Defining the Biological Importance and Conservation of Heme Degrading Enzymes

By

Lisa Jeanne Lojek

Dissertation
Submitted to the Faculty of the
Graduate School of Vanderbilt University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in

Microbiology and Immunology

December 16, 2017

Nashville, Tennessee

Approved:

Timothy L. Cover, M.D.

D. Borden Lacy, Ph.D.

Kevin L. Schey, Ph.D.

David W. Wright, Ph.D.

Eric P. Skaar, Ph.D., M.P.H.

ACKNOWLEDGEMENTS

First, and foremost, I need to thank Eric for all of his help and support during the past five years. He has provided an immense amount of both guidance and patience. He has pushed me to grow and expand as both a scientist and a person. Additionally, he put together a fantastic team and environment that nurtures scientific development and a sense of community.

I have enjoyed being friends and collegues with every member of the Skaar lab, and there are some specific lab members who have been essential to my success within the lab. Allison Farrand started both of the projects that I have worked on during my thesis. She was a fantastic mentor who helped me to learn the ins and outs of the projects and helped orient me to the lab when I first started. Allison has been a wonderful collegue and a fantastic friend. Neal Hammer also took me under his wing very early on and quickly became my scientific big brother. Lillian Juttukonda (nee Johnson) joined the lab at the same time I did and we bonded quickly. Lillian is one of the most driven people I have ever met and I found her enthusiasm and work ethic incredibly inspiring. Jake Choby might be the best graduate student I've ever met. He is incredibly committed and has a near eidetic memory for research articles. Lauren Palmer was always there to help regardless of the situation. Whether my question was about science, personal things, or cooking, Lauren always had the right answer and occasionally a cookie. Finally, I have to thank Will Beavers, who has been my go-to for any question pertaining to chemistry and analytical techniques.

In addition to all of the wonderful people within the lab, I also worked with incredible collaborators. Crysten Blaby-Haas and Sabeeha Merchant worked with me on the LFO1 project.

Jennifer Wisecaver and Antonis Rokas performed the phylogenetic analysis for the IsdG family of heme oxygenases. Heather Kroh and Borden Lacy aided me in attempting to determine the crystal structure of LFO1. Wade Calcutt, Kevin Schey, Jessica Moore, Michelle Reyzer, Joel Harp, Don Stec, and Daniel Ryan all contributed to trying to identify the product of LFO1 heme degradation. Wade Calcutt and Kevin Schey aided in helping me learn how to run and analyze mass spectrometry data. Jessica Moore, Michelle Reyzer, and Daniel Ryan all performed MS experiments to try to determine the LFO1 catabolite. Joel Harp attempted to crystalize the LFO1 heme degradation product and Don Stec performed nuclear magnetic resonance on the purified LFO1 catabolite. Brian Michel developed the COP-1 probe, which was provided by Chris Chang. Shu Xu in Larry Marnett's lab trained me to use the oxygraph to determine if LFO1 is an oxygenase. Andy Weiss created the RNA-sequencing heat map figure.

I would like to thank my thesis committee, my thesis committee chair Tim Cover, D. Borden Lacy, Kevin Schey, and David Wright. My committee has provided profound insight and guidance during my thesis that has hugely improved the work that I have done.

I also need to thank my funding sources, without which I would not have been able to perform the experiments in this thesis. This work was supported by grants from the National Institutes of Health, R01AI069233 (Eric P. Skaar), T32GM008554-21 (Lisa Lojek), GM100753 (Crysten Blaby-Haas), from the National Science Foundation's Plant Genome Research Program IOS-1401682 (Jennifer Wisecaver), and the Division of Chemical Sciences, Geosciences, and Biosciences, Office of Basic Energy Sciences of the US department of Energy DEFD02-04ER15529 (Sabeeha S. Merchant).

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

Figure Page

CHAPTER I

INTRODUCTION

Biological importance and conservation of heme

Heme is a small biological molecule with huge significance. Heme utilization is found in all domains of life (Dailey et al., 2017). Across all domains, heme has essential functionality in a number of enzymatic processes. Heme is essential for energy production by respiration and/or photosynthesis in nearly all organisms (Schafer, Purschke, & Schmidt, 1996). In fact, it has been hypothesized that the ability for organisms to respire allowed for the evolution of multicellularity (Pfeiffer, Schuster, & Bonhoeffer, 2001). Without the presence of heme as a core component in energy production, humans may not be here today. In this thesis, I will focus on two organisms: the freshwater green alga *Chlamydomonas reinhardtii* and the Gram-positive bacterium *Staphylococcus aureus*. For clarity, I will focus on these two organisms during the rest of the introduction, and provide some additional context for related organisms as necessary.

Staphylococcus aureus

Staphylococcus aureus is a Gram-positive facultative anaerobic bacterium. *S. aureus* asymptomatically colonizes about 30% of the population (Klevens et al., 2007). However, *S. aureus* can also cause invasive infections in nearly every organ tissue in the body (DeLeo & Chambers, 2009). In order to obtain this level of pathogenicity, *S. aureus* creates a variety of virulence factors, including hemolysins and leukocidins that target red blood cells (RBCs) and cause lysis (Alonzo & Torres, 2014). *S. aureus* can then use resources from these lysed RBCs, like

heme, to facilitate growth within the host. The import of heme into *S. aureus* has been extensively studied *in vivo*, but the degradation of heme *in vivo* is less well studied.

Chlamydomonas reinhardtii

Chlamydomonas reinhardtii is a single-celled green alga. *C. reinhardtii* is a photoautotroph, meaning that it primarily utilizes photosynthesis for energy production. However, if an alternative carbon source is provided under dark conditions, *C. reinhardtii* can still grow using respiration (Glaesener, Merchant, & Blaby-Haas, 2013). One of the primary reasons for the study of *C. reinhardtii* is to better understand the importance and function of tetrapyrroles and their degradation products in photosynthesis and circadian rhythms in the context of a single-celled organism (Grossman, 2000; Rochaix, 1995). These studies have provided a solid foundation to begin to ascertain the role of heme and its byproducts within *C. reinhardtii*. For these reasons, I have included *C. reinhardtii* as a eukaryotic microbe model to better understand how tetrapyrroles are processed and function within the cell.

Meeting cellular heme requirements

In order for cells to meet their quota of heme, two different strategies are employed. First, cells can import heme from the extracellular milieu or, second, cells containing the appropriate genes can biosynthesize heme intracellularly (Fig. 1, reviewed in (Choby & Skaar, 2016)). Both heme acquisition and heme biosynthesis are observed in organisms across all domains of life (Mansuy & Battioni, 2013).

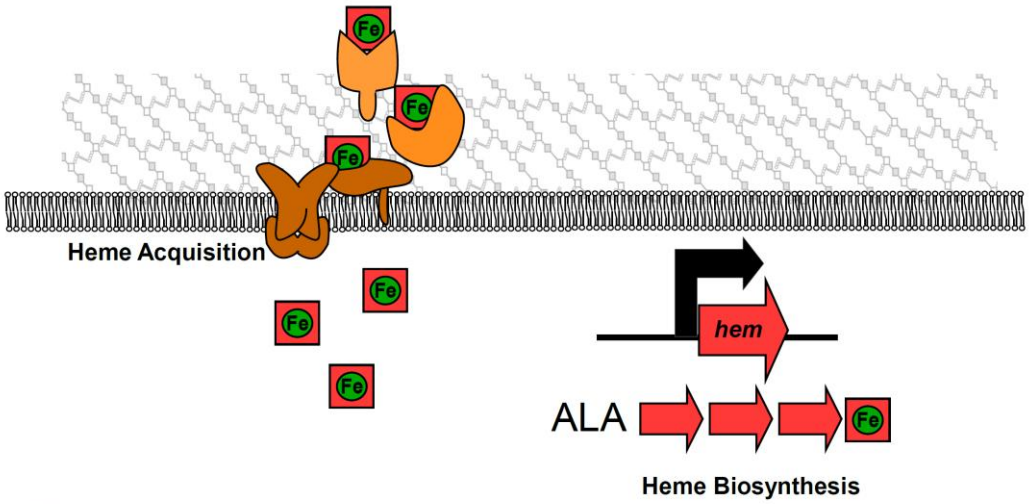


Fig 1. Strategies for meeting cellular heme needs. In order for cells to amass the heme they require, there are two primary mechanisms. First, cells can utilize heme acquisition systems to uptake exogenous heme. Alternatively, cells can biosynthesize the necessary heme endogenously (adapted from Choby, et al. 2016).

Heme acquisition

Heme acquisition in *S. aureus* has been extensively studied. Heme is acquired by *S. aureus* through the iron-regulated surface determinant (Isd) system (Mazmanian et al., 2003). The Isd system is composed of seven protein components that create a high affinity heme import system. Three cell wall anchored proteins are exposed to the extracellular milieu (Mazmanian, Ton-That, Su, & Schneewind, 2002). IsdH binds to haptoglobin-bound hemoglobin, IsdB binds to hemoglobin, and IsdA binds to free heme (Krishna Kumar et al., 2011; Mazmanian et al., 2003; Torres, Pishchany, Humayun, Schneewind, & Skaar, 2006; H. Zhu et al., 2008). Each of these proteins binds to heme or hemoglobin via the near transport (NEAT) domain (Moriwaki et al., 2013). Heme is then transported through the cell wall by IsdC (H. Zhu et al., 2008). The final proteins, IsdDEF, shuttle heme through the cell membrane into the cytoplasm (H. Zhu et al., 2008). The Isd heme transport proteins are expressed from three adjacent operons, all of which are regulated by the ferric uptake regulator (Fur) (Mazmanian et al., 2003). This allows for expression of the heme uptake apparatus only under iron-restricted conditions.

In contrast, far less is known about how *C. reinhardtii* acquires heme. Previous work has shown that exposure to exogenous heme can have profound effects on the cell. This includes increasing growth under low iron conditions and changes in abundance for a variety of transcripts, including heat shock proteins, heme binding proteins, and a heme biosynthesis protein (Duanmu et al., 2012; von Gromoff, Alawady, Meinecke, Grimm, & Beck, 2008; Voss et al., 2011). However, how heme enters *C. reinhardtii* and how it is transported to different organelles within the cell is currently unknown.

Heme biosynthesis

In addition to being able to import extracellularly acquired heme, both *S. aureus* and *C. reinhardtii* can also biosynthesize their own heme. The first committed step of tetrapyrrole biosynthesis begins with the reduction of glutamyl-tRNA by HemA (GtrA, in Gram-positive bacteria) (Fig. 2)(Dailey et al., 2017). Glutamate-1-semialdehyde is then converted to δ-aminolevulinic acid (ALA) by HemL (GsmA) (Fig. 2). The processing of ALA through to coproporphyrinogen III (CPGIII) is conserved in all known cells that biosynthesize heme (Fig. 2) (Dailey et al., 2017). However, Dailey et al. recently discovered that heme biosynthesis in Gram-positive bacteria, specifically Firmicutes and Actinobacteria, differs from that of all other known organisms (Dailey, Gerdes, Dailey, Burch, & Phillips, 2015). Gram-positive organisms oxidize CPGIII to coproporphyrin III (CPIII), which then has an iron atom inserted by a ferrochelatase to produce coproheme III (Fig. 2) (Dailey et al., 2015). Finally, an iron-coproporphyrin decarboxylase cleaves two of the carboxylic acid side chains to vinyl groups, yielding heme (Fig. 2). Heme synthesis in *S. aureus* occurs within the cytoplasm.

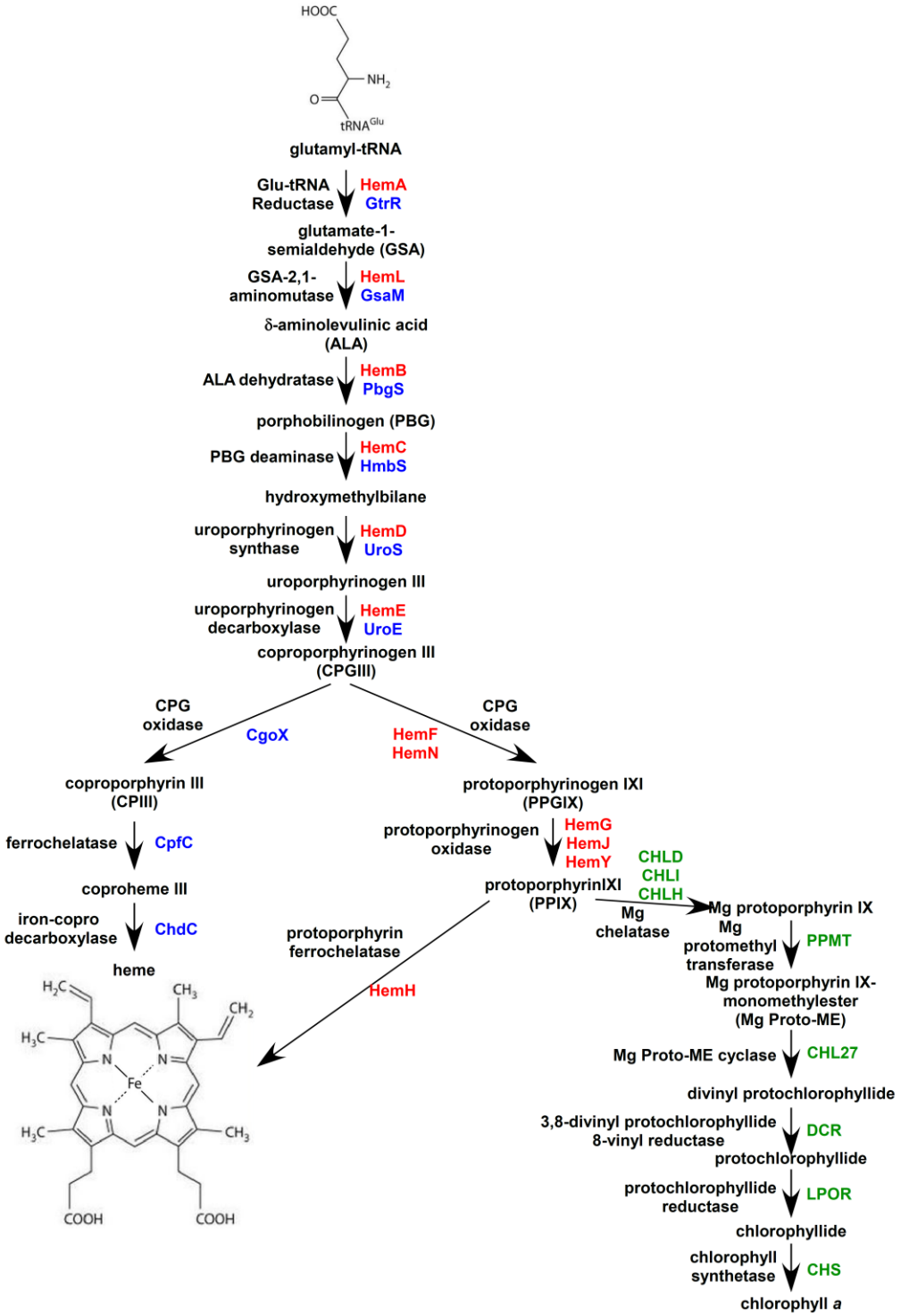


Fig 2. Synthesis pathways for heme and chlorophyll. Tetrapyrrole biosynthesis in both *S. aureus* and *C. reinhardtii* has the same initial steps. However, Gram-positive bacteria have been shown to have divergent heme biosynthesis proteins, which convert coproporphyrinogen oxidase to coprorporphyrin, instead of protoporphyrinogen. In photosynthetic organisms, protoporphyrin IX can have either an iron or magnesium atom inserted into the porphyrin ring. Iron insertion leads to the production of heme, whereas magnesium insertion leads to the chlorophyll synthesis pathway. Enzymes from the classical heme biosynthesis protein are shown in red, enzymes from the non-canonical pathway are in blue, and enzymes in the chlorophyll pathway are in green.

The classical heme biosynthesis pathway, which is utilized by *C. reinhardtii*, progresses from CPGIII to protoporphyrinogen IX (PPGIX) through a coproporphyrinogen oxidase (Fig. 2). PPGIX is then oxidized to protoporphyrin IX (PPIX) via protoporphyrinogen oxidase (Fig. 2) (van Lis, Atteia, Nogaj, & Beale, 2005). Finally, a ferrochelatase inserts an iron atom into the center of the molecule resulting in heme (Fig. 2) (van Lis et al., 2005). Additionally, *C. reinhardtii* also biosynthesizes chlorophyll, which utilizes PPIX and inserts an magnesium ion, instead of iron, before additional processing steps modify the side chains on the tetrapyrrole ring (Fig. 2). Heme synthesis in *C. reinhardtii* occurs solely within the chloroplast (van Lis et al., 2005). This is distinct from heme biosynthesis in higher plants, such as *Pisum sativum*, which have two sets of protoporphyrinogen oxidases and ferrochelatases (Smith, Marsh, & Elder, 1993). One of each protein is localized to the chloroplast and the other to the mitochondria. This allows the initial phases of heme biosynthesis to occur within the chloroplast, however PPGIX is then transferred into the mitochondria, where the final steps of heme biosynthesis can occur. However, since heme synthesis only happens within one organelle in *C. reinhardtii*, heme must be transported into both the cytosol and mitochondria so the heme binding proteins within those compartments can utilize heme (Hoober, 1989; van Lis et al., 2005). The mechanism by which heme is exported out of the chloroplast, transported selectively to the compartment it is required in, and populates heme binding proteins within *C. reinhardtii* is currently unknown.

Heme biosynthesis is regulated in many organisms by modulating the amount of HemA that is present in the cell. In *C. reinhardtii*, heme negatively regulates *hemA* transcript levels in the cell (Chang, Wegmann, & Wang, 1990). Though the mechanism by which heme regulates *hemA* transcript levels is not currently known, it is predicted to be mediated through the binding of a

transcription factor to a putative plasmid response element (PRE) sequence found in the promoter of the *hemA* gene (von Gromoff et al., 2008; Voss et al., 2011). Additionally, the genome uncoupled 4 (GUN4) protein has been shown to affect expression of nearly every gene involved in tetrapyrrole synthesis in *C. reinhardtii* leading to a nearly two-fold increase in the amount of heme in cells with a *gun4* mutation (Brzezowski et al., 2014; Tarahi Tabrizi, Sawicki, Zhou, Luo, & Willows, 2016). It is unclear whether these two mechanisms of heme biosynthesis regulation are related.

How heme biosynthesis is regulated in *S. aureus* is unknown. However, *S. aureus* contains an additional gene within the heme biosynthesis operon, *hemX*, that may be important for heme biosynthesis regulation (Johansson & Hederstedt, 1999). HemX has been shown in *Bacillus subtilis* to post-transcriptionally regulate HemA protein abundance, however the mechanism of action for this is not yet known (Schroder, Johansson, Rutberg, & Hederstedt, 1994). Additionally, it has not yet been determined as to whether HemX has this same role in *S. aureus*.

Heme utilization

There are wide varieties of cellular proteins that require heme in order to function within the cell. One of the main roles of cellular heme is as a cofactor for energy production, which will be more extensively discussed in a following section (Schafer et al., 1996). In addition to energy production, members of enzyme families, such as oxygenases, cytochromes, and peroxidases require heme for their functionality (Li & Poulos, 1994). Heme is also an essential component of catalase in most organisms (Diaz, Loewen, Fita, & Carpena, 2012).

The role of heme in energy production

Heme is an essential part of the electron transport chain, where it acts as an electron acceptor within cytochromes (Schafer et al., 1996). In respiration, the donation of electrons from heme to the terminal electron acceptor, either oxygen (aerobic) or alternative electron acceptors such as nitrite (anaerobic), is the final step. This is true regardless of whether respiration in occurring within the bacterial membrane or in the mitochondria (Schafer et al., 1996). Additionally, heme also binds to cytochromes within the chloroplast, where it acts to transport electrons between photosystems during photosynthesis (Howe & Merchant, 1994).

S. aureus encodes two terminal oxidases, CydAB and QoxABCD (Hammer et al., 2013). Both of these cytochromes bind heme and facilitate electron transfer from menaquinone to the terminal electron acceptor as a component of respiration. When cells cannot produce heme, *S. aureus* forms a small colony variant (SCV) phenotype (von Eiff et al., 1997). Since the cytochromes require heme as a cofactor for their functionality, they are no longer able to function under these conditions. In order to continue to produce energy, the cells switch from respiration to fermentation. However, fermentation yields only two ATP compared to the 36 ATP generated during aerobic respiration (Jurtchuk, 1996). This leads to a significant decrease in growth of SCV strains compared to respiration competent strains.

C. reinhardtii can produce energy through two different mechanisms. When exposed to light, *C. reinhardtii* under goes photosynthesis, however under dark conditions *C. reinhardtii* can utilize aerobic respiration within its mitochondria if they are grown in the presence of acetate as a carbon source (Glaesener et al., 2013; Terauchi, Peers, Kobayashi, Niyogi, & Merchant, 2010). Both of these methods of energy production require heme in order to proceed. Respiration within

the mitochondria of *C. reinhardtii* is conceptually similar to aerobic respiration in bacteria. However, photosynthesis is a vastly different process. Photons of light are captured by light harvesting complex proteins (LHCIIs) which then activate the chlorophyll-protein complexes photosystem II (PSII) or photosystem I (PSI) within the membrane of the thylakoid within the chloroplast (Alberts, Johnson, Lewis, & al., 2002). PSII then uses the energy from the photons to transfer electrons from water to plastoquinone, in what is known as the linear electron flow (LEF) pathway (Finazzi, 2005). From there, electrons are further transferred to the cytochrome b_6f complex and plastocyanin, before they reach PSI. PSI then continues the transfer of electrons to ferrodoxin. Ferrodoxin then is reduced by the ferrodoxin NADPH oxidoreductase to produce NADP+. PSI can also undergo a cyclical electron flow pathway, where it passes electrons back and forth with the cytochrome b_6f complex (Finazzi, 2005). The cytochrome b_6f complex contains eight molecules of heme, and the entirety of both photosystem complexes requires 30 atoms of iron in order to facilitate energy production (Blaby-Haas & Merchant, 2013; Hurt & Hauska, 1981).

Heme as an iron source

In addition to being essential as a cofactor for cellular proteins, heme can also be used an iron source for the cell. Both *S. aureus* and *C. reinhardtii* display a significant decrease in growth under iron deplete conditions (Duanmu et al., 2012; Reniere & Skaar, 2008). However, addition of exogenous heme allows for an increase in growth under these low iron conditions. *C. reinhardtii* also specifically degrades the cytochrome b_6f complex, in addition to PSI and PSII, when in low iron conditions in order to recycle iron so that it can be preferentially used in other proteins, such

as an iron utilizing superoxide dismutase (Fig. 3) (Page et al., 2012). The next section will discuss how these organisms harvest iron from the tetrapyrrole ring.

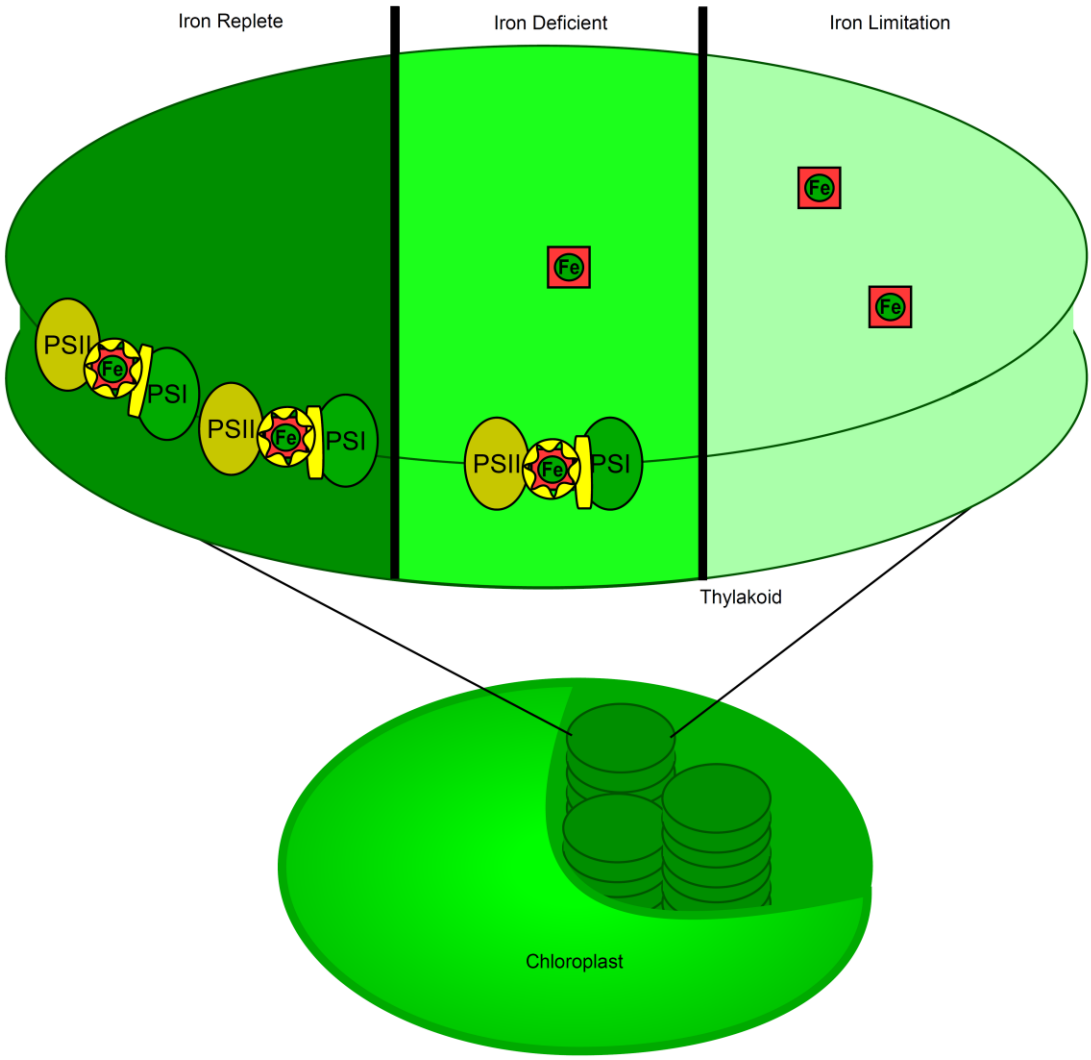


Fig. 3 *C. reinhardtii* response to iron limitation. The photosynthetic machinery requires 30 atoms of iron and 8 hemes in order to function. When grown in iron replete conditions *C. reinhardtii* has abundant iron atoms to facilitate photosynthesis. However, if *C. reinhardtii* experiences iron deficiency or iron limitation, the photosystem complexes and cytochrome b_6f complex are degraded to allow for iron recycling.

