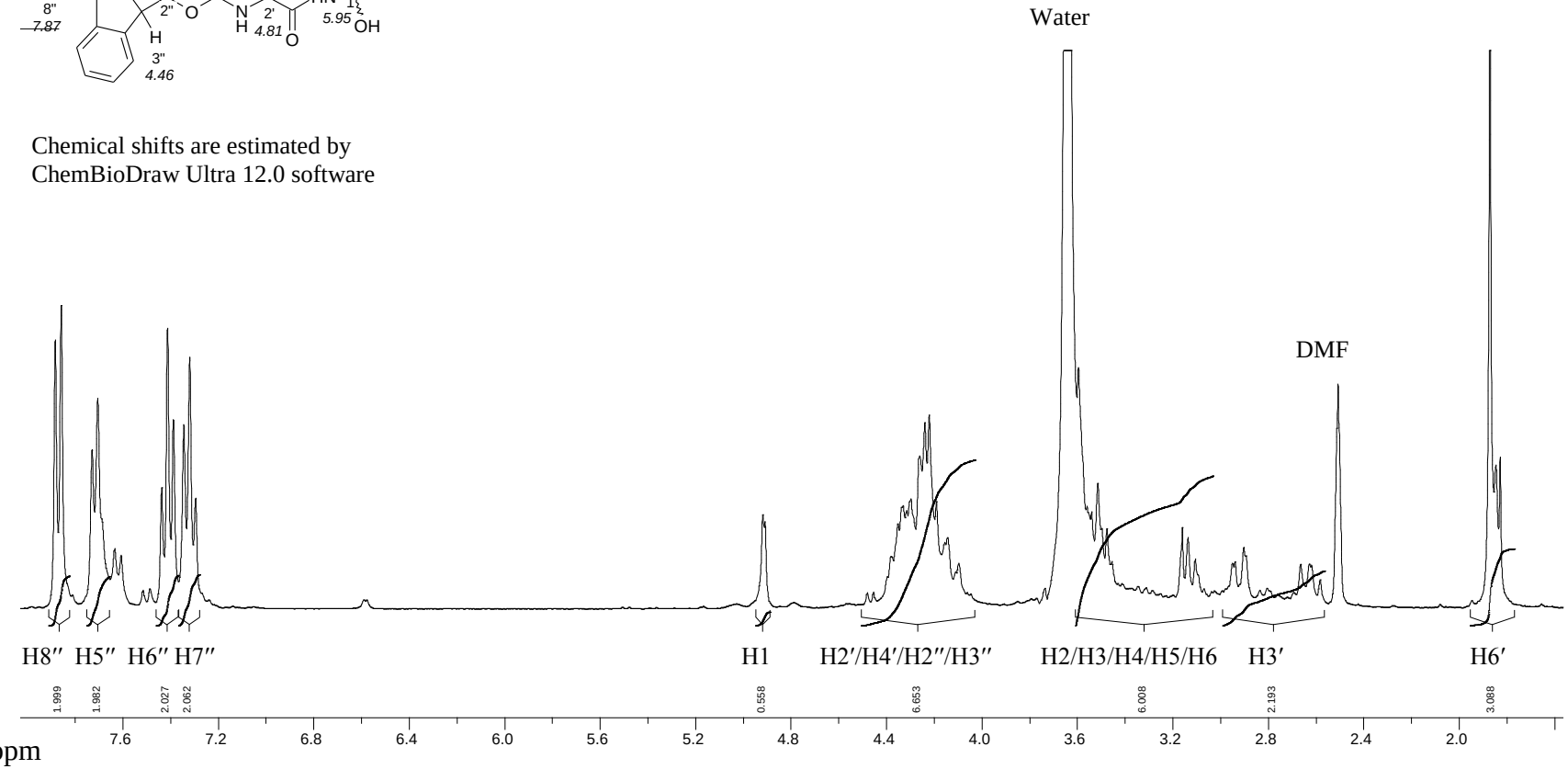
Chemical shifts are estimated by ChemBioDraw Ultra 12.0 software



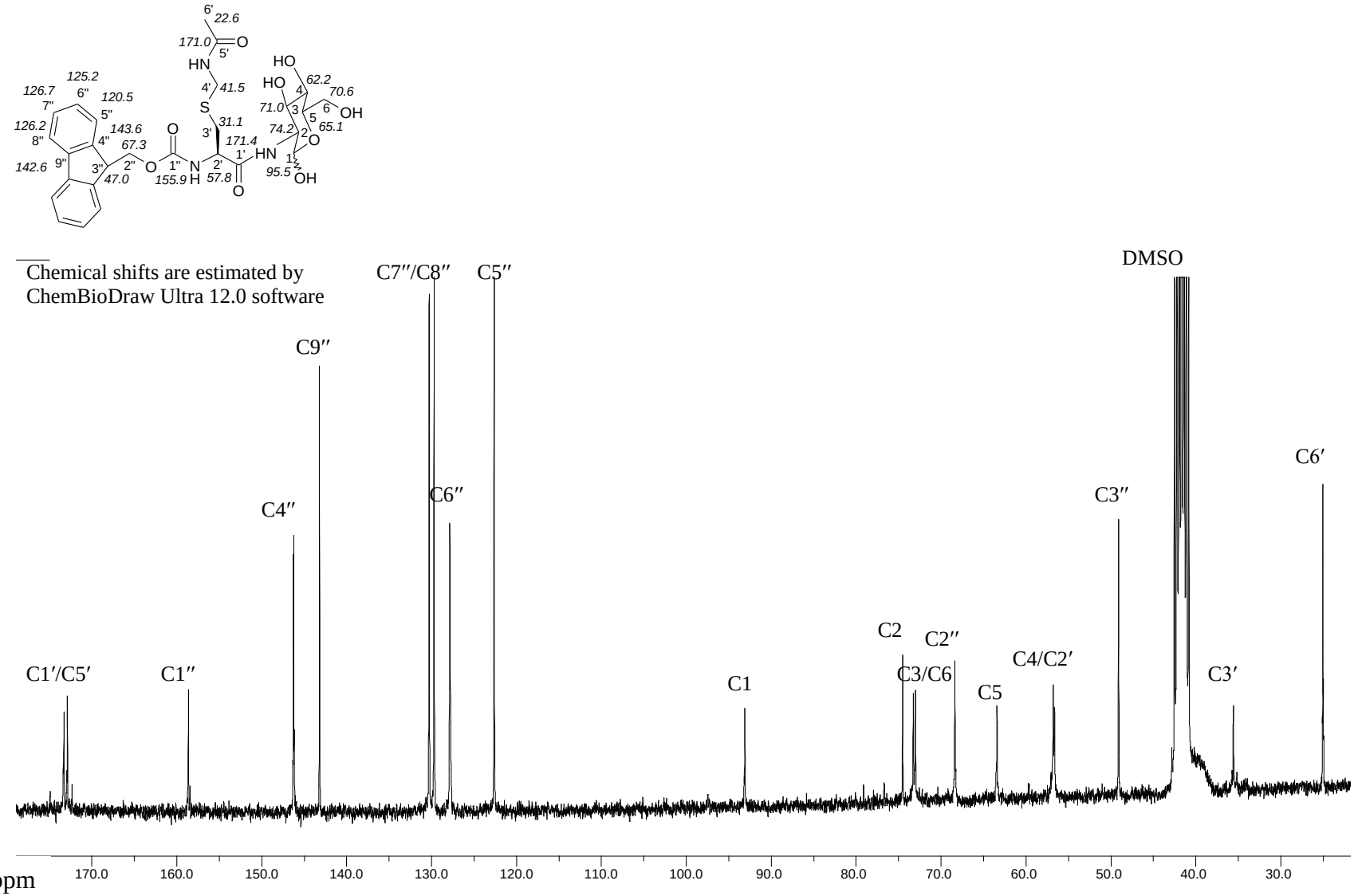
Spectrum A7.9: ^{1}H NMR of Fmoc-Cys(Acm)-GlcN in deuterated DMSO/D_2O



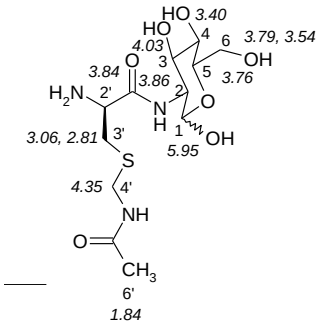
Chemical shifts are estimated by
ChemBioDraw Ultra 12.0 software



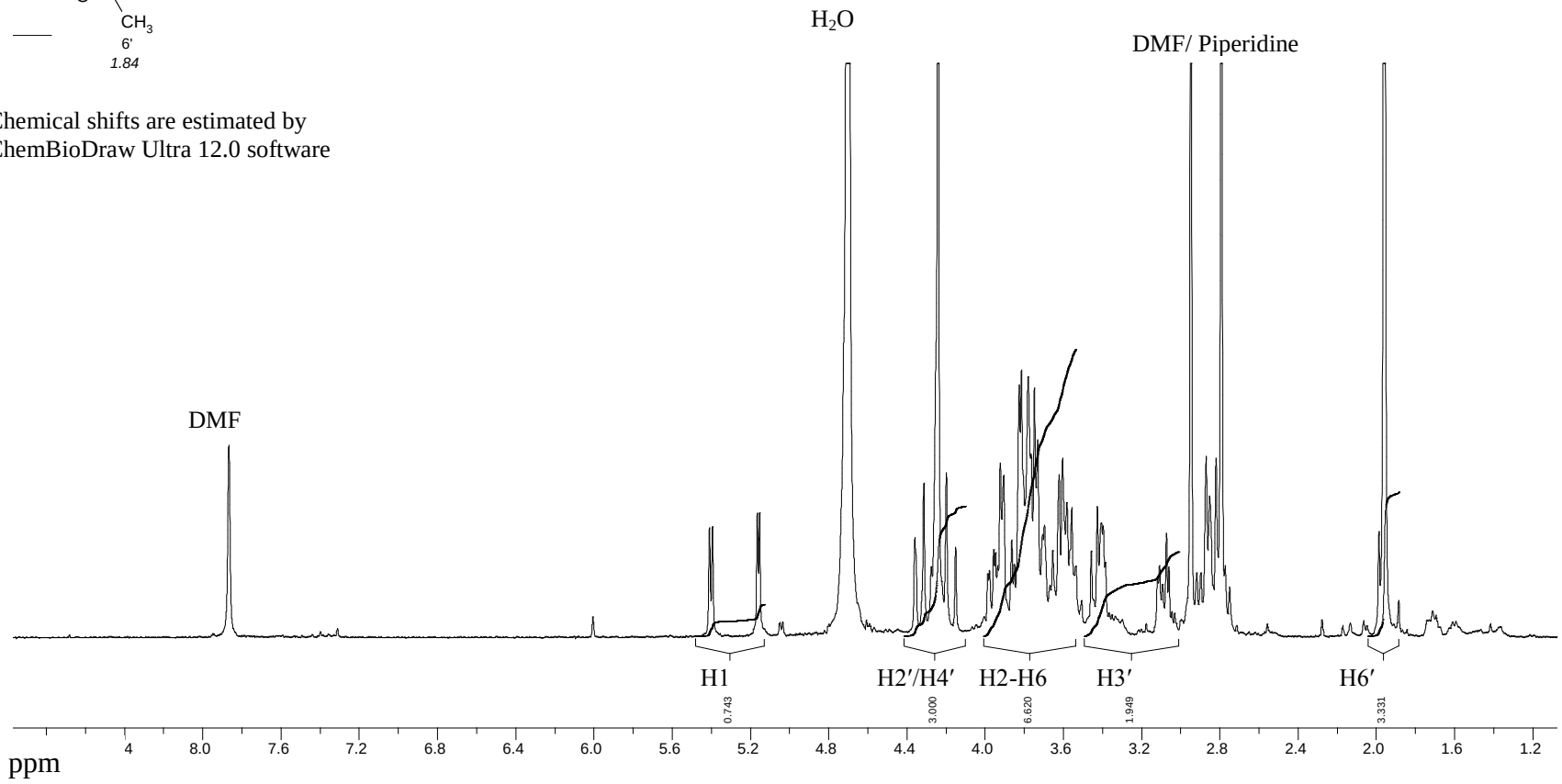
Spectrum A7.10: ^{13}C NMR of Fmoc-Cys(Acm)-GlcN in deuterated DMSO/D_2O



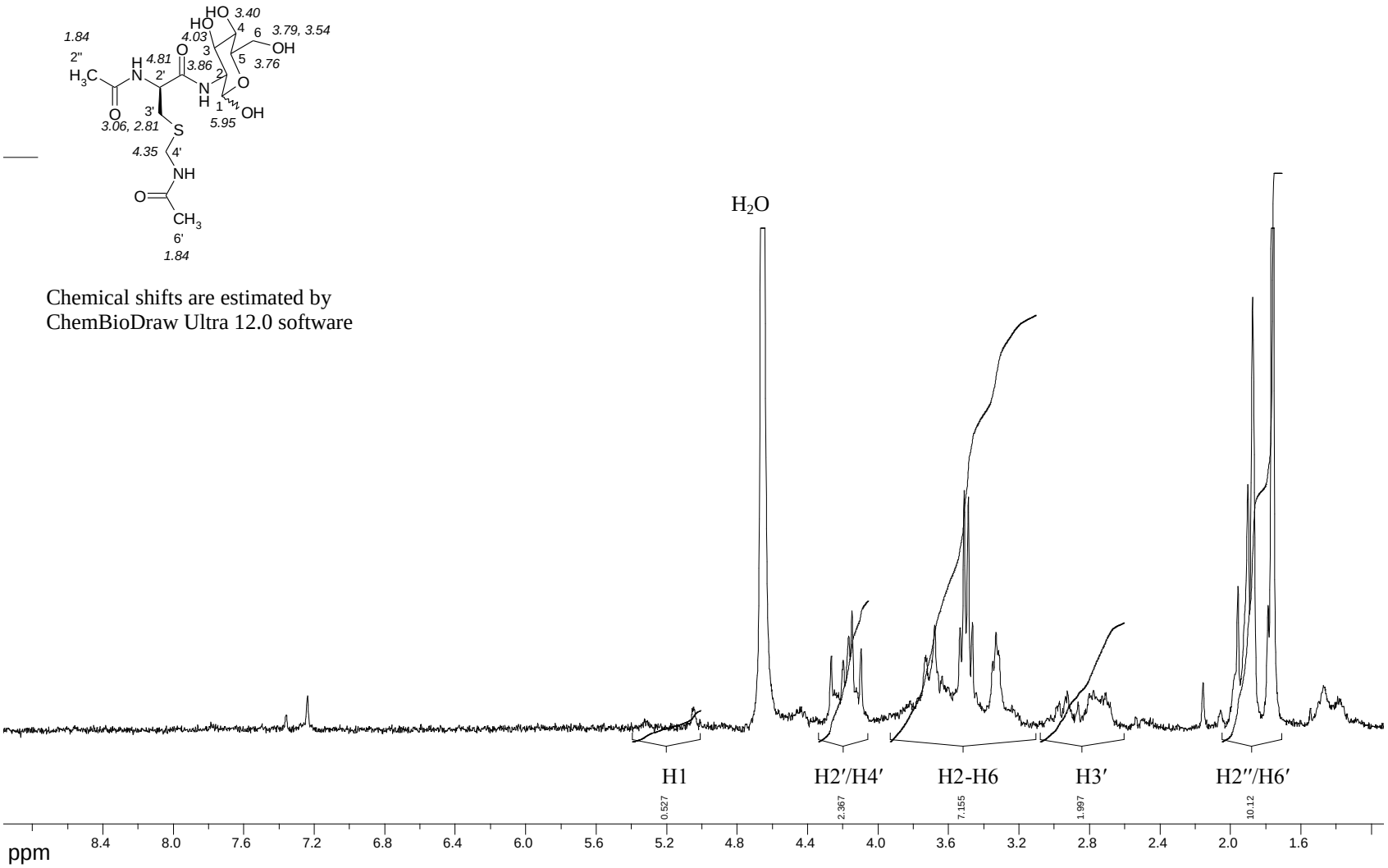
Spectrum A7.11: ^{1}H NMR of NH_2-Cys(Acm)-GlcN in D_2O



Chemical shifts are estimated by ChemBioDraw Ultra 12.0 software

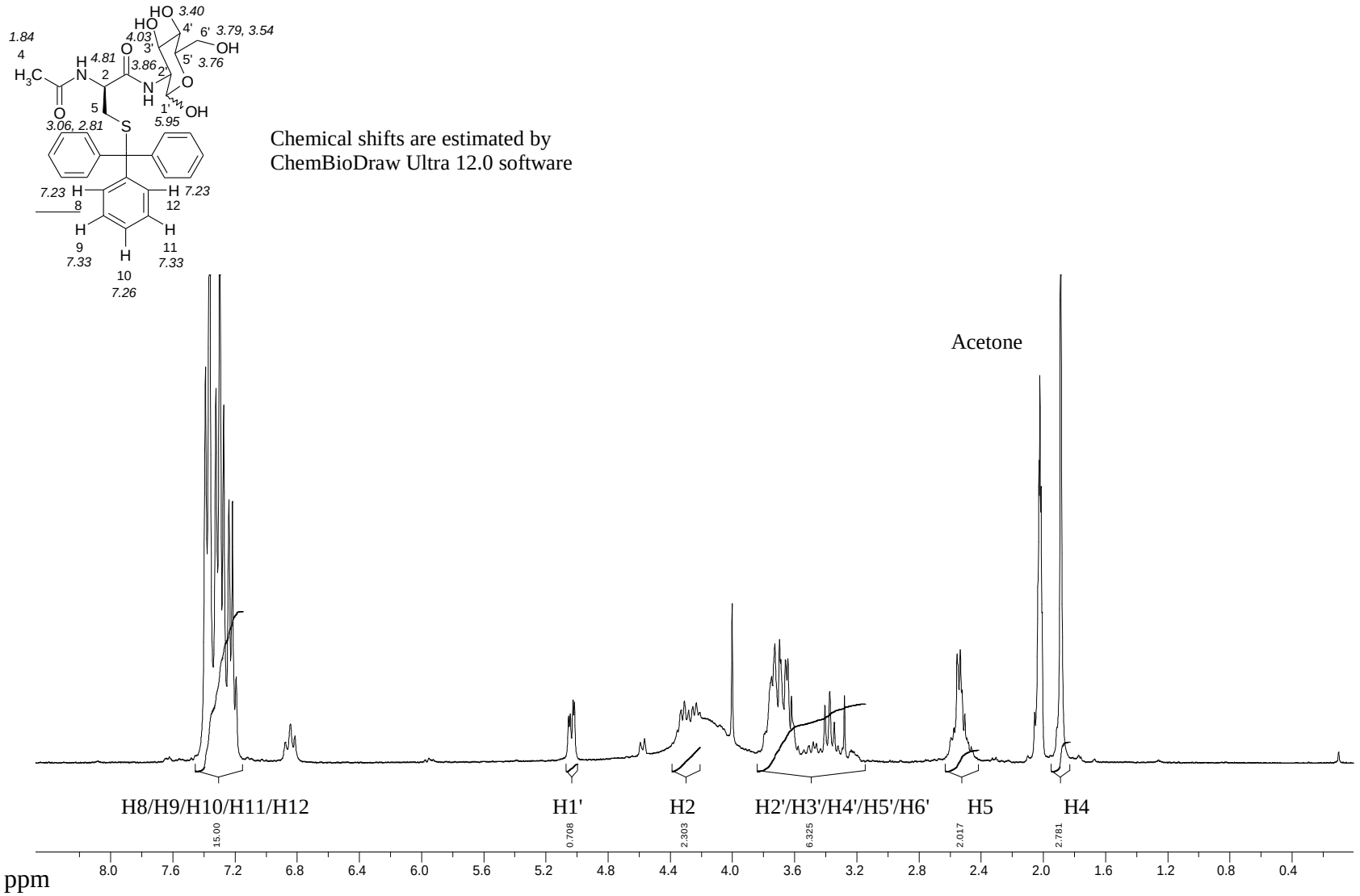


Spectrum A7.12: ^{1}H NMR of Ac-Cys(Acm)-GlcN in D_2O

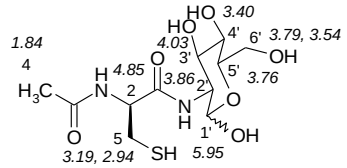


Chemical shifts are estimated by ChemBioDraw Ultra 12.0 software

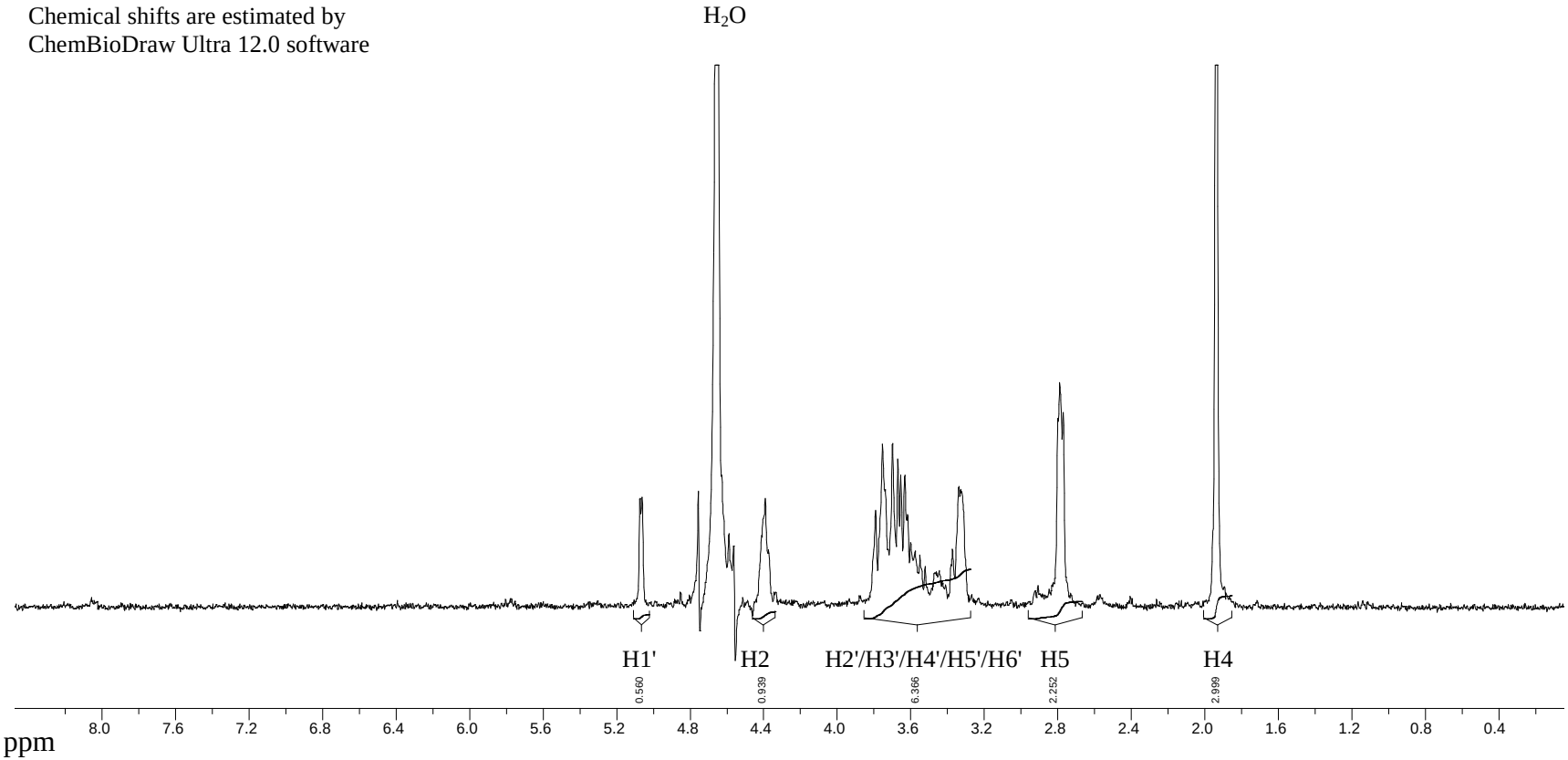
Spectrum A7.13: ^{1}H NMR of Ac-Cys(Trt)-GlcN in acetone-d6



Spectrum A7.14: ^{1}H NMR of free thiol of AcCys-GlcN in CD_3OD



Chemical shifts are estimated by ChemBioDraw Ultra 12.0 software



Appendix 8: The X-ray Crystallographic Protein Structure File for Unpublished Structures of Inhibitor Bound *Escherichia coli* Glyoxalase I

The unpublished X-ray crystallographic structures of *E. coli* GlxI with bound hydroxamate inhibitor (TSI) with the resolution of 2.1 Å and bound *N*-hydroxybenzamide (BHO) with the resolution of 1.6 Å were prepared by Dr. Susan Clugston (the previous Ph. D. student in our laboratory) as well as Dr. Molly He and Dr. Brian Matthews from University of Oregon (Eugene, Oregon).

