MQP JMA AAGD

Bioinformatic and Metal Analysis of Copper Homeostasis in *Sinorhizobium meliloti*

A Major Qualifying Project Report:

Submitted to the Faculty

Of the

WORCESTER POLYTECHNIC INSTITUTE

In partial fulfillment of the requirements for the

Degree of Bachelor of Science

by

Jacquelyn Nemet-Sousa

Date: October 10, 2013

Approved:

Professor José Argüello, Project Advisor

Table of Contents

Abstract

The *Sinorhizobium meliloti* genome encodes five Cu^{+}-ATPases that aid in detoxification and regulate copper homeostasis via the efflux of copper from the cytoplasm to the periplasmic space. Bioinformatic analysis concluded that additional cuproproteins found in *S. meliloti* genome might interact with Cu^{+}-ATPases to aid in copper homeostasis. Atomic Absorption Spectroscopy (AAS) was used to determine the concentration of copper in subcellular fractionation of wild type and mutant strains. Results suggest that copper distribution in different subcellular fractions is linked to the dynamic of copper transporters and cuproproteomes.

Acknowledgements

I would like to thank Professor José Argüello for helping me prepare for my future and for encouraging me to excel. I would also like to thank Sarju Patel for devoting his time to teach and assist me in the lab, along with Teresita Padilla-Benavides, Courtney McCann, and Jessica Collins.

Additionally, I would not have made it this far without the help of my Faculty Advisor, Kristin Wobbe.

1. Introduction

1.1. Copper

Copper is an essential nutrient used in many biochemical processes by bacteria and eukaryotes as an enzymatic cofactor. Such processes include energy transduction, iron mobilization, oxidative stress response, and many others. Copper may also act as a signaling molecule. On the other hand, too much copper is extremely toxic to living organisms. Excess copper damages iron-sulfur clusters, interacts with free thiol groups in proteins, and competes with other metals for binding sites in proteins (2) (3).

Copper is a transition metal that has two oxidative states, Cu^{+} and Cu^{2+}, pertaining to the number of electrons in its outer shell and its ability to accept and donate an electron. The ion participates in Fenton reactions that lead to reactive oxygen species (ROS):

$$Cu^{1+} + H_2O_2 \rightarrow Cu^{2+} + OH^{-} + {}^{\bullet}OH$$

This reaction produces hydroxyl radicals (${}^{\bullet}OH$) that can damage lipids, proteins, and nucleic acids. Therefore, ROS can be detrimental to all life forms (4).

Copper needs a way to move around the cell. Free copper within the cell is toxic and therefore specific chaperone proteins bind copper ions in order to reduce the amount of free copper in the cell. Additionally, copper ATPases help regulate homeostasis by transporting Cu^{+} out of the cytoplasm.

1.2. Copper ATPases

Living cells regulate copper homeostasis in order to prevent toxic levels of copper and its consequences. The proteins that help to regulate the cytosolic copper concentrations are known as copper P-type ATPases. Copper P-type ATPases drive the

efflux of copper from the cytoplasm and translocate the ion across the plasma membrane into the periplasmic compartment of the bacterial cell, via ATP hydrolysis. Additionally, they functionally interact with other cuproproteins located in the periplasmic space or within the plasma membrane. Other proteins involved in copper homeostasis include chaperones, transporters, transcription factors, and metallothioneins. Chaperones and chelators bind to copper within the cytoplasm to reduce the amount of free copper. These cuproproteins transport the copper ion to efflux systems located in the plasma membrane, such as copper ATPases (1).

Copper P-type ATPases have eight transmembrane helicies. Two copper ions bind to the protein in a trigonal planar geometry between helicies six and eight (Fig. 1). More specifically, two cysteines and an asparagine make up the first binding site, while a tyrosine, methionine, and a serine make up the second binding site. Both metal binding domains (MBDs) have to be occupied in order to activate ATP hydrolysis. As previously stated, ATP hydrolysis drives the copper ions across the membrane. The ATP binding domain is located between helicies six and seven. The highly specific phosphorylation motif of DKTGT is prominent in every heavy-metal ATPase's sequence. More specifically the asparagine amino acid is phosphorylated by ATP (1).

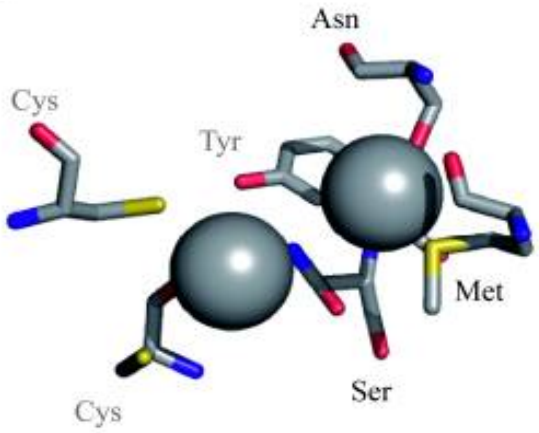

Figure 1: Metal Binding Domain of Copper ATPases. This figure shows the metal binding domain of the copper ATPase efflux system. The figure correctly displays the amino acids associated with the binding of two copper ions, and their conformations.(1)

These ATPases undergo E1/E2 conformational transitions during their transport cycle. In the Albers-Post mechanism, the enzyme adopts two basic conformations (E1⟵⟶ E2), the phosphorylated intermediate (E1P⟵⟶E2P), and occluded ion-bound forms. The proposed catalytic cycle for Cu^+-ATPase is shown in Fig. 2. The central event of unidirectional metal transport is the irreversible ion binding to TM-MBSs. In case of Cu^+-ATPase, amphiphatic kinked TM2b might serve as the docking point for cytoplasmic Cu^+-chaperone (*CopZ*) (5). Previous studies have shown that the ATPase has high affinity for the Cu^+~CopZ complex along with the lack of interaction of the apo-chaperone with the TM-MBS access sites (6). Upon docking of the Cu^+-chaperone, the ion would be transfer unidirectionally to the entry site formed by three invariant residues (Met, Glu and Asp located at the cytoplasmic end of M3, M4 and M5 respectively). The full occupancy of TM-MBSs (2 Cu^+ ions per ATPase) requires the presence of ATP. ATP hydrolysis driven conformational changes are required for metal translocation across the permeability barrier. It has been hypothesized that P_{1B}-ATPases requires specific periplasmic/luminal

"partners" in order to release the metal. Upon metal release, the enzyme phosphorylated E2P form is dephosphorylate to E2 state and enters into the next cycle (1).

