PROTEIN THIOCARBOXYLATES: PROTEOMICS, MECHANISTIC STUDIES AND PATHWAY DISCOVERY

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Kalyanaraman Krishnamoorthy

May 2010

PROTEIN THIOCARBOXYLATES: PROTEOMICS, MECHANISTIC STUDIES AND PATHWAY DISCOVERY

Kalyanaraman Krishnamoorthy, Ph. D.

Cornell University 2010

Protein thiocarboxylates are members of sulfur transfer protein family and have been shown to be involved in a variety of important biosynthetic pathways like vitamin B1, molybdopterin, cysteine, thionucleosides among many others. Despite their importance, there has been lack of systematic efforts towards identifying new thiocarboxylate-forming proteins. In this work, we have taken efforts in developing two strategies to label them in bacterial cell-free extracts using fluorescent tags. In addition, bioinformatics search for new thiocarboxylate-forming proteins using a genomic database, theseed.uchicago.edu and a protein database, Pfam, yielded a new methionine biosynthetic pathway that involves a protein thiocarboxylate as the sulfur donor to make the precursor, homocysteine. This discovery further validated the need for developing methods to identify, in cell-free extracts, proteins carrying this important post-translational modification. The sulfur source for the protein thiocarboxylate involved in the methionine biosynthetic pathway has also been identified as sulfate or sulfite.

BIOGRAPHICAL SKETCH

K.Kalyanaraman (also known as Kalyanaraman Krishnamoorthy) was born in the southern Indian city of Chennai on May 31st 1982 to Mr. N.Krishnamoorthy, a former telecom officer at Chennai telephones and Mrs. K.Radhamani, an officer at Bank of India. He received his high school education at D.A.V. Boys Senior Secondary School, Chennai, India and it was here that he developed his liking for science, chemistry and biology in particular, taught by his two favorite teachers Mrs. Visalakshi and Mrs. Sivagami. Having finished his primary education in Chennai, he decided to pursue Masters in Chemistry at Indian Institute of Technology Bombay, Mumbai, India. His final-year masters thesis was done under the guidance of Prof. Anil K. Singh on synthesis of photolabile bioconjugates. His fascination and liking for bioorganic chemistry grew over the five-year graduate studies period at I.I.T Bombay and he decided to pursue his Ph.D. at a good school in the U.S.A. He was accepted for admission at Cornell University, Ithaca, where he worked with Dr.Tadhg Begley on proteomics and mechanistic enzymology. Kalyanaraman would be joining Dr. PradipSinh K. Rathod in University of Washington, Seattle for his post-doctoral research on Malarial pharmacology.

To my father Mr.N.Krishnamoorthy; my mother, Ms. K.Radhamani; my brother, K.Karthik and my dear fiancée, Anupriya

ACKNOWLEDGMENTS

I would like to thank my mentor Dr. Tadhg Begley for his continued support, patience and guidance during these past six years of my graduate career. Tadhg has been a constant source of encouragement and support both at times of ups and downs. My committee members, Dr. Brian Crane and Dr. Hening Lin, have always been there when I wanted to discuss my projects with them. I would also like to thank them along with Tadhg for helping me in my post-doctoral search. Begley lab has been a wonderful home and has been filled with wonderful people. Grant Sun, Amy Godert, Amy Haas, Kristin Burns and Jennie Sanders have helped me in the first few years learning the science and the ways of the lab. Abhishek Chatterjee and Jeremiah Hanes would always remain a source of inspiration for me when it comes to dedication to science. I also thank Abhishek for being instrumental in bringing me into the lab. David Hilmey has always gone an extra step towards assuring that my synthesis went well whenever I had to do it. Tathagata Mukherjee, Sean O'Leary, Amrita Hazra, BJ, Tomoshige and Lisa Cooper have always helped me scientifically and otherwise in the lab. The first-year graduate students - Rung-Yi Lai, Nilkamal, Angad, Yiquan and Kayla are smart and talented people and it was a wonderful experience working with them, particularly Kayla and Yiquan. Antonio's pizza and Domino's cheesy bread expeditions with Angad and his bubbly enthusiasm about experiments and science would always be remembered and cherished. Dinuka, Udani and Sameh have become good friends over the short span of time they were at the Begley lab. I would like to thank Dinuka for being there whenever I wanted to talk to someone. It was great being a member of the Begley lab and I thank everyone.

Life outside Begley lab was a rewarding experience. I was incredibly lucky to get wonderful friends whom I can proudly call as my Ithaca family. Rusty, Jaya,

Deepti, Lekha and Debamita took good care of me at all times. The good times spent with them would always be cherished and they would remain permanently etched in my memory. I also thank Sarin, Bale, Sudhamsu, Amit, Anand, Aritro, Hariharaputhiran, Hari, Jayashree, Debashree, Parag, Anandarup, Biki, and Tanushree for all the fun and laughter that they shared with me at different points of my life in Ithaca.

Finally, last and the most important, I would like to thank my family - my father, my mother, my brother and my lovely fiancée, Priya, for without their support, encouragement, guidance, love and hard-work, I would not be here writing this acknowledgement.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ACN: Acetonitrile.

AMP: Adenosine-5'-monophosphate.

ATCC: American Type Culture Collection.

ATP: Adenosine-5'-triphosphate.

Cys: L-cysteine.

Da: Dalton.

DMSO: Dimethyl sulfoxide.

DTT: Dithiothreitol.

DXP: 1-deoxy-D-xylulose-5-phosphate.

EDTA: Ethylenediaminetetraacetic acid

ESI-MS: Electrospray ionization mass-spectrometry.

HMP: Hydroxymethyl pyrimidine.

HPLC: High pressure liquid chromatography.

ICP-MS: Inductively coupled plasma mass spectroscopy.

IEF: Isoelectric focusing.

IPTG: Isopropyl β-D-1-thiogalactopyranoside.

mCi: Millicuries.

MALDI: Matrix assisted laser desorption/ionization.

MWCO: Molecular weight cut-off.

Ni-NTA: Nickel-nitrilotriacetic acid.

NMWL: Nominal molecular weight limit.

OD: Optical density.

PDTC: Pyridine dithiocarboxylic acid.

PETG: Polyethylene terephthalate copolyester.

PLP: Pyridoxal phosphate.

PMT: Photomultiplier tube.

rcf: Relative centrifugal force.

rpm: Rotations per minute.

SAM: S-adenosyl-L-methionine.

SDS-PAGE: Sodium dodecyl sulfate-polyacrylamide gel electrophoresis.

Tris-HCl: 2-Amino-2-(hydroxymethyl)-1, 3-propanediol, hydrochloride.

TCEP: Tris(2-carboxyethyl)phosphine.

TEV: Tobacco etch virus.

THF: Tetrahydrofuran.

CHAPTER 1

Sulfur transfer in micro-organisms and the importance of protein thiocarboxylates

1.1 Introduction to sulfur transfer

Sulfur is incorporated by micro-organisms into a variety of biomolecules such as amino-acids (cysteine, methionine), nucleosides (4-thiouridine, 2-thiocytidine) and co-factors that play an important role in catalyzing interesting enzymatic reactions (thiamin, molybdopterin, biotin)[1]. Other examples of sulfur-containing compounds can be seen among siderophores (thioquinolobactin, pyridine dithiocarboxylic acid), that are essential for iron-scavenging and bacterial survival under iron-limiting conditions. The wide prevalence of sulfur in the above mentioned small molecules and many others requires it to be readily available for biosynthesis. However, the concentration of free cellular sulfide is reported to be just enough to allow cysteine biosynthesis and the toxicity of free cysteine makes it necessary for the cellular concentrations of the amino-acid to be maintained at a low steady state level of 100 – 200 μM[2]. These constraints make prokaryotes produce sulfur-carrier proteins that help in the transfer of the element from its source to its destination in a highly regulated fashion. Two prominent means of trafficking sulfur are formation of protein persulfides (R-S-SH) and thiocarboxylates (R-COSH)[1].

1.2 Persulfide forming proteins

Cysteine desulfurases, rhodanese homology domain proteins and mercaptopyruvate sulfur transferases[1,2] are members of this family. Cysteine desulfurases use the co-factor pyridoxal phosphate (PLP) to remove the sulfur from cysteine to form the protein-bound persulfide as shown in **Figure 1.1**[3,4]. Examples include NifS, IscS, SufS

and many others. These proteins have been shown to be involved in the sulfur transfer process in more than one biosynthetic pathway. IscS in *Escherichia coli* has been suggested to be involved in the biosynthesis of thiamin, iron-sulfur clusters and thiouridines.

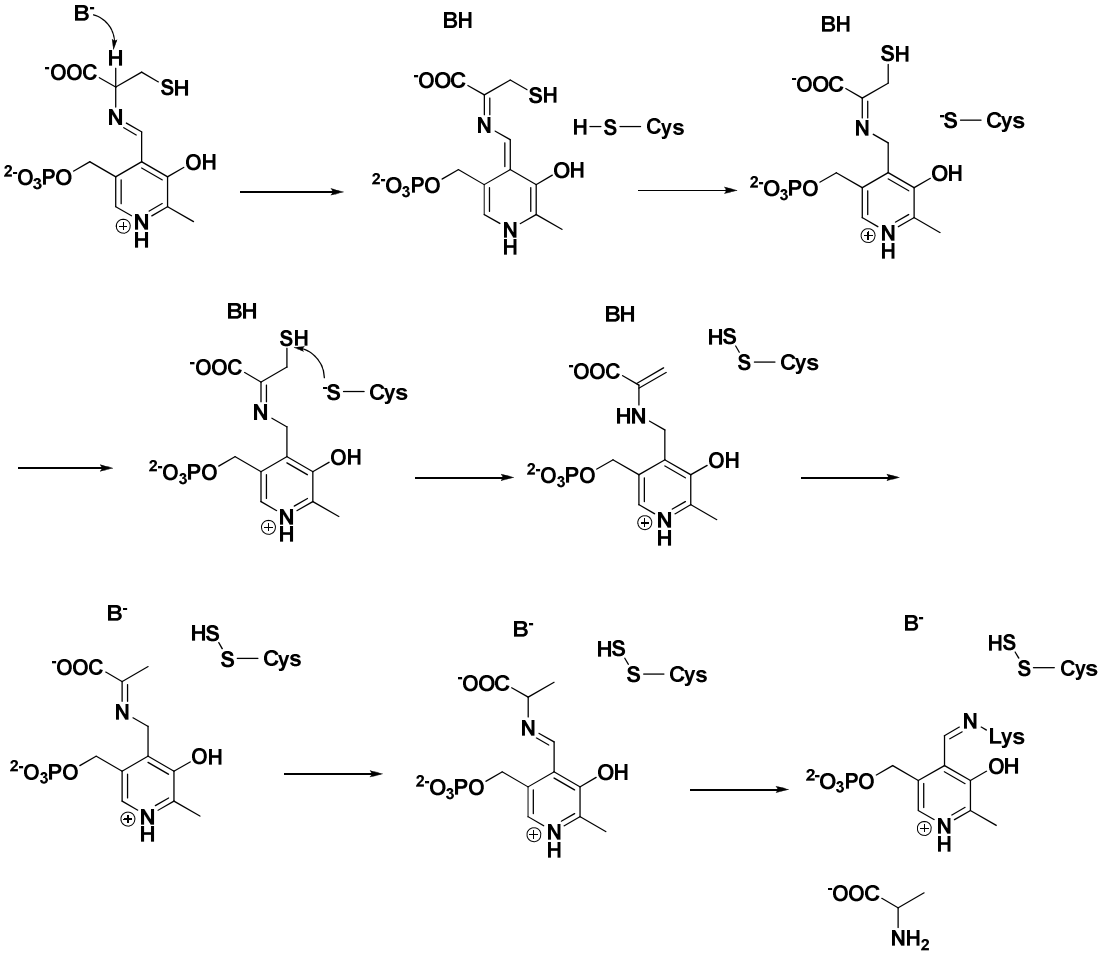


Figure 1.1: Mechanism of PLP-mediated protein persulfide formation.

Rhodanese homology domain proteins form cysteine-persulfide using thiosulfate as the sulfur source in *in-vitro* assays. The physiological role of these proteins is not well understood. However, they have been suggested to be involved in cyanide detoxification[5], iron-sulfur clusters formation[6] and elimination of reactive oxygen

species[7]. A well studied enzyme belonging to this family is the bovine rhodanese [8]. Mercaptopyruvate sulfur transferase converts mercaptopyruvate into pyruvate generating persulfidic sulfur in the process.

1.3 Thiocarboxylate forming proteins

Another form of protein-mediated sulfur trafficking is through C-terminal thiocarboxylation of certain ubiquitin-like proteins. The sulfur source for these important sulfur-trafficking proteins, except for thiamin[9] and molybdopterin[10,11], is not known. Both ThiS and MoaD are postulated to receive sulfur from cysteine through cysteine desulfurases. We have shown later in this work that thiocarboxylate-forming proteins can directly acquire the sulfur through sulfate-assimilation pathway.

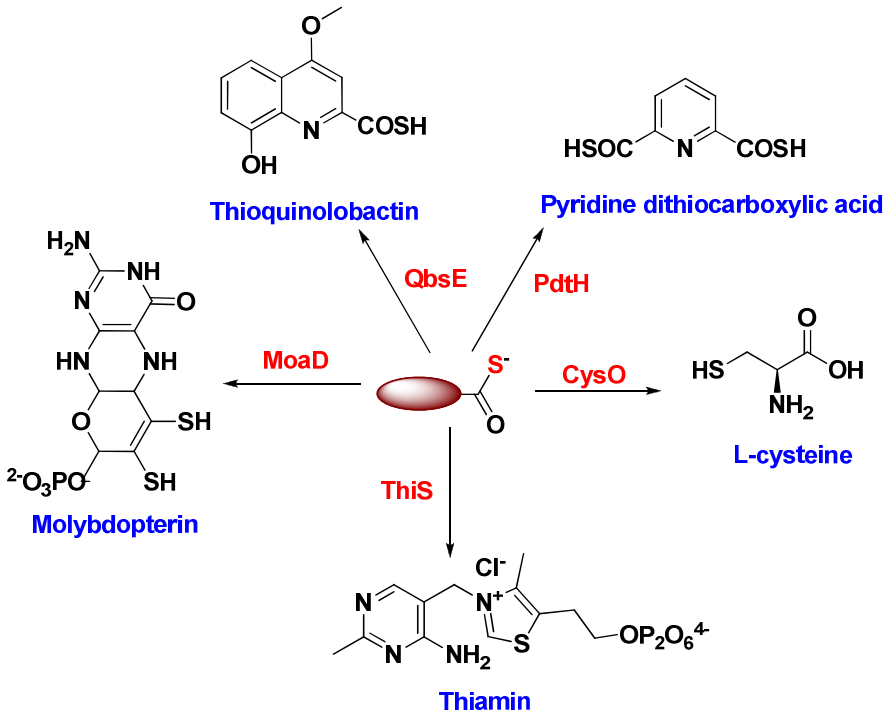
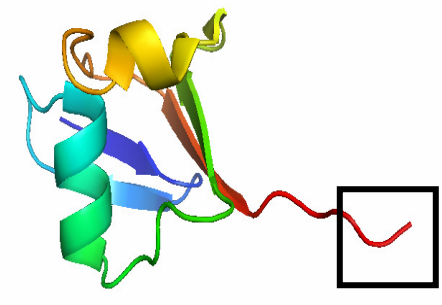


Figure 1.2: Examples of thiocarboxylate-forming proteins (names on the arrows) in certain biosynthetic pathways.

Few examples of such C-terminus thiocarboxylated-proteins include ThiS (involved in thiamin biosynthesis)[12], MoaD (molybdopterin biosynthesis)[11] and CysO (cysteine biosynthesis in *Mycobacterium tuberculosis*)[13] (**Figure 1.2**).

1.4 Physical characteristics of thiocarboxylate-forming proteins

Thiocarboxylate-forming proteins are small proteins. All known examples have a molecular weight of approximately 10 kDa or less. These proteins have a flexible C-terminal tail that ends with diglycyl residues (**Figure 1.3**). These proteins have their carboxy-terminus activated via adenylation for sulfur transfer by an adenylating enzyme (**Figure 1.4**).



MGHHHHHHHHHHSSGHIGGRHMLQLNGKDVKWKKDTGTIQDLLASYQLENKIVIVERNKEIIGKERYHE
VELCDRDVIEIVHFVG**GG**

Figure 1.3: The flexible C-terminal tail (shown in black box) of ThiS, the thiocarboxylate-forming protein involved in vitamin B_1 or thiamin biosynthesis. PDB: 1TYG. The primary sequence of the protein illustrating the diglycyl C-terminus is also shown.

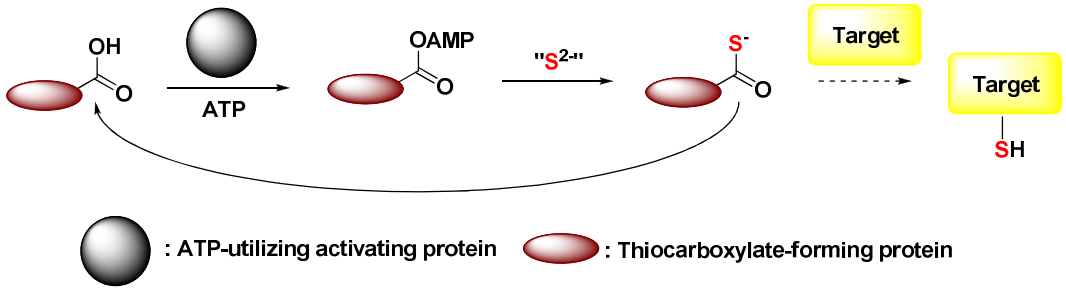


Figure 1.4: Mechanism of protein thiocarboxylate formation.

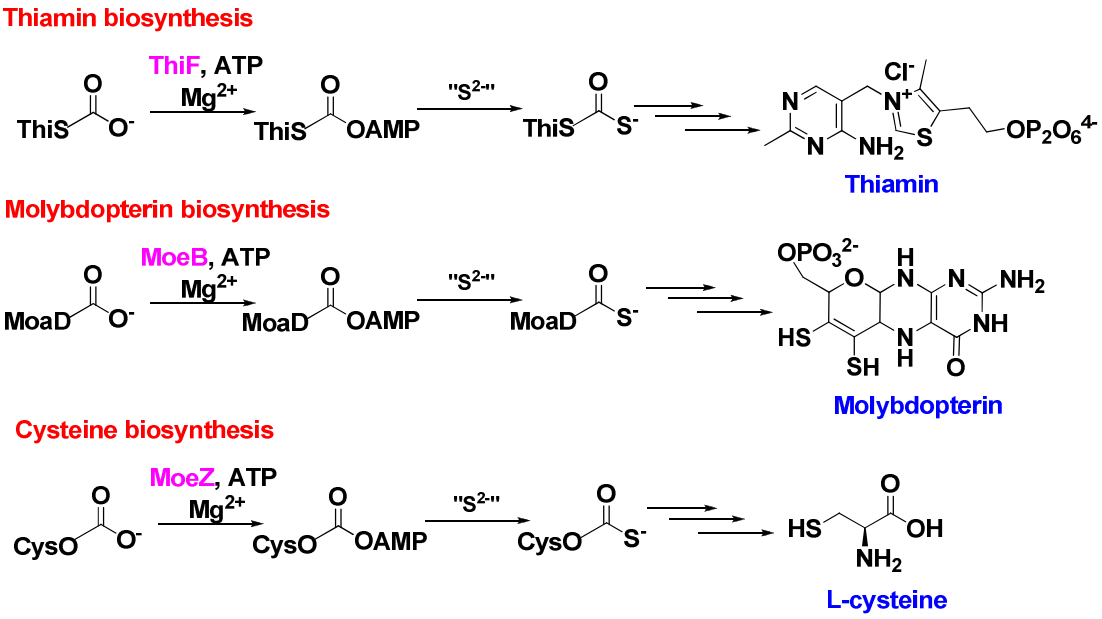
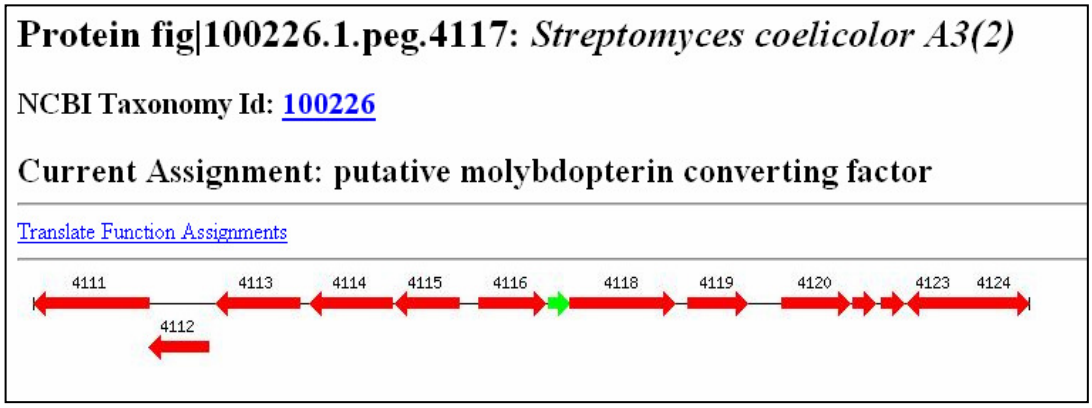


Figure 1.5: Activation of C-terminus of a thiocarboxylate-forming protein as an acyl-adenylate. The activating proteins in thiamin, molybdopterin and cysteine biosynthesis are shown in pink.

1.5 Identification of new thiocarboxylate-forming proteins

The cysteine biosynthetic pathway shown in **Figure 1.5** was uncovered when sequence analysis of genomes for homologs of ThiS and MoaD led to the discovery of *cysO* (Rv1335) in *Mycobacterium tuberculosis*, clustered with *cysM* (Rv1336, cysteine

synthase) and *mec*$^{+}$ gene (Rv1334, a putative hydrolase). Further studies identified the three proteins to be involved in making cysteine. The presence of two other biosynthetic routes to cysteine in this microorganism – the sulfide-dependent pathway and the cystathionine pathway, makes the new pathway look redundant. However, transcriptional profile analysis of *M. tuberculosis* demonstrated that *mec*$^{+}$, *cysO* and *cysM* genes are up-regulated under oxidative stress conditions [14]. This observation and the greater resistance of thiocarboxylates to oxidation than thiols suggests that CysM-CysO-mec^{+} route might be used when the bacteria are in the oxidizing environment of the macrophage.



MAKVTVRYWAAAKAAAGVAEEPYDAATLADALGAVRERHPGELTRVLLRCSFLVDGDPVGTRGHETVRL
AEGGTVEVLPPFA**GG**

Figure 1.6: ThiS-like protein (**4117**) from *Streptomyces coelicolor* present in an unknown gene cluster. **4113**: putative protease, **4116**: putative hydrolase, **4120**: putative thiosulfate sulfur transferase, **4121**: probable sseC protein, **4124**: putative oxidoreductase. Primary sequence of **4117** is also shown.

The discovery of *cysO* by sequence analysis and the subsequent revelation of a new cysteine biosynthetic route is an indication to the presence of more unknown

sulfur transfer pathways in the bacterial world, many of which might be up-regulated only under certain environmental-stress conditions. Sequence similarity search on ThiS-like proteins using Pfam (http://pfam.sanger.ac.uk/) results in the identification of large number of proteins with the characteristic diglycyl C-terminus found in thiocarboxylate-forming proteins. While many of them appear to be involved in already known thiamin, cysteine and molybdopterin biosynthesis, there are few ThiS-like proteins that are present in unknown biosynthetic clusters, as the example in **Figure 1.6** shows. This gene-clustering pattern is seen in many other organisms as well (*Thermomonospora curvata* DSM 43183, *Frankia sp.* EAN1pec, *Streptosporangium roseum* DSM 43021, *Thermobispora bispora* DSM 43833, *Catenulispora acidiphila* DSM 44928, *Frankia sp.* Ccl3) suggesting that they are functionally related. It is possible that **4117** might be activated by a non-specific adenylating enzyme making it ready to accept a sulfur from the thiosulfate-sulfur transferase, **4120**, and form a protein thiocarboxylate.

An analysis of the protein thiocarboxylate dependent biosynthetic pathways shown in **Figure 1.2** suggests the vitality of these proteins for proper bacterial survival and functioning. It would hence be useful to develop a strategy to identify thiocarboxylate proteins in a bacterial proteome and thus to uncover new sulfur transfer pathways. Such a study on virulent bacteria might also aid in designing drugs to inhibit the sulfur transfer process in the organism. Two approaches were taken in this regard to identify potential thiocarboxylate-forming proteins – the proteomics approach and the bioinformatics approach.

1.5.1 Proteomics approach to identify new thiocarboxylate-forming proteins

Proteomics approach involved developing labels containing fluorophores that react specifically with the thiocarboxylate functionality. The two approaches taken in this

study involved a) converting the thiocarboxylate to a reactive thioester using a thiol-reactive reagent followed by nucleophilic substitution of the thioester with a fluorophore, b) treating the thiocarboxylate functionality with a fluorescent sulfonyl azide. In both approaches, the labeled protein was analyzed by SDS-PAGE.

1.5.2 Bioinformatics approach to identify new thiocarboxylate-forming proteins

The primary sequence of a known protein thiocarboxylate like ThiS was used as a template to look for new, similar proteins in genomic and protein databases like The SEED (http://theseed.uchicago.edu/FIG/index.cgi), STRING (http://theseed.uchicago.edu/FIG/index.cgi) and Pfam (http://pfam.sanger.ac.uk/). In the present study, search for sulfur transfer proteins in the genomic database, the SEED and ThiS-like proteins in Pfam led to the discovery of a new protein thiocarboxylate-dependent methionine biosynthetic pathway in *Wolinella succinogenes.* The identity and putative functions of the proteins in the pathway were analyzed by NCBI-BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi).

REFERENCES

1. Mueller, E. G. Trafficking in persulfides: delivering sulfur in biosynthetic pathways. *Nat Chem Biol* **2**, 185-94 (2006).
2. Kessler, D. Enzymatic activation of sulfur for incorporation into biomolecules in prokaryotes. *FEMS Microbiol Rev* **30**, 825-40 (2006).
3. Mihara, H. & Esaki, N. Bacterial cysteine desulfurases: their function and mechanisms. *Appl Microbiol Biotechnol* **60**, 12-23 (2002).
4. Zheng, L., White, R. H., Cash, V. L. & Dean, D. R. Mechanism for the desulfurization of L-cysteine catalyzed by the nifS gene product. *Biochemistry* **33**, 4714-20 (1994).
5. Knowles, C. J. Microorganisms and cyanide. *Bacteriol Rev* **40**, 652-80 (1976).
6. Cerletti, P. Seeking a better job for an under-employed enzyme: rhodanese. *Trends in Biochemical Sciences* **11**, 369-372 (1986).
7. Nandi, D. L., Horowitz, P. M. & Westley, J. Rhodanese as a thioredoxin oxidase. *Int J Biochem Cell Biol* **32**, 465-73 (2000).
8. Westley, J. Rhodanese. *Adv Enzymol Relat Areas Mol Biol* **39**, 327-68 (1973).
9. Begley, T. P., Xi, J., Kinsland, C., Taylor, S. & McLafferty, F. The enzymology of sulfur activation during thiamin and biotin biosynthesis. *Curr Opin Chem Biol* **3**, 623-9 (1999).
10. Leimkuhler, S. & Rajagopalan, K. V. A sulfurtransferase is required in the transfer of cysteine sulfur in the in vitro synthesis of molybdopterin from precursor Z in Escherichia coli. *J Biol Chem* **276**, 22024-31 (2001).
11. Leimkuhler, S., Wuebbens, M. M. & Rajagopalan, K. V. Characterization of Escherichia coli MoeB and its involvement in the activation of molybdopterin

synthase for the biosynthesis of the molybdenum cofactor. *J Biol Chem* **276**, 34695-701 (2001).

12. Dorrestein, P. C., Zhai, H., McLafferty, F. W. & Begley, T. P. The biosynthesis of the thiazole phosphate moiety of thiamin: the sulfur transfer mediated by the sulfur carrier protein ThiS. *Chem Biol* **11**, 1373-81 (2004).
13. Burns, K. E. et al. Reconstitution of a new cysteine biosynthetic pathway in Mycobacterium tuberculosis. *J Am Chem Soc* **127**, 11602-3 (2005).
14. Manganelli, R. et al. Role of the extracytoplasmic-function sigma factor sigma(H) in Mycobacterium tuberculosis global gene expression. *Mol Microbiol* **45**, 365-74 (2002).

CHAPTER 2

Labeling thiocarboxylate proteins using alexafluor 647 cadaverine

2.1 Introduction

The soft and higher nucleophilicity of thioacids as compared to sulfhydryls makes it a good target for labeling reagents like iodoacetamide and N-ethylmaleimide. This property can be exploited in converting a protein thiocarboxylate into a reactive thioester group which can then be hydrolyzed by a nucleophile linked to a fluorophore or resin. The methodology developed involves activating the thiocarboxylate-terminus as an N-ethylmaleimide thioester followed by nucleophilic substitution of the ester with alexafluor 647 cadaverine (**Figure 2.1**). Alexafluor 647 cadaverine has a free amine as the attacking nucleophile, a molecular weight of ~ 1000 Da and has its excitation and emission maxima at 647 and 668 nm respectively.

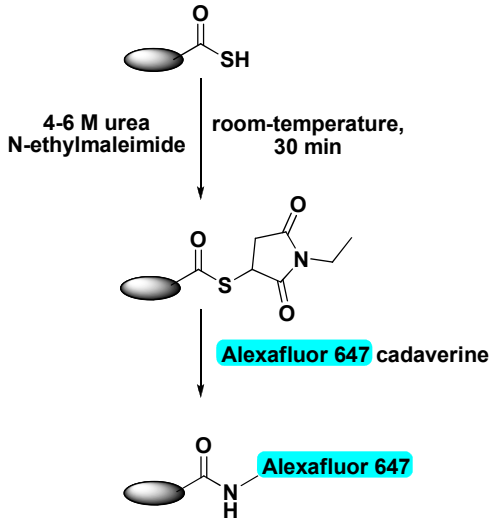


Figure 2.1: Strategy to label thioacids by converting them to reactive thioesters susceptible to nucleophilic substitution.

The strategy of converting thiocarboxylates into a thioester has been used extensively for C-terminal modification of proteins by intein purification chemistry (**Figure 2.2**) where a nucleophilic cysteine in the intein domain undergoes an S-N acyl shift to form the protein-intein thioester link which is then cleaved by using appropriate nucleophile[1].

