STRUCTURAL STUDIES OF SULFUR TRANSFER IN THIOCARBOXYLATE-DEPENDENT METHIONINE BIOSYNTHESIS

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

Megan Christine Kopp

January 2014

STRUCTURAL STUDIES OF SULFUR TRANSFER IN THIOCARBOXYLATE-DEPENDENT METHIONINE BIOSYNTHESIS

Megan Christine Kopp, Ph. D.

Cornell University 2014

Macromolecular crystallography allows for the determination of structures to an atomic resolution and can provide important insight into the function and mechanism of enzymes. Including enzymes involved in interesting reactions and complex formations. In *Wolinella succinogenes* a sulfur carrier protein pathway was discovered to serve as the sulfur source for methionine biosynthesis. During this process, the sulfur carrier protein HcyS must form complexes with a variety of proteins to accomplish the C-terminal thiocarboxylate formation and the biosynthesis of methionine. This ability to form the different complex interactions makes the pathway fascinating in its structural characteristics. Additionally, the methionine pathway has a novel sulfur source when compared to other sulfur carrier protein pathways. The sulfur for the thiocarboxylate is supplied by ferredoxin sulfite reductase (FSR). The structure of FSR has been solved and reveals a more complex function than simply reducing sulfite to sulfide. The presence of a cysteine in the active site suggests the FSR may also serve as a sulfurtransferase to control the formation of the sulfur carrier protein thiocarboxylate. The structure of FSR provides further insight into the formation of complexes by HcyS.

BIOGRAPHICAL SKETCH

Megan Kopp was born and raised in the northern panhandle region of West Virginia near the Pittsburgh area. In her high school years she took many science classes and became interested in both biology and chemistry. Megan attended West Liberty University in West Liberty, WV, graduating magna cum laude with two degrees, Bachelors of Science in Biology and Chemistry in the spring of 2009. While at West Liberty University, Megan had the opportunity to participate in undergraduate research with Dr. Jarrett Aguilar as her advisor. During this time she studied molecular dynamic computer simulation of Cytochrome P450 2C9 ligand binding.

After completing her undergraduate degrees, Megan began her Ph.D. studies in Chemistry and Chemical Biology at Cornell University in the fall of 2009. After the first academic year with the coursework finished, she began studying enzyme structure using X-ray crystallography in Prof. Steven Ealick's lab in the summer of 2010. Her work was focused on the sulfur carrier protein pathway and specifically on the sulfite reductase of this pathway, FSR. Megan's future plans include furthering her scientific knowledge during a postdoctoral position then pursuing a career in the pharmaceutical industry.

For my Grandmother who always believed in me and couldn't stop from telling everyone about it. I miss you.

ACKNOWLEDGMENTS

I would first like to thank Prof. Steven E. Ealick for mentoring me during the past four years. He has provided me with challenging projects to inspire creative scientific thought and curiosity and given me the chance to learn invaluable skills during my time in his lab. I owe great thanks to the former Ealick lab member Dr. Timothy H. Tran for his work on the purification and the identification of initial crystallization hits of FSR. I would also like to thank Prof. Tadhg P. Begley for his collaboration on the HcyS sulfur carrier protein pathway. I owe much thanks to Dr. Bekir Eser, a member of the Begley lab for his work on this project. I would also like to thank the Prof. Hening Lin and Prof. Richard A. Cerione for agreeing to be committee members and offering their advice on my research. Many thanks to Ms. Leslie Kinsland for her help in countless ways and times. I owe many thanks to Dr. Cynthia Kinsland for cloning of plasmids and her advice on protein expression and purification. I would like to thank the beamline staff of NE-CAT and CHESS for their advice on data collection. Finally, I give my thanks to all the current and former Ealick lab members for your advice both practical and scientific in nature, and your support.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

NMR, nuclear magnetic resonance

SAXS, small angle X-ray scattering

Cyro-EM, cyrogenic Electron Microscopy

APS, Advanced Photon Source

NE-CAT, Northeastern Collaborative Access Team

MAD, multiwavelength Anomalous Diffraction

SAD, single-wavelength Anomalous Diffraction

Ub, ubiquitin

ATP, adenosine 5′-triphosphate

DXP, 1-deoxy-D-xylulose-5-phosphate

mnm^5s^2U, 5-methylaminomethyl-2-thiouridine

AdoMet, S-adenosylmethionine

APS, adenosine 5′-phosphosulfate

PAPS, 3'-phosphoadenosine-5'-phosphosulfate

AMP, adenosine 5′-monophosphate

FSR, ferrodoxin sulfite reductase

OAH, *O*-acetylhomoserine

OAHS, *O*-acetylhomoserine sulfhydrylase

PLP, pyridoxal 5′-phosphate

LB, Luria-Bertani

IPTG, isopropylthio-β-D-thiogalactopyranoside

Tris, 2-amino-2-hydroxymethyl-propane-1,3-diol

DTT, dithiothreitol

PEG, polyethylene glycol

SDS-PAGE, sodium dodecyl sulfate polyacrylamide gel electrophoresis

CAPS, 3-(cyclohexylamino)-1-propanesulfonic acid

Sir, sulfite reductase

Nir, nitrite reductase

NADPH, nicotinamide adenine dinucleotide phosphate

SeMet, selenomethionyl

HEPES, 4-(2-hydroxyethyl)piperazine-1-ethanesulfonic acid

MPD, 2-methyl-2,4-pentanediol

MES, 2-(N-morpholino)ethanesulfonic acid

CHESS, Cornell High Energy Synchrotron Source

AMPCPP, α,β-Methyleneadenosine 5′-triphosphate

CHAPTER 1: INTRODUCTION

Structural biology is a field tasked to discover and understand the structural basis of molecular processes, and is based on the idea that to understand it is necessary to see. Several tools are available to gain this structural information including electron microscopy, nuclear magnetic resonance (NMR), small angle X-ray scattering (SAXS), and X-ray crystallography [1, 2]. Electron microscopy can provide low resolution structural information about interactions of macromolecules, while NMR presents dynamic structures for aqueous solutions. Small angle X-ray scattering provides a view of the molecular envelope. X-ray crystallography is among the most powerful of these techniques, as it can offer higher resolution information of a static structure within a crystal lattice [1].

Cryo-electron microscopy (cryo-EM) is a technique that uses a beam of electrons focused by electrical fields to see a magnified image of very large proteins and complexes [3]. This method produces a low resolution structural image of the macromolecule, up to approximately 4 Å though more often in the 10-25 Å range. It can give a snapshot of a dynamic conformational landscape. In cases of high resolution cryo-EM (4 Å), it is possible to build a model into the image, though this is usually limited to proteins with high symmetry and rigid structures, such as virus particles (Figure 1.1) [4, 5]. Staining by metals to more strongly scatter the electron beam is common in other types of EM but the staining method produces distortions in the structure that are not acceptable for high resolution structures. For this reason, cryo-EM samples are not stained and are instead flash-frozen with a cryo-protectant similar to the method used in X-ray crystallography that will be discussed later in this chapter. To obtain higher resolution images, it also makes use of image enhancement, which combines the Fourier transform with electron imaging to increase resolution and decrease the signal to noise

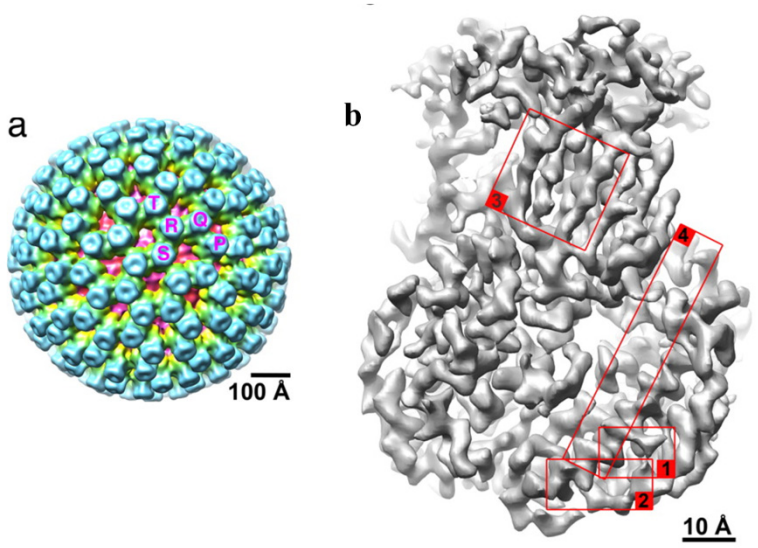

Figure 1.1. (a) Cryo-EM image modified from [4] that shows the cryo-EM image of a rotavirus particle. (b) Cryo-EM of one of proteins that make up the virus particle

ratio making it easier to resolve the image. The use of the Fourier transform, however, averages the images of all the particles yielding a single image that represents the conformation to all particles in the solution. This can be problematic if the solution is heterogeneous, which compromises the averaged image [5]. The method is limited because of the resolution, and provides snapshots of macromolecules of greater than 70 kDa, but image sorting can provide information about conformational changes. All together, these images can provide a more complete picture of dynamic formations, and can provide structural information that can be used to infer function [6].

A high profile example of the power of cyro-EM is shown by studies of the kinesin motor protein, which drives motility [7]. This protein converts ATP energy into unidirectional motion along the microtubule network. They are involved in intracellular transport and control of microtubule dynamics and signal transduction, and are conventionally dimers of identical 120

kDa monomers. Cryo-EM was used to visualize conformational changes of a flexible linker and propose a mechanism for how these motions drive motility along the microtubule [7].

NMR is used to obtain structural information for proteins in solution. It has limitation on the size of proteins whose structure can be determined by this method, and is often limited to proteins with molecular masses less than 30 kDa. The larger the protein, the more peaks in the NMR spectrum, generally creating an upper limit for structure determination by NMR of approximately 50 kDa [8]. Because of the peak overlap, 2D and 3D spectrum techniques are used, but overlap of peaks still occurs for larger proteins (Figure 1.2).

The basic NMR experiment involves placing the sample between the poles of two powerful magnets creating nuclear spin, and a radio frequency pulse is transmitted to the sample. The sample absorbs this energy then relaxes releasing a radio frequency signal to the detector that contains the chemical shift information of the nuclei. This is 1D NMR, which only gives chemical shift information. 2D NMR gives information not only about the chemical shift, but also about coupling between the nuclei of adjacent atoms. This experiment involves a complex series of pulses, delays, and data collection. But it is necessary for larger molecules such as proteins in order to separate the complex shifts that will allow for model building [3].

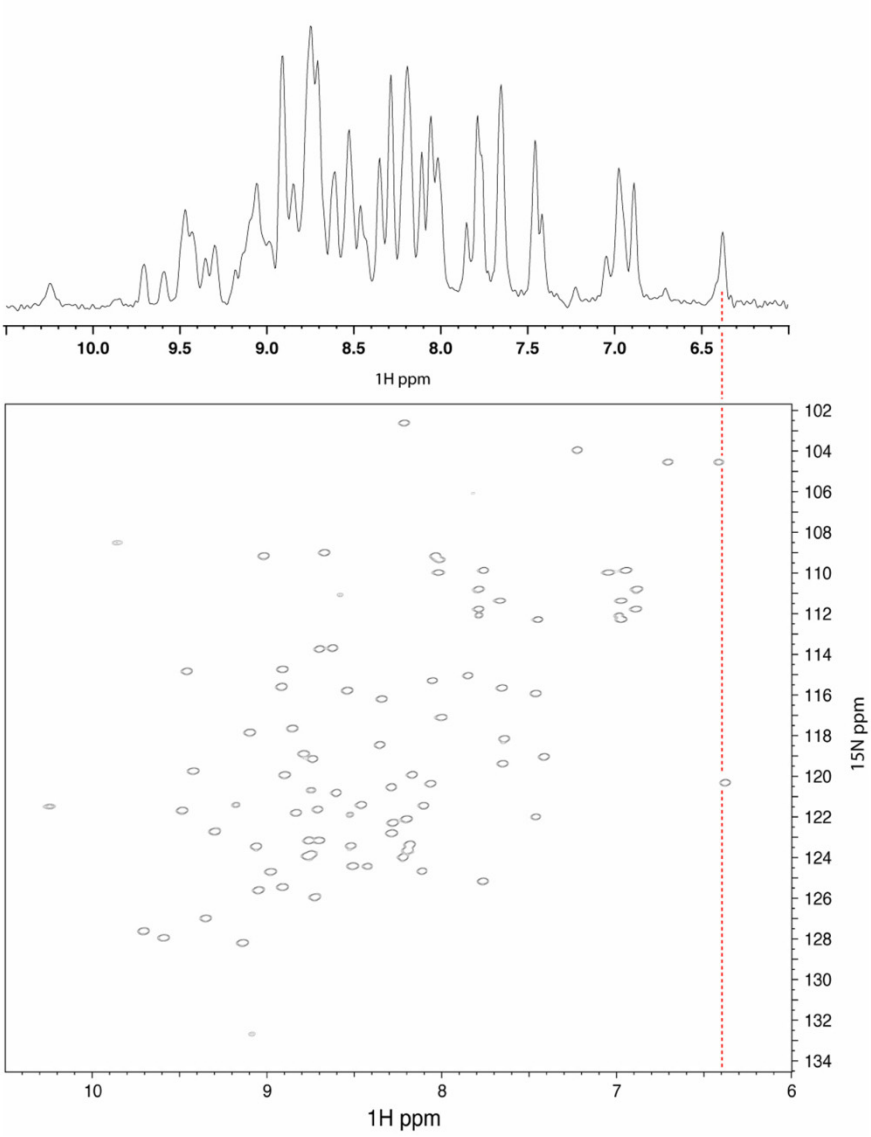

Figure 1.2. Comparison of NMR 1D (top) and 2D (bottom) for the same protein

With 2D NMR, the peaks of the NMR spectrum can provide information on the environment of a residue, and then by looking at this environment the peak can be associated with a specific residue. Once all the peaks are assigned to specific residues, the information can be compiled to build a structural model of the protein [8]. Once the structure is determined, further experimentation can be done with NMR including the addition of ligands, or even other proteins to form a complex. The addition of these molecules can provide a real-time dynamic

representation of the conformational changes that take place when the protein binds a ligand or another protein [9]. NMR solution structure has been applied to many proteins. One high profile example, that is not amenable to X-ray crystallography, is the structure of human prion proteins, which are associated with fatal disease characterized by loss of motor control, dementia, and paralysis [10]. Solution of these structures is very important because of the serious nature of the diseases in which they are involved and can lead to treatment options. NMR was used to solve the structures of human versions of the prion proteins to identify conformational changes that lead to the aggregation of the protein during the onset of the disease (Figure 1.3) [10].

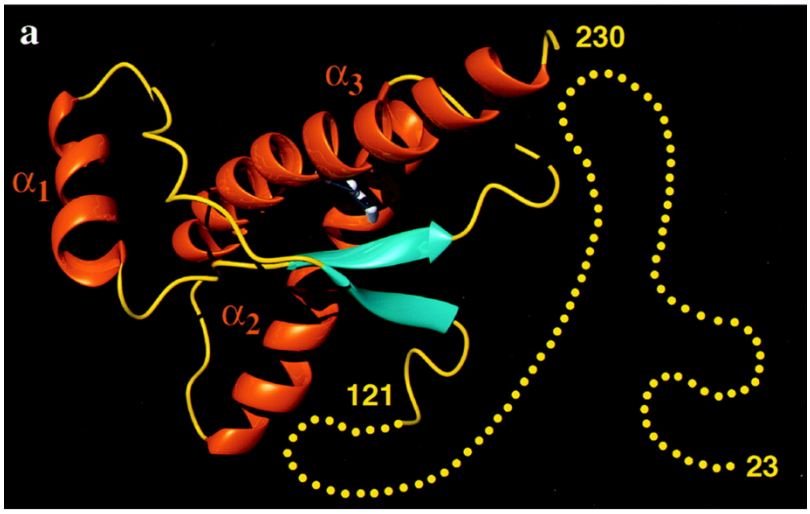

Figure 1.3. Model of NMR solution structure of the human prion protein. Figure adapted from [10]

SAXS is a technique used for structural characterization of non-crystalline materials including macromolecular proteins [11]. A chaotically dispersed protein solution is irradiated with an X-ray beam, and the scattered X-rays are recorded. A scattering profile of a macromolecule can give information about the molecular mass, and consequently the oligomeric state or stochiometry of complex formation. Differences in this scattering profile can also indicate if the protein is unfolded or disordered [12]. First, the solvent scattering is recorded by

using a blank that contains only buffer, then the scattering data is collected for the sample containing the molecule of interest. The solvent scattering is then subtracted from the sample scattering to create a scattering curve for the macromolecule (Figure 1.4) [13].

$$\Delta\rho(r)=\rho(r)-\rho(s)$$

Figure 1.4. Equation of the SAXS scattering: the scattering of the solvent (ρ(s)) is subtracted from the scattering of the solution to determine the scattering of the molecule of interest.

The resulting scattering curve *I(q)* is radially symmetric (isotrophic) due to the randomly oriented molecules in the solution, and is a function of the magnitude of the scattering vector (Figure 1.5). The scattering curve of a homogenous solution can then be derived from the electron distribution of the particle *P(r)* shown in Figure 1.6. The equation used to create the scattering curve is shown in Figure 1.7.

$$q = \frac{4\pi sin\theta}{\lambda}$$

Figure 1.5. Scattering vector equation where 2θ is the scattering angle and λ is the wavelength of incident X-ray beam.

$$P(r) = \frac{r}{2\pi^2}\int_0^\infty I(q)qsin(qr)dq$$

Figure 1.6. Electron distribution of the particle, pair-wise distribution function

$$I(q) = 4\pi\int_0^{Dmax} P(r)\frac{\sin(qr)}{qr}dr$$

Figure 1.7. Scattering curve equation where *P(r)* is the electron distribution particle and Dmax is the maximum distance in the scattering particle

The scattering curve can then be analyzed at different values to extract different parameters about the molecule. At lower q values this includes the radius of gyration (R_g), which gives an estimation of molecular size, and the intensity at the zero angle (I_0), which is necessary for molecular mass calculation. Higher q values hold information about the molecular shape, and

can be particular important for characterization of folded versus unfolded protein, and flexibility of folded proteins with linker regions.

One example of the usefulness of SAXS in structural studies is in exploring the dimerization of a blue-light signal transduction protein, Vivid (VVD) [14]. VVD is dependent on light for signal transduction. Photon absorption causes a conformational change in a domain of VVD and this conformational change leads to a monomer-dimer equilibrium. SAXS was used to determine the dimerization of the VVD, after the excitation of the sample with a laser the *I(0)* increased suggesting a higher molecular weight indicative of dimer formation [14].

X-ray crystallography is the most powerful and most commonly used method for obtaining structural data [3]. The power of X-ray crystallography is based on the idea that function is determined by structure, and that accurate, atomic resolution structural data can provide much information about a protein's function. The atomic details of structures can help elucidate the mechanism of an enzyme, which residues are involved in binding the substrate, and which residues participate in catalysis. X-ray crystallography has supplied a great deal of information for mechanistic enzymology by providing structures of protein complexes with substrates, products, intermediates, or their analogs. Protein complexes can often be co-crystallized and provide information about the residues important for complex formation, and how proteins of the complex contribute to the mechanism of action. It is often possible to trap a ligand in the crystal structure through the use of co-crystallization or soaking techniques. This allows for the identification of key catalytic residues, and a proposal of mechanism, which can suggest further experiments to provide evidence of function. It is also possible that a mutant enzyme will partially catalyze the substrate trapping an intermediate in the crystal structure. This can provide further information for mechanistic proposals.

The enzymes mechanism and ligand binding profile can then be used for structure guided drug design often to design a molecule that will tightly bind in the enzyme's active site and inhibit its function [15]. The role of protein drug target crystallography is increasing as the number of therapeutic drugs on the market that were directly designed using crystallography increases. Among these is the human immunodeficiency virus (HIV) protease inhibitor that was designed using the HIV protease structure [16]. Structure guided drug design can also be seen in the 2012 Nobel Prize in Chemistry, which was awarded to Robert J. Lefkowitz and Brian K. Kobilka for their work on the crystal structure of G-protein coupled receptors (GPCRs) (Figure 1.8). GPCRs are membrane proteins that mediate signals from the outside of the cell, and are of interest as pharmacological targets to treat a variety of conditions. This work has allowed for the design of small molecule inhibitors and agonists that increase or decrease the signaling activity as necessary for the physiological condition being treated [17].

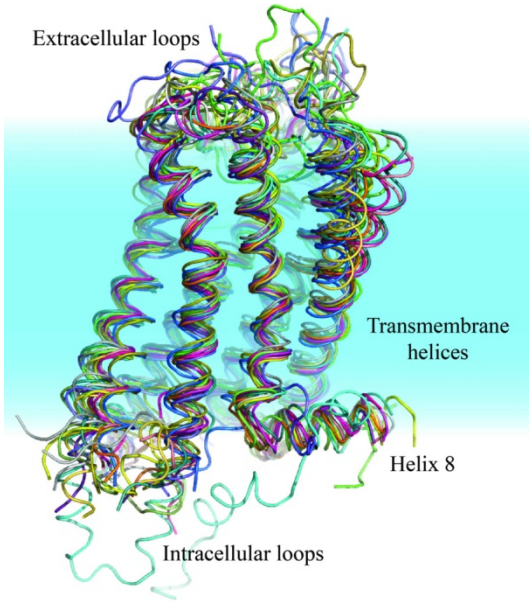

Figure 1.8. GPCRs structures superimposed to show the conservation of the structure. Figure adapted from [18]

Advances in technology have increased the ease of structure determination, and has expanded the use of X-ray crystallography to nonspecialists. Structure determination by X-ray crystallography is relatively straightforward as illustrated in Figure 1.9. However, practical challenges, such as growing protein crystals and obtaining accurate phases can be problematic.

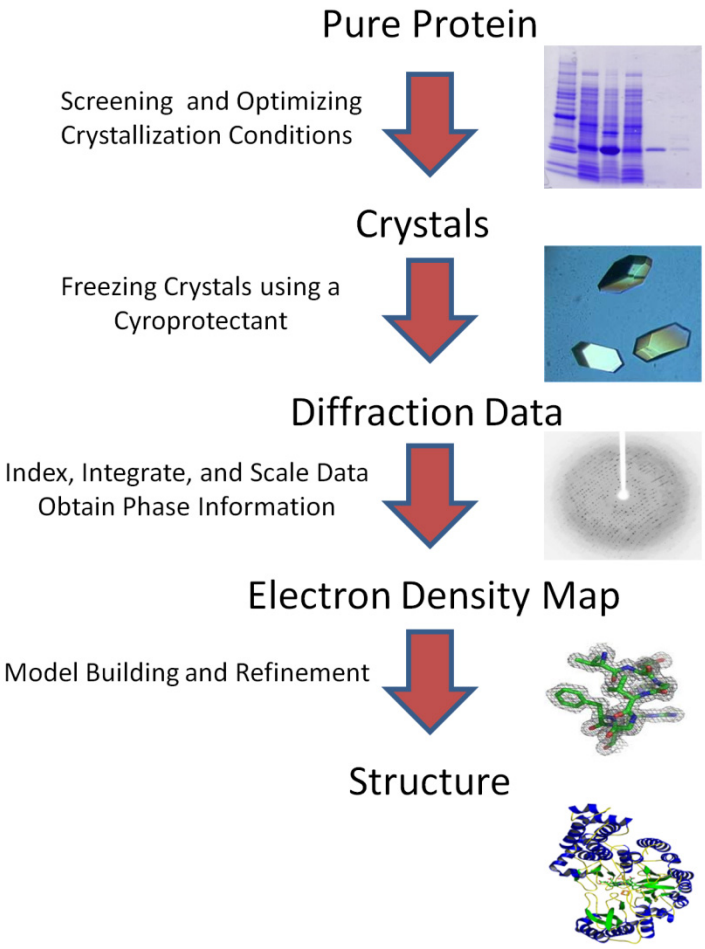

Figure 1.9. Flowchart of the protein crystallography experiment

Preparing and modifying a sample to allow for crystallization can be nontrivial, and often involves careful planning and understanding of the target protein [19]. Crystals must be very well ordered, containing many copies of the protein in identical orientations, to create high resolution X-ray diffraction patterns. Optimizing the crystallization condition to increase the

quality of the crystals can be a difficult process, and in some cases crystal quality will prevent structure determination. In this case, it becomes necessary to start the process again to find a new crystallization condition [3].