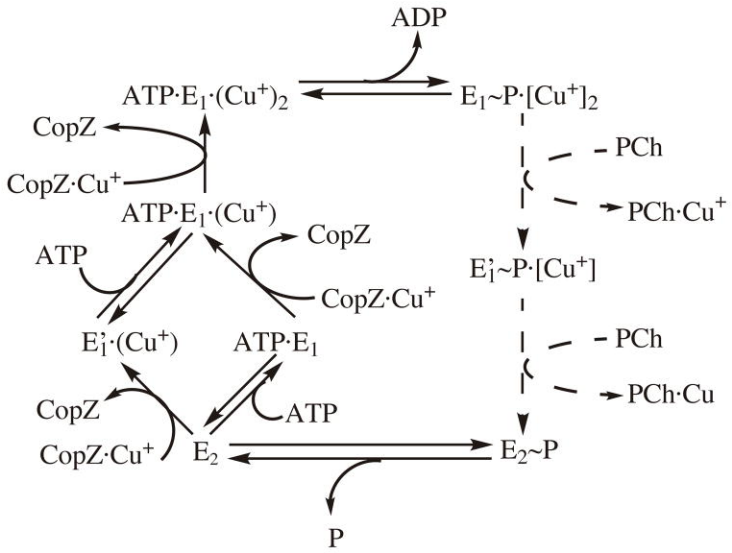

Figure 2: The Catalytic cycle of Copper P_{1b}-type ATPases. This is also known as the Albers-Post transport cycle. Copper ATPases are dependent upon ATP hydrolysis. Copper ions are transported to the copper ATPases via copper chaperones, such as CopZ.

Mutated copper ATPases have major effects that can be detrimental to the organism's existence that include decreased metal resistance due to the enzyme improperly working to excavate copper from within the cell. This can affect the homeostasis of copper between the periplasmic space and the cytoplasm. Limited enzymatic function can also prevent the transfer of copper from a chaperone to an ATPase, depending on where the mutation is. For example, different amino acids may occupy the entrance site and therefore the chaperone may not be able to dock itself in order to transfer the copper ion to the enzyme. In addition, limited enzymatic function may be caused by no phosphorylation of ATP, which also prevents the transport of copper through the membrane.

Several studies have suggested that the host organisms use metal ions (overload of Cu and Zn) as an innate immune response against invading pathogens. Thus, pathogenic

bacteria are required to overcome the host immune responses during infection. Since Cu^+-ATPases translocate metal ions across the plasma membrane, it is likely that Cu^+-ATPases play a role in virulence.

1.3. Symbiosis

S. meliloti is a gram-negative bacterium belonging to the *Rhizobia* family. It lives independently in soil or in nodules on the roots of leguminous plants, such as alfalfa, forming a symbiosis (7). In these nodules the bacterium fixes atmospheric nitrogen into a more applicable form that can be used by the plant. In return, the plant provides the bacterium with energy and a microaerobic environment (8).

S. meliloti is drawn to the root hairs of leguminous plants, such as *Medicago sativa* (alfalfa) by flavonoids, which are aromatic compounds released by the plant. This signal activates the production of nod factors by activating *S. meliloti's nod* genes (9). The plant responds to these nod factors by curling its root hairs that enclose the bacteria. The bacteria gains access into deeper tissues of the plant by creating infection threads that allows the bacteria to successfully invade the plant (10). The bacteria are enclosed in symbiosomes by endocytosis located in the inner cortex of the plant. From here, the symbiosome differentiates into a bacteroid form that fixes nitrogen.

As in any other infection, the host initiates an innate immune response. Relative to the symbiosis of *S. meliloti* and alfalfa plants, the plant host invades the bacterium with excess amounts of copper. This environment becomes extremely toxic to the cells viability. Cu^+-ATPases are programmed defense mechanisms that functions to deplete the excess copper inside the cell. Additionally, plant hosts produce oxidative bursts that lead to ROS. Multiple copies of ATPases encoded in the *S. meliloti* genome are beneficial to bacterial

virulence. In such ways that particular subfamilies of copper ATPases, like CopA2/FixI ATPases, that are associated with cytochrome c oxidases, can help reduce oxygen levels in order to prevent ROS (1).

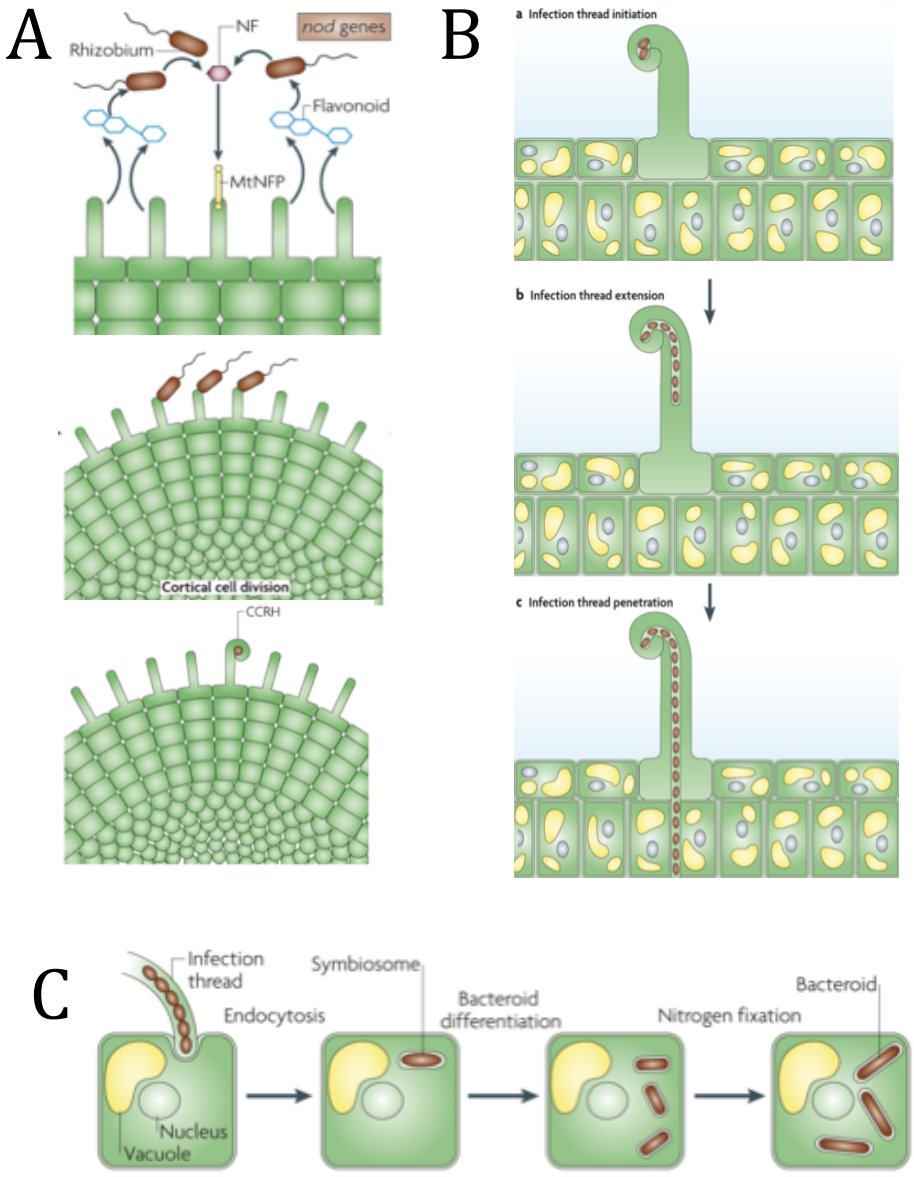

Figure 3: The Rhizobia-Legume symbiosis. Part **A** shows the activation of *nod* genes by the aromatic flavonoid compound, and the initiation of invading the host plant via root hairs. Part **B** shows how the bacteria is taken up by the plant through the curling of the root hairs and the production of the infection thread spreading into the inner tissues of the plant. Part **C** illustrates the endocytosis of the symbiosome and its formation into a nitrogen-fixing bacteroid.