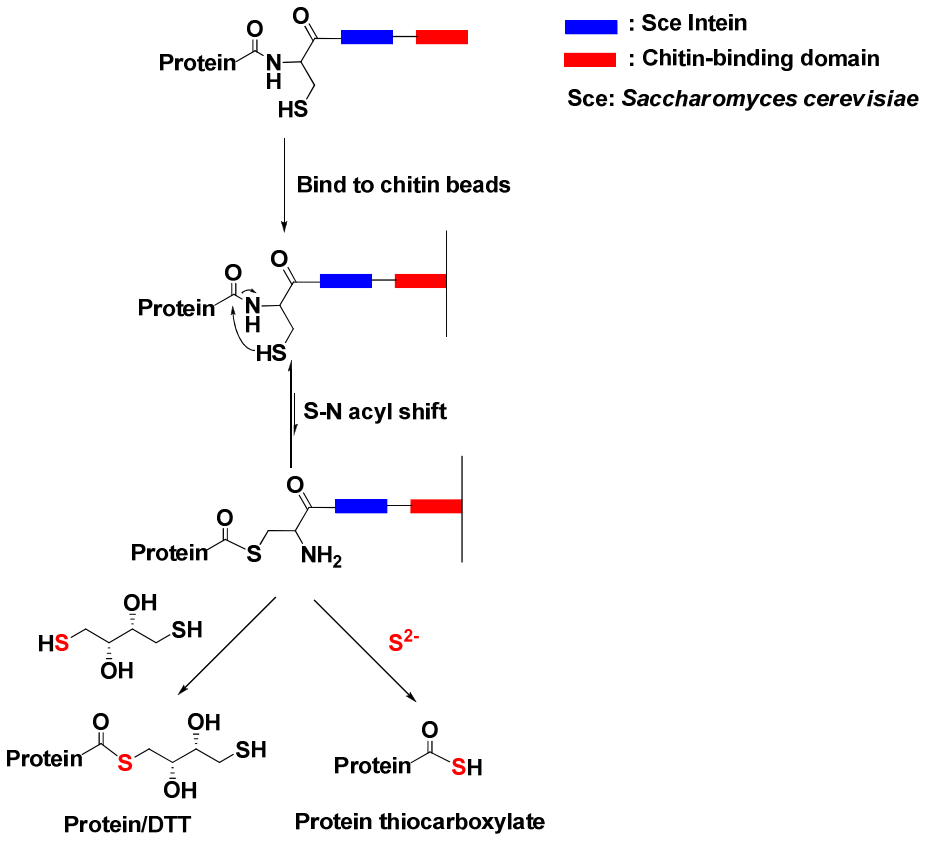


Figure 2.2: Intein purification chemistry. DTT cleaves the protein-intein thioester link to form protein/DTT adduct while cleavage with sulfide yields the protein thiocarboxylate.

2.2 Thiamin biosynthesis through ThiS thiocarboxylate

Biosynthesis of thiamin has been well-studied in micro-organisms[2,3]. Vitamin B_1 or thiamin pyrophosphate (**Figure 2.3**) comprises of two units – a five-membered

thiazole ring and a six-membered hydroxymethylpyrimidine ring, which are made separately and joined together in the final steps of the biosynthesis.

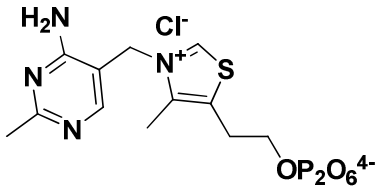


Figure 2.3: Vitamin B1 is made up of five-membered thiazole ring and six-membered hydroxymethyl pyrimidine ring.

The sulfur atom in the thiazole ring is provided by a small thiocarboxylate-forming protein, ThiS, which is adenylated at its C-terminus by the activating enzyme, ThiF. The activated ThiS then receives the sulfur from L-cysteine acted upon by a PLP-dependent cysteine desulfurase, NifS (**Figure 2.4**). The thiamin biosynthetic protein was chosen as the model system to test the labeling strategy. Once it was confirmed that the labeling was specific for the thiocarboxylate moiety, the methodology was used for detecting thiocarboxylated proteins in cell-free extracts of over-expressed systems.

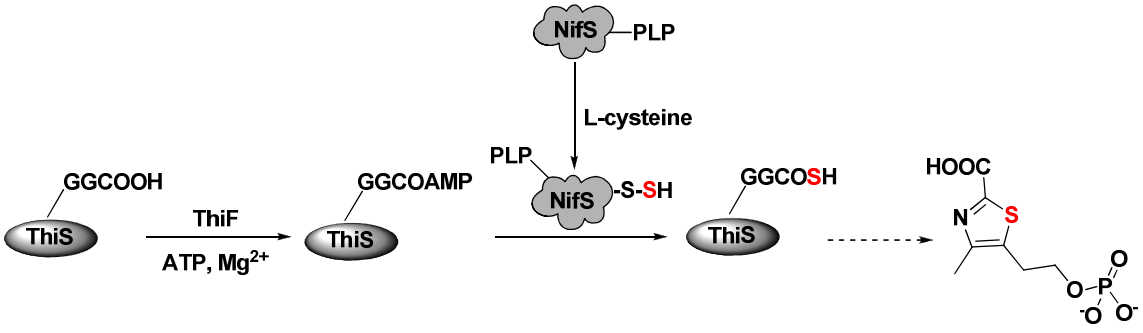


Figure 2.4: Sulfur transfer mechanism in thiamin biosynthesis

2.3 Thioquinolobactin biosynthesis through QbsE thiocarboxylate

Thioquinolobactin is a secondary siderophore produced by *P.fluorescens* (ATCC 17400). The production is enhanced when the biosynthesis of the fluorescent yellow-green primary siderophore, pyoverdine, is knocked out[4,5]. The biosynthesis of quinolobactin/thioquinolobactin is regulated by ferric uptake regulator (fur) protein and involves tryptophan degradation genes and genes analogous to sulfur transfer genes of thiamin biosynthesis[4] (**Figure 2.5**).

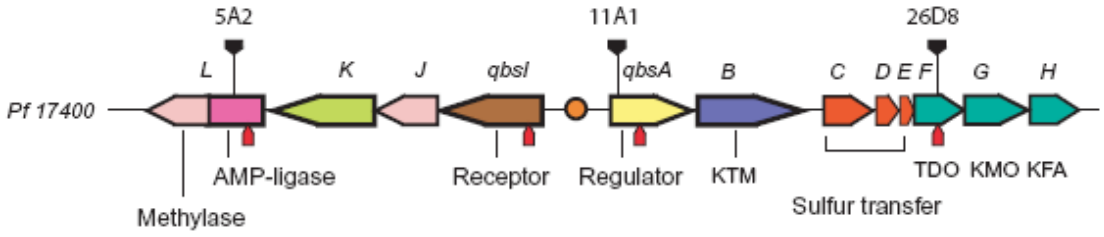


Figure 2.5: Gene clustering of thioquinolobactin biosynthetic proteins. TDO (tryptophan 2,3 dioxygenase), KMO (kynurenine monooxygenase), KTM (kynurenine transaminase) and KFA (kynurenine formamidase) are tryptophan degrading enzymes while QbsC, QbsD and QbsE are speculated to be involved in sulfur transfer.

QbsE, the thiocarboxylate-forming protein in the thioquinolobactin biosynthesis has a cysteine-phenylalanine at the C-terminus ahead of the diglycyl groups (**Figure 2.6**). As mentioned in Chapter 1, all known examples of these proteins end in diglycyl C-terminus. QbsD is a metalloprotease present in the gene-cluster and it has been shown to cleave the last two amino-acids to expose the diglycyl C-terminus making it possible to be activated by a bifunctional enzyme, QbsC, which adenylates QbsE C-terminus and provides a sulfur derived from thiosulfate through a rhodanese cysteine persulfide[6]. This system was chosen to test the labeling strategy on a cell-free extract.

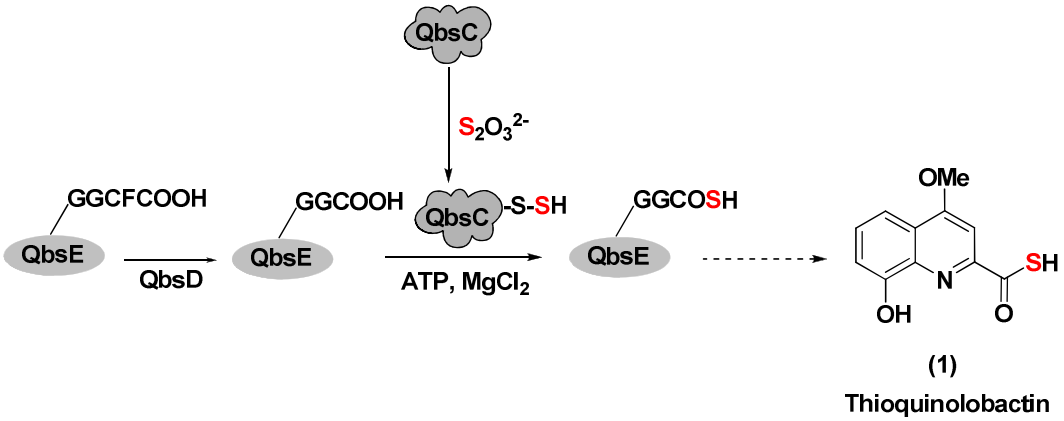


Figure 2.6: Putative sulfur transfer by thioquinolobactin biosynthetic proteins

2.4 Experimental section

Dithiothretiol (DTT), ATP, L-cysteine, urea and N-ethylmaleimide were purchased from Sigma-aldrich (St.Louis, MO). Millipore's YM-10 microcons and 5K NMWL Ultrafree centrifugal filter devices, Tris.HCl and imidazole (Acros chemicals) were obtained from Fisher Scientific (Fairlawn, NJ) while L-^{35}S-cysteine and STORM 860 was got from GE healthcare life sciences (Piscataway, NJ). Ampicillin, kanamycin and isopropyl-β-D-thiogalactopyranoside (IPTG) were obtained from Lab Scientific Inc. (Livingston, NJ). Tris (2-carboxyethyl)phosphine hydrochloride (TCEP) was got from Soltec ventures Inc. (Beverly, MA). Luria-Bertani was purchased from EMD chemicals Inc. (Gibbstown, NJ). Sonication of cell-cultures was done using Misonix sonicator 3000 (Misonix Inc., Farmingdale, NY). Avanti J-E centrifuge (Beckman Coulter, Fullerton, CA) was used for centrifugation purposes. All protein concentrations were measured by Bradford assay[7].

2.4.1 Cloning of thiamin and thioquinolobactin biosynthetic enzymes

qbsC gene was PCR amplified from *P. fluorescens* (ATCC 17400) genomic DNA and cloned into pET-28a vector. *qbsE* was also cloned into pACYCDuet vector. *qbsCE*

over-expression system was prepared by co-transforming *qbsC* (pET-28a vector) along with *qbsE* (pACYCDuet vector). *qbsCDE*, the quinolobactin sulfur transfer genes, were cloned as a single contiguous unit into pET-28a vector with the His-tag on QbsC[6]. *thiSG* and *thiF* were cloned into pET-22b vector while *nifS* was cloned into pET-16b vector[8].

2.4.2 Over-expression and purification of proteins

The thiamin biosynthetic proteins were over-expressed in *E.coli* BL21(DE3) in Luria-Bertani medium and induced at an OD_{600} of 0.6 with a final concentration of 0.5 mM isopropyl-β-D-thiogalactopyranoside (IPTG). ThiSG were over-expressed together as the mutual presence improves the stability and solubility of the proteins. The cells were harvested by centrifugation and lysed by sonication on ice. ThiF, ThiSG and NifS were purified by Ni-NTA purification protocol at 4^{o}C and buffer exchanged into 100 mM Tris, pH 7.8.

2.4.3 Reconstitution of ThiS-COSH

The buffer used for protein samples and reagent stocks is 100 mM Tris, pH 7.8. 91 µL of 93 µM of NifS , 93 µL of 88 µM of ThiF, 336 µL of 6.318 mg/mL of ThiSG, 2 µL of 800 mM $MgCl_2$, 2.5 µL of 200 mM ATP and 1 µL of cysteine stock which contained 0.9 M L-cysteine and 1.8 M DTT were added together and incubated at room temperature for 2 h to reconstitute ThiSCOSH[9].

After incubation, 50 µL of four samples were prepared as shown in **Table 2.1**. ThiSG concentration was the same in all four samples. A final concentration of 5 M urea was added to each sample to denature the proteins and quench the reaction. 7 µL of 100 mM N-ethylmaleimide was then added to samples 2 and 4. The samples were

incubated at room-temperature for 30 min after which 1 µL of 10 mM alexa fluor 647 cadaverine (Invitrogen, Carlsbad, CA) was added and further incubation was carried out in the dark for 5 h. The samples were desalted using Bio-spin 6 Tris columns (Bio-rad laboratories, Hercules, CA) into 10 mM Tris-HCl, 0.02% sodium azide, pH 7.4 before being analyzed on a 15% Tris-glycine SDS-PAGE. The gel was scanned for fluorescence on a STORM 860 imager.

Table 2.1: Controls for ThiS-COSH labeling experiment

Sample number	Samples	Maleimide
1	Pure ThiSG	No
2	Pure ThiSG	Yes
3	Reconstituted ThiSG	No
4	Reconstituted ThiSG	Yes

2.4.4 Rate of ThiS-COSH formation

The assay was done by two different methods – by the use of the fluorescent method developed in this work and by using L-^{35}S-cysteine. All protein samples and substrate stocks were made in 100 mM Tris-HCl, pH 7.8. 35 µL of 121 µM of NifS, 62 µL of 66 µM ThiF, 246.5 µL of 4.3 mg/mL of ThiSG, 1.5 µL of 800 mM $MgCl_2$, 2 µL of 200 mM ATP and 3.5 µL of cysteine stock which contained 200 mM L-cysteine and 400 mM DTT were added together and incubated at room temperature for 3 h. For the radioactive sample, an additional 20 µL of L-^{35}S-cysteine (0.5 mCi/100 µL) was added to the solution. Zero time-point sample was prepared by adding 1.99 µL of 121 µM, 3.5 µL of 66 µM ThiF, 14.1 µL of 4.3 mg/mL ThiSG, 0.2 µL of 800 mM $MgCl_2$, 0.2 µL of 200 mM ATP and 0.2 µL of cysteine stock containing 200 mM L-cysteine and 400 mM DTT to 20 µL of 8 M urea. The corresponding sample for the radioactive

methodology had an additional 1.07 µL of L-^{35}S-cysteine (0.5 mCi/100 µL). The reaction was started by addition of cysteine. 10 µL aliquots were taken out from the reaction mixture after regular time-intervals and quenched with 10 µL of 8 M urea. These quenched samples were then treated as follows.

2.4.4.1 Fluorescent labeling method

7 µL of 100 mM N-ethyl maleimide was added to each sample and incubated at room-temperature for 30 min after which 8 µL of 1 mM alexa fluor 647 cadaverine was added and further incubation for 5 h in the dark at room temperature was carried out. The samples were desalted using YM-10 microcons by multiple washings with 100 mM Tris-HCl, pH 7.8 and analyzed on a 15% Tris-glycine SDS-PAGE.

2.4.4.2 Radioactive method

The samples were desalted using 5K NMWL Ultrafree centrifugal filter devices and then analyzed on a 15% Tris-glycine SDS-PAGE.

Fluorescent and radioactive images of the gel were obtained using STORM 860 imager and the intensities quantified using the Imagequant 5.2 software (Molecular dynamics, Sunnyvale, CA). Date-analysis to obtain the rate was done by non-linear regression using Grafit 5.0.11 (Erithacus software, Surrey, UK).

2.4.5 Labeling of QbsE-COSH in the QbsCDE cell-free extract

Two 100 mL cultures of *E.coli* BL21(DE3) strain over-expressing QbsCDE and QbsCE were induced with 1 mM IPTG at an OD_{600} of 0.8. Post-induction, the cell cultures were then grown at 15^{o}C for 12 h with agitation after which they were harvested using JA-10 rotor in Avanti J-E centrifuge (Beckman Coulter Inc.,

Fullerton, CA) at 13700 rcf. The cell-pellets were then re-suspended in approximately 10-15 mL of lysis buffer (20 mM potassium phosphate, 250 mM NaCl, 20 mM imidazole, pH 8.0). The cells were lysed by sonication (pulse 'on' time 1.0 sec, pulse 'off' time 1.0 sec, output level 0.8, 30 cycles 5 times) on ice. The crude lysate was centrifuged at 39800 rcf using JA-17 rotor in Avanti J-E centrifuge to remove the cell debris. The supernatant in 1 mL aliquots were frozen and stored at -60^{o}C. 2 mL each of CDE and CE cell-free extract were thawed and concentrated to a volume of 200 µL and 72 mg of urea was added to each solution to make a final urea concentration of approximately 5 M. 50 µL of the resulting solution were taken and treated for 45 min at room-temperature with 20 µL of 100 mM N-ethylmaleimide after which 2 µL of 10 mM alexa fluor 647 cadaverine was added and further incubation was carried out in the dark at room-temperature for 2 h. The samples were desalted using YM-10 microcons by multiple washings with the lysis buffer and analyzed on 15% Tris-glycine SDS-PAGE. Fluorescent image of the gel was obtained using STORM 860 imager.

2.5 Results and discussions

The nucleophilicity of thioacid was exploited in increasing the reactivity of the moiety using thiol-reactive probe, N-ethylmaleimide. The thioester so formed is susceptible to nucleophilic substitution with alexa fluor 647 cadaverine (**Figure 2.1**), thereby fluorescently tagging the protein.

As mentioned earlier, thiamin biosynthetic enzymes were used for testing the methodology. ThiS-COSH was formed by treating ThiS-COOH with NifS (PLP-utilizing cysteine desulfurase), L-cysteine, ThiF, ATP and Mg^{2+} ions. Lanes 2 and 3 in the SDS-PAGE shown in **Figure 2.7** contain non-reconstituted ThiS. The carboxylate terminus of ThiS does not react with alexa fluor 647 cadaverine with or without N-

ethylmaleimide treatment. On the other hand, lanes 4 and 5 contain reconstituted ThiS-COSH. There is no labeling in lane 4 because the thioacid has not been activated as a thioester with N-ethylmaleimide. Lane 5 shows labeling because of activation of ThiS-COSH as a thioester with N-ethylmaleimide which makes it efficient for nucleophilic substitution.

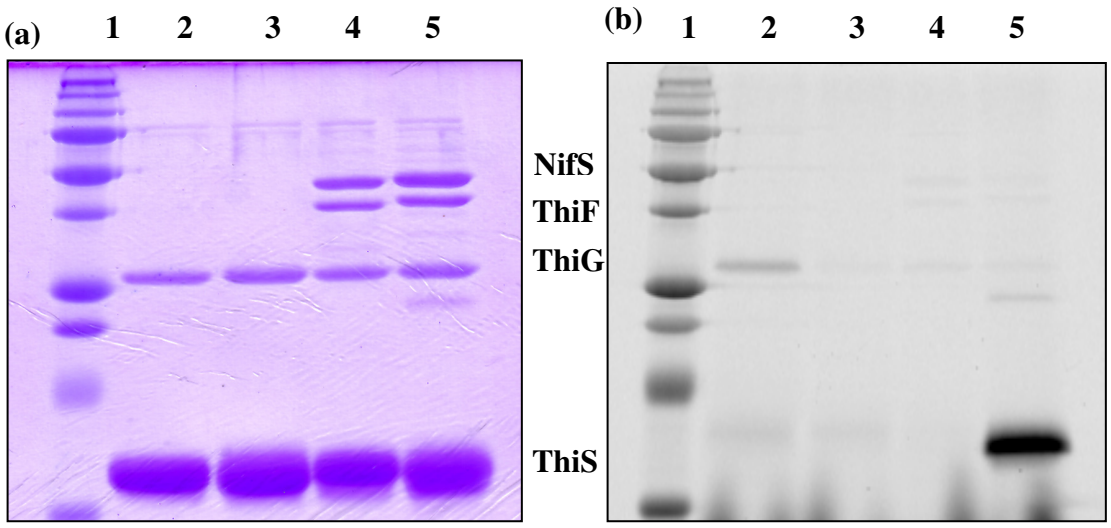


Figure 2.7: Labeling of ThiS-COSH (a) coomassie scan and (b) fluorescent image. Lane 1: Molecular weight markers, Lane 2: ThiSCOOH, no maleimide; Lane 3: ThiSCOOH, with maleimide; Lane 4: ThiSCOSH, no maleimide; Lane 5: ThiSCOSH, with maleimide

The rate of ThiS-COSH formation using cysteine as the sulfur source was studied using the labeling strategy to further check the suitability of the method to assay for protein thiocarboxylates. The rate constant has previously been found to be of the order of 0.02 min^{-1} [8]. The rate was also confirmed using L-S^{35}-cysteine as the sulfur source. ThiS-COSH so produced would be radiolabeled (**Figure 2.8(b)**). The

fluorescent and radioactive intensities can then be directly read out by imaging using STORM 860 (**Figures 2.8(c) and 2.8(d)**).

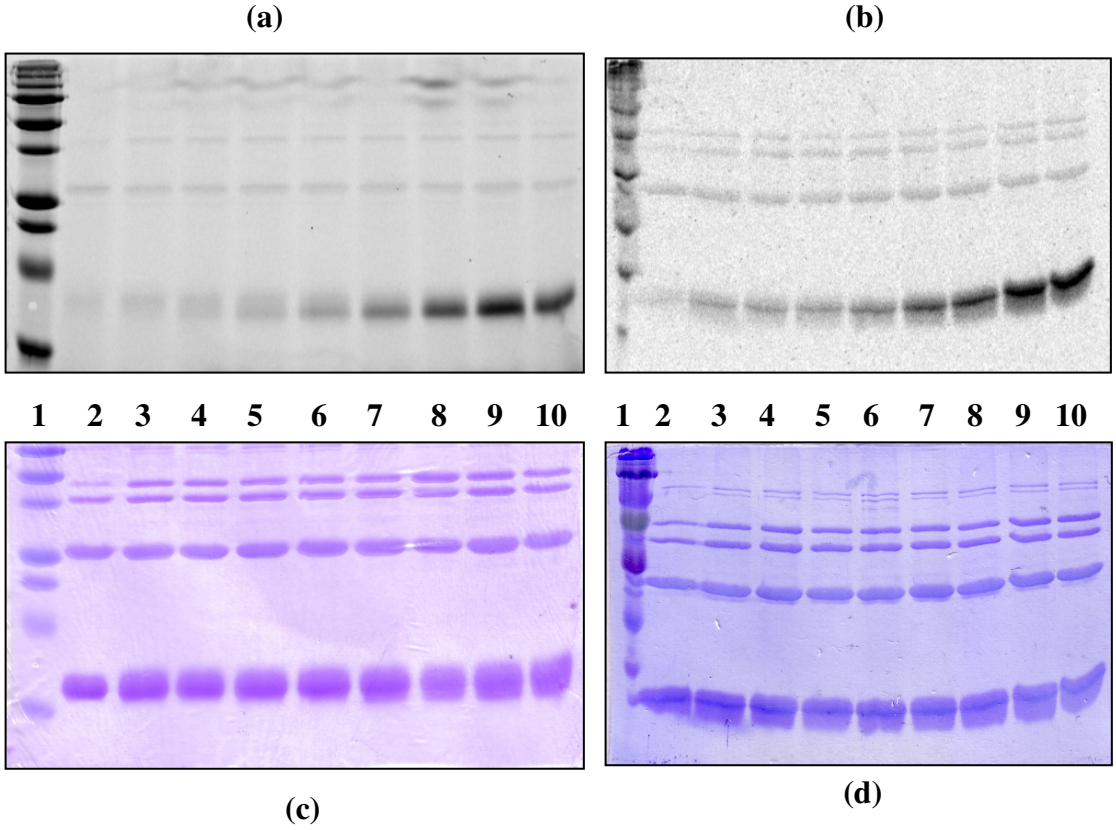


Figure 2.8: Time-dependence of sulfur transfer from cysteine to ThiS. (a) shows labeling of ThiS-COSH formed at different time-points with alexafluor 647 cadaverine (c) is the coomassie staining of the gel. (b) shows formation of radiolabeled ThiS-COSH at various time-intervals. (d) is the corresponding coomassie staining. Lane 1: Molecular weight markers, Lane 2: 0 min, Lane 3: 2 min, Lane 4: 5 min, Lane 5: 8 min, Lane 6: 12 min, Lane 7: 20 min, Lane 8: 30 min, Lane 9: 60 min, Lane 10: 90 min

The fluorescent and radioactive intensities on ThiS-COSH at different time-points obtained from the two methods were plotted using non linear regression with Grafit

5.0.11 (**Figure 2.9**) and identical rate constants of 0.04 min^{-1} was obtained by both ways.

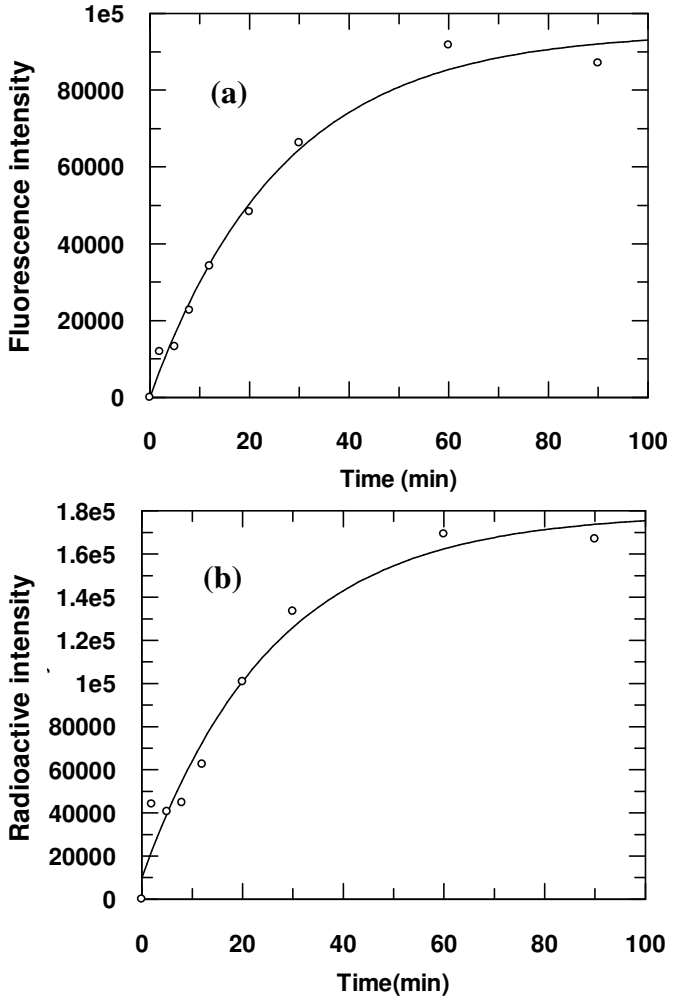


Figure 2.9: Rate of sulfur transfer from cysteine to ThiS (a) by fluorescent labeling of the thiocarboxylate formed (b) by using radioactive cysteine as the sulfur source

To test the strategy on a proteomics scale, thioquinolobactin biosynthetic enzymes were chosen. The proposed mechanism of production of QbsE-COSH is summarized in **Figure 2.6**. QbsCDE over-expression system produces QbsE-COSH. However, the absence of the metalloprotease QbsD in QbsCE over-expression system prevents

sulfur transfer by QbsC to full-length QbsE[6]. Hence, no QbsE-COSH would be formed in the QbsCE over-expression system. The dye was shown to selectively label only QbsE-COSH from the cell-free extract of the QbsCDE system (**Figure 2.10**).

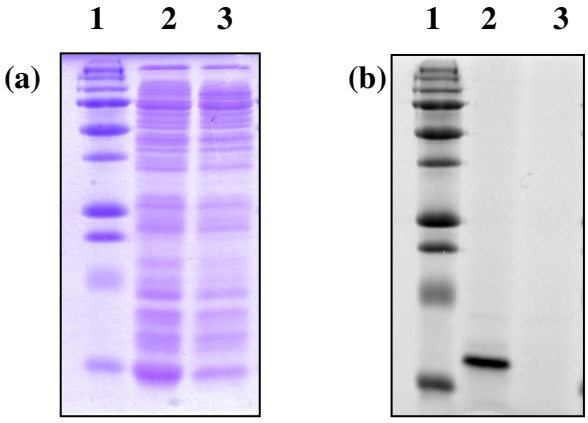


Figure 2.10: Labeling of over-expressed QbsE-COSH in cell-free extract (a) coomassie staining (b) fluorescent image. Lane 1: Molecular weight markers, Lane 2: QbsCDE – cell-free extract, Lane 3: QbsC/QbsE - cell-free extract.

2.6 Conclusions

The methodology developed has demonstrated its ability to label thiocarboxylates in purified protein system as well as in cell-free extract of QbsCDE over-expression system. The main goal of the work is to develop a strategy that can be used as a proteomics tool to identify thiocarboxylate-forming protein in a native non-over-expressed system. While alexafluor 647 cadaverine has proved its mettle in a purified and over-expressed system, it might not be the best tag to use in a native system where efficiency of the tagging reaction would play a key role in pulling out low-copy number putative thiocarboxylate-forming proteins. It is well-known that amines are not nucleophilic enough to achieve quantitative substitution of the protein-N-

ethylmaleimide thioester. Changing the nucleophile to alexafluor 647 hydrazine, which is a better nucleophile than an amine, did not help either as it resulted in extensive non-specific labeling of other proteins. So, a new conjugation technique based on thioacid-azide click chemistry was exploited for proteomics.

REFERENCES

1. Perler, F. B. & Adam, E. Protein splicing and its applications. *Curr Opin Biotechnol* **11**, 377-83 (2000).
2. Dorrestein, P. C., Zhai, H., McLafferty, F. W. & Begley, T. P. The biosynthesis of the thiazole phosphate moiety of thiamin: the sulfur transfer mediated by the sulfur carrier protein ThiS. *Chem Biol* **11**, 1373-81 (2004).
3. Xi, J., Ge, Y., Kinsland, C., McLafferty, F. W. & Begley, T. P. Biosynthesis of the thiazole moiety of thiamin in Escherichia coli: identification of an acyldisulfide-linked protein--protein conjugate that is functionally analogous to the ubiquitin/E1 complex. *Proc Natl Acad Sci U S A* **98**, 8513-8 (2001).
4. Matthijs, S. et al. The Pseudomonas siderophore quinolobactin is synthesized from xanthurenic acid, an intermediate of the kynurenine pathway. *Mol Microbiol* **52**, 371-84 (2004).
5. Mossialos, D. et al. Quinolobactin, a new siderophore of Pseudomonas fluorescens ATCC 17400, the production of which is repressed by the cognate pyoverdine. *Appl Environ Microbiol* **66**, 487-92 (2000).
6. Godert, A. M., Jin, M., McLafferty, F. W. & Begley, T. P. Biosynthesis of the thioquinolobactin siderophore: an interesting variation on sulfur transfer. *J Bacteriol* **189**, 2941-4 (2007).
7. Bradford, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**, 248-54 (1976).
8. Park, J. H. et al. Biosynthesis of the thiazole moiety of thiamin pyrophosphate (vitamin B1). *Biochemistry* **42**, 12430-8 (2003).