Once the crystals are grown, the crystals must be flash frozen in liquid nitrogen to decrease the radiation damage to the crystal during data collection. When freezing the crystals, the crystallization condition must be supplemented with cryo-protectant to prevent ice formation, which will disrupt the arrangement of molecules in the crystal. Typical cryo-protectants include glycerol and ethylene glycol ranging from 10-30% (Figure 1.10).

Figure 1.10. Example picture of a crystal in a loop for data collection

An X-ray diffraction experiment is performed by placing a single crystal in a focused X-ray beam and recording diffraction images. Because X-rays cannot be focused by lenses it is not possible to view an image of a molecule through this method, as is done in optical microscopy [3]. Instead, the electron density of the molecules in the crystal diffracts the X-ray beam to create a diffraction pattern (Figure 1.11). Each diffraction maximum is characterized by a set of Miller indices (h, k, l) and a diffraction angle known as 2θ or the Bragg angle. The value of this angle is determined by the wavelength of the X-rays, the unit cell parameters (a, b, c, α, β, γ), and the

orientation of the crystal.

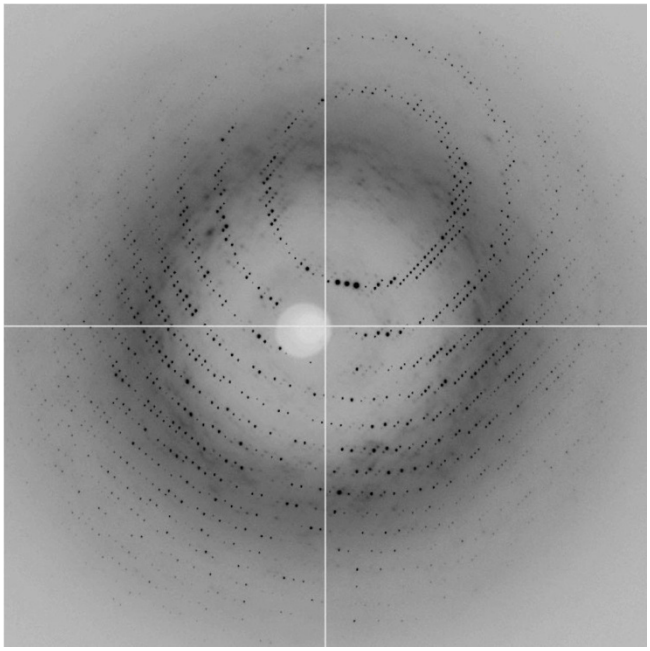

Figure 1.11. Example image of a diffraction pattern

The intensity of each maximum is the square of the structure factor F, which is calculated using the equation shown in Figure 1.12.

$$F(h,k,l) = \sum_j f_j e^{2\pi i(hx_j+ky_j+lz_j)}$$

Figure 1.12. *F(h, k, l)* is the structure factor, h, k, l are the Miller indices of the structure factor, f_j is the resolution-dependent atomic scattering factor, (x, y, z) are the fractional positions of the atom j in the unit cell.

The structure factor intensities of the diffraction image and the phases are combined to reconstruct the electron density by a Fourier transform (Figure 1.13).

$$\rho(x,y,z) = \frac{1}{V}\sum_{h}\sum_{k}\sum_{l} F_{hkl}\, e^{-2\pi i(hx+ky+lz)}$$

Figure 1.13. Fourier transform equation that is used to convert diffraction data to an electron density map

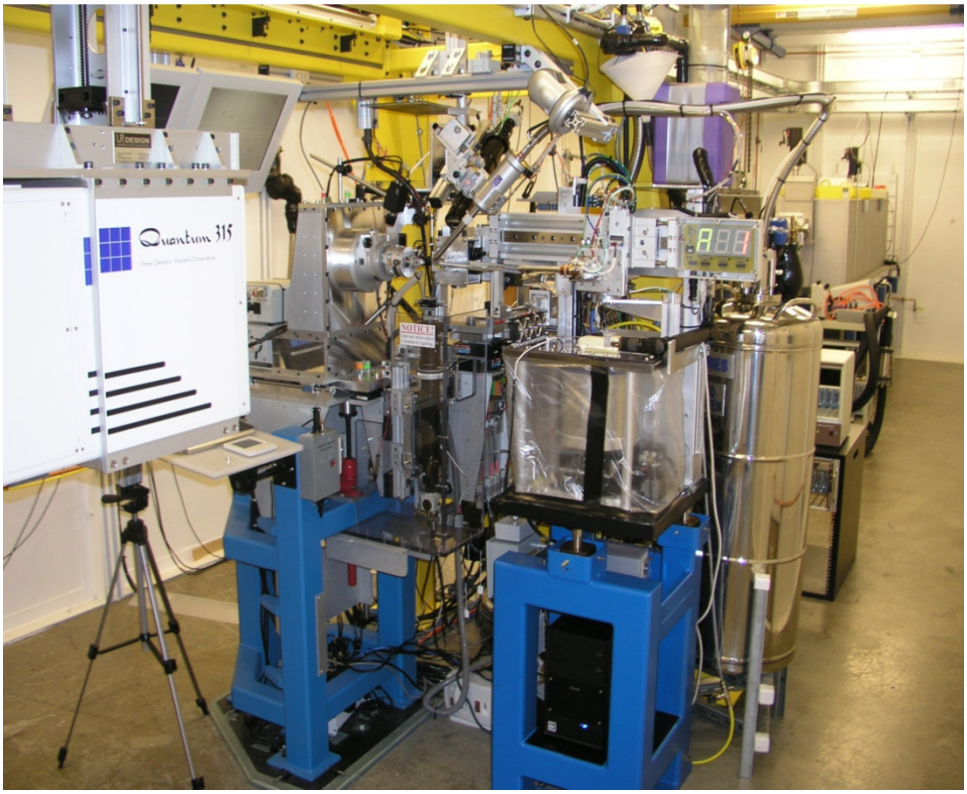

Figure 1.14. The 24-ID-C hutch used for X-ray crystallography data collection at the Northeastern Collaborative Access Team (NE-CAT) beamline at the Advance Photon Source (APS) in Argonne National Labs, Argonne, IL

Data collection usually takes place at a synchrotron beamline configured for macromolecular crystallography (Figure 1.14). Once the data collection is complete, phase information must be generated for calculation of electron density maps. The diffraction images collected for a crystal contains only intensities of the diffractions spots, but not the phase of the diffracted X-ray wave. Several methods can be used to obtain phases, including molecular replacement, multiwavelength anomalous diffraction (MAD), single-wavelength anomalous diffraction

(SAD), and direct methods. In molecular replacement, the phases from structure factors of a known protein structure are used to determine the initial phases by placing the phasing model into the unit cell of the new protein [3]. The molecule used for phasing is identified by sequence alignments to find proteins that are predicted to be structurally similar to the protein of interest. If this method can be used, then the structure can be determined from a single native data set.

MAD and SAD phasing uses incorporation of a heavy atom to perturb the X-ray diffraction pattern of the protein crystal. The heavy atom is often selenium, which can be incorporated by replacing methionine residues with selenomethionine, or by soaking crystals in solution containing metals, which can bind to the protein surface. In MAD phasing at least two data sets are collected from the same crystal but at different wavelengths, one of which is the wavelength of maximum absorbance for the heavy atom and the other is shifted away from this maximum value. The differences in the diffraction intensities for the two wavelengths can be used to obtain the phases [3]. SAD phasing is very similar except that it requires only one data set at the collected at the wavelength of maximum absorption for the heavy atom. SAD is presently the most commonly used technique [19].

Practical application of direct methods is limited to structure determination for molecules with fewer than 200 atoms. It relies upon the conditional probabilities between the normalized structure factor magnitudes and the phases. While these direct methods fail for large protein molecules, they are used extensively for determining the location of the heavy atoms required for SAD or MAD phasing.

The model is then built into the electron density map with many iterations of model building and refinement. The quality of the model agreement with the experimental electron density data can then be evaluated be the use of the R_{factor} (Figure 1.15). Another useful figure

during refinement is R_{free} which can detect overfitting by calculating the R_{factor} for the current model using a set of reflections that are withheld from the refinement.

$$R = \frac{\sum |F_0(h,k,l) - F_c(h,k,l)|}{\sum F_0(h,k,l)}$$

Figure 1.15. Rfactor equation where F_0 is the experimental amplitudes and F_c are the amplitudes calculated from the model.

In X-ray crystallography, the crystal lattice, which limits the freedom of motion, may affect the protein structure [19]. This usually not a serious problem because protein crystals are actually a two-phase system: a crystalline solid and together with aqueous solvent channels that provide a solution-like environment. Comparison to solution NMR structures show similar core structure [20]. In cases where proteins go through large conformational changes, crystallization selects one conformational state. However, different crystal forms can often select different conformations, which allows one to infer information about protein motions and the functional role of conformational changes.

X-ray crystallography is complementary to both a variety of other experimental approaches and to other structural methods. The combination of X-ray crystallography with NMR can provide more information about how proteins move through conformational changes giving dynamic information to the static crystal structure [9]. It can also be combined with electron microscopy since such large proteins and complexes can be difficult to crystallize as a whole. Smaller domains of the large protein can be solved by X-ray crystallography then fitted into the electron microscopy image of the large complex to understand how the smaller parts fit together to make the larger structure [6]. SAXS can be used to characterize the oligomeric states and folding of the protein under different conditions that could be useful to attempts to

crystallize the protein [11]. Additionally, experiments such as mutagenesis, kinetics, circular dichroism, size exclusion chromatography and other techniques can be used to gain more insight into the structural information and to provide a complete story of enzyme function.

REFERENCES

1. Liljas, A., *On the complementarity of methods in structural biology.* Acta Crystallogr D Biol Crystallogr, 2006. **62**(Pt 8): p. 941-5.
2. Campbell, I.D., *Timeline: The march of structural biology.* Nature Reviews Molecular Cell Biology, 2002. **3**(5): p. 377-381.
3. Rhodes, G., *Crystallography Made Crystal Clear: A Guide for Users of Macromolecular Models*. 2010: Elsevier Science.
4. Zhang, X., et al., *Near-atomic resolution using electron cryomicroscopy and single-particle reconstruction.* Proc Natl Acad Sci U S A, 2008. **105**(6): p. 1867-72.
5. Spahn, C.M. and P.A. Penczek, *Exploring conformational modes of macromolecular assemblies by multiparticle cryo-EM.* Curr Opin Struct Biol, 2009. **19**(5): p. 623-31.
6. Lander, G.C., H.R. Saibil, and E. Nogales, *Go hybrid: EM, crystallography, and beyond.* Curr Opin Struct Biol, 2012. **22**(5): p. 627-35.
7. Rice, S., et al., *A structural change in the kinesin motor protein that drives motility.* Nature, 1999. **402**(6763): p. 778-84.
8. Wuthrich, K., *Protein structure determination in solution by NMR spectroscopy.* J Biol Chem, 1990. **265**(36): p. 22059-62.
9. Wuthrich, K., *Six years of protein structure determination by NMR spectroscopy: what have we learned?* Ciba Found Symp, 1991. **161**: p. 136-45; discussion 145-9.
10. Zahn, R., et al., *NMR solution structure of the human prion protein.* Proc Natl Acad Sci U S A, 2000. **97**(1): p. 145-50.
11. Mertens, H.D. and D.I. Svergun, *Structural characterization of proteins and complexes using small-angle X-ray solution scattering.* J Struct Biol, 2010. **172**(1): p. 128-41.

12. Petoukhov, M.V. and D.I. Svergun, *Applications of small-angle X-ray scattering to biomacromolecular solutions.* Int J Biochem Cell Biol, 2013. **45**(2): p. 429-37.

13. Putnam, C.D., et al., *X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution.* Q Rev Biophys, 2007. **40**(3): p. 191-285.

14. Lamb, J.S., et al., *Time-resolved dimerization of a PAS-LOV protein measured with photocoupled small angle X-ray scattering.* J Am Chem Soc, 2008. **130**(37): p. 12226-7.

15. Blundell, T.L., H. Jhoti, and C. Abell, *High-throughput crystallography for lead discovery in drug design.* Nat Rev Drug Discov, 2002. **1**(1): p. 45-54.

16. Greer, J., et al., *Application of the three-dimensional structures of protein target molecules in structure-based drug design.* J Med Chem, 1994. **37**(8): p. 1035-54.

17. Shukla, A.K., et al., *Structure of active beta-arrestin-1 bound to a G-protein-coupled receptor phosphopeptide.* Nature, 2013. **497**(7447): p. 137-41.

18. Kruse, A.C., et al., *Applications of molecular replacement to G protein-coupled receptors.* Acta Crystallogr D Biol Crystallogr, 2013. **69**(Pt 11): p. 2287-92.

19. Rupp, B., *Biomolecular Crystallography: Principles, Practice, and Application to Structural Biology*. 1st ed. 2009, New York, NY: Garland Science. 850.

20. Etter, M.C., *NMR and X-ray Crystallography: Interfaces and Challenges : Proceedings of the*. 1989: Polycrystal Book Service.

CHAPTER 2: SULFUR CARRIER PROTEINS

The main focus of this thesis is the incorporation of sulfur in sulfate-dependent biosynthesis. In sulfate-dependent methionine biosynthesis, sulfur incorporation occurs through a sulfur carrier protein. Sulfur carrier proteins are small proteins (usually less than100 amino acid residues) characterized by a ubiquitin-like β-grasp fold, having a four-stranded β-sheet flanked by at least one α-helix (Figure 2.1) [1].

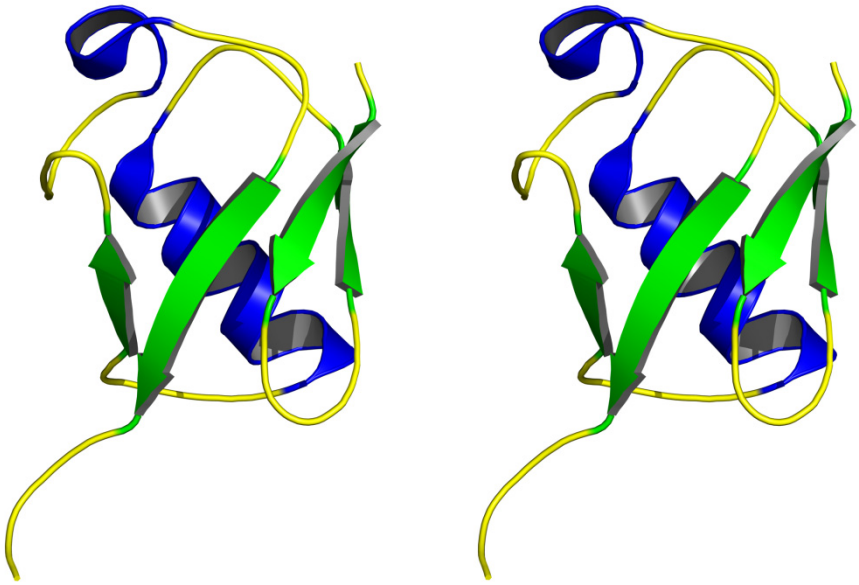

Figure 2.1. Stereoview of ubiquitin showing the β-grasp fold. The α-helices are shown in blue, β-strands in green, and loops in yellow

The sequence identity between sulfur carrier protein homologs is generally very low. Ordinarily, the only conserved feature is a diglycine motif (Gly-Gly) at or near the C-terminus (Figure 2.2) [1, 2]. The structural similarity to ubiquitin and the C-terminal Gly-Gly motif suggest an evolutionary relationship between ubiquitin with the prokaryotic sulfur carrier proteins (Figure 2.3) [1]. In contrast to ubiquitin, sulfur carrier proteins serve as the sulfur source for a variety of

biosynthetic pathways, and are very important to the biosynthesis of several molecules and cofactors necessary for living systems (Figure 2.3).

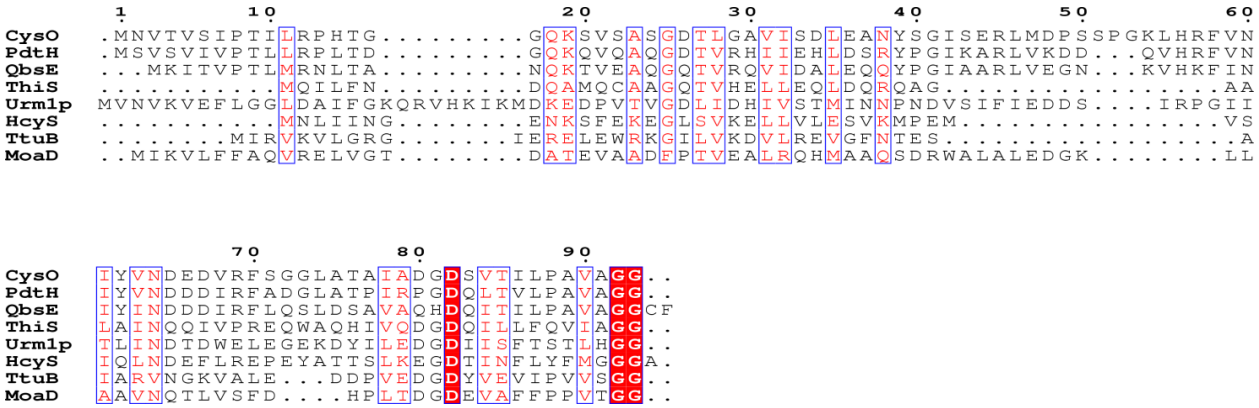

Figure 2.2. Sequence alignment of known sulfur carrier proteins, shows the low sequence identity of this protein type. The C-terminal Gly-Gly motif is an identifying feature that is conserved throughout these protein sequences.

These include vitamin B_1 (ThiS) [3], molybdopterin (MoaD) [4], thioquinolobactin (QbsE) [5], pyrimidine dithiocarboxylic acid (PdtH) [6], cysteine (CysO) [7] , 2-thioribothymidine (TtuB) [8] (Figure 2.4). A eukaryotic sulfur carrier protein involved in the biosynthesis of 5-methoxycarbonylmethyl-2-thiouridine (Urm1p), which is also used in the covalent lysine modification of proteins, has also been characterized, suggesting that Urm1p might be the evolutionary link between ubiquitin and the prokaryotic sulfur carrier proteins [9]. Another sulfur carrier protein has been identified in *Wolinella succinogenes* to be involved in the biosynthesis of methionine, which is discussed in Chapter 3 and 4 [10].

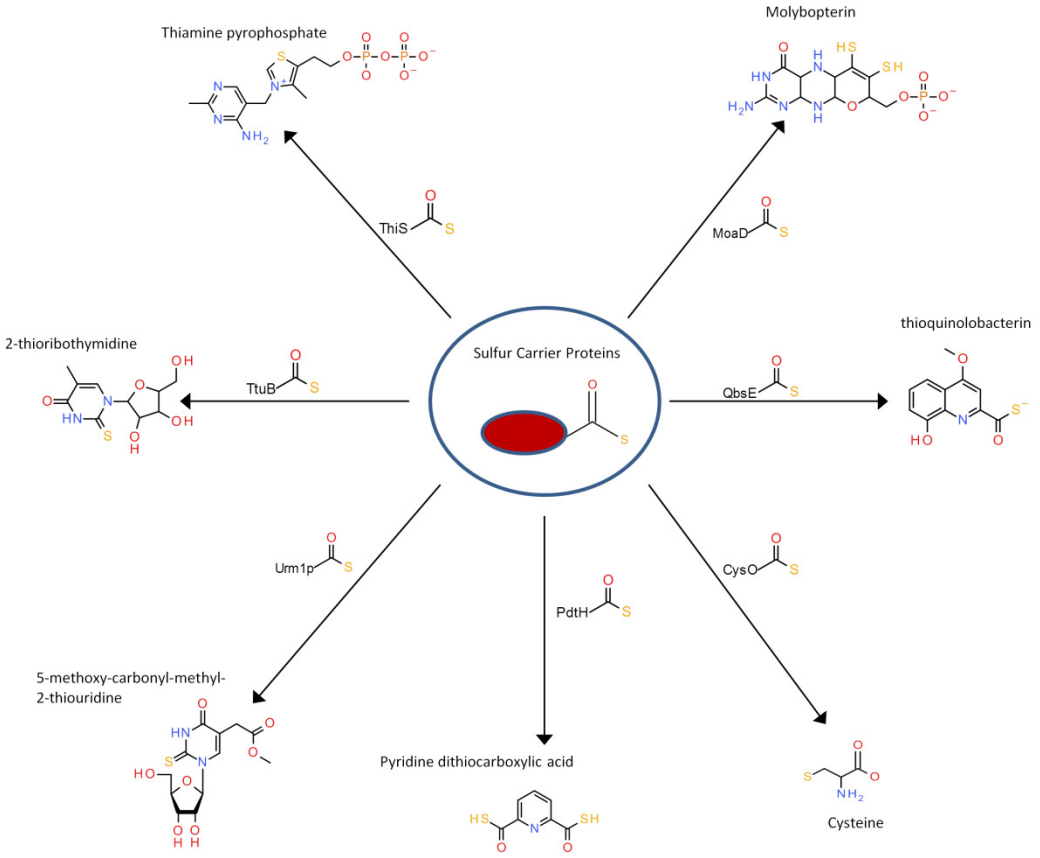

Figure 2.3. Sulfur carrier proteins with the products of their biochemical pathways

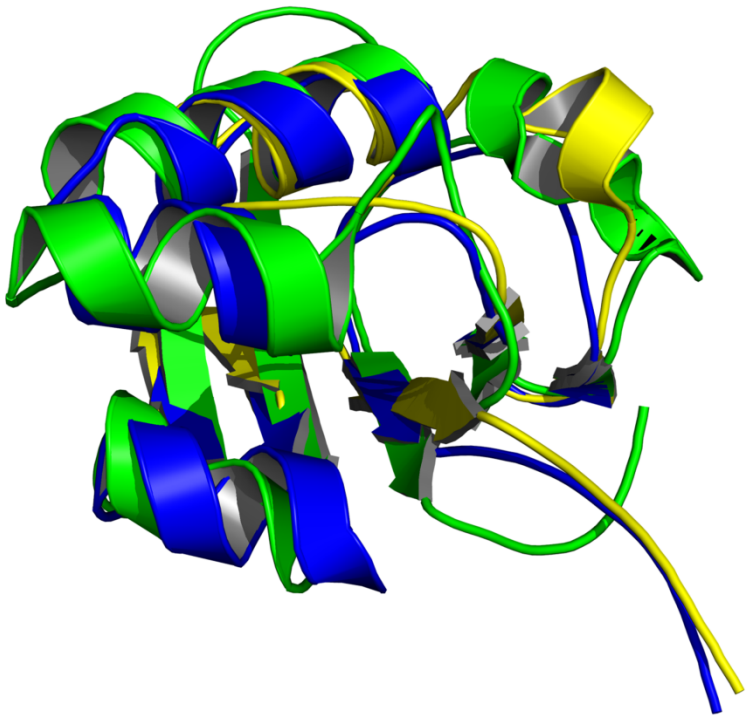

Figure 2.4. Overlay of the structures of MoaD (blue), ThiS (yellow), and Urm1 (green). The proteins as can be seen all have the same fold, the flexible linker shown on the bottom right of the image is the C-terminal containing the Gly-Gly motif.

Ubiquitin (Ub) is a small protein that is added as a covalent modifier to a target protein as a cellular signal. Though it is not a sulfur carrier protein, it is similar in structure and has a similar mechanism of action to the prokaryotic sulfur carrier proteins [1]. Ub is activated by adenylation of its C-terminal by the enzyme E1 similar to the activation of the sulfur carrier proteins. Then E1 forms a thioester with Ub C-terminal using a catalytic cysteine residue as the sulfur source. The thioester Ub-E1 complex interacts with the enzyme E2 and Ub is transferred to a catalytic cysteine of E2 forming another thioester complex. The energy of this thioester bond is then used to conjugate Ub to a lysine residue of its target protein. This is either done directly by E2 or indirectly by the complex formation with another protein, E3 [11]. This use of thioester

formation, instead of thiocarboxylate formation, differs from the mechanism used by the sulfur carrier proteins, but the activation by adenylation of the C-terminal glycine residue is the same. Additionally, sulfur carrier proteins are not used as covalent modifiers of other proteins, but serve as the sulfur source and base for product formation. Except for the protein Urm1p and the recently discovered TtuB, which serve a dual function as sulfur carrier proteins and covalent modifiers. Both of these dual function proteins transfer the sulfur of the thiocarboxylate to a tRNA source.