The PBD file for N-hydroxybenzamide bound E. coli GlxI

```
REMARK   MSI WebLab Viewer PDB file
REMARK   Created:  Fri Mar 02 12:23:41 Eastern Standard Time 2001
CRYST1   49.000   49.000  182.580  90.00  90.00 120.00 P1
ATOM      1  C1  BHO A 200       4.453  38.499  32.229  1.00  0.00
ATOM      2  C2  BHO A 200       5.125  38.873  31.034  1.00  0.00
ATOM      3  C3  BHO A 200       5.776  38.145  30.206  1.00  0.00
ATOM      4  C4  BHO A 200       6.041  36.712  30.234  1.00  0.00
ATOM      5  C5  BHO A 200       5.328  36.221  31.465  1.00  0.00
ATOM      6  C6  BHO A 200       4.673  37.078  32.261  1.00  0.00
ATOM      7  C   BHO A 200       3.565  39.379  32.970  1.00  0.00
ATOM      8  O1  BHO A 200       3.026  38.717  33.904  1.00  0.00
ATOM      9  N   BHO A 200       3.586  40.740  32.996  1.00  0.00
ATOM     10  O2  BHO A 200       2.973  41.292  34.120  1.00  0.00
ATOM     11  CB  MET A   1      -9.340  30.362  38.222  1.00  0.00
ATOM     12  CG  MET A   1     -10.212  30.333  36.976  1.00  0.00
ATOM     13  SD  MET A   1     -11.698  29.337  37.190  1.00  0.00
ATOM     14  CE  MET A   1     -12.565  30.285  38.438  1.00  0.00
ATOM     15  C   MET A   1      -7.439  29.342  39.502  1.00  0.00
ATOM     16  O   MET A   1      -7.643  28.855  40.622  1.00  0.00
ATOM     17  N   MET A   1      -9.311  27.957  38.708  1.00  0.00
ATOM     18  CA  MET A   1      -8.477  29.114  38.410  1.00  0.00
ATOM     19  N   ARG A   2      -6.535  30.279  39.435  1.00  0.00
ATOM     20  CA  ARG A   2      -5.549  30.408  40.498  1.00  0.00
ATOM     21  CB  ARG A   2      -4.474  29.327  40.434  1.00  0.00
ATOM     22  CG  ARG A   2      -3.586  29.394  39.205  1.00  0.00
ATOM     23  CD  ARG A   2      -2.574  28.264  39.208  1.00  0.00
ATOM     24  NE  ARG A   2      -1.657  28.342  38.076  1.00  0.00
ATOM     25  CZ  ARG A   2      -0.336  28.399  38.194  1.00  0.00
ATOM     26  NH1 ARG A   2       0.223  28.386  39.396  1.00  0.00
ATOM     27  NH2 ARG A   2       0.424  28.467  37.109  1.00  0.00
ATOM     28  C   ARG A   2      -4.891  31.779  40.458  1.00  0.00
ATOM     29  O   ARG A   2      -4.880  32.423  39.409  1.00  0.00
ATOM     30  N   LEU A   3      -4.373  32.204  41.598  1.00  0.00
ATOM     31  CA  LEU A   3      -3.684  33.487  41.706  1.00  0.00
ATOM     32  CB  LEU A   3      -3.678  33.948  43.167  1.00  0.00
ATOM     33  CG  LEU A   3      -3.379  35.413  43.491  1.00  0.00
ATOM     34  CD1 LEU A   3      -3.728  35.684  44.948  1.00  0.00
ATOM     35  CD2 LEU A   3      -1.933  35.740  43.226  1.00  0.00
ATOM     36  C   LEU A   3      -2.265  33.211  41.211  1.00  0.00
ATOM     37  O   LEU A   3      -1.560  32.375  41.775  1.00  0.00
ATOM     38  N   LEU A   4      -1.849  33.912  40.158  1.00  0.00
ATOM     39  CA  LEU A   4      -0.528  33.707  39.558  1.00  0.00
ATOM     40  CB  LEU A   4      -0.615  33.979  38.049  1.00  0.00
```

```
ATOM     41  CG  LEU A   4      -1.592  33.102  37.259  1.00  0.00
ATOM     42  CD1 LEU A   4      -1.818  33.679  35.872  1.00  0.00
ATOM     43  CD2 LEU A   4      -1.037  31.691  37.169  1.00  0.00
ATOM     44  C   LEU A   4       0.607  34.529  40.165  1.00  0.00
ATOM     45  O   LEU A   4       1.662  33.992  40.527  1.00  0.00
ATOM     46  N   HIS A   5       0.411  35.838  40.260  1.00  0.00
ATOM     47  CA  HIS A   5       1.445  36.683  40.833  1.00  0.00
ATOM     48  CB  HIS A   5       2.655  36.789  39.884  1.00  0.00
ATOM     49  CG  HIS A   5       2.341  37.385  38.541  1.00  0.00
ATOM     50  CD2 HIS A   5       2.067  38.662  38.173  1.00  0.00
ATOM     51  ND1 HIS A   5       2.329  36.643  37.380  1.00  0.00
ATOM     52  CE1 HIS A   5       2.065  37.435  36.355  1.00  0.00
ATOM     53  NE2 HIS A   5       1.902  38.664  36.808  1.00  0.00
ATOM     54  C   HIS A   5       0.952  38.073  41.178  1.00  0.00
ATOM     55  O   HIS A   5      -0.180  38.447  40.865  1.00  0.00
ATOM     56  N   THR A   6       1.818  38.818  41.855  1.00  0.00
ATOM     57  CA  THR A   6       1.555  40.195  42.243  1.00  0.00
ATOM     58  CB  THR A   6       1.510  40.354  43.783  1.00  0.00
ATOM     59  OG1 THR A   6       0.449  39.547  44.318  1.00  0.00
ATOM     60  CG2 THR A   6       1.272  41.809  44.165  1.00  0.00
ATOM     61  C   THR A   6       2.746  40.956  41.672  1.00  0.00
ATOM     62  O   THR A   6       3.891  40.610  41.953  1.00  0.00
ATOM     63  N   MET A   7       2.469  41.969  40.856  1.00  0.00
ATOM     64  CA  MET A   7       3.520  42.761  40.224  1.00  0.00
ATOM     65  CB  MET A   7       3.123  43.075  38.775  1.00  0.00
ATOM     66  CG  MET A   7       4.213  43.752  37.927  1.00  0.00
ATOM     67  SD  MET A   7       4.461  45.525  38.248  1.00  0.00
ATOM     68  CE  MET A   7       3.575  46.284  36.872  1.00  0.00
ATOM     69  C   MET A   7       3.791  44.060  40.975  1.00  0.00
ATOM     70  O   MET A   7       2.866  44.778  41.341  1.00  0.00
ATOM     71  N   LEU A   8       5.072  44.341  41.207  1.00  0.00
ATOM     72  CA  LEU A   8       5.495  45.560  41.891  1.00  0.00
ATOM     73  CB  LEU A   8       6.103  45.226  43.258  1.00  0.00
ATOM     74  CG  LEU A   8       5.208  44.522  44.279  1.00  0.00
ATOM     75  CD1 LEU A   8       6.005  44.234  45.552  1.00  0.00
ATOM     76  CD2 LEU A   8       4.006  45.397  44.593  1.00  0.00
ATOM     77  C   LEU A   8       6.557  46.242  41.037  1.00  0.00
ATOM     78  O   LEU A   8       7.434  45.576  40.491  1.00  0.00
ATOM     79  N   ARG A   9       6.482  47.563  40.914  1.00  0.00
ATOM     80  CA  ARG A   9       7.482  48.283  40.133  1.00  0.00
ATOM     81  CB  ARG A   9       6.882  49.567  39.551  1.00  0.00
ATOM     82  CG  ARG A   9       5.814  49.316  38.485  1.00  0.00
ATOM     83  CD  ARG A   9       5.218  50.623  37.932  1.00  0.00
ATOM     84  NE  ARG A   9       4.145  50.377  36.964  1.00  0.00
ATOM     85  CZ  ARG A   9       2.865  50.185  37.278  1.00  0.00
ATOM     86  NH1 ARG A   9       2.469  50.220  38.545  1.00  0.00
ATOM     87  NH2 ARG A   9       1.980  49.930  36.316  1.00  0.00
ATOM     88  C   ARG A   9       8.628  48.596  41.092  1.00  0.00
ATOM     89  O   ARG A   9       8.396  48.949  42.246  1.00  0.00
ATOM     90  N   VAL A  10       9.861  48.434  40.630  1.00  0.00
ATOM     91  CA  VAL A  10      11.006  48.699  41.488  1.00  0.00
ATOM     92  CB  VAL A  10      11.769  47.385  41.791  1.00  0.00
ATOM     93  CG1 VAL A  10      10.817  46.406  42.483  1.00  0.00
ATOM     94  CG2 VAL A  10      12.308  46.763  40.504  1.00  0.00
ATOM     95  C   VAL A  10      11.935  49.752  40.896  1.00  0.00
ATOM     96  O   VAL A  10      11.971  49.951  39.683  1.00  0.00
ATOM     97  N   GLY A  11      12.677  50.433  41.764  1.00  0.00
ATOM     98  CA  GLY A  11      13.573  51.486  41.300  1.00  0.00
ATOM     99  C   GLY A  11      14.899  50.998  40.758  1.00  0.00
ATOM    100  O   GLY A  11      15.565  51.698  39.987  1.00  0.00
ATOM    101  N   ASP A  12      15.288  49.796  41.164  1.00  0.00
ATOM    102  CA  ASP A  12      16.540  49.187  40.730  1.00  0.00
ATOM    103  CB  ASP A  12      17.672  49.595  41.678  1.00  0.00
```

```
ATOM    104  CG  ASP A  12      18.993  48.975  41.301  1.00  0.00
ATOM    105  OD1 ASP A  12      19.156  47.766  41.528  1.00  0.00
ATOM    106  OD2 ASP A  12      19.870  49.687  40.763  1.00  0.00
ATOM    107  C   ASP A  12      16.299  47.678  40.752  1.00  0.00
ATOM    108  O   ASP A  12      16.138  47.085  41.813  1.00  0.00
ATOM    109  N   LEU A  13      16.281  47.065  39.570  1.00  0.00
ATOM    110  CA  LEU A  13      15.990  45.642  39.451  1.00  0.00
ATOM    111  CB  LEU A  13      15.941  45.226  37.978  1.00  0.00
ATOM    112  CG  LEU A  13      15.350  43.831  37.734  1.00  0.00
ATOM    113  CD1 LEU A  13      13.939  43.756  38.315  1.00  0.00
ATOM    114  CD2 LEU A  13      15.332  43.530  36.249  1.00  0.00
ATOM    115  C   LEU A  13      16.871  44.664  40.210  1.00  0.00
ATOM    116  O   LEU A  13      16.353  43.829  40.953  1.00  0.00
ATOM    117  N   GLN A  14      18.187  44.741  40.030  1.00  0.00
ATOM    118  CA  GLN A  14      19.078  43.807  40.708  1.00  0.00
ATOM    119  CB  GLN A  14      20.513  43.953  40.201  1.00  0.00
ATOM    120  CG  GLN A  14      21.370  42.750  40.535  1.00  0.00
ATOM    121  CD  GLN A  14      20.889  41.492  39.821  1.00  0.00
ATOM    122  OE1 GLN A  14      20.846  41.445  38.593  1.00  0.00
ATOM    123  NE2 GLN A  14      20.525  40.470  40.590  1.00  0.00
ATOM    124  C   GLN A  14      19.064  43.969  42.223  1.00  0.00
ATOM    125  O   GLN A  14      19.235  43.001  42.955  1.00  0.00
ATOM    126  N   ARG A  15      18.877  45.198  42.689  1.00  0.00
ATOM    127  CA  ARG A  15      18.813  45.453  44.121  1.00  0.00
ATOM    128  CB  ARG A  15      18.597  46.949  44.394  1.00  0.00
ATOM    129  CG  ARG A  15      18.454  47.312  45.864  1.00  0.00
ATOM    130  CD  ARG A  15      18.288  48.817  46.071  1.00  0.00
ATOM    131  NE  ARG A  15      17.071  49.332  45.453  1.00  0.00
ATOM    132  CZ  ARG A  15      16.624  50.575  45.588  1.00  0.00
ATOM    133  NH1 ARG A  15      17.299  51.448  46.327  1.00  0.00
ATOM    134  NH2 ARG A  15      15.498  50.948  44.988  1.00  0.00
ATOM    135  C   ARG A  15      17.631  44.655  44.674  1.00  0.00
ATOM    136  O   ARG A  15      17.736  43.987  45.700  1.00  0.00
ATOM    137  N   SER A  16      16.508  44.709  43.969  1.00  0.00
ATOM    138  CA  SER A  16      15.315  43.997  44.419  1.00  0.00
ATOM    139  CB  SER A  16      14.094  44.469  43.618  1.00  0.00
ATOM    140  OG  SER A  16      13.820  45.829  43.918  1.00  0.00
ATOM    141  C   SER A  16      15.473  42.476  44.357  1.00  0.00
ATOM    142  O   SER A  16      15.126  41.773  45.310  1.00  0.00
ATOM    143  N   ILE A  17      15.997  41.967  43.246  1.00  0.00
ATOM    144  CA  ILE A  17      16.217  40.532  43.108  1.00  0.00
ATOM    145  CB  ILE A  17      16.893  40.203  41.761  1.00  0.00
ATOM    146  CG2 ILE A  17      17.267  38.732  41.715  1.00  0.00
ATOM    147  CG1 ILE A  17      15.966  40.600  40.613  1.00  0.00
ATOM    148  CD1 ILE A  17      16.585  40.503  39.235  1.00  0.00
ATOM    149  C   ILE A  17      17.117  40.036  44.242  1.00  0.00
ATOM    150  O   ILE A  17      16.799  39.053  44.911  1.00  0.00
ATOM    151  N   ASP A  18      18.238  40.721  44.450  1.00  0.00
ATOM    152  CA  ASP A  18      19.169  40.350  45.504  1.00  0.00
ATOM    153  CB  ASP A  18      20.344  41.340  45.563  1.00  0.00
ATOM    154  CG  ASP A  18      21.299  41.190  44.387  1.00  0.00
ATOM    155  OD1 ASP A  18      21.285  40.120  43.739  1.00  0.00
ATOM    156  OD2 ASP A  18      22.080  42.136  44.113  1.00  0.00
ATOM    157  C   ASP A  18      18.486  40.280  46.870  1.00  0.00
ATOM    158  O   ASP A  18      18.803  39.416  47.671  1.00  0.00
ATOM    159  N   PHE A  19      17.557  41.197  47.131  1.00  0.00
ATOM    160  CA  PHE A  19      16.839  41.230  48.399  1.00  0.00
ATOM    161  CB  PHE A  19      15.999  42.509  48.500  1.00  0.00
ATOM    162  CG  PHE A  19      15.124  42.571  49.721  1.00  0.00
ATOM    163  CD1 PHE A  19      15.660  42.854  50.973  1.00  0.00
ATOM    164  CD2 PHE A  19      13.753  42.345  49.618  1.00  0.00
ATOM    165  CE1 PHE A  19      14.845  42.914  52.103  1.00  0.00
ATOM    166  CE2 PHE A  19      12.937  42.404  50.739  1.00  0.00
```

```
ATOM    167  CZ  PHE A  19      13.484  42.690  51.986  1.00  0.00
ATOM    168  C   PHE A  19      15.919  40.024  48.570  1.00  0.00
ATOM    169  O   PHE A  19      16.021  39.282  49.547  1.00  0.00
ATOM    170  N   TYR A  20      15.000  39.843  47.626  1.00  0.00
ATOM    171  CA  TYR A  20      14.061  38.733  47.710  1.00  0.00
ATOM    172  CB  TYR A  20      13.033  38.828  46.575  1.00  0.00
ATOM    173  CG  TYR A  20      12.010  39.914  46.830  1.00  0.00
ATOM    174  CD1 TYR A  20      10.984  39.714  47.745  1.00  0.00
ATOM    175  CE1 TYR A  20      10.086  40.729  48.055  1.00  0.00
ATOM    176  CD2 TYR A  20      12.115  41.168  46.219  1.00  0.00
ATOM    177  CE2 TYR A  20      11.219  42.192  46.522  1.00  0.00
ATOM    178  CZ  TYR A  20      10.212  41.963  47.441  1.00  0.00
ATOM    179  OH  TYR A  20       9.325  42.957  47.760  1.00  0.00
ATOM    180  C   TYR A  20      14.746  37.375  47.707  1.00  0.00
ATOM    181  O   TYR A  20      14.276  36.444  48.361  1.00  0.00
ATOM    182  N   THR A  21      15.870  37.271  47.001  1.00  0.00
ATOM    183  CA  THR A  21      16.592  36.008  46.943  1.00  0.00
ATOM    184  CB  THR A  21      17.393  35.874  45.621  1.00  0.00
ATOM    185  OG1 THR A  21      18.323  36.956  45.497  1.00  0.00
ATOM    186  CG2 THR A  21      16.442  35.901  44.423  1.00  0.00
ATOM    187  C   THR A  21      17.532  35.744  48.124  1.00  0.00
ATOM    188  O   THR A  21      17.381  34.744  48.823  1.00  0.00
ATOM    189  N   LYS A  22      18.490  36.634  48.362  1.00  0.00
ATOM    190  CA  LYS A  22      19.440  36.421  49.454  1.00  0.00
ATOM    191  CB  LYS A  22      20.669  37.321  49.280  1.00  0.00
ATOM    192  CG  LYS A  22      21.524  37.000  48.048  1.00  0.00
ATOM    193  CD  LYS A  22      21.933  35.527  47.987  1.00  0.00
ATOM    194  CE  LYS A  22      22.932  35.265  46.871  1.00  0.00
ATOM    195  NZ  LYS A  22      24.354  35.582  47.239  1.00  0.00
ATOM    196  C   LYS A  22      18.874  36.619  50.852  1.00  0.00
ATOM    197  O   LYS A  22      19.170  35.841  51.757  1.00  0.00
ATOM    198  N   VAL A  23      18.056  37.650  51.031  1.00  0.00
ATOM    199  CA  VAL A  23      17.494  37.936  52.342  1.00  0.00
ATOM    200  CB  VAL A  23      17.175  39.443  52.489  1.00  0.00
ATOM    201  CG1 VAL A  23      16.742  39.739  53.916  1.00  0.00
ATOM    202  CG2 VAL A  23      18.415  40.273  52.145  1.00  0.00
ATOM    203  C   VAL A  23      16.245  37.122  52.666  1.00  0.00
ATOM    204  O   VAL A  23      16.147  36.555  53.752  1.00  0.00
ATOM    205  N   LEU A  24      15.291  37.057  51.740  1.00  0.00
ATOM    206  CA  LEU A  24      14.075  36.295  51.999  1.00  0.00
ATOM    207  CB  LEU A  24      12.847  37.062  51.499  1.00  0.00
ATOM    208  CG  LEU A  24      12.548  38.344  52.288  1.00  0.00
ATOM    209  CD1 LEU A  24      11.272  38.974  51.763  1.00  0.00
ATOM    210  CD2 LEU A  24      12.406  38.020  53.768  1.00  0.00
ATOM    211  C   LEU A  24      14.067  34.856  51.469  1.00  0.00
ATOM    212  O   LEU A  24      13.072  34.139  51.608  1.00  0.00
ATOM    213  N   GLY A  25      15.177  34.436  50.864  1.00  0.00
ATOM    214  CA  GLY A  25      15.292  33.065  50.390  1.00  0.00
ATOM    215  C   GLY A  25      14.501  32.590  49.182  1.00  0.00
ATOM    216  O   GLY A  25      14.291  31.384  49.023  1.00  0.00
ATOM    217  N   MET A  26      14.049  33.512  48.340  1.00  0.00
ATOM    218  CA  MET A  26      13.318  33.133  47.139  1.00  0.00
ATOM    219  CB  MET A  26      12.436  34.283  46.647  1.00  0.00
ATOM    220  CG  MET A  26      11.248  34.597  47.523  1.00  0.00
ATOM    221  SD  MET A  26      10.209  35.892  46.793  1.00  0.00
ATOM    222  CE  MET A  26       8.852  35.982  48.031  1.00  0.00
ATOM    223  C   MET A  26      14.365  32.842  46.075  1.00  0.00
ATOM    224  O   MET A  26      15.568  32.906  46.332  1.00  0.00
ATOM    225  N   LYS A  27      13.901  32.536  44.871  1.00  0.00
ATOM    226  CA  LYS A  27      14.808  32.279  43.767  1.00  0.00
ATOM    227  CB  LYS A  27      14.943  30.778  43.495  1.00  0.00
ATOM    228  CG  LYS A  27      15.824  30.061  44.508  1.00  0.00
ATOM    229  CD  LYS A  27      16.236  28.676  44.041  1.00  0.00
```

```
ATOM    230  CE  LYS A  27      15.058  27.713  44.013  1.00  0.00
ATOM    231  NZ  LYS A  27      15.478  26.332  43.641  1.00  0.00
ATOM    232  C   LYS A  27      14.316  32.974  42.521  1.00  0.00
ATOM    233  O   LYS A  27      13.118  33.216  42.362  1.00  0.00
ATOM    234  N   LEU A  28      15.255  33.307  41.644  1.00  0.00
ATOM    235  CA  LEU A  28      14.913  33.944  40.390  1.00  0.00
ATOM    236  CB  LEU A  28      16.168  34.509  39.729  1.00  0.00
ATOM    237  CG  LEU A  28      15.951  35.271  38.424  1.00  0.00
ATOM    238  CD1 LEU A  28      15.079  36.501  38.671  1.00  0.00
ATOM    239  CD2 LEU A  28      17.294  35.680  37.860  1.00  0.00
ATOM    240  C   LEU A  28      14.307  32.821  39.548  1.00  0.00
ATOM    241  O   LEU A  28      14.908  31.751  39.394  1.00  0.00
ATOM    242  N   LEU A  29      13.105  33.054  39.031  1.00  0.00
ATOM    243  CA  LEU A  29      12.404  32.047  38.243  1.00  0.00
ATOM    244  CB  LEU A  29      10.948  31.948  38.700  1.00  0.00
ATOM    245  CG  LEU A  29      10.667  31.726  40.188  1.00  0.00
ATOM    246  CD1 LEU A  29       9.158  31.743  40.394  1.00  0.00
ATOM    247  CD2 LEU A  29      11.252  30.396  40.675  1.00  0.00
ATOM    248  C   LEU A  29      12.440  32.309  36.743  1.00  0.00
ATOM    249  O   LEU A  29      12.451  31.367  35.942  1.00  0.00
ATOM    250  N   ARG A  30      12.439  33.580  36.361  1.00  0.00
ATOM    251  CA  ARG A  30      12.473  33.925  34.953  1.00  0.00
ATOM    252  CB  ARG A  30      11.195  33.463  34.253  1.00  0.00
ATOM    253  CG  ARG A  30      10.026  34.393  34.457  1.00  0.00
ATOM    254  CD  ARG A  30       9.117  34.391  33.244  1.00  0.00
ATOM    255  NE  ARG A  30       8.231  33.234  33.224  1.00  0.00
ATOM    256  CZ  ARG A  30       7.589  32.805  32.144  1.00  0.00
ATOM    257  NH1 ARG A  30       7.734  33.439  30.983  1.00  0.00
ATOM    258  NH2 ARG A  30       6.800  31.738  32.231  1.00  0.00
ATOM    259  C   ARG A  30      12.653  35.420  34.758  1.00  0.00
ATOM    260  O   ARG A  30      12.416  36.219  35.669  1.00  0.00
ATOM    261  N   THR A  31      13.059  35.799  33.560  1.00  0.00
ATOM    262  CA  THR A  31      13.298  37.201  33.253  1.00  0.00
ATOM    263  CB  THR A  31      14.750  37.603  33.510  1.00  0.00
ATOM    264  OG1 THR A  31      15.597  36.999  32.527  1.00  0.00
ATOM    265  CG2 THR A  31      15.184  37.165  34.901  1.00  0.00
ATOM    266  C   THR A  31      12.961  37.507  31.804  1.00  0.00
ATOM    267  O   THR A  31      12.961  36.602  30.966  1.00  0.00
ATOM    268  N   SER A  32      12.656  38.756  31.520  1.00  0.00
ATOM    269  CA  SER A  32      12.391  39.167  30.146  1.00  0.00
ATOM    270  CB  SER A  32      10.923  38.929  29.791  1.00  0.00
ATOM    271  OG  SER A  32      10.067  39.726  30.592  1.00  0.00
ATOM    272  C   SER A  32      12.761  40.630  29.927  1.00  0.00
ATOM    273  O   SER A  32      12.843  41.408  30.884  1.00  0.00
ATOM    274  N   GLU A  33      13.007  40.977  28.681  1.00  0.00
ATOM    275  CA  GLU A  33      13.351  42.346  28.321  1.00  0.00
ATOM    276  CB  GLU A  33      14.815  42.465  27.879  1.00  0.00
ATOM    277  CG  GLU A  33      15.780  42.830  28.991  1.00  0.00
ATOM    278  CD  GLU A  33      16.397  41.619  29.638  1.00  0.00
ATOM    279  OE1 GLU A  33      17.283  40.998  29.001  1.00  0.00
ATOM    280  OE2 GLU A  33      15.997  41.272  30.771  1.00  0.00
ATOM    281  C   GLU A  33      12.458  42.803  27.181  1.00  0.00
ATOM    282  O   GLU A  33      12.120  42.019  26.288  1.00  0.00
ATOM    283  N   ASN A  34      12.071  44.071  27.223  1.00  0.00
ATOM    284  CA  ASN A  34      11.223  44.647  26.195  1.00  0.00
ATOM    285  CB  ASN A  34       9.825  44.922  26.756  1.00  0.00
ATOM    286  CG  ASN A  34       8.834  45.360  25.690  1.00  0.00
ATOM    287  OD1 ASN A  34       9.048  46.353  24.992  1.00  0.00
ATOM    288  ND2 ASN A  34       7.734  44.623  25.568  1.00  0.00
ATOM    289  C   ASN A  34      11.889  45.946  25.755  1.00  0.00
ATOM    290  O   ASN A  34      11.742  46.986  26.404  1.00  0.00
ATOM    291  N   PRO A  35      12.675  45.884  24.663  1.00  0.00
ATOM    292  CD  PRO A  35      13.062  44.625  24.004  1.00  0.00
```

```
ATOM    293  CA  PRO A  35      13.404  47.014  24.075  1.00  0.00
ATOM    294  CB  PRO A  35      14.133  46.373  22.898  1.00  0.00
ATOM    295  CG  PRO A  35      14.379  44.991  23.368  1.00  0.00
ATOM    296  C   PRO A  35      12.482  48.133  23.620  1.00  0.00
ATOM    297  O   PRO A  35      12.809  49.313  23.752  1.00  0.00
ATOM    298  N   GLU A  36      11.335  47.751  23.078  1.00  0.00
ATOM    299  CA  GLU A  36      10.359  48.706  22.575  1.00  0.00
ATOM    300  CB  GLU A  36       9.195  47.964  21.930  1.00  0.00
ATOM    301  CG  GLU A  36       7.987  48.829  21.679  1.00  0.00
ATOM    302  CD  GLU A  36       7.126  48.323  20.543  1.00  0.00
ATOM    303  OE1 GLU A  36       6.708  47.141  20.582  1.00  0.00
ATOM    304  OE2 GLU A  36       6.864  49.115  19.612  1.00  0.00
ATOM    305  C   GLU A  36       9.834  49.659  23.640  1.00  0.00
ATOM    306  O   GLU A  36       9.962  50.881  23.496  1.00  0.00
ATOM    307  N   TYR A  37       9.252  49.108  24.704  1.00  0.00
ATOM    308  CA  TYR A  37       8.699  49.927  25.783  1.00  0.00
ATOM    309  CB  TYR A  37       7.424  49.291  26.337  1.00  0.00
ATOM    310  CG  TYR A  37       6.259  49.336  25.373  1.00  0.00
ATOM    311  CD1 TYR A  37       6.139  48.404  24.344  1.00  0.00
ATOM    312  CE1 TYR A  37       5.067  48.453  23.445  1.00  0.00
ATOM    313  CD2 TYR A  37       5.284  50.323  25.484  1.00  0.00
ATOM    314  CE2 TYR A  37       4.212  50.384  24.592  1.00  0.00
ATOM    315  CZ  TYR A  37       4.108  49.446  23.580  1.00  0.00
ATOM    316  OH  TYR A  37       3.020  49.490  22.733  1.00  0.00
ATOM    317  C   TYR A  37       9.665  50.230  26.928  1.00  0.00
ATOM    318  O   TYR A  37       9.297  50.868  27.919  1.00  0.00
ATOM    319  N   LYS A  38      10.897  49.763  26.781  1.00  0.00
ATOM    320  CA  LYS A  38      11.954  49.991  27.750  1.00  0.00
ATOM    321  CB  LYS A  38      12.392  51.459  27.676  1.00  0.00
ATOM    322  CG  LYS A  38      12.786  51.861  26.259  1.00  0.00
ATOM    323  CD  LYS A  38      13.491  53.206  26.209  1.00  0.00
ATOM    324  CE  LYS A  38      14.082  53.438  24.822  1.00  0.00
ATOM    325  NZ  LYS A  38      14.899  54.676  24.766  1.00  0.00
ATOM    326  C   LYS A  38      11.685  49.588  29.197  1.00  0.00
ATOM    327  O   LYS A  38      11.478  50.436  30.076  1.00  0.00
ATOM    328  N   TYR A  39      11.700  48.281  29.435  1.00  0.00
ATOM    329  CA  TYR A  39      11.526  47.743  30.778  1.00  0.00
ATOM    330  CB  TYR A  39      10.050  47.792  31.215  1.00  0.00
ATOM    331  CG  TYR A  39       9.084  46.896  30.469  1.00  0.00
ATOM    332  CD1 TYR A  39       8.950  45.547  30.800  1.00  0.00
ATOM    333  CE1 TYR A  39       8.031  44.727  30.140  1.00  0.00
ATOM    334  CD2 TYR A  39       8.276  47.408  29.454  1.00  0.00
ATOM    335  CE2 TYR A  39       7.353  46.599  28.785  1.00  0.00
ATOM    336  CZ  TYR A  39       7.236  45.260  29.135  1.00  0.00
ATOM    337  OH  TYR A  39       6.330  44.455  28.479  1.00  0.00
ATOM    338  C   TYR A  39      12.072  46.324  30.856  1.00  0.00
ATOM    339  O   TYR A  39      12.242  45.655  29.834  1.00  0.00
ATOM    340  N   SER A  40      12.370  45.886  32.073  1.00  0.00
ATOM    341  CA  SER A  40      12.895  44.554  32.314  1.00  0.00
ATOM    342  CB  SER A  40      14.346  44.612  32.784  1.00  0.00
ATOM    343  OG  SER A  40      15.191  45.172  31.796  1.00  0.00
ATOM    344  C   SER A  40      12.048  43.941  33.409  1.00  0.00
ATOM    345  O   SER A  40      11.683  44.622  34.365  1.00  0.00
ATOM    346  N   LEU A  41      11.751  42.655  33.268  1.00  0.00
ATOM    347  CA  LEU A  41      10.941  41.952  34.248  1.00  0.00
ATOM    348  CB  LEU A  41       9.719  41.331  33.569  1.00  0.00
ATOM    349  CG  LEU A  41       8.719  42.284  32.919  1.00  0.00
ATOM    350  CD1 LEU A  41       7.559  41.483  32.347  1.00  0.00
ATOM    351  CD2 LEU A  41       8.228  43.283  33.952  1.00  0.00
ATOM    352  C   LEU A  41      11.736  40.857  34.930  1.00  0.00
ATOM    353  O   LEU A  41      12.662  40.288  34.349  1.00  0.00
ATOM    354  N   ALA A  42      11.370  40.560  36.172  1.00  0.00
ATOM    355  CA  ALA A  42      12.033  39.499  36.914  1.00  0.00
```

```
ATOM    356  CB  ALA A  42      13.202  40.055  37.711  1.00  0.00
ATOM    357  C   ALA A  42      11.009  38.876  37.845  1.00  0.00
ATOM    358  O   ALA A  42      10.318  39.582  38.578  1.00  0.00
ATOM    359  N   PHE A  43      10.889  37.556  37.798  1.00  0.00
ATOM    360  CA  PHE A  43       9.942  36.858  38.655  1.00  0.00
ATOM    361  CB  PHE A  43       9.103  35.863  37.839  1.00  0.00