Heme degradation

In order to access the iron within heme, cells encode heme-degrading enzymes. Two different types of enzymes have been identified that can perform this role: the heme oxygenases, which use oxygen to facilitate degradation, and heme degrading enzymes in anaerobic bacteria, which rely on Fe-S clusters to facilitate radical catalysis (LaMattina, Nix, & Lanzilotta, 2016; Wilks & Heinzl, 2014). Heme oxygenases are widely conserved and are found extensively throughout eukaryotic and bacterial organisms. However, the oxygen independent heme degrading enzymes, such as ChuW form *Escherichia coli*, have only recently been discovered (LaMattina et al., 2016).

Heme oxygenases

Heme oxygenases are heme degrading enzymes that utilize oxygen as an electron acceptor from the central iron atom in heme facilitating the cleavage of the tetrapyrrole ring (Wilks & Heinzl, 2014). This heme degradation reaction results in the production of catabolites and the release of free iron. There are two characterized families of heme oxygenases. These families are differentiated by their structural characteristics, the catalytic residues found in the binding pocket, and the resulting catabolite(s) they produce from degrading heme (Fig. 4).

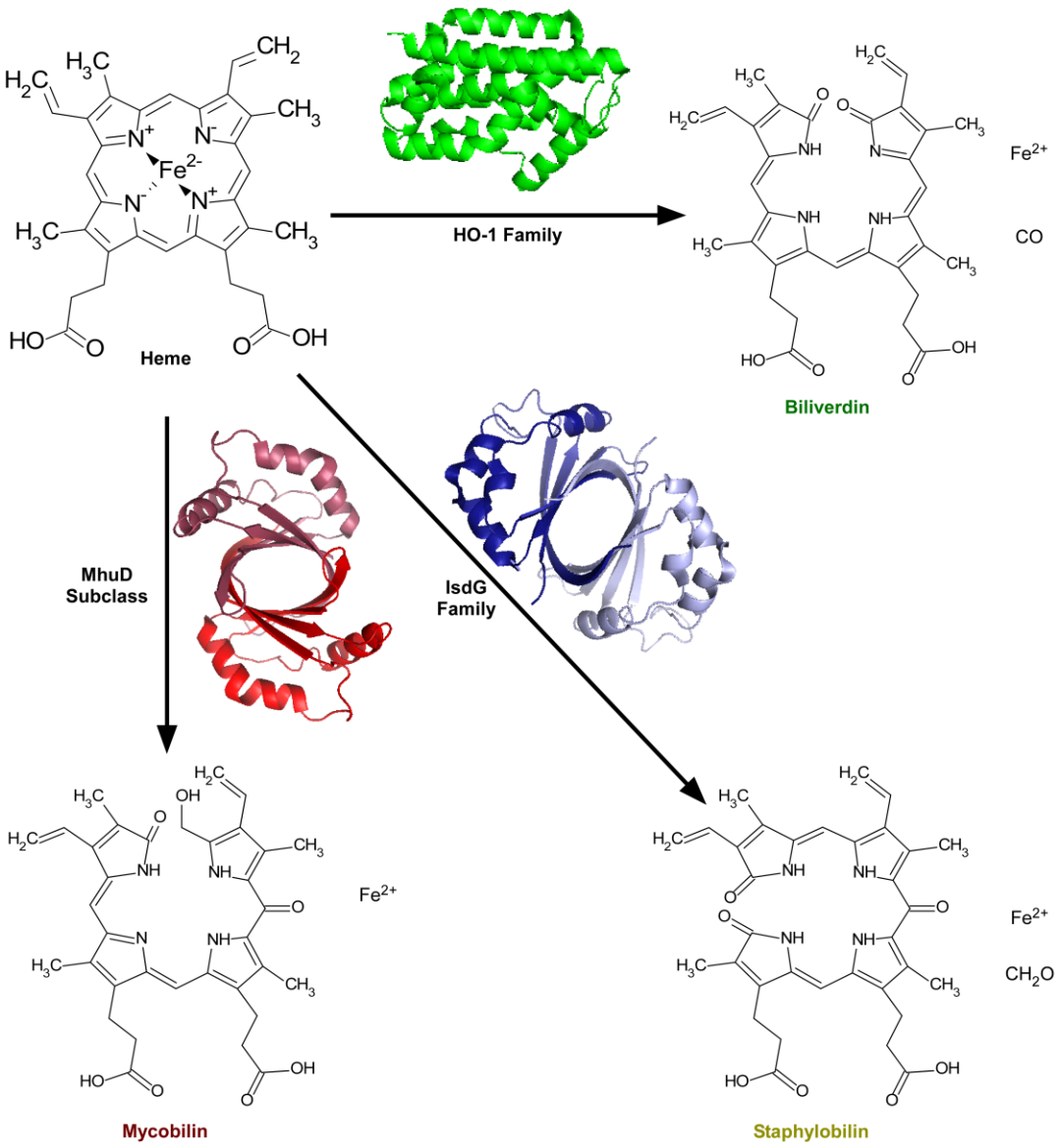


Fig 4. Heme oxygenase families. There are two families of heme oxygenases: the HO-1 family and the IsdG family. HO-1 family heme oxygenases degrade heme to produce biliverdin, carbon monoxide, and free iron. The representative HO-1 family member shown is HmuO from *Corynebacterium diphtheriae*. The IsdG family heme oxygenases degrade heme to staphylobilin, formaldehyde, and free iron. The IsdG family member shown is IsdI from *S. aureus*. MhuD is a subfamily of the IsdG family of heme oxygenases, which degrade heme to mycobilin and free iron. The MhuD enzyme shown is from *Mycobacterium tuberculosis*. Both IsdI and MhuD are homodimeric proteins, each monomer is shown in a different color. The color of the name for each product represents the color of the product.

The HO-1 family of heme oxygenases was the first to be identified and is found throughout eukaryotes and bacterial species (Nakajima, 1963; Wilks & Schmitt, 1998). HO-1 family enzymes act as monomers and are comprised of 10 α-helices (Fig. 3) (Wilks, 2002). Catalysis for HO-1 family heme oxygenases relies on a conserved histidine to degrade heme (Wilks, Ortiz de Montellano, Sun, & Loehr, 1996). Heme degradation by the HO-1 family heme oxygenases results in the products biliverdin and carbon monoxide (Fig. 3) (Wilks & Heinzl, 2014). Additionally, heme is a colored molecule, and all if its degradation products are also chromophores. While heme is rust colored, biliverdin is a vibrant green (Fig. 3).

The second family of heme oxygenases was discovered in *Staphylococcus aureus* in 2004 (Skaar, Gaspar, & Schneewind, 2004). This family, known as the IsdG family, has been predicted to only include 22 proteins, with a small subset of those being functionally characterized (Duong et al., 2014; Haley, Janson, Heilbronner, Foster, & Skaar, 2011; Nambu, Matsui, Goulding, Takahashi, & Ikeda-Saito, 2013; Park, Choi, & Choe, 2012; Puri & O'Brian, 2006; Skaar, Gaspar, et al., 2004; Skaar, Gaspar, & Schneewind, 2006). While the IsdG family enzymes are only found in a small number of bacterial species, the clades that these organisms fall into are incredibly diverse. IsdG enzymes are found in Gram-positive Firmicutes and Actinobacteria, as well as Gram-negative Proteobacteria (Haley et al., 2011). Additionally, IsdG family heme oxygenases are found in both pathogenic and environmental bacterial species (Haley et al., 2011). IsdG family heme oxygenases have a secondary structure distinct from that of the HO-1 family (α–, β–, α–, α–, β–, β–, α–) and form a homodimer, with each monomer binding and degrading one molecule of heme (Fig. 3, each monomer is shown in a different shade) (Wu et al., 2005). Within the binding pocket of IsdG family enzymes are a catalytic triad of amino acids: asparagine,

tryptophan, and histidine, all of which have been shown to be important for heme degradation by these enzymes (Wu et al., 2005). The IsdG family enzymes degrade heme to staphylobilin (yellow) and formaldehyde (Fig. 3) (Matsui et al., 2013; Reniere et al., 2010).

More recently, a third heme oxygenase enzyme was identified in Mycobacterial species. MhuD contains the same secondary structure and catalytic residues as the IsdG family members, however, MhuD degrades heme to only one catabolite, mycobilin (red) (Fig. 3) (Chim, Iniguez, Nguyen, & Goulding, 2010; Matsui et al., 2016; Nambu et al., 2013).

Heme oxygenase regulation

C. reinhardtii encodes two heme oxygenases, HMOX1 and HMOX2, both of which are HO-1 family enzymes. HMOX1 is localized to the chloroplast whereas HMOX2 is located in the cytoplasm (Duanmu et al., 2012). RNA-Sequencing analysis show that both transcripts are constitutively expressed under varying iron levels (Glaesener et al., 2013). Additionally, both *hmox1* and *hmox2* are controlled in a circadian manner (Glaesener et al., 2013).

S. aureus also contains two heme oxygenases, which are expressed in the cytoplasm (Skaar, Gaspar, et al., 2004). Both IsdG and IsdI have similar structures and also degrade heme to the same products, staphylobilin and formaldehyde (Matsui et al., 2013; Reniere et al., 2010; Wu et al., 2005). The *isdG* gene is located within the *isdCDEFsrtBisdG* operon, which is regulated by Fur, resulting in the operon being cotranscribed under low iron conditions (Reniere & Skaar, 2008; Skaar, Gaspar, et al., 2004). *isdI* is within an operon with *orfX*, a gene whose protein functionality is not yet understood (Skaar, Gaspar, et al., 2004). The *orfX* gene contains a Fur binding site, known as a Fur box, within its C-terminus, which leads to *isdI* expression only under iron limiting

conditions (Reniere & Skaar, 2008). While both genes are expressed only during low iron conditions, there are differences in their regulation at the post-translational level. IsdG requires heme binding in order for it to maintain its stability (Reniere & Skaar, 2008). However, in the absence of heme, IsdG is targeted for degradation through cleavage of a flexible loop, which is not contained in IsdI (Reniere, Haley, & Skaar, 2011). Interestingly, deletion of *isdG* and *isdI* have differential effects in the context of pathogenesis. The deletion of both *isdI* and *isdG* lead to a significant decrease in colony forming units (CFUs) within the hearts of mice systemically infected with *S. aureus* compared to a wildtype strain (Reniere & Skaar, 2008). However, only *S. aureus* lacking *isdG* has a decrease in CFUs within the kidneys in the context of systemic infection compared to wildtype (Reniere & Skaar, 2008).

The role of heme degradation products

The products of HO-1 family heme oxygenases are critically important in eukaryotic cells and organisms. In *C. reinhardtii*, biliverdin is produced from heme degradation by both HMOX1 and HMOX2 (Duanmu et al., 2012). Biliverdin in the chloroplast of *C. reinhardtii* is reduced to phycocyanobilin by the biliverdin reductase PCYA (Duanmu et al., 2012). *C. reinhardtii* cells lacking *hmox1* have significantly attenuated growth and less chlorophyll accumulation under photoautotrophic conditions, where they are dependent on photosynthesis for energy production (Duanmu et al., 2012). However, cells lacking *hmox2* had no noticeable phenotypes (Duanmu et al., 2012). Carbon monoxide also has important roles in *C. reinhardtii*. Treating cells with exogenous carbon monoxide leads to an increase in growth rate and chlorophyll abundance in iron starved *C. reinhardtii* (Liping, Hongbo, Xiaohua, & Zhaopu, 2013). This is due to carbon

monoxide treatment leading to an increase in transcription of genes associated with iron acquisition (Liping et al., 2013).

While the role of the HO-1 family heme catabolites have been fairly well established in eukaryotic organisms, very little is known about the functions of the degradation products in bacteria. Within many photosynthetic and some nonphotosynthetic bacteria that contain HO-1 family heme oxygenases, biliverdin can either be modified or directly incorporated into bacteriophytochromes. These protein-phytochrome complexes are important for a variety of functions within the cell, including motility, photosynthetic regulation, stress reponses, virulence, and the production of light absorbing pigments which limit negative effects of light exposure (Purcell & Crosson, 2008). However, the function of biliverdin in bacteria that do not contain bacteriophytochromes and the role of carbon monoxide has yet to be determined. Furthermore, the functional role of the IsdG family heme catabolites inside the cell is entirely unknown. Mice infected with strains of *S. aureus* lacking heme oxygenase activity have reduced bacterial colonization within their livers, compared to mice infected with wildtype bacteria (Reniere & Skaar, 2008). Additionally, *S. aureus* prefers heme as an iron source, which implies that the products of heme degradation may have important cellular roles (Skaar, Humayun, Bae, DeBord, & Schneewind, 2004).

Conclusions

Heme plays important physiological roles in both *C. reinhardtii* and *S. aureus*. Heme homeostasis must be maintained within the cell so that the enzymes requiring heme are able to function. Additionally, the heme degrading enzymes, known as heme oxygenases, are important

for cells that utilize heme as an iron source or in iron recycling, as well as in the production of catabolites which also have important cellular functions. Two families of heme oxygenases have been described, however the IsdG family has only been shown to consist of 22 proteins (Haley et al., 2011). Intriguingly, these proteins are found in very divergent bacterial clades. One aim of this thesis is to determine the conservation of the IsdG family of heme oxygenases to elucidate if the family is larger than previously reported. In addition to interrogating the organisms that contain IsdG family members, this thesis will also seek to identify what role these enzymes have within one specific species, *S. aureus*, and see how their regulation affects heme homeostasis.

A version of the following section (*Chapter II, Chlamydomonas reinhardtii LFO1 is an IsdG family heme oxygenase*) was previously published in *mSphere* 2(4) e00176-17 (eCollection 2017 Jul-Aug) | doi:10.1128/mSphere.00176-17.

CHAPTER II

CHLAMYDOMONAS REINHARDTII LFO1 IS AN ISDG FAMILY HEME OXYGENASE

Introduction

Heme is essential for a variety of cellular processes, including serving as an enzymatic cofactor for catalase and acting as an electron acceptor of the electron transport chain (George, 1948; Igo, Mackler, & Duncan, 1961; Maehly, 1952; Morrison & Stotz, 1955; Tenhunen, Marver, & Schmid, 1968). Additionally, many cells express enzymes to degrade heme. These enzymes, known as heme oxygenases, utilize oxygen to cleave the porphyrin ring of heme to release free iron and secondary catabolites. While many organisms encode heme oxygenases, there is diversity in enzyme structure and catabolite production that may reflect a diversity of function. The first identified family of heme oxygenases is the HO-1 family, which is found in both eukaryotic and bacterial cells (Maines, 1997; Ratliff, Zhu, Deshmukh, Wilks, & Stojiljkovic, 2001; Schuller, Wilks, Ortiz de Montellano, & Poulos, 1999; Tenhunen et al., 1968; Trakshel, Kutty, & Maines, 1986; Wilks & Schmitt, 1998). The HO-1 family degrades heme to biliverdin, carbon monoxide, and free iron (Tenhunen et al., 1968). A second family of heme oxygenases, known as the IsdG family, degrades heme to staphylobilin, formaldehyde, and free iron (Matsui et al., 2013; Reniere et al., 2010). Recent work identified MhuD, a protein with significant secondary structure similarities and conserved catalytic residues with IsdG family members (Chim et al., 2010). However, MhuD degrades heme to distinct products, namely, mycobilin and free iron (Nambu et al., 2013). Finally, an oxygen-independent heme-degrading enzyme, ChuW, was discovered in

Escherichia coli. ChuW uses radical catalysis to degrade heme to the small molecule anaerobilin and free iron (LaMattina et al., 2016).

The IsdG family has only been characterized in a limited number of bacterial species from diverse lineages (Haley et al., 2011). These include the Gram-positive Firmicutes (Duong et al., 2014; Haley et al., 2011; Park et al., 2012; Skaar, Gaspar, et al., 2004; Skaar et al., 2006) and Actinobacteria (Nambu et al., 2013), and the Gram-negative α-Proteobacteria (Puri & O'Brian, 2006). All of the enzymes within this family have conserved secondary structures and catalytic residues, and contain an Antibiotic Biosynthesis Monooxygenase (ABM) domain (Matsui et al., 2016; Wu et al., 2005). Additionally, many of the IsdG family heme oxygenases are also selectively expressed under low iron conditions. Among the seven experimentally characterized proteins within the IsdG family, the amino acid sequence identities range from 20-60% (Haley et al., 2011). Phylogenetic analysis has revealed 22 unique IsdG family heme oxygenases, all within bacterial species (Haley et al., 2011). However, due to the low sequence identity between the IsdG family members and the diversity of the bacteria in which they were identified, I hypothesized that the IsdG family extends beyond this small number of species.

Previous work in *Chlamydomonas reinhardtii* identified a gene, Cre07.g312300, that is more highly expressed in iron-limited conditions compared to iron-replete conditions (Urzica et al., 2012). Analysis of the putative protein encoded by this gene identified an ABM domain and specific secondary structure motifs. Alignment of these structural characteristics with heme oxygenases led us to hypothesize that this putative protein, which I have named LFO1 for Low Fe-responsive Oxygenase 1, belongs to the IsdG family of heme oxygenases. Here I describe the characterization of LFO1 as a functional heme oxygenase that requires catalytic residues

conserved among IsdG family members for activity. Degradation of heme by LFO1 leads to a product with unique chromatographic properties compared to other identified heme catabolites. Using the LFO1 protein sequence, I queried the entirety of known protein sequences for additional IsdG family heme oxygenases. Through this analysis, I identified 866 unique proteins from across the tree of life, significantly expanding the taxonomic distribution of the IsdG family.

Materials and methods

Bacterial Strains and Growth Conditions-Immunoblotting, transformations, plasmid purification, and subcloning were carried out as described previously (Green & Sambrook, 2012). α-His, Rabbit polyclonal antibody was purchased from Santa Cruz. *Escherichia coli* strain DH5α was used for DNA manipulation and BL21 DE3 was used for protein expression of 6His-LFO1, 6His-IsdI, 6His-MhuD, and 6His-HmuO. The production of BL21 DE3 pET15b.*isdI* was previously described (Skaar, Gaspar, et al., 2004). Ampicillin (100 μg/mL) and chloramphenicol (34 μg/mL) were added to the media as required. All strains were incubated at 37°C with shaking.

Construction of Vectors-To create vectors to be used in expression *Lfo1* was codon optimized for expression in *E. coli* and was synthesized into pUC (Genscript). *Lfo1* was cleaved from pUC using BamHI-HF and NdeI (New England Biolabs) and ligated into pET15b.

Expression and Purification of LFO1, IsdI, MhuD, and HmuO-*E.coli* BL21 (DE3) pREL strains containing pET15b.*Lfo1*, pET15b.*isdI*, pET15b.*mhuD*, or pET15b.*hmuO* were grown overnight at 37°C in Terrific broth containing 100 μg/mL ampicillin and 34 μg/mL chloramphenicol. The cells were subcultured into fresh medium and grown at 37°C to mid-log phase, an OD_{600} of 0.4-0.6. Upon reaching mid-log growth, the target genes were induced using 0.1 mM isopropyl-1-thiol-D-galactopyranoside (IPTG). Cell growth was continued for approximately 16 hours at 16°C. Cells were harvested by centrifugation (6,000 x *g* for 15 minutes). Cells were resuspended in lysis buffer (8 M urea in phosphate buffered saline (PBS) with 5 mM (dithiothreitol) DTT and one Pierce Protease Inhibitor Tablet (Thermo Scientific)) and mixed at room temperature for 30 minutes.

Cells were homogenized using a Dounce homogenizer and passed through an EmulsiFlex (Avestin) three times at 20,000 psi. After lysis, lysates were mixed at room temperature for 30 minutes. Lysate was centrifuged at 40,000 x *g* for one hour and filtered with a 0.22 μM filter. The filtered lysate was added to Ni-NTA agarose beads and was incubated at room temperature with rotation for 30 minutes. The lysate and beads were poured into a gravity column. The column was washed with lysis buffer with stepwise decreases in urea, from 8 – 0 M. The column was then washed with 10 column volumes of 5 mM DTT with 100 mM imidazole in PBS. Proteins were eluted with 400 mM imidazole in lysis buffer in 5 column volumes. Proteins were used immediately after purification.

Absorption Spectroscopy-All absorption spectra were obtained using a Varian Cary 50BIO. Hemin from bovine was procured from Sigma Aldrich, and resuspended in 0.1 M sodium hydroxide. Hemin binding studies were performed as previously reported (Wilks & Schmitt, 1998). Specifically, aliquots of hemin from 0.5-30 μM were added to the sample cuvette, containing 10 μM LFO1 and 1 mL of Tris-buffered saline (TBS), and reference cuvettes, containing 1 mL of TBS, at room temperature. Each sample was mixed and allowed to incubate in the dark for 5 minutes, then the spectrum was collected at 300 – 800 nm with 10 nm steps.

Hemin Degradation Assays-Hemin degradation assays were performed as previously described (Skaar, Gaspar, et al., 2004; W. Zhu, Wilks, & Stojiljkovic, 2000). *Reaction with ascorbate*: Specifically, 10 μM LFO1-hemin in TBS were incubated with ascorbate at a final concentration of 10 mM (Ratliff et al., 2001). The spectral changes between 300 – 800 nm were recorded every

10 minutes. The products were extracted and subjected to HPLC as described below. *Reaction in the presence of catalase*: Purified recombinant catalase from *Aspergillus niger* (Sigma), was added to all reaction cuvettes at a ratio of catalase:hemoprotein equal to 1:2 immediately before the addition of ascorbate.

Oxygen Incorporation-The consumption of oxygen was determined by using a Gilson model 5/6 oxygraph (Gilson Medical Electronics, Inc. Middleton, WI) set at 24°C equipped with a Clark electrode and a thermostated cuvette. Triplicate samples of LFO1, IsdI, or no protein at 10 μM were added to TBS with 10 μM heme and 5 μM catalase. 1 mL of each sample was added to the oxygraph and oxygen consumption was measured for 1 min. Degradation was then initialized by adding 10 mM ascorbate and oxygen uptake was measured for 4 additional minutes. The velocity of the oxygen consumption was determined from the first 1 min after ascorbate addition, which were all within the linear portion of the curve. The rate of triplicate samples were averaged and the averages and SEM were graphed.

HPLC of the Heme Degradation Reaction Product-HPLC analysis was performed as previously described (Reniere et al., 2010). Specifically, analysis was performed using 95% water/5% acetonitrile with 0.1% trifluoroacetic acid (TFA) as the mobile phase with a flow rate of 0.5 mL/min on a Microsorb-MV C-18 column. After a 10 minute equilibration period a linear 40 min acetonitrile gradient (5%-80%) was employed and the final concentration was maintained for an additional 20 min. The eluant was monitored using a photodiode array detector from 200–900 nm by reverse-phase chromatography on a Varian ProStar HPLC.

Mass spectrometry methods- Mass spectrometry methods were performed similarly to previously published methods (Reniere et al., 2010). Dried samples were resuspended in 1:1 water:acetonitrile. Resuspended samples were then separated by HPLC, then analyzed by ESI-MS/MS on a Thermo Fisher Orbitrap Elite, as previously published.

COP-1 Fluorescence Detection-COP-1 was utilized as previously described (Michel, Lippert, & Chang, 2012). HmuO, IsdI, and LFO1 at 10 μM were pre-incubated with 10 μM hemin for 5 minutes. An aliquot of 200 μL of the hemin-bound proteins were transferred into a black well 96-well plate. Reactions were incubated with 1 μM COP-1 and background fluorescence at 510 nm was measured. A final concentration of 1 mM ascorbate was then added to each sample and incubated for 5 minutes before measuring fluorescence at 510 nm every 2 minutes for 30 minutes.

Heme Oxygenase Phylogenetic Analysis-The 6 functionally characterized IsdG protein sequences from bacteria were queried against the UniProt reference proteomes using phmmer, a member of the HMMER3 software suite (Finn, Clements, & Eddy, 2011) (web server accessed March 27, 2017), and the full length sequence for each significant hit was downloaded for further analysis (sequence E-value cutoff 0.01; hit E-value cutoff 0.03). Downloaded sequences were filtered based on sequence composition and to reduce redundancy in the dataset using a custom perl script as well as IQ-TREE (Nguyen, Schmidt, von Haeseler, & Minh, 2015) (Appendix 1). Eukaryotic sequences containing two ABM domains were manually split into their individual domains and labeled ‘CTERM’ or ‘NTERM’ accordingly (Appendix 1). Filtered sequences were aligned with

MAFFT v7.023b using the E-INS-i strategy (Katoh & Standley, 2013). The topologies were inferred using maximum likelihood as implemented in IQ-TREE version 1.3.8 using an LG+G4 substitution model (automatically determined to be the best model within IQ-TREE using Bayesian Information Criterion) and ultrafast bootstrapping (1000 replications) (Nguyen et al., 2015). The phylogenies were midpoint rooted and visualized using iTOL version 3.0 (Letunic & Bork, 2016). All trees and alignments are available from the figshare repository (https://doi.org/10.6084/m9.figshare.4810165.v1, last accessed March 31, 2017).

Results

LFO1 exhibits structural similarity to IsdG family heme oxygenases.

Chlamydomonas reinhardtii Cre07.g312300 was identified in an RNA-Sequencing experiment as a transcript that is more highly expressed under iron-limited conditions than under iron-replete conditions (Glaesener et al., 2013). Cre07.g312300 is a nuclear gene that encodes a hypothetical protein of approximately 18 kD (Merchant et al., 2007). Pfam analysis of Cre07.g312300 protein domains identified a putative Antibiotic Biosynthesis Monooxygenase (ABM) domain (Bateman et al., 2002; Sciara et al., 2003) (Fig. 5, red box). Due to its similarity to monooxygenases, I named the gene *Lfo1* for Low Fe-responsive Oxygenase 1. Secondary structure predictions show a secondary structure pattern (Fig 5), consisting of a β-, α-, β-, β-, α-, α-, β-motif (Wu et al., 2005). In addition, there is a 45 amino acid predicted unstructured region between α-helix 3 and β-strand 4 (Fig. 11). The sequence of this region predominantly consists of serine (S), glycine (G), and histidine (H) residues. The N-terminus of LFO1 consists of a predicted chloroplast transit peptide, with a PredAlgo Cscore of 2.32 (Fig. 5) (Tardif et al., 2012), where a Cscore above 0.41 predicts that the protein localizes to the chloroplast.