S. meliloti genome is encoded in a chromosome and two megaplasmids, pSymA and pSymB. The first replicon is a single chromosome that is the largest of the three containing 3.6 million base pairs (Mb). This replicon houses genes that are associated with biosynthetic pathways including metabolic pathways, plant interaction, chemo taxis processes, mobility, and responding to external stress (7).

The pSymB replicon has 1.68 Mb and encodes genes that are involved with the absorption of nutrients from the soil and more importantly the viability of the bacterium (11). The pSymA replicon contains 1.35 Mb of DNA, which encode genes required for nodulation and the symbiosis with leguminous plants. These genes are mainly involved with nitrogen and carbon metabolism, transport, and stress (12).

Using bioinformatics tools we have explored the list of cuproproteins used for copper homeostasis in the *S. meliloti* genome. Experimental procedures involving extracting subcellular fractions of *S. meliloti* and Atomic Absorption Spectroscopy (AAS) show where the copper is directed in correlation to the localization of the cuproproteins, by determining the concentration of copper in each subcellular fraction. As previously stated, free copper is mainly located within the cytoplasm of the cell, it is assumed that the majority of the cuproproteins involved with copper homeostasis reside in the cytoplasm or the cytoplasmic membrane. Thus, these proteins can bind copper and transport it to the periplasmic space, out of the cell, or to other cuproproteins.

2. Materials and Methods

2.1. Bioinformatics

A list of viable copper associated proteins, also known as cuproproteins, was obtained from (13). Additional proteins were obtained from unpublished work. Protein sequences were obtained in FASTA format from the NCBI website using the Protein database.

Homologous proteins were determined by using the BLAST search on the NCBI website (http://blast.ncbi.nlm.nih.gov/Blast.cgi) and the KEGG website (http://www.genome.jp/tools/blast/). The FASTA sequence obtained from the NCBI protein database (http://www.ncbi.nlm.nih.gov/protein/) was inserted into the Query box. The organism chosen was *Sinorhizobium meliloti 1021* for NCBI and *sme* for KEGG. The results were displayed in ascending order of E value, where the smallest E value was at the top and the largest was at the bottom. Only the first five E values that were less than 10^{-3} were considered as homologous proteins.

These sequences were aligned with the initial sequences using the MAFFT alignment software, version 7.055b (http://mafft.cbrc.jp/alignment/software/). The sequences were aligned using the more accurate G-INT-I algorithm (option 4), also known as the Needleman-Wunsch algorithm. This algorithm aligns sequences of similar lengths and places gaps throughout the sequence. The output format was in sorted Clustal format (option 1), where asterisks were displayed when amino acid residues were the same throughout all the aligned sequences. Two dots were displayed if the aligned sequences

had similar amino acid properties and one dot was displayed when a few aligned sequences had similar amino acid side chain properties.

Domains and motifs were identified using the PROSITE database on the ExPASy website (http://prosite.expasy.org). The cellular co-localization was determined using the ClubSub-P database on the Max-Plank website (http://toolkit.tuebingen.mpg.de/clubsubp). This shows the location of the protein within the bacterial cell as well as if it is transported via either the secretory (sec) pathway or the twin-arginine translocation (tat) pathway, or both.

2.2. Bacterial strains and growth conditions

S. meliloti wild type strain WSM419 and mutant strain Δ*actP* was grown in 1 L of TY media at 220 rpm, at 30 °C. Wild type strain RM2011 and mutant strain Δ*Cut5* was grown in 1 L of Rhizobium Defined Medium (RDM) at 200 rpm, at 30 °C. Cells were harvested at 1.5-3.0 OD_{600nm}.

The RDM contained 100 mL of a 10X RDM A stock that included 6 g of KNO_3, 1 g of $CaCl_2$, 2.5 g of $MgSO_4$, and 0.1 g of $FeCl_3$. It also contained 100 mL of a 10X RDM B stock that included 10 g of K_2HPO_4 and 10 g of KH_2PO_4. Additionally the RDM contained 4 mL of Biotin (0.25 mg/mL), 1 mL of Thiamine (10 mg/mL), 5 g Sucrose, and distilled water up to 1 L.

2.3. Subcellular Protein Fractionation

This procedure was adopted from Raimunda, et al. (14) with the following modifications. Two 10 mL whole cell samples were taken from both the wild type and mutant cultures. The cells were sonicated three times each at 8 (unit) for 10 sec in order to whole cell lysate. Both the wild type and mutant cultures were divided so that 70 percent of the culture was used to extract the periplasm and 30 percent was used to extract the

cytosol. The periplasmic fraction was obtained by hypo-osmotic shock as described previously by Raimunda, et al. (15). The periplasmic fraction was concentrated down in 3 kDa Millipore filtering devices.

The pellet left over from the periplasmic extraction was resuspended in 50 mM Tris-Cl pH 7.4. The cells were passed through a French press at 20,000 psi three times. The broken cells were centrifuged at 9000 rpm for 30 min, and the resulting supernatant contained the membrane. The supernatant was ultracentrifuged at 33,000 rpm for 1 hr. The supernatant was decanted and the membrane pellet was resuspended in 0.5 mL of 50 mM Tris-Cl and was transferred to a homogenizer. An additional 0.5 mL of the same buffer was added to the final in the obtained membrane fractions.

Cytosolic fractions were collected after incubating the cells with the Tris-sucrose-EDTA buffer and lysozyme, and centrifuging them at 11000 rpm for 20 min. The supernatant was decanted and the pellet was resuspended in 10 mM Tris-Cl pH 8. The cells incubated for 15 min at room temperature. Following this, 4 mM of MgCl2 and DNAase was added to the tubes preceding an incubation of 10 min. These tubes were centrifuged at 11000 rpm for 20 min and the resulting supernatant contained the cytosolic fraction.

2.4. Copper Concentration Determination

The subcellular fractions obtained previously included the whole cell, membrane, periplasmic, secreted, and cytoplasmic were prepared for copper determination using the Atomic Absorption Spectroscopy (AAS). After determining the concentration of protein of each fraction by Bradford assay (16), it was calculated how much of each fraction to add to the AAS prep sample to ensure a final protein concentration of 50 μg. The volume was brought up to 200 μL with deionized water. Another 200 μL of nitric acid was added to the

samples preceding boiling for 1 hr. The samples cooled down and 100 μL of hydrogen peroxide was added. The samples for the AAS contained 100 μL of the protein/acid/peroxide mixture and were brought up to a final volume of 500 μL with deionized water. The samples and standards were placed into the well of the AAS and the machine ran to determine the concentration of copper in each sample.