Urm1p identified in *Saccharomyces cerevisiea* is a similar protein to ubiquitin as it acts a covalent modifier to proteins such as Ahp1 [12] . Urm1p is activated by the E1-like enzyme Uba4 [13]. Before activating Urm1p, Uba4 first accepts a sulfur from a cysteine desulfurase (Nfs1) to form a persulfide [9]. Uba4 has a amino terminal (N-terminal) domain that adenylates Urm1p C-terminal, and a C-terminal rhodanese-like sulfur transferase domain that uses a cysteine residue to form an acyl disulfide bond with Urm1p. The disulfide bond is reduced and the Urm1p thiocarboxylate is released from Uba4 [12]. Once the thiocarboxylate is formed Urm1p can then form a conjugate to modify a target protein under conditions of oxidative stress. The conjugating enzyme expected to be similar to ubiquitin's E2 has not yet been identified. Alternatively, the Urm1p thiocarboxylate can be used to modify tRNA, making the tRNA modification 5-methoxy-carbonyl-methyl-2-thiouridine (mm^5s^2U). The Ncs2 and Ncs6 proteins are involved in the transfer of the sulfur from the Urm1p thiocarboxylate to the tRNA, but the mechanism of this process is currently unknown [9]. The method of activation and thiocarboxylate formation for Urm1p is more similar to the sulfur carrier protein mechanism than to the mechanism seen with ubiquitin.

In thermophilic organisms, 2-thioribothymidine (s^2T), a 2-thiolated derivative of ribothymidine (rT), is found at position 54 in tRNAs. This thiolation is assumed to provide stability for the tRNA at high temperatures. The biosynthesis of this modified tRNA is known to include several enzymes including the sulfur carrier proteins, TtuB. This sulfur carrier protein is activated by adenylation by the enzyme TtuC. Then it accepts a sulfide from a cysteine desulphurase (IscS or SufS) persulfide through the intermediate acyl-disulfide bond formation. Once reduced and released the thiocarboxylate of TtuB interacts with TtuA and some other unknown factors to transfer the sulfur to make s^2T [8]. However, it has recently been found that TtuB can act similarly to ubiquitin and conjugate to other proteins as a modification. Instead of forming the thiocarboxylate, the adenylated TtuB can form a thioester bond with a cysteine of TtuC. TtuB is then conjugated to proteins involved in thiouridine biosynthesis, changing the amount of s^2T modifications. The conjugating of TtuB is likely to be concerned with regulation of the sulfur transfer [14].

The mechanism of the sulfur carrier proteins with no known covalent modification function uses a similar acyl disulfide bond to a persulfide of another protein that is seen in Urm1p. Then the thiocarboxylate is used to supply the sulfur source for several cofactors, siderophores, and amino acids. This includes the vitamin B_1 or thiamin pyrophosphate biosynthesis in *Bacillus subtilis*. Thiamin is an important cofactor made up of a pyrimidine covalently linked to a thiazole. The two moieties are synthesized separately and then coupled to make thiamin [15]. The sulfur carrier protein ThiS begins the biosynthesis of the thiazole moiety of thiamin [3]. ThiS is activated by adenylation at the C-terminal by the ATP-dependent enzyme ThiF [3, 16]. The thiocarboxylate is next formed by the transfer of a sulfur from the cysteine desulfurase, IscS persulfide [3, 17]. The thiocarboxylate is then incorporated into the thiazole

moiety through a reaction catalyzed by ThiG, the rearrangement of 1-deoxy-D-xylulose-5-phosphate (DXP) and a glycine imine intermediate. The thiazole, which then contains the sulfur transferred from the ThiS thiocarboxylate, is then incorporated into the biosynthesis of thiamin [18].

A sulfur carrier protein is also used in the biosynthesis of another cofactor, called the molybdenum cofactor. Molybdopterin (MPT) is the organic metal binding portion of the molybdenum cofactor (Moco). The Moco is varied by modifications and found in a large number of proteins involved in electron transfer. The biosynthesis of this molecule in *Escherichia coli* is accomplished by a large heterotetramer complex, MPT synthase, that contains two small sulfur carrier protein MoaD subunits and two large MoaE subunits [19]. Before the MPT synthase formation, the MoaD C-terminus is activated by adenylation by MoeB. MoeB is not directly involved in the transfer of sulfur to form the thiocarboxylate [4]. The sulfur comes from the cysteine desulfurase IscS, and is transferred to the MoaD C-terminal the rhodanese-like protein YnjE [20, 21]. The sulfurtransferase TusA is also involved in sulfur transfer and increases the efficiency of MPT biosynthesis; however, it is not known if it is involved in a direct transfer or indirectly transfers the sulfur to another protein, which then transfers it to MoaD [22]. Two MoaD C-terminal thiocarboxylates are present in MPT synthase, which catalyzes the reaction that incorporates the dithiolene moiety into the precursor Z to make MPT [19].

Thioquinolobactin is a siderophore that is produced in fluorescent bacteria *Psuedomonas fluorescens*. There are two pathways for thioquinolobactin biosynthesis: the first involves tryptophan catabolism to hydroxykynurenine, and the second involves the thioquinolobactin from quinolobactin through a sulfur carrier protein mechanism. The sulfur carrier protein for this pathway is QbsE, it has a C-terminal Gly-Gly motif however it is blocked by a cysteine and a

phenylalanine residue at its C-terminal. These two residues are cleaved by a protease QbsD. The exposed C-terminal Gly-Gly motif is then activated by QbsC through adenylation. QbsC forms a persulfide at one of its cysteine residues, the sulfur source of this persulfide formation is unknown. It then transfers a sulfur from the persulfide through the acyl-disulfide intermediate to make the thiocarboxylate, before releasing the QbsE thiocarboxylate after reduction of the disulfide bond. Quinolobactin acyl adenylate is then modified by the QbsE thiocarboxylate to form thioquinolobactin [5]. Another sulfur carrier protein PdtH from *Pseudomonas stutzeri* has been identified by an assay for sulfur carrier proteins [6]. PdtH is thought to be involved in the biosynthesis of a siderophore, pyridine dithiocarboxylate, but has not been characterized.

Finally, the biosynthesis of sulfur containing amino acids also utilizes the sulfur carrier proteins as a sulfur source. Cysteine biosynthesis in *Mycobacterium tuberculosis* can be accomplished using a sulfur carrier protein CysO as a sulfur source [7]. The protein MoeZ activates CysO by adenylation and transfers the sulfur to the CysO thiocarboxylate. MoeZ has a rhodanese-like domain that forms a persulfide at a cysteine residue. The source of the sulfur for persulfide formation is unknown [23]. CysO thiocarboxylate is added to a *O*-phosphoserine to form a thioester adduct in a reaction mediated by cysteine synthase (CysM), a β-elimination sulfhydrylase [24, 25]. After an nitrogen-sulfur acyl shift the CysO-cysteine is complete and the cysteine must only be cleaved from the CysO C-terminal by the protease Mec^{+} [7].

REFERENCES

1. Petroski, M.D., G.S. Salvesen, and D.A. Wolf, *Urm1 couples sulfur transfer to ubiquitin-like protein function in oxidative stress.* Proc Natl Acad Sci U S A, 2011. **108**(5): p. 1749-50.
2. Maupin-Furlow, J.A., *Ubiquitin-like proteins and their roles in archaea.* Trends Microbiol, 2013. **21**(1): p. 31-8.
3. Dorrestein, P.C., et al., *The biosynthesis of the thiazole phosphate moiety of thiamin: the sulfur transfer mediated by the sulfur carrier protein ThiS.* Chem Biol, 2004. **11**(10): p. 1373-81.
4. Leimkuhler, S., M.M. Wuebbens, and K.V. Rajagopalan, *Characterization of Escherichia coli MoeB and its involvement in the activation of molybdopterin synthase for the biosynthesis of the molybdenum cofactor.* J Biol Chem, 2001. **276**(37): p. 34695-701.
5. Godert, A.M., et al., *Biosynthesis of the thioquinolobactin siderophore: an interesting variation on sulfur transfer.* J Bacteriol, 2007. **189**(7): p. 2941-4.
6. Krishnamoorthy, K. and T.P. Begley, *Reagent for the detection of protein thiocarboxylates in the bacterial proteome: lissamine rhodamine B sulfonyl azide.* J Am Chem Soc, 2010. **132**(33): p. 11608-12.
7. Burns, K.E., et al., *Reconstitution of a new cysteine biosynthetic pathway in Mycobacterium tuberculosis.* J Am Chem Soc, 2005. **127**(33): p. 11602-3.
8. Shigi, N., et al., *Common thiolation mechanism in the biosynthesis of tRNA thiouridine and sulphur-containing cofactors.* EMBO J, 2008. **27**(24): p. 3267-78.
9. Wang, F., et al., *The dual role of ubiquitin-like protein Urm1 as a protein modifier and sulfur carrier.* Protein Cell, 2011. **2**(8): p. 612-9.

10. Krishnamoorthy, K. and T.P. Begley, *Protein thiocarboxylate-dependent methionine biosynthesis in Wolinella succinogenes.* J Am Chem Soc, 2011. **133**(2): p. 379-86.

11. Lee, I. and H. Schindelin, *Structural insights into E1-catalyzed ubiquitin activation and transfer to conjugating enzymes.* Cell, 2008. **134**(2): p. 268-78.

12. Schmitz, J., et al., *The sulfurtransferase activity of Uba4 presents a link between ubiquitin-like protein conjugation and activation of sulfur carrier proteins.* Biochemistry, 2008. **47**(24): p. 6479-89.

13. Schlieker, C.D., et al., *A functional proteomics approach links the ubiquitin-related modifier Urm1 to a tRNA modification pathway.* Proc Natl Acad Sci U S A, 2008. **105**(47): p. 18255-60.

14. Shigi, N., *Posttranslational modification of cellular proteins by a ubiquitin-like protein in bacteria.* J Biol Chem, 2012. **287**(21): p. 17568-77.

15. Begley, T.P., et al., *Thiamin biosynthesis in prokaryotes.* Arch Microbiol, 1999. **171**(5): p. 293-300.

16. Lehmann, C., T.P. Begley, and S.E. Ealick, *Structure of the Escherichia coli ThiS-ThiF complex, a key component of the sulfur transfer system in thiamin biosynthesis.* Biochemistry, 2006. **45**(1): p. 11-9.

17. Lauhon, C.T. and R. Kambampati, *The iscS gene in Escherichia coli is required for the biosynthesis of 4-thiouridine, thiamin, and NAD.* J Biol Chem, 2000. **275**(26): p. 20096-103.

18. Dorrestein, P.C., et al., *The biosynthesis of the thiazole phosphate moiety of thiamin (vitamin B1): the early steps catalyzed by thiazole synthase.* J Am Chem Soc, 2004. **126**(10): p. 3091-6.

19. Rudolph, M.J., et al., *Crystal structure of molybdopterin synthase and its evolutionary relationship to ubiquitin activation.* Nat Struct Biol, 2001. **8**(1): p. 42-6.

20. Zhang, W., et al., *IscS functions as a primary sulfur-donating enzyme by interacting specifically with MoeB and MoaD in the biosynthesis of molybdopterin in Escherichia coli.* J Biol Chem, 2010. **285**(4): p. 2302-8.

21. Dahl, J.U., et al., *The identification of a novel protein involved in molybdenum cofactor biosynthesis in Escherichia coli.* J Biol Chem, 2011. **286**(41): p. 35801-12.

22. Dahl, J.U., et al., *The sulfur carrier protein TusA has a pleiotropic role in Escherichia coli that also affects molybdenum cofactor biosynthesis.* J Biol Chem, 2013. **288**(8): p. 5426-42.

23. Voss, M., M. Nimtz, and S. Leimkuhler, *Elucidation of the dual role of Mycobacterial MoeZR in molybdenum cofactor biosynthesis and cysteine biosynthesis.* PLoS One, 2011. **6**(11): p. e28170.

24. Jurgenson, C.T., et al., *Crystal structure of a sulfur carrier protein complex found in the cysteine biosynthetic pathway of Mycobacterium tuberculosis.* Biochemistry, 2008. **47**(39): p. 10354-64.

25. O'Leary, S.E., et al., *O-phospho-L-serine and the thiocarboxylated sulfur carrier protein CysO-COSH are substrates for CysM, a cysteine synthase from Mycobacterium tuberculosis.* Biochemistry, 2008. **47**(44): p. 11606-15.

CHAPTER 3: THIOCARBOXYLATE-DEPENDENT METHIONINE BIOSYNTHESIS

Section 3.1. Methionine Biosynthesis

The process of sulfur based amino acid biosynthesis is well understood in microorganisms. This is due to the importance of cysteine and methionine to a variety of cellular processes including protein synthesis, sulfur transfer, and methylation. The significance of these processes has lead to the characterization of the methionine and cysteine biosynthetic pathways. Methionine is one of two sulfur containing amino acids found in naturally occurring proteins, and is encoded by the start codon (AUG) which begins peptide synthesis. This makes it necessary for the biosynthesis of all proteins by the ribosome, hence methionine levels are necessary in abundance allowing proteins to be available for their respective functions. Methionine is also a precursor to an important molecule, *S*-adenosylmethionine (AdoMet), which serves as a methyl donor for a myriad of biochemical processes. In addition to serving as a methyl donor, AdoMet is a cofactor to the radical SAM enzymes, in which an electron is transferred from an iron-sulfur cluster to SAM to cleave it into methionine and the radical 5′-deoxyadenosine. A hydrogen is then abstracted from the substrate creating a substrate radical [1]. In mammals, the degradation of methionine is used in cysteine biosynthesis. Interestingly, in bacteria the reverse takes place as cysteine can be used as the sulfur source for methionine biosynthesis. Because of the importance of methionine in these functions, the pathways responsible for its biosynthesis are well studied and understood.

Two characterized pathways for methionine biosynthesis are known in prokaryotes (Figure 3.1). The first is the direct sulfhydrylation, which involves the addition of sulfur to *O*-acetylhomoserine 2 to form homocysteine 3. In this pathway, homoserine 1 is modified by the addition of an R group (either acetyl or succinyl group) by a homoserine transferase. The acetyl

or succinyl homoserine 2 is then modified by the addition of sulfur by the enzyme *O*-acetylhomoserine sulfhydrylase to form homocysteine 3. The second pathway is called the transsulfuration pathway and involves the addition of the amino acid cysteine 5 as the sulfur source to *O*-acetylhomoserine 2. This pathway begins when homoserine 1 is modified by the addition of a succinyl group by a homoserine transferase. Then cysteine 5 then reacts with *O*-succinylhomoserine 2 to form cystathionine 6; this reaction is mediated by cystathionine-γ synthase. Then cystathionine 6 is cleaved by the enzyme cystathionine-β lyase to yield homocysteine 3 [2]. Both pathways form homocysteine 3, and finish with the final step of methylation of the homocysteine 3 by methionine synthase to form methionine 4. Depending on the species one of these pathways is usually favored over the other [3].

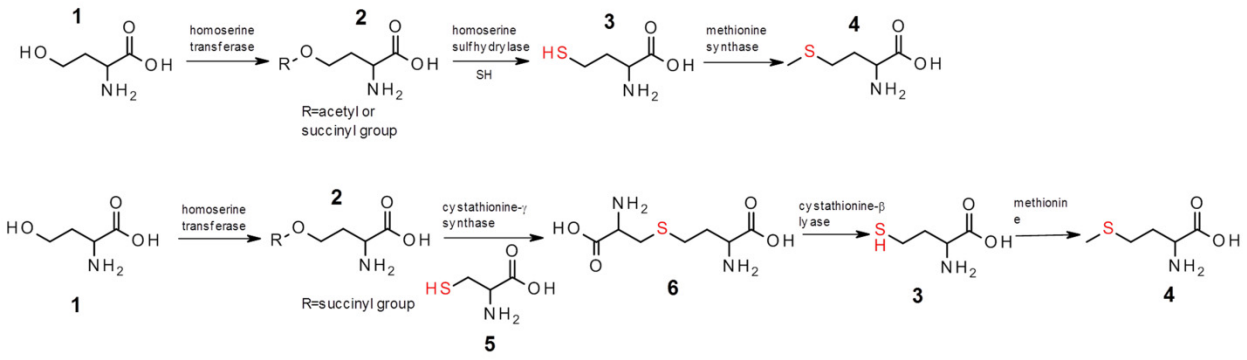

Figure 3.1. Direct sulfhydrylation pathway (Top) and the transulfuration methionine biosynthesis (Bottom).

Section 3.2. Thiocarboxylate-Dependent Methionine Biosynthesis

In *Wolinella succinogenes* the cystathionine transsulfuration pathway is absent. Instead of this transsulfuration pathway, *W. succinogenes* biosynthesizes methionine using both the direct sulhydrylation pathway and a novel pathway that depends on thiocarboxylate formation of a sulfur carrier protein as its sulfur source [4].

The thiocarboxylate-dependent methionine biosynthesis pathway was discovered in *W. succingenes* in an attempt to identify orthologs of ThiS in the SEED database (http://theseed.uchicago.edu/FIG/index.cgi). Despite low sequence identity between the ThiS and the sulfur carrier protein of this pathway (HcyS), both have the glycine-glycine C-terminal motif, characteristic of sulfur carrier proteins. However, in HcyS an alanine residue follows the Gly-Gly motif and must be removed to allow sulfur carrier protein activation. Examination of the gene cluster that contains this putative sulfur carrier protein showed genes for sulfur assimilation and methionine biosynthesis (Figure 3.2) (Table 3.1).

Figure 3.2. The gene cluster containing the sulfur carrier protein (HcyS) and proteins for sulfur carrier protein activation (blue), sulfur assimilation (red), and methionine biosynthesis (black).

Table 3.1. Functions of the proteins found in the sulfur carrier protein gene cluster.

Activation of Sulfur Carrier Protein	
HcyD	Metalloprotease that removes C-terminal alanine of HcyS. Also cleaves the homocysteine from the HcyS-Homocysteine adduct.
HcyF	ATP utilizing enzyme that adenylates HcyS C-terminal
HcyS	Sulfur Carrier Protein
Sulfur Assimilation	
CysD/CysN	Sulfate adenylyltransferase subunits 1 and 2
CysH	Phosphoadenylyl-sulfate reductase/ Adenylyl-sulfate reductase
FSR	Ferredoxin sulfite reductase reduces sulfite to sulfide and is involved in the transfer of sulfide to the HcyS thiocarboxylate
1011	Hypothetical protein probably involved in sulfate reduction
Methionine Biosynthesis	
MetY	*O*-acetylhomoserine sulfydrylase
MetE	Methyltransferase
MetZ	*O*-acetylhomoserine or *O*-succinylhomoserine sulfhydrylase
1014	Putative efflux protein

Methionine biosynthesis starts with the absorption of sulfate 7 from the environment. The sulfate 7 is adenylated by ATP sulfurylase to make adenosine 5′-phosphosulfate (APS, 8) (Figure 3.3a). APS 8 is then phosphorylated in the 5′ position by APS kinase to form 3'-Phosphoadenosine-5'-phosphosulfate (PAPS, 9). Then PAPS reductase forms sulfite 10 from PAPS 9. The PAPS reductase performs this two electron transfer through the use of the cofactor thioredoxin or glutaredoxin as an electon source. This part of the pathway is a common part of the bacterial sulfur utilization pathway (Figure 3.3a) [5]. The sulfur carrier protein pathway (Figure 3.3b) begins with the sulfur carrier protein HcyS, which has the C-terminal Gly-Gly motif and an extra C-terminal alanine residue. The first step is to remove this Ala to expose the Gly-Gly, this is accomplished by the metalloprotease HcyD, in which the two proteins form a complex and HcyD removes the Ala. The purpose of this step is unknown, but could serve a regulatory role. For example, if cleavage of alanine is the rate-limiting step of the pathway, it could provide some level of cellular control for product formation. After the Gly-Gly motif is exposed, HcyF, an ATP utilizing enzyme, adenlyates the C-terminal of HcyS to activate it for thiocarboxylate formation. The sulfite 10 supplied from sulfate reduction is then reduced by ferredoxin sulfite reductase (FSR), which is a large multiple domain protein with a sulfite reductase region and sulfurtransferase region. The sulfide is transferred to HcyS by a nucleophilic attack to form the protein thiocarboxylate. The thiocarboxylate is then used in methionine biosynthesis. MetY catalyzes the pyridoxal 5′-phophate (PLP) dependent condensation of the HcyS thiocarboxylate and *O*-acetylhomoserine (OAH, 11) to form a transient thioester adduct 12. An *S,N*-acyl shift quickly occurs to form HcyS-homocysteine 13. Then the homocysteine 3 is cleaved from HcyS by HcyD. The homocysteine 3 is then methylated by MetE using methyltetrahydrofolate as the cofactor to make methionine 4 (Figure 3.3b).

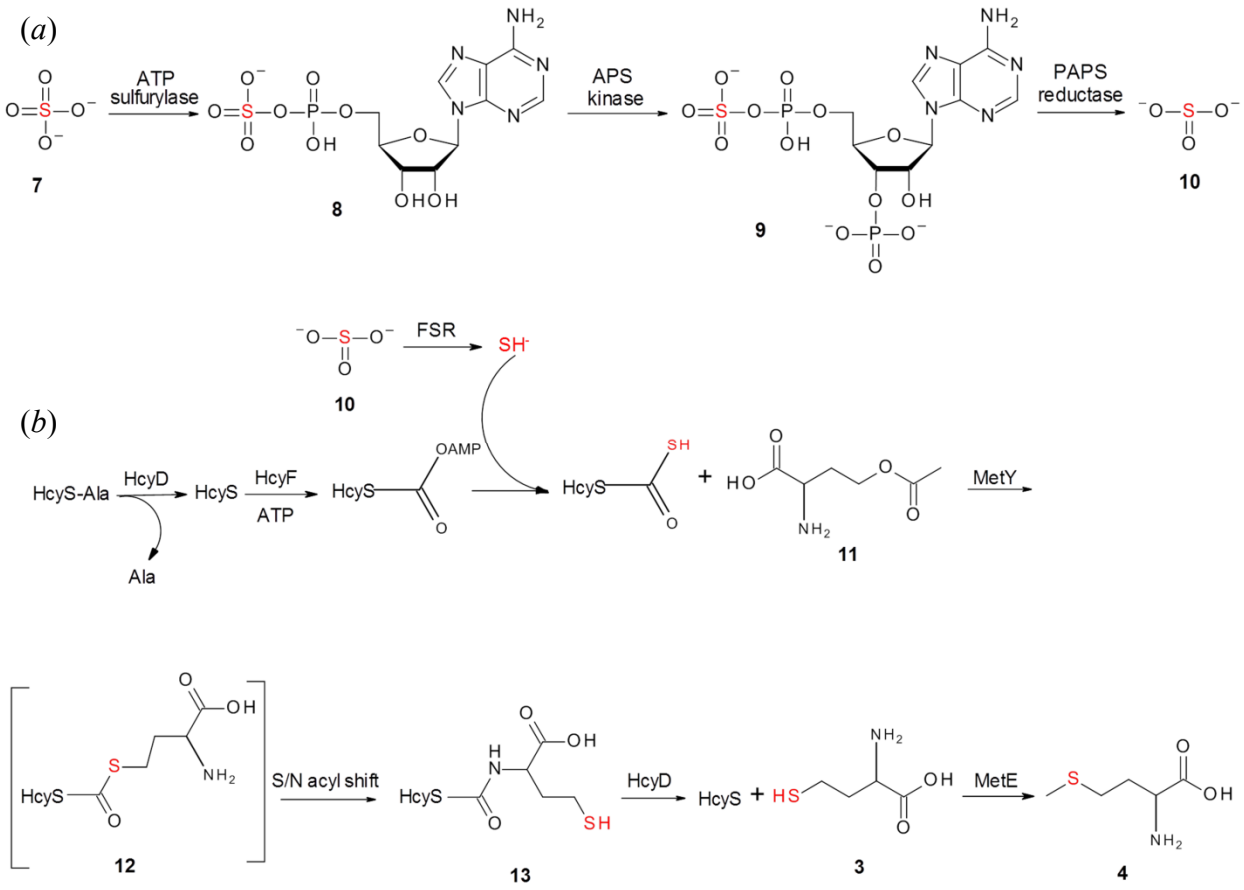

Fig 3.3. (a) Pathway for sulfate reduction to sulfite (b) Thiocarboxylate-dependent methionine biosynthesis using the sulfite incorporated from the sulfate reduction.