The predicted secondary structure of LFO1 is very similar to that of the IsdG family of heme oxygenases (Fig. 5) (Wu et al., 2005). IsdG family heme oxygenases generally have low sequence identity; however, they have highly conserved secondary structures (Matsui et al., 2016; Wu et al., 2005), and all of the IsdG family members contain an ABM domain (Fig. 5). The secondary structure of IsdG family heme oxygenases is quite different from the secondary structure that is seen in HO-1 family heme oxygenases (Fig. 5), which is composed of 10 α-helices (Hirotsu et al., 2004; Schuller et al., 1999; Sugishima, Migita, Zhang, Yoshida, & Fukuyama, 2004;

Sugishima, Sakamoto, Higashimoto, Noguchi, & Fukuyama, 2003). These characteristics led us to hypothesize that LFO1 is an IsdG family heme oxygenase.

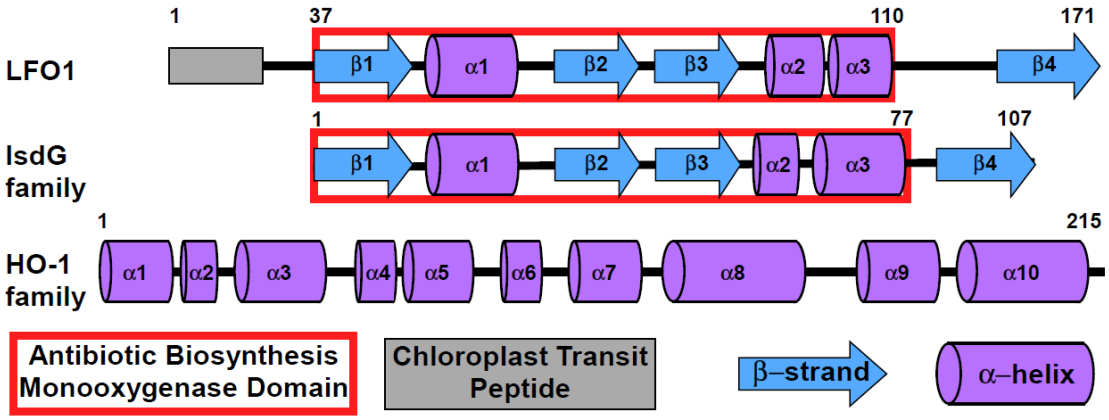


Fig 5. LFO1 secondary structure is more similar to the IsdG family than the HO-1 family of heme oxygenases. The predicted structure of LFO1 is shown compared to a prototypical IsdG family member, IsdG from *S. aureus*, and HO-1 family heme oxygenase, HmuO from *C. diphtheriae*. The blue arrows represent β-strands, the purple cylinders represent α-helices, and the grey rectangle represents the chloroplast transit peptide in LFO1. The red outline represents the predicted Antibiotic Biosynthesis Monooxygenase (ABM) domain.

LFO1 binds and degrades heme.

Heme has a distinct visible spectrum that peaks at ~390 nm (Fig. 6, black dashed line). Upon complexation with heme binding proteins, this peak shifts to what is known as a Soret peak, with an absorption maximum around 400-450 nm (Stryer, 1961; Watanabe & Horie, 1976). Heme oxygenases of both the IsdG and HO-1 families typically exhibit a Soret peak of 405-413 nm (Wilks, Torpey, & Ortiz de Montellano, 1994; W. Zhu et al., 2000). Incubating heme with LFO1 leads to a shift in the peak to 413 nm (Fig. 6), consistent with LFO1 binding heme. The binding ratio of LFO1 to heme is 1:1, based on the inflection point being at 10 μM heme when increasing concentrations of heme were added to LFO1 (Fig. 6, inset). Michaelis-Menten kinetics was used to calculate the binding affinity (K_d) of 2.4 $\pm$ 0.3 μM (Fig. 6, inset). The binding affinity of LFO1 is in the range of binding affinities calculated for various heme oxygenases, which vary from 0.84 + 0.2 μM (Wilks et al., 1996) to 5.0 + 1.5 μM (Skaar, Gaspar, et al., 2004). These data demonstrate that LFO1 binds heme with characteristics consistent with other IsdG family heme oxygenases.

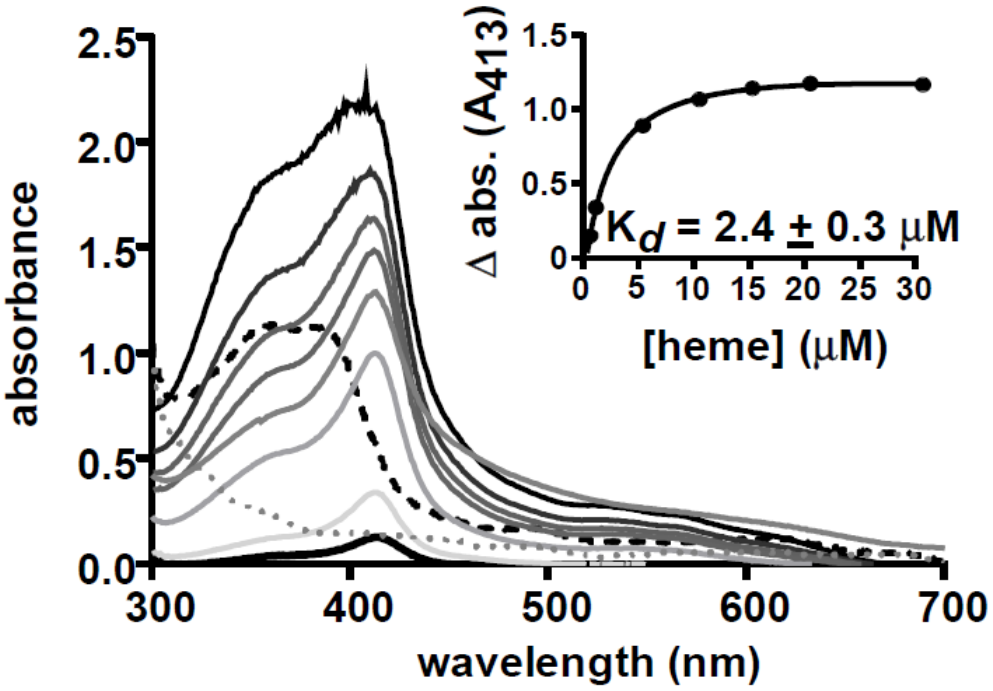


Fig 6. LFO1 binds heme *in vitro*. Absorption spectra of heme binding to LFO1. Increasing amounts of heme (0.5-30 μM) were added to 10 μM protein. The spectrum for 10 μM heme is shown in the dashed black line and LFO1 alone is shown in a dashed grey line, whereas the protein incubated with heme is shown in the solid lines, with increasing heme concentrations leading to an increase in peak height. The inset displays the change in absorbance at 413 nm for LFO1 bound to heme minus heme alone from 1 to 30 μM heme.

To determine whether LFO1 degrades heme, LFO1 was incubated with heme and ascorbate and the visible spectra were collected every 10 minutes for 1 hour. In these conditions, the LFO1-bound heme peak decreased over time, demonstrating that LFO1 degrades heme (Fig. 7). Using the same reaction conditions, I tested the ability of LFO1 to degrade other metallo-protoporphyrins and found that LFO1 can only degrade heme, though it seems to be able to bind most non-iron protoporphyrins (Fig. 8). This finding is consistent with other heme oxygenases, since iron is required to facilitate the cleavage of the porphyrin ring (Lee, Reniere, Skaar, & Murphy, 2008; Streit et al., 2016). The heme degradation reaction was also performed in the presence of catalase, which inhibits heme autoxidation. Once again, the 413 nm peak decreased with time (Fig. 9), indicating that the degradation of heme is due to the catalytic activity of LFO1. These data establish LFO1 as a heme-degrading enzyme.

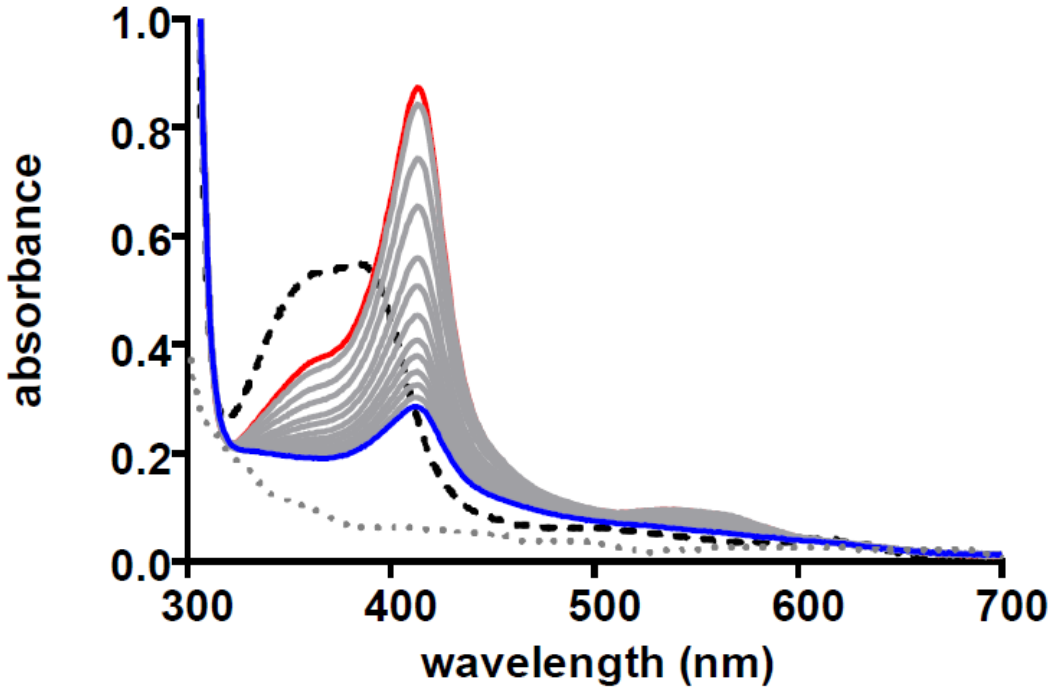


Fig 7. LFO1 degrades heme *in vitro*. Heme degradation reactions with LFO1. Ten µM LFO1 was incubated with equimolar heme, then ascorbic acid was added as the reducing agent. The spectrum was monitored from 0 minutes (red line) to 60 minutes (blue line) with time points obtained every 5 minutes (grey lines).

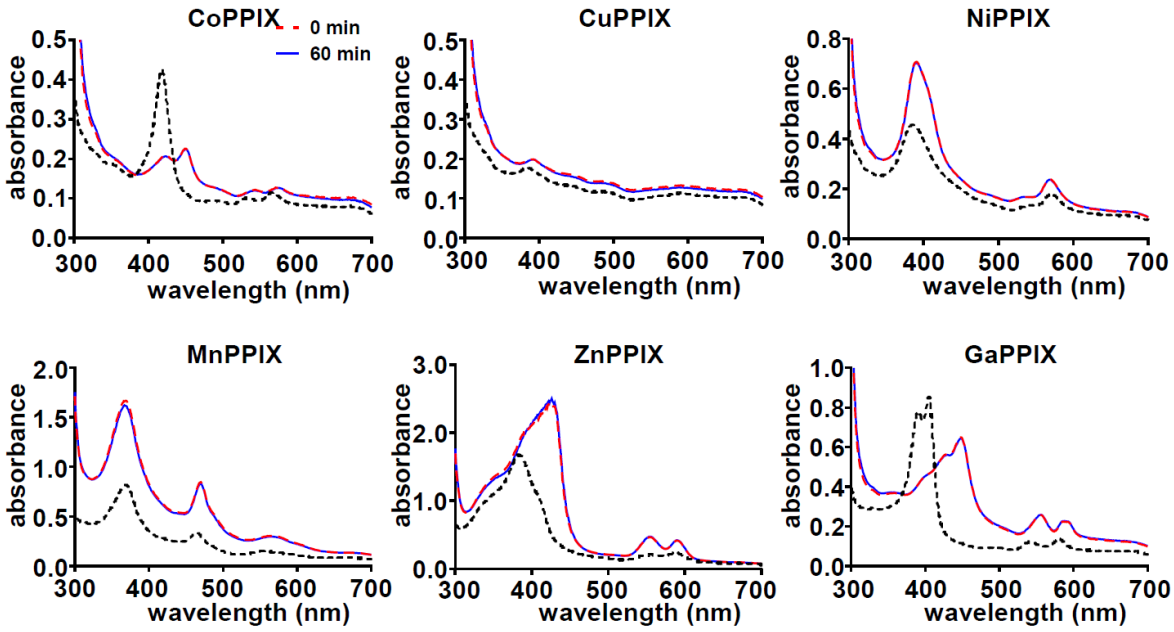


Fig. 8 LFO1 is unable to degrade non-iron protoporphyrin. Heme degradation reactions with LFO1. 10 μM of protein was incubated with equal molar concentrations of various non-iron protoporphyrins: Cobalt protoporphyrin (CoPPIX), Copper protoporphryin (CuPPIX), Nickel protoporphyrin (NiPPIX), Manganese protoporphyrin (MnPPIX), Zinc protoporphyrin (ZnPPIX), and Gallium protoporphyrin (GaPPIX). Ascorbic acid was added as the reducing agent. The spectrum was monitored from 0 minutes (red dashed line) to 60 minutes (blue line), with time points obtained every 10 minutes (grey lines). The dashed black lines represent each protoporphyrin tested alone.

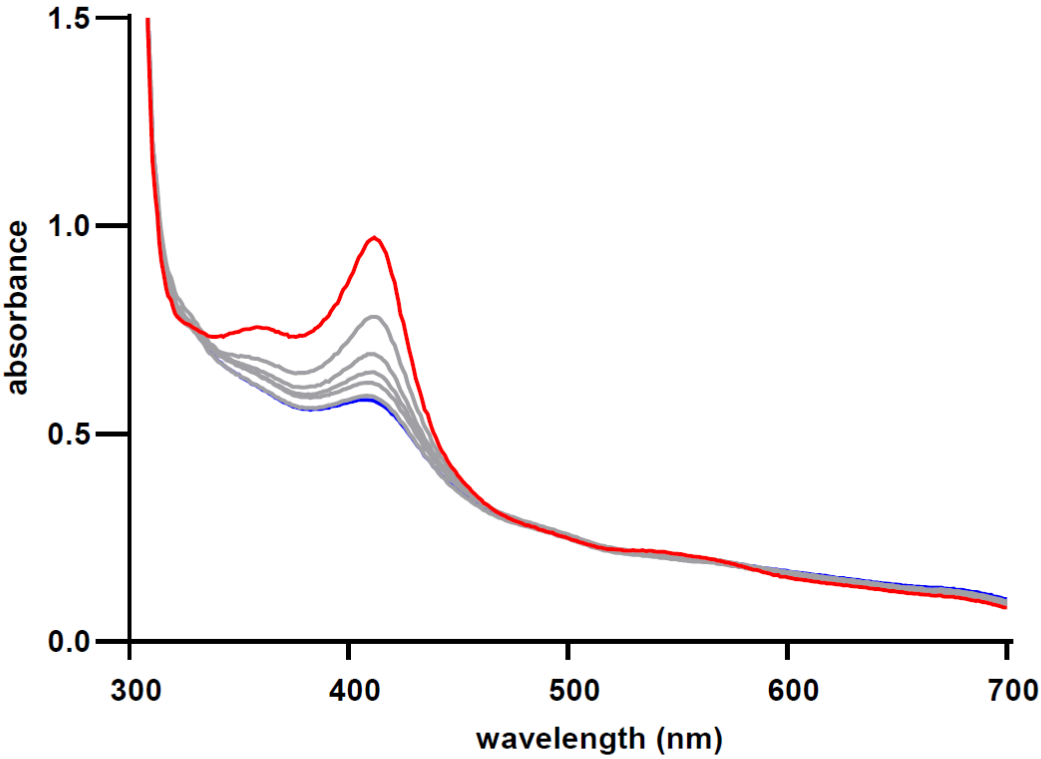


Fig 9. LFO1 degrades heme in the presence of catalase. Ten µM heme was incubated with ten µM LFO1, then catalase and ascorbate was added. The absorbance was monitored from 300-700 nm every 10 minutes for 60 minutes.

LFO1 contains protein sequence similarities to both antibiotic biosynthesis monooxygenases as well as the IsdG family of heme oxygenases, which have been also been identified as monooxygenases (Kendrew, Hopwood, & Marsh, 1997; Streit et al., 2016). However, other non-oxygen dependent heme degrading enzymes have also been recently identified (LaMattina et al., 2016). To test whether LFO1 uses oxygen as a substrate in heme degradation, heme degradation reactions were performed in an oxygraph allowing for the detection of oxygen consumption. When LFO1 was bound to heme in the absence of ascorbate, the rate of oxygen consumption was 6.626 $\pm$ 0.991 nmol of oxygen/μL/min, whereas after ascorbate was added and degradation of heme commensed LFO1 utilized 19.994 $\pm$ 0.116 nmol of oxygen/μL/min (Fig. 10). This rate of oxygen incorporation is similar to that of IsdI (18.366 $\pm$ 3.951 nmol of oxygen/μL/min), yet both were statistically different than the rate of oxygen incorporation of heme alone (7.904 $\pm$ 0.639 nmol of oxygen/μL/min) (Fig. 10). This shows that LFO1 utilizes oxygen during heme degradation, and at a similar rate to other heme monooxygenases.

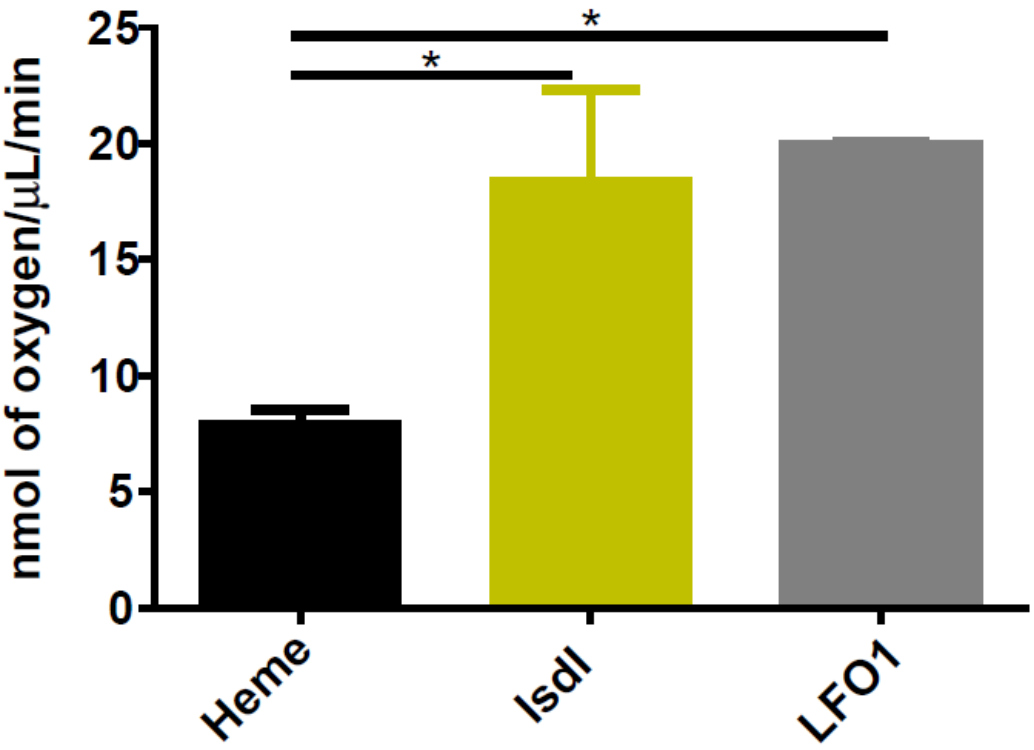


Fig 10. LFO1 consumes oxygen during the degradation of heme. The rate of oxygen consumption for heme alone, IsdI, and LFO1 was monitored via oxygraph. The rates of oxygen utilization for IsdI and LFO1 are significantly different than that of heme alone.

IsdG family heme oxygenase catalytic residues are conserved and important for LFO1 function.

In addition to conserved secondary structural characteristics, IsdG family heme oxygenases are also defined by containing a conserved catalytic triad of amino acids composed of asparagine (N), tryptophan (W), and histidine (H). These residues are required by all functionally characterized IsdG family heme oxygenases to degrade heme (Wu et al., 2005). LFO1 was aligned to protein sequences of seven functionally characterized IsdG family members, revealing the presence of the conserved catalytic triad (Fig. 11).

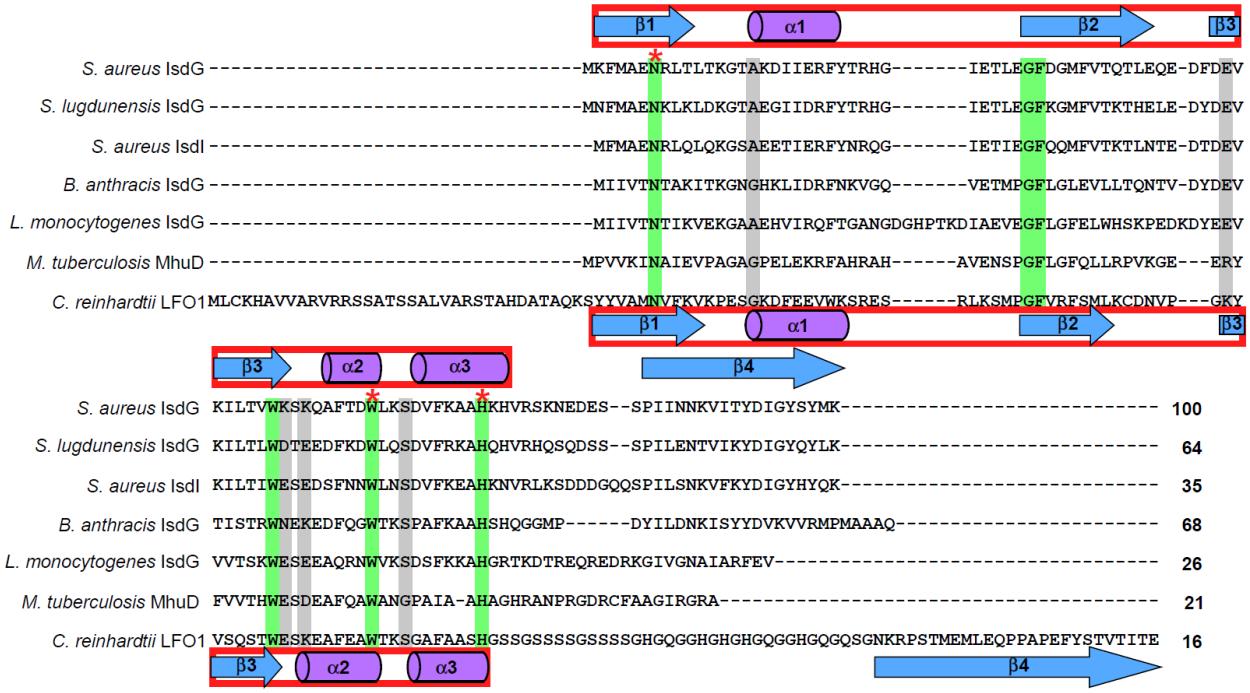


Fig 11. IsdG family catalytic residues are conserved in LFO1 and are required for heme degradation. A, Alignment of the full-length sequence of six functionally characterized IsdG family heme oxygenases with LFO1. Identical residues are highlighted in green and similar residues are highlighted in grey. The IsdG family catalytic residues are starred. Solved secondary structure for *S. aureus* IsdG are displayed above and predicted secondary structure for LFO1 are displayed below the alignment. Percent identity of each sequence compared to *S. aureus* IsdG is shown at the end of the sequence.

To determine if LFO1 catalyzes heme degradation similar to other IsdG family heme, these residues were individually changed to alanine. Mutation of the conserved asparagine (N41A) (Fig. 12), tryptophan (W100A) (Fig. 12), or histidine (H110A) (Fig. 12) decreases the ability of LFO1 to degrade heme. This finding indicates that the ability of LFO1 to degrade heme is dependent on these residues. Interestingly, there are three additional conserved residues across the IsdG family heme oxygenases that are also contained within LFO1: F69, G70, and W91. When these residues were mutated to alanines, the G70A and W91A mutants were unable to degrade heme (Fig. 13). The LFO1 enzyme has both the predicted secondary structure and the conserved triad of catalytic residues categorizing LFO1 as an IsdG family heme oxygenase.

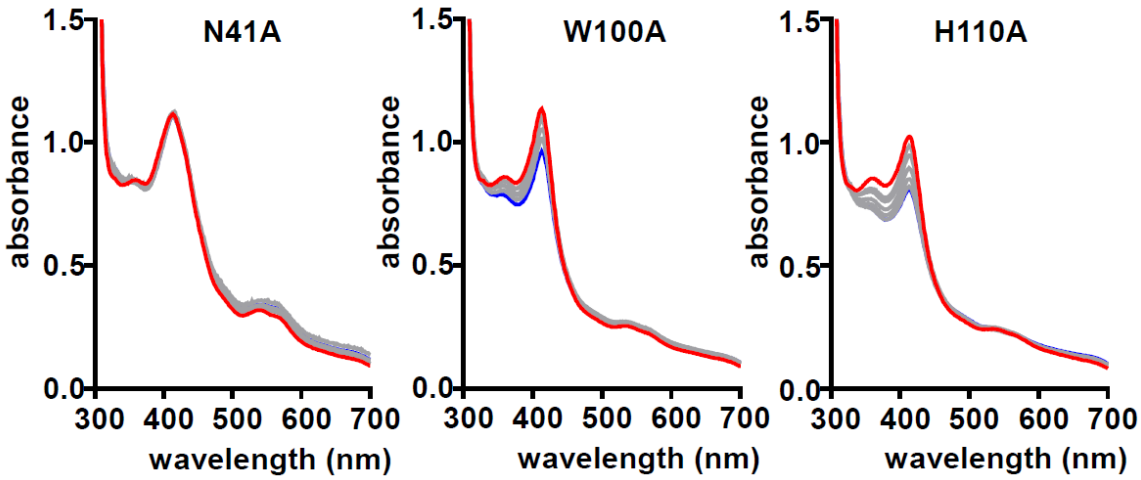


Fig 12. IsdG family catalytic residues are required for heme degradation. Heme degradation reactions, in the presence of catalase, with LFO1 with point mutations in each of the conserved catalytic residues N41A, W100A, and H110A. The spectrum was monitored from 0 minutes (red line) to 60 minutes (blue line) with time points obtained every 5 minutes (grey lines).