3. Results

3.1. Bioinformatics

Table 1: List of Cuproproteins and the organism to which they were derived from and used as a template in order to determine homologous proteins in *S. meliloti*.

Enzymes: redox

#	Gene	Function	Gene Number	Organism	References
2	ndh	NADH dehydrogenase 2	299877527	*Escherichia coli*	(17)
4	NC	nitrosocyanin	499423326	*Nitrosomonas europaea*	(18)
6	nirK	Cu containing nitrite reductase	21623660	*Hyphomicrobium denitrificans*	(19)
7	CotA	laccase	1708638	*Bacillus subtilis*	(20)
8	mel	tyrosinase	153523	*Streptomyces antibioticus*	(21)
13	pcoA	multicopper oxidase	619128	*Escherichia coli*	(22)
14	CueO	multicopper oxidase	388476242	*Escherichia coli*	(23)
15	pcoC	copper resistance protein	619130	*Escherichia coli*	(22)
22	ptrA	transcriptional activator	119524024	*Pseudomonas chlororaphis*	de Kievit et al. unpublished

Chaperones

#	Gene Name	Function	Gene Number	Organism	References
1	ctaB	cytochrome c oxidase	1653266	*Synechocystis PCC 6803*	(24)
3	sodC	superoxide dismutase	485718255	*Escherichia coli*	(25)
5	petE	plastocyanin	47402	*Synechocystis PCC 6803*	(26)
17	copZ	copper chaperone	1652317	*Synechocystis PCC 6803*	(27)
18	copC	copper resistance protein	151190	*Pseudomonas syringae*	(28)
19	copD	copper resistance protein	499586270	*Pseudomonas syringae*	(29)
20	mauC	amicyanin	113693	*Paracoccus denitrificans*	(30)
21	CopG	Copper resistance protein	94152520	*Cupriavidus metallidurans*	(31)

Transporters

#	Gene Name	Function	Gene Number	Organism	References
9	PacS	ATPase	2493001	*Synechocystis PCC 6803*	(32)
10	CopA	ATPase	298280327	*Escherichia coli*	(33)
11	CopB	ATPase	290643	*Enterococcus hirae*	(34)
12	cutC	copper homeostasis protein	388477948	*Escherichia coli*	(35)
16	cusCFBA	copper efflux system	378261079	*Escherichia coli*	(36)

Out of the 31 proteins used to determine homology in the *S. meliloti* genome, only 22 resulted in feasible outcomes as listed in Table 1. Seventy-three percent of these results

had more than one homologous protein as shown in Table 2. The proteins that only had one result were superoxide dismutase, nitrosocyanin, tyrosinase, cutC, pcoC, and copD. PcoC and copD are both copper resistance proteins that are involved in the detoxification of the cell during symbiosis. Both these proteins had the same result, SMa1198. The nitrosocyanin protein derived from *Nitrosomonas europaea* is homologous to nitrous oxide reductase, *nosZ*. *NosZ* catalyzes the redox reaction of nitrous oxide. It also has a cox2 domain in its C-terminus, suggesting it is associated with the cytochrome *c* oxidase protein (37).

Table 2. List of cuproproteins in *S. meliloti* genome.

Gene Name	Function	Gene Number	Cellular localization	Homologous # (Table 1)	GI	Reference
ctaB	Protoheme IX farnesyltransferase	SMc00450	Cytoplasmic membrane	1	15964659	(24)
ubiA	4-hydroxybenzoate polyprenyltransferase	SMc00988	Cytoplasmic membrane	1	15964617	(24)
ndh	NADH dehydrogenase transmembrane protein	SMc04452	Cytoplasm	2	15965832	(7)
	dehydrogenase, oxidoreductase FAD flavoprotein	SMb20861	Cytoplasm	2	16264903	(11)
	oxidoreductase	SMc00914	Cytoplasm	2	15964544	(7)
lpdA1	dihydrolipoamide dehydrogenase	SMc01035	Cytoplasm	2	15965203	(7)
sodC	superoxide dismutase Cu-Zn precursor (bacteriocuprein) transmembrane protein	SMc02597	Cytoplasmic membrane	3	15964850	(7)
nosZ	nitrous oxide reductase	SMa1182	Periplasmic	4	16263096	(37)
azu1	pseudoazurin (blue copper protein)	SMa1243	Periplasmic	5, 20	16263131	(12)
azu2	pseudoazurin (blue copper protein)	SMc04047	Periplasmic	5, 20	15966583	(7)
amcY	amicyanin precursor protein	SMb20594	Periplasmic	5, 20	16265254	(11)
	copper oxidase	SMa1041	Periplasmic	5	16263017	(12)
nirK	nitrite reductase	SMa1250	Periplasmic	6, 13	16263134	(12)
	copper-containing oxidoreductase	SMc02282	Periplasmic	6, 7, 13, 14	15964339	(7)
	copper-containing oxidase	SMa1038	Periplasmic	6, 7, 13, 14	16263016	(12)
	oxidoreductase	SMc01754	Periplasmic	7, 13, 14	15964945	(7)
	oxidoreductase	SMc03287	Periplasmic	8	15966886	(7)
FixI2	(E1/E2 type) ATPase	SMa0621	Cytoplasmic membrane	9, 10, 11	16262778	(12)
fixI1	ATPase	SMa1209	Cytoplasmic	9, 10	16263112	(12)

			membrane			
	heavy metal transporting ATPase	SMc04128	Cytoplasmic membrane	9, 10, 11	15963877	(7)
	cation transport p-type ATPase	SMa1163	Cytoplasmic membrane	9, 10	193782642	(12)
	cation transport p-type ATPase	SMa1155	Cytoplasmic membrane	9, 10	16263078	(12)
	cation transport p-type ATPase	SMa1087	Cytoplasmic membrane	11	16263042	(12)
actP	ATPase	SMa1013	Cytoplasmic membrane	11	16263001	(12)
atcU2	copper-transporting P-type ATPase	SMb21578	Cytoplasmic membrane	11	16264766	(11)
cutC	hypothetical protein	SMc02879	Cytoplasm	12	15963957	(7)
	copper protein	SMa1007	Periplasmic	13	16262996	(12)
	copper export protein	SMa1198	Cytoplasmic membrane	15, 18, 19	16263107	(12)
acrE	acriflavin resistance protein	SMb21497	Cytoplasmic membrane	16	16265075	(11)
mexE1	multidrug efflux system transmembrane protein	SMc1094	Cytoplasmic membrane	16	15964180	(7)
NolF	NolF secretion protein	SMa0876	Cytoplasmic membrane	16	16262934	(38)
	efflux protein	SMb20346	Cytoplasmic membrane	16	16264080	(11)
mexE2	multidrug efflux system transmembrane protein	SMc03972	Cytoplasmic membrane	16	15966509	(7)
	heavy metal binding protein	SMa1009	Cytoplasm	17	193782627	(12)
	hypothetical protein	SMb20560	Cytoplasm	17	16264287	(39)
argD	acetylornithine transaminase	SMc02138	Cytoplasm	18	15964270	(7)
	hypothetical protein	SMc00509	Cytoplasmic membrane	18	15965511	(7)
	hypothetical protein	SMb20881	Periplasmic	21	16264923	(11)
	hypothetical protein	SMa1008	Periplasmic	21	193782626	(12)
	transcriptional regulator	SMc02984	Cytoplasmic	22	15966629	(7)
	transcriptional regulator	SMc00831	Cytoplasmic	22	15964529	(7)
	LysR family transcriptional regulator	SMa1602	Cytoplasmic	22	193782691	(12)
	transcriptional activator	SMa1828	Cytoplasmic	22	16263460	(12)
gstR	gstR-like transcriptional regulator protein	SMb20004	Cytoplasmic	22	16263755	(11)