9. Hazra, A., Chatterjee, A. & Begley, T. P. Biosynthesis of the Thiamin Thiazole in Bacillus subtilis: Identification of the Product of the Thiazole Synthase-Catalyzed Reaction. *J Am Chem Soc* **131**, 3225-9 (2009).

CHAPTER 3

Labeling thiocarboxylate proteins using lissamine rhodamine sulfonyl azide

3.1 Introduction

Thioacids and azides are known to react with each other to form an amide linkage[1,2]. The proposed mechanism of the reaction[1] is shown in **Figure 3.1**. A thiatriazoline intermediate (compound 1, **Figure 3.1**), is proposed to form via either a 2+3 cycloaddition or a stepwise diazo transfer-like mechanism. Decomposition of 6, stepwise or by a retro-[2+3] reaction, would ultimately lead to the amide with the release of nitrogen and sulfur.

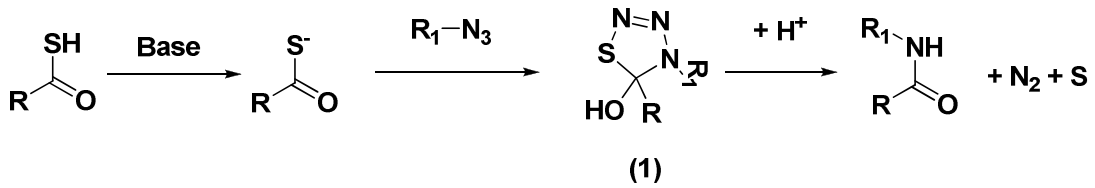
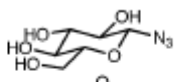
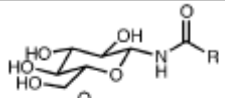
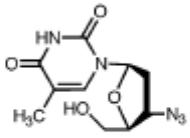
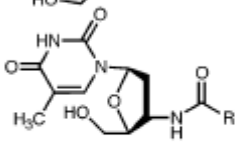
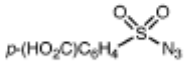
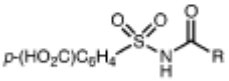


Figure 3.1: Proposed mechanism of the amide formation from thioacid-azide coupling.

The high yields and shorter reaction times of electron-deficient azides like the sulfonyl azides with thioacids (**Table 3.1**) was exploited for labeling thiocarboxylate-forming proteins in bacterial cell-free extracts (**Figure 3.2**). A sensitive and easy-to-use fluorescent tag, Lissamine rhodamine sulfonyl azide, was synthesized and utilized for this purpose.

Table 3.1: Reaction of organic azides with thioacids in aqueous solvent[1].

Entry	Azide	°C/time/solvent	Amide	Yield
1		a) 60/36 h/H_2O b) 60/36 h/H_2O		a) 83% b) 80%
2		a) 60/36 h/H_2O b) 60/36 h/H_2O		a) 68% b) 77%
3		a) 25/1 h/H_2O b) 25/1 h/H_2O		a) 93% b) 98%

[a] Conditions: 0.25−0.040 M azide; 1:1.3−5 azide:thio acid; entry 1, $NaHCO_3$(aq); entry 2, PBS buffer pH 7.4; entry 3, 1.8 equiv of 2,6-lutidine. (a) Thiobenzoic acid, R = C_6H_5. (b) Thioacetic acid, R = CH_3.

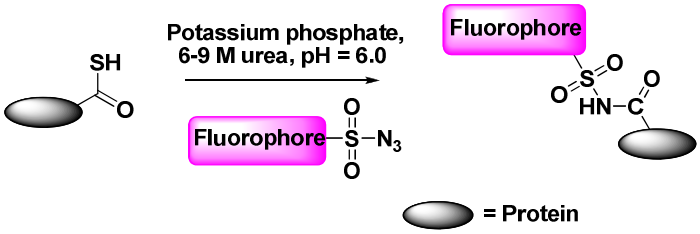


Figure 3.2: Reaction of protein thiocarboxylates with sulfonyl azide.

3.2 Experimental section

Lissamine rhodamine B sulfonyl chloride, EDTA, dithiothreitol (DTT), potassium phosphate, ferrous ammonium sulphate, sterile disposable PETG flask with vented closure and β-mercaptoethanol (Acros organics) were purchased from Fisher Scientific (Fairlawn, NJ), sodium azide, urea, ATP, Tris, bacterial protease inhibitor cocktail, propanedithiol and Tris(2-carboxyethyl)phosphine (TCEP) (Fluka) from Sigma-Aldrich (St.Louis, MO). IPTG was bought from Lab scientific inc. (Livingston, NJ). Luria-Bertani from EMD biosciences (Gibbstown, NJ) and Difco nutrient broth from BD (Franklin lakes, NJ). *Pseudomonas stutzeri* KC (ATCC 55595), *S.coelicolor* (ATCC 10147), *S.erythrea* (ATCC 11635), *S.griesus* (ATCC

23345) and *S.avermitilis* (ATCC 31267) was bought from American Type Culture Collection (Manassas, VA). *B.xenovorans* LB400 was a gift from Dr.James Tiedje (Michigan State University, East Lansing, MI) and *Rhodococcus* sp. RHA1 was provided by Dr.Lindsay Eltis (University of British Columbia, Vancouver, Canada). Chitin beads and pTYB1 vector was obtained from New England Biolabs (Ipswich, MA). Protein concentrations were determined by Bradford assay[3]. All fluorescence gel images were scanned using Typhoon 9400 or Typhoon trio (excitation: 532 nm green laser; emission: 580-nm band-pass filter (580 BP 30)) obtained from GE healthcare biosciences (Piscataway, NJ). Sonication of the cultures was done on Misonix Sonicator 3000 (Misonix Inc., Farmingdale, NY). ESI-MS analysis was performed on Esquire-LC_00146 instrument (Bruker, Billerica, MA). Data for nano-LC-MS/MS analysis of the labeled protein in the cell-free extract *P.stutzeri* KC was provided by Proteomics and Mass-spectrometry facility at Cornell University, Ithaca. NanoLC was carried out by an LC Packings Ultimate integrated capillary HPLC system equipped with a Switchos valve switching unit (Dionex, Sunnyvale, CA). The digested peptides were injected using a Famous auto sampler onto a C18 PepMap trap column (5 µm, 300 µm × 5 mm, Dionex) for on-line desalting and then separated on a PepMap C-18 RP nano column, eluted in a 30-minute gradient of 10% to 40% acetonitrile in 0.1% formic acid at 275 nL/min. The nanoLC was connected in-line to a hybrid triple quadrupole linear ion trap mass spectrometer, 4000 Q Trap from ABI/MDS Sciex (Framingham, MA) equipped with Micro Ion Spray Head II ion source. 2D-gel analysis was done using Biorad IEF protean cell with ReadyStrip IPG strips (pH 4.0 - 7.0). Econo-pac 10 DG desalting columns were also acquired from Bio-rad (Hercules, CA). Dialysis was done using Novagen D-tube dialyzer Maxi MWCO 3.5 kDa (EMD biosciences).

3.2.1 Synthesis of lissamine rhodamine sulfonyl azide

53 mg of lissamine rhodamine sulfonyl chloride (92 mmol) was dissolved in 10 mL acetone in a round-bottomed flask wrapped in aluminium foil. 29 mg of sodium azide (446 mmol, 5 eq.) was then added and the solution was stirred at room-temperature for 24 h. The solvent was removed *in vacuo* and the residue was re-dissolved in dichloromethane and washed with water. The organic layer was dried over anhydrous $MgSO_4$, filtered and removed in vacuo. Yield: 44.6 mg (83 %). Product was dissolved in DMSO to give a concentration of 15 mM. ^{1}H NMR (400 MHz, DMSO) δ 8.43 (d, 1H), 8.23 (t, 1H), 7.68 (dd, 1H), 7.13 – 6.85 (m, 6H), 3.65 (s, 8H), 2.10 (d, 1H), 1.21 (t, 12H), 1.14 (s, 1H). ESI-MS (Positive mode): m/z = 584.

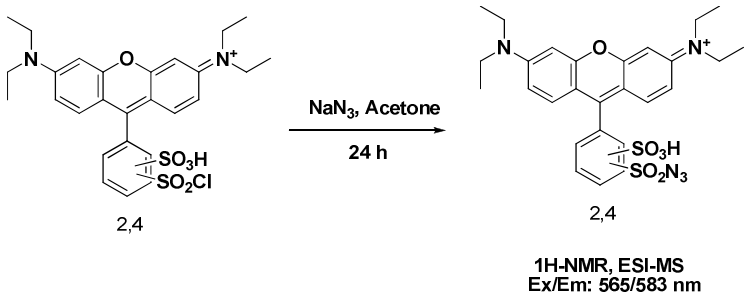


Figure 3.3: Synthesis of lissamine rhodamine sulfonyl azide from lissamine rhodamine sulfonyl chloride.

3.2.2 Over-expression and purification of T.thermophilus ThiS thiocarboxylate (TtThiSCOSH) and the corresponding DTT adduct (TtThiS/DTT)

TtThiS cloned in pTYB1 was obtained from Protein facility, Department of Chemistry and Chemical Biology, Cornell University, Ithaca. The protein was over-expressed in *E.coli* BL21(DE3). 2 L cultures were grown at 37^{o}C in Luria-Bertani till an OD_{600} of 0.6 when the temperature was reduced to 15^{o}C and the cultures were induced with

IPTG to a final concentration of 1 mM. Further growth was carried out at 15^{o}C for 12-16 h with constant agitation. The cultures were harvested by centrifugation and lysed by sonication on ice in 20 mM Tris, 500 mM NaCl, 1 mM EDTA, 0.1% Triton X-100, pH 7.8. The samples were then loaded onto 20 mL chitin beads column at a rate of 0.5 mL/min and washed with 300 mL of 20 mM Tris, 500 mM NaCl, 1 mM EDTA, pH 7.8 at rate of 2 mL/min. Cleavage of the protein was carried out at 4^{o}C for 48 h with 30 mL of 50 mM DTT to give ThiS/DTT and 30 mL of 50 mM Na_2S to yield ThiSCOSH. *T.thermophilus* proteins were buffer-exchanged by dialysis into 100 mM potassium phosphate, pH 8.0 and stored at -80^{o}C as 30% glycerol stock aliquots. No reducing agent was added to the frozen stocks.

3.2.3 Labeling of TtThiSCOSH and TtThiS/DTT

The frozen aliquots of the proteins were thawed and the samples were buffer-exchanged into 50 mM potassium phosphate, 6 M urea, pH 6.0. 50 µL of 184 µM of both the proteins were then treated with 1.5 µL of 15 mM Lissamine rhodamine sulfonyl azide (2.5 eq.). The samples were incubated at room-temperature in the dark for 15 min. 6 µL of 250 mM TCEP (in 1 M potassium phosphate, pH 6.0) was added. The samples were analyzed on 15% Tris-glycine SDS-PAGE and imaged on Typhoon 9400.

3.2.4 Determination of time of incubation

TtThiSCOSH was buffer-exchanged into 50 mM potassium phosphate, 6 M urea, pH 6.0. 100 µL of 93 µM protein was incubated with 2 µL of 15 mM Lissamine rhodamine sulfonyl azide (3 eq., stock made in DMSO). 10 µL aliquots were taken at time-points 0, 5, 10, 15, 30, 45, 60 and 75 min and treated with 10 µL of SDS-PAGE

sample buffer containing 50 mM TCEP and analyzed on a 15% Tris-glycine SDS-PAGE gel before imaging on Typhoon 9400.

3.2.5 Specificity and sensitivity of labelling

E.coli BL21(DE3) cells were grown in 75 mL Luria-Bertani at 37^{o}C till an OD_{600} of 1.6. The cells were harvested by centrifugation and lysed by sonication on ice in 4 mL of 50 mM potassium phosphate, 300 mM NaCl, 2 mM TCEP, pH 8.0. The sample was again centrifuged to obtain the cell-free extract, which was buffer-exchanged using Econo-pac 10 DG desalting columns into 50 mM potassium phosphate, 9 M urea, pH 6.0. Three 90 µL samples of the above extract containing pure TtThiSCOSH (in 100 mM potassium phosphate, 30% glycerol, pH 8.0) to a final concentration of 11 µM, 1.1 µM and 110 nM were made and treated with 20 µL of 15 mM lissamine rhodamine sulfonyl azide for 15 min at room-temperature in the dark followed by treatment with 25 mM TCEP for another 30 min in the dark. The samples were desalted by $CHCl_3$/methanol precipitation and then resuspended into 50 µL of 50 mM potassium phosphate, 9 M urea, pH 6.0. An equal volume of SDS-PAGE sample buffer was added and they were analyzed on 16% Tris-tricine gel for imaging on Typhoon trio (excitation: 532 nm green laser; emission: 580-nm band-pass filter (580 BP 30)).

3.2.6 Labeling different bacterial cell-free extracts

***3.2.6.1 P.stutzeri* KC (ATCC 55595)** was maintained on a nutrient broth agar plate at 4^{o}C and a colony was used to inoculate 100 mL culture of DRM media[4] in sterile disposable PETG flask with vented closure. The culture was grown at 30^{o}C for 48 h with shaking. The culture was harvested by centrifugation. *P.stutzeri KC* cell-pellet

was resuspended in 3 mL of 50 mM potassium phosphate, 300 mM NaCl, 2 mM TCEP, pH 8.0 containing 6.5 mg/mL of bacterial protease inhibitor cocktail. The sample was lysed by sonication on ice and the cell-free extract was obtained by centrifugation. The supernatant was then treated with Na_2S, ATP and $MgCl_2$ to a final concentration of 9 mM, 18 mM and 6 mM respectively and the sample was incubated at room-temperature for 6 h. The sample was dialysed extensively into 50 mM ammonium acetate and was freeze-dried. It was then re-suspended into 1 mL of 50 mM potassium phosphate, 6 M urea, pH 6.0. 100 µL of denatured cell-free extract was treated with 20 µL of 15 mM rhodamine sulfonyl azide for 15 min at room-temperature in the dark and then with 12 µL of 250 mM TCEP(dissolved in water) and further incubated at room-temperature for 30 min in the dark. The sample was precipitated by $CHCl_3$/methanol and subjected to 2D-gel analysis (pH 4-7, 7 cm IEF strip, active rehydration @ 50V for 12 h at 20^{o}C. Four step focussing (20^{o}C): S1: 250V, 15 min, S2: 4000V, linear voltage ramp, 2h, S3: 4000V, rapid voltage ramp, 20000Vh, S4: 500V, hold. Current limit/gel: 50 µA). The lower molecular weight fluorescent spot, imaged on Typhoon trio (excitation: 532 nm green laser; emission: 580-nm band-pass filter (580 BP 30)), was excised out of the gel and subjected to mass-spectrometry analysis.

3.2.6.2 S.coelicolor **(ATCC 10147)** maintained on yeast malt extract glucose agar at 4^{o}C was used to inoculate a 100 mL culture of M9 minimal media containing 400 mg glucose, 2 mM $MgSO_4$ and 100 µM $CaCl_2$ in a sterile disposable PETG flask with vented closure with constant agitation. The culture was incubated at 30^{o}C for 2-3 days till a good cell-growth was observed before being used to inoculate a larger 1.5 L M9 minimal media culture. The larger culture was further incubated at 30^{o}C for 5 days with shaking. The culture was spun-down by centrifugation and the cell-pellet

obtained was resuspended in 45 mL of 50 mM potassium phosphate, 300 mM NaCl, 2 mM TCEP, 55 mg of bacterial protease inhibitor cocktail, pH 8.0. Cells were lysed by sonication on ice and the cell-free extract was treated with 6 mM $MgCl_2$, 9 mM Na_2S and 18 mM ATP for 6 h at room- temperature. The sample was dialyzed extensively against 50 mM ammonium acetate and freeze-dried. The residue was then re-dissolved in 5 mL of 50 mM potassium phosphate, 9 M urea, pH 6.0. 300 µL of denatured cell-free extract was treated with 60 µL of 15 mM rhodamine sulfonyl azide for 15 min at room-temperature in the dark and then with 36 µL of 250 mM TCEP(dissolved in water) and further incubated at room-temperature for 30 min in the dark. The sample was precipitated by $CHCl_3$/methanol and subjected to 2D-gel analysis (pH 4-7, 17 cm IEF strip, active rehydration @ 50V for 12 h at 20^{o}C. Four step focussing (20^{o}C): S1: 250V, 15 min, S2: 10000V, linear voltage ramp, 3h, S3: 10000V, rapid voltage ramp, 60000Vh, S4: 500V, hold. Current limit/gel: 50 µA).

3.2.6.3 Growth conditions for other organisms used in this study

Typical growth conditions for other organisms subjected to the labelling strategy to identify thiocarboxylate-forming proteins in them are shown in **Table 3.2**. All the cultures were harvested by centrifugation and lysed by sonication (in case of *S.avermitilis* and *S.griesus*, this step was preceded by homogenization). The remaining protocol is same as that mentioned in **Sub-section 3.2.6.1**.

3.2.7 Dependence of protein labelling on expression conditions

P.stutzeri KC (ATCC 55595) from a nutrient broth agar plate maintained at 4^{o}C was used to inoculate 100 mL culture of DRM media[4] in sterile disposable PETG flask with vented closure. The culture was grown at 30^{o}C for 48 h with shaking. The culture was harvested by centrifugation. For iron-rich conditions, nutrient broth was used to

Table 3.2: Growth conditions for micro-organisms. All organisms were grown at 30°C in sterile disposable PETG flask with vented closure.

Organism	Liquid culture media	Time of incubation
Burkolderia xenovorans LB400 maintained at 4°C on nutrient broth agar plate	M9 supplemented with 4 g/L glucose, 2 mM $MgSO_4$ and 100 µM $CaCl_2$.	Grown till OD_{600} of 0.6
Rhodococcus sp. RHA1 maintained at 4°C on tryptic soy agar plate	M9 salts with 400 mg glucose, 2 mM $MgSO_4$ and 100 µM $CaCl_2$.	Grown till OD_{600} of 0.6
S.griesus maintained at 4°C on yeast malt extract glucose agar plate	M9 salts with 400 mg glucose, 2 mM $MgSO_4$ and 100 µM $CaCl_2$.	4-5 days
S.avermitilis maintained at 4°C on yeast malt extract glucose agar plate	KH_2PO_4 0.5 g/L, Asparagine 0.5g/L, Glucose 10 g/L, Adjust pH to 7.0, Sterile filtered, 2 mL / L of 1 M $MgSO_4$, 100 µL / L of 1 M $CaCl_2$	4-5 days
S.erythrea maintained at 4°C on yeast malt extract glucose agar plate	M9 salts with 400 mg glucose, 2 mM $MgSO_4$ and 100 µM $CaCl_2$.	4-5 days

grow the bacteria at 30°C till an OD_{600} of 1.0. The cell-pellets were then re-suspended in 3 mL of 50 mM potassium phosphate, 300 mM NaCl, 2 mM TCEP, pH 8.0 containing 6.5 mg/mL of bacterial protease inhibitor cocktail. The samples were lysed

by sonication on ice and the cell-free extract was obtained by centrifugation. Na_2S, ATP and $MgCl_2$ were added to a final concentration of 9 mM, 18 mM and 6 mM respectively to all samples shown in the **Table 3.3** except sample 1 and the samples were then incubated at room-temperature for 6 h. Freshly prepared iodoacetic acid (in 1 M potassium phosphate, pH 8.0) was added to sample 3 to a final concentration of 100 mM. The samples were further incubated at room-temperature for 1 h and then, buffer-exchanged by dialysis into 50 mM ammonium acetate. The samples were freeze-dried and the residue obtained was re-suspended into 50 μL of 50 mM potassium phosphate, 6 M urea, pH 6.0. 2 μL of 15 mM lissamine rhodamine sulfonyl azide was added and the samples were incubated in the dark at room-temperature for 15 min. Equal volumes of 50 mM TCEP in 100 mM potassium phosphate, pH 8.0 were added to the samples and further incubation was carried out in the dark for 30 min. The solutions were then precipitated using $CHCl_3$/methanol and re-suspended in 50 mM potassium phosphate, 6 M urea, pH 6.0. Equal volumes of SDS-PAGE sample buffer containing 50 mM TCEP were added and the samples were analyzed by a 16% Tris-tricine gel, which was imaged on Typhoon 9400.

Table 3.3: Sample preparations for studying the dependence of labelling on bacterial growth condtions.

Sample number	**Treatment with Na_2S/ATP/$MgCl_2$**	**Samples: 1 mL cell-free extract of *P.stutzeri* KC grown in**
1	No	DRM media
2	Yes	DRM media
3	Yes	DRM media
4	Yes	nutrient broth

3.3 Results and discussions

T.thermophilus ThiS thiocarboxylate (TtThiSCOSH) and TtThiS/dithiothreitol adduct were made by cleaving the intein construct with sulfide and DTT respectively. The dye labeled the thiocarboxylate protein and not the DTT adduct suggesting that it is labeling the thioacid functionality specifically (**Figure 3.4(a)**). Approximately, 70% labeling was achieved in 15 min (**Figure 3.4(b)**).

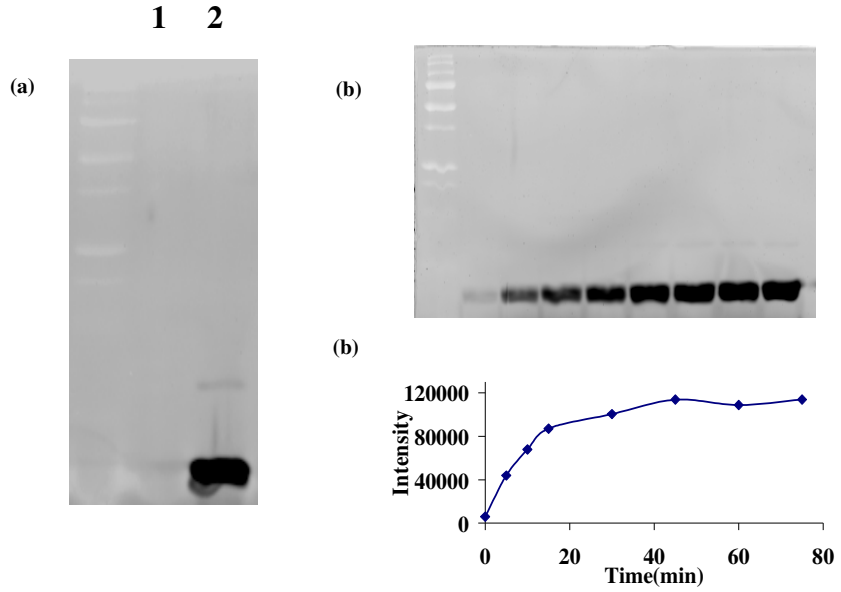


Figure 3.4: (a) Lane 1: ThiS/DTT adduct does not get labeled with the dye while lane 2: TtThiSCOSH does. ***(b)*** Time-dependence of labeling of TtThiSCOSH. Time-points taken: 0, 5, 10, 15, 30, 45, 60 and 75 mins. PMT voltage: 300 V (Typhoon 9400)

Having been able to label pure TtThiSCOSH, the issue of specificity in a cell-free extract was addressed next. TtThiS and human ubiquitin[5] used in a previous study, does not have cysteine. A more general model, *E.coli* BL21(DE3) cell-free extract, was chosen to test the specificity and sensitivity of the sulfonyl azide. The labeling was carried out at pH 6.0 to prevent possible reaction of the dye with the nucleophilic thiol side-chain of cysteine (**Figure 3.5**). This chemistry can be similar to

the reduction of azides to amines using dithiols where sulfenamide adduct (compound 1, **Figure 3.5**) is proposed to be an intermediate[6]. Sulfenamide linkages, however, has been shown to be susceptible to reduction by soft nucleophiles like thiols and phosphines [6,7]. Hence, the labeling reaction is followed by treatment with TCEP/β-mercaptoethanol to break any possible non-specific reaction between the dye and cysteine thiols.

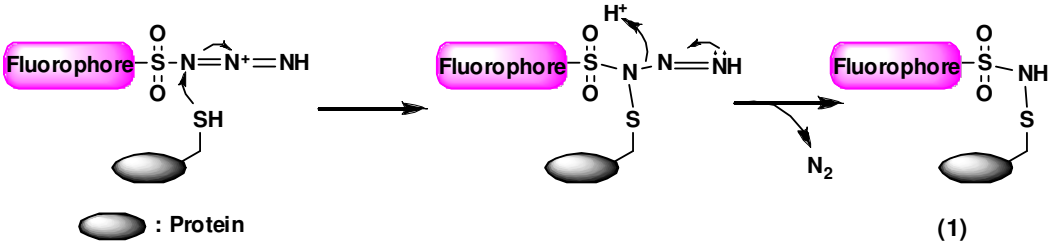


Figure 3.5: Possible side reaction of the sulfonyl azide dye with cysteine sidechains.

Pure TtThiSCOSH was added to the cell-free extract of *E.coli* BL21(DE3) in varying concentrations and labeling was carried out under denaturing conditions (9 M urea) to determine the sensitivity limits of the dye (**Figure 3.6**). TtThiSCOSH could be seen to be clearly labeled up to 9.9 X 10^{-11} moles in 90 µL sample volume (concentration of 1.1 µM), which corresponds to approximately 464 copies of the protein assuming the volume of an *E.coli* cell as 0.7 X 10^{-15} L [8] (One copy of any protein in an *E.coli* cell corresponds to 1/(6.023 X 10^{23}) moles ≡ 2.37 nM. This would imply that a concentration of 1.1 µM would correspond to approximately 464 copies). Low levels of labeling are also observed for the higher molecular-weight proteins in the gel shown in **Figure 3.6**. It still remains to be seen whether they are non-specific or possibly other thiocarboxylate-forming proteins. The fluorescence intensity of the labeled proteins (both the higher molecular-weight and the thiocarboxylated proteins)

increases when the gel is scanned at higher PMT voltages of the instrument as compared to lower PMT.

To test the method on a proteome, *Pseudomonas stutzeri KC* (ATCC 55595) was chosen. *P.stutzeri KC* is known to produce a siderophore, pyridine dithiocarboxylate (PDTC) under iron-limiting conditions[9]. PDTC is responsible for the dechlorination of carbon tetrachloride (CCl_4) displayed by the organism. The biosynthesis of this small molecule requires a protein, PdtH. Sequence analysis of PdtH identified the protein to be small and has the characteristic diglycyl C-terminus, both features suggesting that it could possibly transport sulfur as a thiocarboxylate.

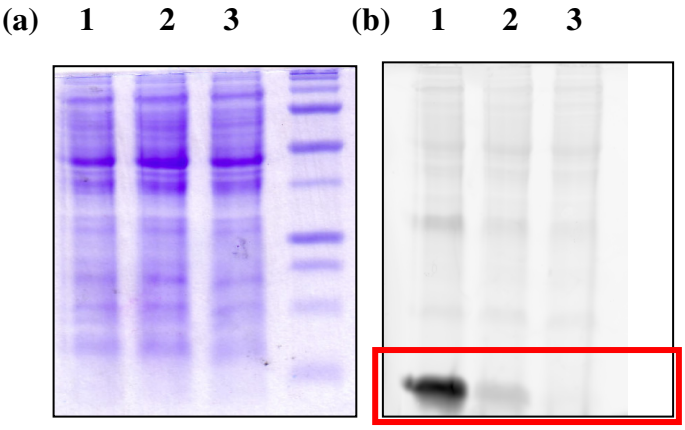


Figure 3.6: Sensitivity of labeling of TtThiSCOSH in the presence of *E.coli* BL21 (DE3) cell-free extract. Lanes 1 (11 µM), 2 (1.1 µM) and 3 (110 nM). 15 µL samples of the concentrations mentioned were loaded in each lane. a) coomassie staining b) fluorescent image. PMT voltage: 400 V (Typhoon trio)

Burkolderia xenovorans LB400 and *Rhodococcus* sp. RHA1, other polychlorinated biphenyl-degrading micro-organisms, were also chosen for this study to see if they harbor any thiocarboxylate-forming protein. *Streptomyces coelicolor* (ATCC 10147) an antibiotic-producing actinomycete, was also tested along with *Streptomyces avermitilis*, *Streptomyces griesus* and *Saccharopolyspora erythrea.* The cell-free

extracts of all the chosen micro-organisms were treated with sulfide and ATP to form protein thiocarboxylate (**Figure 3.7**) before subjecting it to labeling.

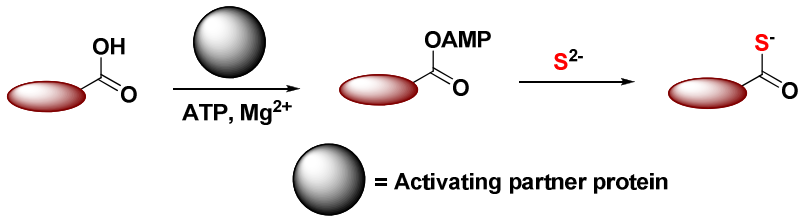


Figure 3.7: Reconstitution of protein thiocarboxylates

P.stutzeri KC and *S.coelicolor* showed protein labeling in the lower-molecular weight region, whose intensity was higher than the background (**Figure 3.8**). 2D-SDS-PAGE was used for analysis of the labeled proteome. This provides better resolution of the proteins according to their isoelectric points and molecular weights.