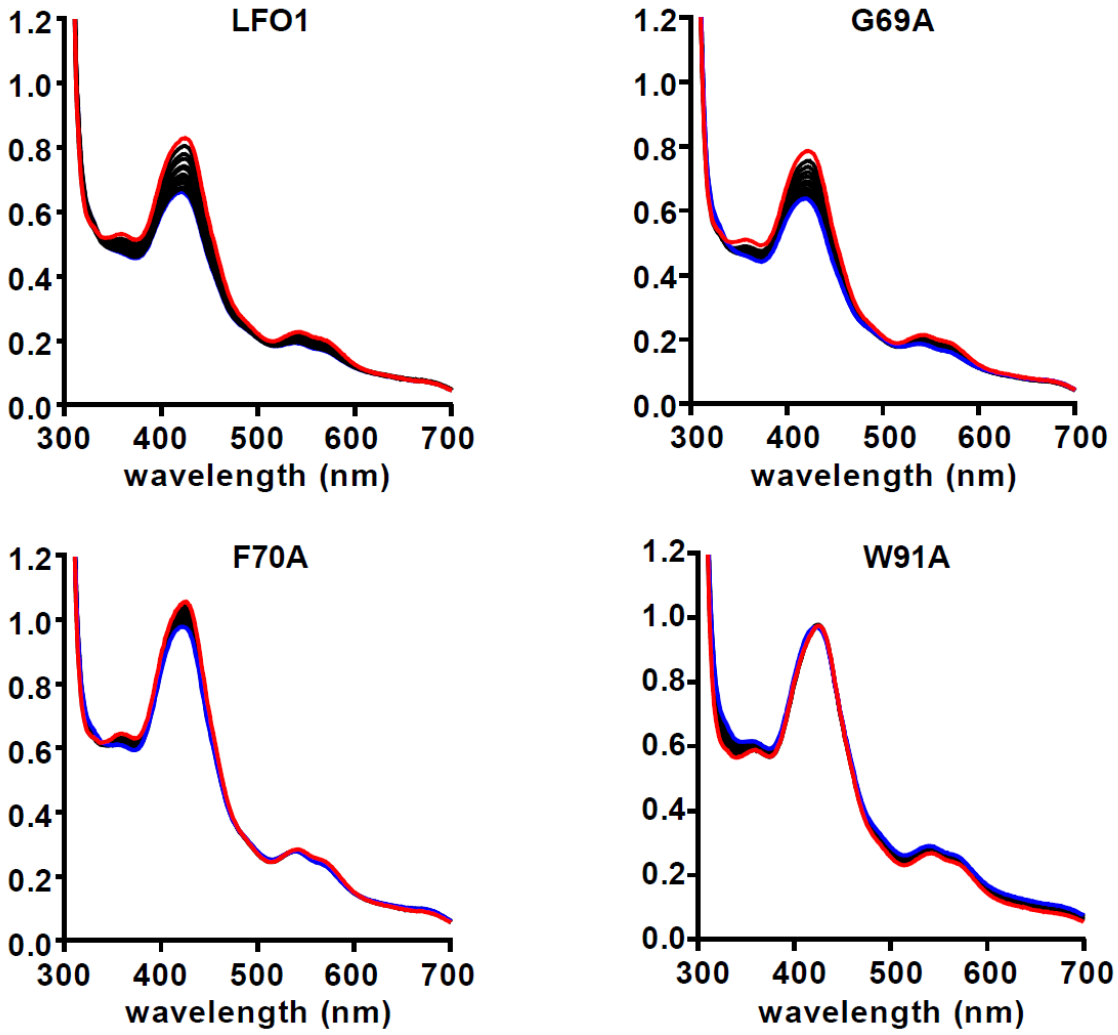


Fig 13. Two IsdG family conserved non-catalytic residues are required for heme degradation by LFO1. Heme degradation reactions, in the presence of catalase, with wildtype LFO1 or LFO1 with point mutations in each of the conserved residues G69A, F70A, and W91A. The spectrum was monitored from 0 minutes (red line) to 60 minutes (blue line) with time points obtained every 5 minutes (black lines).

LFO1 degrades heme to a distinct heme catabolite.

In order to determine if the products of the LFO1 heme degradation reaction were consistent with previously identified heme degradation products, the LFO1 heme degradation reaction was purified and analyzed by High Performance Liquid Chromatography (HPLC). The retention time and spectrum of the LFO1 heme catabolite was compared to staphylobilin, biliverdin, and mycobilin. The retention time of the LFO1 product is 32.57 minutes (Fig. 14), which is distinct from the retention times of known heme oxygenase products. Heme was also run for comparison; however, heme has a retention time greater than that of all of the heme catabolites (~60 min). Additionally, the LFO1 degradation product has a unique absorbance spectrum compared to mycobilin and staphylobilin (Fig. 14). These data indicate that the LFO1 heme catabolite has a different structure that has less hydrophobicity than that of previously identified heme degradation products.

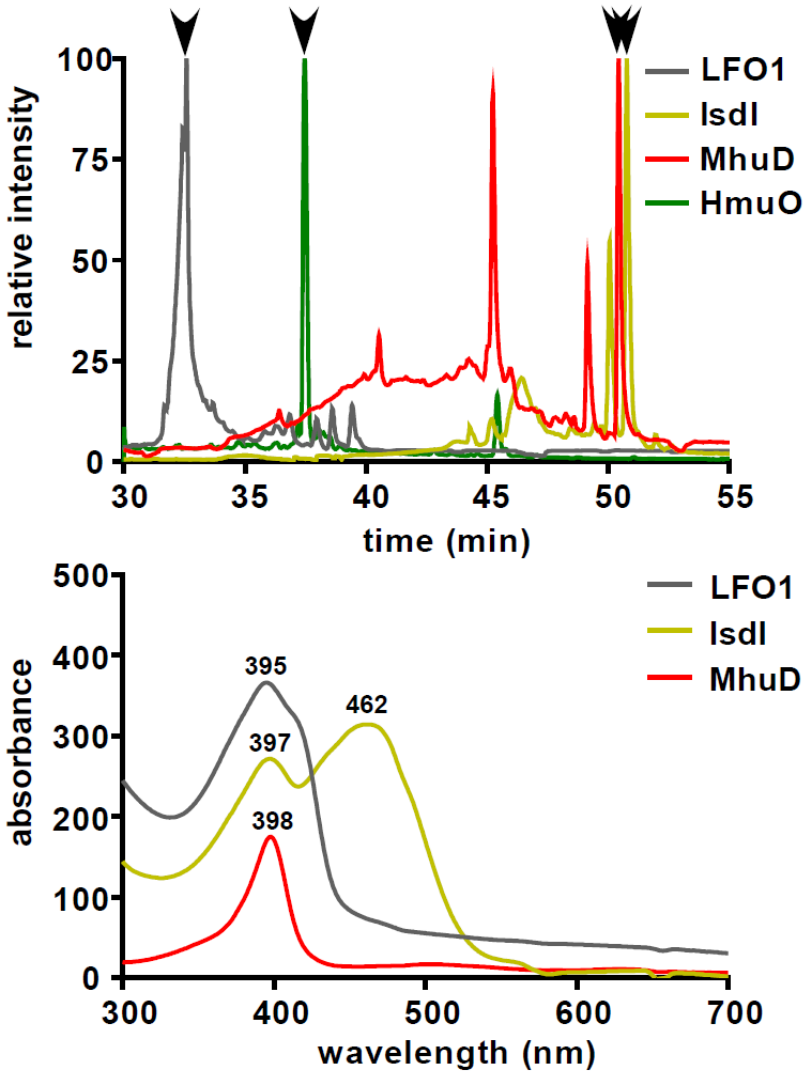


Fig 14. LFO1 degrades heme to a distinct catabolite. A, HPLC chromatograms of mycobilin from MhuD heme degradation (red), biliverdin from HmuO heme degradation (green), staphylobilin from IsdI heme degradation (yellow), and the LFO1 catabolite (grey) monitored at an absorbance wavelength of 400 nm, graphed with the major peak set to 100. The major peak is indicated for each sample with a black arrow. The predominant peak for the LFO1 catabolite has a retention time of 32.57 minutes. B, Spectra for each of the samples were extracted at the retention times of their peak maxima. The LFO1 catabolite has a peak absorbance at 395 nm at 32.57 min.

One of the products of HO-1 family heme degradation is carbon monoxide. Utilizing COP-1, a reaction-based probe which fluoresces selectively in the presence of carbon monoxide (Michel et al., 2012), the ability of LFO1 to produce carbon monoxide as a result of heme degradation was determined. When HmuO from *Corynebacterium diphtheriae* degrades heme in the presence of COP-1 the rate of fluorescence increase is 1215.6 $\pm$ 213.1 fluorescence units/minute. When LFO1 degrades heme in the presence of COP-1 the rate of fluorescence production is significantly less than HmuO (297.8 $\pm$ 117.6 fluorescence units/minute), and more comparable to the rate of fluorescence with IsdI (285.3 $\pm$ 98.6 fluorescence units/min) (Fig. 15). This indicates that LFO1 does not release carbon monoxide upon degradation of heme, confirming that LFO1 is not an HO-1 family heme oxygenase.

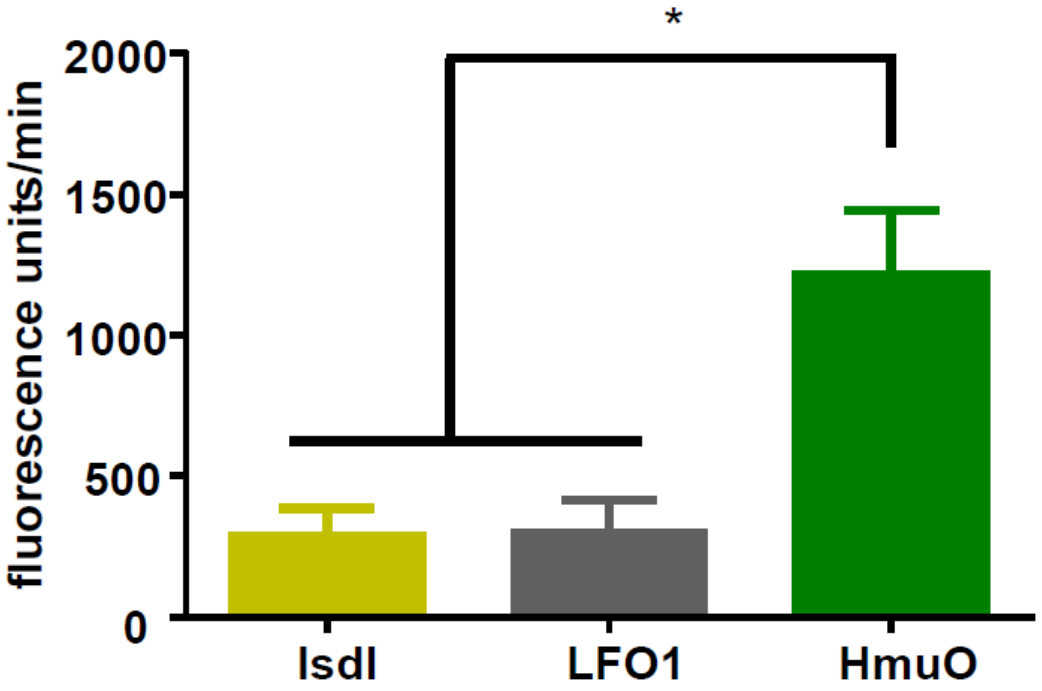


Fig 15. LFO1 does not produce carbon monoxide. Rates of fluorescence units per minute acquired from incubating the heme degradation reactions of IsdI (yellow), LFO1 (grey), or HmuO (green) in the presence of COP-1, a reaction-based probe which fluoresces selectively in the presence of carbon monoxide, ($p=0.0330$).

The LFO1 heme degradation product has a mass spectrum distinct from previously identified catabolites.

In order to definitively determine the catabolite produced from LFO1 heme degradation, I performed a variety of analytical approaches, including mass spectrometry, nuclear magnetic resonance (NMR), and structural determination. Figure 16 displays a schematic of the optimization strategies and analytical techniques that were used to try to ascertain the mass and structure of the LFO1 catabolite.

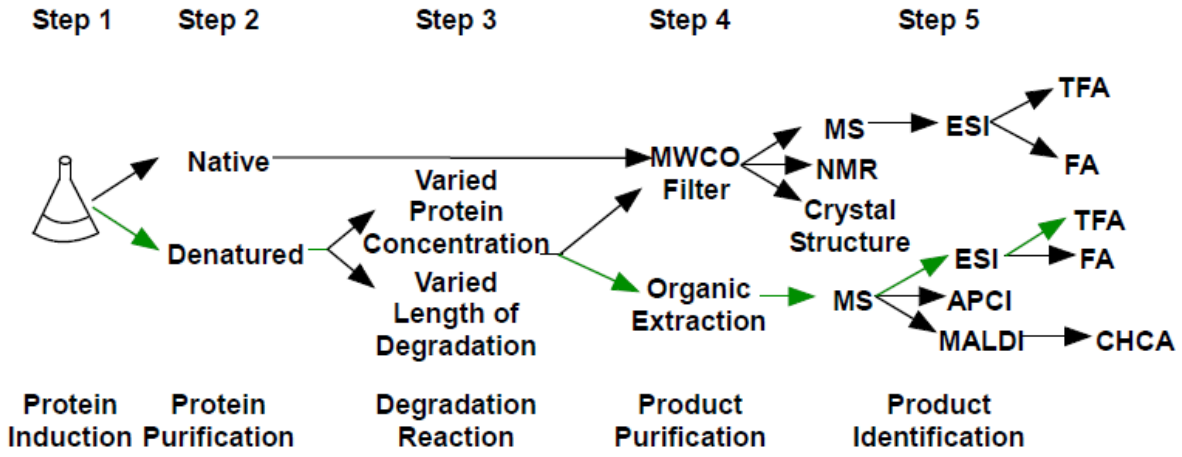


Fig 16. Strategy for identification of the LFO1 heme catabolite. Flowchart for the methods used to identify the heme degradation product of LFO1. Step 1 included varying induction conditions to optimize protein yield. Step 2 focused on determining the best method for protein purification. Step 3 involved optimizing the conditions of the degradation reaction. Step 4 sought to determine the optimal way to purify the product. Finally, step 5 utilized a variety of analytical techniques, ionization strategies, and protocols to identify the product.

Briefly, optimization began with identifying the induction conditions which would provide optimal protein yield (Step 1). After the ideal protein induction parameters were determined, the protein purification procedure was optimized. Since LFO1 is not predicted to be membrane bound within the chloroplast, I initially utilized a protein purification methodology that had been optimized for the cytoplasmic proteins IsdG and IsdI from *S. aureus* (Step 2) (Reniere et al., 2010). However, using that protocol led to a variety of band sizes smaller than that of the predicted 18kD full-length protein band when run on a Coomassie stained polyacrylamide gel. I interpreted this to mean that LFO1 was being degraded during the purification. In order to bypass LFO1 degradation, I switched to a denaturing purification strategy using urea to unfold the protein, then stepwise decreased the urea concentration while the protein was bound to the column to allow for refolding of LFO1 (Step 2). This provided more protein of the predicted 18kD length (Fig. 17). Both natively purified LFO1 and LFO1 purified while denatured and then refolded are able to bind and degrade heme, indacitng that LFO1 is able to refold to its native structure allowing it to be functional.

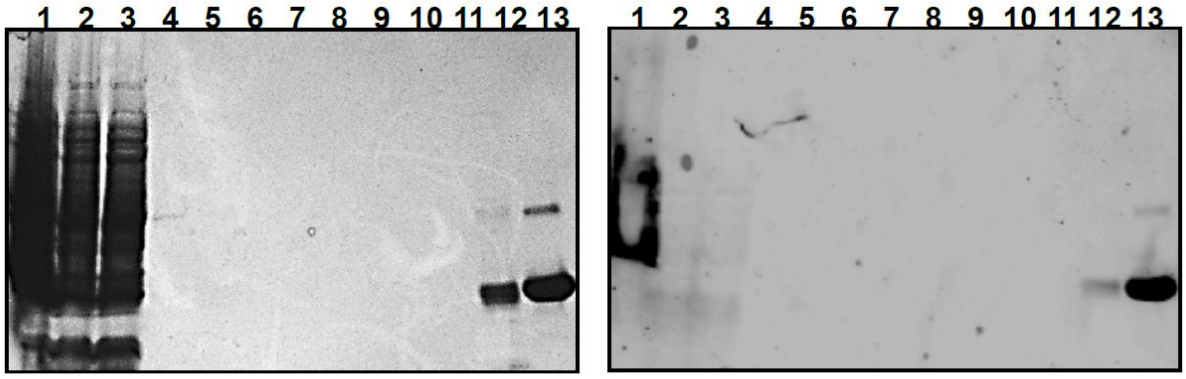


Fig 17. Purification of His-LFO1 from *E. coli*. A, Coomassie stained gel of recombinantly expressed 6His-LFO1 from *E. coli*. B, Western blot for 6His tag. For both gels : lane 1 is the cellular lysate, lane 2 is the column flowthrough, lanes 3-11 are the step-wise decreasing concentrations of urea (from 8 M to 0 M), lane 12 100 mM imidazole wash, and lane 13 is the 400 mM imidazole elution.

I next optimized the heme degradation reaction of LFO1 (Step 3). For this step, I varied the protein concentration and the duration of the degradation reaction to ensure that the protein remained soluble and that the degradation reaction was halted as soon as there was no longer any tetrapyrrole cleavage, as monitored by the spectrophotometer. Step 4 involved purification of the product from the protein. Two different purification processes were attempted during the course of optimization. The initial protocol for staphylobilin purification from IsdI and IsdG utilizes a sizing based separation, where the protein-product complex is concentrated in molecular weight cut off (MWCO) filter centrifuge tubes, then 0.1% trifluoroacetic acid (TFA) was added to denature the protein leading to the release of the product which can then pass through the sizing filter (Reniere et al., 2010). This method was attempted first, however the length of time which it took to concentrate the protein and then separate the product from it was fairly extensive and chromatograms from samples prepared by this method had no clearly distinct major peaks. The second product purification method was based on an organic extraction. The whole heme degradation reaction was added to a separatory funnel, then mixed with ethyl acetate and 0.1% TFA. The organic layer was then removed, dried under inert gas, and resuspended in 1:1 acetonitrile:water before injecting onto either an HPLC, LC-MS/MS, or fixed in matrix and analyzed by matrix associated laser desorption/ionization (MALDI) MS. Finally, five different product identification strategies were attempted (Step 5). Product from natively purified protein was either analyzed by LC-MS/MS using electrospray ionization (ESI), submitted for NMR analysis, or submitted for crystallization. After the protein purification strategy was further optimized using a denaturing protein purification protocol, product identification was attempted with LC-MS/MS using ESI and one of two ion pairing agents (TFA or formic acid (FA)), using

atmospheric pressure chemical ionization (APCI), or using MALDI-MS using α-cyano-4-hydroxycinnamic acid (CHCA) as the matrix. To determine the efficacy of our optimization strategies, I purified IsdI from *S. aureus* alongside LFO1 and analyzed the various analytical techniques for both the LFO1 product and staphylobilin.

In order to determine if the product of LFO1 heme degradation is similar to or distinct from previously identified heme catabolites, I first performed LC-MS/MS analysis comparing purified products of heme degradation reaction using proteins that degrade heme to either staphylobilin or biliverdin, as shown in Figure 18. I positively identified the masses for each heme degradation product, using the methodology originally employed for staphylobilin (Reniere et al., 2010). I was unable to identify any of the previously characterized catabolite masses in the spectrum of the LFO1 degradation product.

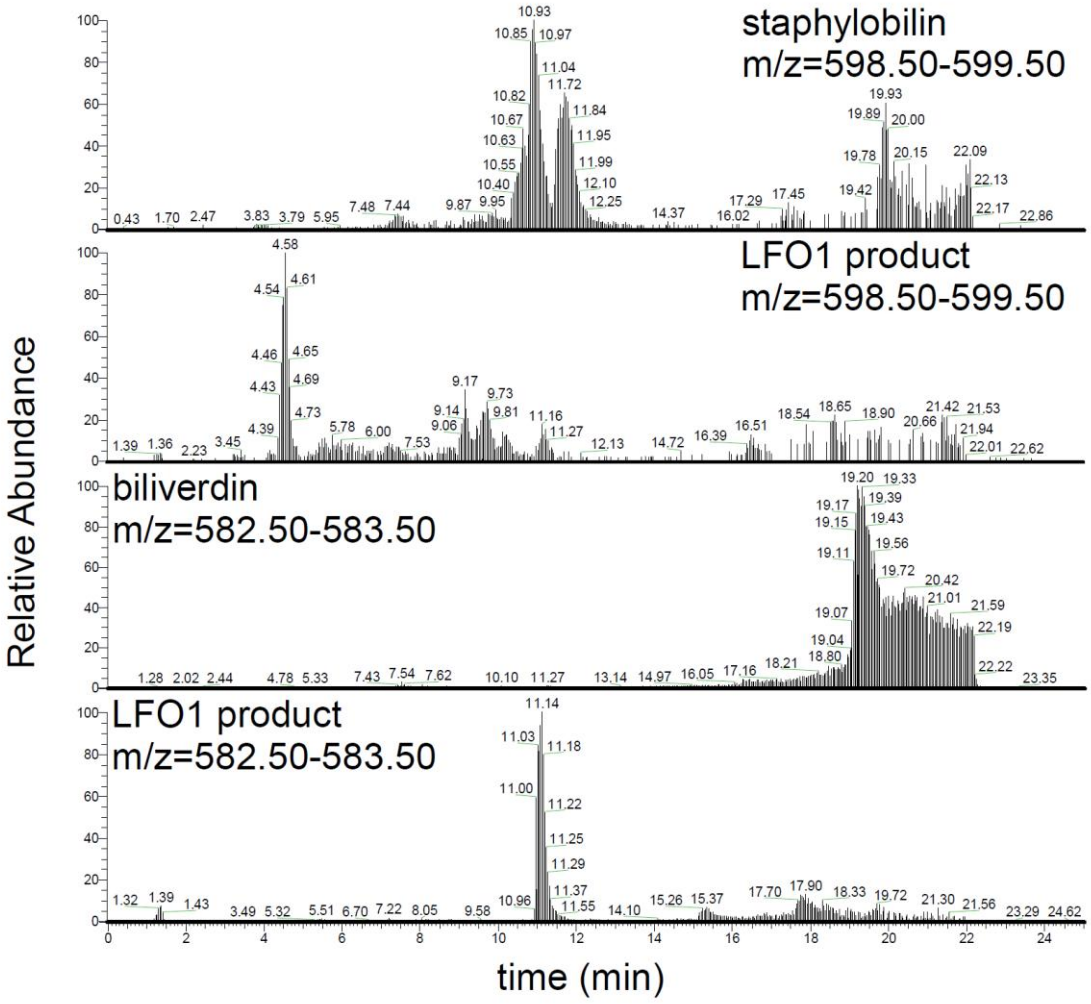


Fig 18. The LFO1 degradation product is distinct from previously identified heme degradation products. Staphylobilin, biliverdin, and LFO1 degradation product were analyzed by LC-MS/MS. Shown are the extracted ion chromatograms (XICs) for two distinct sets of masses. The staphylobilin standard and LFO1 product were interogated for mass/charge ratios (m/z) for 598.50-599.50. The exact mass for staphylobilin is 599.3. The staphylobilin XIC shows two predominant peaks, at 10.93 and 11.72 min, representing β- and δ-staphylobilin. However, the LFO1 product XIC does not contain peaks at these time points, indicating that the LFO1 product is not staphylobilin. The biliverdin XIC contains a peak for m/z of 582.50-583.50 at 19.20 min. The exact mass of biliverdin is 582.56. Once again, the LFO1 product XIC does not contain a peak at the same retention time, indicating that the LFO1 product is not biliverdin.

Using the optimization strategy highlighted in green throughout Figure 16, I performed a time course analysis of LFO1 heme degradation. I performed the degradation reaction over 5 hours, and took samples from the reaction at 1, 3 and 5 hours after the initiation of degradation by the addition of ascorbate. I then purified the products and analyzed the samples by LC-MS/MS. The chromatogram for each sample was collected at 400-500 nm, which provides signal for both heme and the LFO1 degradation product (Fig. 19). Over this time course there is a reduction in heme (retention time 40.24 min), and an increase in a peak at 25.13 min, which is assumed to be the major product peak. However, despite the fact that I was able to verify the heme peak by its mass and fragmentation pattern, I was unable to find a significant mass peak that coincides with the LC peak at 25.13 min for the product and that has a fragmentation pattern that supports it being the product. Two masses were more abundant in the five hour degraded samples at 25.13 min than in any of the other samples, m/z 670.48 and 795.64. However, both of these masses are larger than hemin (exact mass of 651.94) and the MS/MS fragmentation pattern of these masses were inconsistent with cleavage for portions of heme, such as the carboxylic acids. Ultimately, I was unable to identify the product of LFO1 heme degradation.