In the thiocarboxylate-dependent methionine biosynthesis pathway only two protein structures have been determined: FSR, which is discussed in Chapter 5, and MetY. MetY is an *O*-acetylhomoserine sulfhydrylase (OAHS) which uses the cofactor PLP to catalyze the condensation of HcyS thiocarboxylate to (OAH). The structure of MetY reveals a homotetrameric structure (Figure 3.4). Comparision of the MetY to other enzyme structures revealed that it has the same fold as a *Saccharomyces cerevisiae* cystathionine γ-lyase (PDB code 1N8P) [6] and a *Pseudomonas putida* methionine γ-lyase (PDB entry 2O7C) [7]. This information along with active site similarity suggest that MetY is in the Cys/Met metabolism PLP-dependent enzyme family. The Cys/Met family is responsible for a variety of reactions

including γ-replacements, γ-eliminations, and β-eliminations [7]. Published biochemical data [4] along with this structure lead to the identification of MetY as a γ-elimination enzyme. This is also the same enzyme type as CysM, which catalyzes the transfer of the sulfur from the CysO thiocarboxylate to *O*-phophoserine in the sulfur carrier protein-dependent cysteine biosynthesis pathway discussed in Chapter 2 [8]. The structure comparison also helped to identify the active site where PLP was modeled in using water binding in the crystal structure active site to identify key residues important for binding. The adjacent monomers that make up the tetramer have two active sites that face away from the tetramer interface. Each of the monomers contains a putative PLP binding site, though PLP was not present in the crystal structure. The active site is a large cavity that is open to solvent.

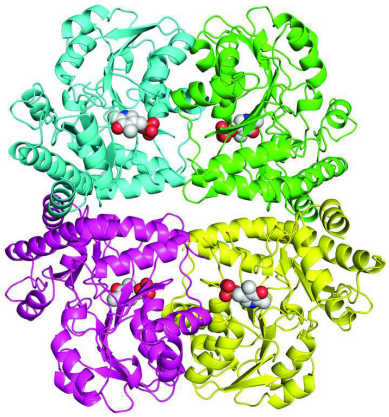

Figure 3.4. The homotetramer structure of MetY with PLP modeled into the active site as spheres (adapted from [9]).

The active site of the MetY is solvent accessible suggesting that the C-terminal of HcyS would have space to interact with the active site. The structure of HcyS and MetY in complex has not been solved so the protein-protein interactions that lead to this PLP-dependent condensation reaction are unknown. The proposed mechanism of the condensation reaction is

shown in Fig 3.5. The mechanism begins with the MetY residue Lys205 covalently bound to PLP 14. A transimination between 14 and OAH 11 forms imine 15. Then a deprotonation occurs to yield 16. A second deprotonation occurs to form the stabilized carbanion 17. The acetyl group is then eliminated from 17 to make 18. These steps are reversed as shown in molecules 19 and 20 to form molecule 21. The product is released through a transimination reaction to reform 14 and the transient HcyS-homocysteine thioester 12. Then a spontaneous S,N acyl shift occurs to form a fused HcyS-homocysteine 13 [8].

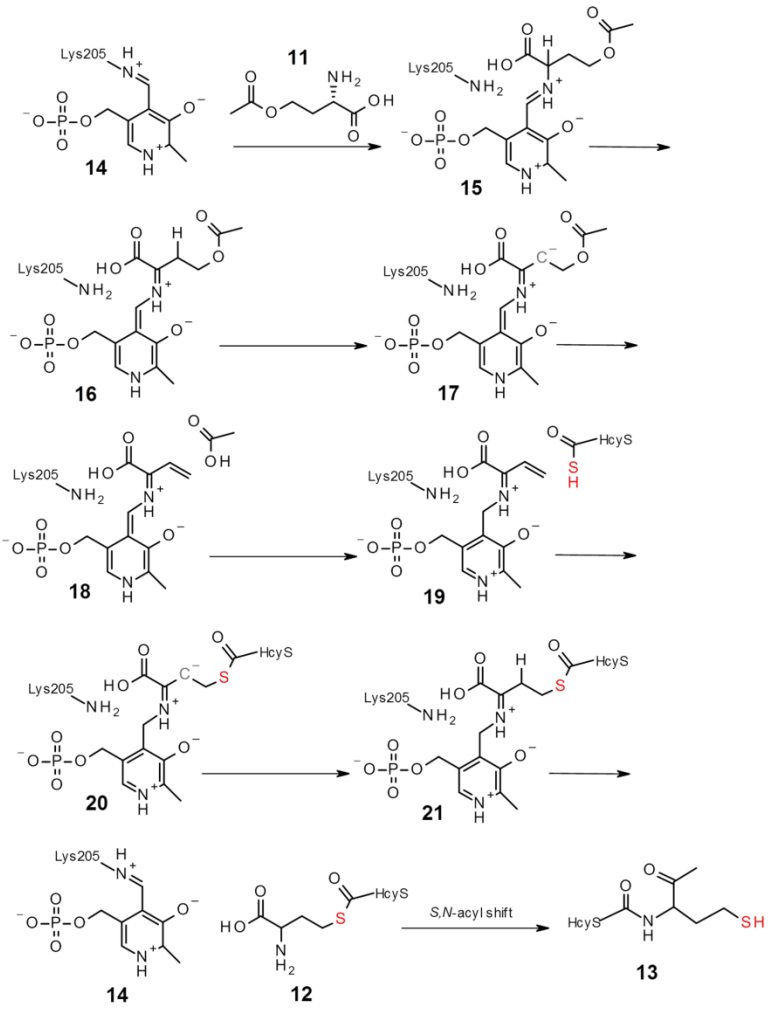

Figure 3.5. Proposed mechanism for MetY catalyzed PLP condensation (adapted from [9]).

REFERENCES

1. Layer, G., et al., *Structure and function of radical SAM enzymes.* Curr Opin Chem Biol, 2004. **8**(5): p. 468-76.
2. Mosharov, E., M.R. Cranford, and R. Banerjee, *The quantitatively important relationship between homocysteine metabolism and glutathione synthesis by the transsulfuration pathway and its regulation by redox changes.* Biochemistry, 2000. **39**(42): p. 13005-11.
3. Belfaiza, J., et al., *Direct sulfhydrylation for methionine biosynthesis in Leptospira meyeri.* J Bacteriol, 1998. **180**(2): p. 250-5.
4. Krishnamoorthy, K. and T.P. Begley, *Protein thiocarboxylate-dependent methionine biosynthesis in Wolinella succinogenes.* J Am Chem Soc, 2011. **133**(2): p. 379-86.
5. Carroll, K.S., et al., *Investigation of the iron-sulfur cluster in Mycobacterium tuberculosis APS reductase: implications for substrate binding and catalysis.* Biochemistry, 2005. **44**(44): p. 14647-57.
6. Messerschmidt, A., et al., *Determinants of enzymatic specificity in the Cys-Met-metabolism PLP-dependent enzymes family: crystal structure of cystathionine gamma-lyase from yeast and intrafamiliar structure comparison.* Biol Chem, 2003. **384**(3): p. 373-86.
7. Kudou, D., et al., *Structure of the antitumour enzyme L-methionine gamma-lyase from Pseudomonas putida at 1.8 A resolution.* J Biochem, 2007. **141**(4): p. 535-44.
8. Burns, K.E., et al., *Reconstitution of a new cysteine biosynthetic pathway in Mycobacterium tuberculosis.* J Am Chem Soc, 2005. **127**(33): p. 11602-3.
9. Tran, T.H., et al., *A novel mechanism of sulfur transfer catalyzed by O-acetylhomoserine*

sulfhydrylase in the methionine-biosynthetic pathway of Wolinella succinogenes. Acta Crystallogr D Biol Crystallogr, 2011. **67**(Pt 10): p. 831-8.

CHAPTER 4: STUDIES OF COMPLEX FORMATION OF THE SULFUR CARRIER PROTEIN HCYS

Section 4.1. Introduction

The sulfur carrier protein HcyS plays a key role in the biosynthesis of methionine is the ε-proteobacteria *Wolinella succinogenes*. Similar to other identified sulfur carrier proteins involved in biosynthesis pathways. HcyS was identified as a sulfur carrier protein when searching for ThiS homologs, and found to be clustered with genes involved in methionine biosynthesis. Like other sulfur carrier proteins, in the course of the creation of the thiocarboxylate and the biosynthesis of methionine, HcyS must interact with a variety of other enzymes.

First the C-terminal alanine residue must be removed by the metalloprotease HcyD to expose the Gly-Gly motif. Then HcyS must interact with the activation enzyme HcyF to adenylate the C-terminal. Next, the thiocarboxylate is formed by the interaction with FSR. The HcyS thiocarboxylate then undergoes a PLP-dependent condensation reaction by interaction with MetY to form the HcyS-homcysteine. Finally, the homocysteine is removed from the HcyS C-terminal by another interaction with HcyD [1]. The structural properties that allow this protein to form such a variety of interactions and complexes are not well understood and are structurally of interest.

Section 4.2. Materials and Methods

Cloning and Overexpression HcyS, HcyF, and HcyD

HcyS, HcyF, and HcyD genes were cloned into the pET-28-based vector (Novagen) with a N-terminal six-histidine tag using standard molecular biology techniques [2,3]. HcyS, HcyF, and HcyD were expressed in *E. coli* B834(DE3) cells. Overnight cultures were grown by

transferring a single colony into 30 mL of LB media supplement with 30 µL of 40 mg/mL kanamycin. LB media (1L) was inoculated with 10 mL of the overnight cultures. The cells were grown at 37 °C to an OD_{600} of 0.4. The cultures were then chilled to 15 °C and when the OD_{600} reaches 0.6, isopropylthio-β-D-thiogalactopyranoside (IPTG) was added to a final concentration of 1 mM. Shaking was continued for approximately 16 hours at 15 °C. Cells were pelleted by centrifugation at 4 °C for 30 minutes at 3300*g*. Cell pellets were stored at -20 °C until purification

Purification of HcyS, HcyD, HcyF

Cells were resuspended in 30 mL of buffer A [20 mM 2-amino-2-hydroxylmethyl-propane-1,3-diol (Tris) (pH 8.0), 333 mM sodium chloride (NaCl), and 10 mM imidazole]. Cells were lysed by sonication, and cell lysate was separated by centrifugation at 4 °C for 45 minutes at 48000*g*. Proteins were purified by nickel affinity chromatography; the clarified lysate was run over 3 mL of nickel resin pre-equilibrated with buffer A at a flow rate of 1 mL/min. The column was washed with ~20 column volumes of buffer A. The protein was eluted with three column volumes of buffer A supplemented with 250 mM imidazole. The eluted protein was further purified by size-exclusion chromatography on an ÄKTAexplorer FPLC with a HiLoad 26/60 Superdex prep-grade G200 column in buffer B [20 mM Tris (pH 8.0), 50 mM NaCl]. After the chromatography step the protein purity was determined by SDS-PAGE.

Cloning and Overexpression of HcyS-Intein

HcyS (without the C-terminal alanine) was inserted between the NdeI/SapI restriction sites in pTYB1, an intein encoding plasmid [1,2]. Overnight cultures were grown by transferring a single colony into 30 mL of LB media supplement with 30 µL of 100 mg/mL ampicillin. LB media (1L) was inoculated with 10 mL of the overnight cultures. The cells were grown in LB

(1L) media at 37 °C to an OD_{600} of 0.4. The cultures were chilled to 15 °C until the OD_{600} reaches 0.6 then the cultures were put on ice for two hours. Then IPTG was added to a final concentration of 0.4 mM. Cells were pelleted by centrifugation at 4 °C for 30 minutes at 3300 *g*. Cell pellets were stored at -20 °C until purification.

Purification of HcyS-Intein

Cells were resuspended in 30 mL of column buffer [20 mM Tris (pH 8.5), 500 mM NaCl]. Cells were lysed by sonication and the cell lysate was separated by centrifugation at 4 °C for 45 minutes at 48000 *g*. The HcyS-intein fusion protein was purified by chitin affinity chromatography. A column of approximately 20 mL of chitin resin was pre-equilibrated with ten column volumes of the column buffer. The clarified cell lysate was run over the pre-equilibrated column at a rate of 0.5 mL/min. The column was washed with approximately ten column volumes of the column buffer. The chitin resin was then quickly washed with three column volumes of the cleavage buffer [20 mM Tris (pH 8.5), 500 mM NaCl, 50 mM dithiothreitol (DTT)]. The column was then incubated in 10 mL of the cleavage buffer for 40 hours at 18 °C. The HcyS protein (without alanine) was then eluted with five column volumes of the column buffer. The eluted protein was further purified by size-exclusion chromatography on an ÄKTAexplorer FPLC with a HiLoad 26/60 Superdex prep-grade G75 column in buffer B [20 mM Tris (pH 8.0), 50 mM NaCl]. After the chromatography step the protein purity was determined by SDS-PAGE.

HcyS complex formation

The proteins HcyS and HcyD were purified as stated above and the elutant from the nickel column purification were mixed. The proteins were incubated together for three hours. The mixed proteins were then purified by size exclusion chromatography on an ÄKTAexplorer

FPLC with a HiLoad 26/60 Superdex prep-grade G75 column in buffer B [20 mM Tris (pH 8.0), 50 mM NaCl]. The peaks of the size exclusion chromatography trace were then tested for complex formation by SDS-PAGE. This procedure was repeated with the purified HcyS and HcyF, and HcyS and FSR. The HcyS/HcyD complex peak was concentrated to 15.0 mg/mL from the Bradford assay [4]. There was no peak for the HcyS/HcyF or the HcyS/FSR complex proteins concentrated separately to 13.0 mg/mL for HcyS, 14.0 mg/mL HcyF and 13.0 mg/mL for FSR from the Bradford assay [4]. All samples were flash frozen and stored at -80 °C.

Crystallization of the HcyS/HcyD complex

Frozen protein samples were thawed at 24 °C and crystallization was carried out by hanging-drop vapor-diffusion method at 22 °C using the sparse-matrix screening solutions from Hampton Research and Emerald Biosystems. Drops contain 1.5 µL of the protein sample and 1.0 µL of the well solution. Screening yielded crystals after five days in a 20% Polyethylene glycol (PEG) 8000, 0.1 M phosphate-citrate buffer pH 4.5, 0.2 M NaCl condition. The crystals were approximately 10 µm thick needles that grow in clusters (Figure 4.1). Attempts to change the precipitant and pH to increase crystal quality were unsuccessful. Crystals were tested at APS NE-CAT 24-ID-C beamline and diffracted to approximately 10 Å.

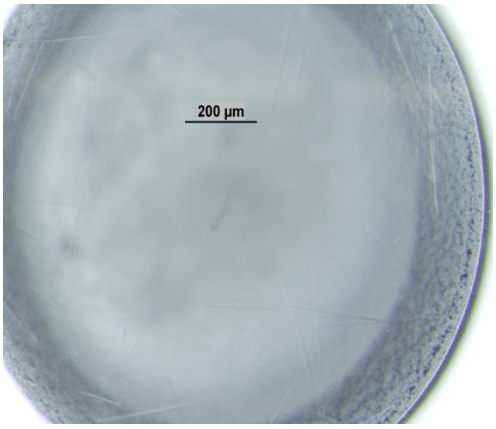

Figure 4.1. HcyS/HcyD complex needle cluster crystals

Crystallization of the HcyS-GG/FSR complex

Frozen protein samples were thawed at 24 °C and the samples were mixed in a 1 FSR: 2 HcyS-GG molar ratio and concentrated to 13.0 mg/ml. Crystallization was carried out by hanging-drop vapor-diffusion method at 22 °C using the sparse-matrix screening solutions from Hampton Research and Emerald Biosystems. Drops contain 1.5 µL of protein and 1.0 µL of the well solution. Screening yielded crystals after one day in a 1.2 M K_2HPO_4, 0.8 M NaH_2PO_4, 0.1 M acetate pH 4.5 condition. The crystals were approximately 20 µm thick needles that grow in clusters (Figure 4.2). Optimization of the condition to lower K_2HPO_4 concentration of 0.9 M produced single crystals of a thickness of approximately 30 µm. Crystals were tested at the Advance Photon Source NE-CAT 24-ID-C beamline and did not diffract.

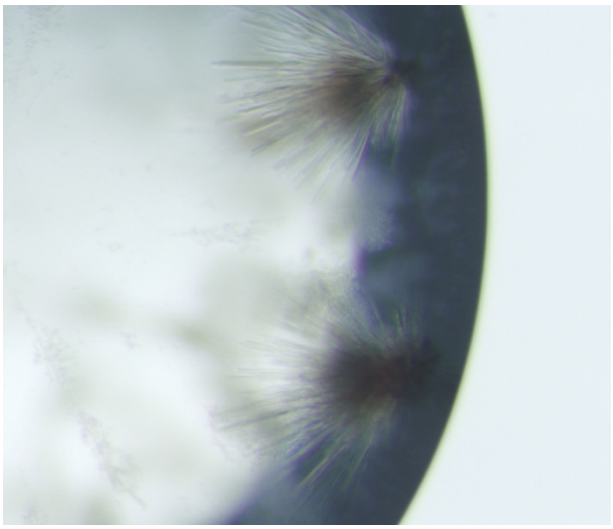

Figure 4.2. FSR/HcyS-GG complex needle cluster crystals

Production of HcyS-COAMP

The reaction mixture contained 2.5 mL of 2.0 mM HcyS-GG, 2.1 mL of 0.47 mM HcyF, 150 µL of 50 mM ATP, and 30 µL of 1M $MgCl_2$. The mixture was incubated for one hour at room temperature. The mixture was then loaded onto a Ni-column pre-equilibrated with buffer

A. The flowthrough was collected and bisulfide was added to the sample to a concentration of 2 mM. The sample was then concentrated by centrifugation to an approximate concentration of 8.0 mg/mL by Bradford assay [3]. The HcyF protein was eluted from the Ni-column with 10 mL of buffer B. Then concentrated by centrifugation to a concentration of 14 mg/mL by Bradford assay [28].

Crystallization of the FSR/HcyS-COAMP complex

FSR and HcyS-COAMP frozen samples were thawed at 24 °C and mixed at a 1:1 molar ratio. Crystallization was carried out by hanging-drop vapor-diffusion method at 22 °C using the sparse-matrix screening solutions from Hampton Research and Emerald Biosystems. Drops contain 1.5 µL of protein and 1.0 µL of the well solution. Needle cluster crystals appeared after one day in a different condition from the FSR/HcyS-GG crystals (Figure 4.3). The new condition is 30% PEG 400, 0.1 M 3-(Cyclohexylamino)-1-propanesulfonic acid (CAPS) pH 10.5.

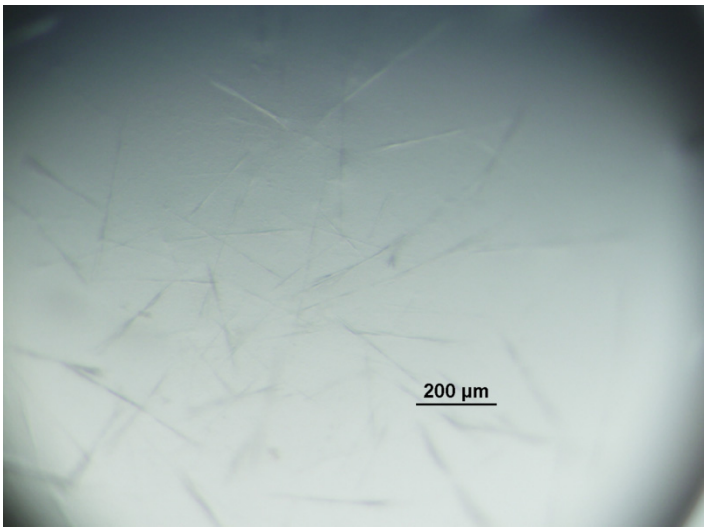

Figure 4.3. HcyS-COAMP/FSR complex crystals

Section 4.3. Results

Purification of the HcyS/HcyD complex

The size exclusion chromatography trace of the mix of HcyS and HcyD is shown in Figure 4.4. It can be seen that there are two peaks the calculated approximate molecular weights of each of these peaks is shown next to the peaks. HcyS and HcyD have a molecular weight of 7.5 kDa and 14.5 kDa respectively. SDS-PAGE analysis shows the peak 2 contains only HcyS. The SDS-PAGE results show that samples of peak 1 contain both HcyD and HcyS (Figure 4.5).

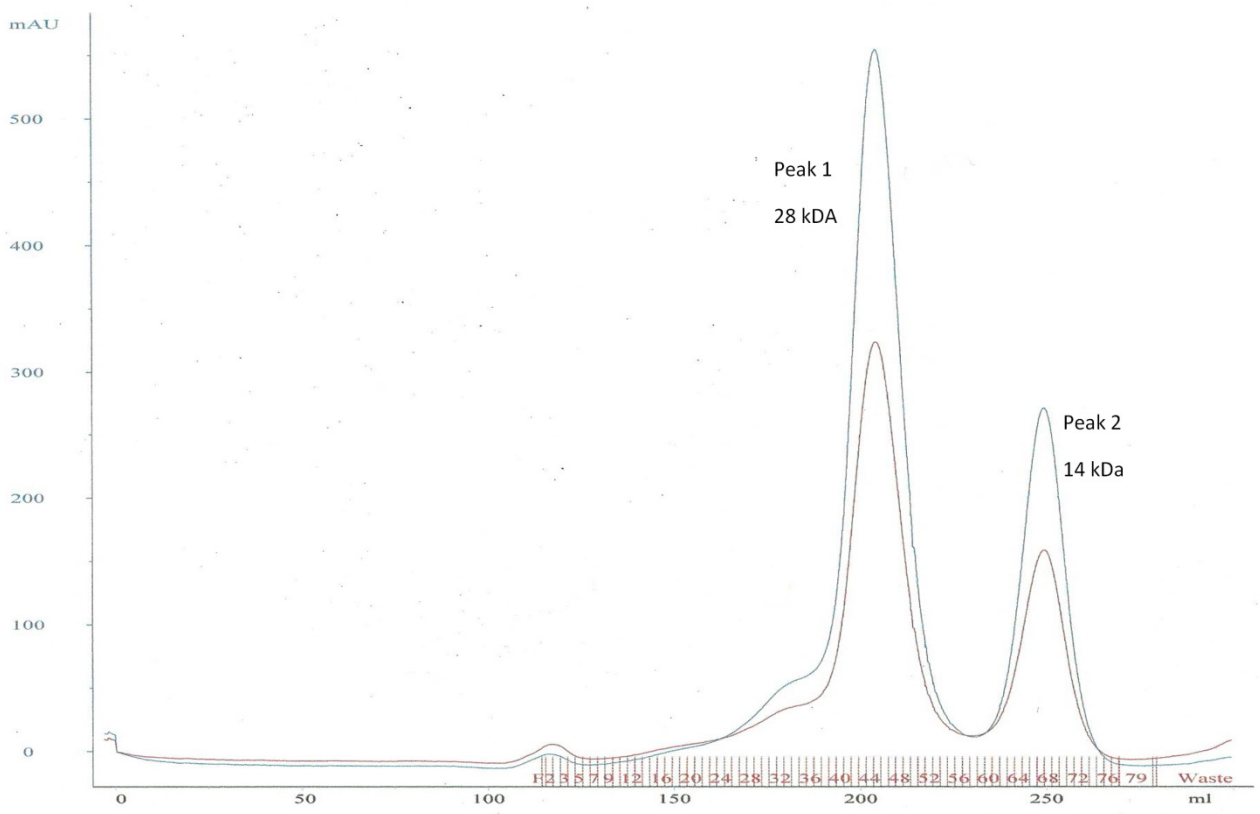

Figure 4.4. The size exclusion chromatography trace of the mix of HcyS and HcyD.