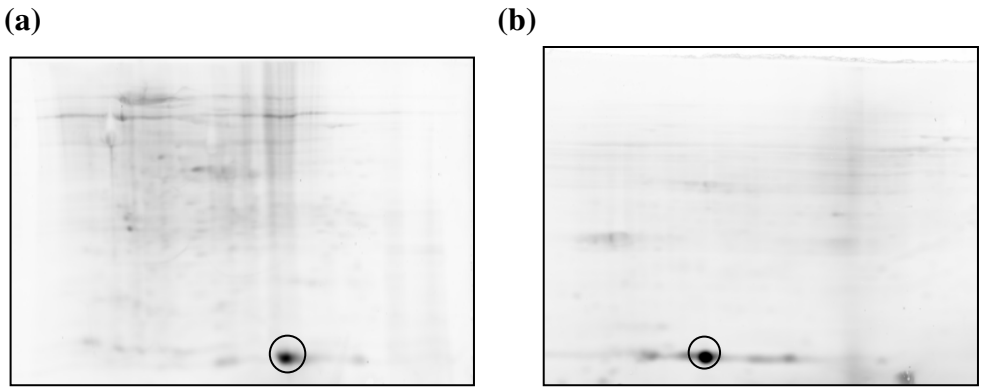
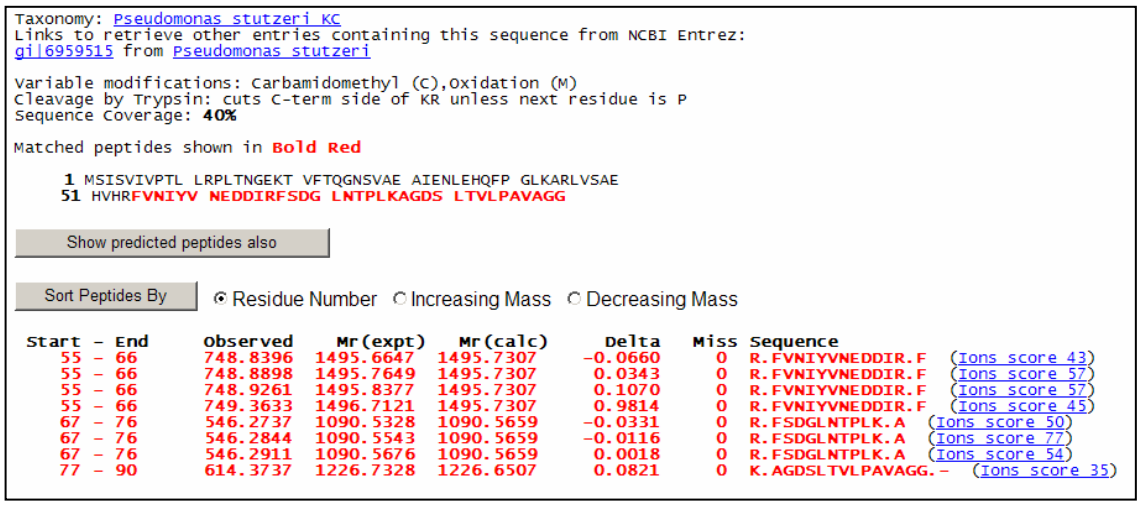


Figure 3.8: 2D-gel of (a) *P.stutzeri* KC cell-free extract, 7 cm IEF strip, pH 4-7, PMT voltage: 400 V (b) *S.coelicolor* cell-free extract, 17 cm IEF strip, pH 4-7, PMT voltage: 400 V (Typhoon trio)

The labeled protein in *P.stutzeri* KC cell-free extract was isolated and subjected to LC-MS/MS analysis (**Figure 3.9** and **Figure 3.10**). The protein was identified as PdtH, the small sulfur-transfer protein that provides sulfur for the biosynthesis of the

siderophore, pyridine dithiocarboxylic acid (PDTC). Identity of the labeled protein in *S.coelicolor* proteome is currently being established. *S.avermitilis* also showed weak labeling in the lower-molecular weight region of the SDS-PAGE. Attempts are being made to isolate a large amount of the labeled protein to identify it. *Burkolderia xenovorans* LB400, *Streptomyces griesus*, *Saccharopolyspors erythrea* and *Rhodococcus* sp. RHA1 did not show any tagged protein in the lower molecular-weight region.



Taxonomy: Pseudomonas stutzeri KC
Links to retrieve other entries containing this sequence from NCBI Entrez:
gi|6959515 from Pseudomonas stutzeri

Variable modifications: Carbamidomethyl (C),Oxidation (M)
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
Sequence Coverage: **40%**

Matched peptides shown in **Bold Red**

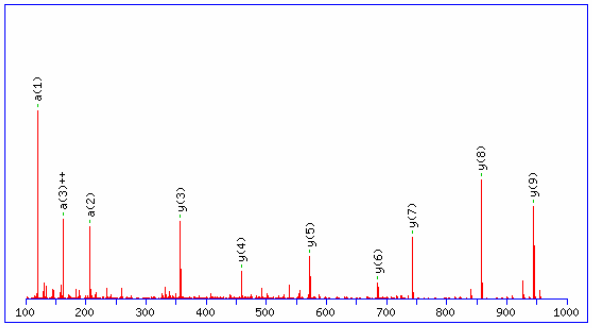
1 MSISVIVPTL LRPLTNGEKT VFTQGNSVAE AIENLEHQFP GLKARLVSAE
51 HVHRFVNIYV NEDDIRFSDG LNTPLKAGDS LTVLPAVAGG

Show predicted peptides also

Sort Peptides By ◉ Residue Number ○ Increasing Mass ○ Decreasing Mass

Start - End	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Sequence	
55 - 66	748.8396	1495.6647	1495.7307	-0.0660	0	R.FVNIYVNEDDIR.F	(Ions score 43)
55 - 66	748.8898	1495.7649	1495.7307	0.0343	0	R.FVNIYVNEDDIR.F	(Ions score 57)
55 - 66	748.9261	1495.8377	1495.7307	0.1070	0	R.FVNIYVNEDDIR.F	(Ions score 57)
55 - 66	749.3633	1496.7121	1495.7307	0.9814	0	R.FVNIYVNEDDIR.F	(Ions score 45)
67 - 76	546.2737	1090.5328	1090.5659	-0.0331	0	R.FSDGLNTPLK.A	(Ions score 50)
67 - 76	546.2844	1090.5543	1090.5659	-0.0116	0	R.FSDGLNTPLK.A	(Ions score 77)
67 - 76	546.2911	1090.5676	1090.5659	0.0018	0	R.FSDGLNTPLK.A	(Ions score 54)
77 - 90	614.3737	1226.7328	1226.6507	0.0821	0	K.AGDSLTVLPAVAGG.-	(Ions score 35)

Figure 3.9: LC-MS/MS analysis of the lower molecular weight protein spot of *P.stutzeri* KC cell-free extract. The protein was subjected to in-gel trypsin digestion followed by oxidation of methionines and carbamidomethylation of the cysteines (with iodoacetamide). 40% sequence coverage was obtained and the matched peptides are shown in red.



Monoisotopic mass of neutral peptide Mr(calc): 1090.5659
Ions Score: 77 Expect: 0.00022
Matches (Bold Red): 10/88 fragment ions using 12 most intense peaks

#	a	a^{++}	a*	a^{*++}	b	b^{++}	b*	b^{*++}	Seq.	y	y^{++}	y*	y^{*++}	#
1	**120.0808**	60.5440			148.0757	74.5415			F					10
2	**207.1128**	104.0600			235.1077	118.0575			S	**944.5047**	472.7560	927.4782	464.2427	9
3	322.1397	**161.5735**			350.1347	175.5710			D	**857.4727**	429.2400	840.4462	420.7267	8
4	379.1612	190.0842			407.1561	204.0817			G	**742.4458**	371.7265	725.4192	363.2132	7
5	492.2453	246.6263			520.2402	260.6237			L	**685.4243**	343.2158	668.3978	334.7025	6
6	606.2882	303.6477	589.2617	295.1345	634.2831	317.6452	617.2566	309.1319	N	**572.3402**	286.6738	555.3137	278.1605	5
7	707.3359	354.1716	690.3093	345.6583	735.3308	368.1690	718.3042	359.6558	T	**458.2973**	229.6523	441.2708	221.1390	4
8	804.3886	402.6980	787.3621	394.1847	832.3836	416.6954	815.3570	408.1821	P	**357.2496**	179.1285	340.2231	170.6152	3
9	917.4727	459.2400	900.4462	450.7267	945.4676	473.2374	928.4411	464.7242	L	260.1969	130.6021	243.1703	122.0888	2
10									K	147.1128	74.0600	130.0863	65.5468	1

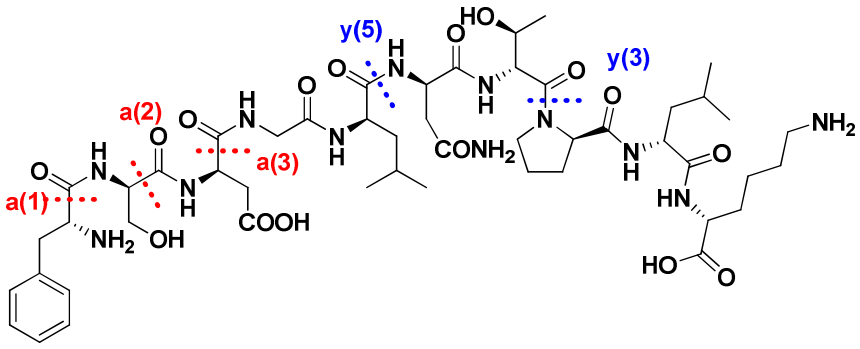


Figure 3.10: MS/MS fragmentation of FSDGLNTPLK (a peptide fragment). The fragmentation of the peptide happens as shown below the mass-spectra. A y-ion is a fragment from the C-terminus that cleaves after the nitrogen of the amide bond while an a-ion is a fragment from the N-terminus that cleaves before the amide carbonyl. ++ indicates a doubly charged fragment. The number in parentheses in the MS spectra (example a(1) or y(9)) indicate the number of amino-acids from the N-terminus (for an a-ion) or C-terminus (for a y-ion) after which cleavage occurs.

Labeling of the proteins was also dependent upon the expression conditions of the micro-organisms. The formation of the fluorescent band in P.stutzeri KC was seen only when the bacteria were cultured in iron-limiting conditions. The band whose molecular weight was less than 10 kDa could not be seen if rich media is used suggesting repression of protein expression under conditions of iron abundance (box, Lane 2, **Figure 3.11**).

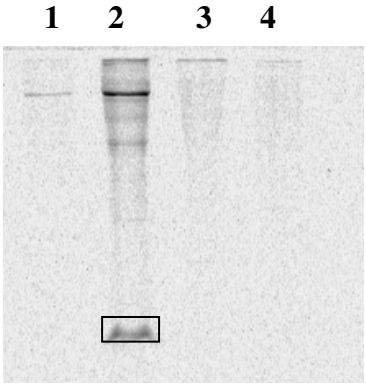


Figure 3.11: Labeling of *P.stutzeri* KC cell-free extract. Lane 1: cell-free extract, Lane 2: cell-free extract treated with sulfide and ATP prior to labeling, Lane 3: cell-free extract treated with sulfide and ATP and treated with iodoacetic acid to cap the thiocarboxylate, Lane 4: cell-free extract of *P.stutzeri KC* grown in nutrient broth. PMT voltage: 250 V (Typhoon 9400)

3.4 Conclusions

The involvement of protein thiocarboxylates in important pathways necessitates the development of a strategy to identify them. We have developed a proteomics tool to label these proteins in cell-free extracts and have successfully detected an important sulfur-transfer protein in the cell-free extract of *P.stutzeri* KC. Formation of PdtH thiocarboxylate has also been demonstrated for the first time. The method can detect thiocarboxylated proteins with copy numbers as low as 500. The labeling was found to

be dependent on the growth conditions of the micro-organism as illustrated by the PdtH protein expression. Finding new thiocarboxylate-forming proteins essential for bacterial survival as in the case of biosynthesis of cofactors like thiamin or amino-acid biosynthesis like cysteine might lead to pathways with interesting biochemistry and open up new vistas for drug discovery.

REFERENCES

1. Shangguan, N., Katukojvala, S., Greenberg, R. & Williams, L. J. The reaction of thio acids with azides: a new mechanism and new synthetic applications. *J Am Chem Soc* **125**, 7754-5 (2003).
2. Kolakowski, R. V., Shangguan, N., Sauers, R. R. & Williams, L. J. Mechanism of thio acid/azide amidation. *J Am Chem Soc* **128**, 5695-702 (2006).
3. Bradford, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**, 248-54 (1976).
4. Lee, C. H., Lewis, T. A., Paszczynski, A. & Crawford, R. L. Identification of an extracellular agent [correction of catalyst] of carbon tetrachloride dehalogenation from Pseudomonas stutzeri strain KC as pyridine-2, 6-bis(thiocarboxylate). *Biochem Biophys Res Commun* **261**, 562-6 (1999).
5. Zhang, X., Li, F., Lu, X. W. & Liu, C. F. Protein C-terminal modification through thioacid/azide amidation. *Bioconjug Chem* **20**, 197-200 (2009).
6. Cartwright, I. L., Hutchinson, D. W. & Armstrong, V. W. The reaction between thiols and 8-azidoadenosine derivatives. *Nucleic Acids Res* **3**, 2331-9 (1976).
7. Ueki, M., Maruyama, H. & Mukaiyama, T. Peptide synthesis by oxidation-reduction condensation. I. Use of NPS peptides as amino component. *Bulletin of the chemical society of japan* **44**, 1108-1111 (1971).
8. Kubitschek, H. E. Cell volume increase in Escherichia coli after shifts to richer media. *J Bacteriol* **172**, 94-101 (1990).
9. Sebat, J. L., Paszczynski, A. J., Cortese, M. S. & Crawford, R. L. Antimicrobial properties of pyridine-2,6-dithiocarboxylic acid, a metal chelator produced by Pseudomonas spp. *Appl Environ Microbiol* **67**, 3934-42 (2001).

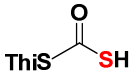
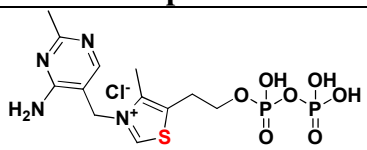
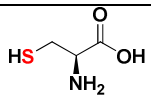
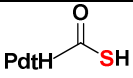
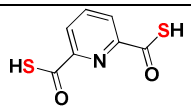
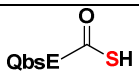
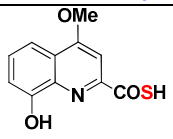
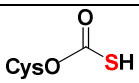
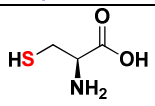
CHAPTER 4

From sulfate to methionine via protein thiocarboxylate

4.1 Introduction

Thiocarboxylate-forming proteins have been shown to be involved in a variety of biosynthetic pathways, prominent among these being vitamin B_1 (ThiS)[1], molybdopterin (MoaD)[2], cysteine (CysO)[3], thioquinolobactin (QbsE)[4], 2-thioribothymidine (TtuB)[5] and 5-methoxy-carbonyl-methyl-2-thiouridine (Urm1p)[6]. Despite their important role in sulfur transfer, there is a lack of information about their sulfur source, except for ThiS and MoaD thiocarboxylate which have been shown to derive their sulfur from L-cysteine[2,7,8] (**Table 4.1**).

Table 4.1: Sulfur source of thiocarboxylate-forming proteins involved in various biosynthetic pathways

Thiocarboxylate protein	End-product	Sulfur source
	 Thiamin	
	 Pyridine dithiocarboxylic acid	??
	 Thioquinolobactin	??
	 L-cysteine	??

A search for sulfur transfer proteins in the genomic database, ***theseed.uchicago.edu***, showed a small putative thiocarboxylate-forming protein clustered with the sulfate assimilation proteins in *Wolinella succinogenes,* a member of the Helicobacteraceae family (**Figure 4.1**).

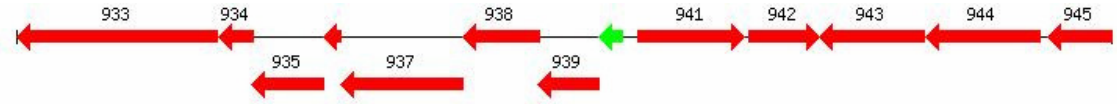


Figure 4.1: The putative thiocarboxylate forming protein, 936, clustered with sulfate assimilation proteins – 933, 937, 938 and 939, and homocysteine biosynthetic proteins, 941 and 944 in *Wolinella succinogenes*. Putative functions of the genes are mentioned in **Table 4.2.**

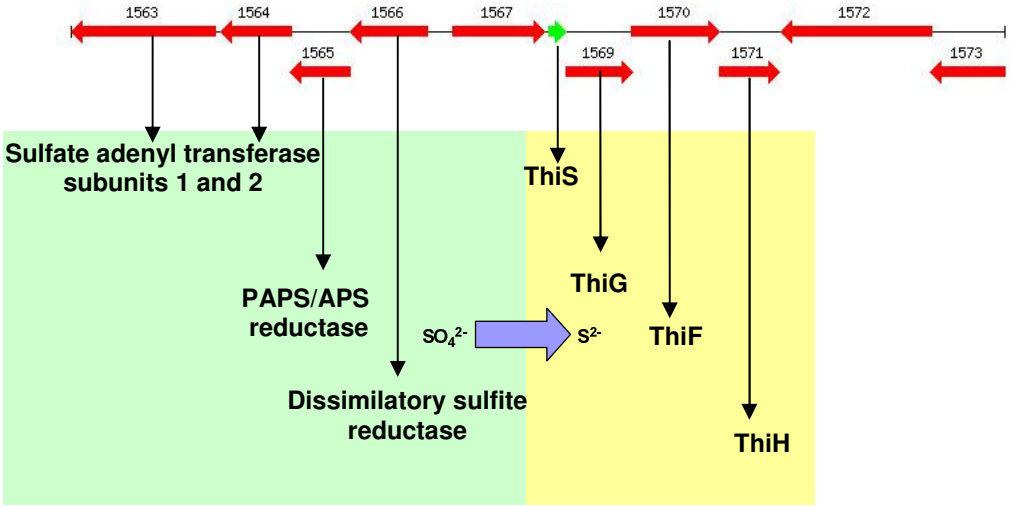
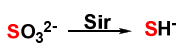
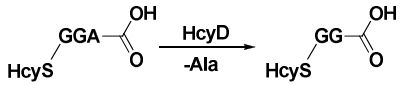
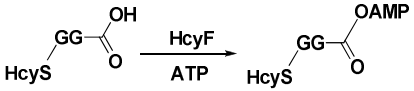
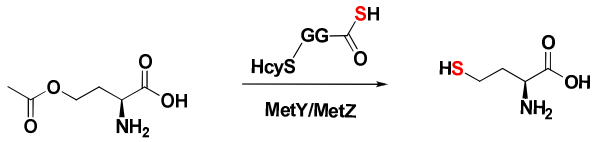
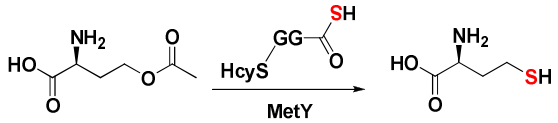
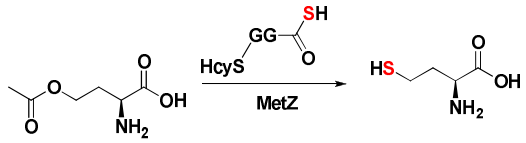


Figure 4.2: Sulfate as possible sulfur source for thiamin biosynthesis in *Acidovorax sp. JS42* and *Pelodictyon luteolum DSM 273*

Table 4.2: Gene names and abbreviations used in this chapter

Gene ID	Gene names	Gene abbreviation	Reactions catalyzed
933	Ferredoxin-sulfite reductase	Sir	
934	Hypothetical protein	HcyD	
935	Molybdpterin biosynthetic protein MoeB	HcyF	
936	Sulfur transfer protein involved in thiamin biosynthesis	HcyS	
941	O-acetylhomoserine sulfhydrylase	MetY	
944	O-acetylhomoserine sulfhydrylase/ O-succinylhomoserine sulfhydrylase	MetZ	

This clustering pattern and variations of it was seen in many other micro-organisms like *Clostridium kluyveri, Clostridium thermocellum, Desulfitobacterium hafniense, Carboxydothermus hydrogenoformans Z-2901, Caldicellulosiruptor saccharolyticus DSM 8903, Alkaliphilus metalliredigens QYMF, Geobacter metallireducens GS-15, Geobacter uraniireducens Rf4, Acidovorax sp. JS42* and *Pelodictyon luteolum DSM 273* suggesting that they are functionally related.

Prevalence of such arrangement of genes in many organisms suggests direct utilization of sulfate as a sulfur source for protein thiocarboxylates. In certain organisms such as *Acidovorax sp. JS42* and *Pelodictyon luteolum DSM 273*, the genes of sulfate assimilation proteins are present in the neighbourhood of thiamin biosynthetic genes (**Figure 4.2**) indicating the possibility of sulfate being the sulfur source of thiamin in these organisms instead of cysteine (**Figure 4.3**).

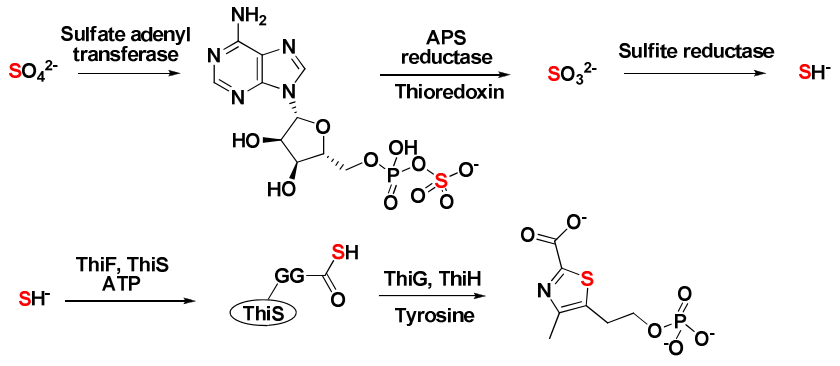


Figure 4.3: Putative mechanism of thiazole formation from sulfate in *Acidovorax sp. JS42* and *Pelodictyon luteolum DSM 273*

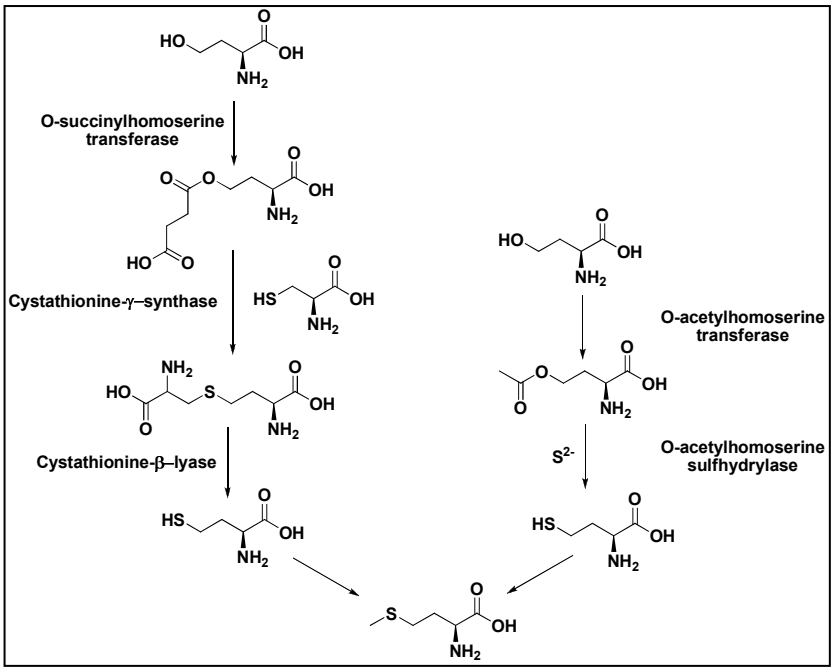


Figure 4.4: Two common pathways for L-methionine biosynthesis

The biosynthesis of homocysteine involves direct sulfhydrylation of *O*-succinylhomoserine or *O*-acetylhomoserine by the corresponding sulfhydrylases, or breakdown of cystathionine by cystathionine-β-lyase[9] (**Figure 4.4**). The homocysteine biosynthesis in *W.succinogenes* is dependent on protein-thiocarboxylate for sulfur transfer rather than a direct sulfhydrylation of O-acetyl-L-homoserine. Methionine is then made by methylation of L-homocysteine. In this work, we have also shown sulfite (derived from sulfate) being utilized directly in *Wolinella succinogene*s to make the protein thiocarboxylate.

4.2 Experimental section

*W.succinogenes g*enomic DNA (ATCC 29543D-5) was purchased from ATCC (Manassas, VA). pTYB1, SapI, NdeI, XhoI and chitin beads were bought from New England Biolabs (Ipswich, MA). EMD biosciences (Gibbstown, NJ) supplied Luria-

Bertani, glycerol, chloroform, methanol, acetonitrile, potassium phosphate and ammonium acetate. IPTG, amplicillin and kanamycin were procured from Lab Scientific (Livingston, NJ). L-cysteine, arabinose, DL-methionine, D-glucose, M9 minimal salts, formic acid, EDTA, urea, D_2O, triton X-100, methyl viologen, TCEP, O-acetyl-L-serine, O-succinyl-L-homoserine, DL-homocysteine, 10% Pt (platinum) on activated carbon, hydroxocobalamin hydrochloride, S-adenosylmethionine chloride, 3-mercaptopropionic acid and o-phthalaldehyde were purchased from Sigma-Aldrich (St.Louis, MO). Ferrous ammonium sulfate, nickel sulfate, chloramphenicol, boric acid, Na_2SO_3 and Tris.HCl were acquired from Fisher Scientific (Fairlawn, NJ). $CaCl_2$, $ZnSO_4$, $MgCl_2$, $MgSO_4$ and NaCl were got from Mallinckrodt. Microcon YM-10 (MWCO 10 kDa), YM-3 (MWCO 3 kDa) and amicon (MWCO 5 kDa) cellulose filters were obtained from Millipore Corporation (Billerica, MA). Aminolevulinic acid, ATP and sodium sulfide were from Acros. O-acetyl-L-homoserine and 5-DL-methyltetrahydrofolic acid calcium salt trihydrate was bought from TRC Canada (North York, Ontario). Typhoon trio and chelating sepharose fast flow (used for Ni-NTA affinity purification) were products of GE healthcare biosciences (Piscataway, NJ). All bacterial cultures were grown on New Brunswick Scientific (Edison, NJ) Excella E25 shaker incubator and lysed by sonication using Misonix sonicator 3000 (Misonix Inc., Farmingdale, NY). Absorbance data was collected on Cary 300 Bio UV-visible spectrophotometer (Varian, Palo Alto, CA). Glove box was made by Coy Laboratory products (Grass lakes, MI). ESI-MS analysis was performed using an Esquire-LC_00146 instrument (Bruker, Billerica, MA) in the positive ion mode. MALDI-MS data was recorded in positive mode on Applied Biosystems Voyager STR (matrix: sinapinic acid). LC-MS data was obtained on an Agilent 1200 capillary HPLC system interfaced to an API QSTAR Pulsar Hybrid QTOF mass spectrometer (Applied Biosystems/MDS Sciex, Framingham, MA) equipped with an electrospray

ionization (ESI) source. Liquid chromatography (LC) separation was achieved using a Phenomenex Jupiter C4 microbore column (150 × 0.50 mm, 300 Å) at a flow rate of 10 μL/min (solvent A: water + 0.1% HCOOH, solvent B: acetonitrile + 0.1% HCOOH, 0 - 5 mins: 2% B; 8 min: 30% B; 52 min: 60% B; 60 min: 95% B, 65 min: 95% B). MALDI-MS and LC-MS data were provided by Laboratory of Biological Mass spectrometry at Texas A&M University, College Station, Texas. All the protein concentrations were measured by Bradford assay [10]. The protein stock samples are in 100 mM Tris, 150 mM NaCl, 2 mM TCEP, 30% glycerol, pH 8.0 unless otherwise mentioned.

4.2.1 Cloning and over-expression of the proteins

sir, *hcyD*, *hcyF*, *hcyS*, *metY* and *metZ* were all cloned from *Wolinella succinogenes* FDC 602W into THT vector (a pET-28 derived vector which allows attachment of a modified 6xHisTag followed by a TEV protease site onto the N-terminus of the expressed protein) between NdeI/XhoI restriction sites. Three methyltransferase genes in the genome - Ws250, Ws942 and Ws1672 that could possibly make methionine from homocysteine were also cloned into THT vector. Sequence analysis suggests Ws250 and Ws1672 are 5-methyltetrahydropteroyltriglutamate-homocysteine methyltransferases, metE. Ws942, the methyltransferase gene present along with the sulfate assimilation and homocysteine biosynthetic genes, is annotated as a SAM-dependent methyltransferase. *Salmonella typhimurium cysG* (siroheme synthase) was cloned into pACYCDuet vector. All the genes except *metZ* (forward primer: CAGCACATGCATATGCCAGCCCACAAAGATGAGACT; Reverse primer: TTATTCCGCTCGAGTTAAGCTTTGGCTAGGGCTTG) and *Ws942* (forward primer: CAGCACATGCATATGCACCCCATGACCCTTGATG; Reverse primer: AATGTTTGCTCTTCCGCTCGAGTCATTTCCTCCATGAGATAAA) were

purchased from Protein facility, Department of Chemistry and Chemical Biology, Cornell University. Methionine-auxotroph *E.coli* B834(DE3) harboring the *iscS*-cluster assembly genes in the vector pDB1282 was a gift from Dr.Steve Ealick (Department of Chemistry and Chemical biology, Cornell University, Ithaca).

All the proteins, except WsSir, were over-expressed in *E.coli* BL21(DE3). Luria-Bertani cultures containing 40 mg kanamycin per liter were grown at 37^{o}C till an OD_{600} of 0.6, cooled to 15^{o}C and then induced with a final concentration of 500 µM IPTG before continuing growth at 15^{o}C for another 12 -16 h.

WsSir was over-expressed along with *S. typhimurium* CysG and *A. vinelandii* IscS cluster in *E.coli* B834(DE3). 1.5 L M9 minimal media cultures supplemented with 30 mL 20% glucose, 3 mL of 1 M $MgSO_4$, 150 µL of 1 M $CaCl_2$, 120 mg DL-methionine, 150 mg ampicillin and 60 mg each of kanamycin and chloramphenicol. The cultures were grown at 37^{o}C till an OD_{600} of 0.1. At this point, 3.75 g of L-arabinose, 88 mg of ferrous ammonium sulfate and 90 mg of L-cysteine were added and the cultures were shaken at 100 rpm till the OD_{600} reached 0.6 when they cooled for 4 h at 4^{o}C. 45 mg of aminolevulinic acid and a final concentration of 0.5 mM IPTG were added and the cultures were further shaken at 100 rpm for 12-16 h at 15^{o}C.