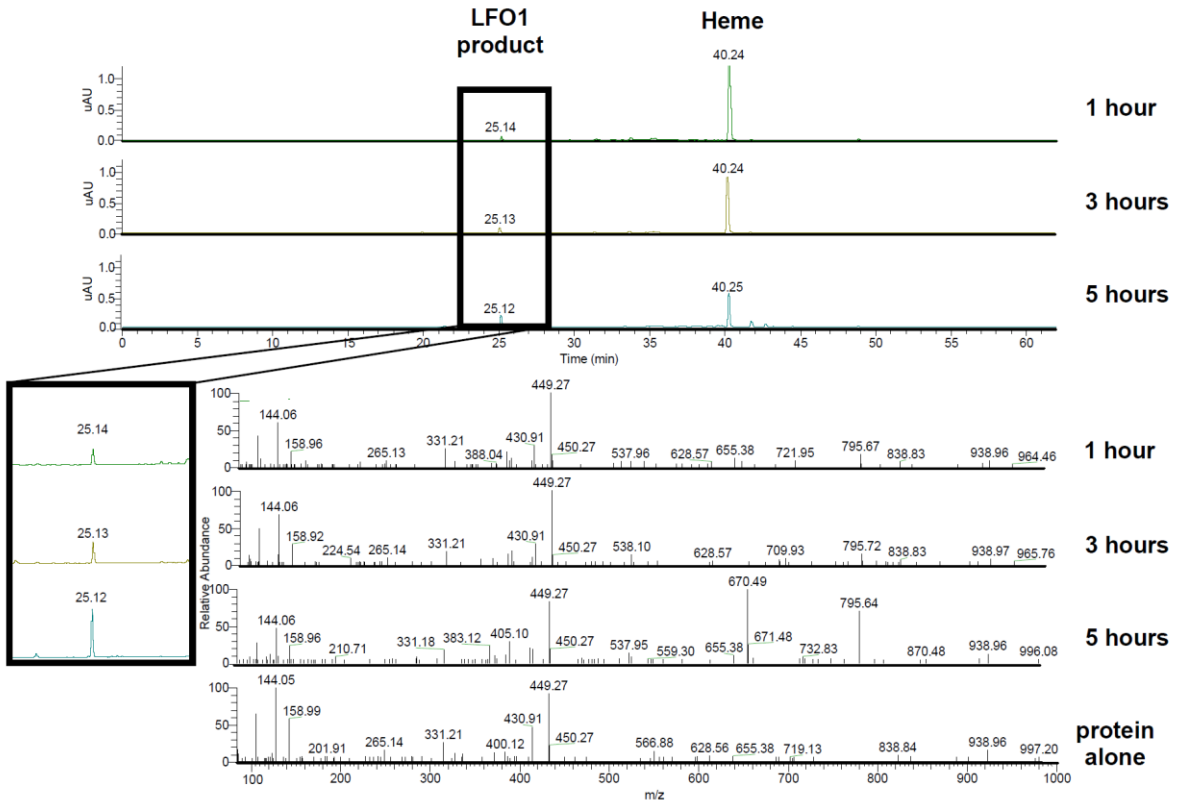


Fig 19. LC-MS analysis of the LFO1 degradation product. an LFO1 heme degradation reaction was set up and 1, 3, and 5 hour points were isolated after initiation of degradation. The product of these reaction conditions were isolated for each time point with ethyl acetate. The isolated products were then analyzed by LC-MS. Top, LC chromatograms of the isolated products observed between 400-500 nm. The 40.25 min peak is heme, whereas the 25.12-25.14 peak is assumed to be the LFO1 degradation product. Boxed, A zoomed in view of the LFO1 heme degradation peak. Below, Mass spectra of the LFO1 product peak at 1, 3, and 5 hours post-intiation of degradation, in addition to a protein alone control.

The IsdG family of heme oxygenases is distributed across all domains of life.

A phylogenetic tree was previously created using the first identified member of the IsdG family of heme oxygenases, IsdG from *S. aureus*, as a seed (Haley et al., 2011). This tree identified a total of 22 proteins as IsdG family members. However, characterization of LFO1 as an IsdG family member shows that this tree does not sufficiently encapsulate all potential IsdG family heme oxygenases. Our search identified 1,204 sequences homologous to characterized IsdG proteins, 866 of which contained the conserved NWH catalytic triad and are hereafter referred to as IsdG-like proteins (Appendix 1). The 852 bacterial IsdG-like proteins were distributed across eight bacterial divisions, with the majority found in Actinobacteria (317 proteins), Proteobacteria (292 proteins), and Firmicutes (206 proteins) (Appendix 1). Our search also identified four archaeal IsdG-like proteins, one from Crenarchaeota and three from Euryarchaeota. In addition to LFO1 from *C. reinhardtii*, our search identified nine IsdG-like proteins in eukaryotes, all from other photosynthetic species including the colonial green algae *Volvox cateri*, prasinophyte green algae in the genera *Ostreococcus* and *Micromonas*, diatoms in the genera *Fragilariopsis* and *Thalassiosira*, and the chromerid *Vitrella brassicaformis* (Appendix 1).

To determine the evolutionary history of LFO1, I constructed a maximum likelihood phylogeny of IsdGs and homologs (Fig. 20). All the eukaryotic sequences, including LFO1 grouped within a large Proteobacterial clade of IsdG-like proteins. *C. reinhardtii* and *V. cateri*, the two fresh water green algae, group with each other but not with the marine prasinophyte green algae, which instead group with marine diatoms (Fig. 20). The small number of eukaryotic IsdG-like sequences and the fact that they do not form a single monophyletic group suggests that these sequences may have been acquired from bacteria via horizontal gene transfer. However, the low

average sequence similarity and small length of this protein resulted in a phylogeny with weak branch support, which makes it challenging or even impossible to infer the number and directionality of possible gene transfer events.

Interestingly, there are multiple organisms identified by these analyses that have previously been described to contain HO-1 family heme oxygenases. This suggests that *C. reinhardtii* is not unique in harboring two distinct families of heme oxygenases.

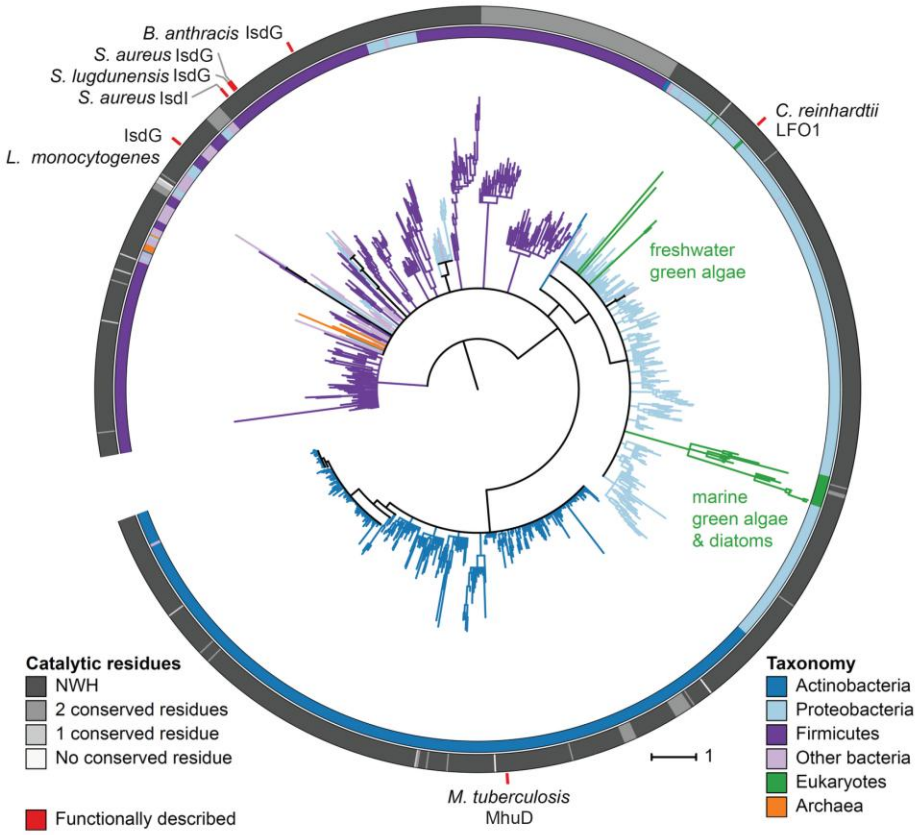


Fig 20. The IsdG family of heme oxygenases is widely distributed in bacteria and present in all three domains of life. Maximum likelihood phylogeny of the IsdG heme oxygenase protein family. The tree was midpoint rooted and visualized using iTOL version 3.0. Color strips along the tree perimeter correspond to the taxonomy (inner strip) and the number of conserved IsdG family catalytic residues contained within each protein sequence (outer strip). Red tick marks indicate the IsdG family enzymes that have been previously functionally characterized.

Discussion

The HO-1 family of heme oxygenases degrade heme to the products biliverdin, carbon monoxide, and free iron (Tenhunen et al., 1968). However, the discovery of the IsdG family indicates that heme can also be degraded to different bilin products (Skaar, Gaspar, et al., 2004). These bilin products could function in as yet undiscovered light signaling or host-microbe signaling pathways. By continuing to explore this family of heme oxygenases, a protein from *C. reinhardtii*, LFO1, was identified. LFO1 has a predicted secondary structure similar to the conserved secondary structure seen in IsdG family heme oxygenases, and distinct from the secondary structure of the HO-1 family (Fig. 5). Functional characterization of LFO1 shows that it is both able to bind and degrade heme (Fig. 6,7), and this heme degradation is dependent on the presence of three catalytic residues, which are conserved in all IsdG family members (Fig. 11, 12). While HO-1 family heme oxygenases create the products biliverdin and carbon monoxide, the IsdG family of heme oxygenases produces multiple products, including staphylobilin and mycobilin. Interestingly, although I was unable to structurally define the product of LFO1-dependent heme catabolism, it appears that LFO1 produces a molecule distinct from the bilin degradation products of previously described heme oxygenases (Fig. 14, 15, 18).

IsdG family heme oxygenases have been functionally characterized in seven bacterial species (Duong et al., 2014; Haley et al., 2011; Nambu et al., 2013; Park et al., 2012; Puri & O'Brian, 2006; Reniere et al., 2010; Skaar et al., 2006), and only 22 enzymes were predicted to be IsdG family members by phylogenetic analysis (Haley et al., 2011). However, due to the low sequence identity between IsdG family members it is not surprising that the previous phylogenetic analysis potentially missed enzymes with similar functionality. Focusing on the secondary

structure and the conservation of the catalytic triad allowed identification of LFO1 as a eukaryotic enzyme that belongs to the IsdG family of heme oxygenases. Using the sequence of LFO1 as a seed, a new phylogenetic analysis was completed. This sequence analysis identified 866 species containing IsdG family members, indicating IsdG heme oxygenases are more widely distributed than previously appreciated (Fig. 20).

Chlamydomonas reinhardtii contains HMOX1 and HMOX2, two HO-1 family heme oxygenases (Duanmu et al., 2012; Elbaz, Wei, Meng, Zheng, & Yang, 2010), and LFO1, an IsdG family heme oxygenase. Interestingly, RNA-sequencing data from *C. reinhardtii* showed that *Lfo1* is differentially expressed as a function of iron nutrition, such that *Lfo1* is upregulated under low iron conditions and down-regulated under iron excess (Glaesener et al., 2013). This regulation of LFO1 is similar to that of bacterial IsdG family members, where the heme oxygenases are important for iron scavenging from heme. However, HMOX1 within the chloroplast and HMOX2 within the cytoplasm are not differentially expressed under these conditions (Glaesener et al., 2013). Photosynthesis is a highly iron-dependent process, requiring approximately 30 iron atoms for linear electron flow (Blaby-Haas & Merchant, 2013). Therefore, many iron-responsive acclimatory processes in *C. reinhardtii* involve the photosynthetic complexes. For instance, if provided a fixed carbon source (acetate), *C. reinhardtii* sacrifices non-essential iron-dependent proteins in the chloroplast, such as cytochrome b_6f, Photosystem I (PSI), and ferredoxin to spare iron and recycle this cofactor to essential iron-dependent proteins (Moseley et al., 2002) (Fig. 3). Cytochrome b_6f, in particular, contains eight hemes per functional dimer (Hurt & Hauska, 1981), suggesting that degradation of this complex during iron-limited photoheterotrophic growth results in a bolus of heme in the chloroplast which must be degraded to prevent heme toxicity and to

enable iron-recycling (Fig. 3). The identification of LFO1 as an iron-regulated heme oxygenase suggests that LFO1 may play a role in heme degradation in order to release its iron to aid in the acclimation of *C. reinhardtii* to low iron conditions. By producing a degradation product unique from the HO-1 family, LFO1 would be able to recycle heme during Fe-deficiency without eliciting the bilin retrograde signalling pathway that involves HMOX1 in the Chlamydomonas chloroplast (Duanmu et al., 2012) (Fig. 21).

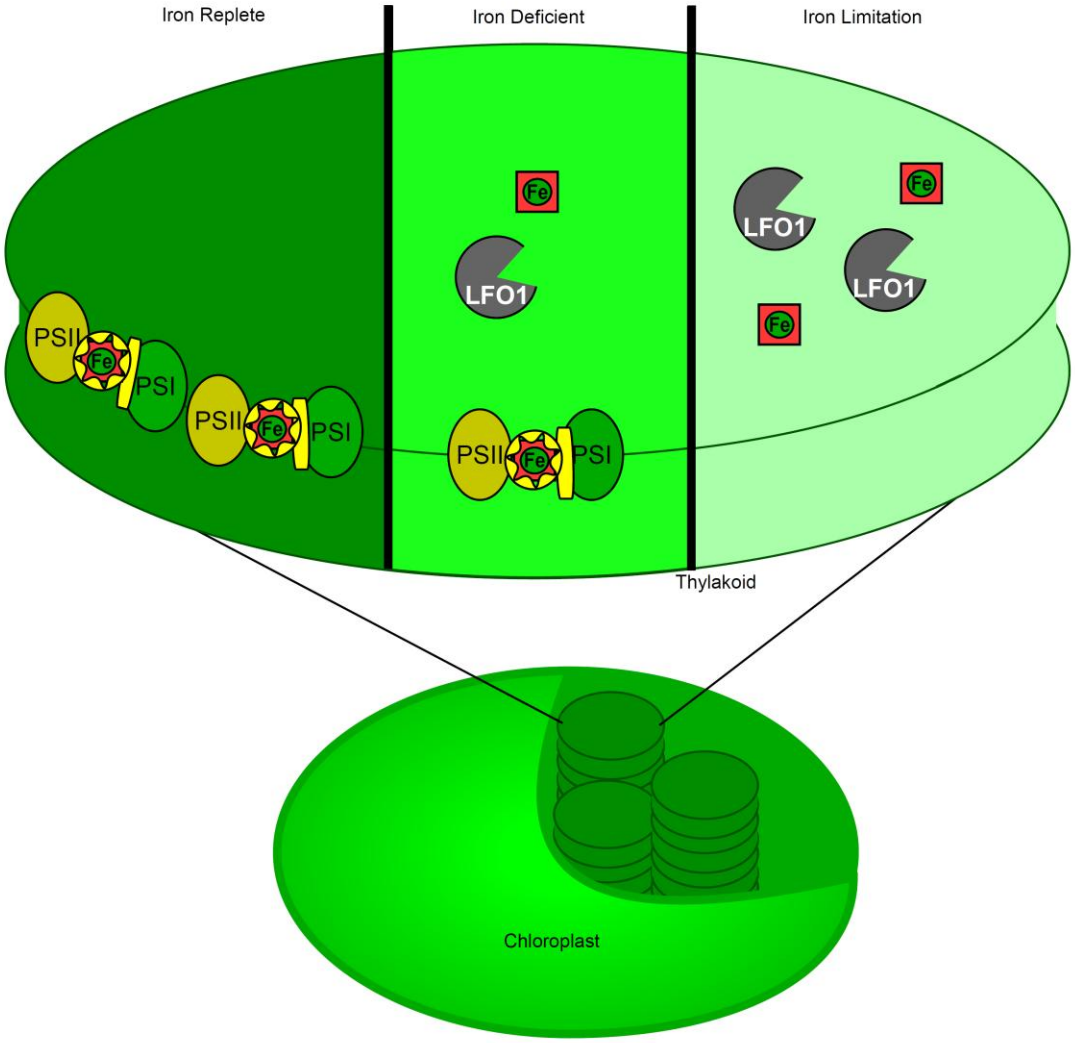


Fig 21. Hypothesis for the physiological role of LFO1 in *C. reinhardtii*. Since the increase in *lfo1* transcript expression levels increases with the degradation of the photosystems and cytochrome b_6f complex, I hypothesize that the role of LFO1 is to degrade the heme released by the cytochrome complex to allow for iron recycling to occur within the chloroplast.

A version of the following section (*Chapter III, Fur regulation of Staphylococcus aureus heme oxygenases is required for heme homeostasis*) has been submitted for publication.

CHAPTER III

FUR REGULATION OF *STAPHYLOCOCCUS AUREUS* HEME OXYGENASES IS REQUIRED FOR HEME HOMEOSTASIS

Introduction

Heme is an important small molecule and an essential cofactor for a variety of enzymes (George, 1948; Igo et al., 1961; Maehly, 1952; Morrison & Stotz, 1955). During cellular respiration, heme populates cytochromes and serves as an electron acceptor in the electron transport chain (Hammer et al., 2013; Morrison & Stotz, 1955). Heme-dependent respiration is critical for many organisms (Musser & Chan, 1998). If heme is unable to populate the cytochromes, either due to genetic inactivation of the cytochromes or a lack of cellular heme, cells are unable to respire and must switch to a fermentative state (von Eiff et al., 1997). Fermentation through glycolysis results in the production of 2 ATP molecules, compared to respiration that can generate up to 38 molecules of ATP per molecule of glucose.

Staphylococcus aureus is a Gram-positive coccoid bacterium and is the leading cause of skin and soft tissue infections (Klevens et al., 2007). In order to meet the cellular requirements for heme, *S. aureus* both biosynthesizes heme and imports heme from the extracellular milieu (Mazmanian et al., 2003; Tien & White, 1968). Heme import is mediated through the high-affinity iron-regulated surface determinant (Isd) heme acquisition system (Mazmanian et al., 2003; Torres et al., 2006). The genes in the *isd* operons are regulated by the Ferric Uptake Regulator (Fur) (Mazmanian et al., 2003; Xiong, Singh, Cabrera, & Jayaswal, 2000). Fur dimerizes when iron is present to bind Fur boxes in the promoter regions of target genes and repress transcription. This

repression is alleviated under iron deplete conditions, when there is insufficient intracellular iron to allow Fur dimerization (Fig. 22) (Bagg & Neilands, 1987). Regulation by Fur is widely conserved throughout bacterial species, and Fur regulates a variety of transcripts associated with pathogenesis (Beall & Sanden, 1995; Litwin, Boyko, & Calderwood, 1992; Tanui, Shyntum, Priem, Theron, & Moleleki, 2017; Torres et al., 2010; Tsolis, Baumler, Stojiljkovic, & Heffron, 1995).

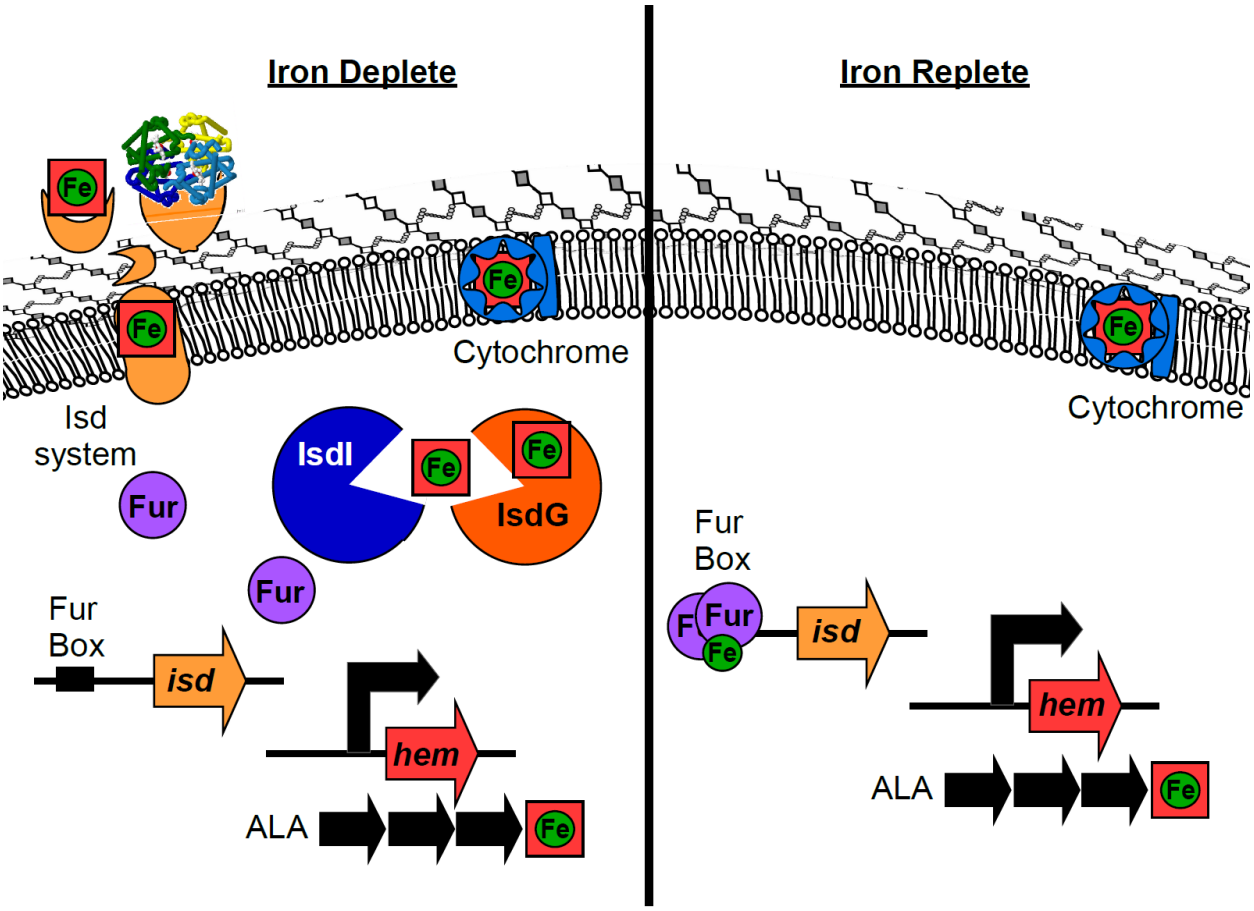


Fig 22. Iron acquisition in *S. aureus* occurs through the Fur regulated Isd system. Heme is imported into *S. aureus* through the iron-regulated surface determinant (Isd) system. The *isd* genes are regulated by the ferric uptake regulator (Fur), such that the genes are expressed under iron depletion conditions and transcription is repressed under iron replete conditions. Additionally, *S. aureus* can biosynthesize its own heme endogenously within the cell through the heme biosynthesis proteins.

In addition to being an important enzymatic cofactor, heme can also be used as a source of iron. Vertebrate-associated microorganisms, especially pathogenic bacteria, exploit host heme as a nutrient source. Aerobically, heme degradation is performed by heme oxygenases, while anaerobic bacteria use enzymes that rely on radical catabolism (LaMattina et al., 2016). Once heme is imported into the cell through the Isd proteins, heme is used to populate heme-binding proteins or heme is degraded by heme degrading enzymes. *S. aureus* encodes two such heme degrading enzymes, the heme oxygenases IsdI and IsdG. IsdG and IsdI facilitate the degradation of heme to produce the secondary catabolites staphylobilin and formaldehyde (Matsui et al., 2013; Reniere et al., 2010). This degraded heme also results in the release of iron, which bacteria can use to meet their iron requirements (Skaar, Humayun, et al., 2004).

Here I describe work initiated to understand the role of the heme degradation products in the context of *S. aureus*. Through an RNA-Sequencing experiment comparing a *S. aureus* strain containing constitutive heme oxygenase activity to one lacking heme degradation, I found a significant increase in a number of transcripts from genes associated with oxygen-independent energy production. This led to the hypothesis that dysregulation of *isdI* causes aberrant degradation of the intracellular heme pool. Further analysis comparing endogenously and constitutively expressed *isdI* and *isdG* showed a significant decrease in intracellular heme levels and heightened fermentation. This work demonstrates the importance of Fur regulation for optimal bacterial growth in low iron conditions.

Materials and methods

Bacterial Strains and Growth Conditions- The *Staphylococcus aureus* strain Newman was used for all experiments (E.S. Duthie & L.L. Lorenz, 1952). The Δ*isdGI* strain was made previously, as well as the Δ*isdGI* strain containing p*lgt.isdI* and p*lgt.isdG* (Reniere & Skaar, 2008). The endogenous promoter for *isdG* was defined as the intergenic region between the *isdA* and *isdB* operons. The endogenous promoter for *isdI* was defined as the intergenic region between *orfX* (NWMN_0112) and *aldA* (NWMN_0113). However, the Fur box for *isdI* is contained within the 3' region of the *orfX* operon, and this region was also included in the constructed promoter for *isdI*. Each of these promoter regions were amplified from Newman and were combined with either the gene encoding *isdG* or *isdI* with the Fur box and 3' region of *orfX*. These amplicons as well as p*OS1* plasmid were combined using Gibson cloning. Bacteria were grown overnight at 37°C with shaking at 180 r.p.m. in tryptic soy broth (TSB; BD) and 10 μg/mL chloramphenicol unless otherwise stated.

RNA-Sequencing- The Δ*isdGI* p*lgt* and Δ*isdGI* p*lgt.isdI* were grown in biological triplicate in chelex (Sigma) treated Roswell Park Memorial Institute (RPMI) medium (Corning) with 0.75 mM Ethylenediamine-N,N-bis(2-hydroxyphenylacetic acid) (EDDHA: LGC Standards), 10 μg/mL chloramphenicol, and 1 μM porcine hemin (Sigma). EDDHA was resuspended in 100% ethanol to a stock concentration of 100 mM. The cultures were grown at 37°C shaking at 180 r.p.m. to an OD600 of 0.4. At the correct density, a 1:1 solution of Acetone:Ethanol was added at an equal volume to each sample. These samples were stored at -80°C until the RNA was extracted. Cells were resuspended in 750 mL of LETs buffer (1 M LiCl, 0.5 M EDTA, 1 M Tris pH 7.4) and added

to tubes holding lysing matrix B (MP Biomedicals). A FastPrep-24 (MP Biomedicals) bead beater was used to lyse cells by bead beating for 45 seconds at 6 m/s. Samples were then heated at 55°C for 5 min then centrifuged for 10 min. The upper phase was transferred to a fresh tube and 1 mL TRIzol (Thermo Scientific) was added. A 200 uL aliquot of chloroform was added to each tube and vortexed. The mixed samples were then centrifuged for 15 minutes to separate the aqueous and organic layers. The aqueous (upper) layer was then aliquoted into a fresh tube. The RNA was precipitated by the addition of 1 mL of isopropyl alcohol and incubated for 10 min at room temperature. Samples were then centrifuged for 10 min and the supernatant was removed. The RNA was washed with 200 µL of 70% ethanol. The RNA pellet was dried for 1 minutes, then resuspended in 100 µL of water. DNA contamination was eliminated through the addition of 8 µL RQ1 DNase, 12 µL of 10x RQ1 buffer, and 2 µL RNase inhibitor (Promega) to the purified RNA. After DNase treatment for 2 hours, the samples were further purified using the RNeasy miniprep RNA cleanup kit (Qiagen). The RNA concentration was determined by using the Synergy 2 with Gen 5 2.1 software (BioTek).