ATOM    362  CG  PHE A  43       8.188  36.511  36.839  1.00  0.00
ATOM    363  CD1 PHE A  43       8.692  37.075  35.669  1.00  0.00
ATOM    364  CD2 PHE A  43       6.815  36.560  37.067  1.00  0.00
ATOM    365  CE1 PHE A  43       7.841  37.678  34.742  1.00  0.00
ATOM    366  CE2 PHE A  43       5.961  37.162  36.146  1.00  0.00
ATOM    367  CZ  PHE A  43       6.475  37.722  34.981  1.00  0.00
ATOM    368  C   PHE A  43      10.708  36.107  39.730  1.00  0.00
ATOM    369  O   PHE A  43      11.635  35.369  39.424  1.00  0.00
ATOM    370  N   VAL A  44      10.339  36.324  40.990  1.00  0.00
ATOM    371  CA  VAL A  44      10.978  35.640  42.110  1.00  0.00
ATOM    372  CB  VAL A  44      11.751  36.627  43.028  1.00  0.00
ATOM    373  CG1 VAL A  44      13.014  37.092  42.331  1.00  0.00
ATOM    374  CG2 VAL A  44      10.877  37.829  43.385  1.00  0.00
ATOM    375  C   VAL A  44       9.935  34.890  42.944  1.00  0.00
ATOM    376  O   VAL A  44       8.778  35.314  43.042  1.00  0.00
ATOM    377  N   GLY A  45      10.342  33.772  43.535  1.00  0.00
ATOM    378  CA  GLY A  45       9.411  32.995  44.331  1.00  0.00
ATOM    379  C   GLY A  45      10.017  31.794  45.029  1.00  0.00
ATOM    380  O   GLY A  45      11.195  31.487  44.855  1.00  0.00
ATOM    381  N   TYR A  46       9.189  31.106  45.807  1.00  0.00
ATOM    382  CA  TYR A  46       9.617  29.944  46.576  1.00  0.00
ATOM    383  CB  TYR A  46       8.920  29.953  47.937  1.00  0.00
ATOM    384  CG  TYR A  46       9.342  31.080  48.852  1.00  0.00
ATOM    385  CD1 TYR A  46      10.646  31.163  49.346  1.00  0.00
ATOM    386  CE1 TYR A  46      11.022  32.188  50.220  1.00  0.00
ATOM    387  CD2 TYR A  46       8.425  32.053  49.252  1.00  0.00
ATOM    388  CE2 TYR A  46       8.789  33.072  50.122  1.00  0.00
ATOM    389  CZ  TYR A  46      10.081  33.137  50.603  1.00  0.00
ATOM    390  OH  TYR A  46      10.430  34.146  51.474  1.00  0.00
ATOM    391  C   TYR A  46       9.348  28.616  45.873  1.00  0.00
ATOM    392  O   TYR A  46       9.614  27.545  46.433  1.00  0.00
ATOM    393  N   GLY A  47       8.827  28.688  44.653  1.00  0.00
ATOM    394  CA  GLY A  47       8.535  27.491  43.889  1.00  0.00
ATOM    395  C   GLY A  47       8.230  27.857  42.454  1.00  0.00
ATOM    396  O   GLY A  47       8.091  29.038  42.139  1.00  0.00
ATOM    397  N   PRO A  48       8.100  26.877  41.554  1.00  0.00
ATOM    398  CD  PRO A  48       8.224  25.427  41.770  1.00  0.00
ATOM    399  CA  PRO A  48       7.806  27.174  40.147  1.00  0.00
ATOM    400  CB  PRO A  48       7.750  25.784  39.503  1.00  0.00
ATOM    401  CG  PRO A  48       8.628  24.958  40.405  1.00  0.00
ATOM    402  C   PRO A  48       6.492  27.940  39.960  1.00  0.00
ATOM    403  O   PRO A  48       5.537  27.753  40.707  1.00  0.00
ATOM    404  N   GLU A  49       6.453  28.797  38.946  1.00  0.00
ATOM    405  CA  GLU A  49       5.256  29.584  38.641  1.00  0.00
ATOM    406  CB  GLU A  49       5.474  30.438  37.401  1.00  0.00
ATOM    407  CG  GLU A  49       6.520  31.503  37.500  1.00  0.00
ATOM    408  CD  GLU A  49       6.449  32.417  36.309  1.00  0.00
ATOM    409  OE1 GLU A  49       5.727  33.436  36.381  1.00  0.00
ATOM    410  OE2 GLU A  49       7.087  32.100  35.285  1.00  0.00
ATOM    411  C   GLU A  49       4.028  28.724  38.365  1.00  0.00
ATOM    412  O   GLU A  49       2.888  29.164  38.537  1.00  0.00
ATOM    413  N   THR A  50       4.263  27.506  37.896  1.00  0.00
ATOM    414  CA  THR A  50       3.161  26.606  37.572  1.00  0.00
ATOM    415  CB  THR A  50       3.649  25.467  36.662  1.00  0.00
ATOM    416  OG1 THR A  50       4.670  24.738  37.349  1.00  0.00
ATOM    417  CG2 THR A  50       4.213  26.020  35.356  1.00  0.00
ATOM    418  C   THR A  50       2.508  25.975  38.808  1.00  0.00
```

```
ATOM    419  O   THR A  50       1.392  25.449  38.727  1.00  0.00
ATOM    420  N   GLU A  51       3.195  26.035  39.943  1.00  0.00
ATOM    421  CA  GLU A  51       2.689  25.444  41.178  1.00  0.00
ATOM    422  CB  GLU A  51       3.723  24.478  41.734  1.00  0.00
ATOM    423  CG  GLU A  51       4.033  23.334  40.794  1.00  0.00
ATOM    424  CD  GLU A  51       5.251  22.549  41.218  1.00  0.00
ATOM    425  OE1 GLU A  51       5.772  21.761  40.394  1.00  0.00
ATOM    426  OE2 GLU A  51       5.692  22.722  42.373  1.00  0.00
ATOM    427  C   GLU A  51       2.301  26.443  42.259  1.00  0.00
ATOM    428  O   GLU A  51       1.565  26.108  43.184  1.00  0.00
ATOM    429  N   GLU A  52       2.793  27.671  42.163  1.00  0.00
ATOM    430  CA  GLU A  52       2.428  28.644  43.173  1.00  0.00
ATOM    431  CB  GLU A  52       3.226  28.391  44.448  1.00  0.00
ATOM    432  CG  GLU A  52       4.733  28.379  44.305  1.00  0.00
ATOM    433  CD  GLU A  52       5.408  27.890  45.576  1.00  0.00
ATOM    434  OE1 GLU A  52       5.282  26.687  45.900  1.00  0.00
ATOM    435  OE2 GLU A  52       6.042  28.708  46.273  1.00  0.00
ATOM    436  C   GLU A  52       2.517  30.104  42.758  1.00  0.00
ATOM    437  O   GLU A  52       3.004  30.432  41.679  1.00  0.00
ATOM    438  N   ALA A  53       2.003  30.971  43.624  1.00  0.00
ATOM    439  CA  ALA A  53       1.998  32.407  43.385  1.00  0.00
ATOM    440  CB  ALA A  53       1.008  33.080  44.326  1.00  0.00
ATOM    441  C   ALA A  53       3.394  32.987  43.587  1.00  0.00
ATOM    442  O   ALA A  53       4.087  32.621  44.530  1.00  0.00
ATOM    443  N   VAL A  54       3.794  33.891  42.697  1.00  0.00
ATOM    444  CA  VAL A  54       5.112  34.510  42.774  1.00  0.00
ATOM    445  CB  VAL A  54       6.010  33.999  41.629  1.00  0.00
ATOM    446  CG1 VAL A  54       6.055  32.480  41.650  1.00  0.00
ATOM    447  CG2 VAL A  54       5.479  34.490  40.286  1.00  0.00
ATOM    448  C   VAL A  54       5.039  36.037  42.706  1.00  0.00
ATOM    449  O   VAL A  54       3.951  36.618  42.641  1.00  0.00
ATOM    450  N   ILE A  55       6.202  36.682  42.727  1.00  0.00
ATOM    451  CA  ILE A  55       6.274  38.137  42.654  1.00  0.00
ATOM    452  CB  ILE A  55       7.124  38.726  43.800  1.00  0.00
ATOM    453  CG2 ILE A  55       7.243  40.235  43.625  1.00  0.00
ATOM    454  CG1 ILE A  55       6.512  38.364  45.154  1.00  0.00
ATOM    455  CD1 ILE A  55       7.277  38.925  46.351  1.00  0.00
ATOM    456  C   ILE A  55       6.926  38.553  41.342  1.00  0.00
ATOM    457  O   ILE A  55       7.976  38.024  40.984  1.00  0.00
ATOM    458  N   GLU A  56       6.287  39.482  40.635  1.00  0.00
ATOM    459  CA  GLU A  56       6.809  39.995  39.373  1.00  0.00
ATOM    460  CB  GLU A  56       5.699  40.104  38.316  1.00  0.00
ATOM    461  CG  GLU A  56       6.206  40.632  36.968  1.00  0.00
ATOM    462  CD  GLU A  56       5.095  40.954  35.980  1.00  0.00
ATOM    463  OE1 GLU A  56       3.925  40.630  36.265  1.00  0.00
ATOM    464  OE2 GLU A  56       5.399  41.534  34.911  1.00  0.00
ATOM    465  C   GLU A  56       7.369  41.383  39.665  1.00  0.00
ATOM    466  O   GLU A  56       6.652  42.259  40.160  1.00  0.00
ATOM    467  N   LEU A  57       8.654  41.568  39.383  1.00  0.00
ATOM    468  CA  LEU A  57       9.314  42.847  39.615  1.00  0.00
ATOM    469  CB  LEU A  57      10.685  42.622  40.268  1.00  0.00
ATOM    470  CG  LEU A  57      10.762  41.805  41.569  1.00  0.00
ATOM    471  CD1 LEU A  57      12.210  41.429  41.858  1.00  0.00
ATOM    472  CD2 LEU A  57      10.179  42.598  42.724  1.00  0.00
ATOM    473  C   LEU A  57       9.482  43.528  38.263  1.00  0.00
ATOM    474  O   LEU A  57       9.996  42.925  37.322  1.00  0.00
ATOM    475  N   THR A  58       9.049  44.783  38.167  1.00  0.00
ATOM    476  CA  THR A  58       9.156  45.521  36.915  1.00  0.00
ATOM    477  CB  THR A  58       7.778  45.983  36.410  1.00  0.00
ATOM    478  OG1 THR A  58       6.932  44.843  36.220  1.00  0.00
ATOM    479  CG2 THR A  58       7.920  46.727  35.089  1.00  0.00
ATOM    480  C   THR A  58      10.025  46.757  37.033  1.00  0.00
ATOM    481  O   THR A  58       9.744  47.653  37.822  1.00  0.00
```

```
ATOM    482  N   TYR A  59      11.077  46.807  36.233  1.00  0.00
ATOM    483  CA  TYR A  59      11.953  47.961  36.236  1.00  0.00
ATOM    484  CB  TYR A  59      13.406  47.540  36.462  1.00  0.00
ATOM    485  CG  TYR A  59      14.397  48.655  36.208  1.00  0.00
ATOM    486  CD1 TYR A  59      14.490  49.741  37.073  1.00  0.00
ATOM    487  CE1 TYR A  59      15.372  50.798  36.814  1.00  0.00
ATOM    488  CD2 TYR A  59      15.206  48.642  35.073  1.00  0.00
ATOM    489  CE2 TYR A  59      16.087  49.686  34.804  1.00  0.00
ATOM    490  CZ  TYR A  59      16.166  50.757  35.676  1.00  0.00
ATOM    491  OH  TYR A  59      17.043  51.789  35.418  1.00  0.00
ATOM    492  C   TYR A  59      11.826  48.678  34.899  1.00  0.00
ATOM    493  O   TYR A  59      12.166  48.125  33.854  1.00  0.00
ATOM    494  N   ASN A  60      11.307  49.899  34.947  1.00  0.00
ATOM    495  CA  ASN A  60      11.154  50.735  33.766  1.00  0.00
ATOM    496  CB  ASN A  60       9.966  51.682  33.946  1.00  0.00
ATOM    497  CG  ASN A  60       8.634  50.968  33.803  1.00  0.00
ATOM    498  OD1 ASN A  60       8.308  50.469  32.728  1.00  0.00
ATOM    499  ND2 ASN A  60       7.865  50.905  34.882  1.00  0.00
ATOM    500  C   ASN A  60      12.463  51.505  33.639  1.00  0.00
ATOM    501  O   ASN A  60      12.861  52.231  34.550  1.00  0.00
ATOM    502  N   TRP A  61      13.136  51.325  32.508  1.00  0.00
ATOM    503  CA  TRP A  61      14.431  51.952  32.262  1.00  0.00
ATOM    504  CB  TRP A  61      14.886  51.655  30.835  1.00  0.00
ATOM    505  CG  TRP A  61      14.968  50.186  30.529  1.00  0.00
ATOM    506  CD2 TRP A  61      15.293  49.585  29.270  1.00  0.00
ATOM    507  CE2 TRP A  61      15.210  48.187  29.441  1.00  0.00
ATOM    508  CE3 TRP A  61      15.646  50.092  28.013  1.00  0.00
ATOM    509  CD1 TRP A  61      14.712  49.156  31.386  1.00  0.00
ATOM    510  NE1 TRP A  61      14.856  47.952  30.743  1.00  0.00
ATOM    511  CZ2 TRP A  61      15.466  47.287  28.404  1.00  0.00
ATOM    512  CZ3 TRP A  61      15.901  49.195  26.980  1.00  0.00
ATOM    513  CH2 TRP A  61      15.809  47.807  27.186  1.00  0.00
ATOM    514  C   TRP A  61      14.456  53.455  32.512  1.00  0.00
ATOM    515  O   TRP A  61      13.680  54.203  31.932  1.00  0.00
ATOM    516  N   GLY A  62      15.358  53.877  33.392  1.00  0.00
ATOM    517  CA  GLY A  62      15.485  55.288  33.704  1.00  0.00
ATOM    518  C   GLY A  62      14.550  55.781  34.791  1.00  0.00
ATOM    519  O   GLY A  62      14.657  56.925  35.228  1.00  0.00
ATOM    520  N   VAL A  63      13.628  54.930  35.228  1.00  0.00
ATOM    521  CA  VAL A  63      12.691  55.319  36.275  1.00  0.00
ATOM    522  CB  VAL A  63      11.242  54.949  35.906  1.00  0.00
ATOM    523  CG1 VAL A  63      10.291  55.463  36.972  1.00  0.00
ATOM    524  CG2 VAL A  63      10.871  55.543  34.547  1.00  0.00
ATOM    525  C   VAL A  63      13.077  54.610  37.566  1.00  0.00
ATOM    526  O   VAL A  63      12.911  53.400  37.699  1.00  0.00
ATOM    527  N   ASP A  64      13.583  55.372  38.524  1.00  0.00
ATOM    528  CA  ASP A  64      14.023  54.795  39.785  1.00  0.00
ATOM    529  CB  ASP A  64      15.489  55.156  40.016  1.00  0.00
ATOM    530  CG  ASP A  64      15.698  56.643  40.249  1.00  0.00
ATOM    531  OD1 ASP A  64      15.042  57.464  39.559  1.00  0.00
ATOM    532  OD2 ASP A  64      16.540  56.990  41.116  1.00  0.00
ATOM    533  C   ASP A  64      13.203  55.181  41.010  1.00  0.00
ATOM    534  O   ASP A  64      13.630  54.943  42.142  1.00  0.00
ATOM    535  N   LYS A  65      12.022  55.752  40.793  1.00  0.00
ATOM    536  CA  LYS A  65      11.179  56.163  41.902  1.00  0.00
ATOM    537  CB  LYS A  65      11.608  57.553  42.374  1.00  0.00
ATOM    538  CG  LYS A  65      11.594  57.733  43.882  1.00  0.00
ATOM    539  CD  LYS A  65      12.690  56.926  44.570  1.00  0.00
ATOM    540  CE  LYS A  65      12.617  57.136  46.080  1.00  0.00
ATOM    541  NZ  LYS A  65      13.761  56.554  46.841  1.00  0.00
ATOM    542  C   LYS A  65       9.703  56.185  41.508  1.00  0.00
ATOM    543  O   LYS A  65       9.371  56.491  40.359  1.00  0.00
ATOM    544  N   TYR A  66       8.834  55.862  42.468  1.00  0.00
```

```
ATOM    545  CA  TYR A  66       7.384  55.860  42.265  1.00  0.00
ATOM    546  CB  TYR A  66       6.864  54.451  41.940  1.00  0.00
ATOM    547  CG  TYR A  66       7.410  53.847  40.675  1.00  0.00
ATOM    548  CD1 TYR A  66       8.561  53.058  40.693  1.00  0.00
ATOM    549  CE1 TYR A  66       9.070  52.514  39.509  1.00  0.00
ATOM    550  CD2 TYR A  66       6.784  54.079  39.451  1.00  0.00
ATOM    551  CE2 TYR A  66       7.284  53.545  38.275  1.00  0.00
ATOM    552  CZ  TYR A  66       8.422  52.769  38.310  1.00  0.00
ATOM    553  OH  TYR A  66       8.922  52.257  37.137  1.00  0.00
ATOM    554  C   TYR A  66       6.691  56.328  43.542  1.00  0.00
ATOM    555  O   TYR A  66       7.314  56.423  44.601  1.00  0.00
ATOM    556  N   GLU A  67       5.396  56.605  43.443  1.00  0.00
ATOM    557  CA  GLU A  67       4.624  57.026  44.604  1.00  0.00
ATOM    558  CB  GLU A  67       3.656  58.138  44.208  1.00  0.00
ATOM    559  CG  GLU A  67       4.332  59.360  43.567  1.00  0.00
ATOM    560  CD  GLU A  67       3.330  60.396  43.061  1.00  0.00
ATOM    561  OE1 GLU A  67       3.751  61.384  42.413  1.00  0.00
ATOM    562  OE2 GLU A  67       2.118  60.216  43.303  1.00  0.00
ATOM    563  C   GLU A  67       3.864  55.780  45.061  1.00  0.00
ATOM    564  O   GLU A  67       3.008  55.273  44.344  1.00  0.00
ATOM    565  N   LEU A  68       4.195  55.293  46.252  1.00  0.00
ATOM    566  CA  LEU A  68       3.574  54.093  46.799  1.00  0.00
ATOM    567  CB  LEU A  68       4.235  53.731  48.130  1.00  0.00
ATOM    568  CG  LEU A  68       3.717  52.478  48.841  1.00  0.00
ATOM    569  CD1 LEU A  68       3.747  51.281  47.881  1.00  0.00
ATOM    570  CD2 LEU A  68       4.571  52.209  50.068  1.00  0.00
ATOM    571  C   LEU A  68       2.069  54.208  46.999  1.00  0.00
ATOM    572  O   LEU A  68       1.338  53.229  46.851  1.00  0.00
ATOM    573  N   GLY A  69       1.607  55.404  47.340  1.00  0.00
ATOM    574  CA  GLY A  69       0.186  55.589  47.560  1.00  0.00
ATOM    575  C   GLY A  69      -0.205  55.043  48.916  1.00  0.00
ATOM    576  O   GLY A  69       0.638  54.548  49.666  1.00  0.00
ATOM    577  N   THR A  70      -1.492  55.102  49.219  1.00  0.00
ATOM    578  CA  THR A  70      -1.989  54.619  50.496  1.00  0.00
ATOM    579  CB  THR A  70      -2.710  55.761  51.243  1.00  0.00
ATOM    580  OG1 THR A  70      -3.806  56.234  50.451  1.00  0.00
ATOM    581  CG2 THR A  70      -1.751  56.912  51.489  1.00  0.00
ATOM    582  C   THR A  70      -2.941  53.434  50.346  1.00  0.00
ATOM    583  O   THR A  70      -3.411  52.887  51.341  1.00  0.00
ATOM    584  N   ALA A  71      -3.202  53.031  49.108  1.00  0.00
ATOM    585  CA  ALA A  71      -4.122  51.931  48.812  1.00  0.00
ATOM    586  CB  ALA A  71      -4.613  52.049  47.376  1.00  0.00
ATOM    587  C   ALA A  71      -3.529  50.546  49.031  1.00  0.00
ATOM    588  O   ALA A  71      -4.076  49.735  49.784  1.00  0.00
ATOM    589  N   TYR A  72      -2.420  50.279  48.352  1.00  0.00
ATOM    590  CA  TYR A  72      -1.751  48.991  48.443  1.00  0.00
ATOM    591  CB  TYR A  72      -0.553  48.946  47.496  1.00  0.00
ATOM    592  CG  TYR A  72       0.258  47.683  47.623  1.00  0.00
ATOM    593  CD1 TYR A  72      -0.245  46.464  47.175  1.00  0.00
ATOM    594  CE1 TYR A  72       0.482  45.283  47.331  1.00  0.00
ATOM    595  CD2 TYR A  72       1.512  47.694  48.230  1.00  0.00
ATOM    596  CE2 TYR A  72       2.246  46.526  48.388  1.00  0.00
ATOM    597  CZ  TYR A  72       1.727  45.325  47.937  1.00  0.00
ATOM    598  OH  TYR A  72       2.474  44.180  48.084  1.00  0.00
ATOM    599  C   TYR A  72      -1.297  48.709  49.864  1.00  0.00
ATOM    600  O   TYR A  72      -0.707  49.570  50.520  1.00  0.00
ATOM    601  N   GLY A  73      -1.574  47.494  50.325  1.00  0.00
ATOM    602  CA  GLY A  73      -1.216  47.110  51.672  1.00  0.00
ATOM    603  C   GLY A  73       0.104  46.379  51.813  1.00  0.00
ATOM    604  O   GLY A  73       1.114  46.975  52.187  1.00  0.00
ATOM    605  N   HIS A  74       0.102  45.083  51.520  1.00  0.00
ATOM    606  CA  HIS A  74       1.316  44.287  51.659  1.00  0.00
ATOM    607  CB  HIS A  74       1.648  44.086  53.144  1.00  0.00
```

```
ATOM    608  CG  HIS A  74       0.548  43.437  53.931  1.00  0.00
ATOM    609  CD2 HIS A  74       0.199  42.137  54.077  1.00  0.00
ATOM    610  ND1 HIS A  74      -0.346  44.159  54.689  1.00  0.00
ATOM    611  CE1 HIS A  74      -1.197  43.333  55.272  1.00  0.00
ATOM    612  NE2 HIS A  74      -0.888  42.098  54.919  1.00  0.00
ATOM    613  C   HIS A  74       1.210  42.909  51.034  1.00  0.00
ATOM    614  O   HIS A  74       0.155  42.498  50.569  1.00  0.00
ATOM    615  N   ILE A  75       2.338  42.208  51.019  1.00  0.00
ATOM    616  CA  ILE A  75       2.381  40.829  50.554  1.00  0.00
ATOM    617  CB  ILE A  75       3.443  40.583  49.460  1.00  0.00
ATOM    618  CG2 ILE A  75       3.766  39.085  49.368  1.00  0.00
ATOM    619  CG1 ILE A  75       2.917  41.058  48.105  1.00  0.00
ATOM    620  CD1 ILE A  75       3.919  40.905  46.983  1.00  0.00
ATOM    621  C   ILE A  75       2.803  40.118  51.829  1.00  0.00
ATOM    622  O   ILE A  75       3.708  40.577  52.517  1.00  0.00
ATOM    623  N   ALA A  76       2.124  39.028  52.159  1.00  0.00
ATOM    624  CA  ALA A  76       2.451  38.278  53.366  1.00  0.00
ATOM    625  CB  ALA A  76       1.180  37.939  54.123  1.00  0.00
ATOM    626  C   ALA A  76       3.182  37.007  52.981  1.00  0.00
ATOM    627  O   ALA A  76       2.816  36.344  52.015  1.00  0.00
ATOM    628  N   LEU A  77       4.226  36.682  53.737  1.00  0.00
ATOM    629  CA  LEU A  77       5.018  35.485  53.485  1.00  0.00
ATOM    630  CB  LEU A  77       6.413  35.868  52.996  1.00  0.00
ATOM    631  CG  LEU A  77       6.587  36.824  51.822  1.00  0.00
ATOM    632  CD1 LEU A  77       8.060  37.146  51.705  1.00  0.00
ATOM    633  CD2 LEU A  77       6.057  36.218  50.526  1.00  0.00
ATOM    634  C   LEU A  77       5.167  34.716  54.785  1.00  0.00
ATOM    635  O   LEU A  77       5.230  35.317  55.858  1.00  0.00
ATOM    636  N   SER A  78       5.215  33.391  54.698  1.00  0.00
ATOM    637  CA  SER A  78       5.393  32.580  55.894  1.00  0.00
ATOM    638  CB  SER A  78       4.836  31.165  55.694  1.00  0.00
ATOM    639  OG  SER A  78       5.596  30.435  54.745  1.00  0.00
ATOM    640  C   SER A  78       6.893  32.518  56.159  1.00  0.00
ATOM    641  O   SER A  78       7.705  32.627  55.240  1.00  0.00
ATOM    642  N   VAL A  79       7.256  32.360  57.423  1.00  0.00
ATOM    643  CA  VAL A  79       8.657  32.282  57.806  1.00  0.00
ATOM    644  CB  VAL A  79       9.179  33.646  58.338  1.00  0.00
ATOM    645  CG1 VAL A  79       8.394  34.075  59.577  1.00  0.00
ATOM    646  CG2 VAL A  79      10.666  33.544  58.667  1.00  0.00
ATOM    647  C   VAL A  79       8.817  31.227  58.890  1.00  0.00
ATOM    648  O   VAL A  79       7.866  30.921  59.606  1.00  0.00
ATOM    649  N   ASP A  80      10.016  30.660  58.987  1.00  0.00
ATOM    650  CA  ASP A  80      10.314  29.645  59.993  1.00  0.00
ATOM    651  CB  ASP A  80      11.663  28.975  59.695  1.00  0.00
ATOM    652  CG  ASP A  80      11.663  28.185  58.394  1.00  0.00
ATOM    653  OD1 ASP A  80      10.701  27.425  58.156  1.00  0.00
ATOM    654  OD2 ASP A  80      12.636  28.313  57.616  1.00  0.00
ATOM    655  C   ASP A  80      10.384  30.293  61.372  1.00  0.00
ATOM    656  O   ASP A  80       9.819  29.794  62.351  1.00  0.00
ATOM    657  N   ASN A  81      11.085  31.421  61.434  1.00  0.00
ATOM    658  CA  ASN A  81      11.276  32.144  62.681  1.00  0.00
ATOM    659  CB  ASN A  81      12.644  31.781  63.260  1.00  0.00
ATOM    660  CG  ASN A  81      12.822  32.257  64.683  1.00  0.00
ATOM    661  OD1 ASN A  81      13.142  33.416  64.932  1.00  0.00
ATOM    662  ND2 ASN A  81      12.602  31.361  65.629  1.00  0.00
ATOM    663  C   ASN A  81      11.209  33.639  62.397  1.00  0.00
ATOM    664  O   ASN A  81      12.132  34.202  61.816  1.00  0.00
ATOM    665  N   ALA A  82      10.115  34.272  62.807  1.00  0.00
ATOM    666  CA  ALA A  82       9.920  35.699  62.575  1.00  0.00
ATOM    667  CB  ALA A  82       8.542  36.121  63.089  1.00  0.00
ATOM    668  C   ALA A  82      11.007  36.578  63.194  1.00  0.00
ATOM    669  O   ALA A  82      11.431  37.561  62.579  1.00  0.00
ATOM    670  N   ALA A  83      11.451  36.238  64.405  1.00  0.00
```

```
ATOM    671  CA  ALA A  83      12.487  37.020  65.083  1.00  0.00
ATOM    672  CB  ALA A  83      12.672  36.527  66.517  1.00  0.00
ATOM    673  C   ALA A  83      13.805  36.933  64.334  1.00  0.00
ATOM    674  O   ALA A  83      14.531  37.915  64.212  1.00  0.00
ATOM    675  N   GLU A  84      14.109  35.741  63.832  1.00  0.00
ATOM    676  CA  GLU A  84      15.339  35.516  63.099  1.00  0.00
ATOM    677  CB  GLU A  84      15.499  34.023  62.820  1.00  0.00
ATOM    678  CG  GLU A  84      16.934  33.555  62.671  1.00  0.00
ATOM    679  CD  GLU A  84      17.667  34.230  61.524  1.00  0.00
ATOM    680  OE1 GLU A  84      16.999  34.602  60.538  1.00  0.00
ATOM    681  OE2 GLU A  84      18.913  34.376  61.603  1.00  0.00
ATOM    682  C   GLU A  84      15.297  36.301  61.795  1.00  0.00
ATOM    683  O   GLU A  84      16.290  36.908  61.388  1.00  0.00
ATOM    684  N   ALA A  85      14.132  36.304  61.153  1.00  0.00
ATOM    685  CA  ALA A  85      13.951  37.020  59.896  1.00  0.00
ATOM    686  CB  ALA A  85      12.519  36.818  59.386  1.00  0.00
ATOM    687  C   ALA A  85      14.252  38.511  60.067  1.00  0.00
ATOM    688  O   ALA A  85      14.937  39.112  59.242  1.00  0.00
ATOM    689  N   CYS A  86      13.741  39.103  61.143  1.00  0.00
ATOM    690  CA  CYS A  86      13.971  40.520  61.397  1.00  0.00
ATOM    691  CB  CYS A  86      13.204  40.972  62.644  1.00  0.00
ATOM    692  SG  CYS A  86      11.388  40.916  62.469  1.00  0.00
ATOM    693  C   CYS A  86      15.452  40.851  61.548  1.00  0.00
ATOM    694  O   CYS A  86      15.893  41.919  61.124  1.00  0.00
ATOM    695  N   GLU A  87      16.216  39.949  62.163  1.00  0.00
ATOM    696  CA  GLU A  87      17.656  40.173  62.334  1.00  0.00
ATOM    697  CB  GLU A  87      18.253  39.145  63.296  1.00  0.00
ATOM    698  CG  GLU A  87      18.019  39.441  64.767  1.00  0.00
ATOM    699  CD  GLU A  87      18.822  40.633  65.256  1.00  0.00
ATOM    700  OE1 GLU A  87      20.068  40.577  65.220  1.00  0.00
ATOM    701  OE2 GLU A  87      18.212  41.628  65.692  1.00  0.00
ATOM    702  C   GLU A  87      18.382  40.098  60.992  1.00  0.00
ATOM    703  O   GLU A  87      19.320  40.852  60.743  1.00  0.00
ATOM    704  N   LYS A  88      17.947  39.189  60.122  1.00  0.00
ATOM    705  CA  LYS A  88      18.544  39.062  58.804  1.00  0.00
ATOM    706  CB  LYS A  88      18.054  37.784  58.125  1.00  0.00
ATOM    707  CG  LYS A  88      18.414  36.513  58.875  1.00  0.00
ATOM    708  CD  LYS A  88      18.105  35.276  58.050  1.00  0.00
ATOM    709  CE  LYS A  88      18.460  34.005  58.806  1.00  0.00
ATOM    710  NZ  LYS A  88      18.168  32.784  58.008  1.00  0.00
ATOM    711  C   LYS A  88      18.245  40.288  57.955  1.00  0.00
ATOM    712  O   LYS A  88      19.043  40.711  57.121  1.00  0.00
ATOM    713  N   ILE A  89      17.070  40.878  58.174  1.00  0.00
ATOM    714  CA  ILE A  89      16.706  42.084  57.445  1.00  0.00
ATOM    715  CB  ILE A  89      15.220  42.429  57.600  1.00  0.00
ATOM    716  CG2 ILE A  89      14.907  43.739  56.893  1.00  0.00
ATOM    717  CG1 ILE A  89      14.353  41.294  57.052  1.00  0.00
ATOM    718  CD1 ILE A  89      12.880  41.444  57.361  1.00  0.00
ATOM    719  C   ILE A  89      17.580  43.261  57.901  1.00  0.00
ATOM    720  O   ILE A  89      18.099  44.015  57.090  1.00  0.00
ATOM    721  N   ARG A  90      17.753  43.352  59.219  1.00  0.00
ATOM    722  CA  ARG A  90      18.567  44.399  59.831  1.00  0.00
ATOM    723  CB  ARG A  90      18.533  44.251  61.355  1.00  0.00
ATOM    724  CG  ARG A  90      18.785  45.522  62.141  1.00  0.00
ATOM    725  CD  ARG A  90      17.666  45.710  63.150  1.00  0.00
ATOM    726  NE  ARG A  90      17.434  44.514  63.953  1.00  0.00