All the cultures were harvested by centrifugation and the cell-pellets were lysed by sonication on ice. The proteins were purified by Ni-NTA purification protocol at 4^{o}C. All buffers contained 1 mM TCEP. After affinity column purification, all proteins were desalted into 100 mM Tris, 150 mM NaCl, 2 mM TCEP, 30% glycerol, pH 8.0 and stored as frozen aliquots at -80^{o}C.

Truncated WsHcyS (with the C-terminus alanine removed) was cloned into an intein construct, pTYB1, between the NdeI/SapI restriction sites (Forward primer: CAGCACATGCATATGAATCTCATCATCAACGGAGAGAATAA; Reverse primers: AATGTTTGCTCTTCCGCAGCCTCCTCCCATGAAATATAAA). To

make truncated WsHcyS-COSH and WsHcyS-(DL)-homocysteine, WsHcyS in pTYB1 was over-expressed in *E.coli* BL21(DE3). Luria-Bertani cultures containing 100 mg of ampicillin per liter were grown at 37^{o}C till an OD_{600} of 0.6-0.8 when the temperature was reduced to 15^{o}C and the cultures were induced with a final concentration of 0.5 mM IPTG. Further growth was carried out at 15^{o}C for 12-16 h with constant agitation. The cultures were harvested by centrifugation and lysed by sonication on ice in 20 mM Tris, 500 mM NaCl, 1 mM EDTA, 0.1% Triton X-100, pH 7.8. The samples were then loaded onto 20 mL chitin beads column at a rate of 0.5 mL/min and washed with 300 mL of 20 mM Tris, 500 mM NaCl, 1 mM EDTA, pH 7.8 at rate of 2 mL/min. Cleavage of the protein was carried out at 4^{o}C for 12-16 h with 30 mL of 50 mM Na_2S to give truncated WsHcyS-COSH and 30 mL of 50 mM DL-homocysteine to yield truncated WsHcyS-(DL)-homocysteine. They were desalted into 100 mM Tris, 150 mM NaCl, 2 mM TCEP, 30% glycerol, pH 8.0 by dialysis using Novagen D-tube dialyzer Maxi (MWCO 3.5 kDa) and stored as frozen aliquots at -80^{o}C.

4.2.2 Activity of gene product WsHcyD (putative metalloprotease)

200 µL of 611 µM WsHcyS was treated with 6 µL of 2.4 mM WsHcyD and 3 µL of 10 mM $ZnSO_4$ at room-temperature for 2 h. The samples were desalted into 200 µL of 50 mM NH_4OAc. 200 µL of acetonitrile was then added along with 2 µL of HCOOH to the sample before being analyzed by positive-mode ESI-MS.

4.2.3 Product of WsHcyD cleavage of WsHcyS-GGA – identification of released alanine

The proteins were desalted twice into 50 mM potassium phosphate, pH 8.0 using Bio-rad bio-spin 6 columns. 85 µL of approximately 28 mM WsHcyS-GGA was treated

with 100 µL of 1.62 mM WsHcyD at room-temperature for 2 h. The sample was freeze-dried and re-dissolved in D_2O. The proteins were filtered off using YM-3 microcons (which had been washed extensively with D_2O to remove glycerol before sample loading) and the filterate was analyzed on Varian 500 MHz NMR.

4.2.4 Activity of gene product WsHcyF (putative MoeB-like protein)

30 µL of 3.98 mM WsHcyS, 23 µL of 1.76 mM WsHcyF, 30 µL of 0.88 mM WsHcyD, 3 µL of 10 mM ATP and 6 µL of 10 mM $MgCl_2$ were added together and incubated at room-temperature for 15 min. The controls (**Figure 4.10**) had the corresponding component missing. The reactions were quenched with equal volumes of 12 M urea (in water) before the proteins were removed using YM-10 microcons and the samples analyzed by HPLC (Agilent 1200 using Supelco supelcosil LC-18-T (15 cm X 4.6 mm, 3 µm)) using the following gradient at a flow rate of 1 mL/min: solvent A is water, solvent B is 100 mM potassium phosphate, pH 6.6, solvent C is methanol. 0 min: 100% B; 7 min: 10% A, 90% B; 12 min: 25% A, 60% B, 15% C; 17 min: 25% A, 10% B, 65% C; 19 min: 100% B, 25 min: 100% B

4.2.5 Full conversion of ATP to AMP by WsHcyF

Two separate samples containing 30 µL of 3.98 mM WsHcyS, 30 µL of 1.33 mM WsHcyF, 15 µL of 1.61 mM WsHcyD, 3 µL of 10 mM ATP and 2 µL of 30 mM $MgCl_2$ were added together and incubated at room-temperature for 15 min and 120 min respectively. The proteins were then removed using YM-3 microcons and analyzed by HPLC (Agilent 1200 using Supelco supelcosil LC-18-T (15 cm X 4.6 mm, 3 µm)) using the following gradient at a flow rate of 1 mL/min: solvent A is water, solvent B is 100 mM potassium phosphate, pH 6.6, solvent C is methanol. 0

min: 100% B; 7 min: 10% A, 90% B; 12 min: 25% A, 60% B, 15% C; 17 min: 25% A, 10% B, 65% C; 19 min: 100% B, 25 min: 100% B

4.2.6 Formation of protein WsHcyS thiocarboxylate from sulfite using WsSir (putative sulfite reductase)

4.2.6.1 Preparation of reduced methyl viologen

149.5 mg of methyl viologen was dissolved in 12.5 mL of 50 mM potassium phosphate, pH 8.0 along with 15 mg of 10% Pt on activated carbon in a 15 mL centrifuge tube. Argon was bubbled through the solution for 5 min after which hydrogen was bubbled through the sample for 30 min. The sample was quickly sealed with parafilm and centrifuged for 5 min to remove the catalyst. It was then taken into the glove-box and the supernatant was transferred to a new 15 mL centrifuge tube. The concentration of the reduced methyl viologen was measured at 600 nm (extinction coefficient 1.3 X 10^4 $M^{-1}cm^{-1}$) [11].

4.2.6.2 Observation by labeling with lissamine rhodamine sulfonyl azide

30 µL of 3.98 mM WsHcyS, 30 µL of 1.33 mM WsHcyF, 10 µL of 1.61 mM WsHcyD and 9 µL of 379 µM WsSir were added together outside the glove-box and moved into the glove-box for 2 h incubation with the lids open (to let the oxygen diffuse out and make the sample anerobic). A similar sample was set-up without WsHcyD as a control. After 2 h, 30 µL of 10 mM ATP, 4 µL of 1 M $MgCl_2$, 2.44 µL of 100 mM Na_2SO_3 and 100 µL of 3.5 mM reduced methyl viologen were added to both the samples inside the glove-box and further incubated within the glove-box for 15 min. The samples were quenched by exposing to air and shaking. They were then desalted using Bio-rad bio-spin 6 columns into 100 µL of 50 mM NH_4OAc, 6 M urea, pH 6.0 treated with 7.4 µL of 15 mM rhodamine sulfonyl azide for 15 min in the dark

at 26°C. They were then desalted by chloroform:methanol precipitation and analyzed by a 15% tris-glycine SDS-PAGE. The fluorescence image of the labeled protein in the gel was obtained on Typhoon Trio imager (excitation: 532 nm green laser; emission: 580-nm band-pass filter (580 BP 30)).

4.2.6.3 Observation of WsHcyS-COSH by LC-MS

Thiocarboxylate formation was also assayed by LC-MS. The same procedure was used as mentioned above. After the quenching, the samples were heated at 100°C for 5 min and left to cool to room-temperature for an hour. The protein precipitate was removed and the filterate was concentrated and desalted using Bio-rad bio-spin 6 columns into 50 mM NH_4OAc, pH 6.0 and analyzed by LC-MS.

4.2.7 Protein thiocarboxylate as sulfur source for homocysteine biosynthesis

95 µL of 159 µM truncated WsHcyS-COSH and 11 µL of 1.4 mM WsMetY were incubated with the appropriate substrate – 1.25 µL of 10 mM O-acetyl-L-serine in one case and 1.25 µL of 10 mM O-acetyl-L-homoserine in the other. The samples were incubated at room-temperature for 1 h. They were then desalted into 50 mM NH_4OAc, pH 6.0 and analyzed by MALDI-MS. The time-period of incubation was then reduced to 2 min when it was realized that the higher molecular-weight 7907 Da adduct was formed.

4.2.8 Homocysteine formation from sodium sulfide and O-acetyl-L-homoserine using WsMetY

80 µL of 1.4 mM WsMetY (desalted into 50 mM potassium phosphate, pH 8.0) was mixed with 320 µL of 50 mM potassium phosphate, 80% D_2O, pH 8.0 along with 50 µL each of 100 mM sodium sulfide and 100 mM O-acetyl-L-homoserine. Sample was

incubated for 1 h at 26°C. Control sample had the above composition except for the protein. Samples were freeze-dried and re-dissolved in 100% D_2O. ^{1}H-NMR taken on Varian 300 MHz.

4.2.9 Reaction of truncated WsHcyS-COSH with WsMetZ

2.85 µL of 1.4 mM WsMetZ and 95 µL of 42 µM truncated WsHcyS-COSH were incubated with the appropriate substrate – 0.4 µL of 10 mM O-acetyl-L-serine, 0.4 µL of 10 mM O-acetyl-L-homoserine and 0.4 µL of 10 mM O-succinyl-L-homoserine for 3 min at room-temperature. The samples were desalted into 50 mM NH_4OAc, pH 6.0.

4.2.10 Homocysteine release from WsHcyS-homocysteine adduct

95 µL of 337 µM truncated WsHcyS-(DL)-homocysteine was mixed with 3.6 µL of 880 µM WsHcyD and incubated at room-temperature for 5 min. The control sample had WsHcyD absent. The samples were then treated with 98.6 µL of 9 M urea, desalted into 50 mM NH_4OAc, pH 6.0 using Bio-rad bio-spin 6 chromatography columns and analyzed by MALDI-MS.

4.2.11 WsHcyD hydrolyzing the WsHcyS-COSH to form WsHcyS-COOH

95 µL of 42 µM truncated WsHcyS-COSH was mixed with 3.6 µL of 88 µM WsHcyD and incubated at room-temperature for 5 min. The control sample had WsHcyD absent. The samples were then treated with 99.5 µL of 9 M urea, desalted into 50 mM NH_4OAc, pH 6.0 using Bio-rad bio-spin 6 chromatography columns and analyzed by MALDI-MS.

4.2.12 Identification of the released product from WsHcyS-homocysteine adduct made from truncated WsHcyS-COSH

190 µL of 42 µM WsHcyS-COSH and 8 µL of 1 mM WsMetY were mixed together, buffer-exchanged and concentrated using YM-5 amicons to 100 µL of 50 mM potassium phosphate, pH 8.0. 0.4 µL of 250 mM TCEP and 0.8 µL of 10 mM O-acetyl-L-homoserine were then added and incubated at room-temperature for 5 min. 10 µL of 76 µM WsHcyD in 50 mM potassium phosphate, pH 8.0 was added and further incubation was carried out for 5 min at room-temperature. The proteins were removed using YM-3 microcons. The samples were then treated with a final concentration of 1 mM final concentration o-phthalaldehyde (dissolved in methanol) and analyzed by HPLC (Agilent 1200 using Supelco supelcosil LC-18-T (15 cm X 4.6 mm, 3 µm)) using the following gradient at a flow rate of 1 mL/min: solvent A is water, solvent B is 100 mM potassium phosphate, pH 6.6, solvent C is methanol. 0 min: 100% B; 7 min: 10% A, 90% B; 12 min: 25% A, 60% B, 15% C; 17 min: 25% A, 10% B, 65% C; 19 min: 100% B, 25 min: 100% B.

4.2.13 Conversion of homocysteine to methionine by the methyltransferases

Two samples containing 100 µL of 169 µM of Ws250 and Ws1672 were incubated with 20 µL of 50 mM potassium phosphate, pH 8.0, 1.69 µL of 5-DL-methyltetrahydrofolate, 1.69 µL of 10 mM $MgSO_4$ and 1.69 µL of DL-homocysteine. MetE is a Zn^{2+} utilizing enzyme that has been reported to require phosphate and its activity is stimulated by divalent ions like Mg^{2+} or Mn^{2+} [12]. Hence, potassium phosphate and $MgSO_4$ was added to the reaction sample. The third sample contained 100 µL of 169 µM Ws942 (volume made upto 120 µL with 100 mM Tris, 150 mM NaCl, 2 mM TCEP, 30% glycerol, pH 8.0) along with 1.69 µL of S-adenosylmethionine, 1.69 µL of DL-homocysteine and 0.48 µL of 250 mM TCEP. All

the samples were incubated at room-temperature for 1 h. The proteins were then removed by passing the samples through YM-10 amicons. 150 µL of the flow-through was treated with 50 µL of o-phthalaldehyde derivatization solution (1 mL of 37 mM o-phthalaldehyde in MeOH, 4 mL of 0.1 M boric acid, pH 9.3, 162 µL of 3-mercaptopropionic acid were mixed together and pH was adjusted to 9.3 with NaOH) for 5 min at room-temperature before being treated with 20 µL of 1 M potassium phosphate, pH 6.0 and analyzed on HPLC (Agilent 1200 using Phenomenex Gemini 5µ C18 110A (15 cm X 4.6 mm, 5 µm)) using the following gradient at a flow rate of 1 mL/min: solvent A is water, solvent B is 100 mM potassium phosphate, pH 6.6, solvent C is methanol. 0 min: 100% B; 7 min: 10% A, 90%B; 12 min: 25% A, 60% B, 15% C; 17 min: 25% A, 10% B, 65% C; 19 min: 100% B, 25 min: 100% B.

4.3 Results and discussion

4.3.1 Growth and over-expression

Genes *sir, hcyD, hcyF, hcyS, metY*, *metZ* and three putative methionine biosynthetic genes (Ws250, Ws942, Ws1672) (**Table 4.2**) were cloned into THT vector. All the proteins except WsSir were over-expressed in *E.coli* BL21 (DE3) in Luria-bertani medium. WsSir was over-expressed in *E.coli* B834 (DE3), a methionine-auxotroph *E.coli*. WsSir, annotated as ferridoxin-sulfite reductase, was over-expressed along with *Salmonella typhimurium* siroheme synthase (cloned into pACYDuet) and *Azotobacter vinelandii* IscS-cluster assembly proteins (cloned into pDB1282) in M9 minimal media. Over-expression with these proteins helps in good reconstitution of the siroheme and [4Fe-4S] cofactors and eliminating either one results in poor reconstitution. The proteins were purified by normal Ni-NTA affinity purification protocol. WsSir was purified outside the glove-box at 4^{o}C. It had the characteristic absorbance of a siroheme-[4Fe-4S] cluster protein [13] (**Figure 4.5**). WsMetY and

WsMetZ, putative PLP-containing proteins, were yellow in color suggestive of bound PLP.

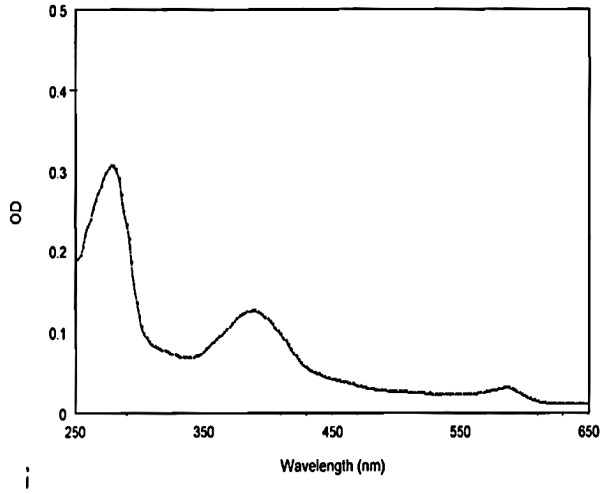


Figure 4.5: Absorbance of WsSir with its characteristic maximums at 388 nm and 590 nm

4.3.2 Activity of gene product WsHcyD (putative metalloprotease)

Thiocarboxylate forming proteins usually end in diglycyl C-terminus (**Figure 4.6**). Sequence analysis of gene product, WsHcyD, suggests that it is a metalloprotease belonging to the Mov34/MPN/PAD-1 protein family.

```
                1       10        20        30        40        50        60        70        80        90  95
                |-------+---------+---------+---------+---------+---------+---------+---------+---------+---|
E.coli                    MQILFNDQAMQCAAGQTVHELLEQLDQRQAGAA----------------LAINQQIVPREQWAQHIVQDGDQILLFQVIAGG
M.tuberculosis  MNVTVSIPTILRPHTGGQKSVSASGDTLGAVISDLEANYSGISERLMDPSSPGKLHRFVNIYVNDEDVRFSGGLATAIADGDSVTILPAVAGG
P.fluorescens     MKITVPTLMRNLTANQKTVEAQGQTVRQVIDALEQQYPGIAARLVEGN---KVHKFINIYINDDDIRFLQSLDSAVAQHDQITILPAVAGGCF
Wolinella                         MNLIINGENKSFEKEGLSVKELLVLESV--KMPEMVSIQLNDEFLREPEYATTSLKEGDTINFLYFMGGGA
Consensus       .........m..l...q....a.gqtv......l*q...g.a..l.......k......i.iN##.vr..q.a...v.#gDq!..l...aGG..
```

Figure 4.6: Sequences of known thiocarboxylate forming proteins. All known examples have diglycyl terminus at their C-terminus in their active form.

Previous members of this family associated with a thiocarboxylate forming proteins are Mec^{+} and QbsD. Mec^{+} cleaves off the product, cysteine from CysO C-terminus in cysteine biosynthesis in *M.tuberculosis* [3]while QbsD removes the last two

amino-acids (CF) from QbsE in thioquinolobactin biosynthesis in *P.fluorescens* (ATCC 17400) [4] (**Figure 4.7**). It is possible that WsHcyD could be cleaving the last amino-acid (alanine) from WsHcyS.

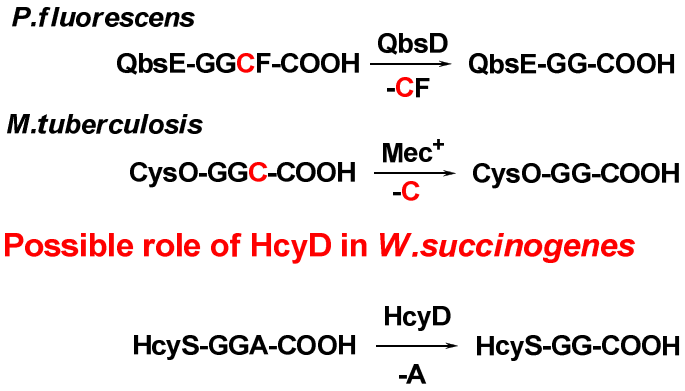


Figure 4.7: Processing of the C-terminus of thiocarboxylate forming proteins by the partner metalloproteases.

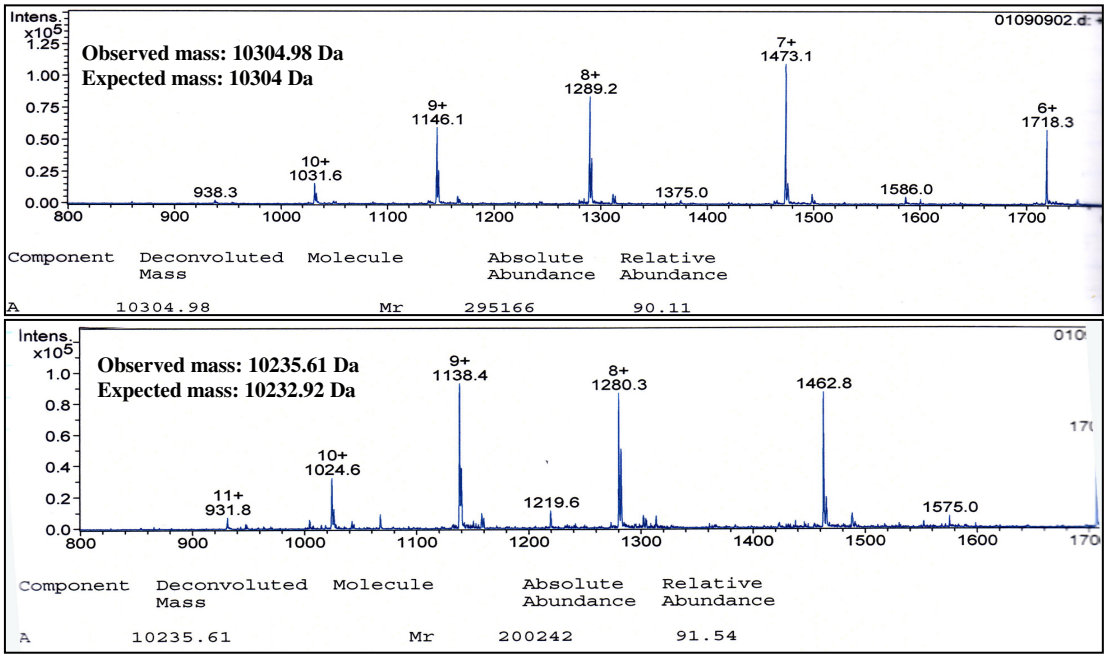


Figure 4.8: ESI-MS showing different charge states of WsHcyS (Top panel) Full-length WsHcyS (Bottom panel) The C-terminal alanine is removed from WsHcyS upon treatment with WsHcyD.

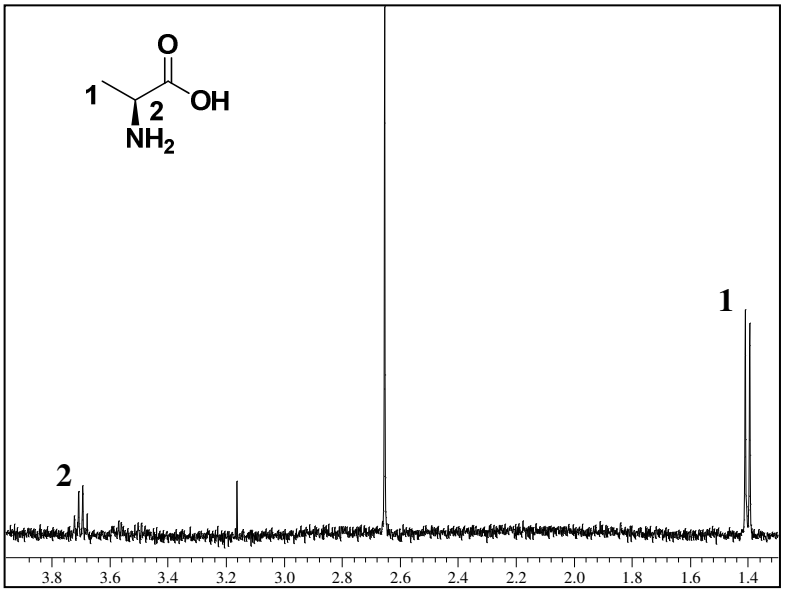


Figure 4.9: ^{1}H-NMR of L-alanine released from WsHcyS C-terminus upon treatment with WsHcyD. No exogenous Zn^{2+} was added.

This proved to be the case when the two proteins were mixed together in the presence of exogenous Zn^{2+} (**Figure 4.8**). ICP-MS analysis suggested sub-stoichiometry occupancy of Zn^{2+} in WsHcyD and hence, the protein was found to be active even in the absence of exogenously added Zn^{2+}. ^{1}H-NMR of the small-molecule product proved conclusively that L-alanine is released from WsHcyS (**Figure 4.9**).

4.3.3 Activity of gene product WsHcyF (putative MoeB-like protein)

Gene WsHcyF is annotated as molybdopterin biosynthetis protein, MoeB. The protein is speculated to activate the C-terminus of truncated gene product WsHcyS as an acyl-adenylate to make it ready for nucleophilic attack by sulfide. The acyl-adenylate, being unstable, would hydrolyze in the absence of any suitable nucleophile to form adenosine monophosphate. The hypothesis was tested and the release of AMP was analyzed by HPLC. AMP formation was dependent on the presence of WsHcyF,

WsHcyS and WsHcyD (**Figure 4.10**). Full conversion of the ATP to AMP occurred within a time-period of 2 h. (**Figure 4.11**).

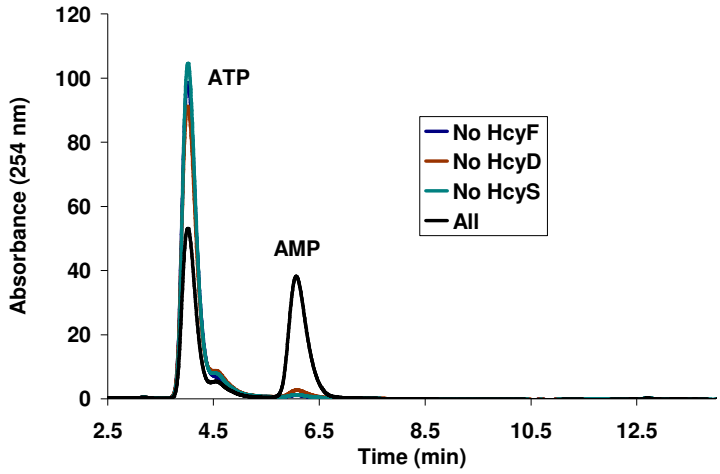


Figure 4.10: Adenylating activity of WsHcyF. WsHcyF adenylates the trunctated WsHcyS terminus. The acyl-adenylate hydrolyses in the absence of the right nucleophile (S^{2-}) to give AMP.

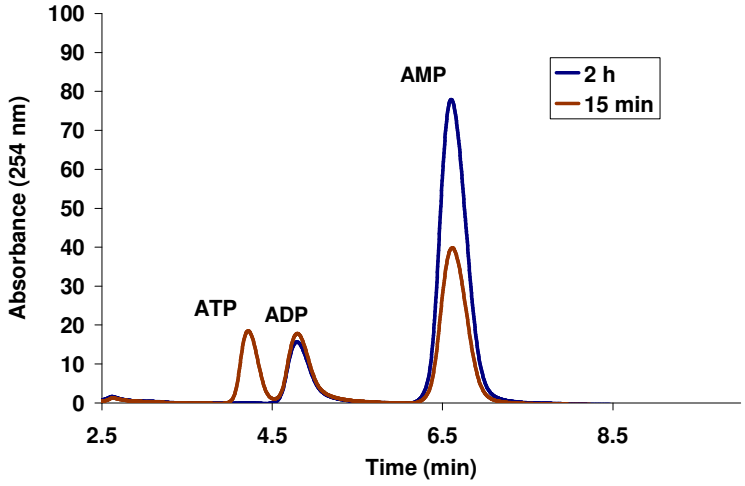


Figure 4.11: Full-conversion of ATP to AMP occurred within 2 h.

4.3.4 Formation of protein WsHcyS thiocarboxylate from sulfite using WsSir (putative sulfite reductase)

The formation of HcyS thiocarboxylate necessitates removal of the alanine first from the C-terminus by HcyD (causing a mass decrease from 10304 Da to 10232.92 Da), followed by activation of the protein by HcyF as an acyl-adenylate. The sulfur for the transformation was provided by reduction of SO_3^{2-} to S^{2-} by WsSir with reduced methyl viologen (obtained by reduction over 10% Pt on activated carbon) as the electron donor (**Figure 4.12**).

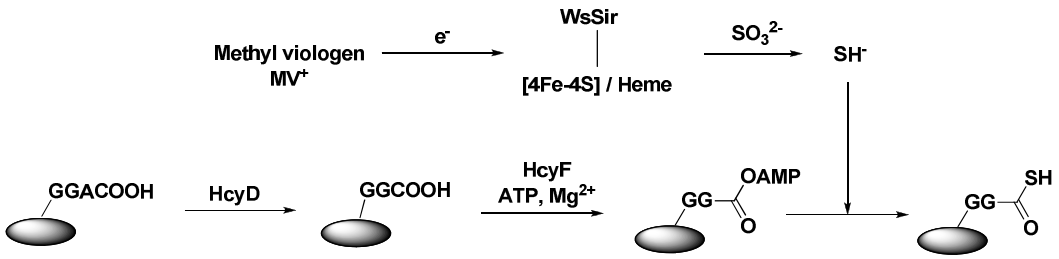


Figure 4.12: Reconstitution of WsHcyS-COSH using methyl viologen as the electron donor and sulfite as the sulfur source.

The conversion of WsHcyS C-terminus (without alanine) from a carboxylate to a thiocarboxylate resulted in a mass-increase of 16 Da (observed mass: 10248.8 Da, expected mass: 10248.92 Da, error: 0.001%) monitored by LC-MS (**Figure 4.13**). The selective transfer to truncated WsHcyS occurs because of the selective activation of the truncated protein by WsHcyF (**Figure 4.10**).

Thiocarboxylate specific labeling strategy was also used to prove the protein thiocarboxylate formation. Small molecule thiocarboxylates have been shown previously to react with sulfonyl azides to form N-acyl-sulfonamide linkages[14,15]. This chemistry has been exploited in a proteomics scale (unpublished results) to identify protein thiocarboxylates in bacterial cell-free extracts and was used in this study to

assay for the formation of WsHcyS-COSH. The sulfur transfer to WsHcyS was tested by click chemistry between a fluorescent sulfonyl azide, lissamine rhodamine

Figure 4.13: Formation of WsHcyS-COSH monitored by LCMS analysis. In the absence of WsHcyD (top panel), the C-terminal is not adenylated by WsHcyF and no WsHcyS thiocarboxylate (10249 Da) is formed.

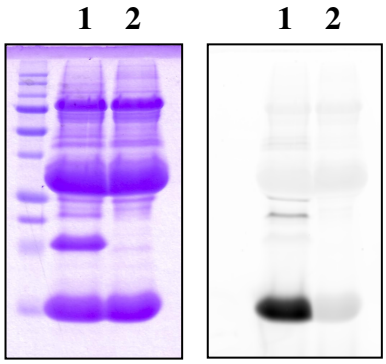


Figure 4.14: SDS-PAGE analysis of WsHcyS-COSH formation after tagging the protein with the fluorescent, lissamine rhodamine sulfonyl azide (a) coomassie image (b) fluorescent image. Lane 1: Sample containing all components – WsHcyD, WsHcyF, WsHcyS, WsSir, SO_3^{2-}, reduced methyl viologen (electron donor), ATP, Mg^{2+} Lane 2: WsHcyD removed from the reaction sample. PMT voltage: 400 V.

sulfonyl azide, and the thiocarboxylate formed (**Figure 3.2, Chapter 3**). Synthesis of the fluorescent dye was carried out as shown in **Figure 3.3, Chapter 3**. The tagged protein was analyzed on SDS-PAGE and scanned for fluorescence using Typhoon trio imager (**Figure 4.14**).