Library creation and RNA-Sequencing was performed by Vanderbilt Technologies for Advanced Genomics Core Facility (VANTAGE) similarly to previous descriptions (Mortensen, Rathi, Chazin, & Skaar, 2014). Briefly, 1.5 µg of RNA from each sample was submitted and an Agilent Bioanalyzer Nano RNA chip was used to determine the quality of the RNA. A Ribo-Zero kit was then used to remove ribosomal RNA by following the manufacturer's protocol. After cDNA creation, samples were pooled for multiplexing on a flow cell then loaded on an Illumina HiSeq 2500 for a 50-bp paired-end run. Raw reads were processed for FASTQ conversion, and only reads passing the pass filter were used for further analysis.

RNA-Sequencing Analysis- Rockhopper was used to analyze RNA-seq data, including reads alignment, quantification of transcript abundance, identification of differential expression and characterization of operon structures (McClure et al., 2013). The genome alignment figure (Fig 2A) was created in Circos. The RPKM values were set to a max of 2000 being shown. The heat map shows the log2-fold changes ranging from -5 to 5. The pathway analysis was performed using Kyoto Encyclopedia of Genes and Genomes (KEGG) Mapper to search the KEGG pathway databases of annotated *S. aureus* genes. Additional annotations were made through literature searches.

qRT-PCR- Samples were grown and RNA was extracted as described for the RNA-Seq experiment. cDNA was synthesized and qRT-PCR was performed as previously described (Mortensen et al., 2014). The assay was repeated to test three biological replicates in triplicate.

Growth curves- *S. aureus* Newman strains were grown in RPMI with 1% Casamino acids, 0.5 mM EDDHA, and 10 μg/mL chloramphenicol. Overnight cultures were diluted 1:50 in chelex treated RPMI with 100 μM $CaCl_2$, 25 μM $ZnCl_2$, 1 mM $MgCl_2$, 25 μM $MnCl_2$, 1% Casamino acids, 0.75 mM EDDHA, and 10 μg/mL chloramphenicol, either with or without 1 μM hemin. Growth curves were setup to a total volume of 200 μL 96 well flat bottom plates. The OD_{600} was monitored every 1 to 2 hours for up to 16 hours. For the glycerol growth curve, Newman strains were grown in biological triplicate overnight as above. Cultures were diluted 1:50 in RPMI without glucose with 22 mM glycerol and 10 μg/mL chloramphenicol. The assay was repeated to test three biological

replicates in triplicate. For the glucose growth curve, Newman strains were grown in biological triplicate overnight as above. Cultures were diluted 1:50 in RPMI without glucose with 11 mM glucose and 10 μg/mL chloramphenicol. The assay was repeated to test three biological replicates in triplicate.

Heme quantification- *S. aureus* Newman strains were grown overnight in TSB. Cultures were diluted 1:200 in TSB and grown aerobically at 37°C while shaking at 180 r.p.m. to an OD_{600} of 0.4 – 0.45. Heme was quantified as previously described (Mike et al., 2013). Briefly, the pellets were resuspended in 20 mM potassium phosphate containing 20 μL of lysostaphin. Samples were incubated at 37°C for 30 minutes, then lysed by sonication. Unbroken protoplasts were removed by centrifugation, the protein concentration of each sample was determined using a Pierce BCA Protein Assay Kit. A 450 μL aliquot of each sample was transferred to a cuvette, then heme was extracted by adding 0.2 M NaOH and 40% pyridine with 200 μM potassium ferricyanide in 450 μL. The absorbance was measured from 540-590 nm. Then 10 μL of 0.5 M DTT was added, then the absorbance was again measured from 540-590 nm. The heme quantity was then calculated using the coefficient of 324 mM^{-1} cm^{-1}.

Lactate assay- *S. aureus* strains were grown to mid-log phase (OD_{600}=0.400) in 20 mL TSB with 10 μg/mL chloramphenicol. Samples were centrifuged and the supernatant was collected. L-Lactate levels were determined using the r-Biopharm D-Lactic acid/L-Lactic acid kit. 13.4 μL of supernatant was added to 134 μL of glycylglycine buffer, 26.8 μL nicotinamide-adenine dinucleotide (NAD), and 2.68 μL of glutamate-pyruvate transaminase suspension to a final volume

of 298 μL in a 96 well plate. These samples were incubated for 5 min then the absorbance at 340 nm was read (A_1). Then 2.68 μL of L-lactacte dehydrogenase solution was added. The samples were then mixed and incubated for 30 min then the absorbance at 340 nm was read a second time (A_2). The concentration of L-lactic acid was determined by subtracting A_1 from A_2 for each samples then multiplying the A_Δ by the final volume times the molecular weight of L-lactic acid, then dividing by the extinction coefficient of NADH at 340 nm.

Gentamicin sensitivity assay- *S. aureus* strains were grown overnight in 5 mL TSB with 10 μg/mL chloramphenicol. 200 μL of overnight culture were plated onto Tryptic Soy Agar (TSA) plates containing 10 μg/mL chloramphenicol. One gentamicin Etest strip (Biomerieux) was applied to each plate. The plates were incubated for ~16 hours at 37°C, then the minimum inhibitory concentration was determined by the concentration at which there was no zone of clearing.

Table 1. Bacterial strains.

Bacterial Strain	Description	Reference
Wild-type	*Staphylococcus aureus* strain Newman	(E. S. Duthie & L. L. Lorenz, 1952)
Δ*isdGI*	*S. aureus* strain Newman with *isdI* replaced with *tetM*, and *isdG* replaced with *erm*	(Reniere & Skaar, 2008)

Table 2. Primers.

Primer Name	Purpose	Sequence
Native promoter *isdI* F	Gibson cloning 5’ – pOS1 3’ – *isdI* promoter	GCCTTAAAGACGATCCGGGGATTATCTCTC CTTTGTTTATTGG
Native promoter *isdI* R	Gibson cloning 5’ – pOS1 3’ – *isdI* promoter	TTACAATGTATGAAAAAACTCCATTCTAGTA CTAATG
Fur box *isdI* F	Gibson cloning 5’ – pOS1 3’ – *isdI* fur box	AGTTTTTTCATACATTGTAATAGGTTTTTCA ATTTATATG
Fur box *isdI* R	Gibson cloning 5’ – pOS1 3’ – *isdI* fur box	CTTGTTTGGATCCTCGAGCACTAGGGGTTT TTAATATATTTATTTTTGATAG
Native promoter *isdG* F	Gibson cloning 5’ – pOS1 3’ – *isdG* promoter	GCCTTAAAGACGATCCGGGGCATGTTGTTT TCCTCCTAAG
Native promoter *isdG* R	Gibson cloning 5’ – pOS1 3’ – *isdG* promoter	ATAAATTTCATAACATAATCCTCCTTTTTAT GATTG
nirR F	qPCR	CGT CCG TGC CTC ATG CCA TGT GC

nirR R	qPCR	GGC ACA CGA AGA GTT ACA AAT ATG CGC TC
nirB F	qPCR	GGC ACA TGA TAT TTC TTA GCC ACT TCT CC
nirB R	qPCR	CGG TGG TAA TGG TGG TAC AGA AG
nirD F	qPCR	CCTTTTTTCCAATTAGGGGTGTTAATTC
nirD R	qPCR	GTT TAT GAG GTA GAA GTT ACA GAC GGG

Results

Expression of heme oxygenases in S. aureus is essential for growth with heme as the sole source of iron.

The function of the heme degradation products in *S. aureus* are unclear. To try to elucidate the roles of these catabolites in the cell, I created two strains of *S. aureus*; one lacking heme oxygenase activity (Δ*isdGI* p*lgt*) and one containing the endogenous heme oxygenase *isdI* added back *in trans* under the constitutive expression of the promoter for lipoprotein diacylglycerol transferase (*lgt)* (Selvan & Sankaran, 2008) (Δ*isdGI* p*lgt.isdI*). Providing *isdI in trans* complemented the growth defect of Δ*isdGI* p*lgt* when grown in the presence of heme as the sole source of iron (Fig. 23). These data indicate that the expression of IsdI is necessary for growth when heme is the sole source of iron, and provide an experimental condition where heme is actively degraded in the presence of the heme oxygenases but it is not degraded in their absence.

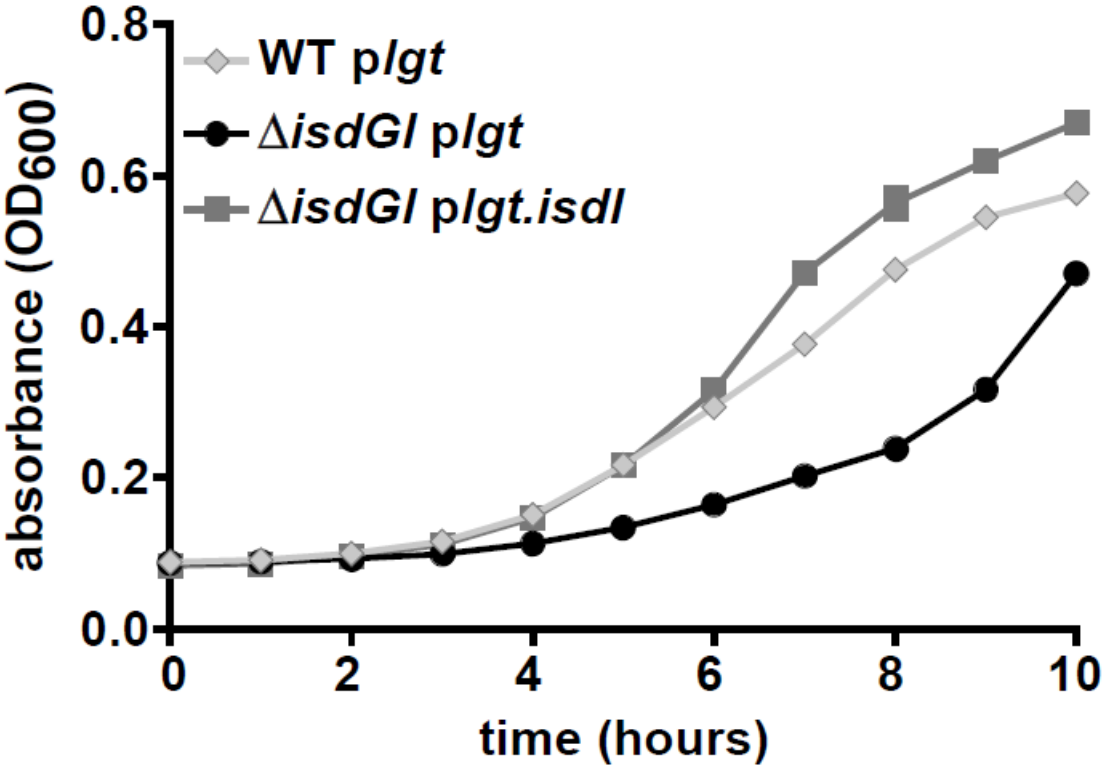


Fig 23. Expression of isdI is required for growth of *S. aureus* with heme as the sole source of iron. Growth of *S. aureus* strains in chelex treated RPMI with 0.75 mM EDDHA and 1 μM heme. The addition of p*lgt.isdI* to a strain lacking heme oxygenase activity (Δ*isdGI* p*lgt*) complements the growth of WT plgt.

Constitutive expression of isdI leads to an increase in expression of transcripts associated with oxygen-independent energy production.

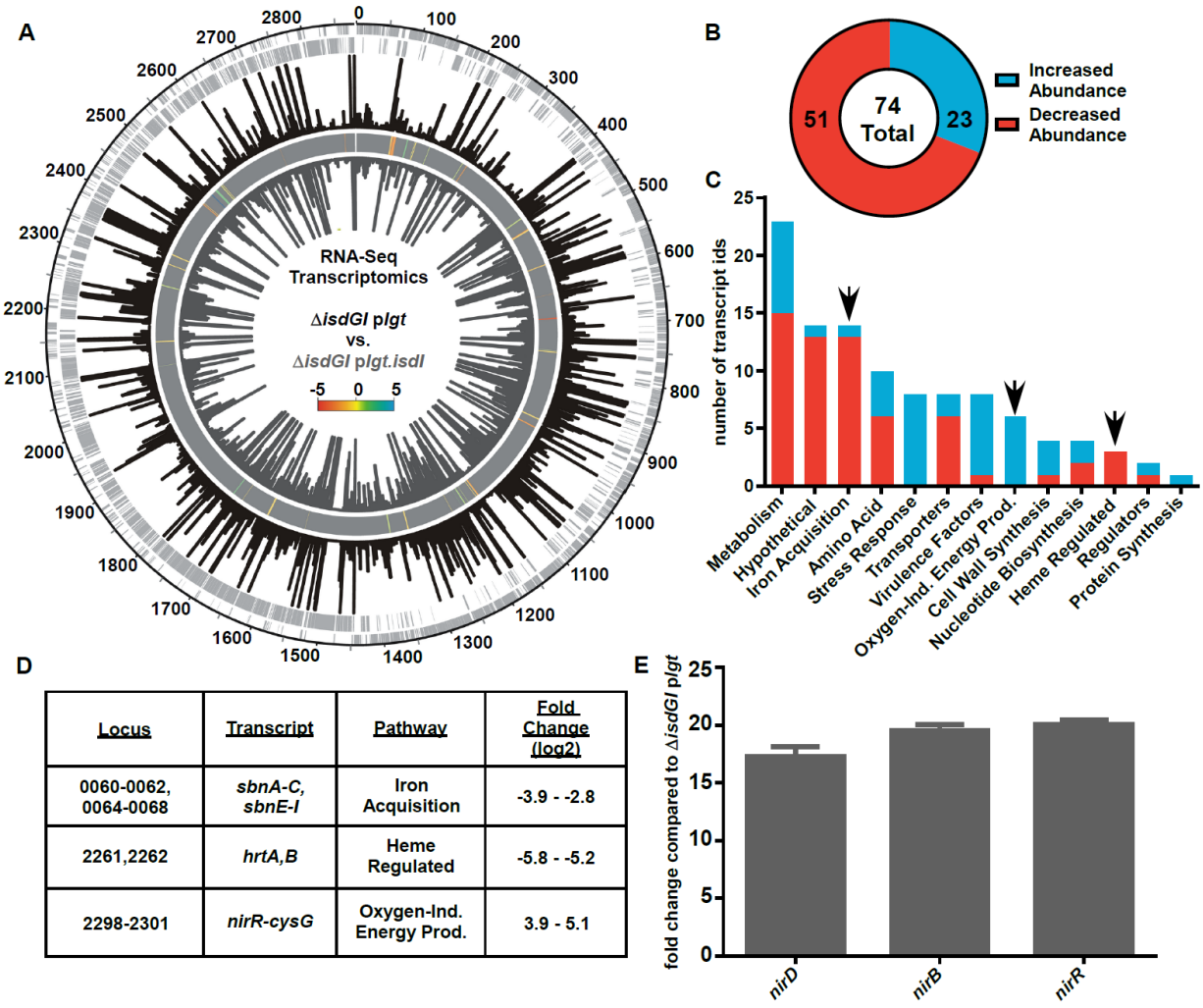
In order to elucidate the role of heme catabolites in *S. aureus*, Δ*isdGI* p*lgt* and Δ*isdGI* p*lgt.isdI* were grown with heme as the sole iron source. Samples were collected for RNA isolation at mid-log phase growth and RNA from these samples was submitted for RNA-Sequencing analysis. In total, the abundance of 74 transcripts were significantly different between Δ*isdGI* p*lgt* and Δ*isdGI* p*lgt.isdI* ($q \leq 0.05$) (Fig. 24A). Of these, 23 were increased in abundance and 51 were decreased in abundance (Fig. 24B). As expected, there were significant changes in abundance of transcripts for genes associated with iron acquisition and genes known to be regulated by heme, such as the heme regulated transporter (*hrtAB*) genes, which are regulated by the heme sensing two component system (HssRS) (Torres et al., 2010; Torres et al., 2007) (Fig. 24C). However, there was also a significant increase in the expression of genes associated with oxygen-independent energy production (Fig. 24C,D, Table 3). This result was surprising because heme oxygenases require oxygen to function (Streit et al., 2016) and the cultures were grown aerobically. To confirm the results of the RNA-Seq, quantitative Real Time Polymerase Chain Reaction (qRT-PCR) was performed on transcripts from a subset of the genes linked with oxygen-independent energy production. The nitrite reductase operon contains three genes (*nirD, nirB,* and *nirR*) known to be upregulated under anaerobic conditions in various bacteria, including the closely related *Staphylococcus carnosus* (Goldman & Roth, 1993; Jayaraman, Peakman, Busby, Quincey, & Cole, 1987; Neubauer, Pantel, & Gotz, 1999). Analysis by qRT-PCR confirmed that these transcripts had a 17 – 20 fold increase in transcript abundance in the Δ*isdGI* p*lgt.isdI* strain

compared to Δ*isdGI* p*lgt* strain (Fig. 24E). Due to this increased expression of genes associated with oxygen-independent energy production, I hypothesized that the over-expression of *isdI* under the control of *lgt* promoter may lead to inappropriate production of the heme oxygenase, unregulated heme degradation, and disruption of heme homeostasis.

Table 3. RNA-Sequencing Δ*isdGI* p*lgt* vs Δ*isdGI* p*lgt.isdI* differentially abundant transcripts.

Gene	Activity	Fold Change	q Value
NWMN_0060	O-acetyl serine sulfhydrylase	0.356012549	5.57908E-34
NWMN_0061	ornithine cyclodeaminase	0.346277367	4.89746E-40
NWMN_0062	siderophore biosynthesis IucC family protein	0.293208737	1.0597E-63
NWMN_0064	siderophore biosynthesis IucA family protein	0.275476279	1.39277E-71
NWMN_0065	siderophore biosynthesis IucC family protein	0.25348987	4.1114E-100
NWMN_0066	2-dehydro-3-deoxyglucarate aldolase	0.277392368	2.40852E-69
NWMN_0067	diaminopimelate decarboxylase	0.267943366	1.50468E-84
NWMN_0068	hypothetical protein	0.267943366	8.63096E-86
NWMN_0094	bifunctional acetaldehyde-CoA/alcohol dehydrogenase	3.758090997	2.56647E-39
NWMN_0111	heme-degrading monooxygenase IsdI	16.44982123	0
NWMN_0113	aldehyde dehydrogenase-like protein	0.346277367	3.78535E-43
NWMN_0135	N-acetylmuramic acid-6-phosphate etherase	2.0139111	1.75074E-12
NWMN_0196	murein hydrolase regulator LrgA	2.265767771	4.2533E-08
NWMN_0197	antiholin-like protein LrgB	2.173469725	0.011439691
NWMN_0208	hypothetical protein	0.291183397	5.76115E-40
NWMN_0390	superantigen-like protein	2.313376368	1.52214E-21
NWMN_0391	superantigen-like protein	2.808889751	1.00768E-29
NWMN_0418	NADH dehydrogenase subunit 5	0.216134308	1.947E-100
NWMN_0419	hypothetical protein	0.323088208	1.87197E-42
NWMN_0484	transcriptional regulator CtsR	0.320856474	4.54721E-48
NWMN_0485	UvrB/UvrC motif-containing protein	0.299369676	1.56726E-68
NWMN_0486	ATP:guanido phosphotransferase	0.397768242	5.17828E-21
NWMN_0542	hypothetical protein	0.160428237	9.48E-288
NWMN_0601	MntA	0.10153155	0
NWMN_0602	MntC	0.10881882	0
NWMN_0603	MntB	0.102237757	0
NWMN_0676	hypothetical protein	2.128740365	9.74931E-07
NWMN_0679	hypothetical protein	0.463294031	7.37849E-17
NWMN_0680	6-pyruvoyl tetrahydropterin synthase	0.456915725	2.72517E-14
NWMN_0681	hypothetical protein	0.450625231	1.47336E-16
NWMN_0831	argininosuccinate lyase	0.476318999	1.46232E-12
NWMN_0832	argininosuccinate synthase	0.473028823	1.16534E-19

Gene	Activity	Fold Change	q Value
NWMN_1040	iron-regulated heme-iron binding protein IsdB	0.193445624	1.7197E-172
NWMN_1041	iron-regulated heme-iron binding protein IsdA	0.38958229	2.09204E-16
NWMN_1046	NPQTN-specific sortase B	0.5	7.70631E-12
NWMN_1047	heme-degrading monooxygenase IsdG	0.463294031	2.28279E-11
NWMN_1067	formyl peptide receptor-like 1 inhibitory protein	2.428389769	6.79307E-17
NWMN_1075	superantigen-like protein	2.549121255	5.66651E-22
NWMN_1076	superantigen-like protein	3.052518418	6.05852E-35
NWMN_1077	superantigen-like protein	3.458148925	6.66887E-49
NWMN_1084	anti protein	4.563054863	4.14097E-84
NWMN_1109	bifunctional pyrimidine regulatory protein PyrR uracil phosphoribosyltransferase	0.381564802	9.15781E-43
NWMN_1110	uracil permease	0.368567304	2.13945E-37
NWMN_1198	pyruvate flavodoxin ferredoxin oxidoreductase, alpha subunit	0.453759578	3.50607E-12
NWMN_1199	2-oxoglutarate ferredoxin oxidoreductase subunit beta	0.435275282	2.69048E-19
NWMN_1240	homoserine dehydrogenase	2.713208655	5.80463E-28
NWMN_1241	threonine synthase	2.084931522	4.05317E-11
NWMN_1392	ferredoxin	0.395020656	1.86594E-28
NWMN_1483	molecular chaperone DnaK	0.438302861	7.15287E-20
NWMN_1484	heat shock protein GrpE	0.466516496	4.71487E-19
NWMN_1485	heat-inducible transcription repressor	0.466516496	1.48426E-19
NWMN_1552	hypothetical protein	0.40332088	5.6114E-26
NWMN_1576	threonyl-tRNA synthetase	3.89061979	1.6749E-41
NWMN_1873	truncated beta-hemolysin	2.67585511	1.18205E-13
NWMN_1937	chaperonin GroEL	0.378929142	2.78788E-20
NWMN_1938	co-chaperonin GroES	0.40332088	1.68956E-24
NWMN_2056	glucosamine--fructose-6-phosphate aminotransferase	2.602683711	3.97904E-15
NWMN_2089	osmoprotectant transporter	0.420448208	1.36754E-23
NWMN_2110	alpha-acetolactate decarboxylase	0.450625231	3.45974E-14
NWMN_2111	acetolactate synthase	0.482968164	6.36497E-10
NWMN_2261	ABC transporter ATP-binding protein	0.192109398	1.2644E-125
NWMN_2262	hypothetical protein	0.173138684	7.7562E-191
NWMN_2265	hypothetical protein	0.318640157	1.90579E-42
NWMN_2266	hypothetical protein	0.236514412	2.05048E-74
NWMN_2288	nitrite transport protein	16.79546694	0
NWMN_2298	uroporphyrin-III C-methyl transferase	3.91768119	8.0484E-99
NWMN_2299	assimilatory nitrite reductase [NAD(P)H], small subunit	4.169863043	4.51604E-55
NWMN_2300	assimilatory nitrite reductase [NAD(P)H], large subunit	5.13370359	8.9742E-134
NWMN_2301	nitrite reductase transcriptional regulator NirR	4.856779538	1.5873E-99
NWMN_2315	phosphoglyceromutase	0.429282718	2.42433E-21
NWMN_2316	cation efflux family protein	2.549121255	2.84146E-20
NWMN_2318	gamma-hemolysin component A	0.473028823	3.84037E-16
NWMN_2456	hypothetical protein	0.175555609	3.3767E-138
NWMN_2587	hypothetical protein	0.26425451	5.70974E-83



Locus	Transcript	Pathway	Fold Change (log2)
0060-0062, 0064-0068	*sbnA-C, sbnE-I*	Iron Acquisition	-3.9 - -2.8
2261,2262	*hrtA,B*	Heme Regulated	-5.8 - -5.2
2298-2301	*nirR-cysG*	Oxygen-Ind. Energy Prod.	3.9 - 5.1

Fig. 24 Constitutive expression of isdI leads to an increase in transcripts associated with non-aerobic growth. RNA-Sequencing analysis comparing RNA from Δ*isdGI* p*lgt.isdI* to Δ*isdGI* p*lgt*. A, RPKM mapping of RNA-Seq data for Δ*isdGI* p*lgt* (black bars, fourth strip) and Δ*isdGI* p*lgt.isdI* (grey bars, sixth/inner strip) mapped relative to the chromosomal location (outer strip) for each of the positive strand (second strip) and negative strand (third strip) genes. A heat map (fifth strip) is included showing the log2-fold changes for transcripts with significant q-values. B, Pie chart showing the number of total differentially abundant transcripts, as well as the number with increased (blue) and decreased (red) abundance. C, Pathway analysis was performed for the differentially abundant transcripts with KEGG Pathway. The number of transcript identifications which fall into each pathway bin is shown, as well as whether the transcript has increased (blue) or decreased (red) abundance. D, Fold change values for a subset of transcripts identified in the RNA-Seq analysis from the iron acquisition, heme regulated, and oxygen-independent energy production pathways. E, qRT-PCR analysis of the nitrite reductase operon transcripts confirming their increased abundance in Δ*isdGI* p*lgt.isdI* compared to Δ*isdGI* p*lgt* (Student's T-test, $p<0.05$).

Constitutive expression of S. aureus heme oxygenase leads to heme-dependent growth inhibition.