There was only one hit for a superoxide dismutase protein, *sodC*. Superoxide dismutase is also a part of the organism's defense mechanism during symbiosis. The enzyme destroys radicals that have been produced by ROS by the host-plant. It is located within the cytoplasmic membrane (7). The *sodC* protein sequence was aligned with homologous proteins of two different organisms, *E. coli* and *Pseudomondas syringae*. When

these sequences were imputed into the PROSITE database on the ExPASy website (http://prosite.expasy.org), only one domain became prominent. The Cu/Zn superoxide dismutase domain (ref: PS00332) is highlighted in yellow in Figure 4. Within this domain Arg197 and Cys200, highlighted in blue, are important for function. Additionally, copper ions bind to the protein via histidine residues, highlighted in green.

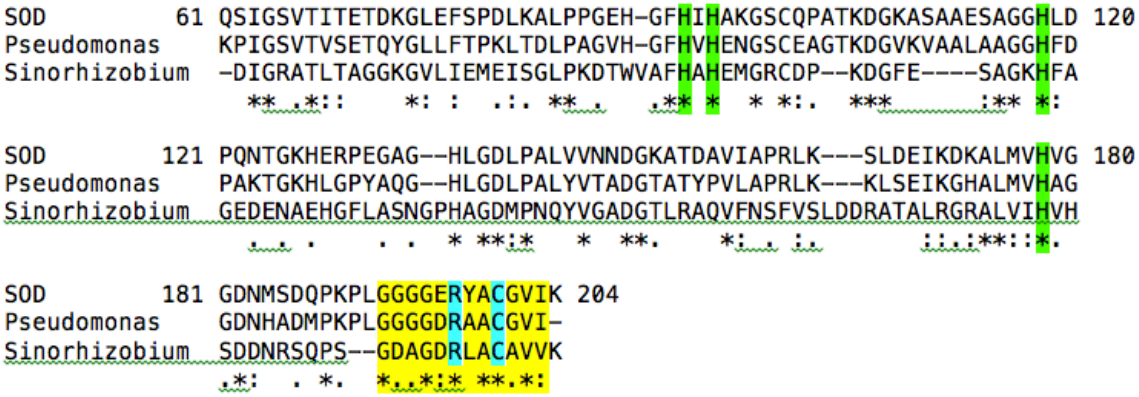

Figure 4: Protein sequence alignment of superoxide dismutase in *Escherichia coli* (SOD), in *Pseudomonas syringae*, and in *S. meliloti*. Copper binding sites are highlighted in green. The Cu/Zn superoxide dismutase domain is highlighted in yellow. A disulfide bond is highlighted in blue.

Significant copper transport proteins, such as pacS, copA, and copB, from organisms *Synechocystis PCC 6803, E. coli, and E. hirae*, respectively, have obtained similar outcomes of homologous proteins. These proteins are known as copper ATPases, and it is known that the *S. meliloti* genome encodes only five. However, the outcomes from the BLAST database resulted in eight possible ATPase genes. Both *pacS* and *copA* had the exact same BLAST results that included *FixI1*, *FixI2*, SMc04128, SMa1163, and SMa1155. Furthermore, *copB* had three other possible outcomes that *pacS* and *copA* did not. These three proteins are SMa1087, *actP*, and *atcU2*. Some of these proteins are labeled as cation or heavy metal ATPases, which can be construed as not only transporting copper but also other metals like zinc, iron, and magnesium.

One characteristic that *copB* lacks compared to *pacS* and *copA* is the heavy metal associated (HMA) domain. On the other hand, it still contains the specific phosphorylation site bearing the DKTGT motif, which characterizes the protein as a P-type ATPase. The heavy metal associated domain is relatively seventy residues long and contains two cysteine residues that bind copper.

There are other proteins that contain the HMA domain, but are not characterized as ATPases. *CopZ* of *Synechocystis PCC 6803* is a chaperone located within the cytoplasm of the bacterial cell. This protein binds free copper and transports it to copper ATPases. Therefore, as a result of copper chaperones and copper ATPases having similar domains suggests they interact with one another. The HMA domain is practically displayed as the entire sequence, including residues two through seventy-four. The alignment of *copZ* and the two *S. meliloti* homologous proteins are shown in Figure 5. The metal binding site is highlighted in green. The copper ion binds to both cys12 cys15. The alignment shows how similar the protein sequences of SMa1009 and SMb20560.

```
copZ          1 MTIQLTVPTIACEACAEAVTKAVQNEDAQATVQVDLTSKKVTITSALGEEQLRTAIASAG 60
SMa1009         M-YHLNIPEMTCGHCVSAVEKAVKSVDPNAKVVVNLEAKTASIDSQAASEAFVAAIEDAG
SMb20560        M-YHLNVEDVTCGHCAATVEKAVKAADPKAKVAVNLEAKTASIESDIGPDVFIAAIEDAG
                *  :*.:  ::*  *. :* ***:  *.:*.* *:* :*..:* *  . : : :** .**

copZ         61 HE----------VE 74
SMa1009         YKASFAKSCCSHVA
SMb20560        YRASFKKSCCSHVD
                :.          *
```

Figure 5: Protein sequence alignment of *copZ* from *Synechocystis sp. PCC 6803* and genes SMa1009 and SMb20560 from *S. meliloti*. The copper binding site is highlighted in green.

The BLAST results of a cytochrome *c* oxidase protein only produced two hits. Proteins *ctaB* and *ubiA* are homologous to cytochrome *c* oxidase in the organism *Synechocystis PCC 6803*. These proteins all encode the same ubiA prenyltransferase family domain, highlighted in yellow. This protein is involved with the transfer of electrons, and

reduces oxygen to water (24). This protein, along with nitrosocyanin, belongs to a super family of proteins called cupredoxins. Other proteins included in this family are plastocyanins (*azu1, azu2, amcY*, and SMa1041), and multicopper oxidases (*nirK*, SMc02282, SMa1038, and SMc01754).