4.3.5 Protein thiocarboxylate as sulfur source for homocysteine biosynthesis

It would be interesting to know the final destination of the sulfur on WsHcyS-COSH. The presence of two genes, O-acetylhomoserine sulfhydrylase (WsMetY) and O-succinyl/O-acetylhomoserine sulfhydrylase (WsMetZ), suggests the possibility of the thiocarboxylate acting as a sulfur source to make homocysteine. WsMetY was tested as a possible homocysteine biosynthetic protein. Thiocarboxylate form of truncated WsHcyS (with the last alanine removed) was made by cleaving the protein-intein construct with Na_2S and used as a substrate along with O-acetyl-L-homoserine for WsMetY.

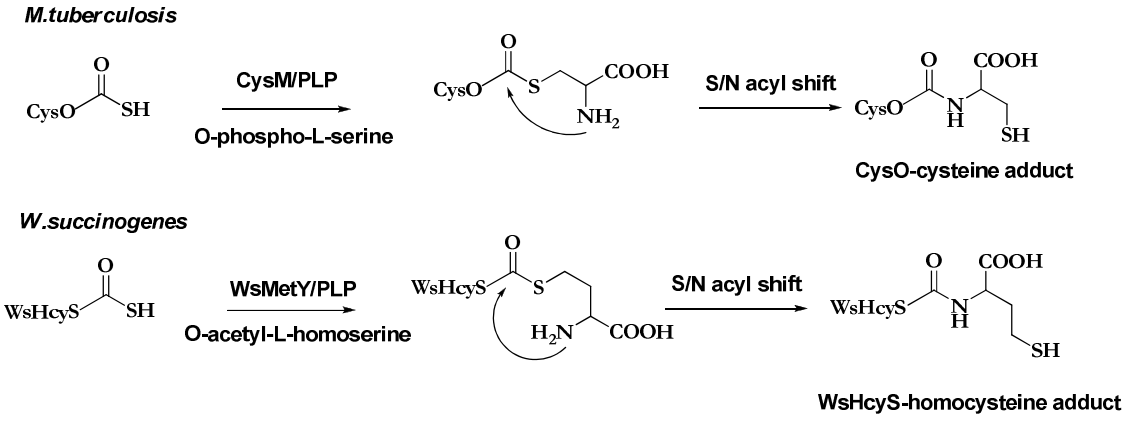


Figure 4.15: Formation of CysO-cysteine adduct in M.tuberculosis and a similar formation of WsHcyS-homocysteine adduct in *W.succinogenes*

In cysteine biosynthetic pathway in *M.tuberculosis*, CysO-thiocarboxylate is converted to CysO-cysteine adduct upon treatment with PLP-dependent O-phospho-L-serine sulfhydrylase, CysM and O-phospho-L-serine[3,16] (**Figure 4.15**). WsHcyS may similarly form WsHcyS-homocysteine adduct in the presence of WsMetY and O-acetyl-L-homoserine.

WsMetY is selective to O-acetyl-L-homoserine as a substrate (**Figure 4.16 (a),** observed mass: 7809.83 Da, expected mass: 7805.96 Da, error: 0.05%) and does not act on O-acetyl-L-serine even after extended incubation time (**Figure 4.16 (b)**, observed mass: 7709.83 Da, expected mass: 7805.96 Da). Longer incubation times or higher concentration of O-acetyl-L-homoserine results in a higher molecular weight, WsHcyS adduct being formed (m/z = 7911 Da), the identity of which has not been established but could possibly be homolanthionine (**Figure 4.17**). Homolanthionine has been previously observed as a side-product of MetB, cystathionine-γ-synthase in *Corynebacterium glutamicum* ATCC 13032 where methionine and cysteine biosynthesis repressor protein (McbR) was knocked-out[17]. WsMetZ, on the other hand, could not utilize truncated WsHcyS-COSH as a sulfur source for O-acetyl-L-homoserine to form WsHcyS-homocysteine adduct (**Figure 4.18,** observed mass: 7712.63 Da, expected mass: 7805.96 Da).

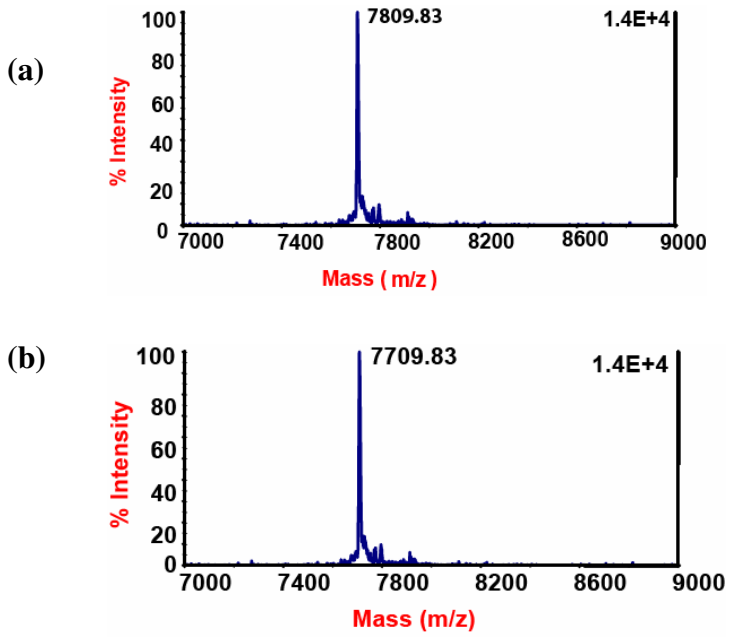


Figure 4.16: MALDI-MS analysis of WsHcyS-COSH in the presence of WsMetY and **(a)** O-acetyl-L-homoserine, 2 min incubation time **(b)** O-acetyl-L-serine, 1 h incubation time

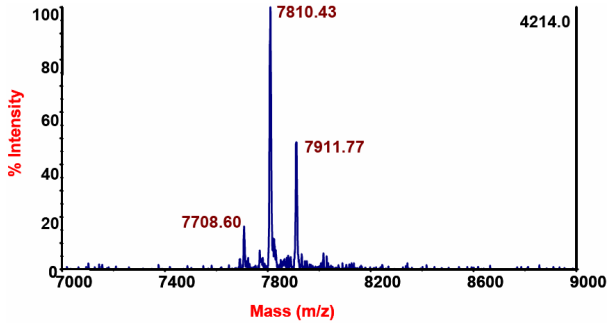


Figure 4.17: WsHcyS-COSH in the presence of WsMetY and O-acetyl-L-homoserine, 1 h incubation time.

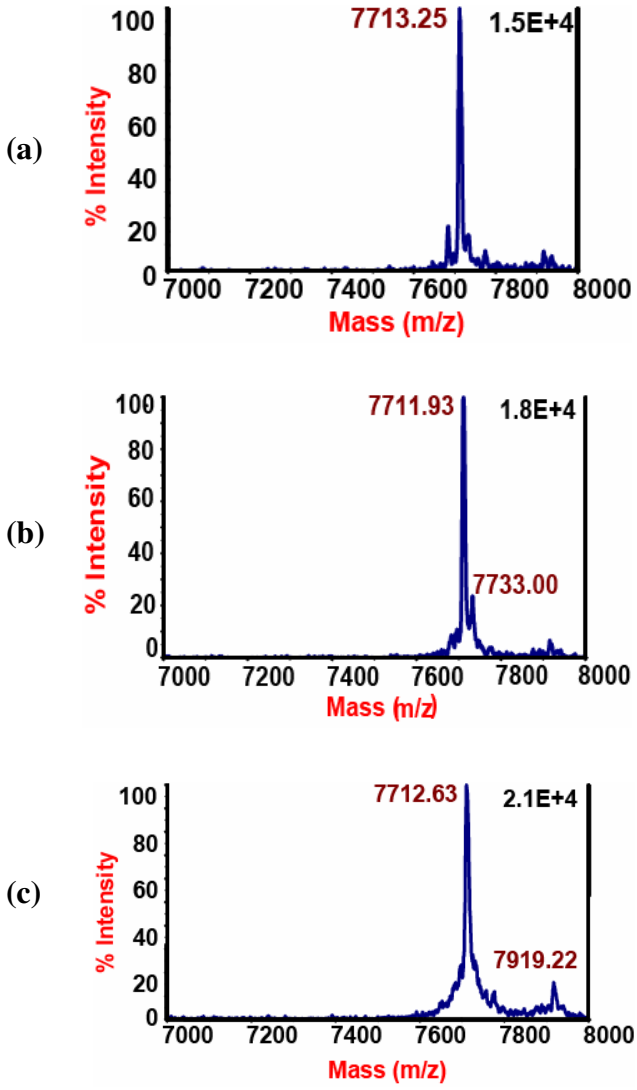


***Figure 4.18*:** MALDI-MS analysis of (a) WsHcyS-COSH in the presence of WsMetZ and O-acetyl-L-serine, (b) WsHcyS-COSH in the presence of WsMetZ and O-acetyl-L-homoserine (c) WsHcyS-COSH in the presence of WsMetZ and O-succinyl-L-homoserine. No 7805.96 Da adduct corresponding to WsHcyS-Homocysteine adduct is formed in any of these cases.

WsMetY can also use S^{2-} to convert O-acetyl-L-homoserine to homocysteine (**Figure 4.19**).

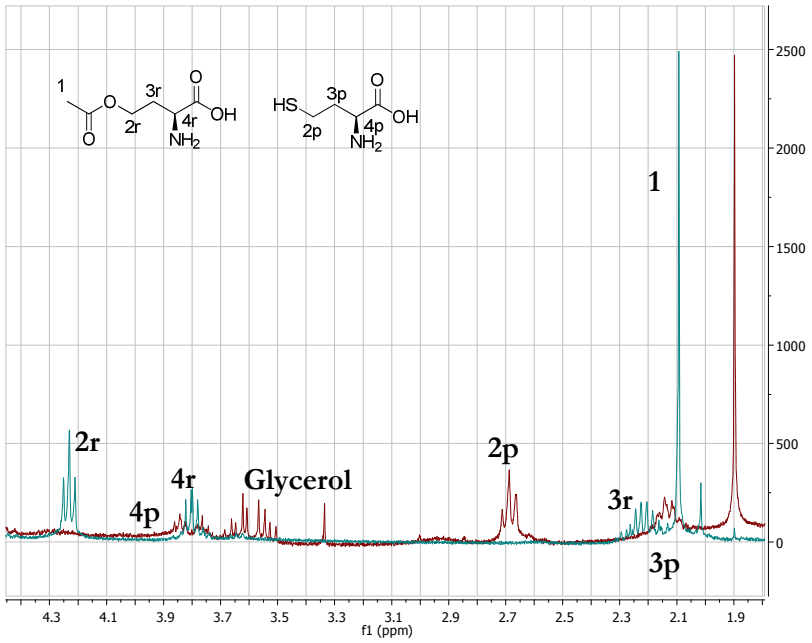


Figure 4.19: ^{1}H-NMR of L-homocysteine made by WsMetY (red) using O-acetyl-L-homoserine and Na_2S as the sulfur source, 1 h incubation time. No homocysteine is formed in the absence of the enzyme (blue).

4.3.6 Release of homocysteine from WsHcyS-DL-homocysteine adduct

The release of homocysteine from WsHcyS-DL-homocysteine adduct might require a protease and WsHcyD was again tested as the possible candidate to free the terminal amino-acid. WsHcyS-DL-homocysteine adduct was again made by cleaving truncated WsHcyS-intein construct with DL-homocysteine. Product release was monitored by changes in the mass of WsHcyS-DL-homocysteine by MALDI-MS. WsHcyD indeed was seen to cleave homocysteine from the protein (**Figure 4.20**, observed mass: 7695.42 Da, expected mass: 7688.82 Da, error: 0.09%).

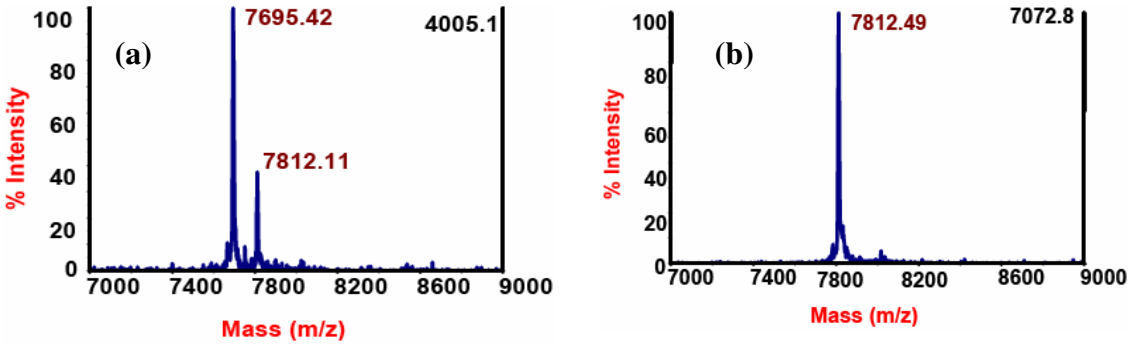


Figure 4.20: MALDI-MS analysis of (a) WsHcyS-homocysteine in the presence of WsHcyD (b) WsHcyS-homocysteine in the absence of WsHcyD.

WsHcyD was also seen to cleave the sulfide off WsHcyS thiocarboxylate (**Figure 4.21**). It seems to be non-specific with respect to the identity of the amino-acid residue at the C-terminus of WsHcyS.

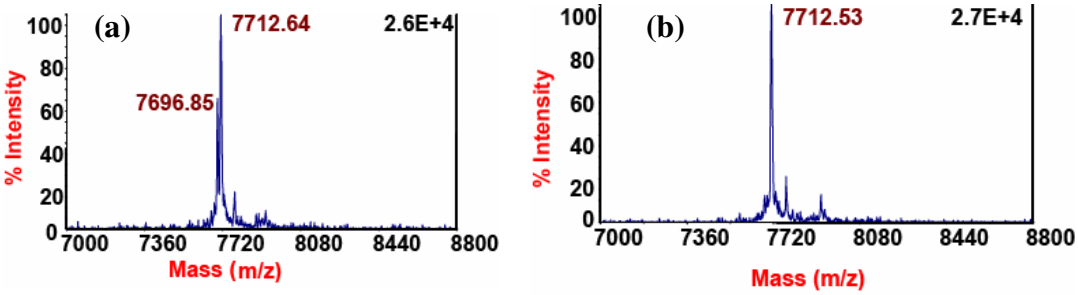


Figure 4.21: MALDI-MS analysis of (a) WsHcyS-COSH in the presence of WsHcyD (b) WsHcyS-COSH in the absence of WsHcyD.

4.3.7 Identification of the released product from WsHcyS-homocysteine adduct made from truncated WsHcyS-COSH

The treatment of truncated WsHcyS-COSH with O-acetyl-L-homoserine and WsMetY results in a 7805 Da adduct, the putative WsHcyS-DL-homocysteine adduct. To conclusively prove that the protein-bound component is homocysteine, the sample was

treated with the metalloprotease, WsHcyD after the reaction to release the C-terminal amino-acid and the released product was derivatized with o-phthalaldehyde to yield a fluorescent compound (**Figure 4.22**)[18-20] which was analyzed by HPLC (**Figure 4.23**). Co-elution with the standard homocysteine:o-phthalaldehyde derivative identified the released amino-acid as homocysteine.

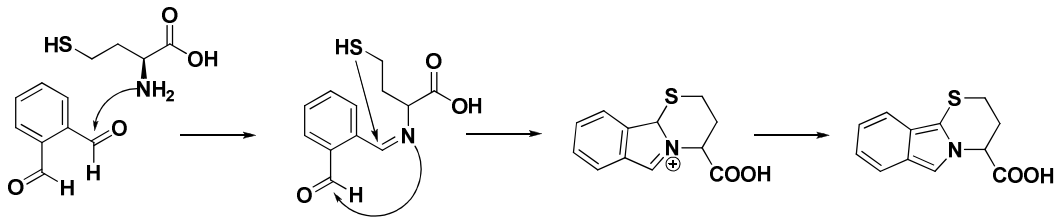


Figure 4.22: o-Phthalaldehyde derivatization of homocysteine released from WsHcyS-COSH reaction with WsMetY.

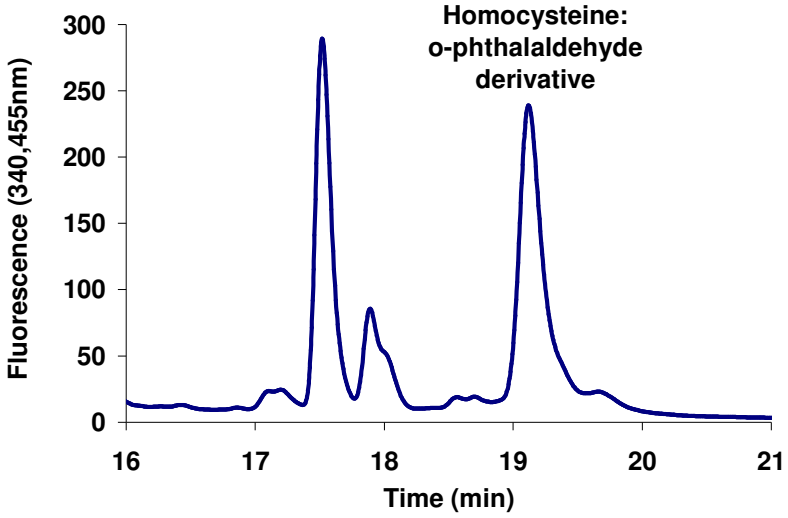


Figure 4.23: o-Phthalaldehyde derivatization of homocysteine released from WsHcyS-COSH reaction with WsMetY.

4.3.8 Conversion of homocysteine to methionine

W.succinogenes has four genes in the genome that could possibly biosynthesize methionine from homocysteine, a non-proteogenic amino-acid - *250* (5-methyltetrahydropteroyltriglutamate-homocysteine methyltransferases, MetE), *942* (SAM-dependent methyltransferase present in the gene neighbourhood), *1672* (5-methyltetrahydropteroyltriglutamate-homocysteine methyltransferases, MetE) and *1141* (B12-dependent metH).

SAM-dependent methyltransferase, Ws942, was analyzed first for methionine biosynthetic activity. Ws942 was unable to convert DL-homocysteine to methionine using S-adenosylmethionine as the methyl donor (**Figure 4.24**). Methionine production was assayed by derivatizing with o-phthalaldehyde and mercaptopropionic acid in borate buffer, pH 9.3.

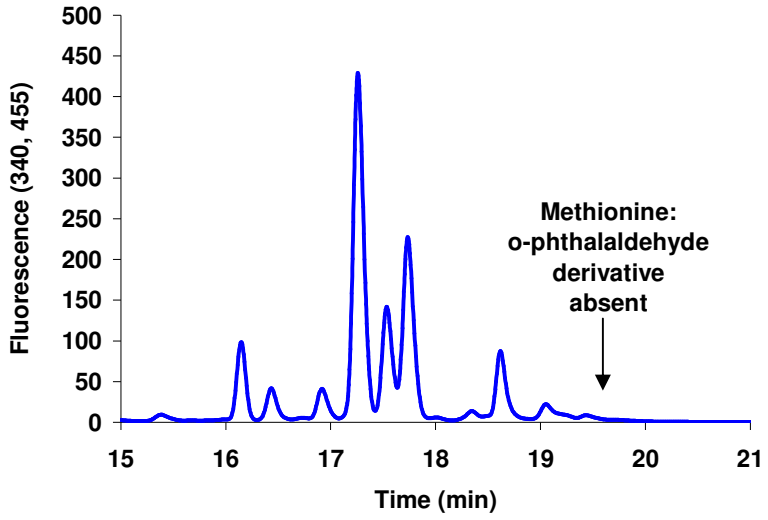


Figure 4.24: Methionine biosynthetic activity of methyltransferases, Ws942. Standard methionine:o-phthalaldehyde derivative elutes at 19.5 min in reverse-phase C18 column.

Activities of the two WsMetE proteins (250 and 1672) were assayed next. Ws250, but not Ws1672, was found to be biosynthesize methionine (**Figure 4.25**). Sequence analysis of the two *W.succinogenes* MetE (250 and 1672) with the *E.coli* and *T.maritima* MetE enzymes suggests that while 1672 has the residues necessary to bind to Zn^{2+} (red boxes, **Figure 4.26**), the protein lacks the N-terminal domain, which aids in the 5-methyltetrahydrofolate binding, present in other MetE proteins. A similar protein sequence has previously been observed in *M.thermoautotrophicum* MetE and the enzyme was not able to utilize CH_3-H_4PteGlu or its analog N^5-methyltetrahydromethanopterin for homocysteine methylation[21,22]. Even though activity was obtained by using methylcobalamin and metylcobinamide as the methyl donors, the physiological methyl donor has still not been ascertained [21].

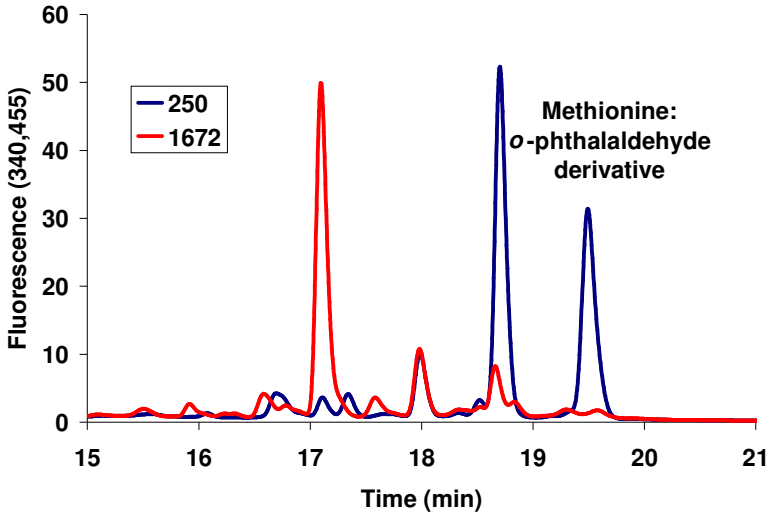


Figure 4.25: Methionine biosynthetic activity of the two WsMetE – Ws250 and Ws1672. The peak at 19.5 min co-migrated with the standard methionine:o-phthalaldehyde derivative.

The activity of cobalamin-dependent methionine synthase, WsMetH (1141) was also tested but did not show methionine production conclusively. It has, however, been

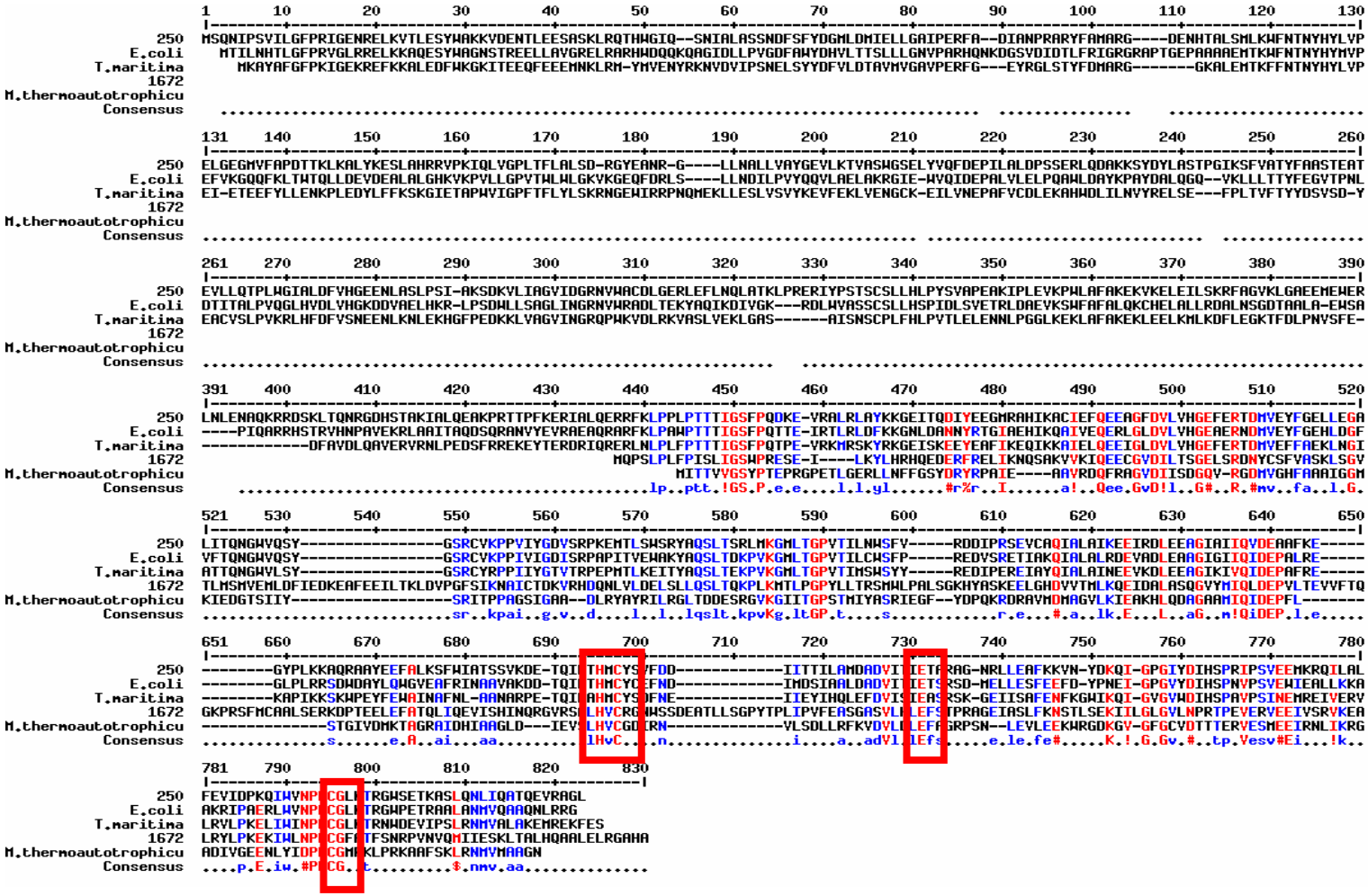


***Figure 4.26**:* Sequence alignment of the two WsMetE (250 and 1672) from *W.succinogenes* with *T.maritima*, *M.thermoautotrophicum* and *E.coli* MetE using Multalign. The Zn-binding sites are highlighted in the red boxes (His, Cys, Glu and Cys).

proved by the activity of WsMetE (250) that *W.succinogenes* has the capability to convert homocysteine formed in the new pathway into methionine.

4.4 Conclusions

The occurrence of the thiocarboxylate-forming protein along with other biosynthetic enzymes, in the gene neighborhood of sulfate assimilation proteins answers an interesting question regarding the sulfur source of these proteins. It has been shown conclusively in this work that sulfate or sulfite can also act as alternate source of sulfur for this important protein modification. A new biosynthetic route for L-methionine that utilizes a protein thiocarboxylate to make the precursor, L-homocysteine, has been discovered in *Wolinella succinogenes*. Previously known route for cystathionine-independent production of L-homocysteine involved direct sulfhydrylation of O-acetyl/O-succinyl-L-homoserine by the corresponding sulfhydrylases. In this work, it has been shown that the sulfide is first loaded onto a small protein, forming a C-terminal thiocarboxylate, before being utilized for making homocysteine. The advantage of loading a sulfur atom onto a protein as a thiocarboxylate is to prevent its possible oxidation under conditions of oxidative stress. This was observed in the case of cysteine biosynthesis in *M.tuberculosis* where the cysO, the thiocarboxylate-forming protein, was expressed at higher levels when the microbe was subjected to oxidative stress [23]. The identification of a new protein thiocarboxylate-dependent methionine biosynthetic pathway further lays stress on the importance of these interesting sulfur transfer proteins and the necessity for systematic approaches to further discover newer proteins and pathways.

REFERENCES

1. Dorrestein, P. C., Zhai, H., McLafferty, F. W. & Begley, T. P. The biosynthesis of the thiazole phosphate moiety of thiamin: the sulfur transfer mediated by the sulfur carrier protein ThiS. *Chem Biol* **11**, 1373-81 (2004).
2. Leimkuhler, S., Wuebbens, M. M. & Rajagopalan, K. V. Characterization of Escherichia coli MoeB and its involvement in the activation of molybdopterin synthase for the biosynthesis of the molybdenum cofactor. *J Biol Chem* **276**, 34695-701 (2001).
3. Burns, K. E. et al. Reconstitution of a new cysteine biosynthetic pathway in Mycobacterium tuberculosis. *J Am Chem Soc* **127**, 11602-3 (2005).
4. Godert, A. M., Jin, M., McLafferty, F. W. & Begley, T. P. Biosynthesis of the thioquinolobactin siderophore: an interesting variation on sulfur transfer. *J Bacteriol* **189**, 2941-4 (2007).
5. Shigi, N., Sakaguchi, Y., Asai, S., Suzuki, T. & Watanabe, K. Common thiolation mechanism in the biosynthesis of tRNA thiouridine and sulphur-containing cofactors. *Embo J* **27**, 3267-78 (2008).
6. Leidel, S. et al. Ubiquitin-related modifier Urm1 acts as a sulphur carrier in thiolation of eukaryotic transfer RNA. *Nature* **458**, 228-32 (2009).
7. Begley, T. P., Xi, J., Kinsland, C., Taylor, S. & McLafferty, F. The enzymology of sulfur activation during thiamin and biotin biosynthesis. *Curr Opin Chem Biol* **3**, 623-9 (1999).
8. Leimkuhler, S. & Rajagopalan, K. V. A sulfurtransferase is required in the transfer of cysteine sulfur in the in vitro synthesis of molybdopterin from precursor Z in Escherichia coli. *J Biol Chem* **276**, 22024-31 (2001).

9. Belfaiza, J., Martel, A., Margarita, D. & Saint Girons, I. Direct sulfhydrylation for methionine biosynthesis in Leptospira meyeri. *J Bacteriol* **180**, 250-5 (1998).

10. Bradford, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**, 248-54 (1976).

11. Thorneley, R. N. A convenient electrochemical preparation of reduced methyl viologen and a kinetic study of the reaction with oxygen using an anaerobic stopped-flow apparatus. *Biochim Biophys Acta* **333**, 487-96 (1974).

12. Gonzalez, J. C., Peariso, K., Penner-Hahn, J. E. & Matthews, R. G. Cobalamin-independent methionine synthase from Escherichia coli: a zinc metalloenzyme. *Biochemistry* **35**, 12228-34 (1996).