Since the heme oxygenases of *S. aureus* are regulated by Fur, additional strains were created expressing both of the heme oxygenases, *isdI* and *isdG*, on the pOS1 plasmid controlled by either the *lgt* promoter (p*lgt*) or their endogenous *isdI*/*isdG* promoters (p*isd*). The strains were grown in heme as the sole source of iron. Under these conditions, the strains expressing the heme oxygenases from their native promoters (p*isd.isdI* and p*isd.isdG*) grew to slightly higher optical densities than the strains containing the constitutively expressed heme oxygenases (p*lgt.isdI* and p*lgt.isdG*) (Fig. 25B). Additionally, when the strains were grown under iron limiting media lacking heme, the strain expressing the constitutively expressed *isdI* grew worse than all of the other strains (Fig. 25A). Unlike the strain constitutively expressing *isdI*, the strain constitutively expressing *isdG* did not exhibit reduced growth under these conditions. This is consistent with the fact that IsdG requires heme for protein stability, and in the absence of heme, IsdG is rapidly degraded post-translationally (Reniere & Skaar, 2008). These data indicate that while heme oxygenases are active under iron limited conditions, constitutive *isdI* expression results in reduced growth.

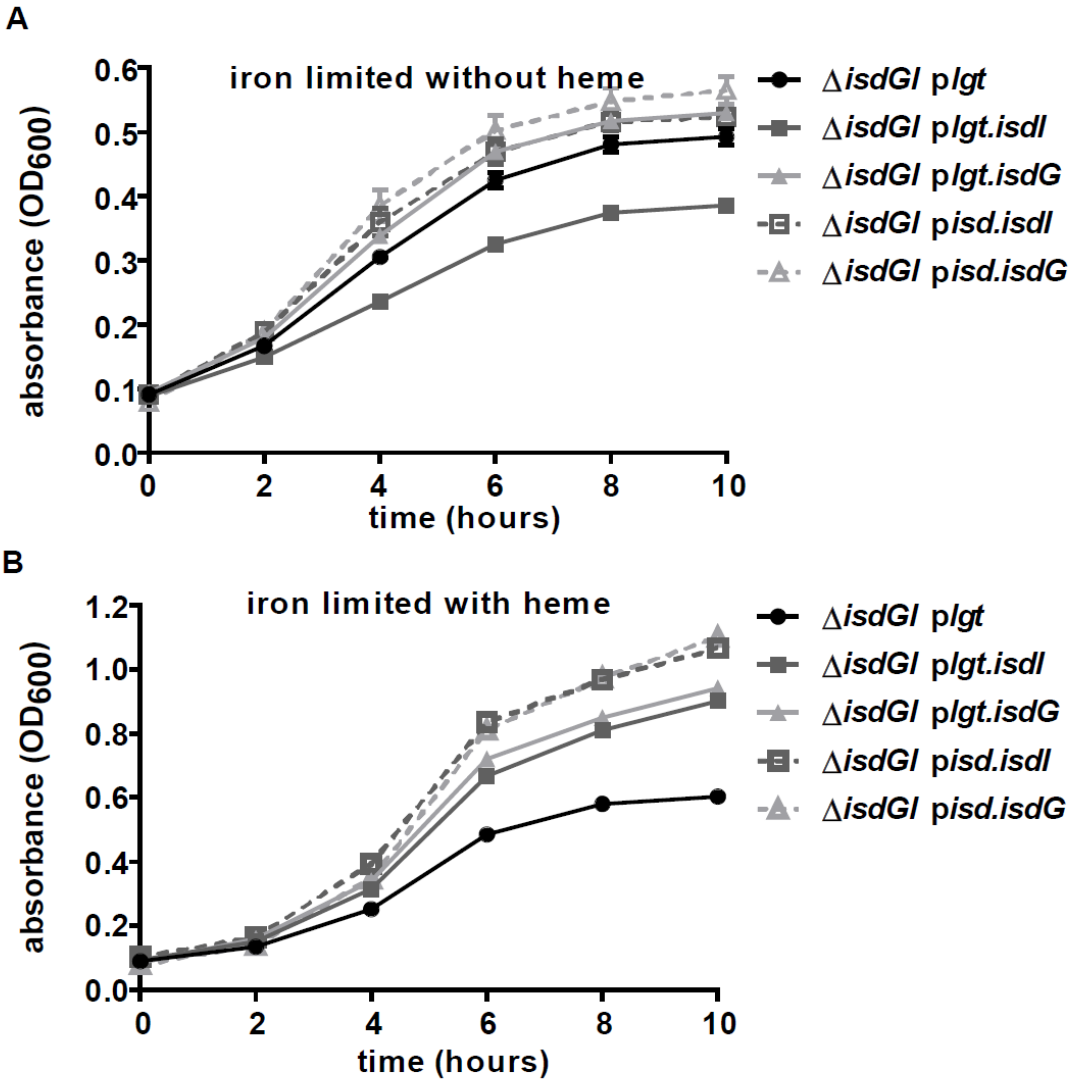


Fig 25. Misregulation of *isdI* expression leads to heme-dependent growth inhibition. Growth of *S. aureus* strains in chelex treated RPMI with 0.75 mM EDDHA either in the absence (A) or presence (B) of 1 μM heme. A, Under iron deplete conditions lacking heme, Δ*isdGI* p*lgt.isdI* (dark grey) exhibits reduced growth compared with any of the other strains. B, When grown in the iron limited conditions in the presence of heme, the endogenously expressed strains (Δ*isdGI* p*isd.isdI* and Δ*isdGI* p*isd.isdG*) exhibit enhanced growth as compared to constitutively expressed strains (Δ*isdGI* p*lgt.isdI* and Δ*isdGI* p*lgt.isdG*).

Heme levels are decreased in strains with constitutively expressed heme oxygenases.

I hypothesized that the decreased growth in the constitutively expressed heme oxygenase-containing strain is due to inappropriate heme degradation. This hypothesis predicts that these strains contain lower levels of cellular heme. However, the levels of heme in these strains under the low iron conditions were below the limit of detection of our assays. Therefore, the strains were grown under iron replete conditions. Since the heme oxygenases are expressed from pOS1, a high copy number plasmid, I performed immunoblot analysis to ensure that the enzymes expressed from their endogenous promoters were suppressed by Fur under these iron replete conditions. There is a significant increase in IsdI (2.4 fold) and IsdG (2.5 fold) abundance in Δ*isdGI* p*lgt.isdI* and Δ*isdGI* p*lgt.isdG* compared to Δ*isdGI* p*isd.isdI* and Δ*isdGI* p*isd.isdG*, respectively, suggesting that Fur is actively repressing in these conditions (Fig. 26). However, since I did not detect protein expression of heme oxygenases expressed from their endogenous promoters, the amount of Fur in the cell may be insufficient to fully repress transcription.

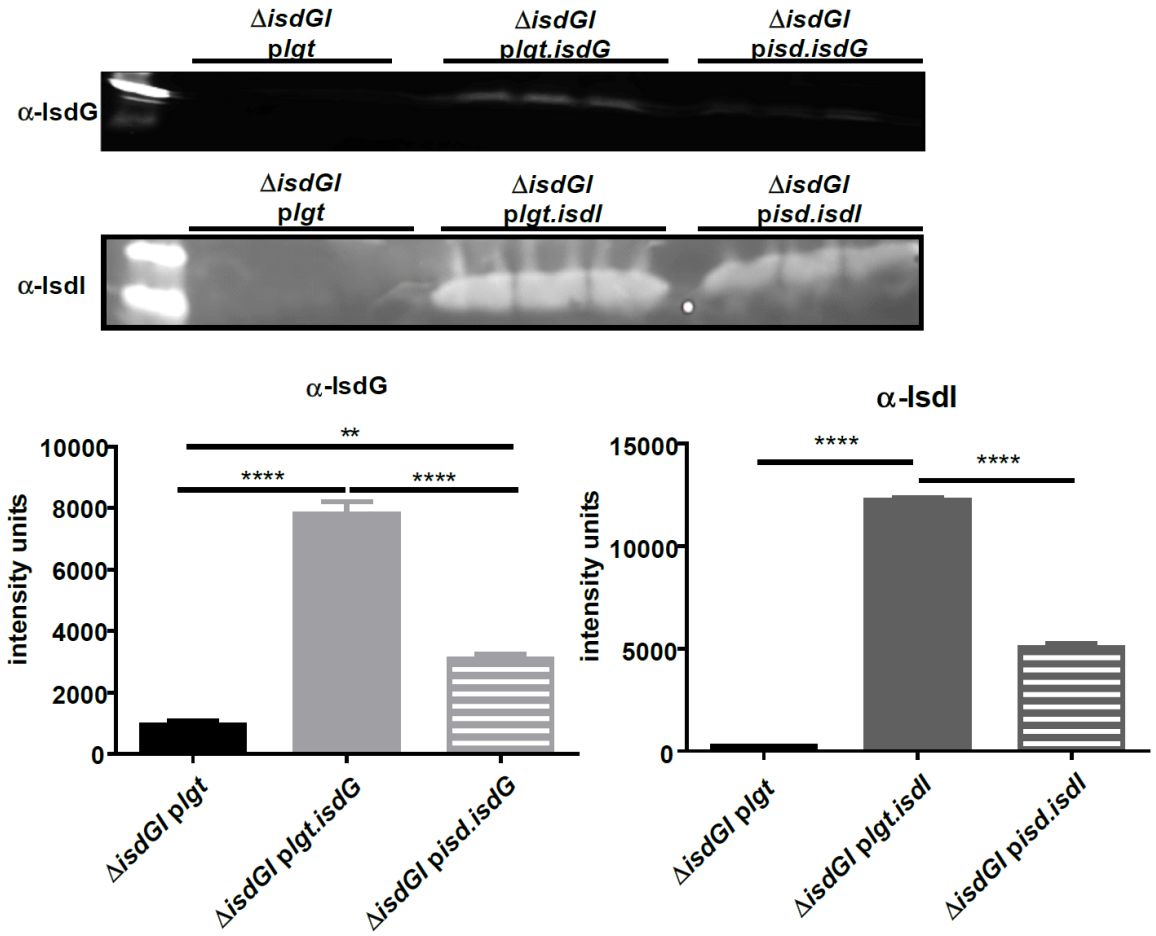


Fig 26. Western blots for IsdI and IsdG. For both gels, 40 µg of cytoplasmic fraction were run for each strain : lane 1 is the ladder, lane 2-4 are Δ*isdGI* p*lgt*, lanes 6-8 are the constitutively expressed heme oxygenase containing strains (Δ*isdGI* p*lgt.isdG* Top, Δ*isdGI* p*lgt.isdI* Bottom), lane 10-12 are the endogenously expressed heme oxygenase containing strains (Δ*isdGI* p*isd.isdG* Top, Δ*isdGI* p*isd.isdI* Bottom). Each of the bands were quantified in ImageJ, and the quantification is shown below (1way-ANOVA, $p<0.05$).

Using the iron replete conditions established above, I performed heme quantification on each of the strains harvested at mid-log phase growth. Both Δ*isdGI* p*lgt.isdI* and Δ*isdGI* p*lgt.isdG* contained significantly less heme than Δ*isdGI* p*lgt* (32.1-51.2%), as well as Δ*isdGI* p*isd.isdI* and Δ*isdGI* p*isd.isdG*, respectively (Fig. 27). These data indicate that dysregulation of heme oxygenase results in significantly less heme in *S. aureus*, potentially leading to alterations in cell physiology.

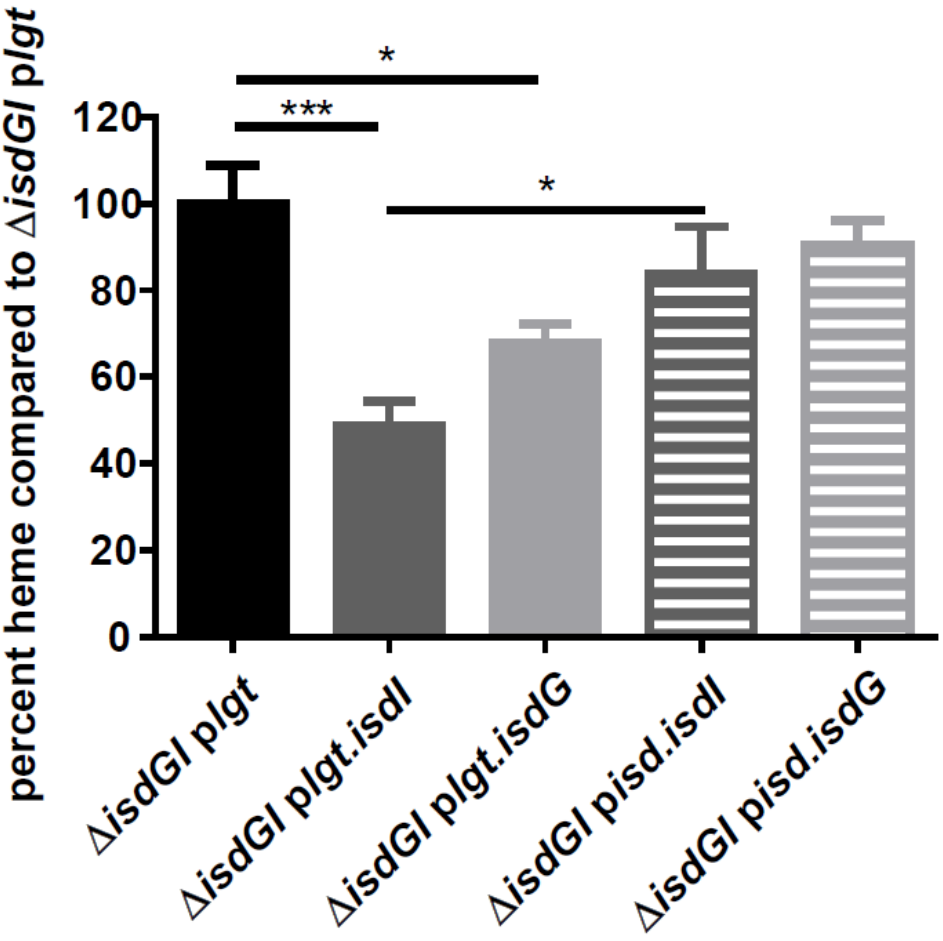


Fig 27. Heme levels are decreased in strains with constitutive heme oxygenase expression. Quantification of cellular heme levels of the heme oxygenase strains compared to the strain lacking heme oxygenase activity (Δ*isdGI* p*lgt*). The Δ*isdGI* p*lgt.isdI* (dark grey, solid) strains contains significantly less heme than Δ*isdGI* p*lgt* (black) and Δ*isdGI* p*isd.isdI* (dark grey, dashed) strains. Additionally, the Δ*isdGI* p*lgt.isdG* (light grey, solid) strain contains significantly less heme than Δ*isdGI* p*lgt* (1way ANOVA, $p<0.05$).

Strains with constitutively expressed heme oxygenases undergo fermentation.

Since cells lacking heme are unable to respire, and there were lower levels of heme in Δ*isdGI* p*lgt.isdI* and Δ*isdGI* p*lgt.isdG*, I hypothesized that these strains may be forced to rely on fermentation for energy production. To determine if the strains containing the constitutively expressed heme oxygenases generate energy through fermentation, the fermentation product lactate was measured from the supernatant of cells grown under iron replete conditions. Δ*isdGI* p*lgt.isdI* and Δ*isdGI* p*lgt.isdG* produce higher levels of lactate (23.7-26.5% increase) than Δ*isdGI* p*lgt*, whereas Δ*isdGI* p*isd.isdI* and Δ*isdGI* p*isd.isdG* do not have significant increases in lactate production (Fig. 28).

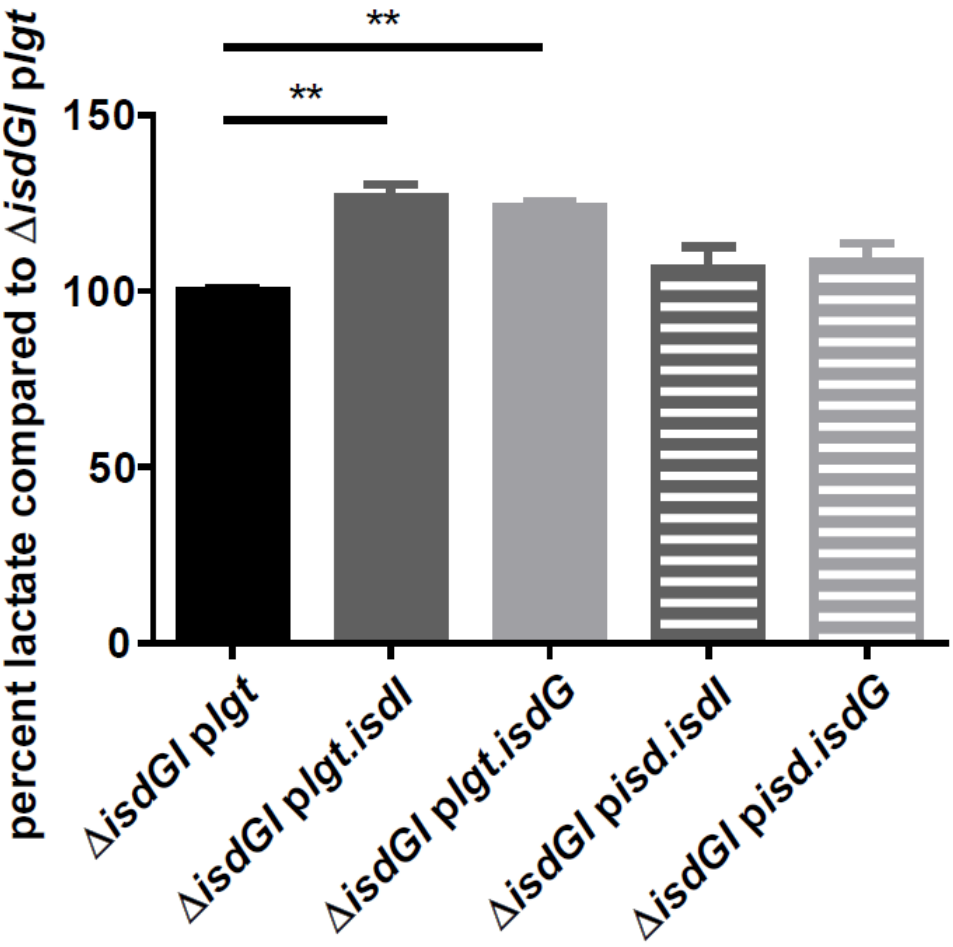


Fig 28. Constitutive expression of the heme oxygenases leads to an increase in extracellular lactate levels. Supernatant lactate levels were quantified and graphed as percent lactate compared to Δ*isdGI* p*lgt*. Both Δ*isdGI* p*lgt.isdI* (dark grey, solid) and Δ*isdGI* p*lgt.isdG* (light grey, solid) supernatants contain significantly more lactate than the Δ*isdGI* p*lgt* (black) supernatant (2way-ANOVA, $p<0.05$).

In addition to lactate production, fermenting cells exhibit decreased proton motive force (Bryan & Kwan, 1983). Gentamicin requires functional respiration within *S. aureus* to be bactericidal (Bryan & Kwan, 1983). When exposed to gentamicin, Δ*isdGI* p*lgt.isdI* had a 2 fold increased minimum inhibitory concentration (MIC) than Δ*isdGI* p*lgt* and Δ*isdGI* p*isd.isdI* (Fig. 29), suggesting constitutive expression of the heme oxygenases reduces the proton motive force within the cell.

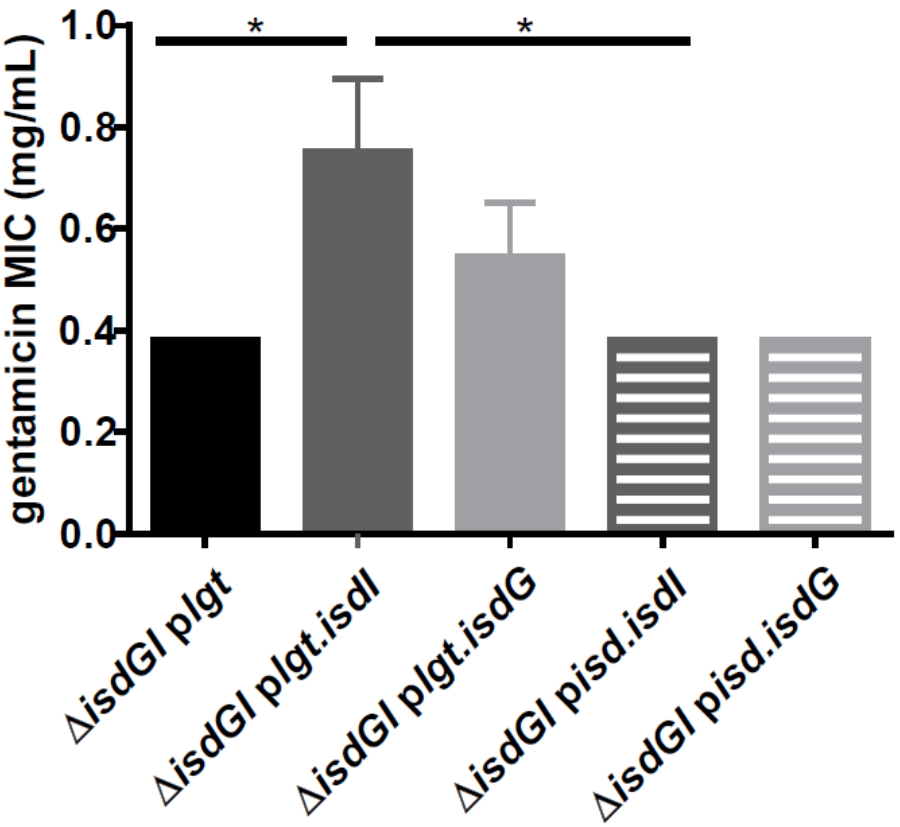


Fig 29. Unregulated *isdI* expression leads to an increase in resistance to gentamicin. Strains were plated onto TSA 10 g/mL chloramphenicol plates, then a gentamicin E-test strip was added. The minimum inhibitory concentration (MIC) were then determined the next day. The gentamicin MIC was significantly greater for Δ*isdGI* p*lgt.isdI* (dark grey, solid) than either Δ*isdGI* p*lgt* (black) or Δ*isdGI* p*isd.isdI* (dark grey, dashed) (Student's T-test, $p<0.05$).

If the strains which are constitutively expressing heme oxygenases are less able to respire to produce energy, they would have reduced growth using a carbon source which is unable to be fermented. When grown in the non-fermentable carbon source glycerol, there is a significant decrease in growth of Δ*isdGI* p*lgt.isdI* and Δ*isdGI* p*lgt.isdG* strains after 6 hours (Fig. 30, 31). However, when these same strains were grown with glucose as the primary carbon source, there is no decrease in growth (Fig. 30, 31). This indicates that the constitutively expressed heme oxygenase strains grow less well in glycerol because they are not able to ferment. Taken together, these data are consistent with a model where unregulated heme oxygenase activity decreases respiration and increases reliance on fermentation for energy production in *S. aureus.*

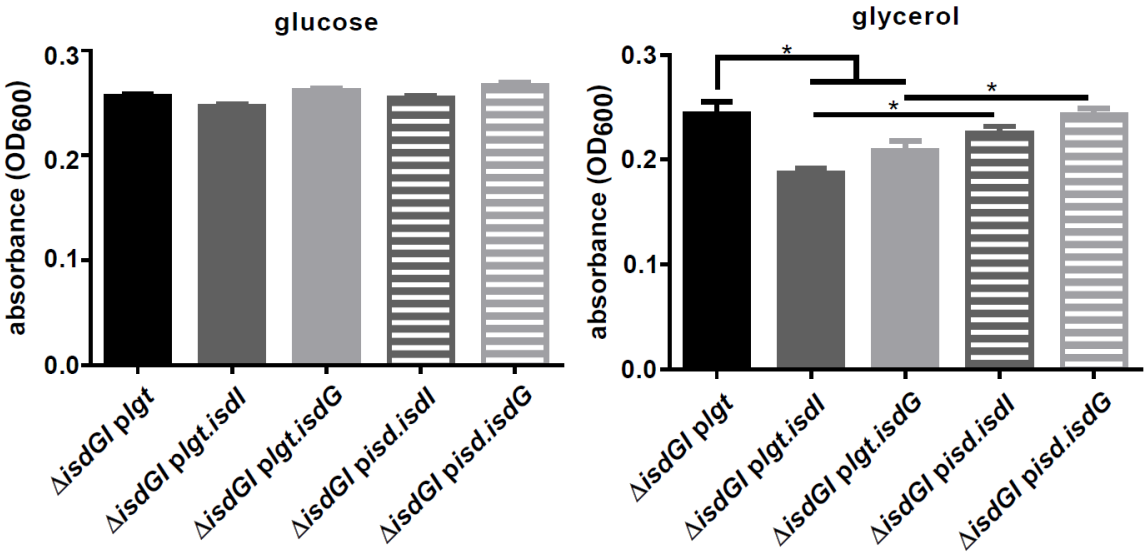


Fig 30. Strains containing constitutively expressed heme oxygenases have decreased growth on a non-fermentable carbon source. *S. aureus* strains were grown in RPMI without glucose containing 22 mM glycerol as the carbon source. Both the Δ*isdGI* p*lgt.isdI* (dark grey, solid) and Δ*isdGI* p*lgt.isdG* (light grey, solid) had significantly less growth at 6 hours than Δ*isdGI* p*lgt* (black) or Δ*isdGI* p*isd.isdI* (dark grey, dashed) and Δ*isdGI* p*isd.isdG* (light grey, dashed), respectively (1way ANOVA, $p<0.05$).