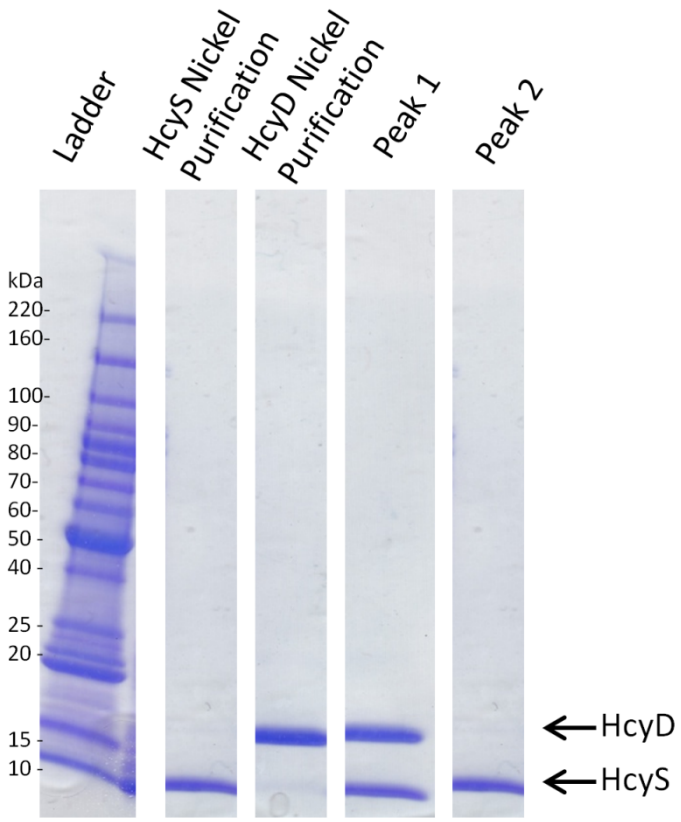

Figure 4.5. SDS-PAGE analysis of the size exclusion chromatography peaks shown in Figure 4.4

Purification of the HcyF/HcyS complex

The size exclusion chromatography trace of the mixture of HcyS and HcyF is shown in Figure 4.6. This trace also has two peaks and the calculated approximate molecular weight of each peak is shown next to each peak. HcyF has an approximate molecular weight of 30 kDa, and HcyS has a molecular weight of 7.5 kDa. The approximate molecular weight of peak 2 is approximately 14 kDa and is shown through SDS-PAGE results to contain only HcyS. Peak 1 has an approximate molecular weight of 60 kDa and SDS-PAGE analysis of the sample shows that it contains only HcyF (Figure 4.7)

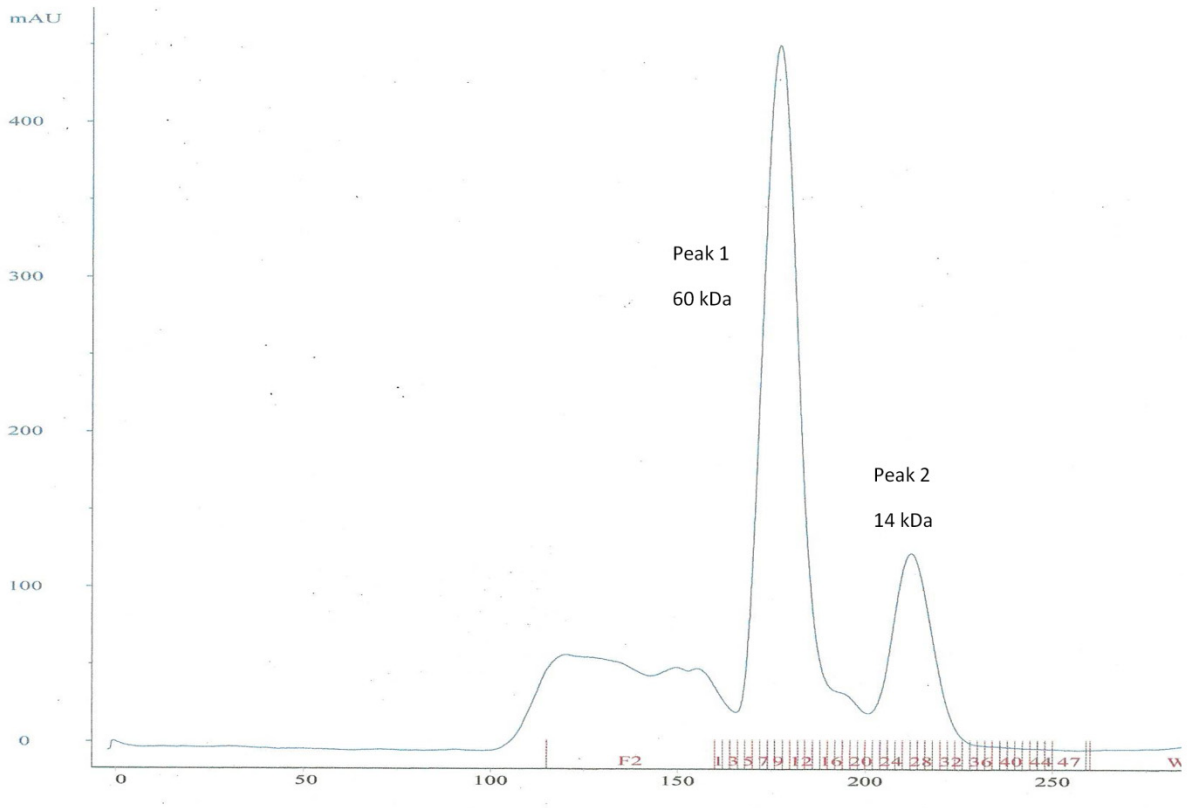

Figure 4.6. Size exclusion chromatography trace of the mix of HcyS and HcyF

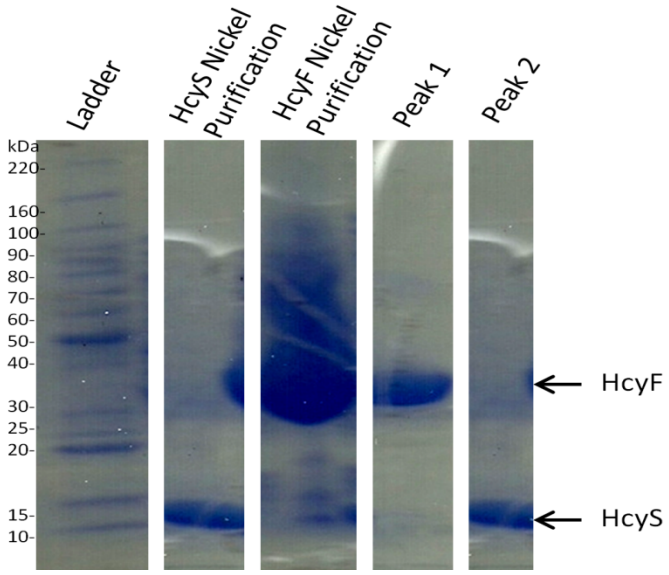

Figure 4.7. SDS-PAGE analysis of HcyF and HcyS peaks shown Figure 4.6.

Purification of the FSR/HcyS complex

The size exclusion chromatography trace of the mix of FSR and HcyS is shown in Figure 4.8. This trace has two peaks and the calculated approximate molecular weight of each peak is shown next to each peak. FSR has an approximate molecular weight of 86 kDa, and HcyS has a molecular weight of 7.5 kDa. The approximate molecular weight of peak 2 is approximately 14 kDa and is shown through SDS-PAGE to contain only HcyS. Peak 1 has an approximate molecular weight of 170 kDa and contains only FSR by SDS-PAGE analysis (Figure 4.9).

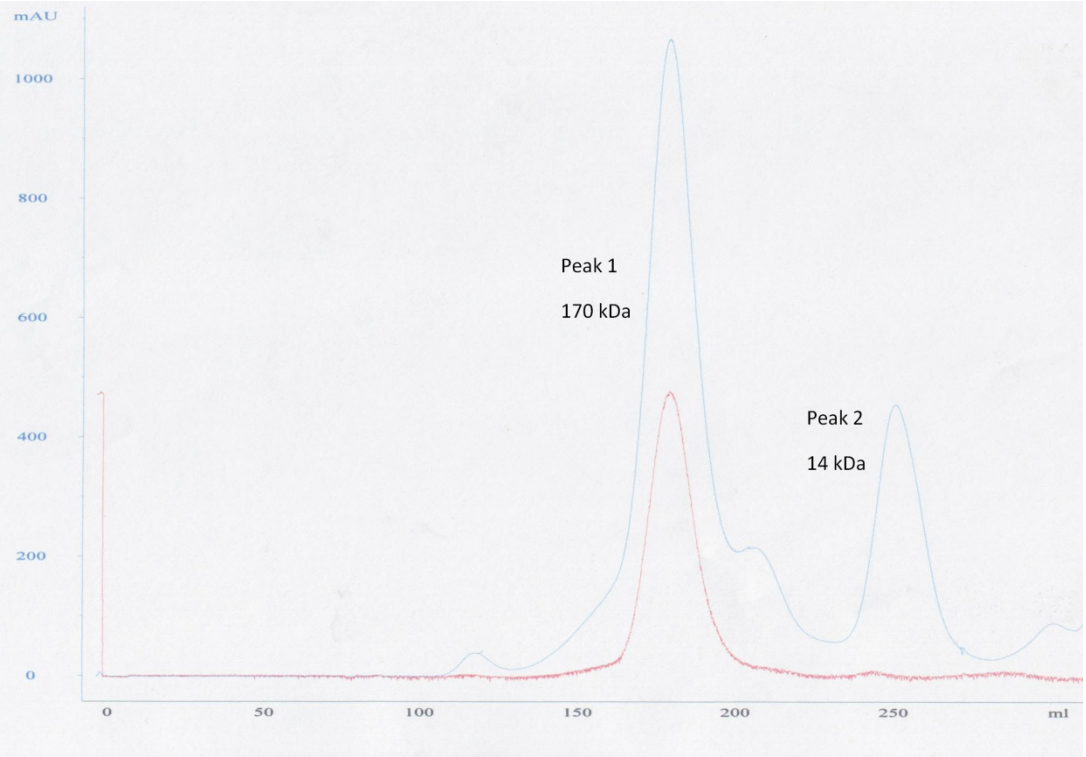

Figure 4.8. Size exclusion chromatography trace of the mix of HcyS and FSR.

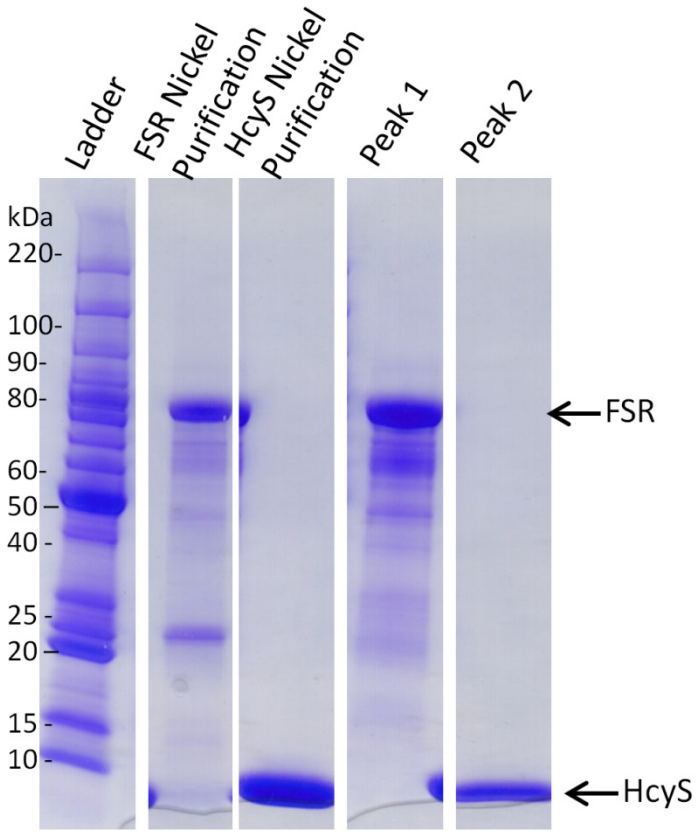

Figure 4.9. SDS-PAGE analysis of FSR and HcyS peaks shown Figure 4.8.

Section 4.4. Discussion

The approximate molecular weight of peak 2 (Figure 4.4, 4.6, and 4.8) for all of the size exclusion chromatography traces suggest that HcyS forms a dimer in solution. This is unusual when compared to the other sulfur carrier proteins. However, the size exclusion chromatography trace is not precise for such low molecular weights, so the dimer formation cannot be confirmed. In the HcyS/HcyD mix peak 1 could suggest that HcyD forms a dimer or that HcyD as a monomer forms a complex with the dimer of HcyS (Figure 4.4). SDS-PAGE analysis revealed that peak 2 contains HcyS and HcyD supporting the second possibility (Figure 4.5). These results show that HcyD and HcyS form a stable complex, which allowed crystallization. The overall oligomeric state of this complex cannot be determined from these results as the molecular weight is only estimated.

Peak 1 of Figure 4.6 with an estimated molecular weight of 60 kDa could suggest a dimer of HcyF (30 kDa) or some complicated oligomeric state of HcyS and HcyF complex formation. However, the SDS-PAGE analysis shows the peak 1 contains only HcyF suggesting HcyF dimer formation (Figure 4.7). The formation of a dimer is expected due to the structures of other activating enzymes of this type, which form a dimer. This includes ThiF and MoeB discussed in Chapter 2 [5,6]. This experiment was repeated with the HcyF and HcyS-GG, which is the state of the sulfur carrier protein that would form this complex (results not shown). The loss of the HcyS C-terminal alanine residue did not seem to have any effect on the results and there was still no evidence of complex formation. This suggests that HcyS and HcyF complex formation is not stable and is perhaps transient or that the ligands of the reaction ATP and Mg are needed to make the complex stable.

The FSR and HcyS mix peak 1 has an approximate molecular weight of 170 kDa, which is the same as the molecular weight of the FSR dimer seen at pH 8 and discussed in Chapter 5 (Figure 4.8). Because of the large difference in molecular weight between the FSR and HcyS dimers it can be difficult to detect complex formation by molecular weight. However, SDS-PAGE analysis of the peaks shown in Figure 4.9 shows that peak 1 contains only FSR showing no complex formation. If FSR and HcyS do form a complex it is likely not stable. The HcyS-GG and FSR mix was analyzed (results not shown), and the results were the same as shown in Figure 4.8 and Figure 4.9. The HcyS-COAMP is the state of the sulfur carrier protein that would possibly interact with FSR. This mixture was not examined with size exclusion chromatography analysis because of the instability of the adenylated HcyS. It is still unknown whether HcyS and FSR form a complex or if another sulfur transfer protein accepts the sulfur from FSR and then transfers it to the HcyS C-terminal.

REFERENCES

1. Krishnamoorthy, K. and T.P. Begley, *Protein thiocarboxylate-dependent methionine biosynthesis in Wolinella succinogenes.* J Am Chem Soc, 2011. **133**(2): p. 379-86.
2. Ausubel, F.M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl. Short protocols in molecular biology. 2002, New York, NY: John Wiley & Sons.
3. Maniatis, T., E. F. Fritsch, and J. Sambrook. Molecular cloning, a laboratory manual. 1982, Cold Spring, NY: Cold Spring Harbor Laboratory
4. 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, 1976. **72**: p. 248-54.
5. Lehmann, C., T.P. Begley, and S.E. Ealick, *Structure of the Escherichia coli ThiS-ThiF complex, a key component of the sulfur transfer system in thiamin biosynthesis.* Biochemistry, 2006. **45**(1): p. 11-9
6. Rudolph, M.J., et al., *Crystal structure of molybdopterin synthase and its evolutionary relationship to ubiquitin activation.* Nat Struct Biol, 2001. **8**(1): p. 42-6

CHAPTER 5: STRUCTURAL STUDIES OF FERREDOXIN SULFITE REDUCTASE

Section 5.1. Introduction

Compounds containing sulfur are distributed throughout a variety of natural products including amino acids, carbohydrates, nucleic acids, cofactors [1]. The synthesis of these compounds draws from one of three sulfur sources in bacteria: free sulfide (cysteine biosynthesis), protein persulfides (iron sulfur cluster biosynthesis), and protein thiocarboxylates [2]. Protein thiocarboxylates or sulfur carrier proteins are part of a family of sulfide donors that are involved in a variety of biosynthetic pathways. This family includes ThiS (vitamin B_1) [3], CysO (cysteine) [4], and MoaD (molybdopterin) [5]. Sulfur carrier proteins are generally of low molecular weight less than 10,000 Da with a Gly-Gly C-terminal motif and an ubiquitin-like fold. In some cases the Gly-Gly must be revealed by cleaving one or more amino acids present after the Gly-Gly motif.

A search for ThiS orthologs uncovered a new sulfur carrier protein gene in *Wolinella succinogenes*, which is a Gram-positive bacterium and a member of the Heliobacteraceae group of ε-protobacterium. This group of bacterium contains several pathogens of humans and animals including *Heliobacter pylori* and *Campylobacter jejuni* [6]. Bioinformatics analysis of this *W. succinogenes* putative sulfur carrier protein gene showed that it was clustered with genes for methionine biosynthesis and sulfur assimilation. In this gene cluster, ferredoxin sulfite reductase (FSR) is responsible for the reduction of sulfite to sulfide, which is then transferred to the sulfur carrier protein in the thiocarboxylate-dependent methionine biosynthesis pathway (Figure 5.1). Other proteins identified in this pathway are a sulfur carrier protein (HcyS), a metalloprotease (HcyD), an adenylating enzyme (HcyF), and several proteins for methionine biosynthesis. The pathway begins when HcyS with a C-terminal alanine residue (HcyS-Ala) is cleaved by the

protease HcyD to reveal the C-terminal Gly-Gly motif (HcyS-GlyGly) that is characteristic of thiocarboxylate-dependent sulfur carrier proteins. HcyS-GlyGly is then adenylated by HcyF to yield HcyS-COAMP. The sulfite 10 is reduced to sulfide and then transferred by FSR to HcyS to replace the AMP making the thiocarboxylate HcyS-COSH [7]. Then OAH 11 is added to HcyS-COSH through a condensation reaction is catalyzed by MetY to form 12 [7, 8]. An S,N-acyl shift occurs to form HcyS-homocysteine 13. The homocysteine 3 is then cleaved from HcyS by HcyD. Finally, methylation of homocysteine 3 by MetE using 5-methyltetrahydrofolate as a cofactor finishes the methionine 4 biosynthesis [7].

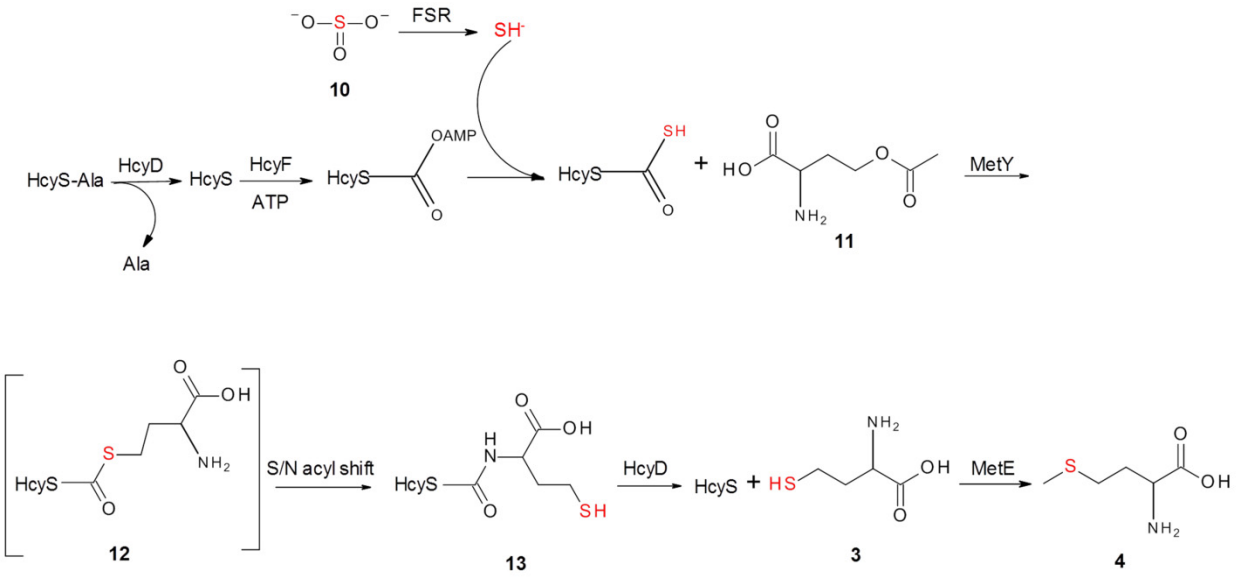

Figure 5.1. Thiocarboxylate-dependent methionine biosynthesis pathway

The enzyme that provides the sulfur source for this pathway, FSR, is a sulfite reductase. Several structures of these sulfite or the very similar nitrite reductases have been solved and the mechanism of this reduction has been characterized [9]. This characterization is due to the necessity of the reduction of sulfite and nitrite to sulfide and ammonia respectively for the biological utilization of sulfur and nitrogen. Sulfite reductases (SiRs) and nitrite reductases (NiRs) use a novel prosthetic group assembly to perform a six-electron reduction. The active site

centers contain a reduced prophyrin of the isobacteriochlorin class called a siroheme that is covalently linked by a cysteine residue to a Fe_4S_4 cluster. This forms an electronically integrated metallo-cofactor for delivering electrons to the substrate. SiRs and NiRs are grouped into two classes assimilatory and dissimilatory. Assimilatory sulfite reductases (aSiRs) are found in bacteria, fungi, and plants, but not in animals and generate sulfide for incorporation into sulfur-contains amino acids and cofactors. Dissimilatory sulfite reductases (dSiRs) are found primarily in diverse sulfate-reducing eubacteria, and thermophilic archaebacteria. The bacteria couples the reduction of sulfite with the terminal redox of the electron transport chain. The dSiRs can be distinguished from aSiRs by the tendency to produce incompletely reduced sulfur in the forms of trithionate ($S_3O_6^{2-}$) and thiosulfate ($S_2O_3^{2-}$) [10].

The structures of sulfite reductases have been solved from several organisms, including *Escherichia coli* (1AOP) [11], *Spinacia oleracea* (2AKJ) [12], *Nicotiana tabacum* (3B0L) [13], and *Mycobacterium tuberculosis* (1ZJ8) [14]. In *E. coli* SiR is part of a large protein complex including two parts referred to as the flavoprotein subunit and the hemoprotein subunit. The holoenzyme is made up of eight 66 kDa flavoprotein subunits, which accept electrons from NADPH and four 64 kDa hemoprotein subunits which accept electrons from the flavoprotein subunit and catalyzes the reduction of sulfite. However, the hemoprotein subunit is capable of reducing sulfite to sulfide, nitrite to ammonia, and hydroxylamine to ammonia without the flavoprotein being present as long as an electron source is provided [11]. Other SiRs are structurally similar to this hemoprotein subunit, with an r.m.s.d. of approximately 2.0 Å with three domains folding around the prosthetic groups in the center (Figure 5.2). Biologically, many of these SiRs use the protein ferredoxin as a source of electrons. Like the *E.coli* hemoprotein the

ferredoxin SiRs can catalyze the sulfite reduction without ferredoxin if another electron source is provided [14].