ATOM    727  CZ  ARG A  90      16.239  44.155  64.420  1.00  0.00
ATOM    728  NH1 ARG A  90      15.177  44.904  64.160  1.00  0.00
ATOM    729  NH2 ARG A  90      16.096  43.053  65.141  1.00  0.00
ATOM    730  C   ARG A  90      20.015  44.339  59.345  1.00  0.00
ATOM    731  O   ARG A  90      20.585  45.357  58.939  1.00  0.00
ATOM    732  N   GLN A  91      20.606  43.150  59.381  1.00  0.00
ATOM    733  CA  GLN A  91      21.993  42.985  58.959  1.00  0.00
```

```
ATOM    734  CB  GLN A  91      22.469  41.555  59.235  1.00  0.00
ATOM    735  CG  GLN A  91      21.922  40.525  58.288  1.00  0.00
ATOM    736  CD  GLN A  91      22.310  39.107  58.679  1.00  0.00
ATOM    737  OE1 GLN A  91      21.868  38.588  59.710  1.00  0.00
ATOM    738  NE2 GLN A  91      23.153  38.477  57.862  1.00  0.00
ATOM    739  C   GLN A  91      22.180  43.313  57.483  1.00  0.00
ATOM    740  O   GLN A  91      23.260  43.711  57.058  1.00  0.00
ATOM    741  N   ASN A  92      21.127  43.160  56.694  1.00  0.00
ATOM    742  CA  ASN A  92      21.246  43.462  55.281  1.00  0.00
ATOM    743  CB  ASN A  92      20.422  42.464  54.465  1.00  0.00
ATOM    744  CG  ASN A  92      21.151  41.138  54.273  1.00  0.00
ATOM    745  OD1 ASN A  92      21.998  41.008  53.383  1.00  0.00
ATOM    746  ND2 ASN A  92      20.839  40.156  55.117  1.00  0.00
ATOM    747  C   ASN A  92      20.880  44.905  54.952  1.00  0.00
ATOM    748  O   ASN A  92      20.664  45.250  53.788  1.00  0.00
ATOM    749  N   GLY A  93      20.832  45.738  55.993  1.00  0.00
ATOM    750  CA  GLY A  93      20.544  47.158  55.847  1.00  0.00
ATOM    751  C   GLY A  93      19.091  47.560  55.696  1.00  0.00
ATOM    752  O   GLY A  93      18.797  48.695  55.310  1.00  0.00
ATOM    753  N   GLY A  94      18.180  46.651  56.028  1.00  0.00
ATOM    754  CA  GLY A  94      16.764  46.940  55.885  1.00  0.00
ATOM    755  C   GLY A  94      16.042  47.375  57.146  1.00  0.00
ATOM    756  O   GLY A  94      16.562  47.248  58.260  1.00  0.00
ATOM    757  N   ASN A  95      14.829  47.885  56.965  1.00  0.00
ATOM    758  CA  ASN A  95      14.026  48.349  58.088  1.00  0.00
ATOM    759  CB  ASN A  95      13.286  49.641  57.736  1.00  0.00
ATOM    760  CG  ASN A  95      14.173  50.664  57.069  1.00  0.00
ATOM    761  OD1 ASN A  95      15.323  50.858  57.461  1.00  0.00
ATOM    762  ND2 ASN A  95      13.634  51.344  56.060  1.00  0.00
ATOM    763  C   ASN A  95      12.998  47.334  58.553  1.00  0.00
ATOM    764  O   ASN A  95      12.424  46.590  57.760  1.00  0.00
ATOM    765  N   VAL A  96      12.766  47.326  59.857  1.00  0.00
ATOM    766  CA  VAL A  96      11.777  46.446  60.461  1.00  0.00
ATOM    767  CB  VAL A  96      12.410  45.567  61.557  1.00  0.00
ATOM    768  CG1 VAL A  96      11.334  44.792  62.298  1.00  0.00
ATOM    769  CG2 VAL A  96      13.414  44.610  60.924  1.00  0.00
ATOM    770  C   VAL A  96      10.763  47.408  61.083  1.00  0.00
ATOM    771  O   VAL A  96      11.029  48.004  62.123  1.00  0.00
ATOM    772  N   THR A  97       9.607  47.567  60.445  1.00  0.00
ATOM    773  CA  THR A  97       8.604  48.493  60.957  1.00  0.00
ATOM    774  CB  THR A  97       7.599  48.940  59.861  1.00  0.00
ATOM    775  OG1 THR A  97       6.909  47.805  59.333  1.00  0.00
ATOM    776  CG2 THR A  97       8.323  49.638  58.736  1.00  0.00
ATOM    777  C   THR A  97       7.826  47.977  62.161  1.00  0.00
ATOM    778  O   THR A  97       7.230  48.768  62.887  1.00  0.00
ATOM    779  N   ARG A  98       7.823  46.663  62.366  1.00  0.00
ATOM    780  CA  ARG A  98       7.137  46.059  63.512  1.00  0.00
ATOM    781  CB  ARG A  98       5.678  45.728  63.160  1.00  0.00
ATOM    782  CG  ARG A  98       4.755  45.476  64.364  1.00  0.00
ATOM    783  CD  ARG A  98       4.848  44.055  64.881  1.00  0.00
ATOM    784  NE  ARG A  98       3.738  43.713  65.767  1.00  0.00
ATOM    785  CZ  ARG A  98       3.562  42.516  66.323  1.00  0.00
ATOM    786  NH1 ARG A  98       4.429  41.537  66.094  1.00  0.00
ATOM    787  NH2 ARG A  98       2.505  42.288  67.089  1.00  0.00
ATOM    788  C   ARG A  98       7.892  44.789  63.906  1.00  0.00
ATOM    789  O   ARG A  98       7.830  43.782  63.208  1.00  0.00
ATOM    790  N   GLU A  99       8.610  44.849  65.022  1.00  0.00
ATOM    791  CA  GLU A  99       9.396  43.703  65.483  1.00  0.00
ATOM    792  CB  GLU A  99      10.143  44.047  66.777  1.00  0.00
ATOM    793  CG  GLU A  99      11.614  44.400  66.596  1.00  0.00
ATOM    794  CD  GLU A  99      12.428  43.272  65.978  1.00  0.00
ATOM    795  OE1 GLU A  99      12.349  42.117  66.464  1.00  0.00
ATOM    796  OE2 GLU A  99      13.164  43.545  65.015  1.00  0.00
```

```
ATOM    797  C   GLU A  99       8.567  42.451  65.713  1.00  0.00
ATOM    798  O   GLU A  99       7.409  42.520  66.134  1.00  0.00
ATOM    799  N   ALA A 100       9.181  41.305  65.443  1.00  0.00
ATOM    800  CA  ALA A 100       8.534  40.015  65.616  1.00  0.00
ATOM    801  CB  ALA A 100       9.534  38.899  65.383  1.00  0.00
ATOM    802  C   ALA A 100       7.952  39.914  67.015  1.00  0.00
ATOM    803  O   ALA A 100       8.615  40.240  68.008  1.00  0.00
ATOM    804  N   GLY A 101       6.703  39.472  67.079  1.00  0.00
ATOM    805  CA  GLY A 101       6.022  39.326  68.349  1.00  0.00
ATOM    806  C   GLY A 101       4.617  38.808  68.129  1.00  0.00
ATOM    807  O   GLY A 101       4.067  38.944  67.035  1.00  0.00
ATOM    808  N   PRO A 102       4.005  38.191  69.144  1.00  0.00
ATOM    809  CD  PRO A 102       4.419  38.035  70.547  1.00  0.00
ATOM    810  CA  PRO A 102       2.649  37.687  68.947  1.00  0.00
ATOM    811  CB  PRO A 102       2.297  37.107  70.316  1.00  0.00
ATOM    812  CG  PRO A 102       3.094  37.956  71.254  1.00  0.00
ATOM    813  C   PRO A 102       1.722  38.816  68.546  1.00  0.00
ATOM    814  O   PRO A 102       1.967  39.979  68.876  1.00  0.00
ATOM    815  N   VAL A 103       0.675  38.472  67.806  1.00  0.00
ATOM    816  CA  VAL A 103      -0.320  39.448  67.386  1.00  0.00
ATOM    817  CB  VAL A 103      -1.332  38.806  66.428  1.00  0.00
ATOM    818  CG1 VAL A 103      -2.501  39.766  66.198  1.00  0.00
ATOM    819  CG2 VAL A 103      -0.657  38.452  65.108  1.00  0.00
ATOM    820  C   VAL A 103      -1.022  39.877  68.670  1.00  0.00
ATOM    821  O   VAL A 103      -1.115  39.088  69.610  1.00  0.00
ATOM    822  N   LYS A 104      -1.503  41.116  68.723  1.00  0.00
ATOM    823  CA  LYS A 104      -2.183  41.607  69.922  1.00  0.00
ATOM    824  CB  LYS A 104      -2.579  43.079  69.757  1.00  0.00
ATOM    825  CG  LYS A 104      -3.132  43.713  71.033  1.00  0.00
ATOM    826  CD  LYS A 104      -3.533  45.168  70.811  1.00  0.00
ATOM    827  CE  LYS A 104      -4.159  45.773  72.054  1.00  0.00
ATOM    828  NZ  LYS A 104      -5.333  44.971  72.504  1.00  0.00
ATOM    829  C   LYS A 104      -3.424  40.774  70.232  1.00  0.00
ATOM    830  O   LYS A 104      -4.443  40.881  69.541  1.00  0.00
ATOM    831  N   GLY A 105      -3.324  39.946  71.269  1.00  0.00
ATOM    832  CA  GLY A 105      -4.438  39.098  71.666  1.00  0.00
ATOM    833  C   GLY A 105      -4.233  37.618  71.394  1.00  0.00
ATOM    834  O   GLY A 105      -4.567  36.778  72.240  1.00  0.00
ATOM    835  N   GLY A 106      -3.688  37.286  70.224  1.00  0.00
ATOM    836  CA  GLY A 106      -3.485  35.887  69.875  1.00  0.00
ATOM    837  C   GLY A 106      -2.103  35.331  70.166  1.00  0.00
ATOM    838  O   GLY A 106      -1.390  35.841  71.037  1.00  0.00
ATOM    839  N   THR A 107      -1.734  34.269  69.445  1.00  0.00
ATOM    840  CA  THR A 107      -0.427  33.631  69.610  1.00  0.00
ATOM    841  CB  THR A 107      -0.564  32.189  70.164  1.00  0.00
ATOM    842  OG1 THR A 107      -1.256  31.354  69.221  1.00  0.00
ATOM    843  CG2 THR A 107      -1.329  32.203  71.482  1.00  0.00
ATOM    844  C   THR A 107       0.401  33.594  68.320  1.00  0.00
ATOM    845  O   THR A 107       1.618  33.359  68.364  1.00  0.00
ATOM    846  N   THR A 108      -0.257  33.807  67.178  1.00  0.00
ATOM    847  CA  THR A 108       0.419  33.830  65.882  1.00  0.00
ATOM    848  CB  THR A 108      -0.549  34.209  64.727  1.00  0.00
ATOM    849  OG1 THR A 108      -1.665  33.311  64.703  1.00  0.00
ATOM    850  CG2 THR A 108       0.166  34.140  63.383  1.00  0.00
ATOM    851  C   THR A 108       1.502  34.910  65.961  1.00  0.00
ATOM    852  O   THR A 108       1.212  36.051  66.314  1.00  0.00
ATOM    853  N   VAL A 109       2.740  34.541  65.649  1.00  0.00
ATOM    854  CA  VAL A 109       3.858  35.478  65.690  1.00  0.00
ATOM    855  CB  VAL A 109       5.181  34.776  66.058  1.00  0.00
ATOM    856  CG1 VAL A 109       6.283  35.818  66.270  1.00  0.00
ATOM    857  CG2 VAL A 109       4.993  33.917  67.312  1.00  0.00
ATOM    858  C   VAL A 109       4.015  36.101  64.311  1.00  0.00
ATOM    859  O   VAL A 109       4.045  35.389  63.304  1.00  0.00
```

```
ATOM    860  N   ILE A 110       4.107  37.427  64.271  1.00  0.00
ATOM    861  CA  ILE A 110       4.266  38.145  63.010  1.00  0.00
ATOM    862  CB  ILE A 110       2.930  38.745  62.492  1.00  0.00
ATOM    863  CG2 ILE A 110       1.820  37.701  62.523  1.00  0.00
ATOM    864  CG1 ILE A 110       2.545  39.960  63.348  1.00  0.00
ATOM    865  CD1 ILE A 110       1.326  40.710  62.845  1.00  0.00
ATOM    866  C   ILE A 110       5.230  39.310  63.167  1.00  0.00
ATOM    867  O   ILE A 110       5.657  39.646  64.274  1.00  0.00
ATOM    868  N   ALA A 111       5.564  39.917  62.035  1.00  0.00
ATOM    869  CA  ALA A 111       6.435  41.079  61.992  1.00  0.00
ATOM    870  CB  ALA A 111       7.897  40.660  62.113  1.00  0.00
ATOM    871  C   ALA A 111       6.205  41.772  60.661  1.00  0.00
ATOM    872  O   ALA A 111       5.706  41.163  59.715  1.00  0.00
ATOM    873  N   PHE A 112       6.539  43.053  60.596  1.00  0.00
ATOM    874  CA  PHE A 112       6.415  43.798  59.354  1.00  0.00
ATOM    875  CB  PHE A 112       5.404  44.930  59.482  1.00  0.00
ATOM    876  CG  PHE A 112       3.986  44.493  59.306  1.00  0.00
ATOM    877  CD1 PHE A 112       3.254  44.002  60.376  1.00  0.00
ATOM    878  CD2 PHE A 112       3.376  44.573  58.055  1.00  0.00
ATOM    879  CE1 PHE A 112       1.926  43.600  60.204  1.00  0.00
ATOM    880  CE2 PHE A 112       2.055  44.176  57.877  1.00  0.00
ATOM    881  CZ  PHE A 112       1.332  43.690  58.957  1.00  0.00
ATOM    882  C   PHE A 112       7.773  44.377  59.024  1.00  0.00
ATOM    883  O   PHE A 112       8.447  44.911  59.904  1.00  0.00
ATOM    884  N   VAL A 113       8.169  44.250  57.760  1.00  0.00
ATOM    885  CA  VAL A 113       9.448  44.756  57.290  1.00  0.00
ATOM    886  CB  VAL A 113      10.472  43.601  57.158  1.00  0.00
ATOM    887  CG1 VAL A 113      10.559  42.859  58.474  1.00  0.00
ATOM    888  CG2 VAL A 113      10.064  42.635  56.033  1.00  0.00
ATOM    889  C   VAL A 113       9.242  45.429  55.936  1.00  0.00
ATOM    890  O   VAL A 113       8.170  45.314  55.340  1.00  0.00
ATOM    891  N   GLU A 114      10.252  46.153  55.465  1.00  0.00
ATOM    892  CA  GLU A 114      10.138  46.813  54.171  1.00  0.00
ATOM    893  CB  GLU A 114      10.361  48.315  54.290  1.00  0.00
ATOM    894  CG  GLU A 114       9.494  49.016  55.293  1.00  0.00
ATOM    895  CD  GLU A 114       9.652  50.517  55.216  1.00  0.00
ATOM    896  OE1 GLU A 114      10.791  50.983  55.004  1.00  0.00
ATOM    897  OE2 GLU A 114       8.649  51.236  55.374  1.00  0.00
ATOM    898  C   GLU A 114      11.164  46.281  53.204  1.00  0.00
ATOM    899  O   GLU A 114      12.233  45.840  53.602  1.00  0.00
ATOM    900  N   ASP A 115      10.841  46.327  51.921  1.00  0.00
ATOM    901  CA  ASP A 115      11.814  45.887  50.944  1.00  0.00
ATOM    902  CB  ASP A 115      11.118  45.177  49.761  1.00  0.00
ATOM    903  CG  ASP A 115      10.590  46.115  48.700  1.00  0.00
ATOM    904  OD1 ASP A 115      10.486  47.335  48.931  1.00  0.00
ATOM    905  OD2 ASP A 115      10.259  45.594  47.608  1.00  0.00
ATOM    906  C   ASP A 115      12.607  47.144  50.564  1.00  0.00
ATOM    907  O   ASP A 115      12.348  48.229  51.090  1.00  0.00
ATOM    908  N   PRO A 116      13.590  47.025  49.665  1.00  0.00
ATOM    909  CD  PRO A 116      14.097  45.820  48.993  1.00  0.00
ATOM    910  CA  PRO A 116      14.387  48.196  49.288  1.00  0.00
ATOM    911  CB  PRO A 116      15.358  47.623  48.259  1.00  0.00
ATOM    912  CG  PRO A 116      15.508  46.205  48.713  1.00  0.00
ATOM    913  C   PRO A 116      13.651  49.417  48.765  1.00  0.00
ATOM    914  O   PRO A 116      14.144  50.539  48.882  1.00  0.00
ATOM    915  N   ASP A 117      12.476  49.201  48.195  1.00  0.00
ATOM    916  CA  ASP A 117      11.699  50.290  47.633  1.00  0.00
ATOM    917  CB  ASP A 117      11.066  49.834  46.318  1.00  0.00
ATOM    918  CG  ASP A 117      12.106  49.555  45.248  1.00  0.00
ATOM    919  OD1 ASP A 117      12.660  50.525  44.685  1.00  0.00
ATOM    920  OD2 ASP A 117      12.380  48.367  44.979  1.00  0.00
ATOM    921  C   ASP A 117      10.638  50.823  48.582  1.00  0.00
ATOM    922  O   ASP A 117       9.874  51.727  48.232  1.00  0.00
```

```
ATOM    923  N   GLY A 118      10.589  50.259  49.782  1.00  0.00
ATOM    924  CA  GLY A 118       9.626  50.712  50.765  1.00  0.00
ATOM    925  C   GLY A 118       8.336  49.927  50.805  1.00  0.00
ATOM    926  O   GLY A 118       7.417  50.289  51.538  1.00  0.00
ATOM    927  N   TYR A 119       8.246  48.854  50.025  1.00  0.00
ATOM    928  CA  TYR A 119       7.038  48.033  50.034  1.00  0.00
ATOM    929  CB  TYR A 119       7.066  47.019  48.894  1.00  0.00
ATOM    930  CG  TYR A 119       6.786  47.591  47.526  1.00  0.00
ATOM    931  CD1 TYR A 119       5.505  48.011  47.171  1.00  0.00
ATOM    932  CE1 TYR A 119       5.240  48.499  45.895  1.00  0.00
ATOM    933  CD2 TYR A 119       7.797  47.681  46.573  1.00  0.00
ATOM    934  CE2 TYR A 119       7.544  48.170  45.298  1.00  0.00
ATOM    935  CZ  TYR A 119       6.264  48.573  44.967  1.00  0.00
ATOM    936  OH  TYR A 119       6.014  49.036  43.703  1.00  0.00
ATOM    937  C   TYR A 119       6.943  47.271  51.353  1.00  0.00
ATOM    938  O   TYR A 119       7.941  46.751  51.850  1.00  0.00
ATOM    939  N   LYS A 120       5.740  47.194  51.910  1.00  0.00
ATOM    940  CA  LYS A 120       5.540  46.478  53.162  1.00  0.00
ATOM    941  CB  LYS A 120       4.323  47.037  53.900  1.00  0.00
ATOM    942  CG  LYS A 120       4.563  48.363  54.587  1.00  0.00
ATOM    943  CD  LYS A 120       5.470  48.180  55.800  1.00  0.00
ATOM    944  CE  LYS A 120       5.587  49.436  56.649  1.00  0.00
ATOM    945  NZ  LYS A 120       6.207  50.573  55.912  1.00  0.00
ATOM    946  C   LYS A 120       5.375  44.970  52.970  1.00  0.00
ATOM    947  O   LYS A 120       4.618  44.512  52.113  1.00  0.00
ATOM    948  N   ILE A 121       6.108  44.204  53.771  1.00  0.00
ATOM    949  CA  ILE A 121       6.039  42.748  53.725  1.00  0.00
ATOM    950  CB  ILE A 121       7.399  42.125  53.334  1.00  0.00
ATOM    951  CG2 ILE A 121       7.284  40.610  53.289  1.00  0.00
ATOM    952  CG1 ILE A 121       7.850  42.671  51.978  1.00  0.00
ATOM    953  CD1 ILE A 121       9.235  42.198  51.557  1.00  0.00
ATOM    954  C   ILE A 121       5.654  42.246  55.112  1.00  0.00
ATOM    955  O   ILE A 121       6.305  42.574  56.101  1.00  0.00
ATOM    956  N   GLU A 122       4.588  41.460  55.179  1.00  0.00
ATOM    957  CA  GLU A 122       4.097  40.903  56.436  1.00  0.00
ATOM    958  CB  GLU A 122       2.562  40.856  56.392  1.00  0.00
ATOM    959  CG  GLU A 122       1.851  40.289  57.611  1.00  0.00
ATOM    960  CD  GLU A 122       0.336  40.434  57.493  1.00  0.00
ATOM    961  OE1 GLU A 122      -0.279  39.780  56.624  1.00  0.00
ATOM    962  OE2 GLU A 122      -0.253  41.221  58.260  1.00  0.00
ATOM    963  C   GLU A 122       4.672  39.498  56.576  1.00  0.00
ATOM    964  O   GLU A 122       4.448  38.652  55.711  1.00  0.00
ATOM    965  N   LEU A 123       5.432  39.256  57.641  1.00  0.00
ATOM    966  CA  LEU A 123       6.020  37.931  57.859  1.00  0.00
ATOM    967  CB  LEU A 123       7.478  38.062  58.309  1.00  0.00
ATOM    968  CG  LEU A 123       8.406  38.851  57.382  1.00  0.00
ATOM    969  CD1 LEU A 123       9.811  38.882  57.963  1.00  0.00
ATOM    970  CD2 LEU A 123       8.396  38.225  55.993  1.00  0.00
ATOM    971  C   LEU A 123       5.199  37.211  58.919  1.00  0.00
ATOM    972  O   LEU A 123       5.018  37.723  60.026  1.00  0.00
ATOM    973  N   ILE A 124       4.696  36.026  58.575  1.00  0.00
ATOM    974  CA  ILE A 124       3.865  35.247  59.489  1.00  0.00
ATOM    975  CB  ILE A 124       2.473  34.974  58.867  1.00  0.00
ATOM    976  CG2 ILE A 124       1.625  34.134  59.819  1.00  0.00
ATOM    977  CG1 ILE A 124       1.780  36.311  58.552  1.00  0.00
ATOM    978  CD1 ILE A 124       0.422  36.166  57.888  1.00  0.00
ATOM    979  C   ILE A 124       4.537  33.923  59.831  1.00  0.00
ATOM    980  O   ILE A 124       4.843  33.121  58.952  1.00  0.00
ATOM    981  N   GLU A 125       4.742  33.691  61.122  1.00  0.00
ATOM    982  CA  GLU A 125       5.400  32.477  61.583  1.00  0.00
ATOM    983  CB  GLU A 125       5.883  32.690  63.010  1.00  0.00
ATOM    984  CG  GLU A 125       6.898  31.683  63.492  1.00  0.00
ATOM    985  CD  GLU A 125       7.468  32.082  64.833  1.00  0.00
```

```
ATOM    986  OE1 GLU A 125       8.335  32.981  64.886  1.00  0.00
ATOM    987  OE2 GLU A 125       7.038  31.505  65.846  1.00  0.00
ATOM    988  C   GLU A 125       4.469  31.274  61.520  1.00  0.00
ATOM    989  O   GLU A 125       3.326  31.339  61.986  1.00  0.00
ATOM    990  N   GLU A 126       4.963  30.178  60.948  1.00  0.00
ATOM    991  CA  GLU A 126       4.161  28.969  60.832  1.00  0.00
ATOM    992  CB  GLU A 126       4.035  28.558  59.358  1.00  0.00
ATOM    993  CG  GLU A 126       3.272  29.577  58.503  1.00  0.00
ATOM    994  CD  GLU A 126       2.613  28.963  57.269  1.00  0.00
ATOM    995  OE1 GLU A 126       3.328  28.511  56.348  1.00  0.00
ATOM    996  OE2 GLU A 126       1.365  28.932  57.215  1.00  0.00
ATOM    997  C   GLU A 126       4.708  27.808  61.662  1.00  0.00
ATOM    998  OT1 GLU A 126       5.871  27.884  62.132  1.00  0.00
ATOM    999  OXT GLU A 126       3.956  26.826  61.841  1.00  0.00
TER
ATOM   1000 NI+2 NI2 N   1       2.087  40.223  35.439  1.00  0.00
TER
ATOM   1990  OH2 WAT W   1     -11.390  51.191  53.835  1.00  0.00
ATOM   1991  OH2 WAT W   2      -2.515  31.690  52.442  1.00  0.00
ATOM   1992  OH2 WAT W   3      -3.374  48.313  40.324  1.00  0.00
ATOM   1993  OH2 WAT W   4      -2.475  50.747  39.050  1.00  0.00
ATOM   1994  OH2 WAT W   5      -1.322  52.311  46.943  1.00  0.00
ATOM   1995  OH2 WAT W   6      -4.828  44.280  41.898  1.00  0.00
ATOM   1996  OH2 WAT W   7      -5.239  49.474  56.807  1.00  0.00
ATOM   1997  OH2 WAT W   8     -15.309  41.139  57.809  1.00  0.00
ATOM   1998  OH2 WAT W   9     -14.892  48.326  47.302  1.00  0.00
ATOM   1999  OH2 WAT W  10      -6.718  42.335  42.662  1.00  0.00
ATOM   2000  OH2 WAT W  11      -3.157  33.795  54.124  1.00  0.00
ATOM   2001  OH2 WAT W  12      -6.525  31.534  45.444  1.00  0.00
ATOM   2002  OH2 WAT W  13     -15.662  38.542  60.911  1.00  0.00
ATOM   2003  OH2 WAT W  14      -0.258  51.991  40.142  1.00  0.00
ATOM   2004  OH2 WAT W  15      -1.128  36.742  46.788  1.00  0.00
ATOM   2005  OH2 WAT W  16     -15.215  33.004  37.238  1.00  0.00
ATOM   2006  OH2 WAT W  17       0.053  46.670  35.861  1.00  0.00
ATOM   2007  OH2 WAT W  18      -8.625  51.334  47.117  1.00  0.00
ATOM   2008  OH2 WAT W  19      -0.634  49.156  37.604  1.00  0.00
ATOM   2009  OH2 WAT W  20     -10.398  35.482  61.035  1.00  0.00
ATOM   2010  OH2 WAT W  21     -19.456  44.365  43.415  1.00  0.00
ATOM   2011  OH2 WAT W  22      -4.702  46.128  22.695  1.00  0.00
ATOM   2012  OH2 WAT W  23      -6.211  56.657  43.558  1.00  0.00
ATOM   2013  OH2 WAT W  24     -17.990  32.435  48.969  1.00  0.00
ATOM   2014  OH2 WAT W  25      -8.327  41.504  20.706  1.00  0.00
ATOM   2015  OH2 WAT W  26     -19.298  43.012  40.991  1.00  0.00
ATOM   2016  OH2 WAT W  27     -17.926  52.636  52.065  1.00  0.00
ATOM   2017  OH2 WAT W  28     -13.608  32.797  31.695  1.00  0.00
ATOM   2018  OH2 WAT W  29     -16.706  50.941  41.207  1.00  0.00
ATOM   2019  OH2 WAT W  30       3.097  53.421  37.255  1.00  0.00
ATOM   2020  OH2 WAT W  31     -16.847  50.750  36.334  1.00  0.00
ATOM   2021  OH2 WAT W  32     -19.210  50.489  53.474  1.00  0.00
ATOM   2022  OH2 WAT W  33      -5.940  48.509  59.378  1.00  0.00
ATOM   2023  OH2 WAT W  34      -0.196  44.836  22.640  1.00  0.00
ATOM   2024  OH2 WAT W  35      -2.262  45.828  24.149  1.00  0.00
ATOM   2025  OH2 WAT W  36     -18.984  45.448  24.296  1.00  0.00
ATOM   2026  OH2 WAT W  37      -6.489  44.037  23.078  1.00  0.00
ATOM   2027  OH2 WAT W  38      -9.243  28.834  53.794  1.00  0.00
ATOM   2028  OH2 WAT W  39     -13.597  29.149  44.014  1.00  0.00
ATOM   2029  OH2 WAT W  40     -17.735  48.179  54.099  1.00  0.00
ATOM   2030  OH2 WAT W  41      -4.474  30.284  43.767  1.00  0.00
ATOM   2031  OH2 WAT W  42     -17.466  46.416  58.486  1.00  0.00
ATOM   2032  OH2 WAT W  43       1.841  30.836  46.987  1.00  0.00
ATOM   2033  OH2 WAT W  44     -19.988  37.854  51.939  1.00  0.00
ATOM   2034  OH2 WAT W  45     -17.139  43.559  36.830  1.00  0.00
ATOM   2035  OH2 WAT W  46     -13.070  50.241  38.548  1.00  0.00
```

```
ATOM   2036  OH2 WAT W  47     -14.785  39.836  25.092  1.00  0.00
ATOM   2037  OH2 WAT W  48      -5.138  44.152  57.101  1.00  0.00
ATOM   2038  OH2 WAT W  49      -9.947  34.273  37.097  1.00  0.00
ATOM   2039  OH2 WAT W  50     -18.341  44.922  38.898  1.00  0.00
ATOM   2040  OH2 WAT W  51     -16.579  53.896  54.224  1.00  0.00
ATOM   2041  OH2 WAT W  52     -15.669  29.431  40.696  1.00  0.00
ATOM   2042  OH2 WAT W  53      -9.108  52.247  60.812  1.00  0.00
ATOM   2043  OH2 WAT W  54      -7.465  53.007  63.237  1.00  0.00
ATOM   2044  OH2 WAT W  55     -14.327  35.721  35.349  1.00  0.00
ATOM   2045  OH2 WAT W  56     -10.859  54.196  43.119  1.00  0.00
ATOM   2046  OH2 WAT W  57     -12.537  33.663  36.366  1.00  0.00
ATOM   2047  OH2 WAT W  58     -11.661  32.882  23.277  1.00  0.00
ATOM   2048  OH2 WAT W  59     -12.257  31.208  34.250  1.00  0.00
ATOM   2049  OH2 WAT W  60     -15.655  30.627  28.994  1.00  0.00
ATOM   2050  OH2 WAT W  61      -9.561  53.964  46.413  1.00  0.00
ATOM   2051  OH2 WAT W  62     -12.108  42.008  68.743  1.00  0.00
ATOM   2052  OH2 WAT W  63     -25.321  39.883  32.264  1.00  0.00
ATOM   2053  OH2 WAT W  64     -11.149  52.459  33.182  1.00  0.00
ATOM   2054  OH2 WAT W  65     -14.102  48.200  37.099  1.00  0.00
ATOM   2055  OH2 WAT W  66     -12.106  58.616  53.504  1.00  0.00
ATOM   2056  OH2 WAT W  67       6.978  27.089  48.495  1.00  0.00
ATOM   2057  OH2 WAT W  68     -12.170  28.514  46.706  1.00  0.00
ATOM   2058  OH2 WAT W  69      -5.834  27.676  35.883  1.00  0.00
ATOM   2059  OH2 WAT W  70     -10.148  34.484  24.975  1.00  0.00
ATOM   2060  OH2 WAT W  71     -25.333  45.086  32.737  1.00  0.00
ATOM   2061  OH2 WAT W  72      -2.795  31.875  26.670  1.00  0.00
ATOM   2062  OH2 WAT W  73     -18.818  42.428  54.344  1.00  0.00
ATOM   2063  OH2 WAT W  74     -16.351  47.045  39.108  1.00  0.00
ATOM   2064  OH2 WAT W  75      -2.364  29.058  56.542  1.00  0.00
ATOM   2065  OH2 WAT W  76       0.224  42.809  24.594  1.00  0.00
ATOM   2066  OH2 WAT W  77      -4.370  48.216  31.013  1.00  0.00
ATOM   2067  OH2 WAT W  78     -16.632  38.564  23.435  1.00  0.00
ATOM   2068  OH2 WAT W  79      -7.363  26.888  55.121  1.00  0.00
ATOM   2069  OH2 WAT W  80      -4.669  30.129  25.187  1.00  0.00
ATOM   2070  OH2 WAT W  81     -16.962  53.090  64.913  1.00  0.00
ATOM   2071  OH2 WAT W  82      -7.113  35.835  21.080  1.00  0.00
ATOM   2072  OH2 WAT W  83     -18.630  29.982  34.780  1.00  0.00
ATOM   2073  OH2 WAT W  84      -9.435  36.175  22.934  1.00  0.00
ATOM   2074  OH2 WAT W  85     -15.346  50.593  67.406  1.00  0.00
ATOM   2075  OH2 WAT W  86       4.896  57.477  37.934  1.00  0.00
ATOM   2076  OH2 WAT W  87      -5.992  49.609  25.858  1.00  0.00
ATOM   2077  OH2 WAT W  88     -20.760  36.038  47.100  1.00  0.00
ATOM   2078  OH2 WAT W  89     -11.240  38.164  21.761  1.00  0.00
ATOM   2079  OH2 WAT W  90      -4.071  32.456  20.887  1.00  0.00
```