13. Koguchi, O. & Tamura, G. Ferredoxin-sulfite Reductase from a Cyanobacterium, Spirulina platensis. *Agric. Biol. Chem.* **52**, 373-380 (1988).

14. Shangguan, N., Katukojvala, S., Greenberg, R. & Williams, L. J. The reaction of thio acids with azides: a new mechanism and new synthetic applications. *J Am Chem Soc* **125**, 7754-5 (2003).

15. Kolakowski, R. V., Shangguan, N., Sauers, R. R. & Williams, L. J. Mechanism of thio acid/azide amidation. *J Am Chem Soc* **128**, 5695-702 (2006).

16. O'Leary, S. E., Jurgenson, C. T., Ealick, S. E. & Begley, T. P. O-phospho-L-serine and the thiocarboxylated sulfur carrier protein CysO-COSH are substrates for CysM, a cysteine synthase from Mycobacterium tuberculosis. *Biochemistry* **47**, 11606-15 (2008).

17. Kromer, J. O., Heinzle, E., Schroder, H. & Wittmann, C. Accumulation of homolanthionine and activation of a novel pathway for isoleucine biosynthesis

in Corynebacterium glutamicum McbR deletion strains. *J Bacteriol* **188**, 609-18 (2006).

18. Okumura, H. S., Philmus, B., Portmann, C. & Hemscheidt, T. K. Homotyrosine-containing cyanopeptolins 880 and 960 and anabaenopeptins 908 and 915 from Planktothrix agardhii CYA 126/8. *J Nat Prod* **72**, 172-6 (2009).
19. Puri, R. N. & Roskoski, R., Jr. Reaction of low molecular weight aminothiols with o-phthalaldehyde. *Anal Biochem* **173**, 26-32 (1988).
20. Benson, J. R. & Hare, P. E. O-phthalaldehyde: fluorogenic detection of primary amines in the picomole range. Comparison with fluorescamine and ninhydrin. *Proc Natl Acad Sci U S A* **72**, 619-22 (1975).
21. Schroder, I. & Thauer, R. K. Methylcobalamin:homocysteine methyltransferase from Methanobacterium thermoautotrophicum. Identification as the metE gene product. *Eur J Biochem* **263**, 789-96 (1999).
22. Ferrer, J. L., Ravanel, S., Robert, M. & Dumas, R. Crystal structures of cobalamin-independent methionine synthase complexed with zinc, homocysteine, and methyltetrahydrofolate. *J Biol Chem* **279**, 44235-8 (2004).
23. Manganelli, R. et al. Role of the extracytoplasmic-function sigma factor sigma(H) in Mycobacterium tuberculosis global gene expression. *Mol Microbiol* **45**, 365-74 (2002).

CHAPTER 5

Cross-linking thiazole synthase (ThiG) with ThiS, the thiocarboxylate-forming protein involved in thiamin biosynthesis

5.1 Introduction to thiamin biosynthesis

Thiamin is an essential vitamin that humans need to uptake through their diet. Micro-organisms, on the other hand, are capable of meeting their thiamin requirements through de novo biosynthesis. The two units constituting the thiamin – hydroxymethylpyrimidine (HMP) and thiazole are made separately and joined together to form thiamin. HMP is made by a single protein ThiC in prokaryotes[1], plants and algae, and THI5p in fungi and eukaryotes[2].

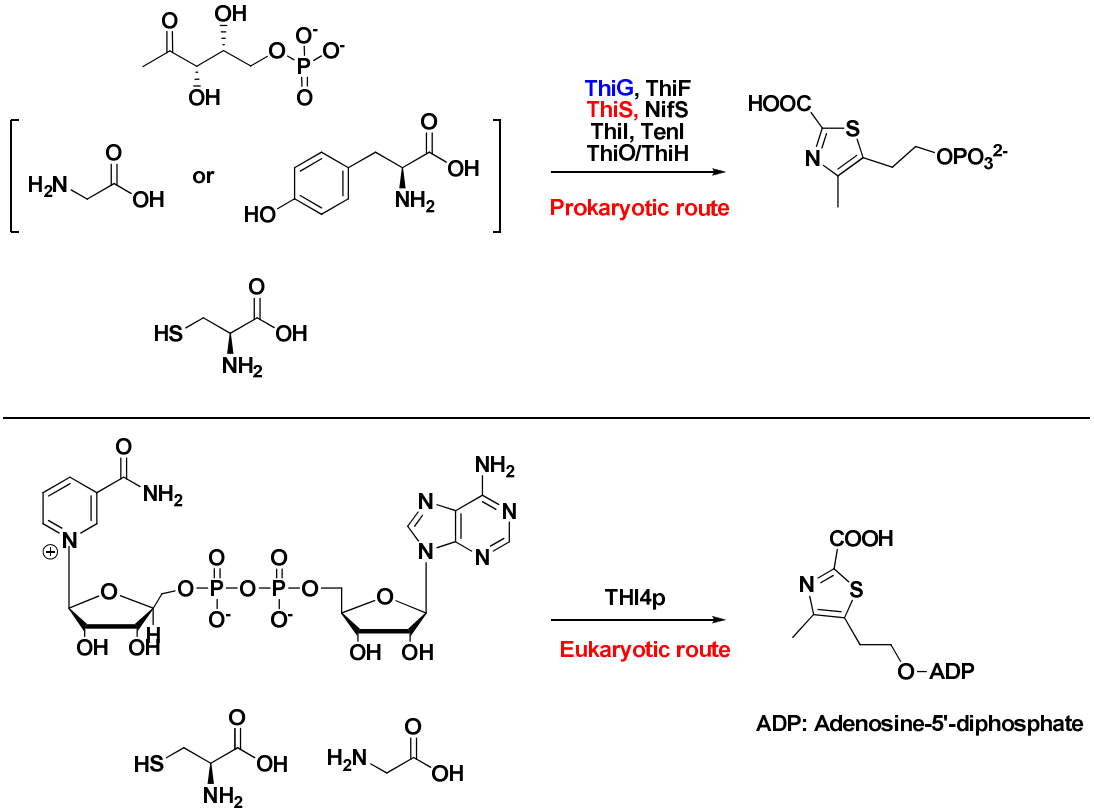


***Figure 5.1*:** Biosynthesis of the thiazole moiety of thiamin in prokaryotes and eukaryotes.

The mode of biosynthesis of thiazole ring is different in prokaryotes and eukaryotes. **Figure 5.1** shows the two routes of biosynthesizing the thiazole moiety of thiamin in these two classes of organisms. While eukaryotes utilize a single enzyme, THI4p, to perform the complicated procedure for assembling the thiazole ring[3], prokaryotes make use of seven enzymes, including ThiS thiocarboxylate, for this purpose [4].

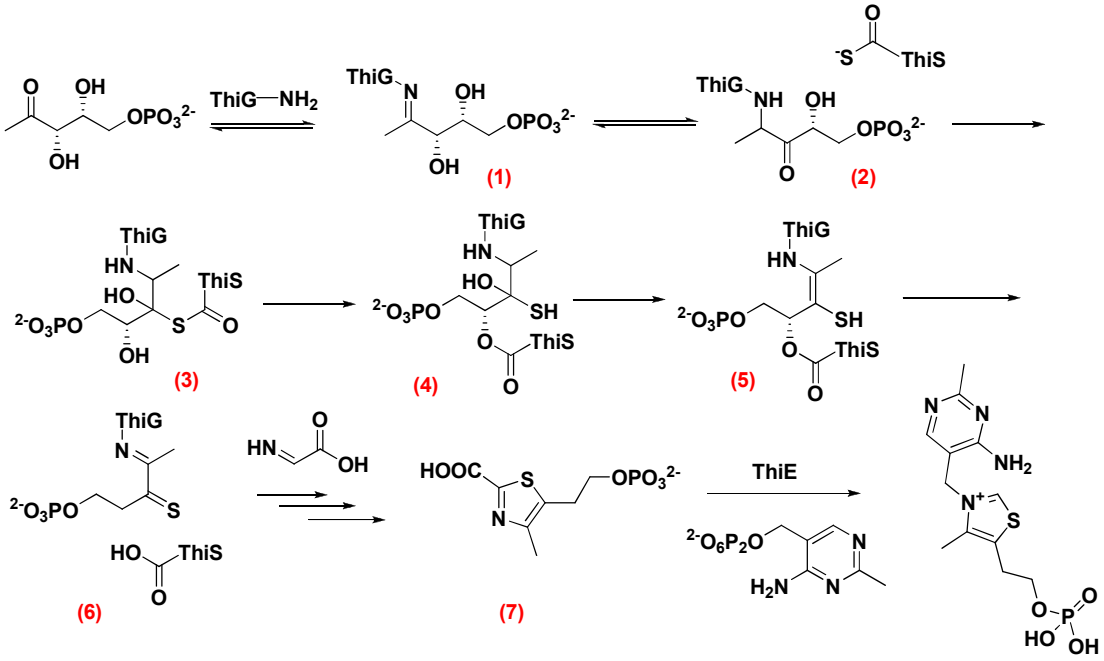


Figure 5.2: Mechanism of sulfur transfer from ThiS thiocarboxylate to the small five-carbon sugar, 1-deoxy-D-xylulose-5-phosphate, bound in the active-site of thiazole synthase, ThiG.

The mechanism of sulfur transfer from ThiS thiocarboxylate is shown in **Figure 5.2.** ThiG catalyzes an amadori-type rearrangement of DXP using simple acid-base chemistry. DXP binds to the active-site lysine via an imine bond **(1)**. A tautomerization **(2)** followed by nucleophilic addition of the ThiSCOSH onto the C3 ketone yields **(3)**, which undergoes S/O acyl shift followed by water loss to give thioketone **(6)**. This is then taken over a series of steps to yield **(7)**, one of the

precursor to thiamin. All the steps in the biosynthesis have been characterized using techniques like ^{1}H-NMR, radioactive assays and isotope studies Previous crystallographic studies on these proteins (PDB ID: 1TYG, 3.15 Å resolution) showed the C-terminus of *Bacillus subtilis* ThiS (BsThiS) interacting with the active-site of BsThiG [5] (**Figure 5.3**). However, the low-resolution crystal structure had only a phosphate ion bound in the active-site.

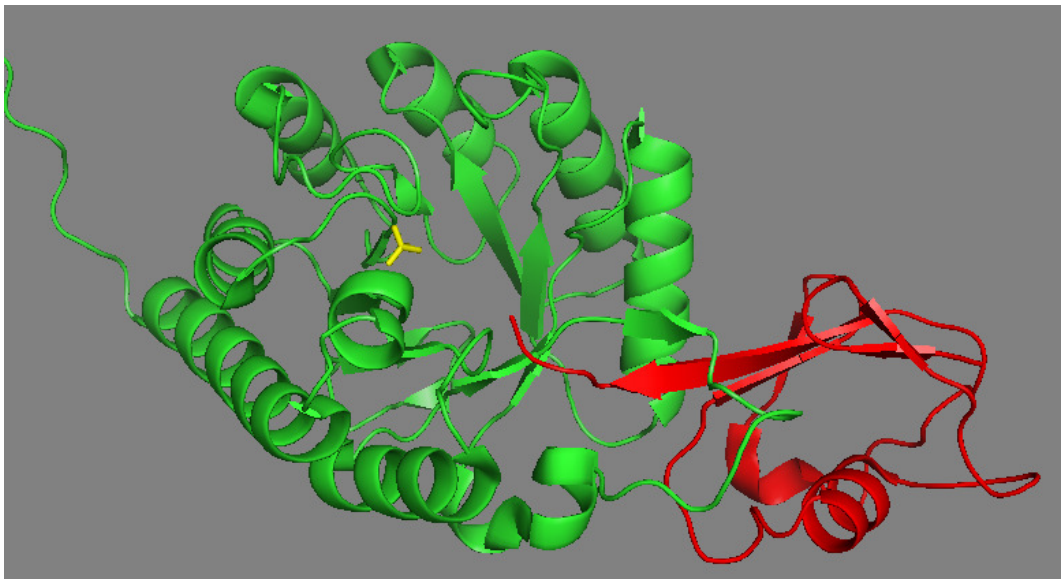


Figure 5.3: X-ray crystal structure of BsThiS (red) and BsThiG (green). Note the C-terminus of BsThiS reaching into the active-site of BsThiG that has a phosphate (yellow) bound.

Substrate analogs (**Figure 5.4**) that can be used for co-crystallization or soaking studies with these two proteins were made with the goal of getting some insight into the interactions of the substrate or intermediate with ThiG active-site residues and ThiS C-terminus. Analogs **8** and **9** were synthesized by the schemes shown below (**Figure 5.5** and **Figure 5.6**). Analog **10** was synthesized by cleaving the ThiS-intein

construct with cysteine followed by labeling the cysteine with iodoacetamide (**Figure 5.7**).

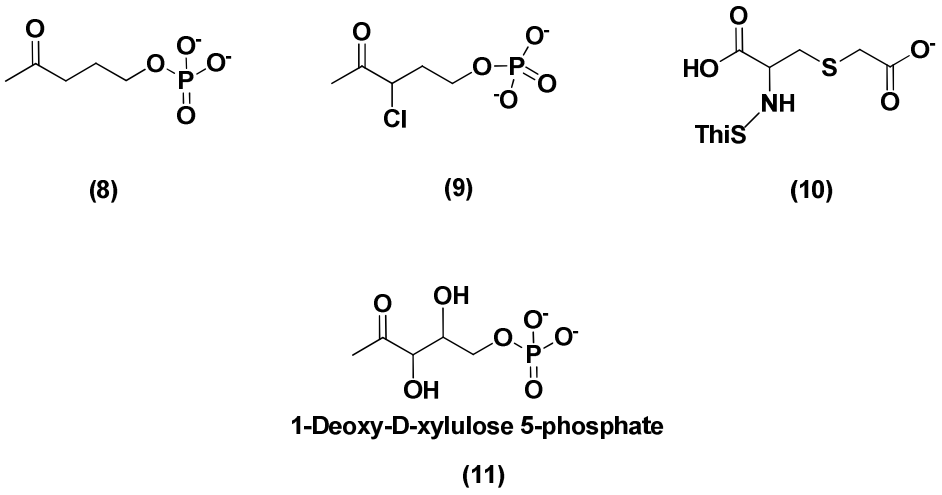


Figure 5.4: DXP analogs (8, 9 and 10) synthesized for crystallization studies.

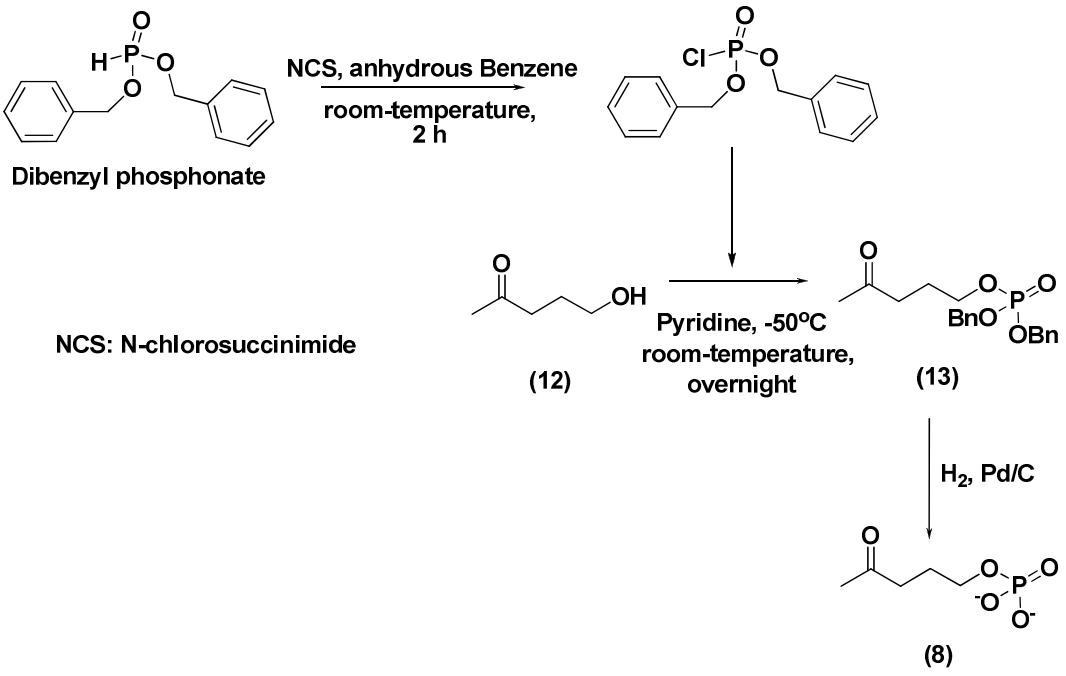


Figure 5.5: Synthesis of analog **8**

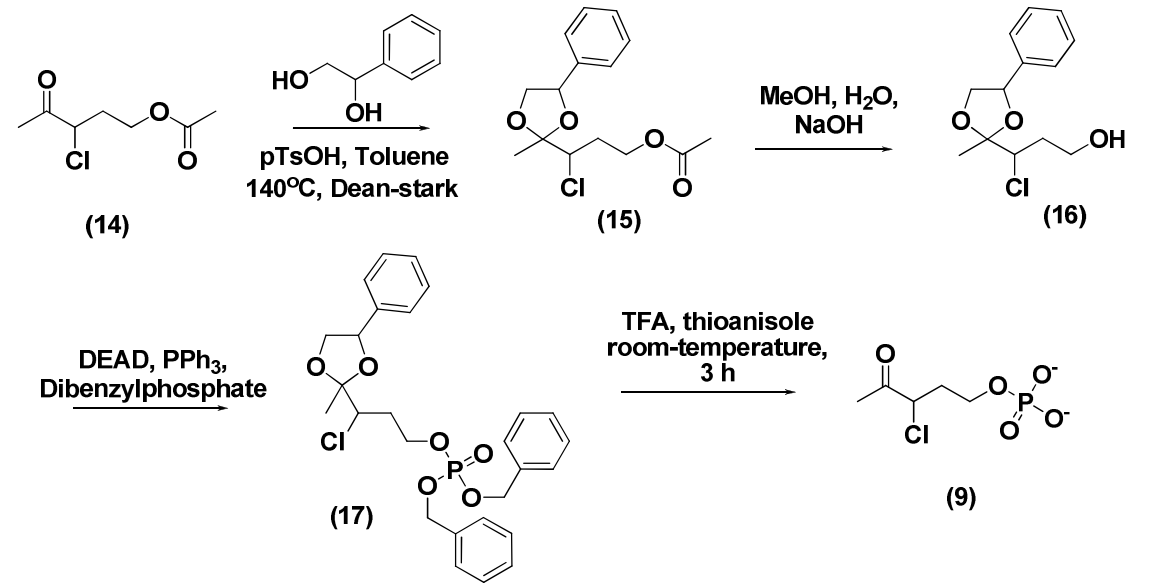


Figure 5.6: Synthesis of analog **9**

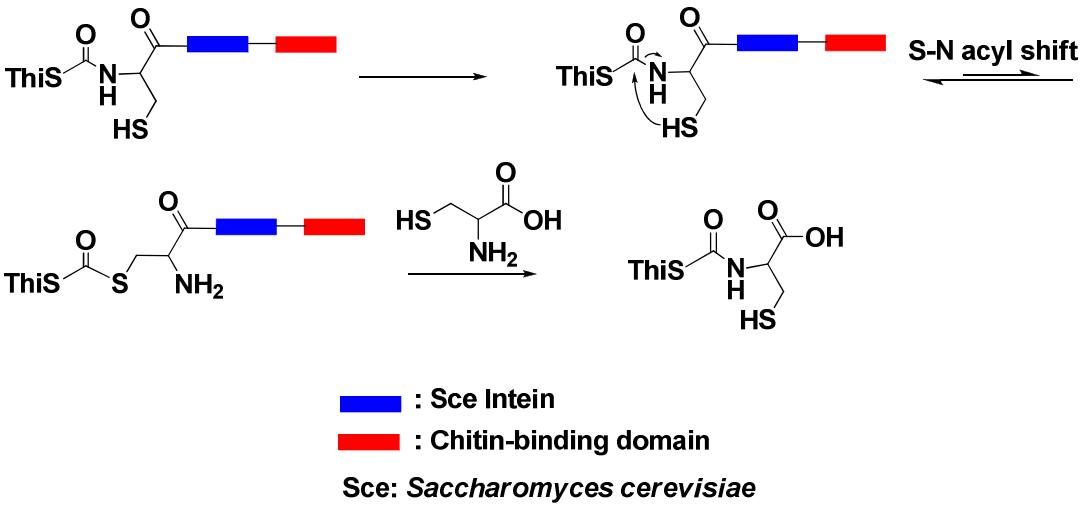


Figure 5.7: Synthesis of analog **10**

5.2 Experimental section

All chemicals except for compound 14 were acquired from Sigma-aldrich (St.Louis, MO) or Fisher scientific (Fairlawn, NJ). ESI-MS analysis of the carbamidomethylated BsThiS-cysteine adduct was performed in positive mode using an Esquire-LC_00146 instrument (Bruker, Billerica, MA). EMD biosciences (Gibbstown, NJ) supplied Luria-Bertani. Chitin beads was obtained in New England Biolabs (Ipswich, MA). IPTG was procured from Lab Scientific (Livingston, NJ). Sonication was done using Misonix sonicator 3000 (Misonix Inc., Farmingdale, NY).

5.2.1 Synthetic route for analog 8

5.2.1.1 Synthesis of compound 13

3.28 g of dibenzylphosphite and 1.32 g of N-chlorosuccinimide were mixed together in anhydrous benzene and stirred at room-temperature for 2 h under argon. The solution was filtered directly into a solution of anhydrous pyridine (10 mL) and 3-acetyl-1-propanol (204 µL) stirring at -50^{o}C under argon. After the addition was over, the cold bath was removed and stirring was continued at room-temperature overnight. The compound was purified by column chromatography (60:40:: hexane:EtOAc). Yield: 425 mg (59%). ^{1}H NMR (300 MHz, $CDCl_3$) δ 7.33 (d, 10H), 5.13 – 4.95 (m, 4H), 3.98 (dd, 2H), 2.76 (s, 6H), 2.45 (t, 2H), 2.08 (s, 3H), 1.84 (dd, 2H).

5.2.1.2 Synthesis of compound 8

425 mg of compound 6 was dissolved in 10 mL anhydrous THF. The solution was purged with argon before adding 85 mg of Pd (palladium) / 10 wt % activated carbon. Hydrogen was filled in a balloon and bubbled through the solution for 19 h. The solution was filtered to remove the metal and the solvent was removed by vacuum. The solid obtained was dissolved in water; pH was adjusted to approximately 7.0 and

washed with dichloromethane/ethylacetate. The aqueous layer was lyophilized. Yield: 230 mg (95 %). ^{1}H NMR (300 MHz, D_2O) 3.75 (q, 2H), 2.65 (t, 2H), 2.21 (s, 3H), 1.88 – 1.72 (m, 2H). ESI-MS (negative mode): 181 g/mol.

5.2.2 Synthetic route for analog 9

5.2.2.1 Synthesis of compound 15

205 mg of compound 14, 608 mg of 1-phenyl-1,2-ethanediol, 64 mg of p-toluenesulfonic acid were dissolved in 10 mL toluene and refluxed for 2 h 15 min at 140^{o}C using Dean-Stark apparatus. Toluene was removed by vacuum and the compound was purified by silica gel chromatography (90:10::hexane:EtOAc). Yield: 155 mg (46.4 %). ^{1}H NMR (300 MHz, $CDCl_3$) δ 7.45 – 7.28 (m, 5H), 5.20 – 5.08 (m, 1H), 4.47 – 4.02 (m, 4H), 3.82 – 3.69 (m, 1H), 2.52 – 2.26 (m, 1H), 2.18 – 1.88 (m, 4H), 1.69 – 1.53 (m, 3H).

5.2.2.2 Synthesis of compound 16

140 mg of compound 15 was dissolved in 10 mL MeOH, 1 mL water and 40 mg of NaOH (2 eq.) and stirred at room-temperature for 30 min. 20 mL double distilled H_2O was added to the solution and the compound was extracted into chloroform. Chloroform layer was dried with $MgSO_4$, filtered and removed by vacuum. Yield: Quantitative conversion. ^{1}H NMR (300 MHz, $CDCl_3$) δ 7.47 – 7.28 (m, 5H), 5.23 – 5.07 (m, 1H), 4.53 – 3.56 (m, 5H), 2.41 – 2.17 (m, 1H), 2.05 – 1.82 (m, 1H), 1.65 – 1.56 (m, 3H).

5.2.2.3 Synthesis of compound 17

131 mg of compound 16, 241 µL of diethyl azodicarboxylate (DEAD), 425 mg of dibenzylphosphate and 401 mg of PPh_3 were dissolved in 10 mL of anhydrous THF at

0°C. DEAD was added the last. The reaction was allowed to warm to room-temperature overnight (12 h). The solvent was removed by vacuum and the compound purified by column chromatography (75:25::hexane:EtOAc). The final product had DEAD impurities. Yield: Quantitative conversion. ^{1}H NMR (300 MHz, $CDCl_3$) δ 7.46 – 7.20 (m, 15H), 5.19 – 4.96 (m, 5H), 4.39 – 4.00 (m, 8H, has DEAD peaks), 3.79 – 3.59 (m, 1H), 2.52 – 2.27 (m, 1H), 1.97 – 1.77 (m, 1H), 1.63 – 1.46 (m, 3H).

5.2.2.4 Synthesis of compound 9

Approximately 100 mg of compound 17 was dissolved in 5:1::TFA:thioanisole and the reaction was stirred at room-temperature for 3 h. The solvents were removed by vacuum. The oily residue was dissolved in water, pH was made to 7.5, washed with chloroform and the aqueous layer was lyophilized. The NMR spectrum had some aromatic impurities. The compound was washed with absolute ethanol and the solid filtered, re-dissolved in water and lyophilized again. Yield: 9.2 mg. ^{1}H NMR (300 MHz, D_2O) δ 4.04 – 3.82 (m, 3H), 2.46 – 2.21 (m, 4H), 2.11 – 1.94 (m, 1H).

5.2.3 Preparation of analog 10

20 mL of chitin bead slurry was loaded onto a column and the beads were allowed to settle. The column was then equilibrated with column buffer (20 mM Tris, 500 mM NaCl, 1mM EDTA, pH = 7.5). Meanwhile, *E.coli* ThiS-intein-chitin binding domain (pCLK413 in pET-22b) was transformed into *E.coli* Tuner(DE3). The cultures were grown at 37°C in Luria-Bertani till an OD_{600} of 0.4 when the temperature was reduced to 15°C. Induction was done at an OD_{600} of 0.5 with 0.4 mM IPTG and further growth at 15°C was carried out for 15 h. The cultures were harvested and lysed in lysis buffer (column buffer + 0.1% Triton X) by sonication. The crude lysate was centrifuged and the clarified lysate was loaded onto the column at a rate of 0.5 mL/min at 4°C. The

column was then washed with 300 mL of column buffer at a rate of 1.5 mL/min at 4°C. 30 mL of column buffer containing 50 mM DL-cysteine was then loaded onto the column and drained at a rate of 2 mL/min till a small amount of buffer remained at the top of the bed. The column was then capped at both ends and incubated at 4°C for 48 h. After 48 h, the column was washed with column buffer and the fractions collected were concentrated and buffer-exchanged into 50 mM Tris, 150 mM NaCl, pH = 7.8.

5.2.3.1 Labeling the C-terminal cysteine with iodoacetamide/iodoacetic acid: 160 µL of 157 µM ThiS-Cys was treated with 5 µL of 250 mM TCEP and different concentrations of iodoacetamide as shown in **Table 5.1**. The samples were incubated at room-temperature for 2 h. After incubation, the samples were desalted and concentrated to 20 µL of 50 mM ammonium acetate (not pH adjusted) and diluted into 100 µL of 50:45:5::ACN:H_2O:HCOOH. The sample was analyzed by positive mode ESI-MS.

Table 5.1: Iodoacetamide concentrations in each sample

Sample	Volume of 250 mM iodoacetamide (µL)	Vol. of buffer added (µL)	Final concentration (mM)
1	8	42	9
2	16	34	19
3	24	26	28
4	32	18	37
5	40	10	47

5.3 Results and discussions

The substrate analogs synthesized resemble the natural substrate, DXP, as shown in **Figure 5.8**. The five-carbon backbone is maintained in substrate analog **8**. The phosphate might bind to the phosphate-binding cavity of the thiazole synthase while

the ketone can anchor the analog to the active-site lysine. This is also the case for the substrate analog **9**. An additional chloride on C3 of the backbone is a replacement for the C3-hydroxy of the natural substrate, DXP. Attempts were made to synthesize **9** and **9a** from **8** by the transformations shown in **Figure 5.9** but were not successful. The synthetic route shown in **Figure 5.6** worked efficiently.

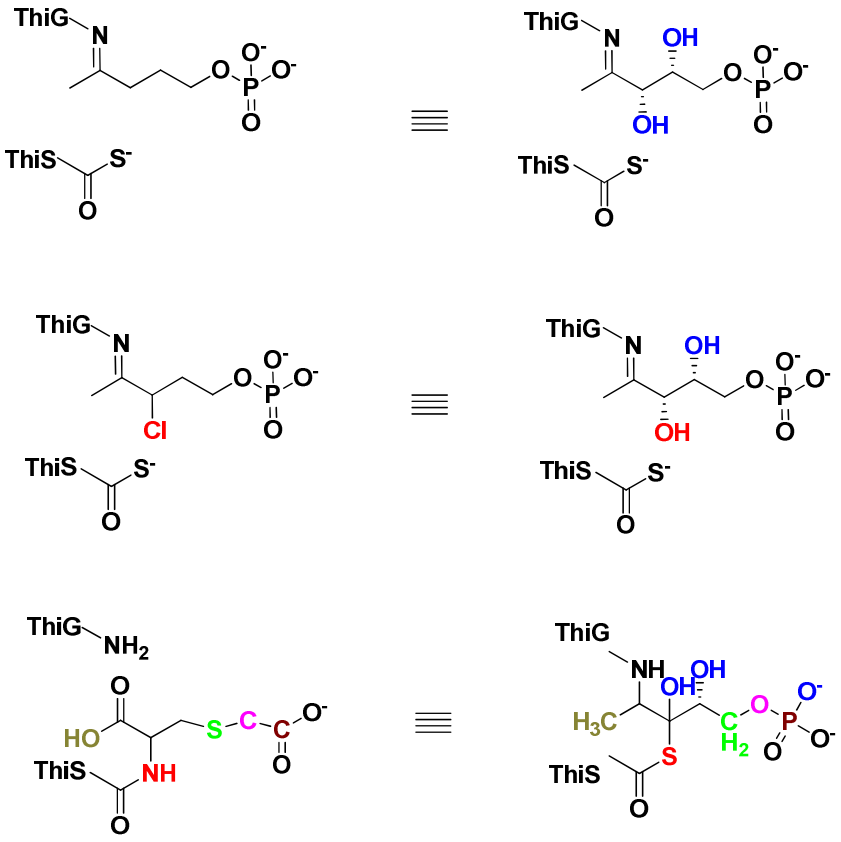


Figure 5.8: Substrate analogs and their structural similarity to the actual substrate, DXP. Atoms shown in blue on DXP are missing from the analogs and the other colored atoms on the analogs represent the corresponding colored atoms on DXP.