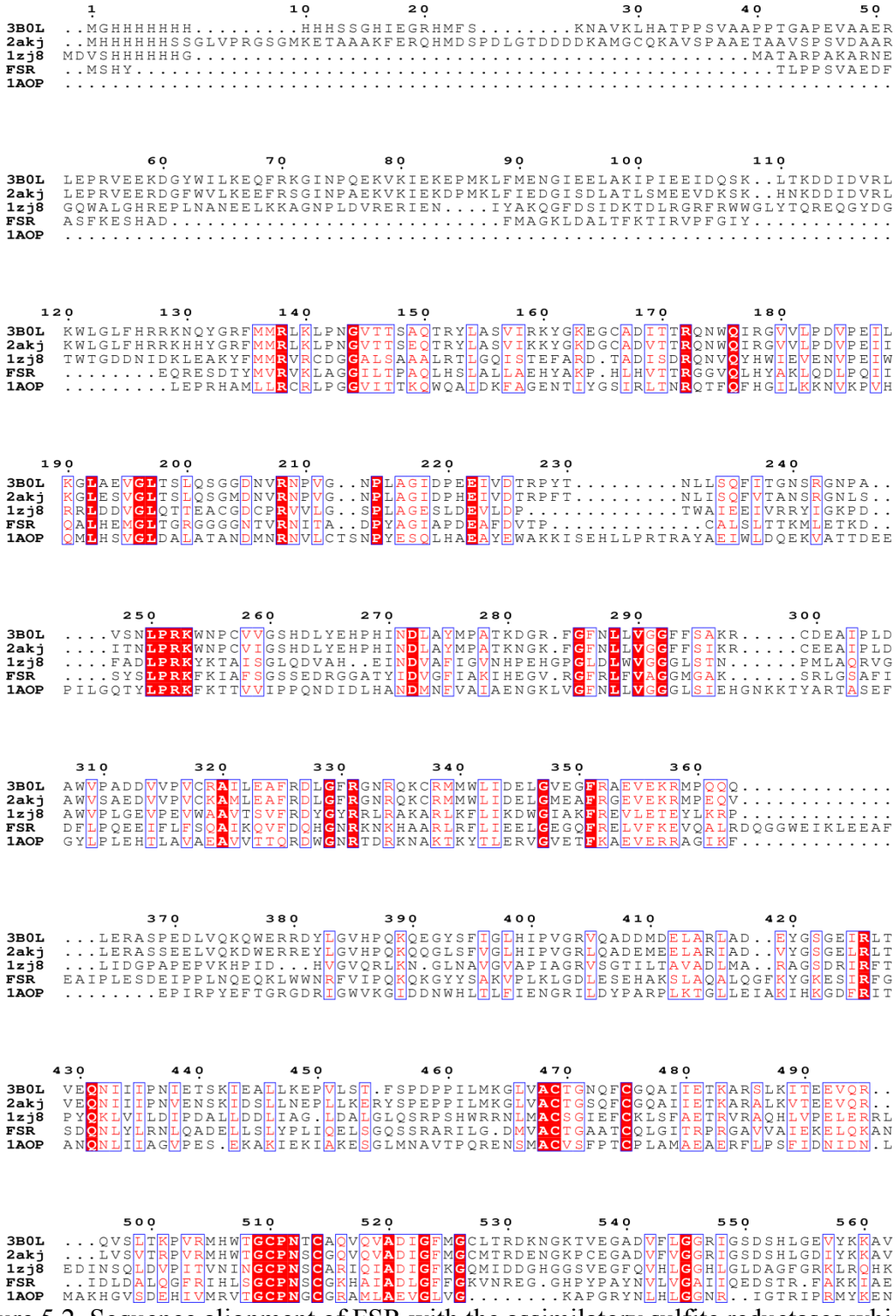

Figure 5.2. Sequence alignment of FSR with the assimilatory sulfite reductases which have structure deposited in the PDB database [11, 12, 13, 14]

The mechanism of the reduction has been well studied in the *E.coli* hemoprotein, but due to a high level of structural similarity to the ferredoxin SiRs, the same mechanism is associated with both. The overall reaction is a six electron transfer and is shown in Figure 5.3 and the mechanism is shown in Figure 5.4.

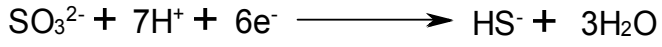

Figure 5.3. Overall reaction of the sulfite reduction to sulfide

The Fe_4S_4 cluster is in an oxidized overall 2+ state to start, the iron of the siroheme is also oxidized at an Fe (III) state 22 (Figure 5.4). The mechanism starts with a transfer of two electrons to the Fe_4S_4 cluster, one of which is transferred through the cysteine that covalently links the cluster to the siroheme iron. This changes the state of the cluster to 1+ and the siroheme iron to Fe (II) 23 (Figure 5.4). After this reduction of the cofactors, the substrate, sulfite, binds to the Fe (II) of the siroheme 24 (Figue 5.4). The basic residues above the siroheme stabilize the binding of the sulfite by forming hydrogen bonds with the oxygen atoms of the substrate. The electrons held by the Fe_4S_4 cluster and the siroheme iron are transferred to the sulfite, which destabilizes one of the oxygen sulfur bonds leading it to break 25 (Figure 5.4). Protons from the basic residues are donated to make H_2O from the released oxygen, and the cofactors are returned to the cluster 2+ and siroheme Fe(III) states. This process is repeated twice as four electrons are transferred through the cofactor to the sulfite, the other two oxygens are released and form two H_2O molecules 26 and 27 (Figure 5.4). The sulfide is then released from the cofactors, leaving them in the original $[Fe_4S_4]^{2+}$ and Fe (III) state 22 (Figure 5.4) [9].

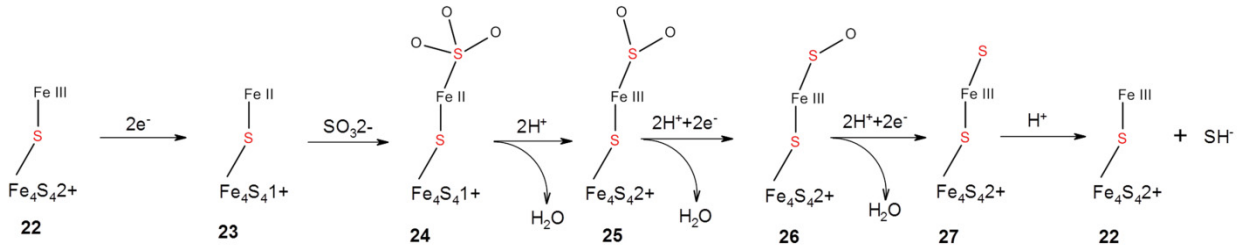

Figure 5.4. Mechanism of sulfite reductase, the Fe_4S_4 represents the iron sulfur cluster and the Fe III is the iron of the siroheme. The sulfur that is the covalent link between the two is from a protein cysteine residue.

The sulfite reductase of the thiocarboxylate-dependent methionine biosynthesis pathway, FSR has 764 residues making it larger than other sulfite reductases, which have approximately 550 residues. The longer sequence of FSR suggests that it might be a fused multiple domain protein (Figure 5.5). The last 90 residues of FSR have sequence similarity to a sulfurtransferase, TusA [15].

Figure 5.5. FSR sequence diagram divided into domains identified by sequence alignments

TusA is the first protein of a sulfur relay pathway that involves the transfer of sulfur to the wobble uridine of several tRNAs (Figure 5.6). The first position of the anticodon or the wobble position participates in the identification of the genetic code mediated by the codon-anitcodon interaction. The uridine of the wobble position can receive a thio modification to make 5-methylaminomethyl-2-thiouridine (mnm^5s^2U), which is important for specific recognition by

glutaminyl-tRNA synthetase [15]. A cysteine sulfur 5 is transferred to an IscS active site cysteine to make a persulfide through a PLP dependent reaction. TusA accepts the sulfur of IscS persulfide at its own active site cysteine residue [16]. The sulfide is then transferred to a cysteine residue of TusD, which is part of a complex with TusB and TusC. This transfer from the TusA persulfide to the TusD cysteine is stimulated by the interaction of TusE with the TusBCD complex. Then TusE accepts the sulfide from TusD at a cysteine residue forming a persulfide, which is then transferred to the protein MnmA and finally to tRNA to form the mnm^{5}s^{2}U 26 [15].

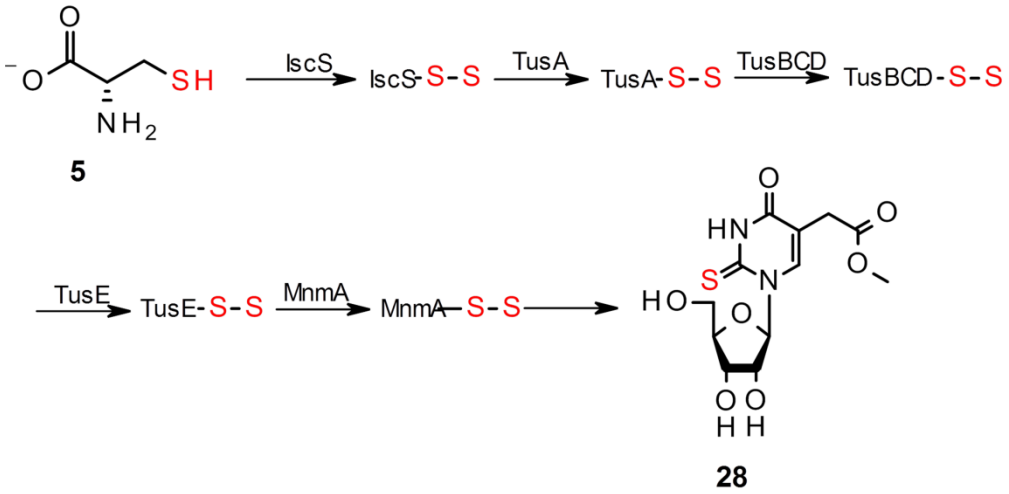

Figure 5.6. The sulfur transfer pathway to the tRNA, mnm^{5}s^{2}U that includes the sulfurtransferase TusA

In addition to this function, studies have also suggested that TusA might be involved in the distribution of sulfur by the enzyme cysteine desulfurase IsCS. A part of this distribution is to the sulfur transfer pathway of the molybdenum cofactor (Moco) biosynthesis [17]. The biosynthesis of Moco is the final step of the molybdopterin (MPT) biosynthesis pathway discussed in Chapter 2, and involves the transfer of sulfur by a sulfur carrier protein (MoaD) [18]. MoaD is very similar to HcyS, the sulfur carrier protein of the methionine biosynthesis

pathway of which FSR is a part. They have a sequence identity of 38 %, and both have the Gly-Gly motif that forms the thiocarboxylate for sulfur transfer [7, 18]. According to TusA knockout studies, without TusA, IscS favors sulfur transfer to enzymes of iron-sulfur cluster biosynthesis. In these knockouts Moco biosynthesis is still accomplished but at a decreased efficiency. It seems that without TusA the rhodanese-like sulfur transferase YnjE preferentially accepts the sulfur from IscS and then transfers it to MoaD/MoaE complex that is called MPT synthase [17]. It had been shown in earlier studies that YnjE is capable of the transfer of sulfur to MPT synthase [19]. However, it is shown by the decrease of the Moco biosynthesis efficiency in the TusA knockout that TusA is somehow involved in the sulfur transfer to MPT synthase. It is currently unknown if it acts as YnjE accepting the sulfur from IscS and directly transferring it to MPT synthase, or if it transfers it to another sulfurtransferase to then transfer to MPT synthase [17].

In this study we report a 2.4 Å resolution structure of FSR, the first documented structure of a sulfite reductase that is fused to a domain that has the same fold as HEPN domains and a disordered sulfur transferase domain. It transfers the sulfide to form a thiocarboxylate of a sulfur carrier protein. Comparing this structure to other proteins, we found that FSR residues 1-560 has the same fold of other sulfite/nitrite reductases, residues 580-670 has the same fold as HEPN domains. The structural similarities seen in this study and the published biochemical data supports the hypothesis that FSR reduces sulfite to sulfide and participates in the transfer of the sulfide to the sulfur carrier protein HcyS.

Section 5.2. Materials and Methods

Cloning and Overexpression of Selenomethioninyl (SeMet) FSR

FSR and FSR N-terminal domain were cloned into a pET-28-based vector (Novagen)

with an N-terminal six histidine tag using standard molecular biology techniques [20, 21]. The FSR vector, *Salmonella typhimurium* cysG (siroheme synthase) vector, and *Azotobacter vinelandii* IscS cluster vector were coexpressed in *Escherichia coli* B834 (DE3) cells. An overnight culture was grown by the transfer of a colony to 30 mL of LB media supplemented with 30 μL of 40 mg/mL kanamycin, 30 μL of 100 mg/mL ampicillin, and 30 μL of 40 mg/mL chloramphenicol at 37 °C. The overnight cultures were pelleted at 4 °C by centrifugation for 10 minutes at 4000*g*. The resulting cell pellet was resuspended in 10 mL of M9 minimal salts. The 10 mL resuspended pellet was used to inoculate M9 minimal media (1L) supplemented with 10 mL 40% glucose, 2 mL of 1M $MgSO_4$, 1 mL of 0.1 M $CaCl_2$, 10 mL of MEM vitamin solution, 50 mg L-selenomethionine (SeMet), 100 mg of ampicillin, and 40 mg each of kanamycin and chloramphenicol. The 1 L cultures were grown at 37 °C to an OD_{600} of 0.4. Then 2.5 g of arabinose, 58.7 g of ferrous ammonium sulfate, and 90 mg of L-cysteine were added and the cultures was cooled to 15 °C while continuing to shake at 180 rpm. When the OD_{600} reached 0.6, 30 mg of aminolevulinic acid and IPTG to a final concentration of 1 mM were added and shaking continued at 15 °C for 12-16 hours. Native FSR cells were grown in the same way except with media supplemented with 80 mg of DL-methionine instead of L-selenomethionine. Cells were pelleted by centrifugation at 4 °C for 30 minutes at 3300*g*. Cell pellets were stored at -20 °C until purification.

Purification

Cells were resuspended in 30 mL of buffer A [20 mM Tris pH 8.0, 333 mM NaCl, 10 mM Imidazole, and 1 mM DTT]. Cells were lysed by sonication, and cell lysate was separated by centrifugation at 4 °C for 45 minutes at 48000*g*. FSR was purified from the cell lysate by nickel affinity chromatography. The clarified lysate was incubated with 3 mL of nickel resin for

one hour; the mixture was then filtered through a column. The nickel resin was then washed with 20 column volumes of buffer A. The protein was eluted with 3 column volumes of buffer A supplemented with 250 mM imidazole. The eluted protein was further purified by size-exclusion chromatography on an ÄKTAexplorer FPLC with a HiLoad 26/60 Superdex prep-grade G200 column with buffer B [20 mM Tris (pH 8), 50 mM NaCl, 1 mM DTT]. The protein purity was determined by SDS-PAGE after the chromatography step. The SeMet protein was then concentrated to 20 mg/ml as determined by Bradford assay [22] using centrifugal concentrator and aliquots were flash-frozen and stored at -80 °C. The native FSR was concentrated to 14 mg/ml as determined by a Bradford assay [22].

Crystallization SeMet full-length FSR

Frozen protein was thawed at 24 °C, and crystallization of SeMet FSR was carried out by hanging-drop vapor-diffusion method at 22 °C using sparse-matrix screening solutions from Hampton Research and Emerald BioSystems. Drops contained 1.5 µL of protein solution and 1.0 µL of well solution. Screenings yielded crystals that appeared after approximately one month in a 20% PEG 8000, 0.1 M 4-(2-Hydroxyethyl)piperazine-1-ethanesulfonic acid (HEPES) pH 7.5 condition. The addition of 0.2 M potassium chloride to the condition grew higher quality crystals.

Crystallization of FSR N-terminal domain

Frozen protein was thawed at 24 °C and crystallization of FSR N-terminal domain was carried out by hanging-drop vapor-diffusion method at 22 °C using sparse-matrix screening solutions from Hampton Research and Emerald BioSystems. Drops contained 1.5 µL of protein solution and 1.0 µL of well solution. Screening yielded crystals in the same condition as SeMet FSR with the same crystal morphology after one day. However, a 35% 2-methyl-2,4-pentanediol

(MPD), 0.2 M Li_2SO_4, 0.1 M 2-(N-Morpholino)ethanesulfonic acid (MES) pH 6.0 condition yielded higher quality crystals.

Crystallization FSR co-crystallized with sulfide

Crystallization of native FSR was carried out by hanging-drop vapor-diffusion method at 22 °C using sparse-matrix screening solution from Hampton Research and Emerald BioSystems. Sodium sulfide was added to the protein sample to a final concentration of 2 μM and let incubate for 15 minutes before setting up the screens. Screening yielded crystals that appeared after one day in a 1.2 M sodium phosphate monobasic, 0.8 M potassium phosphate dibasic, 0.2 M Li_2SO_4, and 0.1 M CAPS pH 10.5 condition.

Data Collection

The crystals were dipped into a solution consisting of 20% PEG 8000, 0.2 M potassium chloride, 0.1 M HEPES pH 7.5 and 20% glycerol were flash-frozen in liquid nitrogen. Data was collected on a single SeMet crystal at the NE-CAT beamline 24-ID-C at the Advanced Photon Source (Argonne National Laboratory, Argonne, IL) using a Quantum 315 detector. To maximize the anomalous signal, a single-wavelength anomalous diffraction experiment was conducted at peak f for selenium. Two datasets were collected on this SeMet FSR crystal to 3.0 Å resolution using a 1° oscillation range over 120 frames for each dataset. These two datasets were indexed, integrated, scaled and merged using the HKL2000 program suite [23]. Data-collection statistics are shown in Table 5.1.

FSR N-terminal domain crystals were not dipped in a cryogenic solution and were taken from the drop solution, 35% MPD, 0.2 M Li_2SO_4, 0.1 M MES pH 6.0. The data were collected at the Cornell High Energy Synchrotron Source (CHESS) A1 beamline. One data set was collected from this crystal to 2.4 Å using a 1° oscillation range for 180 frames. Data were indexed,

integrated, and scaled using the HKL2000 program suite [23]. Data-collection statistics are shown in Table 5.1.

FSR sulfide co-crystallization crystals were dipped into a solution consisting of 0.8 M NaH_2PO_4, 0.8 M potassium K_2HPO_4, 0.2 M Li_2SO_4, 0.1 M CAPS pH 10.5, 10% glycerol, and 10% ethylene glycol. Data was collected at the NE-CAT beamline 24-ID-C at the Advanced Photon Source. One data set was collected from this crystal to 2.4 Å using a 1° oscillation range over 120 frames. Data was indexed, integrated, and scaled using the HKL2000 program suite [23]. Data collection statistics are shown in Table 5.1.

Table 5.1. Summary of data collection statistics

	SeMet FSR N-terminal	Native FSR N-terminal	FSR 2 mM Na_2S
Beamline	APS NE-CAT 24-ID-C	CHESS A1	APS NE-CAT 24-ID-C
Wavelength (Å)	0.97918	0.97590	0.97910
Space group	$P2_12_12_1$	$P2_12_12_1$	$P4_12_12_1$
a (Å)	70.0	71.5	89.0
b (Å)	127.4	127.1	89.0
c (Å)	167.2	170.3	258.4
Chains per asymmetric unit	2	2	1
Resolution (Å)	3.1	2.4	2.4
Total number of reflections	282631	322561	294868
Number of unique reflections[a]	57259(5726)[b]	60702(2913)[b]	41932(2060)[b]
Redundancy	9.2 (8.8)[b]	6.0 (5.7)[b]	7.0 (6.9)[b]
R_{sym}[c](%)	10.8 (71.1)[b]	7.0 (45.2)[b]	6.7 (47.6)[b]
I/σ	21.9 (2.6)[b]	30.6(4.1)[b]	21.2 (3.3)[b]
Completeness (%)	99.4(99.4)[b]	98.4 (96.1)[b]	99.8 (100.0)[b]

[a]Unique reflections include Bijvoet pairs. [b]Values in parentheses refer to the highest resolution shell. $R_{sym} = \Sigma\Sigma_i|I_i - (I)|/\Sigma(I)$, where (I) is the mean intensity of N reflections with intensities I_i and common indices h, k, l.

Structure Determination

FSR contains 13 selenomethionine residues and the crystallographic asymmetric unit contains two monomers for a total of 26 selenomethionine residues. A total of 12 Se positions were located using hkl2map. Initial electron density maps were calculated using SAD phasing

and the position of the Se atom. A model was placed in the SeMet electron density map using MOLREP [24] real space molecular replacement using *Spinacia oleracea* nitrite reductase structure (PDB entry 2AKJ) as model [12]. *CHAINSAW* [25] was used to remove atoms that were not common to the two sequences. The FSR N-terminal domain data were used to extend the resolution of this structure to 2.4 Å resolution. The siroheme, Fe_4S_4 cluster, and 254 waters were added to the structure during refinement. Refinement was performed using *PHENIX.REFINE* [26] and alternated with successive manual model building using *COOT* [27] guided by F_o - F_c and $2F_o$ - F_c maps. The geometry of FSR was validated using *PROCHECK* [28]. Refinement statistics are shown in Table 3.2.

The FSR/sulfide complex structure was determined using molecular replacement as implemented in the program PHENIX.PHASER [26] with the SeMet FSR structure as the search model. The model corresponding to residues 570-673 was built with PHENIX.AUTOBUILD [26]. Refinement was performed using PHENIX.REFINE [26] and alternated with manual model building using *COOT* [27], and 306 water molecules were added to the structure. The geometry of FSR was validated using *PROCHECK* [28]. Refinement statistics are shown in Table 5.2.

Figure Production

All figures were made with PyMol [29], ESCript 2.2, and Accelrys Draw [Accelrys].

Table 5.2 Summary of data refinement statistics

	FSR N-terminal Residues (1-560)	FSR Residues (4-673)
Resolution (Å)	2.4	2.4
No. of protein atoms	8600	5036
No of ligand atoms	193	96
No. of water atoms	254	308
Reflections in the working set	59150	41859
Reflections in the test set	3007	1543
R factor[a] (%)	18.5	15.4
R_{free} (%)	24.1	19.2
R.m.s.d from ideals		
Bond lengths (Å)	0.013	0.010
Angles (Å)	1.52	1.42
Average Bfactor ($Å^3$)	37.4	36.37
Favored (%)	97.4	97.3
Allowed (%)	2.6	2.7
Disallowed (%)	0.0	0.0

[a]R factor = $\sum hkl||Fobs| - k|Fcal||/\sum hkl|Fobs|$, where Fobs and Fcal are observed and calculated structure factors, respectively. [b]Rfree, the sum is extended over a subset of reflections (5%) excluded from all stages of refinement.

Section 5.3. Results

The Matthews coefficient [30] for the FSR N-terminal crystals was calculated to be 2.17 $Å^3$/Da, assuming two monomers in the asymmetric unit, which corresponds to a solvent content of 43%. The final model included two monomers and had an R_{factor} and R_{free} of 18.5% and 24.1%, respectively. Each monomer contains residues 7-560 of 764 possible amino acid residues. The C-terminal region, residues 561-764, has no electron density. The crystals were analyzed by SDS-

PAGE and the molecular weight was seen to be approximately 60 kDa (the full length FSR protein is 86 kDa), which suggests that the electron density of the C-terminal region is missing due to proteolysis. Proteolysis is further supported by crystallization of the FSR N-terminal domain protein, which only contains residues (1-560), in the same condition, space group, and approximate unit cell parameters. Each monomer in the asymmetric unit contains one active site, each of which contains a Fe_4S_4 cluster and a siroheme. This N-terminal region of FSR is structurally similar to other sulfite and nitrite reductases and contains three domains. According to size exclusion chromatography results at a pH of 8.0, FSR N-terminal is a monomer in solution.

The Matthews coefficient [30] for the FSR/sulfide complex crystals was calculated to be 2.98 $Å^3\ Da^{-1}$, which assuming one monomer in the asymmetric unit corresponds to a solvent content of 58.79%. The final model included one monomer and had an R_{factor} and R_{free} of 15.5% and 19.5% respectively. Of 764 possible amino acid residues, the structure contains residues 4-673. The C-terminal region residues 674-764 are disordered and the electon density is missing. This monomer contains the same three domains as the FSR N-terminal domain structure and one extra domain.