The PBD file for transition state analogue, hydroxamate (TSI), bound E. coli GlxI

```
REMARK   MSI WebLab Viewer PDB file
REMARK   Created:  Fri Mar 02 12:23:19 Eastern Standard Time 2001
CRYST1   45.179   45.179  106.574  90.00  90.00  90.00 P1
ATOM      1  N1  TSI A 300      21.869  43.266  22.220  1.00  0.00
ATOM      2  CA1 TSI A 300      22.040  42.077  22.405  1.00  0.00
ATOM      3  C1  TSI A 300      20.829  41.032  22.925  1.00  0.00
ATOM      4  O11 TSI A 300      20.021  41.876  23.007  1.00  0.00
ATOM      5  O12 TSI A 300      20.872  40.246  21.915  1.00  0.00
ATOM      6  CB1 TSI A 300      23.434  41.726  22.203  1.00  0.00
ATOM      7  CG1 TSI A 300      23.947  40.537  22.408  1.00  0.00
ATOM      8  CD1 TSI A 300      24.525  40.091  21.089  1.00  0.00
ATOM      9  OE1 TSI A 300      24.732  41.109  19.991  1.00  0.00
ATOM     10  N2  TSI A 300      24.917  38.540  20.894  1.00  0.00
ATOM     11  CA2 TSI A 300      25.531  38.336  19.382  1.00  0.00
ATOM     12  C2  TSI A 300      24.448  38.631  18.307  1.00  0.00
```

```
ATOM     13  O2  TSI A 300      22.971  38.373  18.328  1.00  0.00
ATOM     14  CB2 TSI A 300      25.873  36.792  19.333  1.00  0.00
ATOM     15  CG2 TSI A 300      27.103  36.507  20.593  1.00  0.00
ATOM     16  N3  TSI A 300      24.820  39.241  17.080  1.00  0.00
ATOM     17  CA3 TSI A 300      24.227  39.679  15.865  1.00  0.00
ATOM     18  C3  TSI A 300      24.979  39.797  14.695  1.00  0.00
ATOM     19  O31 TSI A 300      24.478  40.172  13.614  1.00  0.00
ATOM     20  O32 TSI A 300      26.191  39.488  14.855  1.00  0.00
ATOM     21  CD2 TSI A 300      26.889  35.020  20.881  1.00  0.00
ATOM     22  OE2 TSI A 300      25.783  34.705  21.387  1.00  0.00
ATOM     23  NE2 TSI A 300      27.889  34.015  20.562  1.00  0.00
ATOM     24  OZ2 TSI A 300      27.563  33.149  21.612  1.00  0.00
ATOM     25  CZ1 TSI A 300      28.820  34.296  19.275  1.00  0.00
ATOM     26  N   MET A   1      28.110  18.256  14.928  1.00  0.00
ATOM     27  CA  MET A   1      27.158  19.169  15.553  1.00  0.00
ATOM     28  CB  MET A   1      25.863  19.222  14.744  1.00  0.00
ATOM     29  CG  MET A   1      26.046  19.671  13.313  1.00  0.00
ATOM     30  SD  MET A   1      24.465  19.872  12.481  1.00  0.00
ATOM     31  CE  MET A   1      24.525  21.596  12.100  1.00  0.00
ATOM     32  C   MET A   1      26.832  18.746  16.976  1.00  0.00
ATOM     33  O   MET A   1      26.737  17.553  17.279  1.00  0.00
ATOM     34  N   ARG A   2      26.649  19.731  17.844  1.00  0.00
ATOM     35  CA  ARG A   2      26.320  19.479  19.242  1.00  0.00
ATOM     36  CB  ARG A   2      27.560  18.990  19.986  1.00  0.00
ATOM     37  CG  ARG A   2      28.665  20.030  20.039  1.00  0.00
ATOM     38  CD  ARG A   2      29.947  19.484  20.657  1.00  0.00
ATOM     39  NE  ARG A   2      30.955  20.533  20.771  1.00  0.00
ATOM     40  CZ  ARG A   2      31.219  21.202  21.888  1.00  0.00
ATOM     41  NH1 ARG A   2      30.557  20.932  23.006  1.00  0.00
ATOM     42  NH2 ARG A   2      32.139  22.158  21.883  1.00  0.00
ATOM     43  C   ARG A   2      25.836  20.785  19.858  1.00  0.00
ATOM     44  O   ARG A   2      26.070  21.856  19.300  1.00  0.00
ATOM     45  N   LEU A   3      25.158  20.708  20.997  1.00  0.00
ATOM     46  CA  LEU A   3      24.687  21.921  21.651  1.00  0.00
ATOM     47  CB  LEU A   3      23.558  21.619  22.637  1.00  0.00
ATOM     48  CG  LEU A   3      22.127  21.697  22.109  1.00  0.00
ATOM     49  CD1 LEU A   3      21.893  20.571  21.143  1.00  0.00
ATOM     50  CD2 LEU A   3      21.144  21.600  23.264  1.00  0.00
ATOM     51  C   LEU A   3      25.838  22.580  22.391  1.00  0.00
ATOM     52  O   LEU A   3      26.518  21.942  23.187  1.00  0.00
ATOM     53  N   LEU A   4      26.049  23.863  22.115  1.00  0.00
ATOM     54  CA  LEU A   4      27.116  24.630  22.739  1.00  0.00
ATOM     55  CB  LEU A   4      27.730  25.582  21.706  1.00  0.00
ATOM     56  CG  LEU A   4      28.483  24.885  20.560  1.00  0.00
ATOM     57  CD1 LEU A   4      28.783  25.851  19.444  1.00  0.00
ATOM     58  CD2 LEU A   4      29.765  24.262  21.102  1.00  0.00
ATOM     59  C   LEU A   4      26.619  25.408  23.967  1.00  0.00
ATOM     60  O   LEU A   4      27.295  25.466  24.991  1.00  0.00
ATOM     61  N   HIS A   5      25.437  26.004  23.872  1.00  0.00
ATOM     62  CA  HIS A   5      24.911  26.746  25.005  1.00  0.00
ATOM     63  CB  HIS A   5      25.836  27.921  25.361  1.00  0.00
ATOM     64  CG  HIS A   5      25.960  28.959  24.285  1.00  0.00
ATOM     65  CD2 HIS A   5      25.093  29.905  23.854  1.00  0.00
ATOM     66  ND1 HIS A   5      27.109  29.127  23.541  1.00  0.00
ATOM     67  CE1 HIS A   5      26.945  30.134  22.702  1.00  0.00
ATOM     68  NE2 HIS A   5      25.731  30.624  22.872  1.00  0.00
ATOM     69  C   HIS A   5      23.503  27.280  24.816  1.00  0.00
ATOM     70  O   HIS A   5      22.904  27.173  23.742  1.00  0.00
ATOM     71  N   THR A   6      22.983  27.842  25.900  1.00  0.00
ATOM     72  CA  THR A   6      21.678  28.470  25.906  1.00  0.00
ATOM     73  CB  THR A   6      20.714  27.770  26.863  1.00  0.00
ATOM     74  OG1 THR A   6      20.415  26.465  26.349  1.00  0.00
ATOM     75  CG2 THR A   6      19.425  28.572  27.003  1.00  0.00
```

```
ATOM     76  C   THR A   6      21.978  29.879  26.390  1.00  0.00
ATOM     77  O   THR A   6      22.653  30.064  27.400  1.00  0.00
ATOM     78  N   MET A   7      21.500  30.870  25.655  1.00  0.00
ATOM     79  CA  MET A   7      21.767  32.247  26.024  1.00  0.00
ATOM     80  CB  MET A   7      22.209  33.047  24.794  1.00  0.00
ATOM     81  CG  MET A   7      22.691  34.466  25.095  1.00  0.00
ATOM     82  SD  MET A   7      21.387  35.665  25.475  1.00  0.00
ATOM     83  CE  MET A   7      21.121  36.378  23.815  1.00  0.00
ATOM     84  C   MET A   7      20.548  32.880  26.659  1.00  0.00
ATOM     85  O   MET A   7      19.421  32.693  26.197  1.00  0.00
ATOM     86  N   LEU A   8      20.794  33.612  27.739  1.00  0.00
ATOM     87  CA  LEU A   8      19.753  34.302  28.478  1.00  0.00
ATOM     88  CB  LEU A   8      19.459  33.560  29.779  1.00  0.00
ATOM     89  CG  LEU A   8      19.036  32.093  29.655  1.00  0.00
ATOM     90  CD1 LEU A   8      19.011  31.471  31.031  1.00  0.00
ATOM     91  CD2 LEU A   8      17.676  31.979  28.972  1.00  0.00
ATOM     92  C   LEU A   8      20.256  35.706  28.797  1.00  0.00
ATOM     93  O   LEU A   8      21.406  35.882  29.215  1.00  0.00
ATOM     94  N   ARG A   9      19.403  36.704  28.592  1.00  0.00
ATOM     95  CA  ARG A   9      19.782  38.082  28.888  1.00  0.00
ATOM     96  CB  ARG A   9      18.975  39.062  28.033  1.00  0.00
ATOM     97  CG  ARG A   9      19.145  38.887  26.529  1.00  0.00
ATOM     98  CD  ARG A   9      18.296  39.903  25.771  1.00  0.00
ATOM     99  NE  ARG A   9      18.377  39.725  24.324  1.00  0.00
ATOM    100  CZ  ARG A   9      17.816  38.721  23.656  1.00  0.00
ATOM    101  NH1 ARG A   9      17.116  37.795  24.296  1.00  0.00
ATOM    102  NH2 ARG A   9      17.975  38.631  22.345  1.00  0.00
ATOM    103  C   ARG A   9      19.467  38.312  30.362  1.00  0.00
ATOM    104  O   ARG A   9      18.377  37.968  30.824  1.00  0.00
ATOM    105  N   VAL A  10      20.423  38.877  31.096  1.00  0.00
ATOM    106  CA  VAL A  10      20.244  39.153  32.519  1.00  0.00
ATOM    107  CB  VAL A  10      21.288  38.397  33.380  1.00  0.00
ATOM    108  CG1 VAL A  10      21.178  36.902  33.114  1.00  0.00
ATOM    109  CG2 VAL A  10      22.700  38.905  33.090  1.00  0.00
ATOM    110  C   VAL A  10      20.344  40.650  32.810  1.00  0.00
ATOM    111  O   VAL A  10      20.984  41.389  32.069  1.00  0.00
ATOM    112  N   GLY A  11      19.710  41.087  33.893  1.00  0.00
ATOM    113  CA  GLY A  11      19.734  42.499  34.244  1.00  0.00
ATOM    114  C   GLY A  11      20.941  42.927  35.058  1.00  0.00
ATOM    115  O   GLY A  11      21.242  44.117  35.159  1.00  0.00
ATOM    116  N   ASP A  12      21.629  41.954  35.638  1.00  0.00
ATOM    117  CA  ASP A  12      22.804  42.202  36.455  1.00  0.00
ATOM    118  CB  ASP A  12      22.393  42.303  37.929  1.00  0.00
ATOM    119  CG  ASP A  12      23.506  42.829  38.823  1.00  0.00
ATOM    120  OD1 ASP A  12      24.674  42.431  38.631  1.00  0.00
ATOM    121  OD2 ASP A  12      23.208  43.632  39.738  1.00  0.00
ATOM    122  C   ASP A  12      23.733  41.003  36.236  1.00  0.00
ATOM    123  O   ASP A  12      23.457  39.898  36.703  1.00  0.00
ATOM    124  N   LEU A  13      24.824  41.221  35.512  1.00  0.00
ATOM    125  CA  LEU A  13      25.773  40.154  35.216  1.00  0.00
ATOM    126  CB  LEU A  13      26.842  40.649  34.246  1.00  0.00
ATOM    127  CG  LEU A  13      27.743  39.540  33.707  1.00  0.00
ATOM    128  CD1 LEU A  13      26.942  38.651  32.738  1.00  0.00
ATOM    129  CD2 LEU A  13      28.946  40.156  33.012  1.00  0.00
ATOM    130  C   LEU A  13      26.452  39.582  36.459  1.00  0.00
ATOM    131  O   LEU A  13      26.548  38.367  36.617  1.00  0.00
ATOM    132  N   GLN A  14      26.934  40.450  37.338  1.00  0.00
ATOM    133  CA  GLN A  14      27.589  39.985  38.550  1.00  0.00
ATOM    134  CB  GLN A  14      28.063  41.173  39.388  1.00  0.00
ATOM    135  CG  GLN A  14      28.747  40.763  40.672  1.00  0.00
ATOM    136  CD  GLN A  14      30.056  40.048  40.422  1.00  0.00
ATOM    137  OE1 GLN A  14      30.319  39.001  41.007  1.00  0.00
ATOM    138  NE2 GLN A  14      30.890  40.616  39.556  1.00  0.00
```

```
ATOM    139  C   GLN A  14      26.602  39.141  39.351  1.00  0.00
ATOM    140  O   GLN A  14      26.899  38.005  39.737  1.00  0.00
ATOM    141  N   ARG A  15      25.424  39.702  39.596  1.00  0.00
ATOM    142  CA  ARG A  15      24.366  39.013  40.332  1.00  0.00
ATOM    143  CB  ARG A  15      23.084  39.845  40.270  1.00  0.00
ATOM    144  CG  ARG A  15      21.893  39.264  41.009  1.00  0.00
ATOM    145  CD  ARG A  15      20.710  40.241  40.958  1.00  0.00
ATOM    146  NE  ARG A  15      20.140  40.383  39.615  1.00  0.00
ATOM    147  CZ  ARG A  15      19.219  41.289  39.287  1.00  0.00
ATOM    148  NH1 ARG A  15      18.762  42.141  40.196  1.00  0.00
ATOM    149  NH2 ARG A  15      18.738  41.343  38.053  1.00  0.00
ATOM    150  C   ARG A  15      24.103  37.624  39.749  1.00  0.00
ATOM    151  O   ARG A  15      23.947  36.646  40.487  1.00  0.00
ATOM    152  N   SER A  16      24.049  37.539  38.422  1.00  0.00
ATOM    153  CA  SER A  16      23.799  36.265  37.746  1.00  0.00
ATOM    154  CB  SER A  16      23.568  36.485  36.254  1.00  0.00
ATOM    155  OG  SER A  16      22.268  36.985  36.032  1.00  0.00
ATOM    156  C   SER A  16      24.918  35.261  37.937  1.00  0.00
ATOM    157  O   SER A  16      24.660  34.098  38.249  1.00  0.00
ATOM    158  N   ILE A  17      26.157  35.701  37.739  1.00  0.00
ATOM    159  CA  ILE A  17      27.304  34.819  37.921  1.00  0.00
ATOM    160  CB  ILE A  17      28.626  35.537  37.613  1.00  0.00
ATOM    161  CG2 ILE A  17      29.806  34.597  37.879  1.00  0.00
ATOM    162  CG1 ILE A  17      28.621  36.021  36.163  1.00  0.00
ATOM    163  CD1 ILE A  17      29.870  36.764  35.763  1.00  0.00
ATOM    164  C   ILE A  17      27.338  34.325  39.364  1.00  0.00
ATOM    165  O   ILE A  17      27.580  33.147  39.615  1.00  0.00
ATOM    166  N   ASP A  18      27.102  35.227  40.314  1.00  0.00
ATOM    167  CA  ASP A  18      27.089  34.847  41.723  1.00  0.00
ATOM    168  CB  ASP A  18      26.721  36.040  42.611  1.00  0.00
ATOM    169  CG  ASP A  18      27.861  37.030  42.782  1.00  0.00
ATOM    170  OD1 ASP A  18      28.965  36.798  42.248  1.00  0.00
ATOM    171  OD2 ASP A  18      27.649  38.052  43.466  1.00  0.00
ATOM    172  C   ASP A  18      26.083  33.716  41.972  1.00  0.00
ATOM    173  O   ASP A  18      26.370  32.786  42.715  1.00  0.00
ATOM    174  N   PHE A  19      24.909  33.791  41.351  1.00  0.00
ATOM    175  CA  PHE A  19      23.892  32.753  41.544  1.00  0.00
ATOM    176  CB  PHE A  19      22.552  33.178  40.929  1.00  0.00
ATOM    177  CG  PHE A  19      21.513  32.082  40.924  1.00  0.00
ATOM    178  CD1 PHE A  19      20.751  31.815  42.060  1.00  0.00
ATOM    179  CD2 PHE A  19      21.309  31.298  39.785  1.00  0.00
ATOM    180  CE1 PHE A  19      19.802  30.787  42.065  1.00  0.00
ATOM    181  CE2 PHE A  19      20.364  30.266  39.779  1.00  0.00
ATOM    182  CZ  PHE A  19      19.609  30.009  40.920  1.00  0.00
ATOM    183  C   PHE A  19      24.292  31.403  40.952  1.00  0.00
ATOM    184  O   PHE A  19      24.281  30.390  41.642  1.00  0.00
ATOM    185  N   TYR A  20      24.637  31.394  39.673  1.00  0.00
ATOM    186  CA  TYR A  20      25.014  30.161  38.996  1.00  0.00
ATOM    187  CB  TYR A  20      25.132  30.410  37.485  1.00  0.00
ATOM    188  CG  TYR A  20      23.783  30.514  36.809  1.00  0.00
ATOM    189  CD1 TYR A  20      22.959  29.394  36.687  1.00  0.00
ATOM    190  CE1 TYR A  20      21.673  29.492  36.156  1.00  0.00
ATOM    191  CD2 TYR A  20      23.291  31.742  36.371  1.00  0.00
ATOM    192  CE2 TYR A  20      22.002  31.856  35.838  1.00  0.00
ATOM    193  CZ  TYR A  20      21.198  30.728  35.738  1.00  0.00
ATOM    194  OH  TYR A  20      19.911  30.835  35.251  1.00  0.00
ATOM    195  C   TYR A  20      26.275  29.480  39.513  1.00  0.00
ATOM    196  O   TYR A  20      26.384  28.256  39.443  1.00  0.00
ATOM    197  N   THR A  21      27.221  30.255  40.031  1.00  0.00
ATOM    198  CA  THR A  21      28.454  29.671  40.547  1.00  0.00
ATOM    199  CB  THR A  21      29.679  30.594  40.303  1.00  0.00
ATOM    200  OG1 THR A  21      29.457  31.864  40.923  1.00  0.00
ATOM    201  CG2 THR A  21      29.909  30.788  38.814  1.00  0.00
```

```
ATOM    202  C   THR A  21      28.358  29.376  42.041  1.00  0.00
ATOM    203  O   THR A  21      28.770  28.312  42.498  1.00  0.00
ATOM    204  N   LYS A  22      27.797  30.317  42.791  1.00  0.00
ATOM    205  CA  LYS A  22      27.656  30.170  44.237  1.00  0.00
ATOM    206  CB  LYS A  22      27.509  31.548  44.883  1.00  0.00
ATOM    207  CG  LYS A  22      28.763  32.401  44.814  1.00  0.00
ATOM    208  CD  LYS A  22      29.727  32.050  45.943  1.00  0.00
ATOM    209  CE  LYS A  22      29.159  32.441  47.313  1.00  0.00
ATOM    210  NZ  LYS A  22      28.915  33.912  47.418  1.00  0.00
ATOM    211  C   LYS A  22      26.497  29.283  44.680  1.00  0.00
ATOM    212  O   LYS A  22      26.665  28.453  45.564  1.00  0.00
ATOM    213  N   VAL A  23      25.323  29.461  44.076  1.00  0.00
ATOM    214  CA  VAL A  23      24.151  28.669  44.455  1.00  0.00
ATOM    215  CB  VAL A  23      22.848  29.477  44.274  1.00  0.00
ATOM    216  CG1 VAL A  23      21.655  28.652  44.728  1.00  0.00
ATOM    217  CG2 VAL A  23      22.925  30.769  45.063  1.00  0.00
ATOM    218  C   VAL A  23      24.001  27.332  43.722  1.00  0.00
ATOM    219  O   VAL A  23      23.696  26.323  44.352  1.00  0.00
ATOM    220  N   LEU A  24      24.216  27.324  42.406  1.00  0.00
ATOM    221  CA  LEU A  24      24.083  26.091  41.627  1.00  0.00
ATOM    222  CB  LEU A  24      23.448  26.381  40.260  1.00  0.00
ATOM    223  CG  LEU A  24      22.006  26.909  40.214  1.00  0.00
ATOM    224  CD1 LEU A  24      21.499  26.873  38.775  1.00  0.00
ATOM    225  CD2 LEU A  24      21.108  26.068  41.093  1.00  0.00
ATOM    226  C   LEU A  24      25.387  25.312  41.417  1.00  0.00
ATOM    227  O   LEU A  24      25.407  24.308  40.703  1.00  0.00
ATOM    228  N   GLY A  25      26.472  25.784  42.017  1.00  0.00
ATOM    229  CA  GLY A  25      27.744  25.088  41.899  1.00  0.00
ATOM    230  C   GLY A  25      28.472  25.105  40.568  1.00  0.00
ATOM    231  O   GLY A  25      29.378  24.295  40.361  1.00  0.00
ATOM    232  N   MET A  26      28.100  26.004  39.662  1.00  0.00
ATOM    233  CA  MET A  26      28.768  26.082  38.365  1.00  0.00
ATOM    234  CB  MET A  26      27.879  26.807  37.339  1.00  0.00
ATOM    235  CG  MET A  26      26.767  25.939  36.746  1.00  0.00
ATOM    236  SD  MET A  26      25.776  26.749  35.457  1.00  0.00
ATOM    237  CE  MET A  26      24.190  25.994  35.746  1.00  0.00
ATOM    238  C   MET A  26      30.117  26.795  38.450  1.00  0.00
ATOM    239  O   MET A  26      30.422  27.448  39.443  1.00  0.00
ATOM    240  N   LYS A  27      30.921  26.661  37.400  1.00  0.00
ATOM    241  CA  LYS A  27      32.223  27.313  37.326  1.00  0.00
ATOM    242  CB  LYS A  27      33.272  26.395  36.683  1.00  0.00
ATOM    243  CG  LYS A  27      33.420  25.001  37.280  1.00  0.00
ATOM    244  CD  LYS A  27      34.279  24.129  36.351  1.00  0.00
ATOM    245  CE  LYS A  27      34.486  22.711  36.884  1.00  0.00
ATOM    246  NZ  LYS A  27      35.405  22.648  38.062  1.00  0.00
ATOM    247  C   LYS A  27      32.043  28.514  36.408  1.00  0.00
ATOM    248  O   LYS A  27      31.207  28.485  35.507  1.00  0.00
ATOM    249  N   LEU A  28      32.811  29.572  36.637  1.00  0.00
ATOM    250  CA  LEU A  28      32.733  30.736  35.767  1.00  0.00
ATOM    251  CB  LEU A  28      33.156  32.016  36.495  1.00  0.00
ATOM    252  CG  LEU A  28      32.711  33.365  35.903  1.00  0.00
ATOM    253  CD1 LEU A  28      33.627  34.454  36.434  1.00  0.00
ATOM    254  CD2 LEU A  28      32.751  33.348  34.386  1.00  0.00
ATOM    255  C   LEU A  28      33.768  30.379  34.718  1.00  0.00
ATOM    256  O   LEU A  28      34.965  30.574  34.929  1.00  0.00
ATOM    257  N   LEU A  29      33.304  29.827  33.602  1.00  0.00
ATOM    258  CA  LEU A  29      34.189  29.410  32.524  1.00  0.00
ATOM    259  CB  LEU A  29      33.383  28.711  31.424  1.00  0.00
ATOM    260  CG  LEU A  29      32.712  27.419  31.901  1.00  0.00
ATOM    261  CD1 LEU A  29      31.799  26.863  30.809  1.00  0.00
ATOM    262  CD2 LEU A  29      33.792  26.403  32.287  1.00  0.00
ATOM    263  C   LEU A  29      34.980  30.562  31.936  1.00  0.00
ATOM    264  O   LEU A  29      36.207  30.524  31.905  1.00  0.00
```

```
ATOM    265  N   ARG A  30      34.279  31.589  31.474  1.00  0.00
ATOM    266  CA  ARG A  30      34.935  32.739  30.887  1.00  0.00
ATOM    267  CB  ARG A  30      35.581  32.359  29.551  1.00  0.00
ATOM    268  CG  ARG A  30      34.716  31.472  28.657  1.00  0.00
ATOM    269  CD  ARG A  30      35.242  31.444  27.230  1.00  0.00
ATOM    270  NE  ARG A  30      35.126  32.770  26.632  1.00  0.00
ATOM    271  CZ  ARG A  30      35.996  33.293  25.776  1.00  0.00
ATOM    272  NH1 ARG A  30      37.068  32.600  25.400  1.00  0.00
ATOM    273  NH2 ARG A  30      35.800  34.524  25.315  1.00  0.00
ATOM    274  C   ARG A  30      33.961  33.877  30.665  1.00  0.00
ATOM    275  O   ARG A  30      32.751  33.715  30.809  1.00  0.00
ATOM    276  N   THR A  31      34.502  35.033  30.316  1.00  0.00
ATOM    277  CA  THR A  31      33.691  36.208  30.053  1.00  0.00
ATOM    278  CB  THR A  31      33.804  37.221  31.181  1.00  0.00
ATOM    279  OG1 THR A  31      35.170  37.631  31.288  1.00  0.00
ATOM    280  CG2 THR A  31      33.358  36.609  32.508  1.00  0.00
ATOM    281  C   THR A  31      34.249  36.848  28.793  1.00  0.00
ATOM    282  O   THR A  31      35.444  36.734  28.511  1.00  0.00
ATOM    283  N   SER A  32      33.383  37.519  28.046  1.00  0.00
ATOM    284  CA  SER A  32      33.771  38.184  26.811  1.00  0.00
ATOM    285  CB  SER A  32      33.161  37.464  25.609  1.00  0.00
ATOM    286  OG  SER A  32      33.500  36.089  25.626  1.00  0.00
ATOM    287  C   SER A  32      33.267  39.615  26.846  1.00  0.00
ATOM    288  O   SER A  32      32.401  39.967  27.653  1.00  0.00
ATOM    289  N   GLU A  33      33.789  40.437  25.948  1.00  0.00
ATOM    290  CA  GLU A  33      33.395  41.829  25.918  1.00  0.00
ATOM    291  CB  GLU A  33      34.457  42.650  26.629  1.00  0.00
ATOM    292  CG  GLU A  33      34.023  43.096  27.981  1.00  0.00
ATOM    293  CD  GLU A  33      33.124  44.292  27.903  1.00  0.00