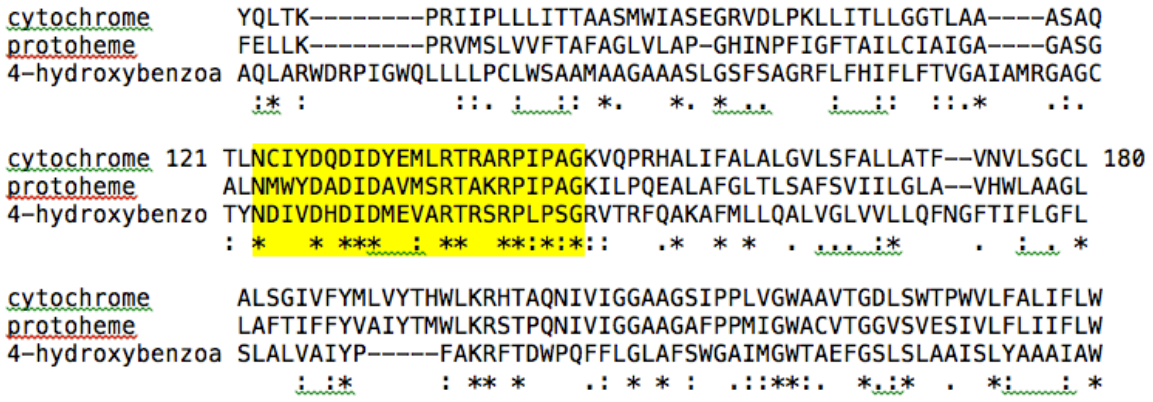

Figure 6: Protein sequence alignment of cytochrome c oxidase from *Synechocystis PCC 6803*, and *ctaB* and *ubiA* in *S. meliloti*. The ubiA prenyltransferase family domain is highlighted in yellow.

Multicopper oxidases are involved with electron transfer within the periplasmic space of the bacterial cell. These proteins can bind copper ions, given three characteristic copper centers, type 1, type 2, and type 3. Type 1 copper centers display an electron absorption band near 600 nm, emiting a blue light. These blue copper proteins, also known as pseudoazurins, genes *azu1* and *azu2*, transfer electrons to and from nitrite reductase (*nirK*). Other homologous proteins encoded in the *S. meliloti* genome include *amcY* and SMa1041. These proteins were derived from plastocyanin (*petE*) of *Synechocystis PCC 6803* and amicyanin (*mauC*) of *Paracoccus denitrificans* (40).

Nitrite reductase (*nirK*) catalyzes the reduction of nitrite to nitric oxide. The copper ion binds to the type 1 center from which nitrite binds to the copper ion. These ions are transferred to the type 2 center where nitrite is reduced. These ions are bound to histidine rich residues, which have positively charged side chains in order to bond with the

negatively charged Cu^{+} ion. This protein contains a plastocyanin domain because it accepts electrons from such proteins. Additionally, this protein accepts electrons from cytochrome *c* oxidase (*ctaB* and *ubiA*) (19).

These multicopper oxidase proteins are derived from nirK of *Hyphomicrobium denitrificans*, *pcoA* and *cueO* of *Escherichia coli*, and *cotA* of *Bacillus subtilis*. It is proposed that these proteins are involved with either translocating cytoplasmic Cu(I) to the periplasmic space via the TAT pathway, or by reducing Cu^{+} ions to Cu^{2+}, which is less toxic for the cell. Additionally, these proteins may be involved with preventing the formation of damaging free radicals due to the ROS species produced by the host plant.

3.2. Copper concentration determination

The concentration of copper in the membrane, periplasmic, cytosolic, and secreted fractions would add up to the concentration of copper in the whole cell fraction. Moreover, it is proposed the concentration of copper of the cytosolic fraction of the copper ATPase mutant strain (Δ*actP*) would be greater than that of the wild type strain (WSM419). Possibly due to the copper ATPases of the mutant strain having a lower enzymatic function, and are incompetent of transporting the copper in the cytoplasm to the periplasmic space.

The WSM419 and Δ*actP* strains were initially grown in TY media. The cells inoculated to an OD_{600nm} of 1.5 and 1.7, respectively. The fractions that were used in determining the copper concentration were the whole cell, membrane, secreted, periplamic space, and the cytoplasm. According to Figure 7, the secreted fraction of the Δ*actP* strain had almost three times as much copper than the whole cell fraction, and the WSM419 strain had at least double the amount. Additionally, there was more copper in the periplasmic space for both strains compared to half that amount in the cytoplasmic

fraction. It is questionable as to why the whole cell fraction and the cytoplasmic fraction have the lowest copper concentrations compared to the other cell fractions in both strains, when they should have the highest.

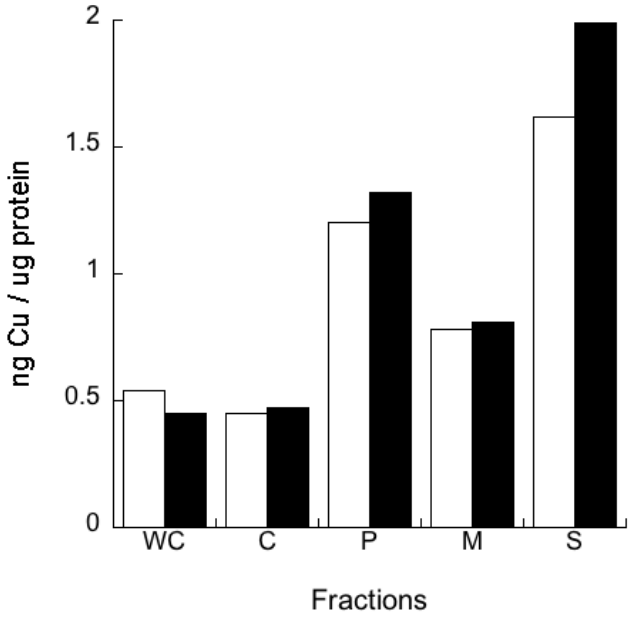

Figure 7: The amount of copper per microgram of protein for each subcellular fraction of the WSM419 (white bars) strain and the ΔactP (black bars) strain. Fractions include whole cell (WC), cytoplasmic (C), periplasmic space (P), membrane (M), and secreted (S). TY media was used to harvest cells.

It was proposed that the media had to do with these backwards values. The media contained a high concentration of protein due to the yeast extract. These proteins can bind to the available copper in the media, causing the higher amount of copper in the secreted fraction than the copper acquired by bacteria. The next step was to harvest the cells in a Rhizobium Defined Medium (RDM), which is a minimal media. The only disadvantage to this new medium is that cells take longer to grow. On the other hand, this had reduced the amount of copper accumulating in the secreted fraction.