Structure of FSR

FSR likely contains five structural domains. The first three domains are visible in the SeMet FSR structure, and the first four domains are visible in the FSR/sulfide complex. Domain 1 contains two β-sheets: the first contains five antiparallel β-strands flanked by five α-helices, and the second contains four antiparallel β-strands flanked by two α-helices. Domain 2 contains a four stranded mostly parallel β-sheet (β5↓β6↓β7↑β8↓) surrounded by four solvent exposed α-helices. Domain 3 has a four-stranded mostly parallel β-sheet (β14↓β15↓β16↑β17↓) flanked by

seven solvent exposed α-helices. Domain 4 contains five α-helices (Figure 5.7). Residues 632-764 likely form domain 5, the domains is missing because of proteolysis in the SeMet FSR structure, and disordered in the FSR/sulfide complex (Figure 5.8).

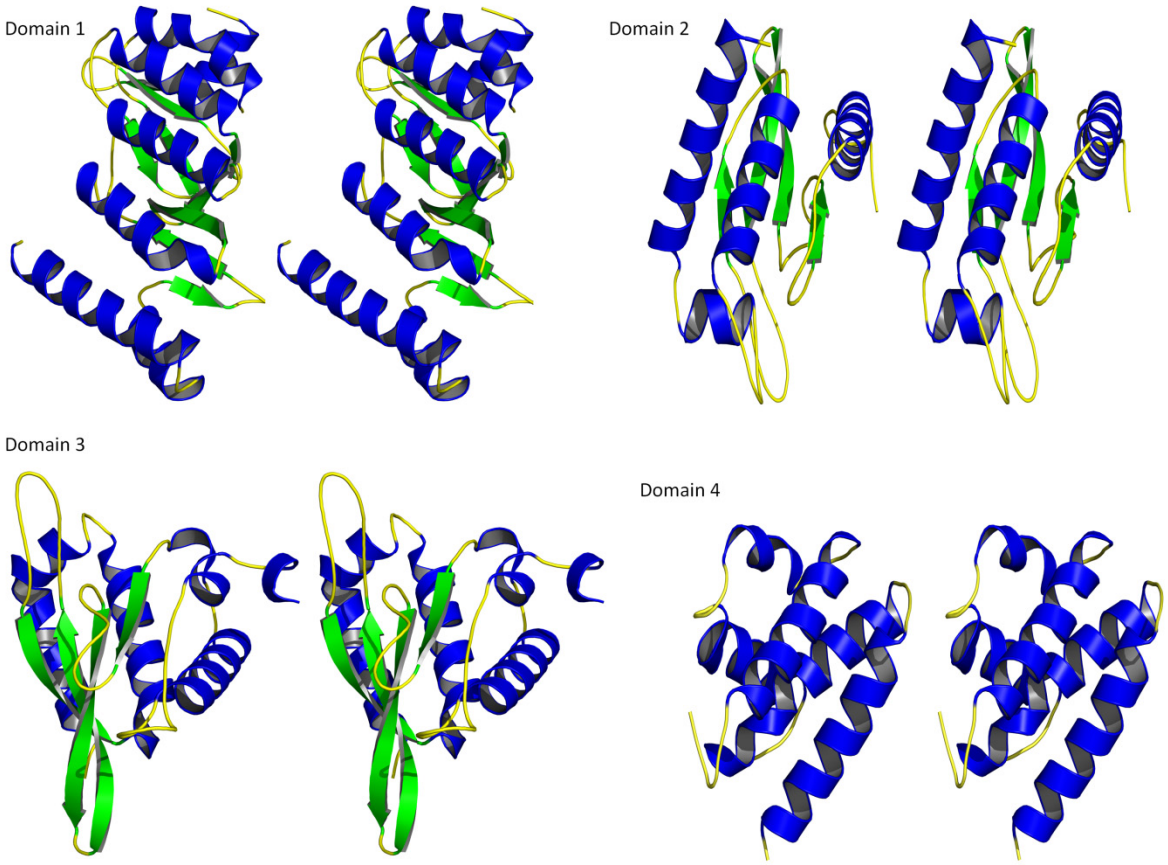

Figure 5.7. Stereodiagram of each of the four domains with α-helices colored in blue, β-strands colored in green, and loops colored in yellow.

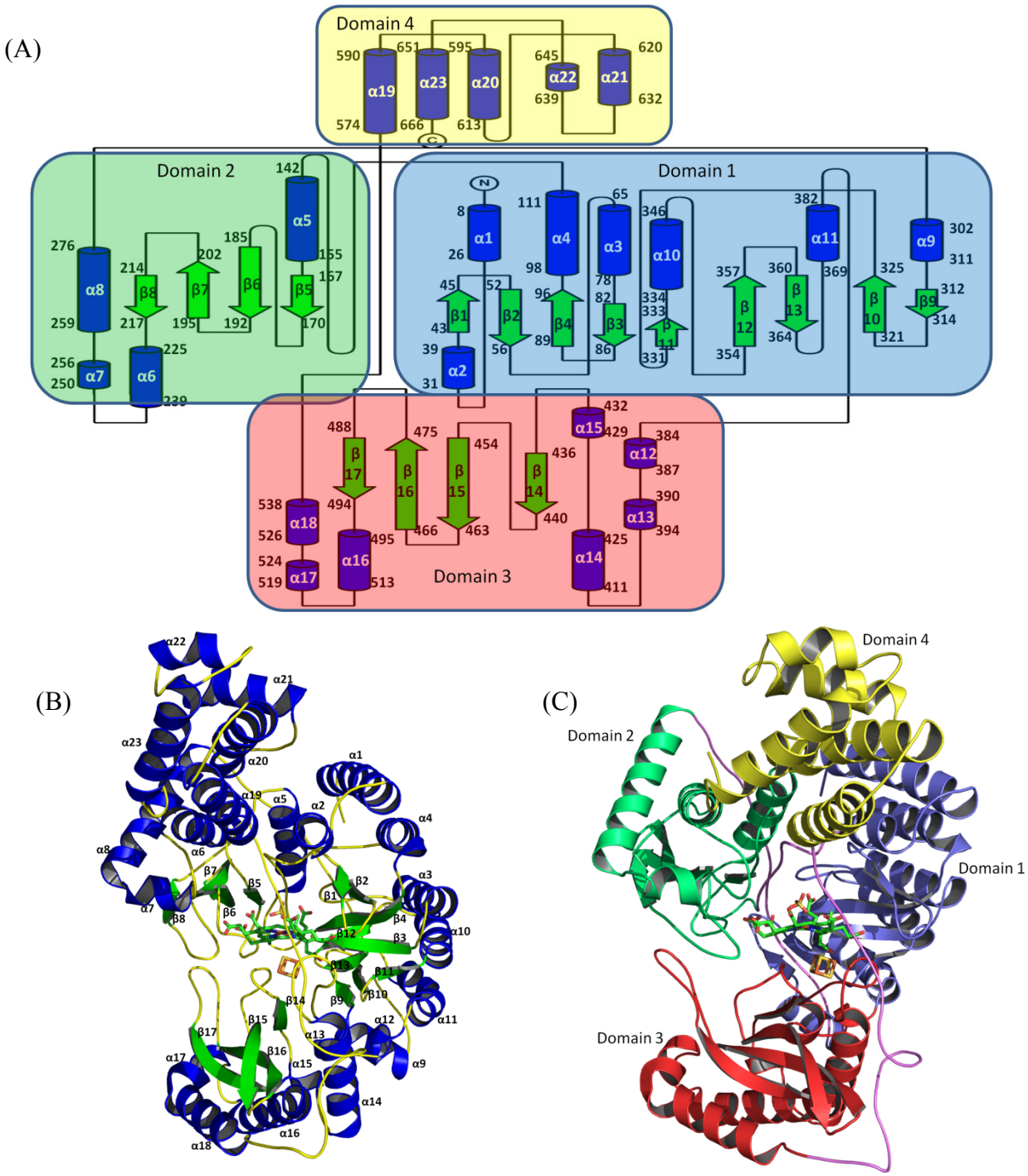

Figure 5.8. (A) Topology diagram of FSR /sulfide complex structure with α-helices shown in blue and β-strands shown in green. Colored boxes represents the secondary structure that belongs to each domain: Domain 1 (blue), Domain 2 (green), Domain 3 (red), Domain 4 (yellow). (B) Monomer of FSR structure with α-helices labeled and shown in blue and β-strands labeled and shown in green. (C) Monomer of FSR structure domains colored to correspond to the colored boxes of the topology diagram. Linker regions between domains are colored in purple.

FSR Dimer

Size exclusion chromatography at pH 8.0 suggests that the full-length FSR is a dimer in solution. Although the N-terminal FSR structure contains two molecules in the asymmetric unit the two molecules do not form a dimer, which agrees with the N-terminal monomer size exclusion chromatography results. The FSR/sulfide complex structure has one molecule in the asymmetric unit and no dimer is present in the crystallographic symmetry. This FSR structure is a monomer which disagrees with the size exclusion chromatography results for full-length FSR. However, when the size exclusion chromatography experiment was repeated at the same pH of the crystallization condition (phosphate buffer of pH 6.2), the estimation of the molecular weight suggests that FSR is a monomer at this pH.

Siroheme group

The siroheme prosthetic group (tetrahydroporphyrin of the isobacteriochlorin class) that consists of a metalloporphyrin that contains an Fe (III) bound in the center of the ring (Figure 5.8). The Fe (III) is responsible for the electron transfer in the sulfite reduction reaction, and is coordinated to the sulfur of the Cys446, which links it to the Fe_4S_4 cluster. In the FSR structure, the Fe (III) of the siroheme is bound to sulfate, which is in the sulfite binding site. The distal face of the siroheme is surrounded by several basic residues Arg164, Lys167, Arg124, Arg55, Arg88, and Lys 165. These form hydrogen bonds with the siroheme to stabilize its binding to FSR (Figure 5.9)

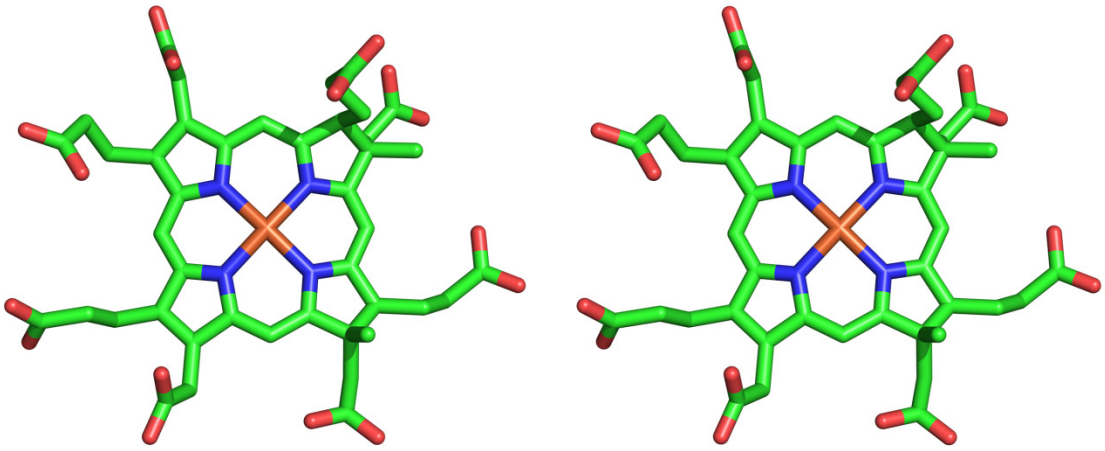

Figure 5.9. Stereodiagram of the siroheme cofactor

Fe_4S_4 cluster

The Fe_4S_4 cluster of sulfite reductases start in an overall (2+) state. The Fe_4S_4 cluster has four iron atoms that are coordinated to the four sulfur atoms of the cluster and the four sulfur atoms of four cysteine residues. Cys396, Cys404, Cys442, and Cys446 in the FSR structures (Figure 5.10).

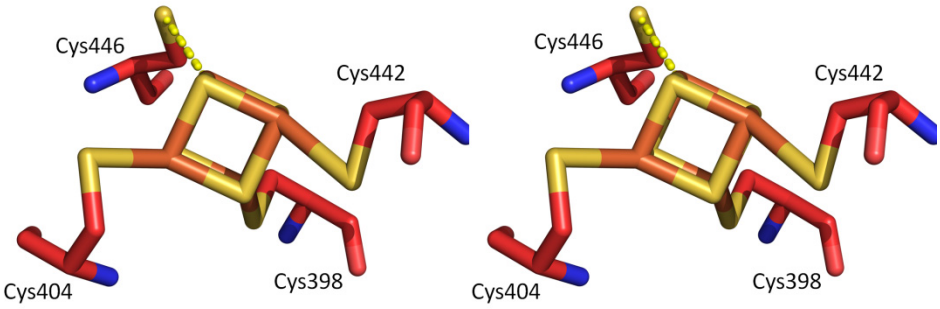

Figure 5.10. Stereodiagram of the iron sulfur cluster from FSR structure coordinated with the cysteines (Cys398, Cys404, Cys442, Cys446)

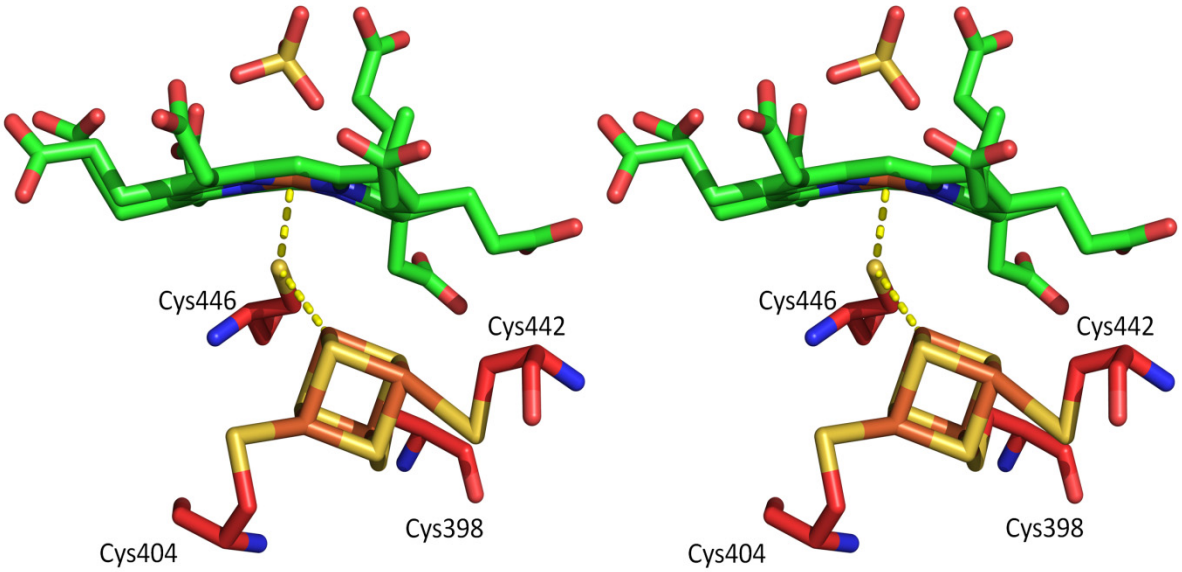

Figure 5.11. Stereodiagram of FSR active site. Residues labeled and colored by domain. Domain 3 is red

Sulfate binding site

Both the N-terminal and FSR/sulfide complex structure have a negatively charged ion, sulfate, bound in active site. The sulfate forms hydrogen bonds with the residue Arg55, Arg124, Lys165, and Lys 167 (Figure 5.11). In the FSR/sulfide complex structure, above the distal face of the siroheme, Cys570 extends into the active site from a linker between domains 3 and 4 (Figure 5.12). The sulfur of Cys570 is 3.9 Å from the sulfur atom of the bound sulfate. This residue's position suggests that the Cys570 would accept the product of the sulfite reduction sulfide for sulfur transfer to the sulfur carrier protein HcyS. Despite being necessary for the crystallization condition, sulfide is not seen in the structure.

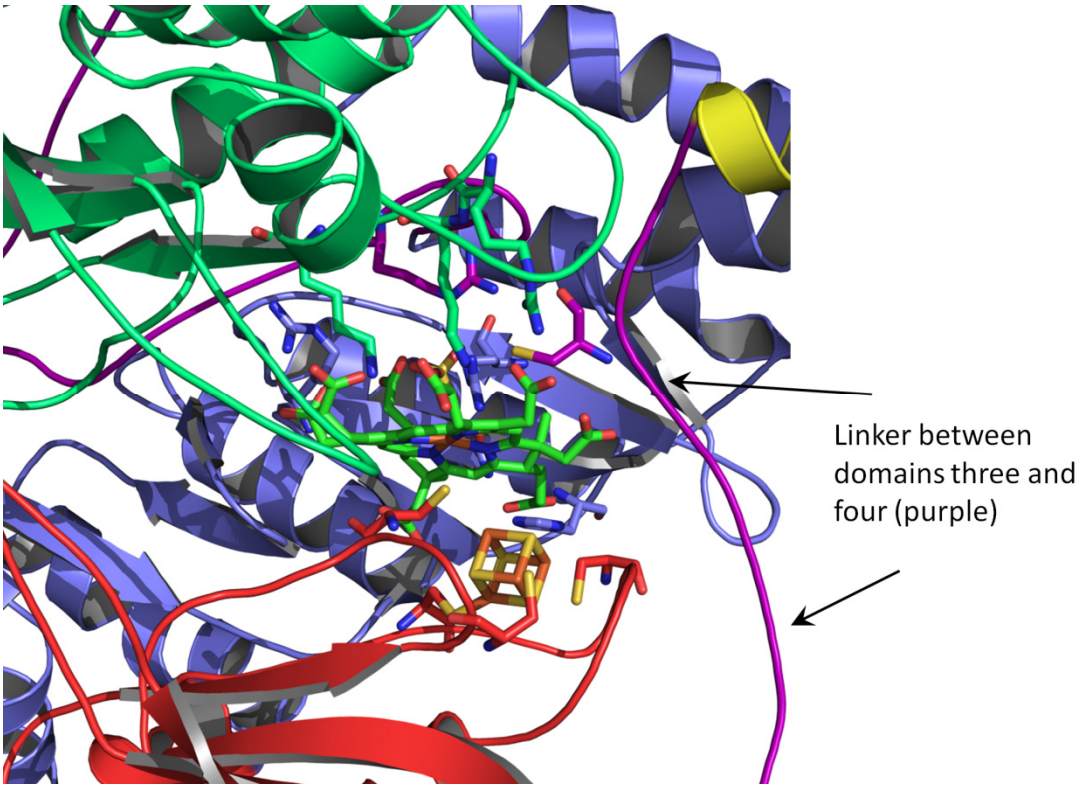

Figure 5.12. FSR structure residues (4-673) colored by domains: domain 1(blue), domain 2 (green), domain 3 (red), domain 4 (yellow). Linker region shown in purple. The linker between domain 3 and 4 that overlays the active site and inserts a cysteine into the active site is labeled.

Section 5.4. Discussion

Comparison of FSR and homologs

A DALI search of the FSR N-terminal structure and the sulfide co-crystallization structure identified several structural homologs [31]. The top hit for the sulfide co-crystallization is an assimilatory sulfite reductase (aSir) from *Mycobacterium tuberculosis* (PDB code 1ZJ9) [14] with a Z score of 36.3 and an aligned sequence identity of 25%. All the top hits are nitrite or sulfite reductases and have approximately 540 residues and three α/β domains (Table 5.3). Domain 1 forms hydrogen bonds with the siroheme and provides basic residues for the sulfite binding site. Additionally, domain 1 has two parts separated in the sequence by domain 2.

Domain 2 forms the other half of the siroheme and sulfite binding site interactions. Then domain 3 interacts with the Fe_4S_4 cluster. When the FSR N-terminal structure is superimposed with the three domains of the sulfite reductase (1ZJ9), the r.m.s.d. is 1.6 Å (Figure 5.13). No other assimilatory sulfite or nitrite reductase structure currently in the PDB database is fused with other domain types like is seen in the FSR/sulfide complex structure.

The active sites of these sulfite and nitrite reductase homologs are similar to FSR. The top hits all include a Fe_4S_4 cluster coordinated to four cysteine residues with one cysteine forming a bridge between the Fe_4S_4 cluster and the iron atom of the siroheme. The distal face of the siroheme is surrounded by basic residue to provide protons for the reduction and to stabilize the binding of the negatively charged sulfite ion. Also these basic residues form hydrogen bonds with the siroheme to stabilize the binding of the prosthetic group. However, none of these structures have a cysteine residue to accept the sulfide for transfer as FSR Cys570.

Table 5.3 Summary of top results of the FSR N-terminal DALI search

PDB code	Protein	Z-score	R.m.s.d	Sequence identity (%)
1zj9	Sulfite reductase (*Mycobacterium tuberculosis)*	36.3	2.3	25
2akj	Nitrite reductase (*Spinacia oleracea*)	35.5	2.3	26
2gep	Sulfite reductase (*Escherichia coli)*	32.0	2.7	19

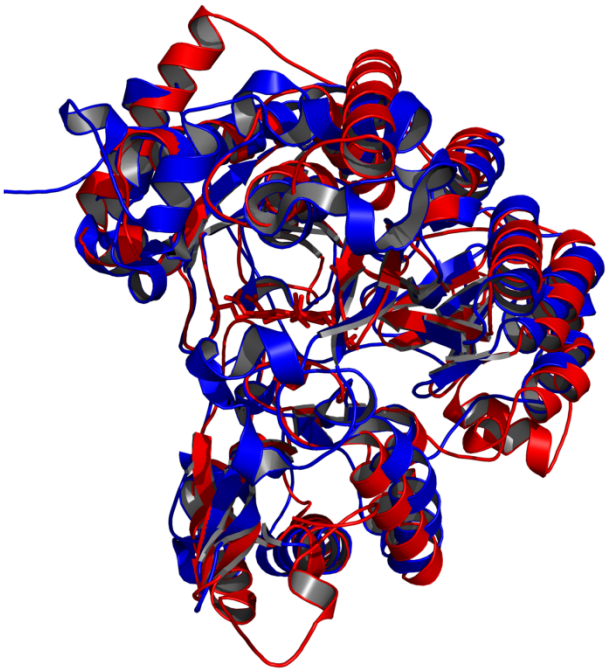

Figure 5.13. FSR N-terminal structure (red) overlaid with sulfite reductase (PDB code 1zj9) (blue) [14]. Structures have the same fold and the r.m.s.d is 1.6 Å.

Oligomeric structure and domain organization in comparison to homologs

These other sulfite reductases are all monomers, which makes the dimer formation at pH 8.0 suggested by the size exclusion chromatography results more unusual. However, since FSR is a monomer at the pH of the crystallization condition, the functional differences created by the dimer formation cannot be predicted.

A DALI search of the FSR domain four (residues 570-674) yielded structural homologs [32]. The top hit is a hypothetical protein (NP_069135.1; PDB code 2HSB) [results not yet published] with a Z score 8.6 and a sequence alignment of 9% (Figure 5.14).

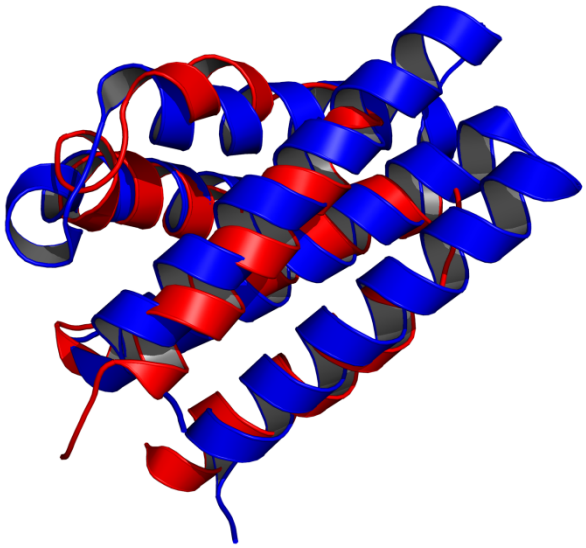

Figure 5.14. FSR domain four structure (red) overlaid with the 2HSB (HEPN domain) structure (blue) [34]. The r.m.s.d is 2.1 Å and shows that domain four has the same fold as the HEPN domain.