ATOM    294  OE1 GLU A  33      33.651  45.408  27.677  1.00  0.00
ATOM    295  OE2 GLU A  33      31.890  44.116  28.052  1.00  0.00
ATOM    296  C   GLU A  33      33.159  42.385  24.530  1.00  0.00
ATOM    297  O   GLU A  33      33.986  42.218  23.630  1.00  0.00
ATOM    298  N   ASN A  34      32.025  43.054  24.357  1.00  0.00
ATOM    299  CA  ASN A  34      31.707  43.643  23.072  1.00  0.00
ATOM    300  CB  ASN A  34      30.477  42.966  22.466  1.00  0.00
ATOM    301  CG  ASN A  34      30.466  43.038  20.952  1.00  0.00
ATOM    302  OD1 ASN A  34      30.790  44.074  20.370  1.00  0.00
ATOM    303  ND2 ASN A  34      30.092  41.938  20.308  1.00  0.00
ATOM    304  C   ASN A  34      31.460  45.141  23.230  1.00  0.00
ATOM    305  O   ASN A  34      30.316  45.589  23.336  1.00  0.00
ATOM    306  N   PRO A  35      32.540  45.939  23.242  1.00  0.00
ATOM    307  CD  PRO A  35      33.947  45.516  23.163  1.00  0.00
ATOM    308  CA  PRO A  35      32.447  47.395  23.387  1.00  0.00
ATOM    309  CB  PRO A  35      33.913  47.834  23.481  1.00  0.00
ATOM    310  CG  PRO A  35      34.636  46.591  23.948  1.00  0.00
ATOM    311  C   PRO A  35      31.742  48.047  22.205  1.00  0.00
ATOM    312  O   PRO A  35      31.028  49.036  22.363  1.00  0.00
ATOM    313  N   GLU A  36      31.949  47.485  21.016  1.00  0.00
ATOM    314  CA  GLU A  36      31.357  48.025  19.794  1.00  0.00
ATOM    315  CB  GLU A  36      31.888  47.263  18.572  1.00  0.00
ATOM    316  CG  GLU A  36      31.643  47.959  17.231  1.00  0.00
ATOM    317  CD  GLU A  36      32.407  49.271  17.091  1.00  0.00
ATOM    318  OE1 GLU A  36      33.223  49.599  17.981  1.00  0.00
ATOM    319  OE2 GLU A  36      32.198  49.974  16.084  1.00  0.00
ATOM    320  C   GLU A  36      29.825  48.008  19.798  1.00  0.00
ATOM    321  O   GLU A  36      29.201  48.942  19.292  1.00  0.00
ATOM    322  N   TYR A  37      29.217  46.963  20.361  1.00  0.00
ATOM    323  CA  TYR A  37      27.757  46.895  20.402  1.00  0.00
ATOM    324  CB  TYR A  37      27.247  45.673  19.642  1.00  0.00
ATOM    325  CG  TYR A  37      27.543  45.763  18.170  1.00  0.00
ATOM    326  CD1 TYR A  37      28.833  45.547  17.689  1.00  0.00
ATOM    327  CE1 TYR A  37      29.130  45.696  16.339  1.00  0.00
```

```
ATOM    328  CD2 TYR A  37      26.552  46.130  17.264  1.00  0.00
ATOM    329  CE2 TYR A  37      26.836  46.285  15.911  1.00  0.00
ATOM    330  CZ  TYR A  37      28.127  46.064  15.457  1.00  0.00
ATOM    331  OH  TYR A  37      28.416  46.195  14.126  1.00  0.00
ATOM    332  C   TYR A  37      27.176  46.928  21.811  1.00  0.00
ATOM    333  O   TYR A  37      26.034  46.527  22.036  1.00  0.00
ATOM    334  N   LYS A  38      27.990  47.402  22.747  1.00  0.00
ATOM    335  CA  LYS A  38      27.613  47.592  24.145  1.00  0.00
ATOM    336  CB  LYS A  38      26.777  48.869  24.258  1.00  0.00
ATOM    337  CG  LYS A  38      27.542  50.154  23.916  1.00  0.00
ATOM    338  CD  LYS A  38      28.070  50.147  22.489  1.00  0.00
ATOM    339  CE  LYS A  38      28.722  51.474  22.122  1.00  0.00
ATOM    340  NZ  LYS A  38      29.903  51.748  22.994  1.00  0.00
ATOM    341  C   LYS A  38      26.911  46.476  24.916  1.00  0.00
ATOM    342  O   LYS A  38      25.714  46.570  25.216  1.00  0.00
ATOM    343  N   TYR A  39      27.662  45.436  25.259  1.00  0.00
ATOM    344  CA  TYR A  39      27.128  44.330  26.039  1.00  0.00
ATOM    345  CB  TYR A  39      26.102  43.535  25.222  1.00  0.00
ATOM    346  CG  TYR A  39      26.669  42.683  24.107  1.00  0.00
ATOM    347  CD1 TYR A  39      27.160  41.401  24.357  1.00  0.00
ATOM    348  CE1 TYR A  39      27.634  40.597  23.316  1.00  0.00
ATOM    349  CD2 TYR A  39      26.673  43.141  22.795  1.00  0.00
ATOM    350  CE2 TYR A  39      27.140  42.353  21.759  1.00  0.00
ATOM    351  CZ  TYR A  39      27.615  41.083  22.022  1.00  0.00
ATOM    352  OH  TYR A  39      28.026  40.295  20.976  1.00  0.00
ATOM    353  C   TYR A  39      28.276  43.434  26.499  1.00  0.00
ATOM    354  O   TYR A  39      29.347  43.417  25.882  1.00  0.00
ATOM    355  N   SER A  40      28.055  42.712  27.593  1.00  0.00
ATOM    356  CA  SER A  40      29.060  41.815  28.148  1.00  0.00
ATOM    357  CB  SER A  40      29.461  42.264  29.552  1.00  0.00
ATOM    358  OG  SER A  40      29.981  43.578  29.532  1.00  0.00
ATOM    359  C   SER A  40      28.510  40.402  28.226  1.00  0.00
ATOM    360  O   SER A  40      27.303  40.205  28.315  1.00  0.00
ATOM    361  N   LEU A  41      29.407  39.424  28.197  1.00  0.00
ATOM    362  CA  LEU A  41      29.007  38.027  28.264  1.00  0.00
ATOM    363  CB  LEU A  41      29.263  37.332  26.921  1.00  0.00
ATOM    364  CG  LEU A  41      28.578  37.856  25.655  1.00  0.00
ATOM    365  CD1 LEU A  41      29.101  37.087  24.445  1.00  0.00
ATOM    366  CD2 LEU A  41      27.062  37.705  25.771  1.00  0.00
ATOM    367  C   LEU A  41      29.756  37.269  29.350  1.00  0.00
ATOM    368  O   LEU A  41      30.915  37.554  29.636  1.00  0.00
ATOM    369  N   ALA A  42      29.078  36.303  29.954  1.00  0.00
ATOM    370  CA  ALA A  42      29.693  35.459  30.964  1.00  0.00
ATOM    371  CB  ALA A  42      29.265  35.894  32.363  1.00  0.00
ATOM    372  C   ALA A  42      29.222  34.035  30.673  1.00  0.00
ATOM    373  O   ALA A  42      28.044  33.809  30.387  1.00  0.00
ATOM    374  N   PHE A  43      30.141  33.078  30.728  1.00  0.00
ATOM    375  CA  PHE A  43      29.795  31.689  30.469  1.00  0.00
ATOM    376  CB  PHE A  43      30.646  31.116  29.326  1.00  0.00
ATOM    377  CG  PHE A  43      30.443  31.813  28.006  1.00  0.00
ATOM    378  CD1 PHE A  43      30.930  33.103  27.799  1.00  0.00
ATOM    379  CD2 PHE A  43      29.756  31.185  26.971  1.00  0.00
ATOM    380  CE1 PHE A  43      30.734  33.752  26.587  1.00  0.00
ATOM    381  CE2 PHE A  43      29.556  31.832  25.752  1.00  0.00
ATOM    382  CZ  PHE A  43      30.046  33.118  25.561  1.00  0.00
ATOM    383  C   PHE A  43      29.982  30.848  31.721  1.00  0.00
ATOM    384  O   PHE A  43      31.061  30.823  32.318  1.00  0.00
ATOM    385  N   VAL A  44      28.923  30.147  32.106  1.00  0.00
ATOM    386  CA  VAL A  44      28.952  29.309  33.290  1.00  0.00
ATOM    387  CB  VAL A  44      28.060  29.921  34.390  1.00  0.00
ATOM    388  CG1 VAL A  44      28.646  31.259  34.846  1.00  0.00
ATOM    389  CG2 VAL A  44      26.654  30.133  33.860  1.00  0.00
ATOM    390  C   VAL A  44      28.506  27.876  32.972  1.00  0.00
```

```
ATOM    391  O   VAL A  44      27.619  27.657  32.141  1.00  0.00
ATOM    392  N   GLY A  45      29.132  26.908  33.633  1.00  0.00
ATOM    393  CA  GLY A  45      28.803  25.511  33.403  1.00  0.00
ATOM    394  C   GLY A  45      29.521  24.567  34.351  1.00  0.00
ATOM    395  O   GLY A  45      30.337  24.992  35.172  1.00  0.00
ATOM    396  N   TYR A  46      29.210  23.279  34.240  1.00  0.00
ATOM    397  CA  TYR A  46      29.815  22.252  35.081  1.00  0.00
ATOM    398  CB  TYR A  46      28.775  21.189  35.447  1.00  0.00
ATOM    399  CG  TYR A  46      27.638  21.710  36.289  1.00  0.00
ATOM    400  CD1 TYR A  46      27.836  22.055  37.626  1.00  0.00
ATOM    401  CE1 TYR A  46      26.795  22.556  38.399  1.00  0.00
ATOM    402  CD2 TYR A  46      26.367  21.877  35.746  1.00  0.00
ATOM    403  CE2 TYR A  46      25.318  22.382  36.511  1.00  0.00
ATOM    404  CZ  TYR A  46      25.540  22.718  37.830  1.00  0.00
ATOM    405  OH  TYR A  46      24.516  23.241  38.570  1.00  0.00
ATOM    406  C   TYR A  46      30.969  21.580  34.364  1.00  0.00
ATOM    407  O   TYR A  46      31.527  20.606  34.856  1.00  0.00
ATOM    408  N   GLY A  47      31.305  22.084  33.186  1.00  0.00
ATOM    409  CA  GLY A  47      32.389  21.503  32.426  1.00  0.00
ATOM    410  C   GLY A  47      32.781  22.439  31.311  1.00  0.00
ATOM    411  O   GLY A  47      32.110  23.442  31.090  1.00  0.00
ATOM    412  N   PRO A  48      33.865  22.144  30.588  1.00  0.00
ATOM    413  CD  PRO A  48      34.817  21.038  30.800  1.00  0.00
ATOM    414  CA  PRO A  48      34.304  23.013  29.491  1.00  0.00
ATOM    415  CB  PRO A  48      35.707  22.508  29.199  1.00  0.00
ATOM    416  CG  PRO A  48      35.589  21.034  29.507  1.00  0.00
ATOM    417  C   PRO A  48      33.401  22.972  28.265  1.00  0.00
ATOM    418  O   PRO A  48      32.894  21.920  27.885  1.00  0.00
ATOM    419  N   GLU A  49      33.215  24.134  27.654  1.00  0.00
ATOM    420  CA  GLU A  49      32.377  24.281  26.472  1.00  0.00
ATOM    421  CB  GLU A  49      32.444  25.733  25.997  1.00  0.00
ATOM    422  CG  GLU A  49      31.847  26.726  26.992  1.00  0.00
ATOM    423  CD  GLU A  49      32.559  28.068  26.994  1.00  0.00
ATOM    424  OE1 GLU A  49      33.779  28.097  27.255  1.00  0.00
ATOM    425  OE2 GLU A  49      31.900  29.097  26.748  1.00  0.00
ATOM    426  C   GLU A  49      32.782  23.344  25.332  1.00  0.00
ATOM    427  O   GLU A  49      31.995  23.066  24.426  1.00  0.00
ATOM    428  N   THR A  50      34.014  22.857  25.383  1.00  0.00
ATOM    429  CA  THR A  50      34.511  21.978  24.336  1.00  0.00
ATOM    430  CB  THR A  50      36.044  21.893  24.362  1.00  0.00
ATOM    431  OG1 THR A  50      36.466  21.342  25.617  1.00  0.00
ATOM    432  CG2 THR A  50      36.651  23.272  24.183  1.00  0.00
ATOM    433  C   THR A  50      33.961  20.574  24.490  1.00  0.00
ATOM    434  O   THR A  50      34.053  19.764  23.571  1.00  0.00
ATOM    435  N   GLU A  51      33.375  20.298  25.648  1.00  0.00
ATOM    436  CA  GLU A  51      32.841  18.972  25.912  1.00  0.00
ATOM    437  CB  GLU A  51      33.567  18.367  27.105  1.00  0.00
ATOM    438  CG  GLU A  51      35.040  18.143  26.843  1.00  0.00
ATOM    439  CD  GLU A  51      35.785  17.735  28.086  1.00  0.00
ATOM    440  OE1 GLU A  51      35.283  16.853  28.811  1.00  0.00
ATOM    441  OE2 GLU A  51      36.874  18.292  28.332  1.00  0.00
ATOM    442  C   GLU A  51      31.350  18.929  26.165  1.00  0.00
ATOM    443  O   GLU A  51      30.689  17.955  25.809  1.00  0.00
ATOM    444  N   GLU A  52      30.814  19.973  26.788  1.00  0.00
ATOM    445  CA  GLU A  52      29.389  19.995  27.073  1.00  0.00
ATOM    446  CB  GLU A  52      29.117  19.403  28.454  1.00  0.00
ATOM    447  CG  GLU A  52      29.770  20.143  29.612  1.00  0.00
ATOM    448  CD  GLU A  52      29.142  19.763  30.938  1.00  0.00
ATOM    449  OE1 GLU A  52      27.943  20.086  31.159  1.00  0.00
ATOM    450  OE2 GLU A  52      29.842  19.132  31.755  1.00  0.00
ATOM    451  C   GLU A  52      28.769  21.384  26.977  1.00  0.00
ATOM    452  O   GLU A  52      29.473  22.391  26.845  1.00  0.00
ATOM    453  N   ALA A  53      27.439  21.421  27.031  1.00  0.00
```

```
ATOM    454  CA  ALA A  53      26.700  22.671  26.934  1.00  0.00
ATOM    455  CB  ALA A  53      25.217  22.383  26.655  1.00  0.00
ATOM    456  C   ALA A  53      26.842  23.521  28.188  1.00  0.00
ATOM    457  O   ALA A  53      26.957  23.008  29.311  1.00  0.00
ATOM    458  N   VAL A  54      26.829  24.834  27.986  1.00  0.00
ATOM    459  CA  VAL A  54      26.943  25.768  29.094  1.00  0.00
ATOM    460  CB  VAL A  54      28.328  26.449  29.118  1.00  0.00
ATOM    461  CG1 VAL A  54      29.416  25.398  29.142  1.00  0.00
ATOM    462  CG2 VAL A  54      28.487  27.364  27.913  1.00  0.00
ATOM    463  C   VAL A  54      25.864  26.829  28.967  1.00  0.00
ATOM    464  O   VAL A  54      25.075  26.813  28.023  1.00  0.00
ATOM    465  N   ILE A  55      25.832  27.741  29.932  1.00  0.00
ATOM    466  CA  ILE A  55      24.865  28.823  29.944  1.00  0.00
ATOM    467  CB  ILE A  55      24.173  28.938  31.308  1.00  0.00
ATOM    468  CG2 ILE A  55      23.188  30.092  31.289  1.00  0.00
ATOM    469  CG1 ILE A  55      23.498  27.607  31.652  1.00  0.00
ATOM    470  CD1 ILE A  55      22.792  27.583  32.990  1.00  0.00
ATOM    471  C   ILE A  55      25.592  30.124  29.653  1.00  0.00
ATOM    472  O   ILE A  55      26.576  30.469  30.309  1.00  0.00
ATOM    473  N   GLU A  56      25.097  30.833  28.654  1.00  0.00
ATOM    474  CA  GLU A  56      25.663  32.098  28.236  1.00  0.00
ATOM    475  CB  GLU A  56      25.733  32.132  26.707  1.00  0.00
ATOM    476  CG  GLU A  56      26.025  33.489  26.098  1.00  0.00
ATOM    477  CD  GLU A  56      26.253  33.396  24.601  1.00  0.00
ATOM    478  OE1 GLU A  56      27.388  33.073  24.195  1.00  0.00
ATOM    479  OE2 GLU A  56      25.300  33.619  23.830  1.00  0.00
ATOM    480  C   GLU A  56      24.777  33.216  28.772  1.00  0.00
ATOM    481  O   GLU A  56      23.600  33.319  28.416  1.00  0.00
ATOM    482  N   LEU A  57      25.341  34.035  29.652  1.00  0.00
ATOM    483  CA  LEU A  57      24.610  35.144  30.250  1.00  0.00
ATOM    484  CB  LEU A  57      24.956  35.263  31.742  1.00  0.00
ATOM    485  CG  LEU A  57      24.695  34.013  32.600  1.00  0.00
ATOM    486  CD1 LEU A  57      25.313  34.193  33.978  1.00  0.00
ATOM    487  CD2 LEU A  57      23.197  33.748  32.716  1.00  0.00
ATOM    488  C   LEU A  57      24.987  36.428  29.516  1.00  0.00
ATOM    489  O   LEU A  57      26.153  36.809  29.475  1.00  0.00
ATOM    490  N   THR A  58      23.994  37.083  28.933  1.00  0.00
ATOM    491  CA  THR A  58      24.216  38.318  28.192  1.00  0.00
ATOM    492  CB  THR A  58      23.468  38.303  26.837  1.00  0.00
ATOM    493  OG1 THR A  58      23.989  37.257  26.010  1.00  0.00
ATOM    494  CG2 THR A  58      23.617  39.645  26.124  1.00  0.00
ATOM    495  C   THR A  58      23.733  39.527  28.978  1.00  0.00
ATOM    496  O   THR A  58      22.605  39.548  29.473  1.00  0.00
ATOM    497  N   TYR A  59      24.589  40.529  29.104  1.00  0.00
ATOM    498  CA  TYR A  59      24.208  41.740  29.808  1.00  0.00
ATOM    499  CB  TYR A  59      25.109  41.973  31.026  1.00  0.00
ATOM    500  CG  TYR A  59      24.952  43.359  31.602  1.00  0.00
ATOM    501  CD1 TYR A  59      23.777  43.733  32.249  1.00  0.00
ATOM    502  CE1 TYR A  59      23.607  45.020  32.748  1.00  0.00
ATOM    503  CD2 TYR A  59      25.965  44.311  31.466  1.00  0.00
ATOM    504  CE2 TYR A  59      25.809  45.607  31.961  1.00  0.00
ATOM    505  CZ  TYR A  59      24.624  45.954  32.600  1.00  0.00
ATOM    506  OH  TYR A  59      24.447  47.231  33.078  1.00  0.00
ATOM    507  C   TYR A  59      24.309  42.937  28.875  1.00  0.00
ATOM    508  O   TYR A  59      25.391  43.264  28.404  1.00  0.00
ATOM    509  N   ASN A  60      23.175  43.573  28.596  1.00  0.00
ATOM    510  CA  ASN A  60      23.157  44.749  27.737  1.00  0.00
ATOM    511  CB  ASN A  60      21.805  44.894  27.029  1.00  0.00
ATOM    512  CG  ASN A  60      21.663  43.959  25.843  1.00  0.00
ATOM    513  OD1 ASN A  60      22.391  44.076  24.859  1.00  0.00
ATOM    514  ND2 ASN A  60      20.725  43.025  25.931  1.00  0.00
ATOM    515  C   ASN A  60      23.408  45.966  28.622  1.00  0.00
ATOM    516  O   ASN A  60      22.581  46.314  29.466  1.00  0.00
```

```
ATOM    517  N   TRP A  61      24.566  46.589  28.435  1.00  0.00
ATOM    518  CA  TRP A  61      24.944  47.764  29.209  1.00  0.00
ATOM    519  CB  TRP A  61      26.064  48.522  28.490  1.00  0.00
ATOM    520  CG  TRP A  61      27.336  47.735  28.322  1.00  0.00
ATOM    521  CD2 TRP A  61      28.546  48.181  27.693  1.00  0.00
ATOM    522  CE2 TRP A  61      29.494  47.144  27.829  1.00  0.00
ATOM    523  CE3 TRP A  61      28.920  49.359  27.030  1.00  0.00
ATOM    524  CD1 TRP A  61      27.594  46.480  28.793  1.00  0.00
ATOM    525  NE1 TRP A  61      28.885  46.119  28.503  1.00  0.00
ATOM    526  CZ2 TRP A  61      30.798  47.246  27.328  1.00  0.00
ATOM    527  CZ3 TRP A  61      30.221  49.461  26.530  1.00  0.00
ATOM    528  CH2 TRP A  61      31.142  48.409  26.684  1.00  0.00
ATOM    529  C   TRP A  61      23.758  48.704  29.431  1.00  0.00
ATOM    530  O   TRP A  61      23.074  49.091  28.486  1.00  0.00
ATOM    531  N   GLY A  62      23.510  49.051  30.688  1.00  0.00
ATOM    532  CA  GLY A  62      22.423  49.961  31.007  1.00  0.00
ATOM    533  C   GLY A  62      21.025  49.372  31.091  1.00  0.00
ATOM    534  O   GLY A  62      20.119  50.016  31.625  1.00  0.00
ATOM    535  N   VAL A  63      20.838  48.165  30.561  1.00  0.00
ATOM    536  CA  VAL A  63      19.528  47.511  30.576  1.00  0.00
ATOM    537  CB  VAL A  63      19.241  46.837  29.216  1.00  0.00
ATOM    538  CG1 VAL A  63      17.886  46.148  29.244  1.00  0.00
ATOM    539  CG2 VAL A  63      19.299  47.868  28.108  1.00  0.00
ATOM    540  C   VAL A  63      19.509  46.450  31.672  1.00  0.00
ATOM    541  O   VAL A  63      20.344  45.553  31.687  1.00  0.00
ATOM    542  N   ASP A  64      18.556  46.541  32.589  1.00  0.00
ATOM    543  CA  ASP A  64      18.508  45.575  33.677  1.00  0.00
ATOM    544  CB  ASP A  64      18.970  46.244  34.965  1.00  0.00
ATOM    545  CG  ASP A  64      18.305  47.571  35.192  1.00  0.00
ATOM    546  OD1 ASP A  64      17.058  47.603  35.352  1.00  0.00
ATOM    547  OD2 ASP A  64      19.036  48.586  35.205  1.00  0.00
ATOM    548  C   ASP A  64      17.165  44.911  33.904  1.00  0.00
ATOM    549  O   ASP A  64      16.887  44.428  34.994  1.00  0.00
ATOM    550  N   LYS A  65      16.343  44.862  32.869  1.00  0.00
ATOM    551  CA  LYS A  65      15.030  44.256  32.996  1.00  0.00
ATOM    552  CB  LYS A  65      14.097  45.249  33.699  1.00  0.00
ATOM    553  CG  LYS A  65      12.651  44.820  33.827  1.00  0.00
ATOM    554  CD  LYS A  65      12.469  43.664  34.789  1.00  0.00
ATOM    555  CE  LYS A  65      10.985  43.396  35.018  1.00  0.00
ATOM    556  NZ  LYS A  65      10.226  43.275  33.736  1.00  0.00
ATOM    557  C   LYS A  65      14.485  43.881  31.616  1.00  0.00
ATOM    558  O   LYS A  65      14.630  44.638  30.657  1.00  0.00
ATOM    559  N   TYR A  66      13.875  42.705  31.517  1.00  0.00
ATOM    560  CA  TYR A  66      13.303  42.247  30.249  1.00  0.00
ATOM    561  CB  TYR A  66      14.182  41.165  29.608  1.00  0.00
ATOM    562  CG  TYR A  66      15.590  41.614  29.323  1.00  0.00
ATOM    563  CD1 TYR A  66      16.548  41.653  30.333  1.00  0.00
ATOM    564  CE1 TYR A  66      17.827  42.136  30.085  1.00  0.00
ATOM    565  CD2 TYR A  66      15.950  42.063  28.053  1.00  0.00
ATOM    566  CE2 TYR A  66      17.219  42.545  27.796  1.00  0.00
ATOM    567  CZ  TYR A  66      18.149  42.582  28.811  1.00  0.00
ATOM    568  OH  TYR A  66      19.401  43.086  28.557  1.00  0.00
ATOM    569  C   TYR A  66      11.922  41.661  30.483  1.00  0.00
ATOM    570  O   TYR A  66      11.477  41.535  31.627  1.00  0.00
ATOM    571  N   GLU A  67      11.245  41.310  29.394  1.00  0.00
ATOM    572  CA  GLU A  67       9.939  40.682  29.498  1.00  0.00
ATOM    573  CB  GLU A  67       8.987  41.184  28.410  1.00  0.00
ATOM    574  CG  GLU A  67       8.278  42.499  28.749  1.00  0.00
ATOM    575  CD  GLU A  67       7.629  42.497  30.140  1.00  0.00
ATOM    576  OE1 GLU A  67       6.930  41.512  30.491  1.00  0.00
ATOM    577  OE2 GLU A  67       7.815  43.492  30.885  1.00  0.00
ATOM    578  C   GLU A  67      10.218  39.200  29.315  1.00  0.00
ATOM    579  O   GLU A  67      10.605  38.762  28.236  1.00  0.00
```

```
ATOM    580  N   LEU A  68      10.035  38.435  30.383  1.00  0.00
ATOM    581  CA  LEU A  68      10.296  37.004  30.354  1.00  0.00
ATOM    582  CB  LEU A  68      10.151  36.438  31.770  1.00  0.00
ATOM    583  CG  LEU A  68      10.866  35.119  32.088  1.00  0.00
ATOM    584  CD1 LEU A  68      10.980  34.961  33.597  1.00  0.00
ATOM    585  CD2 LEU A  68      10.113  33.947  31.474  1.00  0.00
ATOM    586  C   LEU A  68       9.395  36.249  29.368  1.00  0.00
ATOM    587  O   LEU A  68       9.825  35.271  28.752  1.00  0.00
ATOM    588  N   GLY A  69       8.154  36.704  29.216  1.00  0.00
ATOM    589  CA  GLY A  69       7.231  36.055  28.296  1.00  0.00
ATOM    590  C   GLY A  69       6.680  34.719  28.768  1.00  0.00