The strains used to harvest in the RDM were *S. meliloti* RM2011 wild type and Δ*cut5* mutant strain. The Δ*cut5* cells were inoculated to an OD_{600nm} of 1.3 and the RM2011 cells

were inoculated to an OD_{600nm} of 1.8. The secreted fractions were null and void when preparing the samples for the AAS. This was due to cells growing in the secreted fraction of the Δ*cut5* strain making it extremely difficult to concentrate the protein in a Millipore filtering tube. The cells were pelleting down in the filter clogging the holes. This fraction was centrifuged at 7000 rpm for 10 min in order to pellet down any cells that were left. After this there were still problems trying to concentrate down the protein, even after cleaning the filter. Therefore, it was decided not to use that fraction for the AAS. Likewise, the secreted fraction of the RM2011 strain was unable to be measured by the AAS due to a small concentration of protein. There was not enough protein to make a 50 μg sample for the AAS. Therefore, it was decided not to use the fraction and also in part that there was nothing to compare it to.

The concentration of copper in the RM2011 strain was greater than the Δ*cut5* strain, as shown in Figure 8. This is probably due to the fact there were more wild type cells used in the preparation than the mutant. Furthermore, the concentrations of copper for each fraction of the Δ*cut5* strain are relatively the same values. Compared to the previous figure, the whole cell fraction still has the lowest concentration for the mutant strain. On the other hand, the whole cell fraction of the wild type strain has the highest concentration of copper compared to the other fractions.

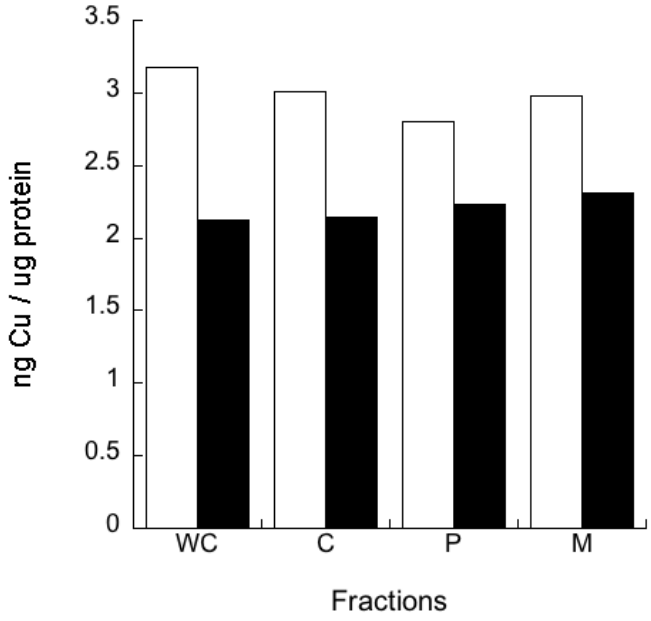

Figure 8: The amount of copper per microgram of protein for each subcellular fraction of the RM2011 (white bars) strain and the Δcut5 (black bars) strain. Fractions include whole cell (WC), cytoplasmic (C), periplasmic space (P), and membrane (M). RDM media was used to harvest cells.

Additionally, the concentrations of copper of the membrane, periplasmic, and cytoplasmic fractions should add up to the concentration of copper of the whole cell fraction, in this case, they do not. It is uncertain as to why these values are very similar.

These results with the RDM media support the hypothesis more than those of the TY media, because the concentration of copper of the cytoplasmic fraction for the RM2011 strain is slightly greater than the periplasmic fraction. However, that is not the case for the Δ*cut5* strain. The concentration of copper of the cytoplasmic fraction is slightly less than the periplasmic fraction. This shows that the other copper ATPases found in *S. meliloti* genome are functioning efficiently by transporting the copper from the cytoplasm to the periplasmic space.

4. Discussion

Symbiotic organism like *S. meliloti* encoded for large array of cuproproteins. These cuproproteins are involved in copper homeostasis during symbiosis and thus has a large impact on the outcome of host-pathogen interaction. It was hypothesized that the majority of the cuproproteins in Table 2 were subcellular localized in the cytoplasm or the periplasm, because free copper mainly exists within those areas of a gram-negative bacterial cell (14). Copper ions are bound to chaperones, like SMa1009 and SMb20560 of *S. meliloti*, which transport the ion to copper ATPases, like *ActP*, bound in the cytoplasmic membrane. This interaction is distinguished during the rhizobia-legume infection process, in which these proteins help detoxify the bacterial cell of copper when exposed to excess amounts.

Additionally, cytochrome c oxidases, like *ctaB* and *ubiA*, are bound in the cytoplasmic membrane and interact with proteins of the cupredoxin family. These proteins include plastocyanins and multicopper oxidases. Cytochromes also interact with proteins that are involved with nitrogen metabolism, *nosZ*. These proteins interact with one another to donate or accept electrons. These proteins work together during the infection process to protect the bacterial cell of damaging radicals from oxidative bursts produced by the plant-hosts innate immune response.

In this experiment of the copper determination in different subcellular fractions, we observed the same amount of copper distributed in the different fractions. It is probable that the copper within the cell had equilibrated due to long process time for the fraction extraction. Additionally, another probable cause could be correlated to the amount of

protein extracted from the different fractions. The protein concentrations could have been low, making it difficult for the AAS to analyze the amount of copper in each cellular fraction.

Other problems that arose were cells continuing to grow in the secreted fraction upon the storage. This made it extremely difficult to concentrate down in the Millipore filtering tubes because the cells were being pelleted down and clogged the filtering holes. For the future, it would probably be best not to store the fraction for long, but to either concentrate the fraction right away, or centrifuge the secreted fraction at higher speed to remove the cell debris from protein fraction.

Further studies of proteomic analysis include 2D electrophoresis and X-ray fluorescence. 2D electrophoresis separates proteins by molecular weight and isoelectric point. This makes it easier to pick out the protein of interest, in this case copper ATPases, or other cuproproteins. Further analysis of the 2D gel with X-ray fluorescence can reveal if the protein of interest is bound to copper.