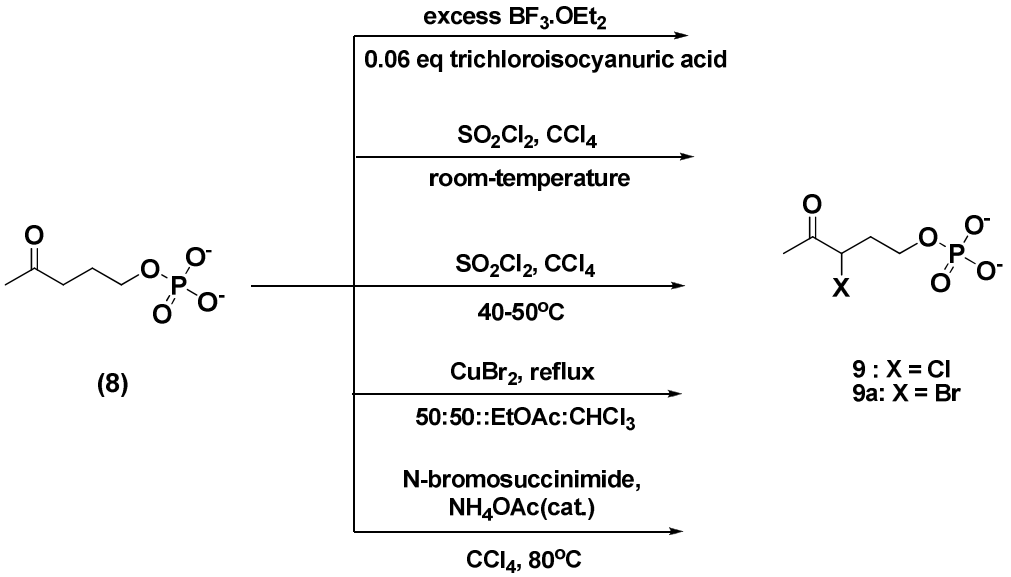


Figure 5.9: Attempts to synthesize **9** and **9a** from **8**.

The five-carbon backbone in adduct **10** is provided by the carbamidomethylated C-terminal cysteine (**Figure 5.8**). The cysteine carboxyl is analogous to the ketone and may bind to the active-site lysine of ThiG. To provide the negatively charged group on the cysteine as a substitute for phosphate, the cysteine can be labeled with iodoacetic acid (in this work, the labeling was tried with iodoacetamide). This step yielded two *E.coli* ThiS adducts. *E.coli* ThiS has two cysteines in the primary sequence – the native cysteine of the protein and the C-terminal cysteine that was added during the cleavage step in the intein purification as shown in **Figure 5.10**. The major product has the C-terminal cysteine labeled with iodoacetamide while the minor product has both cysteines carbamidomethylated (**Figure 5.11**).

To solve the problem of dicarbamidomethylation, *Thermus thermophilus* HB8 ThiS (in intein construct) and ThiG have been cloned. These proteins do not have any cysteines in the primary sequence. The same procedure of labeling with

iodoacetamide/iodoacetic acid can be repeated with the *Thermus thermophilus HB8* constructs to yield substrate analog 3.

a) MQILFNDQAMQ**C**AAGQTVHELLEQLDQRQAGAALAINQQIVPREQWAQHIVQDGDQILLFQVIAGG

b) MQILFNDQAMQ**C**AAGQTVHELLEQLDQRQAGAALAINQQIVPREQWAQHIVQDGDQILLFQVIAGG**C**

Figure 5.10: a) pre-intein *E.coli* primary sequence b) post-intein *E.coli* primary sequence after cleaving the intein construct with cysteine. Cysteines are highlighted in **red**.

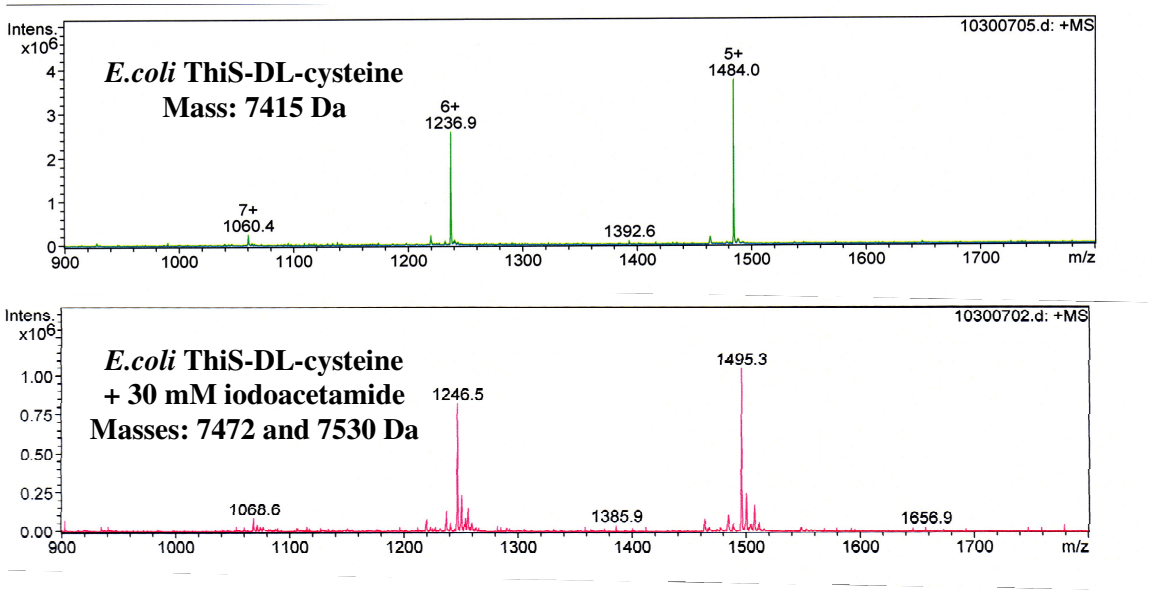


Figure 5.11: ESI-MS of *E.coli* ThiS-DL-cysteine (top panel) treated with 30 mM iodoacetamide (bottom panel).

5.4 Conclusions

Crystallography experiments with these compounds still need to be done. This important sulfur transfer process has not been characterized structurally and these substrate analogs and the natural substrate can be a good starting point.

REFERENCES

1. Chatterjee, A. et al. Reconstitution of ThiC in thiamine pyrimidine biosynthesis expands the radical SAM superfamily. *Nat Chem Biol* **4**, 758-65 (2008).
2. Wightman, R. & Meacock, P. A. The THI5 gene family of Saccharomyces cerevisiae: distribution of homologues among the hemiascomycetes and functional redundancy in the aerobic biosynthesis of thiamin from pyridoxine. *Microbiology* **149**, 1447-60 (2003).
3. Chatterjee, A., Schroeder, F. C., Jurgenson, C. T., Ealick, S. E. & Begley, T. P. Biosynthesis of the thiamin-thiazole in eukaryotes: identification of a thiazole tautomer intermediate. *J Am Chem Soc* **130**, 11394-8 (2008).
4. Begley, T. P., Chatterjee, A., Hanes, J. W., Hazra, A. & Ealick, S. E. Cofactor biosynthesis--still yielding fascinating new biological chemistry. *Curr Opin Chem Biol* **12**, 118-25 (2008).
5. Settembre, E. C. et al. Thiamin biosynthesis in Bacillus subtilis: structure of the thiazole synthase/sulfur carrier protein complex. *Biochemistry* **43**, 11647-57 (2004).

CHAPTER 6

Summary and Outlook

6.1 Summary

Two protein thiocarboxylate-labeling strategies using the thioester-alexa fluor 647 cadaverine nucleophilic displacement chemistry and the sulfonyl azide-thioacid coupling chemistry have been developed. The alexa fluor 647 cadaverine tagging methodology was able to identify an over-expressed thiocarboxylate protein in a cell-free extract. To improve upon the tagging efficiency, the sulfonyl azide tagging chemistry was developed and it was shown to successfully label PdtH in *P.stutzeri* KC cell-free extract and an unknown protein in the *S.coelicolor* system. The necessity for developing a proteomics strategy for identifying new thiocarboxylate-forming proteins was further reiterated by the discovery of a methionine biosynthetic pathway involving these proteins. This raises the possibility of many other pathways still waiting to be identified. The methionine biosynthetic pathway also proved that sulfate, apart from cysteine, could be a source of sulfur for protein thiocarboxylate formation.

6.2 Labeling of protein thiocarboxylates in cell-free extracts

6.2.1 Development of new labeling reagents

Thiocarboxylated proteins could possibly be the gateway to pathways with very interesting sulfur transfer chemistry and efforts to identify more of these special proteins in micro-organisms would be worthwhile. While the labeling method using lissamine rhodamine sulfonyl azide yielded impressive results with the *P.stutzeri KC* and *S.coelicolor* systems and the dye could detect proteins whose copy number is as low as 500 (**Chapter 3**), there still exists some scope for improvement in terms of the sensitivity of the dye. To increase the sensitivity, we propose to make resin linkers

with the sulfonyl azide moiety on it. The advantage of developing a resin as a tag instead of the fluorophore/SDS-PAGE assay method is that there would be no restrictions to the amount of protein sample that can be loaded onto a resin with respect to SDS-PAGE further enhancing the sensitivity of detection.

Figure 6.1 shows one possible resin modification that result in sulfonyl azide functionality at the linker-end and the mode of cleavage of the protein from the resin after tagging. The ester linkage formed would be susceptible to hydrolysis by appropriate nucleophiles.

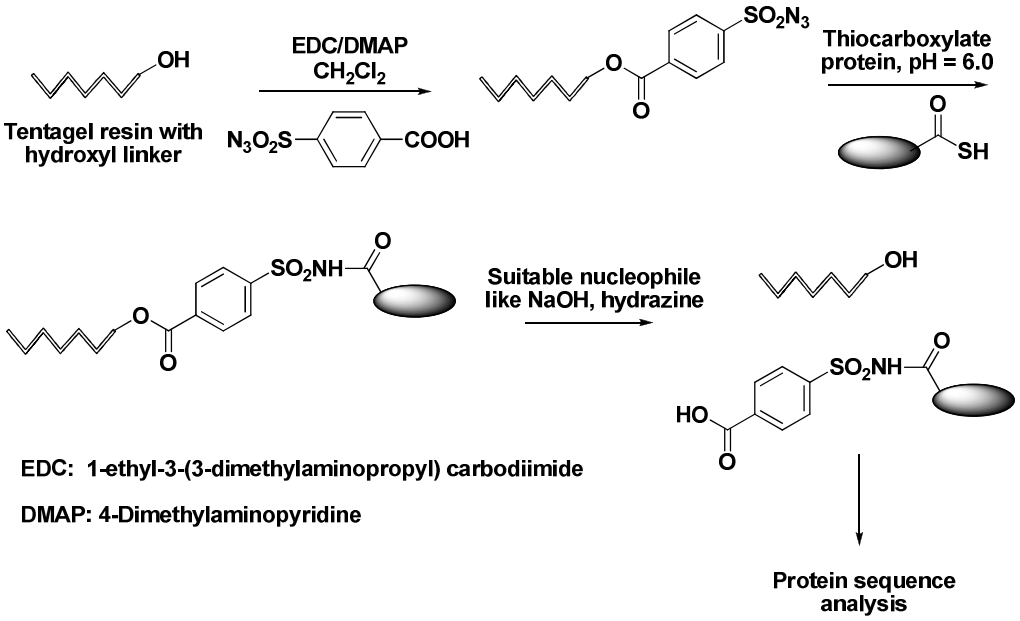


Figure 6.1: Tentagel-hydroxyl resin modification using 4-carboxybenzenesulfonyl azide. The resulting ester formed can be possibly hydrolyzed using a base or a strong nucleophile like hydrazine.

Another possible way to increase the sensitivity of the labeling strategy using a fluorophore is to use one with high fluorescence intensity like alexa fluor 647 cadaverine and form a sulfonamide with 2, 4-dinitrobenzenesulfonyl chloride. This

can then be used to label thiocarboxylate-forming proteins by a mechanism shown in **Figure 6.2.** This strategy has been used previously to synthesize glycosylated peptides[1,2].

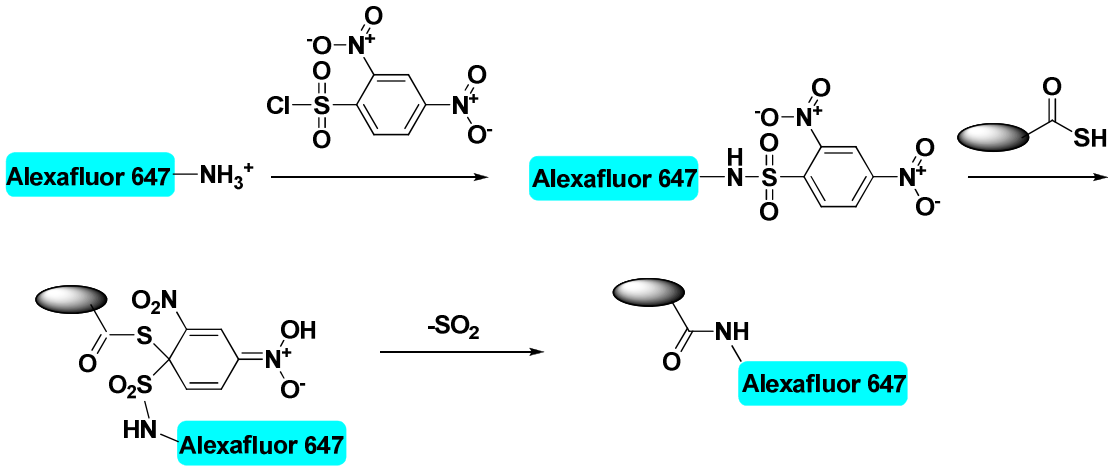


Figure 6.2: Labeling of protein thiocarboxylates using a fluorescent sulfonamide.

6.3 Methionine biosynthesis through protein thiocarboxylate

6.3.1 Mechanism of sulfide transfer from sulfite reductase to WsHcyS-COAMP

The mechanism of sulfur transfer from *W.succinogenes* ferredoxin-sulfite reductase (WsSir) siroheme center to WsHcyS-AMP to form WsHcyS thiocarboxylate would be interesting to study. Sequence alignment of *W.succinogenes* sulfite reductase with *E.coli* sulfite reductase shows that the former has an extra protein overhang with two cysteines (cysteines 570 and 703) that are conserved in all organisms having the extra primary sequence (**Figure 6.3**). Preliminary results suggest that mutating the two cysteines to serines results in complete loss of protein expression/solubility in the case of C703S mutant (analyzed by SDS-PAGE, data not shown) and lower activity in case of C570S mutant. **Figure 6.4** shows the absorption spectrum of the two cysteine mutants.

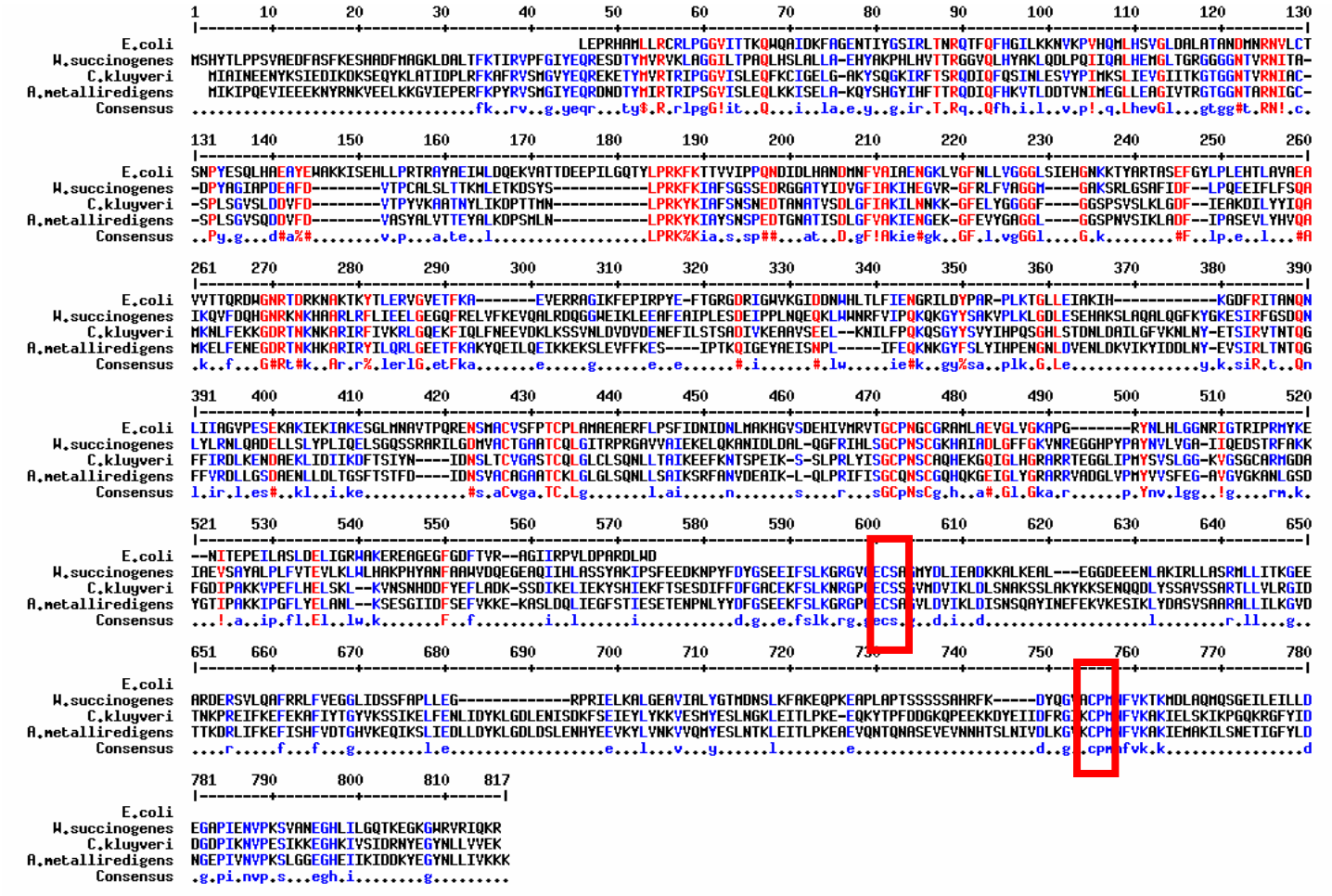


Figure 6.3: Sequence analysis of sulfite reductase from *E.coli, W.succinogenes, C.kluyveri and A.metalliredegens* using Multalign.

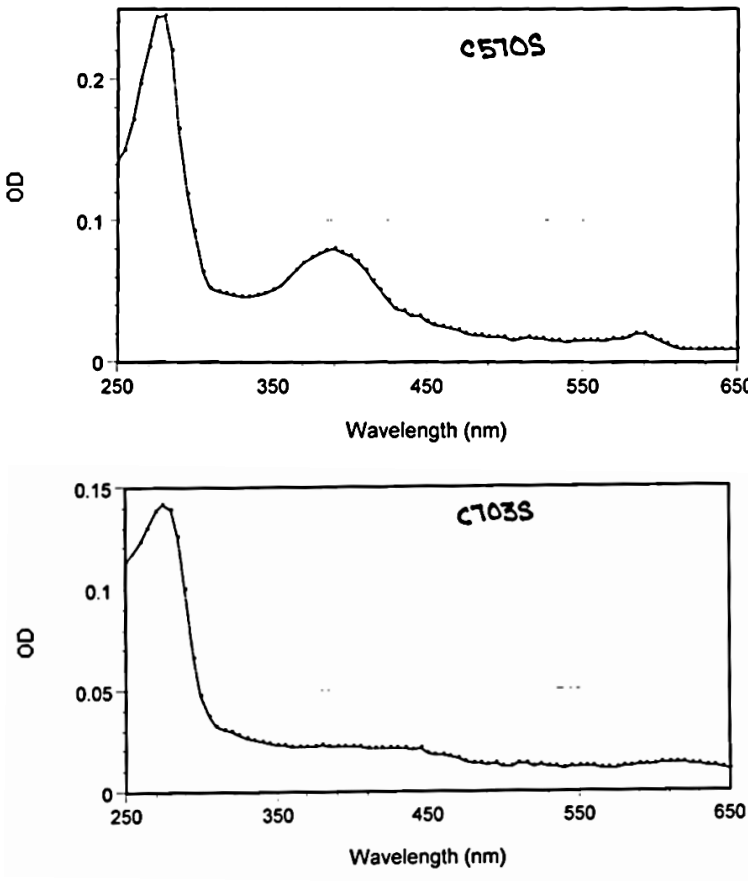


Figure 6.4: Absorbance of cysteine to serine (C570S and C703S) mutants of ferredoxin-sulfite reductase of *W.succinogenes.*

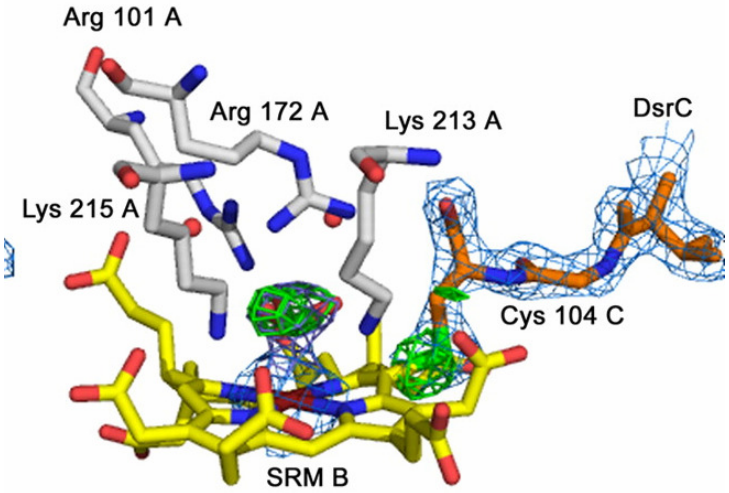


Figure 6.5: Cysteine 104 of *D.vulgaris* DsrC bound to siroheme cofactor of DsrA (adapted from *J.Biol.Chem.,* 283, 34141-34149, **2008**)

A recent report on the crystal structure of *Desulfovibrio vulgaris* dissimilatory sulfite reductase bound to DsrC, a 11 kDa protein, shows the penultimate cysteine (C104) of DsrC bound to the siroheme co-factor of its partner[3]. This cysteine is said to be conserved among all family members (including YccK/TusE, a sulfur-transfer DsrC-homolog that was shown recently to be involved in thio-modifications of bacterial tRNA wobble positions). A second cysteine of DsrC, Cys-93, is conserved only in DsrC proteins involved in dissimilatory sulfur metabolism. The C-terminal arm of DsrC extends into the active-site of DsrA and DsrB subunits (the sulfite reductases) reaching to the siroheme center. This raises the question whether the siroheme-cysteine cross-linking is mechanistically relevant. It might be possible that Cys-703 of *W.succionogenes* is necessary to form the cross-link with the siroheme co-factor. This structural issue could be a reason for the low-expression/solubility of the C703S mutant of ferredoxin-sulfite reductase of *W.succinogenes.* Cys-570 could, on the other hand, be involved in transferring the sulfide from the siroheme-Fe center to the activated WsHcyS-AMP through a persulfide linkage as shown in **Figure 6.6**.

6.3.2 Crystal structure of the proteins

X-ray crystallography studies on the proteins involved in the methionine biosynthetic pathway would help us understand the interactions between the proteins. For the first time, a thiocarboxylate-forming protein (in this pathway, WsHcyS) has been found along with its sulfur source (WsSir). Co-crystallization of WsHcyS, both its full-length and the truncated form, with its partner proteins - WsSir, WsHcyD, WsHcyF and WsMetY would shed some light on the mechanism of sulfur transfer. A structure of WsSir would also help us understand the role played by the conserved cysteines in sulfur reduction and transfer chemistry. It would be interesting to see if C703 also cross-links to the siroheme co-factor.

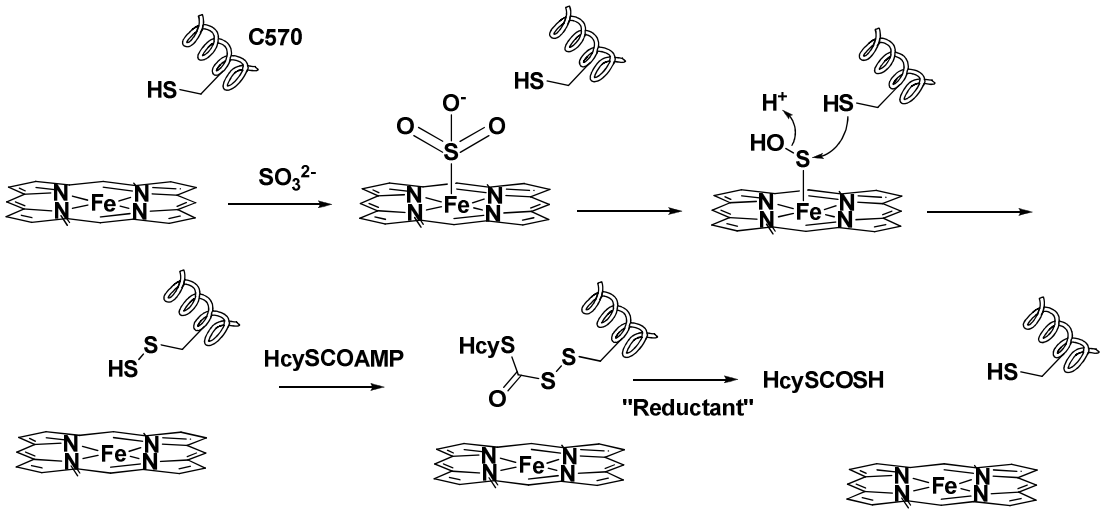


Figure 6.6: Plausible role of Cys-570 of WsSir in the sulfur transfer to WsHcyS-AMP adduct to form the protein thiocarboxylate.

6.3.3 Fate of the putative homolanthionine

Figure 4.17 (**Chapter 4**) shows a higher molecular-weight WsHcyS adduct formed upon treating WsMetY with truncated WsHcyS-COSH and excess O-acetyl-L-homoserine. A possible mechanism for formation of this adduct is shown (**Figure 6.7**).

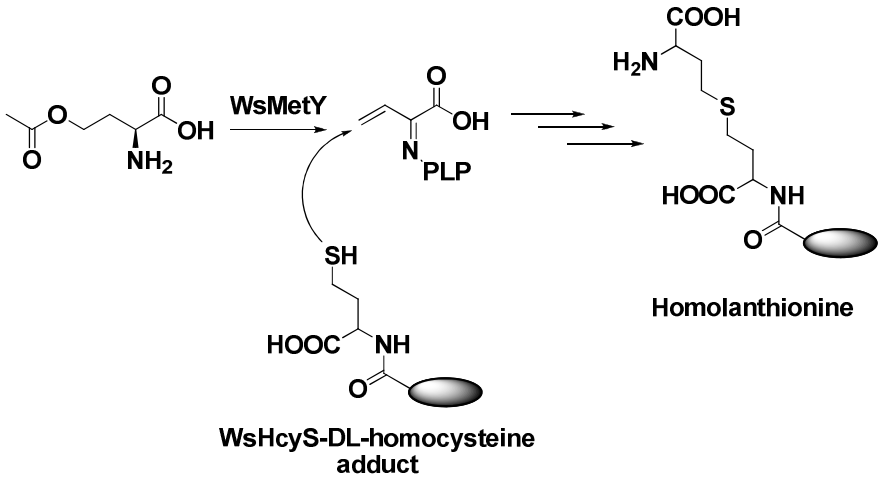


Figure 6.7: Plausible mechanism for the higher molecular-weight WsHcyS adduct formation.

In a previous study, it was reported that homolanthionine and homocysteine were present in elevated levels when the methionine and cysteine biosynthesis repressor protein (McbR) in *Corynebacterium glutamicum* ATCC 13032, which releases almost all enzymes of methionine biosynthesis and sulfate assimilation from transcriptional regulation, was deleted[4]. No overproduction of methionine was seen. It was also seen that cystathionine-γ-synthase (MetB), an enzyme that condenses cysteine with O-acetyl-L-homoserine to make cystathionine, produced homolanthionine as a side reaction. Deletion of metB completely prevented homolanthionine accumulation. The cell was seen to cleave homolanthionine at low rates via cystathionine-β-lyase (MetC). The by-product of MetC-dependent cleavage of homolanthionine, 2-oxobutanoate, could be used for isoleucine biosynthesis through a novel threonine-independent pathway, an hypothesis supported by an increased intracellular isoleucine level. Again, the question regarding the relevance of the higher molecular weight WsHcyS adduct still remains unanswered.

REFERENCES

1. Crich, D., Sasaki, K., Rahaman, M. Y. & Bowers, A. A. One-pot syntheses of dissymmetric diamides based on the chemistry of cyclic monothioanhydrides. Scope and limitations and application to the synthesis of glycodipeptides. *J Org Chem* 74, 3886-93 (2009).
2. Crich, D., Sana, K. & Guo, S. Amino acid and peptide synthesis and functionalization by the reaction of thioacids with 2,4-dinitrobenzenesulfonamides. *Org Lett* 9, 4423-6 (2007).
3. Oliveira, T. F. et al. The crystal structure of Desulfovibrio vulgaris dissimilatory sulfite reductase bound to DsrC provides novel insights into the mechanism of sulfate respiration. *J Biol Chem* 283, 34141-9 (2008).
4. Kromer, J. O., Heinzle, E., Schroder, H. & Wittmann, C. Accumulation of homolanthionine and activation of a novel pathway for isoleucine biosynthesis in Corynebacterium glutamicum McbR deletion strains. *J Bacteriol* 188, 609-18 (2006).