ATOM    591  O   GLY A  69       6.861  34.335  29.919  1.00  0.00
ATOM    592  N   THR A  70       6.013  34.001  27.871  1.00  0.00
ATOM    593  CA  THR A  70       5.428  32.706  28.219  1.00  0.00
ATOM    594  CB  THR A  70       3.891  32.734  28.065  1.00  0.00
ATOM    595  OG1 THR A  70       3.560  32.843  26.677  1.00  0.00
ATOM    596  CG2 THR A  70       3.295  33.926  28.806  1.00  0.00
ATOM    597  C   THR A  70       5.950  31.561  27.354  1.00  0.00
ATOM    598  O   THR A  70       5.570  30.408  27.547  1.00  0.00
ATOM    599  N   ALA A  71       6.813  31.874  26.398  1.00  0.00
ATOM    600  CA  ALA A  71       7.335  30.848  25.512  1.00  0.00
ATOM    601  CB  ALA A  71       7.827  31.478  24.217  1.00  0.00
ATOM    602  C   ALA A  71       8.443  30.027  26.155  1.00  0.00
ATOM    603  O   ALA A  71       8.296  28.822  26.328  1.00  0.00
ATOM    604  N   TYR A  72       9.543  30.684  26.518  1.00  0.00
ATOM    605  CA  TYR A  72      10.682  30.007  27.124  1.00  0.00
ATOM    606  CB  TYR A  72      11.768  31.020  27.509  1.00  0.00
ATOM    607  CG  TYR A  72      12.962  30.389  28.180  1.00  0.00
ATOM    608  CD1 TYR A  72      13.786  29.498  27.487  1.00  0.00
ATOM    609  CE1 TYR A  72      14.855  28.874  28.113  1.00  0.00
ATOM    610  CD2 TYR A  72      13.248  30.642  29.522  1.00  0.00
ATOM    611  CE2 TYR A  72      14.315  30.019  30.160  1.00  0.00
ATOM    612  CZ  TYR A  72      15.112  29.137  29.451  1.00  0.00
ATOM    613  OH  TYR A  72      16.162  28.507  30.077  1.00  0.00
ATOM    614  C   TYR A  72      10.274  29.201  28.351  1.00  0.00
ATOM    615  O   TYR A  72       9.518  29.675  29.197  1.00  0.00
ATOM    616  N   GLY A  73      10.790  27.981  28.436  1.00  0.00
ATOM    617  CA  GLY A  73      10.461  27.115  29.547  1.00  0.00
ATOM    618  C   GLY A  73      11.466  27.088  30.679  1.00  0.00
ATOM    619  O   GLY A  73      11.202  27.609  31.756  1.00  0.00
ATOM    620  N   HIS A  74      12.620  26.476  30.442  1.00  0.00
ATOM    621  CA  HIS A  74      13.648  26.371  31.471  1.00  0.00
ATOM    622  CB  HIS A  74      13.113  25.614  32.689  1.00  0.00
ATOM    623  CG  HIS A  74      12.619  24.235  32.371  1.00  0.00
ATOM    624  CD2 HIS A  74      13.282  23.065  32.196  1.00  0.00
ATOM    625  ND1 HIS A  74      11.286  23.949  32.174  1.00  0.00
ATOM    626  CE1 HIS A  74      11.148  22.664  31.892  1.00  0.00
ATOM    627  NE2 HIS A  74      12.344  22.106  31.898  1.00  0.00
ATOM    628  C   HIS A  74      14.858  25.613  30.972  1.00  0.00
ATOM    629  O   HIS A  74      14.907  25.167  29.825  1.00  0.00
ATOM    630  N   ILE A  75      15.833  25.480  31.863  1.00  0.00
ATOM    631  CA  ILE A  75      17.049  24.729  31.604  1.00  0.00
ATOM    632  CB  ILE A  75      18.316  25.599  31.747  1.00  0.00
ATOM    633  CG2 ILE A  75      19.540  24.699  31.941  1.00  0.00
ATOM    634  CG1 ILE A  75      18.488  26.488  30.507  1.00  0.00
ATOM    635  CD1 ILE A  75      19.492  27.600  30.685  1.00  0.00
ATOM    636  C   ILE A  75      17.046  23.677  32.695  1.00  0.00
ATOM    637  O   ILE A  75      16.772  23.998  33.851  1.00  0.00
ATOM    638  N   ALA A  76      17.326  22.430  32.339  1.00  0.00
ATOM    639  CA  ALA A  76      17.342  21.360  33.329  1.00  0.00
ATOM    640  CB  ALA A  76      16.522  20.159  32.831  1.00  0.00
ATOM    641  C   ALA A  76      18.770  20.934  33.654  1.00  0.00
ATOM    642  O   ALA A  76      19.610  20.788  32.764  1.00  0.00
```

```
ATOM    643  N   LEU A  77      19.037  20.748  34.944  1.00  0.00
ATOM    644  CA  LEU A  77      20.351  20.340  35.423  1.00  0.00
ATOM    645  CB  LEU A  77      20.957  21.427  36.306  1.00  0.00
ATOM    646  CG  LEU A  77      21.026  22.846  35.736  1.00  0.00
ATOM    647  CD1 LEU A  77      21.630  23.775  36.785  1.00  0.00
ATOM    648  CD2 LEU A  77      21.850  22.868  34.455  1.00  0.00
ATOM    649  C   LEU A  77      20.204  19.065  36.234  1.00  0.00
ATOM    650  O   LEU A  77      19.199  18.872  36.915  1.00  0.00
ATOM    651  N   SER A  78      21.210  18.200  36.168  1.00  0.00
ATOM    652  CA  SER A  78      21.166  16.946  36.909  1.00  0.00
ATOM    653  CB  SER A  78      22.012  15.889  36.198  1.00  0.00
ATOM    654  OG  SER A  78      23.392  16.173  36.331  1.00  0.00
ATOM    655  C   SER A  78      21.688  17.139  38.333  1.00  0.00
ATOM    656  O   SER A  78      22.638  17.874  38.546  1.00  0.00
ATOM    657  N   VAL A  79      21.046  16.483  39.298  1.00  0.00
ATOM    658  CA  VAL A  79      21.447  16.540  40.707  1.00  0.00
ATOM    659  CB  VAL A  79      20.469  17.370  41.582  1.00  0.00
ATOM    660  CG1 VAL A  79      20.434  18.816  41.117  1.00  0.00
ATOM    661  CG2 VAL A  79      19.092  16.749  41.547  1.00  0.00
ATOM    662  C   VAL A  79      21.430  15.104  41.218  1.00  0.00
ATOM    663  O   VAL A  79      20.816  14.229  40.603  1.00  0.00
ATOM    664  N   ASP A  80      22.084  14.849  42.341  1.00  0.00
ATOM    665  CA  ASP A  80      22.120  13.491  42.860  1.00  0.00
ATOM    666  CB  ASP A  80      23.419  13.275  43.639  1.00  0.00
ATOM    667  CG  ASP A  80      24.644  13.346  42.735  1.00  0.00
ATOM    668  OD1 ASP A  80      24.592  12.761  41.628  1.00  0.00
ATOM    669  OD2 ASP A  80      25.656  13.978  43.116  1.00  0.00
ATOM    670  C   ASP A  80      20.909  13.068  43.688  1.00  0.00
ATOM    671  O   ASP A  80      20.591  11.874  43.762  1.00  0.00
ATOM    672  N   ASN A  81      20.207  14.034  44.282  1.00  0.00
ATOM    673  CA  ASN A  81      19.046  13.696  45.100  1.00  0.00
ATOM    674  CB  ASN A  81      19.505  13.443  46.541  1.00  0.00
ATOM    675  CG  ASN A  81      18.352  13.251  47.490  1.00  0.00
ATOM    676  OD1 ASN A  81      17.938  14.184  48.174  1.00  0.00
ATOM    677  ND2 ASN A  81      17.809  12.035  47.529  1.00  0.00
ATOM    678  C   ASN A  81      17.862  14.674  45.083  1.00  0.00
ATOM    679  O   ASN A  81      16.770  14.340  45.551  1.00  0.00
ATOM    680  N   ALA A  82      18.053  15.870  44.545  1.00  0.00
ATOM    681  CA  ALA A  82      16.964  16.846  44.474  1.00  0.00
ATOM    682  CB  ALA A  82      15.788  16.267  43.712  1.00  0.00
ATOM    683  C   ALA A  82      16.502  17.342  45.840  1.00  0.00
ATOM    684  O   ALA A  82      16.590  18.543  46.131  1.00  0.00
ATOM    685  N   ALA A  83      15.988  16.438  46.674  1.00  0.00
ATOM    686  CA  ALA A  83      15.532  16.829  48.006  1.00  0.00
ATOM    687  CB  ALA A  83      15.110  15.602  48.816  1.00  0.00
ATOM    688  C   ALA A  83      16.703  17.522  48.681  1.00  0.00
ATOM    689  O   ALA A  83      16.564  18.610  49.232  1.00  0.00
ATOM    690  N   GLU A  84      17.862  16.877  48.615  1.00  0.00
ATOM    691  CA  GLU A  84      19.083  17.404  49.198  1.00  0.00
ATOM    692  CB  GLU A  84      20.146  16.311  49.214  1.00  0.00
ATOM    693  CG  GLU A  84      21.521  16.770  49.656  1.00  0.00
ATOM    694  CD  GLU A  84      22.595  16.324  48.682  1.00  0.00
ATOM    695  OE1 GLU A  84      22.690  16.936  47.587  1.00  0.00
ATOM    696  OE2 GLU A  84      23.335  15.360  49.003  1.00  0.00
ATOM    697  C   GLU A  84      19.600  18.617  48.429  1.00  0.00
ATOM    698  O   GLU A  84      20.049  19.595  49.027  1.00  0.00
ATOM    699  N   ALA A  85      19.543  18.550  47.102  1.00  0.00
ATOM    700  CA  ALA A  85      20.007  19.654  46.272  1.00  0.00
ATOM    701  CB  ALA A  85      19.784  19.337  44.798  1.00  0.00
ATOM    702  C   ALA A  85      19.290  20.949  46.642  1.00  0.00
ATOM    703  O   ALA A  85      19.898  22.011  46.668  1.00  0.00
ATOM    704  N   CYS A  86      17.996  20.861  46.928  1.00  0.00
ATOM    705  CA  CYS A  86      17.229  22.048  47.299  1.00  0.00
```

```
ATOM    706  CB  CYS A  86      15.736  21.734  47.361  1.00  0.00
ATOM    707  SG  CYS A  86      14.964  21.396  45.776  1.00  0.00
ATOM    708  C   CYS A  86      17.666  22.594  48.652  1.00  0.00
ATOM    709  O   CYS A  86      17.809  23.809  48.824  1.00  0.00
ATOM    710  N   GLU A  87      17.873  21.705  49.617  1.00  0.00
ATOM    711  CA  GLU A  87      18.274  22.153  50.941  1.00  0.00
ATOM    712  CB  GLU A  87      18.402  20.953  51.888  1.00  0.00
ATOM    713  CG  GLU A  87      17.149  20.063  51.895  1.00  0.00
ATOM    714  CD  GLU A  87      17.085  19.092  53.079  1.00  0.00
ATOM    715  OE1 GLU A  87      16.272  18.140  53.036  1.00  0.00
ATOM    716  OE2 GLU A  87      17.840  19.279  54.059  1.00  0.00
ATOM    717  C   GLU A  87      19.588  22.912  50.800  1.00  0.00
ATOM    718  O   GLU A  87      19.749  24.011  51.330  1.00  0.00
ATOM    719  N   LYS A  88      20.511  22.333  50.048  1.00  0.00
ATOM    720  CA  LYS A  88      21.807  22.945  49.791  1.00  0.00
ATOM    721  CB  LYS A  88      22.627  21.985  48.929  1.00  0.00
ATOM    722  CG  LYS A  88      23.976  22.491  48.445  1.00  0.00
ATOM    723  CD  LYS A  88      24.721  21.362  47.720  1.00  0.00
ATOM    724  CE  LYS A  88      23.826  20.669  46.679  1.00  0.00
ATOM    725  NZ  LYS A  88      24.508  19.558  45.926  1.00  0.00
ATOM    726  C   LYS A  88      21.611  24.290  49.078  1.00  0.00
ATOM    727  O   LYS A  88      22.254  25.293  49.413  1.00  0.00
ATOM    728  N   ILE A  89      20.715  24.311  48.096  1.00  0.00
ATOM    729  CA  ILE A  89      20.443  25.536  47.358  1.00  0.00
ATOM    730  CB  ILE A  89      19.446  25.265  46.195  1.00  0.00
ATOM    731  CG2 ILE A  89      18.897  26.572  45.640  1.00  0.00
ATOM    732  CG1 ILE A  89      20.162  24.462  45.100  1.00  0.00
ATOM    733  CD1 ILE A  89      19.318  24.192  43.872  1.00  0.00
ATOM    734  C   ILE A  89      19.906  26.617  48.307  1.00  0.00
ATOM    735  O   ILE A  89      20.354  27.767  48.258  1.00  0.00
ATOM    736  N   ARG A  90      18.971  26.242  49.182  1.00  0.00
ATOM    737  CA  ARG A  90      18.413  27.190  50.146  1.00  0.00
ATOM    738  CB  ARG A  90      17.279  26.547  50.957  1.00  0.00
ATOM    739  CG  ARG A  90      15.932  26.628  50.266  1.00  0.00
ATOM    740  CD  ARG A  90      14.809  26.117  51.149  1.00  0.00
ATOM    741  NE  ARG A  90      14.783  24.657  51.159  1.00  0.00
ATOM    742  CZ  ARG A  90      13.895  23.922  50.498  1.00  0.00
ATOM    743  NH1 ARG A  90      12.947  24.513  49.781  1.00  0.00
ATOM    744  NH2 ARG A  90      13.968  22.596  50.542  1.00  0.00
ATOM    745  C   ARG A  90      19.485  27.719  51.088  1.00  0.00
ATOM    746  O   ARG A  90      19.542  28.923  51.357  1.00  0.00
ATOM    747  N   GLN A  91      20.320  26.817  51.594  1.00  0.00
ATOM    748  CA  GLN A  91      21.417  27.184  52.486  1.00  0.00
ATOM    749  CB  GLN A  91      22.344  25.983  52.705  1.00  0.00
ATOM    750  CG  GLN A  91      22.236  25.294  54.058  1.00  0.00
ATOM    751  CD  GLN A  91      23.065  25.984  55.122  1.00  0.00
ATOM    752  OE1 GLN A  91      22.766  27.111  55.531  1.00  0.00
ATOM    753  NE2 GLN A  91      24.122  25.319  55.570  1.00  0.00
ATOM    754  C   GLN A  91      22.224  28.300  51.829  1.00  0.00
ATOM    755  O   GLN A  91      22.555  29.302  52.467  1.00  0.00
ATOM    756  N   ASN A  92      22.540  28.105  50.549  1.00  0.00
ATOM    757  CA  ASN A  92      23.319  29.070  49.782  1.00  0.00
ATOM    758  CB  ASN A  92      23.829  28.422  48.491  1.00  0.00
ATOM    759  CG  ASN A  92      24.805  27.299  48.756  1.00  0.00
ATOM    760  OD1 ASN A  92      25.668  27.409  49.626  1.00  0.00
ATOM    761  ND2 ASN A  92      24.682  26.211  48.001  1.00  0.00
ATOM    762  C   ASN A  92      22.547  30.350  49.456  1.00  0.00
ATOM    763  O   ASN A  92      23.081  31.267  48.831  1.00  0.00
ATOM    764  N   GLY A  93      21.290  30.409  49.882  1.00  0.00
ATOM    765  CA  GLY A  93      20.485  31.596  49.642  1.00  0.00
ATOM    766  C   GLY A  93      19.769  31.642  48.308  1.00  0.00
ATOM    767  O   GLY A  93      19.458  32.726  47.809  1.00  0.00
ATOM    768  N   GLY A  94      19.512  30.472  47.728  1.00  0.00
```

```
ATOM    769  CA  GLY A  94      18.821  30.405  46.453  1.00  0.00
ATOM    770  C   GLY A  94      17.326  30.284  46.670  1.00  0.00
ATOM    771  O   GLY A  94      16.879  29.858  47.738  1.00  0.00
ATOM    772  N   ASN A  95      16.544  30.652  45.662  1.00  0.00
ATOM    773  CA  ASN A  95      15.087  30.580  45.772  1.00  0.00
ATOM    774  CB  ASN A  95      14.436  31.783  45.084  1.00  0.00
ATOM    775  CG  ASN A  95      14.814  33.104  45.731  1.00  0.00
ATOM    776  OD1 ASN A  95      14.862  33.217  46.955  1.00  0.00
ATOM    777  ND2 ASN A  95      15.070  34.115  44.908  1.00  0.00
ATOM    778  C   ASN A  95      14.496  29.310  45.185  1.00  0.00
ATOM    779  O   ASN A  95      14.408  29.165  43.960  1.00  0.00
ATOM    780  N   VAL A  96      14.091  28.391  46.056  1.00  0.00
ATOM    781  CA  VAL A  96      13.473  27.142  45.619  1.00  0.00
ATOM    782  CB  VAL A  96      13.756  25.986  46.606  1.00  0.00
ATOM    783  CG1 VAL A  96      12.942  24.764  46.223  1.00  0.00
ATOM    784  CG2 VAL A  96      15.236  25.636  46.591  1.00  0.00
ATOM    785  C   VAL A  96      11.972  27.403  45.540  1.00  0.00
ATOM    786  O   VAL A  96      11.319  27.659  46.548  1.00  0.00
ATOM    787  N   THR A  97      11.436  27.334  44.327  1.00  0.00
ATOM    788  CA  THR A  97      10.026  27.602  44.069  1.00  0.00
ATOM    789  CB  THR A  97       9.902  28.499  42.840  1.00  0.00
ATOM    790  OG1 THR A  97      10.455  27.821  41.705  1.00  0.00
ATOM    791  CG2 THR A  97      10.682  29.780  43.052  1.00  0.00
ATOM    792  C   THR A  97       9.145  26.369  43.856  1.00  0.00
ATOM    793  O   THR A  97       7.926  26.484  43.704  1.00  0.00
ATOM    794  N   ARG A  98       9.754  25.191  43.830  1.00  0.00
ATOM    795  CA  ARG A  98       8.988  23.965  43.626  1.00  0.00
ATOM    796  CB  ARG A  98       8.783  23.736  42.120  1.00  0.00
ATOM    797  CG  ARG A  98       7.728  22.697  41.739  1.00  0.00
ATOM    798  CD  ARG A  98       8.182  21.276  42.036  1.00  0.00
ATOM    799  NE  ARG A  98       7.308  20.292  41.414  1.00  0.00
ATOM    800  CZ  ARG A  98       7.415  18.981  41.595  1.00  0.00
ATOM    801  NH1 ARG A  98       8.361  18.497  42.383  1.00  0.00
ATOM    802  NH2 ARG A  98       6.573  18.155  40.998  1.00  0.00
ATOM    803  C   ARG A  98       9.789  22.835  44.258  1.00  0.00
ATOM    804  O   ARG A  98      10.721  22.316  43.655  1.00  0.00
ATOM    805  N   GLU A  99       9.423  22.480  45.488  1.00  0.00
ATOM    806  CA  GLU A  99      10.105  21.437  46.252  1.00  0.00
ATOM    807  CB  GLU A  99       9.384  21.202  47.584  1.00  0.00
ATOM    808  CG  GLU A  99       9.302  22.414  48.485  1.00  0.00
ATOM    809  CD  GLU A  99      10.660  23.003  48.818  1.00  0.00
ATOM    810  OE1 GLU A  99      11.537  22.261  49.306  1.00  0.00
ATOM    811  OE2 GLU A  99      10.856  24.219  48.599  1.00  0.00
ATOM    812  C   GLU A  99      10.244  20.108  45.520  1.00  0.00
ATOM    813  O   GLU A  99       9.387  19.730  44.721  1.00  0.00
ATOM    814  N   ALA A 100      11.332  19.402  45.809  1.00  0.00
ATOM    815  CA  ALA A 100      11.610  18.113  45.190  1.00  0.00
ATOM    816  CB  ALA A 100      12.837  17.480  45.817  1.00  0.00
ATOM    817  C   ALA A 100      10.428  17.179  45.333  1.00  0.00
ATOM    818  O   ALA A 100       9.800  17.109  46.386  1.00  0.00
ATOM    819  N   GLY A 101      10.134  16.454  44.263  1.00  0.00
ATOM    820  CA  GLY A 101       9.033  15.517  44.281  1.00  0.00
ATOM    821  C   GLY A 101       8.818  14.979  42.883  1.00  0.00
ATOM    822  O   GLY A 101       9.325  15.558  41.917  1.00  0.00
ATOM    823  N   PRO A 102       8.081  13.873  42.733  1.00  0.00
ATOM    824  CD  PRO A 102       7.312  13.098  43.721  1.00  0.00
ATOM    825  CA  PRO A 102       7.874  13.360  41.382  1.00  0.00
ATOM    826  CB  PRO A 102       7.136  12.046  41.618  1.00  0.00
ATOM    827  CG  PRO A 102       6.321  12.346  42.836  1.00  0.00
ATOM    828  C   PRO A 102       7.057  14.368  40.598  1.00  0.00
ATOM    829  O   PRO A 102       6.127  14.975  41.129  1.00  0.00
ATOM    830  N   VAL A 103       7.431  14.554  39.337  1.00  0.00
ATOM    831  CA  VAL A 103       6.748  15.484  38.460  1.00  0.00
```

```
ATOM    832  CB  VAL A 103       7.314  15.374  37.020  1.00  0.00
ATOM    833  CG1 VAL A 103       7.181  13.938  36.522  1.00  0.00
ATOM    834  CG2 VAL A 103       6.602  16.346  36.092  1.00  0.00
ATOM    835  C   VAL A 103       5.257  15.168  38.463  1.00  0.00
ATOM    836  O   VAL A 103       4.860  13.996  38.469  1.00  0.00
ATOM    837  N   LYS A 104       4.433  16.210  38.492  1.00  0.00
ATOM    838  CA  LYS A 104       2.984  16.035  38.467  1.00  0.00
ATOM    839  CB  LYS A 104       2.276  17.396  38.477  1.00  0.00
ATOM    840  CG  LYS A 104       2.442  18.200  39.760  1.00  0.00
ATOM    841  CD  LYS A 104       1.651  19.503  39.694  1.00  0.00
ATOM    842  CE  LYS A 104       1.646  20.240  41.035  1.00  0.00
ATOM    843  NZ  LYS A 104       2.988  20.769  41.424  1.00  0.00
ATOM    844  C   LYS A 104       2.626  15.288  37.182  1.00  0.00
ATOM    845  O   LYS A 104       2.993  15.714  36.086  1.00  0.00
ATOM    846  N   GLY A 105       1.923  14.167  37.317  1.00  0.00
ATOM    847  CA  GLY A 105       1.541  13.396  36.145  1.00  0.00
ATOM    848  C   GLY A 105       2.590  12.394  35.701  1.00  0.00
ATOM    849  O   GLY A 105       2.438  11.763  34.660  1.00  0.00
ATOM    850  N   GLY A 106       3.652  12.242  36.488  1.00  0.00
ATOM    851  CA  GLY A 106       4.704  11.296  36.152  1.00  0.00
ATOM    852  C   GLY A 106       5.405  10.773  37.391  1.00  0.00
ATOM    853  O   GLY A 106       4.890  10.914  38.496  1.00  0.00
ATOM    854  N   THR A 107       6.584  10.181  37.219  1.00  0.00
ATOM    855  CA  THR A 107       7.335   9.630  38.347  1.00  0.00
ATOM    856  CB  THR A 107       7.547   8.109  38.205  1.00  0.00
ATOM    857  OG1 THR A 107       8.362   7.851  37.054  1.00  0.00
ATOM    858  CG2 THR A 107       6.221   7.396  38.049  1.00  0.00
ATOM    859  C   THR A 107       8.710  10.261  38.511  1.00  0.00
ATOM    860  O   THR A 107       9.315  10.169  39.581  1.00  0.00
ATOM    861  N   THR A 108       9.210  10.889  37.452  1.00  0.00
ATOM    862  CA  THR A 108      10.519  11.531  37.493  1.00  0.00
ATOM    863  CB  THR A 108      10.801  12.247  36.152  1.00  0.00
ATOM    864  OG1 THR A 108      10.773  11.280  35.092  1.00  0.00
ATOM    865  CG2 THR A 108      12.169  12.939  36.168  1.00  0.00
ATOM    866  C   THR A 108      10.579  12.535  38.649  1.00  0.00
ATOM    867  O   THR A 108       9.682  13.358  38.812  1.00  0.00
ATOM    868  N   VAL A 109      11.634  12.459  39.454  1.00  0.00
ATOM    869  CA  VAL A 109      11.798  13.355  40.594  1.00  0.00
ATOM    870  CB  VAL A 109      12.639  12.687  41.701  1.00  0.00
ATOM    871  CG1 VAL A 109      12.857  13.666  42.850  1.00  0.00
ATOM    872  CG2 VAL A 109      11.922  11.437  42.211  1.00  0.00
ATOM    873  C   VAL A 109      12.458  14.671  40.190  1.00  0.00
ATOM    874  O   VAL A 109      13.614  14.696  39.788  1.00  0.00
ATOM    875  N   ILE A 110      11.720  15.770  40.327  1.00  0.00
ATOM    876  CA  ILE A 110      12.230  17.082  39.945  1.00  0.00
ATOM    877  CB  ILE A 110      11.530  17.582  38.668  1.00  0.00
ATOM    878  CG2 ILE A 110      11.628  16.528  37.580  1.00  0.00
ATOM    879  CG1 ILE A 110      10.063  17.900  38.974  1.00  0.00
ATOM    880  CD1 ILE A 110       9.315  18.549  37.813  1.00  0.00
ATOM    881  C   ILE A 110      12.051  18.165  41.013  1.00  0.00
ATOM    882  O   ILE A 110      11.481  17.928  42.074  1.00  0.00
ATOM    883  N   ALA A 111      12.556  19.354  40.704  1.00  0.00
ATOM    884  CA  ALA A 111      12.449  20.518  41.574  1.00  0.00
ATOM    885  CB  ALA A 111      13.467  20.443  42.706  1.00  0.00
ATOM    886  C   ALA A 111      12.704  21.744  40.703  1.00  0.00
ATOM    887  O   ALA A 111      13.224  21.624  39.591  1.00  0.00