5. Bibliography

1. Arguello, J. M., Gonzalez-Guerrero, M., and Raimunda, D. (2011) *Biochemistry* **50**, 9940-9949
2. Rademacher, C., and Masepohl, B. (2012) *Microbiology-Sgm* **158**, 2451-2464
3. Festa, R. A., and Thiele, D. J. (2012) *Plos Pathogens* **8**
4. Hodgkinson, V., and Petris, M. J. (2012) *J Biol Chem* **287**, 13549-13555
5. Gourdon, P., Liu, X. Y., Skjorringe, T., Morth, J. P., Moller, L. B., Pedersen, B. P., and Nissen, P. (2011) *Nature* **475**, 59-64
6. Gonzalez-Guerrero, M., Hong, D., and Arguello, J. M. (2009) *The Journal of biological chemistry* **284**, 20804-20811
7. Capela, D., Barloy-Hubler, F., Gouzy, J., Bothe, G., Ampe, F., Batut, J., Boistard, P., Becker, A., Boutry, M., Cadieu, E., Dreano, S., Gloux, S., Godrie, T., Goffeau, A., Kahn, D., Kiss, E., Lelaure, V., Masuy, D., Pohl, T., Portetelle, D., Puhler, A., Purnelle, B., Ramsperger, U., Renard, C., Thebault, P., Vandenbol, M., Weidner, S., and Galibert, F. (2001) *Proc Natl Acad Sci U S A* **98**, 9877-9882
8. Nanda, A. K., Andrio, E., Marino, D., Pauly, N., and Dunand, C. (2010) *Journal of Integrative Plant Biology* **52**, 195-204
9. Perret, X., Staehelin, C., and Broughton, W. J. (2000) *Microbiol Mol Biol Rev* **64**, 180-201
10. Gage, D. J. (2004) *Microbiol Mol Biol Rev* **68**, 280-300
11. Finan, T. M., Weidner, S., Wong, K., Buhrmester, J., Chain, P., Vorholter, F. J., Hernandez-Lucas, I., Becker, A., Cowie, A., Gouzy, J., Golding, B., and Puhler, A. (2001) *Proc Natl Acad Sci U S A* **98**, 9889-9894

12. Barnett, M. J., Fisher, R. F., Jones, T., Komp, C., Abola, A. P., Barloy-Hubler, F., Bowser, L., Capela, D., Galibert, F., Gouzy, J., Gurjal, M., Hong, A., Huizar, L., Hyman, R. W., Kahn, D., Kahn, M. L., Kalman, S., Keating, D. H., Palm, C., Peck, M. C., Surzycki, R., Wells, D. H., Yeh, K. C., Davis, R. W., Federspiel, N. A., and Long, S. R. (2001) *Proc Natl Acad Sci U S A* **98**, 9883-9888

13. Ridge, P. G., Zhang, Y., and Gladyshev, V. N. (2008) *Plos One* **3**

14. Raimunda, D., Padilla-Benavides, T., Vogt, S., Boutigny, S., Tomkinson, K. N., Finney, L. A., and Arguello, J. M. (2013) *Metallomics* **5**, 144-151

15. Raimunda, D., Khare, T., Giometti, C., Vogt, S., Arguello, J. M., and Finney, L. (2012) *Metallomics* **4**, 921-927

16. Bradford, M. M. (1976) *Anal Biochem* **72**, 248-254

17. Young, I. G., Rogers, B. L., Campbell, H. D., Jaworowski, A., and Shaw, D. C. (1981) *Eur J Biochem* **116**, 165-170

18. Lieberman, R. L., Arciero, D. M., Hooper, A. B., and Rosenzweig, A. C. (2001) *Biochemistry* **40**, 5674-5681

19. Yamaguchi, K., Kataoka, K., Kobayashi, M., Itoh, K., Fukui, A., and Suzuki, S. (2004) *Biochemistry* **43**, 14180-14188

20. Borriss, R., Porwollik, S., and Schroeter, R. (1996) *Microbiology* **142 (Pt 11)**, 3027-3031

21. Bernan, V., Filpula, D., Herber, W., Bibb, M., and Katz, E. (1985) *Gene* **37**, 101-110

22. Brown, N. L., Barrett, S. R., Camakaris, J., Lee, B. T., and Rouch, D. A. (1995) *Mol Microbiol* **17**, 1153-1166

23. Grass, G., and Rensing, C. (2001) *J Bacteriol* **183**, 2145-2147

24. Cavet, J. S., Borrelly, G. P., and Robinson, N. J. (2003) *FEMS Microbiol Rev* **27**, 165-181

25. Imlay, K. R., and Imlay, J. A. (1996) *J Bacteriol* **178**, 2564-2571

26. Briggs, L. M., Pecoraro, V. L., and McIntosh, L. (1990) *Plant Mol Biol* **15**, 633-642

27. Banci, L., Bertini, I., Ciofi-Baffoni, S., Su, X. C., Borrelly, G. P. M., and Robinson, N. J. (2004) *J Biol Chem* **279**, 27502-27510

28. Cha, J. S., and Cooksey, D. A. (1991) *Proc Natl Acad Sci U S A* **88**, 8915-8919

29. Koay, M., Zhang, L. Y., Yang, B. S., Maher, M. J., Xiao, Z. G., and Wedd, A. G. (2005) *Inorg Chem* **44**, 5203-5205

30. Zhu, Z., Cunane, L. M., Chen, Z., Durley, R. C., Mathews, F. S., and Davidson, V. L. (1998) *Biochemistry* **37**, 17128-17136

31. Borremans, B., Hobman, J. L., Provoost, A., Brown, N. L., and van Der Lelie, D. (2001) *J Bacteriol* **183**, 5651-5658

32. Badarau, A., and Dennison, C. (2011) *Proc Natl Acad Sci U S A* **108**, 13007-13012

33. Rensing, C., Fan, B., Sharma, R., Mitra, B., and Rosen, B. P. (2000) *Proc Natl Acad Sci U S A* **97**, 652-656

34. Solioz, M., and Odermatt, A. (1995) *The Journal of biological chemistry* **270**, 9217-9221

35. Gupta, S. D., Lee, B. T., Camakaris, J., and Wu, H. C. (1995) *J Bacteriol* **177**, 4207-4215

36. Mealman, T. D., Blackburn, N. J., and McEvoy, M. M. (2012) *Curr Top Membr* **69**, 163-196

37. Holloway, P., McCormick, W., Watson, R. J., and Chan, Y. K. (1996) *J Bacteriol* **178**, 1505-1514

38. Baev, N., Endre, G., Petrovics, G., Banfalvi, Z., and Kondorosi, A. (1991) *Mol Gen Genet* **228**, 113-124

39. Galibert, F., Finan, T. M., Long, S. R., Puhler, A., Abola, P., Ampe, F., Barloy-Hubler, F., Barnett, M. J., Becker, A., Boistard, P., Bothe, G., Boutry, M., Bowser, L., Buhrmester, J., Cadieu, E., Capela, D., Chain, P., Cowie, A., Davis, R. W., Dreano, S., Federspiel, N. A., Fisher, R. F., Gloux, S., Godrie, T., Goffeau, A., Golding, B., Gouzy, J., Gurjal, M., Hernandez-Lucas, I., Hong, A., Huizar, L., Hyman, R. W., Jones, T., Kahn, D., Kahn, M. L., Kalman, S., Keating, D. H., Kiss, E., Komp, C., Lelaure, V., Masuy, D., Palm, C., Peck, M. C., Pohl, T. M., Portetelle, D., Purnelle, B., Ramsperger, U., Surzycki, R., Thebault, P., Vandenbol, M., Vorholter, F. J., Weidner, S., Wells, D. H., Wong, K., Yeh, K. C., and Batut, J. (2001) *Science* **293**, 668-672

40. Adman, E. T. (2011) Pseudoazurin. in *Encyclopedia of Inorganic and Bioinorganic Chemistry*