This structure, as well as the other top hits of the DALI search, is identified as a HEPN (higher eukaryotic and prokaryotic nucleotide binding) domain (Table 5.4). The function of this type of domain is not completely known, however it is suggested that it will bind monophosphate nucleotide molecules [34]. This implies that this domain in FSR would be involved in binding of the AMP molecule of the adenylated HcyS.

Table 5.4. Summary of top DALI search results for FSR domain four

PDB code	Protein	Z-score	R.m.s.d.	Sequence identity (%)
2hsb	HEPN domain (*Archaeoglobus fulgidus*)	8.6	2.1	9
3o10	Sacsin HEPN domain (*Homo sapien*)	7.2	2.6	6
1o3u	HEPN domain (*Thermatoga maritima*)	6.9	2.5	10
1kny	Kanamycin Nucleotidyltransferase (*Staphylococcus aureus*)	6.0	6.8	11

FSR domain four was superimposed with the kanamycin nucleotidyltransferase (KNTase) (PDB code 1KNY) [35], which contains an HEPN domain and a nucleotidyltransferase domain and the structure contains the nucleotide diphosphomethylphosphonic acid adenosyl ester (AMPCPP). Comparison of the KNTase nuceleotide binding site to FSR shows some conservation of the residues involved in the binding (Figure 5.15). These conserved interactions are made up by a dimer of the HEPN domain in KNTase [33]. Two of the four AMPCPP binding interactions of the HEPN domain are conserved in FSR domain 4. This HEPN dimer might be recreated in the FSR dimer (at biological pH) and bind the AMP of adenylated HcyS. However, because the sequence identity is very low between the FSR domain four and KNTase, and only part of the interactions are conserved, it is possible that domain four only has the same fold as the HEPN domain. It might be that domain four does not function in the binding of the adenylated HcyS, and is only present to anchor the linker between domains three and four and allow Cys570 to be inserted into the active site.

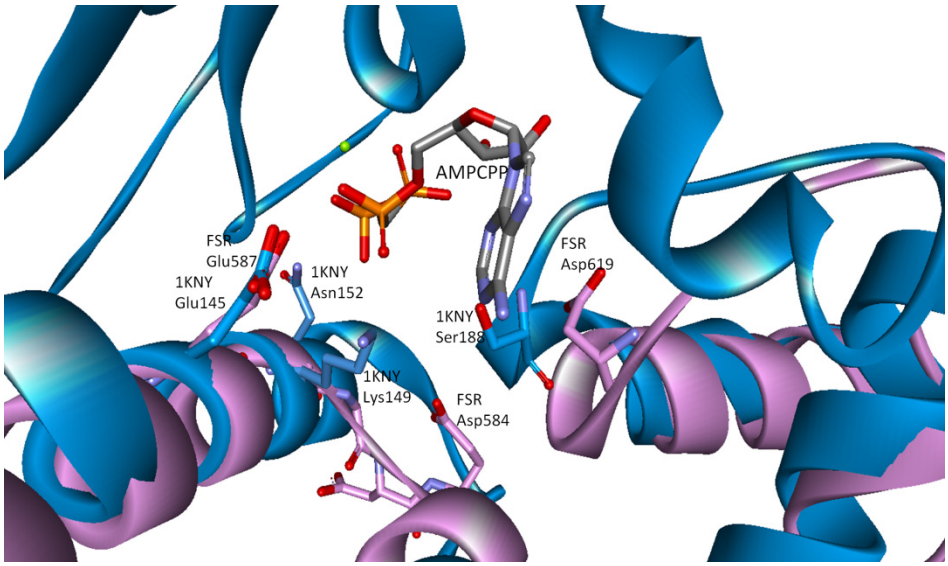

Figure 5.15. FSR domain 4 (pink) superimposed with the 1KNY structure (blue). The 1KNY is a dimer with a HEPN domain in each monomer. Two copies of the FSR domain 4 is overlaid with each of the HEPN domains. The AMPCPP is bound between the HEPN domains of the dimer with interactions from both chains. The 1KNY Ser18 form a hydrogen bond with the adenine ring, In FSR there is Asp619 close to the same position the side chain is pointed up in this structure but a different rotamer could allow it to form the same hydrogen bond. The second interaction is 1KNY Glu145 which form a hydrogen bond with the oxygen of the first phosphate of AMPCPP. The FSR Glu587 is in the same position and could form the same interaction. The other two interaction of 1KNY are Lys149 and Asn152 with phosphate two and three, which do not overlay with any similar residues in FSR. In FSR the α-helix is shorter and forms a loop at this position with Gly591, Gly592, Asp593, and a Glu594, which are not in position to interact with AMPCPP. This could be because the FSR would bind AMP where 1KNY binds ATP [33].

FSR Disordered Domain

The disordered region (residues 674-764) in the FSR/sulfide complex structure shows 50% identity with a putative sulfurtransferase TusA from *Candidatus Nitrospira defluvii* (YP_003798379), and a 27% identity with a TusA ortholog from *E. coli* from which the structure has been solved (PDB code 1DCJ) [34]. Biochemical studies of TusA show that $Cys19^{TusA}$ accepts a sulfide from IscS and forms a persulfide. The sulfide is then transferred

from TusA to another protein TusD and eventually to tRNA to make the mnm^5s^2U modification. Genomic analysis of bacterial sulfurtransferases has identified large proteins that contain TusA sulfurtransferases, different from rhodanese homology domains, at their C-terminal in Episilion (*W. succinogenes*) and Delta proteobacteria and Clostridia [35]. A multiple sequence alignment (Figure 5.16) of FSR C terminal region, 1DCJ (TusA), and other sulfurtransferases of this type identifies $Cys703^{FSR}$ as important for sulfide transfer. Because of the disorder of the region containing Cys703, further study is needed to establish the role of Cys703 and the C-terminal putative sulfertransferase region.

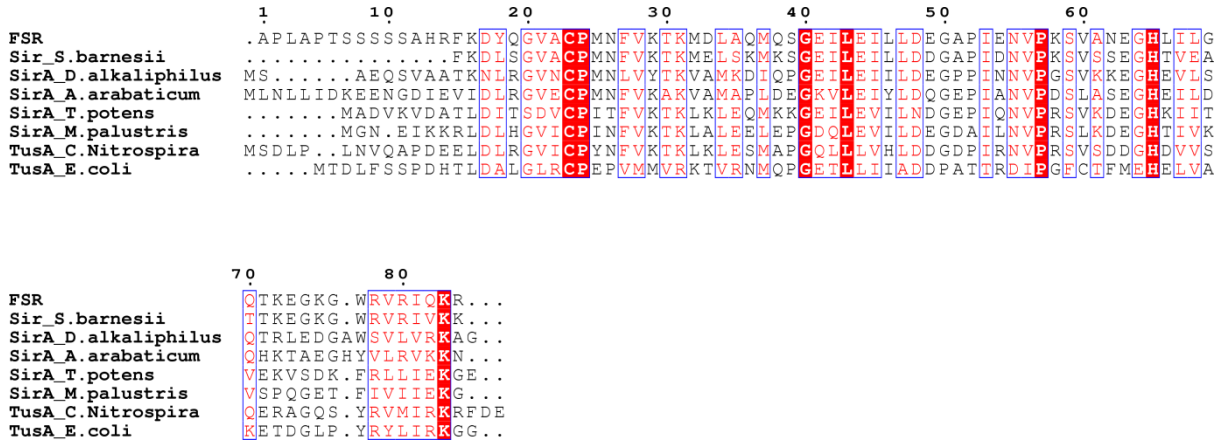

Figure 5.16. Multiple sequence alignment for FSR disordered domain (680-764). The Sir protein is another approximately 760 residue putative sulfite reductase fused with other domains that has not been characterized, TusA and SirA proteins are small (about 100 residues) sulfurtransferases. The conserved cysteine at FSR position 23 is Cys703 in the complete protein sequence. This conserved cysteine in TusA *E.coli* is the active site cysteine that forms the persulfide

Mechanism of FSR

The Fe_4S_4 cluster is in an oxidized overall 2+ state to start, the iron of the siroheme is also oxidized at an Fe (III) state 22 (Figure 5.17). The mechanism starts with a transfer of two electrons to the Fe_4S_4 cluster, one of which is transferred through the cysteine that covalently

links the cluster to the siroheme iron. This makes the state of the cluster to 1+ and the siroheme iron Fe (II) 23. After the reduction of the cofactors the substrate, sulfite, binds to the Fe (II) of the siroheme 24. The basic residues above the siroheme stabilize the binding of the sulfite by forming hydrogen bonds with the oxygen atoms of the substrate. The electrons held by the Fe_4S_4 cluster and the siroheme iron are transferred to the sulfite, which destabilizes one of the oxygen sulfur bonds leading it to break 25. Protons from the basic residues are donated to make H_2O from the released oxygen, and the cofactors are returned to the cluster 2+ and siroheme Fe(III) states. This process is repeated twice as four electrons are transferred through the cofactor to the sulfite, the other two oxygen bonds are broken and form two H_2O molecules 26 and 27. The sulfide is then released from the cofactors, leaving them in the original $[Fe_4S_4]^{2+}$ and Fe (III) state 22 [9]. Based on the cysteine in the active site, the sulfide would then form a persulfide with Cys 570. The sulfide is then transferred to the sulfur carrier protein HcyS to form the C-terminal thiocarboxylate, whether this involves another intermediate persulfide transfer is currently unclear.

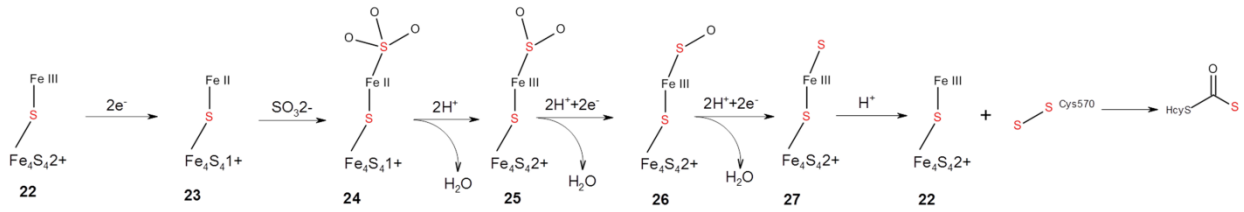

Figure 5.17. Mechanism of FSR, the Fe_4S_4 represents the iron sulfur cluster and the Fe III is the iron of the siroheme. The sulfur that is the covalent link between the two is from a protein cysteine residue Cys446.

Section 5.5. Conclusions

Unlike previously characterized sulfite reductases, which have approximately 560 residues and contain three structural domains, *W. succinogenes* FSR contains approximately 760 residues with five structural dontains. The first three structural domains are homologous to those of conventional sulfite reductases. Domains 4 and 5 were missing in the first FSR structure as the result of proteolysis. A new condition containing sulfide resulted in a crystal structure in which domain 4 was ordered, but domain 5 was not, even though no proteolysis had occurred. Surprisingly, domain 4, which showed no sequence similarity to any structure in the PDB, is structurally homologous to HEPN nucleotide binding domains, suggesting that domain 4 might bind adenylated HcyS for sulfur transfer and thiocarboxylate formation. In addition, Cys570 in the linker between domains 3 and 4 was found in the FSR active site about 4 Å from the predicted sulfite sulfur atom position, suggesting the possibility of a sulfur transfer mechanism involving Cys570 persulfide formation. Interestingly, domain 5 is homologous to TusA, which has been shown to aid sulfur transfer in other sulfur carrier protein pathways through cysteine (equivalent to Cys703 in HcyS) persulfide formation. Whether Cys570 transfers the sulfide directly to HcyS, or transfers the sulfide to Cys703 in domain 5, which then subsequently transfers it to the HcyS is currently unknown. Mutagenesis and kinetic studies are currently underway to unravel this question.

REFERENCES

1. Kessler, D., *Enzymatic activation of sulfur for incorporation into biomolecules in prokaryotes.* FEMS Microbiol Rev, 2006. **30**(6): p. 825-40.
2. Mueller, E.G., *Trafficking in persulfides: delivering sulfur in biosynthetic pathways.* Nat Chem Biol, 2006. **2**(4): p. 185-94.
3. Dorrestein, P.C., et al., *The biosynthesis of the thiazole phosphate moiety of thiamin: the sulfur transfer mediated by the sulfur carrier protein ThiS.* Chem Biol, 2004. **11**(10): p. 1373-81.
4. Burns, K.E., et al., *Reconstitution of a new cysteine biosynthetic pathway in Mycobacterium tuberculosis.* J Am Chem Soc, 2005. **127**(33): p. 11602-3.
5. Leimkuhler, S., M.M. Wuebbens, and K.V. Rajagopalan, *Characterization of Escherichia coli MoeB and its involvement in the activation of molybdopterin synthase for the biosynthesis of the molybdenum cofactor.* J Biol Chem, 2001. **276**(37): p. 34695-701.
6. Mileni, M., et al., *Heterologous production in Wolinella succinogenes and characterization of the quinol:fumarate reductase enzymes from Helicobacter pylori and Campylobacter jejuni.* Biochem J, 2006. **395**(1): p. 191-201.
7. Krishnamoorthy, K. and T.P. Begley, *Protein thiocarboxylate-dependent methionine biosynthesis in Wolinella succinogenes.* J Am Chem Soc, 2011. **133**(2): p. 379-86.
8. Tran, T.H., et al., *A novel mechanism of sulfur transfer catalyzed by O-acetylhomoserine sulfhydrylase in the methionine-biosynthetic pathway of Wolinella succinogenes.* Acta Crystallogr D Biol Crystallogr, 2011. **67**(Pt 10): p. 831-8.
9. Crane, B.R., L.M. Siegel, and E.D. Getzoff, *Probing the catalytic mechanism of sulfite reductase by X-ray crystallography: structures of the Escherichia coli hemoprotein in*

complex with substrates, inhibitors, intermediates, and products. Biochemistry, 1997. **36**(40): p. 12120-37.

10. Crane, B.R. and E.D. Getzoff, *The relationship between structure and function for the sulfite reductases.* Curr Opin Struct Biol, 1996. **6**(6): p. 744-56.
11. Crane, B.R., L.M. Siegel, and E.D. Getzoff, *Sulfite reductase structure at 1.6 A: evolution and catalysis for reduction of inorganic anions.* Science, 1995. **270**(5233): p. 59-67.
12. Swamy, U., et al., *Structure of spinach nitrite reductase: implications for multi-electron reactions by the iron-sulfur:siroheme cofactor.* Biochemistry, 2005. **44**(49): p. 16054-63.
13. Nakano, S., et al., *Structure-function relationship of assimilatory nitrite reductases from the leaf and root of tobacco based on high-resolution structures.* Protein Sci, 2012. **21**(3): p. 383-95.
14. Schnell, R., et al., *Siroheme- and [Fe4-S4]-dependent NirA from Mycobacterium tuberculosis is a sulfite reductase with a covalent Cys-Tyr bond in the active site.* J Biol Chem, 2005. **280**(29): p. 27319-28.
15. Ikeuchi, Y., et al., *Mechanistic insights into sulfur relay by multiple sulfur mediators involved in thiouridine biosynthesis at tRNA wobble positions.* Mol Cell, 2006. **21**(1): p. 97-108.
16. Shi, R., et al., *Structural basis for Fe-S cluster assembly and tRNA thiolation mediated by IscS protein-protein interactions.* PLoS Biol, 2010. **8**(4): p. e1000354.
17. Dahl, J.U., et al., *The Sulfur Carrier Protein TusA Has a Pleiotropic Role in Escherichia coli That Also Affects Molybdenum Cofactor Biosynthesis.* Journal of Biological Chemistry, 2013. **288**(8): p. 5426-5442.

18. Schmitz, J., et al., *Role of the C-terminal Gly-Gly motif of Escherichia coli MoaD, a molybdenum cofactor biosynthesis protein with a ubiquitin fold.* Biochemistry, 2007. **46**(3): p. 909-16.

19. Dahl, J.U., et al., *The Identification of a Novel Protein Involved in Molybdenum Cofactor Biosynthesis in Escherichia coli.* Journal of Biological Chemistry, 2011. **286**(41): p. 35801-35812.

20. Maniatis, T., E. F. Fritsch, and J. Sambrook. Molecular cloning, a laboratory manual. 1982, Cold Spring, NY: Cold Spring Harbor Laboratory.

21. Ausubel, F.M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl. Short protocols in molecular biology. 2002, New York, NY: John Wiley & Sons.

22. 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, 1976. **72**: p. 248-54.

23. Otwinowski, Z. and W. Minor, *Processing of X-ray diffraction data collected in oscillation mode.* Macromolecular Crystallography, Pt A, 1997. **276**: p. 307-326.

24. Vagin, A. and A. Teplyakov, *An approach to multi-copy search in molecular replacement.* Acta Crystallogr D Biol Crystallogr, 2000. **56**(Pt 12): p. 1622-4.

25. Schwarzenbacher, R., et al., *The importance of alignment accuracy for molecular replacement.* Acta Crystallogr D Biol Crystallogr, 2004. **60**(Pt 7): p. 1229-36.

26. Adams, P.D., et al., *PHENIX: a comprehensive Python-based system for macromolecular structure solution.* Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 2): p. 213-21.

27. Emsley, P. and K. Cowtan, *Coot: model-building tools for molecular graphics.* Acta Crystallogr D Biol Crystallogr, 2004. **60**(Pt 12 Pt 1): p. 2126-32.

28. Laskowski, R.A., D.S. Moss, and J.M. Thornton, *Main-chain bond lengths and bond angles in protein structures.* J Mol Biol, 1993. **231**(4): p. 1049-67.

29. DeLano, W.L., *The PyMOL molecular graphics system.* 2002.

30. Matthews, B.W., *Solvent content of protein crystals.* J Mol Biol, 1968. **33**(2): p. 491-7.

31. Holm, L. and P. Rosenstrom, *Dali server: conservation mapping in 3D.* Nucleic Acids Res, 2010. **38**(Web Server issue): p. W545-9.

32. Grynberg, M., H. Erlandsen, and A. Godzik, *HEPN: a common domain in bacterial drug resistance and human neurodegenerative proteins.* Trends Biochem Sci, 2003. **28**(5): p. 224-6.

33. Pedersen, L.C., M.M. Benning, and H.M. Holden, *Structural investigation of the antibiotic and ATP-binding sites in kanamycin nucleotidyltransferase.* Biochemistry, 1995. **34**(41): p. 13305-11.

34. Katoh, E., et al., *High precision NMR structure of YhhP, a novel Escherichia coli protein implicated in cell division.* J Mol Biol, 2000. **304**(2): p. 219-29.

35. Kotera, M., et al., *Comprehensive genomic analysis of sulfur-relay pathway genes.* Genome Inform, 2010. **24**(1): p. 104-15.

36. Mansuy, D. and J.L. Boucher, *Alternative nitric oxide-producing substrates for NO synthases.* Free Radic Biol Med, 2004. **37**(8): p. 1105-21.

CHAPTER 6: CONCLUSION

Using X-ray crystallography to examine structure has provided insight into protein function and mechanism for decades. The act of growing crystals, collecting diffraction patterns, and determining a protein structure at an atomic resolution has altered our understanding of molecular machinery of cellular pathways. In a protein structure knowing the position of every atom can provide key catalytic residues that provide insight into a proteins function and mechanism.

For example, the structure of ferredoxin sulfite reductase (FSR) reported in Chapter 5 provides new information about the transfer of sulfur in pathways of sulfur carrier proteins. Sulfur carrier protein pathways have been significantly studied over that last few years and despite the extensive study of the pathways including thiamin [1], cysteine [2], and molybdopterin [3] the new variation in sulfur transfer methods are still being discovered. Many of the thiocarboxylates of the sulfur transfer proteins are made through the transfer of a sulfur from a persulfide created by a cysteine desulferase, like IscS [4]. However, sometimes there is another protein containing a rhodanese-like domain that accepts the persulfide to transfer it to the sulfur carrier protein [5]. There is also evidence another type of sulfurtransferase TusA could also transfer the sulfur to make the thiocarboxylate [6]. In the sulfur carrier protein dependent methionine biosynthesis the sulfur source is unusually supplied by FSR. The sulfide would then react with the C-terminal of the activated sulfur carrier protein HcyS in a nucleophilic attack [7]. FSR was thought to only be involved in the reduction of sulfite providing an excess amount of sulfide. However, the structure of FSR provided insight into a more complex function. The structure demonstrates that it is a multiple domain protein that seems to be a fusion of several smaller proteins with different functions.

The crystal structure of FSR residues (4-673) presented in Chapter 5 was determined to a resolution of 2.4 Å, using molecular replacement from the FSR N-terminal structure which solved by SAD phasing. It is similar to other sulfite reductases, except it contains another domain. This other domain (domain four) is separated from domain three by a long linker that overlays the active site. The presence of the Cys570 residue inserted into the active site from this linker suggests that FSR is involved in the formation of a persulfide for eventual sulfur transfer to the HcyS thiocarboxylate. The last 90 residues of FSR structure are disordered, but the region has significant sequence identity with the sulfurtransferase TusA further suggesting the function of FSR to be involved in the transfer of the sulfide after the reduction. This is the only structure of a sulfite reductase currently solved with a cysteine positioned to accept the sulfide for sulfur transfer.

The sulfur carrier protein HcyS of this same pathway is also of interest for the structural elements it uses to form complexes. Since HcyS must complex with several other proteins the structural properties that allow this variety of protein-protein interactions are very interesting. The studies of HcyS complex formation are presented in Chapter 4. The only complex that is stable enough to observe through size exclusion chromatography is the HcyS/HcyD complex, which is the first step of the methionine biosynthesis pathway. The other complex formation HcyS/HcyF and HcyS/FSR might be sensitive to state of ligand binding and the state of the HcyS C-terminal, which is the catalytic region of the protein and is modified by each protein in the pathway. It is possible that complex formation is highly sensitive to the state of these modifications as only the HcyS/HcyD complex, the first step of the pathway requiring no modification, is stably formed. The attempts to tests the HcyS-GG ability to form a stable complex with HcyF did not show any change in complex formation. This might be due to the

need for the ligands for the reaction ATP and Mg^{2+} for stable complex formation. Overall, the complex formation seems to be rather unstable due to changes in HcyS C-terminal and ligand binding.

The sulfur carrier protein dependent methionine biosynthesis pathway provides new opportunities to probe structural elements and functions of interesting enzymes. A collaboration of structural biologists and mechanistic enzymologists will continue to work to elucidate these structural questions.

REFERENCES

1. Dorrestein, P.C., et al., *The biosynthesis of the thiazole phosphate moiety of thiamin: the sulfur transfer mediated by the sulfur carrier protein ThiS.* Chem Biol, 2004. **11**(10): p. 1373-81.
2. Burns, K.E., et al., *Reconstitution of a new cysteine biosynthetic pathway in Mycobacterium tuberculosis.* J Am Chem Soc, 2005. **127**(33): p. 11602-3.
3. Leimkuhler, S., M.M. Wuebbens, and K.V. Rajagopalan, *Characterization of Escherichia coli MoeB and its involvement in the activation of molybdopterin synthase for the biosynthesis of the molybdenum cofactor.* J Biol Chem, 2001. **276**(37): p. 34695-701.
4. Lauhon, C.T. and R. Kambampati, *The iscS gene in Escherichia coli is required for the biosynthesis of 4-thiouridine, thiamin, and NAD.* J Biol Chem, 2000. **275**(26): p. 20096-103.
5. Dahl, J.U., et al., *The identification of a novel protein involved in molybdenum cofactor biosynthesis in Escherichia coli.* J Biol Chem, 2011. **286**(41): p. 35801-12.
6. Dahl, J.U., et al., *The sulfur carrier protein TusA has a pleiotropic role in Escherichia coli that also affects molybdenum cofactor biosynthesis.* J Biol Chem, 2013. **288**(8): p. 5426-42.
7. Krishnamoorthy, K. and T.P. Begley, *Protein thiocarboxylate-dependent methionine biosynthesis in Wolinella succinogenes.* J Am Chem Soc, 2011. **133**(2): p. 379-86.