ATOM    888  N   PHE A 112      12.296  22.913  41.181  1.00  0.00
ATOM    889  CA  PHE A 112      12.512  24.156  40.434  1.00  0.00
ATOM    890  CB  PHE A 112      11.208  24.722  39.852  1.00  0.00
ATOM    891  CG  PHE A 112      10.756  24.070  38.574  1.00  0.00
ATOM    892  CD1 PHE A 112       9.840  23.032  38.592  1.00  0.00
ATOM    893  CD2 PHE A 112      11.223  24.521  37.346  1.00  0.00
ATOM    894  CE1 PHE A 112       9.391  22.452  37.409  1.00  0.00
```

```
ATOM    895  CE2 PHE A 112      10.778  23.947  36.161  1.00  0.00
ATOM    896  CZ  PHE A 112       9.861  22.911  36.193  1.00  0.00
ATOM    897  C   PHE A 112      13.089  25.216  41.356  1.00  0.00
ATOM    898  O   PHE A 112      12.709  25.315  42.520  1.00  0.00
ATOM    899  N   VAL A 113      14.013  25.997  40.820  1.00  0.00
ATOM    900  CA  VAL A 113      14.623  27.100  41.539  1.00  0.00
ATOM    901  CB  VAL A 113      16.090  26.800  41.934  1.00  0.00
ATOM    902  CG1 VAL A 113      16.137  25.611  42.893  1.00  0.00
ATOM    903  CG2 VAL A 113      16.913  26.504  40.711  1.00  0.00
ATOM    904  C   VAL A 113      14.580  28.250  40.543  1.00  0.00
ATOM    905  O   VAL A 113      14.550  28.017  39.339  1.00  0.00
ATOM    906  N   GLU A 114      14.542  29.482  41.031  1.00  0.00
ATOM    907  CA  GLU A 114      14.515  30.618  40.131  1.00  0.00
ATOM    908  CB  GLU A 114      13.259  31.459  40.347  1.00  0.00
ATOM    909  CG  GLU A 114      11.970  30.718  40.059  1.00  0.00
ATOM    910  CD  GLU A 114      10.850  31.645  39.627  1.00  0.00
ATOM    911  OE1 GLU A 114      10.902  32.855  39.958  1.00  0.00
ATOM    912  OE2 GLU A 114       9.909  31.155  38.964  1.00  0.00
ATOM    913  C   GLU A 114      15.745  31.472  40.344  1.00  0.00
ATOM    914  O   GLU A 114      16.140  31.733  41.481  1.00  0.00
ATOM    915  N   ASP A 115      16.353  31.901  39.248  1.00  0.00
ATOM    916  CA  ASP A 115      17.540  32.724  39.342  1.00  0.00
ATOM    917  CB  ASP A 115      18.391  32.571  38.065  1.00  0.00
ATOM    918  CG  ASP A 115      17.822  33.304  36.865  1.00  0.00
ATOM    919  OD1 ASP A 115      18.485  33.263  35.804  1.00  0.00
ATOM    920  OD2 ASP A 115      16.736  33.913  36.961  1.00  0.00
ATOM    921  C   ASP A 115      17.136  34.180  39.609  1.00  0.00
ATOM    922  O   ASP A 115      15.952  34.513  39.602  1.00  0.00
ATOM    923  N   PRO A 116      18.114  35.062  39.850  1.00  0.00
ATOM    924  CD  PRO A 116      19.563  34.808  39.804  1.00  0.00
ATOM    925  CA  PRO A 116      17.854  36.477  40.129  1.00  0.00
ATOM    926  CB  PRO A 116      19.246  37.094  40.059  1.00  0.00
ATOM    927  CG  PRO A 116      20.113  35.985  40.557  1.00  0.00
ATOM    928  C   PRO A 116      16.877  37.172  39.182  1.00  0.00
ATOM    929  O   PRO A 116      16.184  38.105  39.578  1.00  0.00
ATOM    930  N   ASP A 117      16.833  36.720  37.932  1.00  0.00
ATOM    931  CA  ASP A 117      15.964  37.320  36.928  1.00  0.00
ATOM    932  CB  ASP A 117      16.718  37.427  35.604  1.00  0.00
ATOM    933  CG  ASP A 117      17.850  38.422  35.669  1.00  0.00
ATOM    934  OD1 ASP A 117      17.564  39.625  35.763  1.00  0.00
ATOM    935  OD2 ASP A 117      19.024  38.007  35.640  1.00  0.00
ATOM    936  C   ASP A 117      14.631  36.608  36.705  1.00  0.00
ATOM    937  O   ASP A 117      13.911  36.929  35.763  1.00  0.00
ATOM    938  N   GLY A 118      14.310  35.642  37.562  1.00  0.00
ATOM    939  CA  GLY A 118      13.052  34.931  37.432  1.00  0.00
ATOM    940  C   GLY A 118      13.083  33.707  36.537  1.00  0.00
ATOM    941  O   GLY A 118      12.091  32.984  36.445  1.00  0.00
ATOM    942  N   TYR A 119      14.209  33.470  35.872  1.00  0.00
ATOM    943  CA  TYR A 119      14.326  32.312  35.000  1.00  0.00
ATOM    944  CB  TYR A 119      15.683  32.293  34.298  1.00  0.00
ATOM    945  CG  TYR A 119      15.851  33.324  33.214  1.00  0.00
ATOM    946  CD1 TYR A 119      15.097  33.262  32.040  1.00  0.00
ATOM    947  CE1 TYR A 119      15.318  34.166  30.993  1.00  0.00
ATOM    948  CD2 TYR A 119      16.819  34.319  33.327  1.00  0.00
ATOM    949  CE2 TYR A 119      17.049  35.227  32.294  1.00  0.00
ATOM    950  CZ  TYR A 119      16.306  35.144  31.131  1.00  0.00
ATOM    951  OH  TYR A 119      16.579  36.008  30.097  1.00  0.00
ATOM    952  C   TYR A 119      14.188  31.054  35.841  1.00  0.00
ATOM    953  O   TYR A 119      14.677  30.996  36.971  1.00  0.00
ATOM    954  N   LYS A 120      13.526  30.046  35.290  1.00  0.00
ATOM    955  CA  LYS A 120      13.335  28.798  36.016  1.00  0.00
ATOM    956  CB  LYS A 120      11.948  28.224  35.714  1.00  0.00
ATOM    957  CG  LYS A 120      10.830  29.106  36.236  1.00  0.00
```

```
ATOM    958  CD  LYS A 120       9.461  28.607  35.817  1.00  0.00
ATOM    959  CE  LYS A 120       9.127  27.294  36.478  1.00  0.00
ATOM    960  NZ  LYS A 120       7.914  26.688  35.863  1.00  0.00
ATOM    961  C   LYS A 120      14.412  27.775  35.698  1.00  0.00
ATOM    962  O   LYS A 120      14.777  27.562  34.540  1.00  0.00
ATOM    963  N   ILE A 121      14.942  27.165  36.748  1.00  0.00
ATOM    964  CA  ILE A 121      15.968  26.145  36.602  1.00  0.00
ATOM    965  CB  ILE A 121      17.249  26.496  37.383  1.00  0.00
ATOM    966  CG2 ILE A 121      18.355  25.523  37.016  1.00  0.00
ATOM    967  CG1 ILE A 121      17.680  27.933  37.078  1.00  0.00
ATOM    968  CD1 ILE A 121      17.949  28.181  35.630  1.00  0.00
ATOM    969  C   ILE A 121      15.394  24.865  37.181  1.00  0.00
ATOM    970  O   ILE A 121      15.076  24.805  38.370  1.00  0.00
ATOM    971  N   GLU A 122      15.245  23.847  36.339  1.00  0.00
ATOM    972  CA  GLU A 122      14.700  22.579  36.796  1.00  0.00
ATOM    973  CB  GLU A 122      13.879  21.917  35.697  1.00  0.00
ATOM    974  CG  GLU A 122      13.094  20.712  36.183  1.00  0.00
ATOM    975  CD  GLU A 122      12.500  19.918  35.048  1.00  0.00
ATOM    976  OE1 GLU A 122      11.435  19.306  35.254  1.00  0.00
ATOM    977  OE2 GLU A 122      13.096  19.896  33.953  1.00  0.00
ATOM    978  C   GLU A 122      15.791  21.616  37.222  1.00  0.00
ATOM    979  O   GLU A 122      16.803  21.474  36.546  1.00  0.00
ATOM    980  N   LEU A 123      15.578  20.958  38.352  1.00  0.00
ATOM    981  CA  LEU A 123      16.534  19.982  38.850  1.00  0.00
ATOM    982  CB  LEU A 123      16.803  20.206  40.341  1.00  0.00
ATOM    983  CG  LEU A 123      17.272  21.604  40.763  1.00  0.00
ATOM    984  CD1 LEU A 123      17.470  21.636  42.283  1.00  0.00
ATOM    985  CD2 LEU A 123      18.556  21.962  40.047  1.00  0.00
ATOM    986  C   LEU A 123      15.939  18.596  38.622  1.00  0.00
ATOM    987  O   LEU A 123      14.766  18.362  38.898  1.00  0.00
ATOM    988  N   ILE A 124      16.742  17.687  38.089  1.00  0.00
ATOM    989  CA  ILE A 124      16.294  16.325  37.835  1.00  0.00
ATOM    990  CB  ILE A 124      16.122  16.067  36.318  1.00  0.00
ATOM    991  CG2 ILE A 124      15.738  14.614  36.069  1.00  0.00
ATOM    992  CG1 ILE A 124      15.045  16.998  35.752  1.00  0.00
ATOM    993  CD1 ILE A 124      14.912  16.952  34.248  1.00  0.00
ATOM    994  C   ILE A 124      17.322  15.346  38.395  1.00  0.00
ATOM    995  O   ILE A 124      18.508  15.399  38.051  1.00  0.00
ATOM    996  N   GLU A 125      16.872  14.457  39.271  1.00  0.00
ATOM    997  CA  GLU A 125      17.784  13.479  39.847  1.00  0.00
ATOM    998  CB  GLU A 125      17.040  12.534  40.777  1.00  0.00
ATOM    999  CG  GLU A 125      16.181  13.242  41.794  1.00  0.00
ATOM   1000  CD  GLU A 125      15.825  12.346  42.950  1.00  0.00
ATOM   1001  OE1 GLU A 125      15.634  11.135  42.724  1.00  0.00
ATOM   1002  OE2 GLU A 125      15.731  12.858  44.081  1.00  0.00
ATOM   1003  C   GLU A 125      18.406  12.698  38.705  1.00  0.00
ATOM   1004  O   GLU A 125      17.701  12.221  37.807  1.00  0.00
ATOM   1005  N   GLU A 126      19.732  12.591  38.738  1.00  0.00
ATOM   1006  CA  GLU A 126      20.475  11.885  37.706  1.00  0.00
ATOM   1007  CB  GLU A 126      21.929  11.672  38.161  1.00  0.00
ATOM   1008  CG  GLU A 126      22.104  11.488  39.680  1.00  0.00
ATOM   1009  CD  GLU A 126      21.530  10.177  40.210  1.00  0.00
ATOM   1010  OE1 GLU A 126      22.088   9.104  39.870  1.00  0.00
ATOM   1011  OE2 GLU A 126      20.523  10.212  40.965  1.00  0.00
ATOM   1012  C   GLU A 126      19.825  10.552  37.370  1.00  0.00
ATOM   1013  O   GLU A 126      19.639  10.220  36.195  1.00  0.00
ATOM   1014  N   LYS A 127      19.475   9.800  38.410  1.00  0.00
ATOM   1015  CA  LYS A 127      18.848   8.492  38.246  1.00  0.00
ATOM   1016  CB  LYS A 127      18.264   8.015  39.587  1.00  0.00
ATOM   1017  CG  LYS A 127      17.260   6.883  39.448  1.00  0.00
ATOM   1018  CD  LYS A 127      16.673   6.429  40.782  1.00  0.00
ATOM   1019  CE  LYS A 127      17.658   5.568  41.553  1.00  0.00
ATOM   1020  NZ  LYS A 127      16.985   4.803  42.660  1.00  0.00
```

```
ATOM   1021  C   LYS A 127      17.754   8.491  37.186  1.00  0.00
ATOM   1022  O   LYS A 127      17.524   7.468  36.543  1.00  0.00
ATOM   1023  N   ASP A 128      17.088   9.630  36.996  1.00  0.00
ATOM   1024  CA  ASP A 128      15.993   9.711  36.025  1.00  0.00
ATOM   1025  CB  ASP A 128      14.757  10.341  36.678  1.00  0.00
ATOM   1026  CG  ASP A 128      14.260   9.547  37.863  1.00  0.00
ATOM   1027  OD1 ASP A 128      14.113   8.312  37.727  1.00  0.00
ATOM   1028  OD2 ASP A 128      14.010  10.153  38.932  1.00  0.00
ATOM   1029  C   ASP A 128      16.317  10.462  34.736  1.00  0.00
ATOM   1030  O   ASP A 128      15.445  10.660  33.898  1.00  0.00
ATOM   1031  N   ALA A 129      17.569  10.867  34.581  1.00  0.00
ATOM   1032  CA  ALA A 129      18.013  11.591  33.398  1.00  0.00
ATOM   1033  CB  ALA A 129      19.517  11.844  33.485  1.00  0.00
ATOM   1034  C   ALA A 129      17.687  10.826  32.111  1.00  0.00
ATOM   1035  O   ALA A 129      18.150   9.694  31.913  1.00  0.00
ATOM   1036  N   GLY A 130      16.881  11.440  31.249  1.00  0.00
ATOM   1037  CA  GLY A 130      16.514  10.810  29.988  1.00  0.00
ATOM   1038  C   GLY A 130      15.371   9.821  30.098  1.00  0.00
ATOM   1039  O   GLY A 130      15.238   8.914  29.268  1.00  0.00
ATOM   1040  N   ARG A 131      14.548   9.986  31.127  1.00  0.00
ATOM   1041  CA  ARG A 131      13.404   9.101  31.351  1.00  0.00
ATOM   1042  CB  ARG A 131      13.305   8.702  32.823  1.00  0.00
ATOM   1043  CG  ARG A 131      14.461   7.881  33.347  1.00  0.00
ATOM   1044  CD  ARG A 131      14.555   6.512  32.698  1.00  0.00
ATOM   1045  NE  ARG A 131      15.481   5.673  33.454  1.00  0.00
ATOM   1046  CZ  ARG A 131      15.297   5.330  34.727  1.00  0.00
ATOM   1047  NH1 ARG A 131      14.222   5.738  35.384  1.00  0.00
ATOM   1048  NH2 ARG A 131      16.203   4.603  35.356  1.00  0.00
ATOM   1049  C   ARG A 131      12.115   9.803  30.971  1.00  0.00
ATOM   1050  O   ARG A 131      11.036   9.205  31.024  1.00  0.00
ATOM   1051  N   GLY A 132      12.225  11.075  30.599  1.00  0.00
ATOM   1052  CA  GLY A 132      11.036  11.830  30.245  1.00  0.00
ATOM   1053  C   GLY A 132      10.236  12.027  31.517  1.00  0.00
ATOM   1054  O   GLY A 132      10.817  12.279  32.579  1.00  0.00
ATOM   1055  N   LEU A 133       8.916  11.887  31.430  1.00  0.00
ATOM   1056  CA  LEU A 133       8.056  12.052  32.595  1.00  0.00
ATOM   1057  CB  LEU A 133       6.596  12.183  32.165  1.00  0.00
ATOM   1058  CG  LEU A 133       6.240  13.495  31.454  1.00  0.00
ATOM   1059  CD1 LEU A 133       4.785  13.459  31.029  1.00  0.00
ATOM   1060  CD2 LEU A 133       6.498  14.682  32.377  1.00  0.00
ATOM   1061  C   LEU A 133       8.205  10.925  33.614  1.00  0.00
ATOM   1062  O   LEU A 133       7.763  11.052  34.756  1.00  0.00
ATOM   1063  N   GLY A 134       8.827   9.820  33.207  1.00  0.00
ATOM   1064  CA  GLY A 134       9.030   8.716  34.134  1.00  0.00
ATOM   1065  C   GLY A 134       8.071   7.547  34.023  1.00  0.00
ATOM   1066  O   GLY A 134       7.383   7.186  34.988  1.00  0.00
ATOM   1067  N   ASN A 135       8.005   6.954  32.840  1.00  0.00
ATOM   1068  CA  ASN A 135       7.139   5.802  32.638  1.00  0.00
ATOM   1069  CB  ASN A 135       6.616   5.753  31.196  1.00  0.00
ATOM   1070  CG  ASN A 135       5.153   6.160  31.094  1.00  0.00
ATOM   1071  OD1 ASN A 135       4.477   5.827  30.123  1.00  0.00
ATOM   1072  ND2 ASN A 135       4.660   6.889  32.097  1.00  0.00
ATOM   1073  C   ASN A 135       7.911   4.523  32.963  1.00  0.00
ATOM   1074  OT1 ASN A 135       8.743   4.557  33.908  1.00  0.00
ATOM   1075  OXT ASN A 135       7.659   3.498  32.285  1.00  0.00
TER
ATOM   1076 NI+2 NI2 N   1      25.439  32.795  21.931  1.00  0.00
TER
ATOM   2127  OH2 WAT W   1       1.720  24.167  22.908  1.00  0.00
ATOM   2128  OH2 WAT W   2      12.313  37.301  14.794  1.00  0.00
ATOM   2129  OH2 WAT W   3      24.390  37.036  23.375  1.00  0.00
ATOM   2130  OH2 WAT W   4       9.636  25.324  16.967  1.00  0.00
ATOM   2131  OH2 WAT W   5      21.821  40.032  13.895  1.00  0.00
```

```
ATOM   2132  OH2 WAT W   6      22.559  17.580  21.358  1.00  0.00
ATOM   2133  OH2 WAT W   7      21.643  37.906  20.480  1.00  0.00
ATOM   2134  OH2 WAT W   8      13.114  27.953   9.830  1.00  0.00
ATOM   2135  OH2 WAT W   9       7.939   9.403  29.589  1.00  0.00
ATOM   2136  OH2 WAT W  10       9.188  33.244  26.598  1.00  0.00
ATOM   2137  OH2 WAT W  11       1.395  19.282  18.291  1.00  0.00
ATOM   2138  OH2 WAT W  12      33.487  45.173  20.097  1.00  0.00
ATOM   2139  OH2 WAT W  13      16.216  28.903  20.851  1.00  0.00
ATOM   2140  OH2 WAT W  14      14.835  32.984  20.328  1.00  0.00
ATOM   2141  OH2 WAT W  15      35.948  40.653  20.773  1.00  0.00
ATOM   2142  OH2 WAT W  16      -0.340  30.446  25.247  1.00  0.00
ATOM   2143  OH2 WAT W  17      32.711  41.962  12.829  1.00  0.00
ATOM   2144  OH2 WAT W  18      15.982  45.543  25.324  1.00  0.00
ATOM   2145  OH2 WAT W  19      19.597  36.020  20.473  1.00  0.00
ATOM   2146  OH2 WAT W  20      26.333  37.970   7.867  1.00  0.00
ATOM   2147  OH2 WAT W  21      29.039  23.082   8.154  1.00  0.00
ATOM   2148  OH2 WAT W  22       5.661  24.349  16.440  1.00  0.00
ATOM   2149  OH2 WAT W  23      27.272  41.113  18.159  1.00  0.00
ATOM   2150  OH2 WAT W  24      25.904  30.592   3.082  1.00  0.00
ATOM   2151  OH2 WAT W  25       0.283  21.190  32.526  1.00  0.00
ATOM   2152  OH2 WAT W  26       5.399  35.642  20.496  1.00  0.00
ATOM   2153  OH2 WAT W  27      -1.002  23.007  17.139  1.00  0.00
ATOM   2154  OH2 WAT W  28      17.793   8.166  18.807  1.00  0.00
ATOM   2155  OH2 WAT W  29       9.682  35.125  13.490  1.00  0.00
ATOM   2156  OH2 WAT W  30      16.993  36.394  20.953  1.00  0.00
ATOM   2157  OH2 WAT W  31      11.658  35.919  17.048  1.00  0.00
ATOM   2158  OH2 WAT W  32      24.881  43.416  18.704  1.00  0.00
ATOM   2159  OH2 WAT W  33      17.305  15.430  29.260  1.00  0.00
ATOM   2160  OH2 WAT W  34      18.124  14.931   9.038  1.00  0.00
ATOM   2161  OH2 WAT W  35      19.428   9.384  23.383  1.00  0.00
ATOM   2162  OH2 WAT W  36      20.192  22.586  26.554  1.00  0.00
ATOM   2163  OH2 WAT W  37      26.186  14.556  37.110  1.00  0.00
ATOM   2164  OH2 WAT W  38      23.123  43.140  10.032  1.00  0.00
ATOM   2165  OH2 WAT W  39      14.929  11.101  14.006  1.00  0.00
ATOM   2166  OH2 WAT W  40      22.024   9.433  19.349  1.00  0.00
ATOM   2167  OH2 WAT W  41      11.343  26.107   8.550  1.00  0.00
ATOM   2168  OH2 WAT W  42      26.912  41.330  10.514  1.00  0.00
ATOM   2169  OH2 WAT W  43      14.510   8.558  16.198  1.00  0.00
ATOM   2170  OH2 WAT W  44      24.635  13.558  15.991  1.00  0.00
ATOM   2171  OH2 WAT W  45      10.040  31.359  11.823  1.00  0.00
ATOM   2172  OH2 WAT W  46      26.356  37.178   3.508  1.00  0.00
ATOM   2173  OH2 WAT W  47      -6.411  15.058  34.654  1.00  0.00
ATOM   2174  OH2 WAT W  48       8.600  22.180  10.254  1.00  0.00
ATOM   2175  OH2 WAT W  49       7.430  34.801  17.202  1.00  0.00
ATOM   2176  OH2 WAT W  50      28.011  20.824  10.515  1.00  0.00
ATOM   2177  OH2 WAT W  51      14.462  27.698   2.479  1.00  0.00
ATOM   2178  OH2 WAT W  52      24.561  39.510   8.340  1.00  0.00
ATOM   2179  OH2 WAT W  53       2.221  17.492  16.828  1.00  0.00
ATOM   2180  OH2 WAT W  54      15.369  22.595  -2.149  1.00  0.00
ATOM   2181  OH2 WAT W  55      30.700  20.055   2.228  1.00  0.00
ATOM   2182  OH2 WAT W  56      13.483  30.782   4.176  1.00  0.00
ATOM   2183  OH2 WAT W  57       5.472   7.437  27.757  1.00  0.00
ATOM   2184  OH2 WAT W  58       6.787  20.189   8.923  1.00  0.00
ATOM   2185  OH2 WAT W  59      31.581  28.630   2.030  1.00  0.00
ATOM   2186  OH2 WAT W  60       8.711  43.825  25.064  1.00  0.00
ATOM   2187  OH2 WAT W  61      16.524  14.315  32.212  1.00  0.00
ATOM   2188  OH2 WAT W  62      20.750  15.405  28.157  1.00  0.00
ATOM   2189  OH2 WAT W  63       3.035  21.104   9.719  1.00  0.00
ATOM   2190  OH2 WAT W  64       3.609  29.577  29.375  1.00  0.00
ATOM   2191  OH2 WAT W  65      30.995  30.483   6.880  1.00  0.00
ATOM   2192  OH2 WAT W  66      32.817  21.788  18.148  1.00  0.00
ATOM   2193  OH2 WAT W  67       6.368  38.701  22.476  1.00  0.00
ATOM   2194  OH2 WAT W  68      10.639  22.680   8.955  1.00  0.00
```

```
ATOM   2195  OH2 WAT W  69      25.080  32.497   4.726  1.00  0.00
ATOM   2196  OH2 WAT W  70      12.048  44.137  27.080  1.00  0.00
ATOM   2197  OH2 WAT W  71      36.625  33.008  17.259  1.00  0.00
ATOM   2198  OH2 WAT W  72       4.117  29.627  18.531  1.00  0.00
ATOM   2199  OH2 WAT W  73      -3.083  23.153  18.986  1.00  0.00
ATOM   2200  OH2 WAT W  74      13.692   6.811  20.204  1.00  0.00
ATOM   2201  OH2 WAT W  75       6.774  29.452  20.986  1.00  0.00
ATOM   2202  OH2 WAT W  76      13.271  36.674  23.152  1.00  0.00
ATOM   2203  OH2 WAT W  77      42.132  29.903  21.736  1.00  0.00
ATOM   2204  OH2 WAT W  78      12.244  45.242  23.530  1.00  0.00
ATOM   2205  OH2 WAT W  79      21.741  40.832  11.570  1.00  0.00
ATOM   2206  OH2 WAT W  80      23.287   6.235  25.095  1.00  0.00
ATOM   2207  OH2 WAT W  81      10.668  20.083   7.540  1.00  0.00
ATOM   2208  OH2 WAT W  82      25.298  38.638   5.500  1.00  0.00
ATOM   2209  OH2 WAT W  83      44.292  38.262  16.432  1.00  0.00
ATOM   2210  OH2 WAT W  84      23.934  23.383  14.293  1.00  0.00
ATOM   2211  OH2 WAT W  85      29.338  40.325   9.928  1.00  0.00
ATOM   2212  OH2 WAT W  86      31.341  32.888   6.992  1.00  0.00
ATOM   2213  OH2 WAT W  87      -3.336  23.294  28.092  1.00  0.00
ATOM   2214  OH2 WAT W  88      18.616  40.161  19.320  1.00  0.00
ATOM   2215  OH2 WAT W  89      14.236  13.212  29.359  1.00  0.00
ATOM   2216  OH2 WAT W  90      -3.555  20.852  34.395  1.00  0.00
ATOM   2217  OH2 WAT W  91      37.843  32.572  13.044  1.00  0.00
ATOM   2218  OH2 WAT W  92       1.366  27.868  15.733  1.00  0.00
ATOM   2219  OH2 WAT W  93      28.719  17.817  23.827  1.00  0.00
ATOM   2220  OH2 WAT W  94      14.996   5.087  17.007  1.00  0.00
ATOM   2221  OH2 WAT W  95      18.752  35.916   6.512  1.00  0.00
ATOM   2222  OH2 WAT W  96      19.431  43.759  21.209  1.00  0.00
ATOM   2223  OH2 WAT W  97      14.228  14.734  31.279  1.00  0.00
ATOM   2224  OH2 WAT W  98      21.821  21.352   9.155  1.00  0.00
ATOM   2225  OH2 WAT W  99      16.448  30.453   4.488  1.00  0.00
ATOM   2226  OH2 WAT W 100      16.059  26.110  20.120  1.00  0.00
ATOM   2227  OH2 WAT W 101      22.592  25.762  -1.313  1.00  0.00
ATOM   2228  OH2 WAT W 102      27.120  34.809   4.534  1.00  0.00
ATOM   2229  OH2 WAT W 103      -1.219  34.618  20.819  1.00  0.00
ATOM   2230  OH2 WAT W 104      38.492  26.602  19.394  1.00  0.00
ATOM   2231  OH2 WAT W 105      13.098  48.337  24.647  1.00  0.00
ATOM   2232  OH2 WAT W 106      23.516   5.537  22.680  1.00  0.00
ATOM   2233  OH2 WAT W 107      28.476  20.900   0.794  1.00  0.00
ATOM   2234  OH2 WAT W 108      -4.627  11.219  25.413  1.00  0.00
```