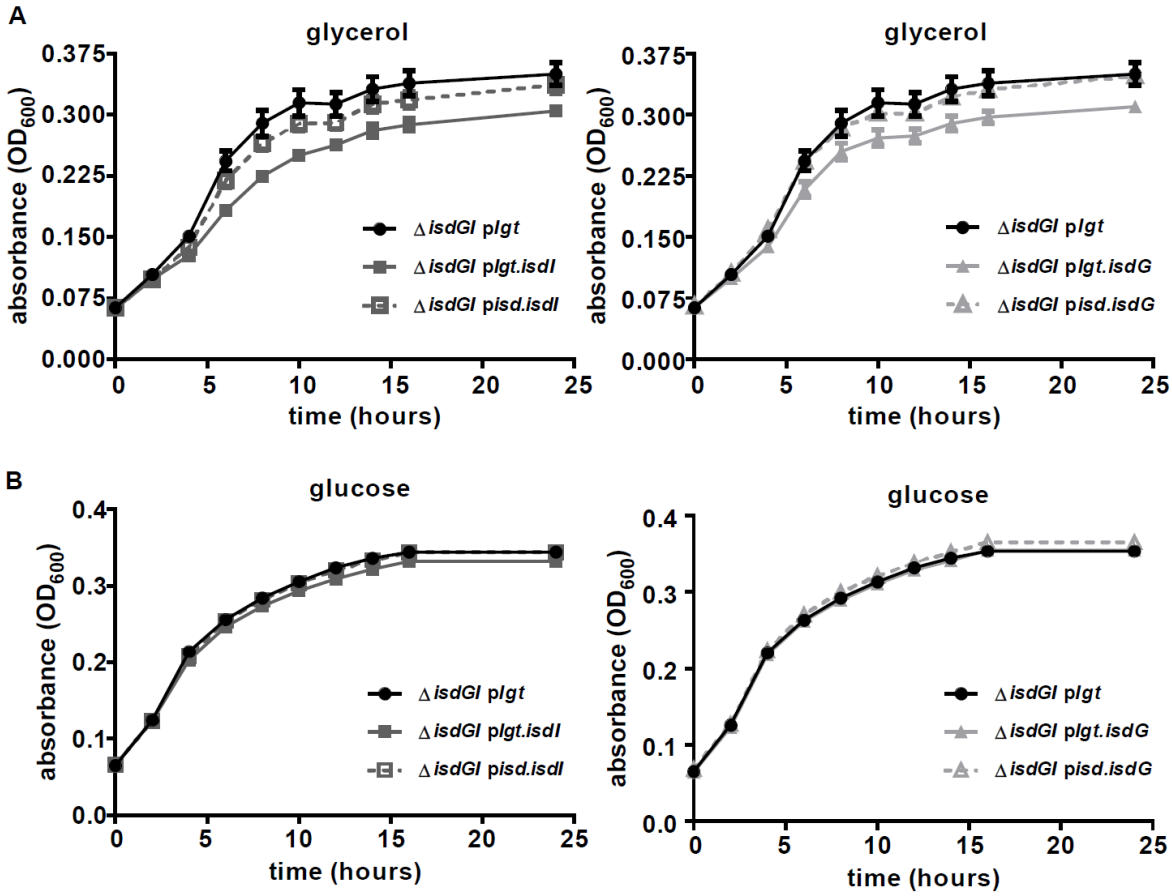


Fig 31. Alternative carbon source growth curves. Each of the strains were grown in RPMI with no glucose. Then either A, 22 mM glycerol or B, 11 mM glucose. There is no change in growth of the strains using glucose as the primary carbon source. In the presence of glycerol as the primary carbon source each of the constitutively expressed heme oxygenases (Δ*isdGI* p*Igt.isdI* and Δ*isdGI* p*Igt.isdG*) exhibit reduced growth as compared to the endogenously expressed strains (Δ*isdGI* p*isd.isdI* and Δ*isdGI* p*isd.isdG*).

Discussion

Heme is a cofactor for a variety of cellular processes, including cytochromes where it facilitates respiration by acting as an electron acceptor (Hammer et al., 2013; Hurt & Hauska, 1981; Morrison & Stotz, 1955). If *S. aureus* is unable to make or acquire heme, then the cell must depend on fermentation for energy production, which generates less ATP than respiration (Jurtchuk, 1996). The production of less energy leads to significantly less growth and to changes in cellular physiology.

In addition to the importance of heme as an enzymatic cofactor, heme also acts as an important iron source. In order to access the iron within heme, *S. aureus* encodes heme degrading enzymes known as heme oxygenases (Skaar, Gaspar, et al., 2004; Wu et al., 2005). The two heme oxygenases encoded by *S. aureus*, IsdG and IsdI, are essential for growth under iron deplete conditions where heme is the sole source of iron (Fig. 23). I initially sought to investigate the role of heme degradation products in *S. aureus*. However, upon performing RNA-Sequencing analysis, I identified that expressing *isdI* under the constitutively expressed *lgt* promoter led to an increase in expression of genes linked to oxygen-independent energy production (Fig. 24). Since the cultures for these experiments had been grown aerobically, I hypothesized that the overproduction of *isdI* led to excess heme degradation, resulting in an insufficient amount of heme present to populate the cytochromes required for efficient aerobic respiration. To test this hypothesis, I developed strains of *S. aureus* expressing each heme oxygenase either under the endogenous promoters (Δ*isdGI* p*isd.*) or the constitutively expressed *lgt* promoter (Δ*isdGI* p*lgt.*). When these strains were grown under iron limitation with heme as the sole source of iron, the endogenously expressed heme oxygenases grew better than the constitutively expressed heme oxygenases (Fig.

25). I next investigated the effect of the heme oxygenase transcriptional regulation on cellular heme abundance and found that the strains constitutively expressing *isdI* and *isdG* contained less heme (Fig. 27). As a result of this decrease in cellular heme, the strains constitutively expressing the heme oxygenases produce higher levels of the fermentation end-product lactate (Fig. 28). Additionally, these strains are also more resistant to gentamicin, which requires active electron transport for entry into the cell (Fig. 29). Finally, the constitutively expressing heme oxygenase strains had a slight growth defect compared to a strain lacking heme oxygenase activity when grown on the non-fermentable carbon source glycerol (Fig. 30, 32).

S. aureus strains that are unable to respire, due to lack of either heme or menadione, form a small colony variant (SCV) phenotype (Bates et al., 2003; von Eiff et al., 2006). These SCVs have increased resistance to aminoglycoside antibiotics that contributes to their ability to create persisting infections (Kahl et al., 1998; Musher, Baughn, & Merrell, 1979). While constitutively expressing heme oxygenases in strains of *S. aureus* do not have the same colony morphology as SCVs, the dysregulation of the heme oxygenases leads to a significant decrease in cellular heme. The reduced heme levels exhibited by the *S. aureus* strains constitutively expressing heme oxygenases lower overall fitness. Therefore, connecting heme oxygenase expression to iron availability through Fur may ensure optimal heme levels for bacterial fitness.

This work indicates that Fur regulation of the heme oxygenases is necessary to ensure that there is enough cellular heme to allow for respiration and optimal fitness under both iron limiting and iron replete conditions. Because so many organisms encode heme oxygenases, this may be a widely conserved paradigm that organisms must balance iron acquisition and heme homeostasis, or control oxygenase expression in the context of iron limitation or heme uptake.

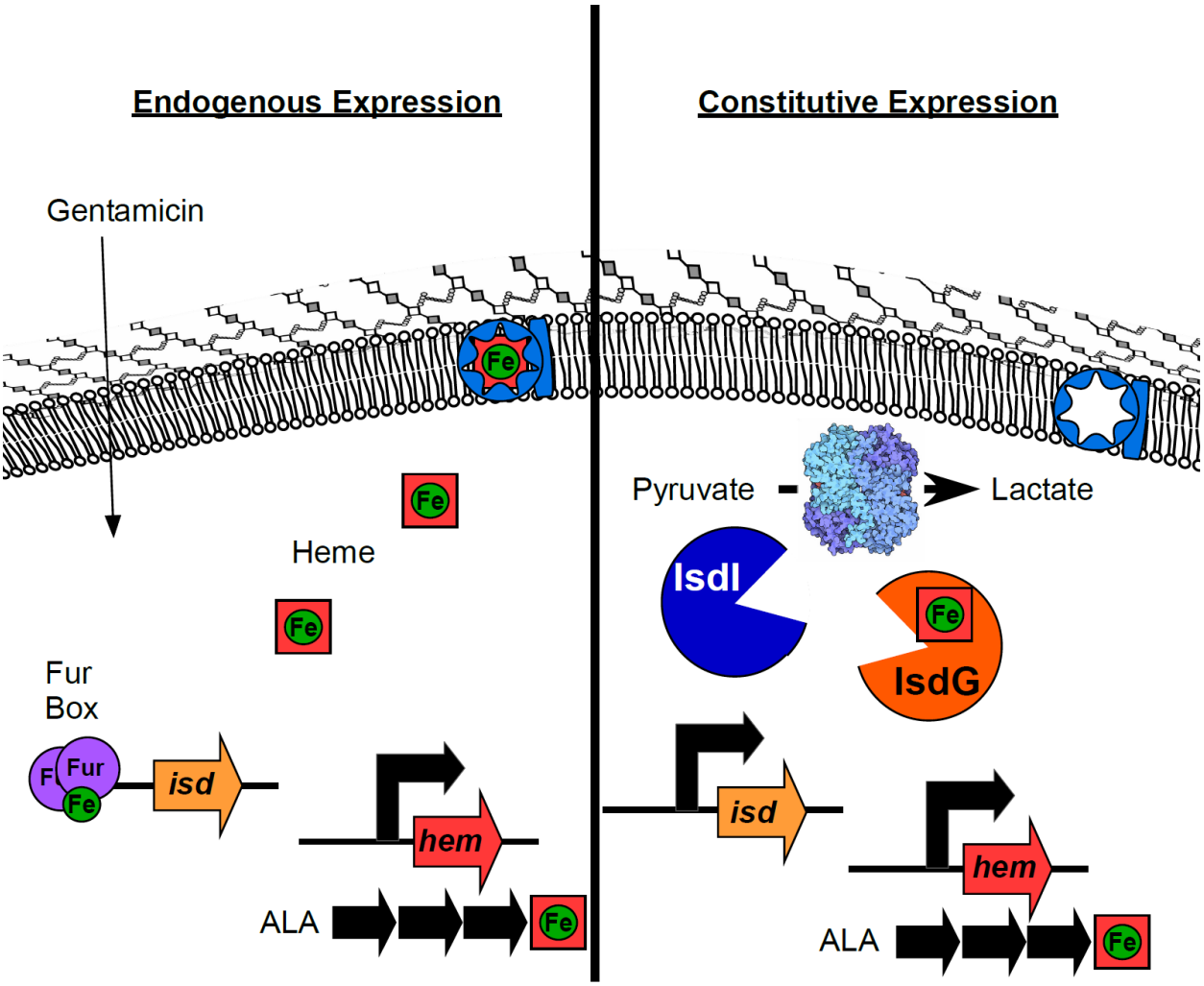


Fig 32. Constitutive expression of heme oxygenases in *S. aureus* leads to a decrease in cellular heme and an increase in fermentation. Unregulated expression of the heme oxygenases in *S. aureus* has a variety of physiological effects on the cell. Primarily, inappropriate expression of the heme oxygenase enzymes leads to a significant decrease in cellular heme levels. Additionally, there is a build up of lactate in the supernatant of cell containing the constitutively expressed heme oxygenases compared to cells lacking heme oxygenase activity. Finally, constitutive expression of *isdI* leads to an increased gentamicin MIC when compared to a strain expressing the endogenously expressed heme oxygenase.

CHAPTER IV

SUMMARY AND SIGNIFICANCE

IsdG family heme oxygenases are widely conserved.

A previous phylogenetic study of the IsdG family of heme oxygenases identified 22 putative members that were all contained within bacteria (Haley et al., 2011). The bacterial phyla that the organisms containing these proteins fell into were Firmicutes, Actinobacteria, and Proteobacteria. Seven of the predicted IsdG family members have been functionally characterized as belonging to the IsdG family (Duong et al., 2014; Haley et al., 2011; Matsui et al., 2016; Park et al., 2012; Skaar, Gaspar, et al., 2004; Skaar et al., 2006). However, this phylogenetic analysis used the entire length of the IsdG protein as a seed to probe for additional IsdG family members (Haley et al., 2011). Since the IsdG family of heme oxygenases have a wide range of sequence identity (20-60%), and the majority of conserved residues are concentrated in the antibiotic biosynthesis monooxygenase (ABM) domain, I hypothesized that the previous bioinformatic analysis did not fully encapsulate all potential IsdG family heme oxygenases. This was reinforced by my identification of LFO1 as the first eukaryotic IsdG family heme oxygenase, which had not been included in the previous phylogenetic analysis.

In order to test whether the previous phylogenetic analysis was incomplete a new phylogeny was created. However instead of using the entirety of the protein to query the known protein database, I focused on the ABM domain of LFO1. This focus on the ABM domain was due to the ABM domain being conserved in all functionally characterized IsdG family heme oxygenases and this domain contains the highest percentage of conserved amino acids and the

catalytic triad. Through this strategy, I created a new phylogenetic tree that broadened the IsdG family of heme oxygenases from 22 to 866 proteins.

In addition to the increase in proteins, this phylogeny expands the IsdG family of heme oxygenases to being in organisms from all domains of life. Within the bacterial domain, 845 proteins were identified over 12 phyla, including both pathogenic and environmental bacteria. Eight proteins were identified within eight archaeal species across two phyla. Finally, 13 proteins were identified, each within separate organisms, across five phyla. Intriguingly, all of the eukaryotes are either green algae or diatoms. Also, the marine green algae and diatoms cluster separately from the freshwater green algae within the Proteobacterial section of the phylogenetic tree.

Unfortunately, since there is fairly low sequence identity between the IsdG family heme oxygenases, I was not able to determine how the IsdG family enzymes were acquired by the eukaryotic and archaeal organisms. However, I hypothesize that the eukaryotic organisms, including *C. reinhardtii*, acquired IsdG family heme oxygenases through horizontal gene transfer. A large portion of the bacteria containing IsdG family enzymes are environmental, including photosynthetic bacteria within the Cyanobacteria, Chlorobi, and Chloroflexi phyla. Organisms within these phyla can live within similar niches as the green algae, meaning that transfer between these organisms is possible. Additionally, a recent publication identified bacteriophages containing HO-1 family heme oxygenases, so bacteriophages may also have picked up IsdG family heme oxygenases and acted as a vector to transfer these heme oxygenases into these eukaryotic organisms (Dammeyer, Bagby, Sullivan, Chisholm, & Frankenberg-Dinkel, 2008).

IsdG family heme oxygenases degrade heme to a variety of products.

The HO-1 family of heme oxygenases were the first identified heme-degrading enzymes. All of the HO-1 family heme oxygenases that have been functionally characterized thus far have been shown to degrade heme to biliverdin, carbon monoxide, and free iron (Wilks, 2002). However, in 2010, IsdG and IsdI from *S. aureus* were shown to degrade heme to the novel catabolite staphylobilin (Reniere et al., 2010). The predicted IsdG family member MhuD degrades heme to a third primary product, mycobilin (Nambu et al., 2013). This has lead us to define the IsdG family of heme oxygenases based on their structure, the presence of the ABM domain, and the presence of the conserved catalytic triad, and not based on the products made during heme degradation.

The identification of LFO1 as the first eukaryotic IsdG family member and the degradation of heme by LFO1 to a new heme catabolite reinforces our classification system for defining IsdG family heme oxygenases. However, the discovery that LFO1 degrades heme to a new catabolite leads to the question as to why the IsdG family enzymes degrade heme to many different products. From a physical perspective, the lack of a high level of sequence identity may allow for a wider array of degradation chemistries. Additionally, from an evolutionary perspective, the production of different catabolites may fill different roles within the cell. Different evolutionary pressures could lead to mutations within the proteins after acquisition to create products that most benefit the cell. This hypothesis is especially intriguing since *C. reinhardtii* has both families of heme oxygenases, potentially two different families within the chloroplast that make two different products.

Regulation of heme oxygenase expression is necessary for cellular heme homeostasis.

Heme is required for the most efficient form of energy production, aerobic respiration (Jurtchuk, 1996). It is of critical importance to a cell to ensure that they are producing as much energy as possible within the niche that they reside so they can continue the process of replication. This underscores the importance of heme within the cell. In order to fill the cell's need for heme, cells can either biosynthesize heme intracellularly or they can import exogenous heme.

S. aureus utilizes both heme biosynthesis and heme acquisition, however these processes are differentially regulated (Dailey et al., 2015; Mazmanian et al., 2003). The mechanisms of regulation for the heme biosynthesis genes is currently unknown, however a previous two-dimensional difference in-gel electrophoresis (2D-DIGE) data set revealed that these proteins are not regulated by Fur (Torres et al., 2010). Furthermore, within the RNA-Seq data contained in Chapter III, which was perfomed on cells with different levels of heme, the heme biosynthesis genes were not differentially transcribed. These data indicate that regulation of heme biosynthesis within *S. aureus* may differ from regulation of homologous genes in other organisms. However, the iron-regulated surface determinant (Isd) system for heme acquisition, which is the only heme acquisition system identified in *S. aureus*, is regulated by Fur (Mazmanian et al., 2003). In addition to the heme acquisition system being regulated by Fur, the heme oxygenases within *S. aureus* are also regulated by Fur (Reniere & Skaar, 2008).

In Chapter III, I showed that dysregulation of the heme oxygenases in *S. aureus* leads to degradation of intracellularly biosynthesized heme. If this were to happen under endogenous conditions, the cells would not grow as well due to a significant decrease in ATP resulting from

the switch from aerobic respiration to fermentative growth. This serious deficit in energy production is likely why the heme oxygenases in *S. aureus*, as well those in other organisms, are only expressed under low iron conditions where heme is also being imported into the cell (Barker, Barkovits, & Wilks, 2012; Farrand et al., 2015; Reniere & Skaar, 2008; Schmitt, 1997). The concerted expression of heme import and heme degradation potentially allows the cell to only degrade the exogenously imported heme and not the heme being endogenously produced.

Summary

The goal of this thesis was to investigate the role of the IsdG family of heme oxygenases within microbial cells. In Chapter II, I identified the first IsdG family member in a eukaryotic organism and expanded the IsdG family from 22 protein members to 866. *C. reinhardtii* is the first organism which contains both HO-1 and IsdG family heme oxygenases. Intriguingly, LFO1 is the only heme oxygenase whose transcription is upregulated under low iron conditions, a condition where the cytochrome associated with the photosynthetic complexes in the chloroplast are degraded and heme is released into the chloroplast cytoplasm. I have hypothesized that LFO1 is important for degradation of this heme and in iron recycling under these iron limited conditions. Additionally, during the process of trying to determine the role of the IsdG family heme catabolites within *S. aureus*, in Chapter III I elucidated the importance of Fur regulation of the heme oxygenases. Removing the inhibition of transcription by Fur under iron replete conditions, the heme oxygenases led to degradation of endogenously biosynthesized heme and an increased reliance on fermentation for energy production. In total, these studies have displayed the

importance of the IsdG family of heme oxygenases for iron recycling and heme homeostasis within organisms from different domains of life.

CHAPTER V

FUTURE DIRECTIONS

Determine localization and protein expression of LFO1 in *C. reinhardtii*

LFO1 contains a putative N-terminal chloroplast localization sequence, however I have not yet determined whether LFO1 is localized to or active within the chloroplast. Additionally, LFO1 is highly unstable when purified *in vitro*, and the stability of LFO1 *in vivo* is currently unknown. LFO1 also contains a predicted metal binding site within the C-terminal unstructured sequence, which may affect both the stability and chemistry of the enzyme. However, the ability of the site to bind a metal and the identity of the metal that may bind are currently unknown.

Sabeeha Merchant's group has been actively working on both tagging the LFO1 protein with a hexa-Histidine tag and knocking out *lfo1* in *C. reinhardtii*. Using the tagged version of LFO1, I would be able to definitively determine the cellular localization of LFO1 in order to confirm the N-terminal chloroplast localization sequence. Tagged LFO1 could then be purified directly from *C. reinhardtii*. This would enable the determination of the stability of the protein *in vivo* after pulse chase with ^{35}S-methionine. Additionally a purified tagged protein would allow for determination of whether LFO1 binds a metal ion, and identification of the ion, by ICP-MS. Finally, the creation of the *C. reinhardtii* cells lacking *lfo1* would allow experiments to determine the importance of LFO1 in iron recycling during iron limitation conditions by monitoring the growth of the mutant and wildtype *C. reinhardtii* under a variety of iron concentrations.

Identify the product of LFO1 heme catabolism

An extensive amount of work during my thesis was focused on determining the product of LFO1 heme degradation. However, despite exhaustive efforts I have not been able to identify the catabolite produced by LFO1.

There are some additional experiments that could be done to identify the product. First, NMR and crystal structure analysis was not performed on the product after the protein purification and the product purification had been optimized. Since these assays are not based on ionization of the molecules analyzed, performing them with a pure catabolite sample may allow for product identification. Since one of the predicted problems with the LFO1 catabolite is that I have not been able to identify an ionization strategy that has worked for MS, a non-ionization based method for product identification should be re-addressed.

Alternatively, MS approaches can be continued utilizing different ion pairing agents. One of the problems with using TFA as an ion pairing agent is that it can be ion suppressive, which may contribute to the lack of ionization that I see for the LFO1 catabolite (Jessome & Volmer, 2006). Unfortunately, TFA is the only ion pairing agent that I have identified that allows for retention of the LFO1 product on the C18 column. Though I also utilized FA, which has less ion suppression associated with it than TFA, the LFO1 product was not retained on the C18 column. There are additional ion pairing agents which could be utilized, such as heptafluorobutyric acid (HFBA). Finally, while MALDI-MS was attempted, a more extensive array of matrices, including 2-hydroxy-1-naphthoic acid which has been successfully used for identification of porphyrins in

the past, could be used to identify the LFO1 heme degradation product (Bartlett, Busch, Wells, & Schey, 1996).

Determine if the LFO1 catabolite has a defined function in *C. reinhardtii*

C. reinhardtii utilizes the products of HO-1 family heme degradation, biliverdin and carbon monoxide, for specific roles within the cell (Duanmu et al., 2012; Liping et al., 2013). Biliverdin has been shown to be important in regulating transcription of specific genes involved in high light stress response (Duanmu et al., 2012). This is intriguing since *C. reinhardtii* does not contain phycobiliproteins, which usually bind to bilins to mediate transcriptional regulation. Exogenous treatment with carbon monoxide leads to a decrease in chlorosis through its regulation of iron acquisition genes (Liping et al., 2013). These activities allow *C. reinhardtii* to respond to cellular stresses.

Because the heme catabolites of HO-1 family heme oxygenases are so impactful in the physiology of *C. reinhardtii*, I would predict that the product of LFO1 heme degradation may also have an important, but potentially different, role within the cell. To begin to elucidate what this potential role may be, RNA-Seq analysis should be performed comparing the transcriptomes of strains of *C. reinhardtii* both containing and lacking both *hmox1* and *lfo1*. This would allow for direct comparison of the roles of both biliverdin and the LFO1 product on transcriptional regulation in *C. reinhardtii*. It would also be interesting to know if the LFO1 product could be targeted by the biliverdin reductase, PCYA (Duanmu et al., 2012). This would indicate that a modified form of the LFO1 product is what is active in the cell. Additionally, identifying which compartment contains the LFO1 product would be informative. The mechanism of how either

biliverdin or its reduced form, phycocyanobilin, facilitates transcriptional changes is currently unknown. It is possible that either product is imported into the nucleus, however how it would be targeted there from the chloroplast is also unknown. Localizing the LFO1 product, either by LC-MS or through using an alkyne labelled heme which can be purified using click chemistry followed by a pull down, to a distinct compartment would give some indication as to where the product is most active within the cell (Nierth & Marletta, 2014). However, if the LFO1 product is seen in multiple compartments this would indicate that the LFO1 product is being targeted for a specific activity.

Elucidate the chemistry of the LFO1 heme degradation reaction

The product of LFO1 heme degradation is a unique product, inconsistent with previously identified heme degradation catabolites. The IsdG family of heme oxygenases, of which LFO1 is a member, contains low sequence identity between the various IsdG family members (Haley et al., 2011). This low level of sequence identity may lead to unique heme degradation reactions within various family members. This has already been seen with MhuD from *M. tuberculosis* (Nambu et al., 2013).

The degradation intermediates for heme degradation reactions of IsdG from *S. aureus* and MhuD from *M. tuberculosis* have been previously identified (Matsui et al., 2016; Streit et al., 2016). Briefly, both enzymes utilize oxygenation of the central iron atom to initiate cleavage of the pyrrole ring. However, while IsdG has shown to be a true monooxygenase, consuming one molecule of oxygen per reaction cycle, MhuD has been shown to have both monooxygenation and dioxygenation capabilities within the same catalytic site, making it the first enzyme identified with

such a capacity (Matsui et al., 2016). Additionally, the cleavage sites of both enzymes are different. IsdG cleaves heme at either the β- or δ-carbon, leading to β- and δ-staphylobilin (Reniere et al., 2010). MhuD cleaves heme at the α-carbon, however, MhuD still results in two catabolite products, mycobilin-*a* or –*b* (Nambu et al., 2013). The proposed reason for this difference in cleavage sites is due to the degree and timing of heme ruffling that occurs within the active site of each enzyme, which allows for different intermediate products after the formation of the hydroxyheme intermediate (Graves et al., 2014; Lee et al., 2008).

MhuD has 21% sequence identity to IsdG, whereas LFO1 has 16% identity to IsdG. The low sequence identity of these proteins likely allows for unique conformations in the protein structure. These changes in the structure of the protein support different orientations of heme within the binding pocket, thus resuling in unique chemistries and the catabolism of heme to distinct products. Since LFO1 degrades heme to a product different than those previously identified, I would hypothesize that the chemistry performed by LFO1 is likely also unique. In order to determine how LFO1 degrades heme, the product of heme degradation will first need to be solved. After the structure of the LFO1 heme degradation product is determined, a variety of methods could be used to determine how LFO1 degrades heme to its catabolite. Jen DuBois perfected a method of monitoring heme degradation by IsdG using *in proteo* mass spectrometry (Streit et al., 2016). This methodology is appealing since the mass of each intermediate can be determined during the course of the degradation reaction. However, since LFO1 is a very unstable protein with a degradation product that does not ionize, this approach may not be feasible. First, the number of oxygen atoms incorporated into the LFO1 product should be determined using $^{18}O_2$ labelled oxygen. Both mass spectrometry and Raman spectroscopy can be used to determine the

number of oxygen atoms incorporated in the LFO1 product. After the oxygenation is confirmed, the production of hydroxyheme, the common heme degradation intermediate should be determined by LFO1-dependent heme degradation in the presence on H_2O_2. If LFO1 produces hydroxyheme as an initial intermediate, LFO1 can then be reconstituted with hydroxyheme and either O_2 or H_2O_2 can be titrated into the reaction in the presence of deferozamine. Intermediates can then be purified from the protein and identified by mass spectrometry.

Identify the role of staphylobilin and formaldehyde in *S. aureus*

Previous work has shown that heme is a more important source of iron than iron acquired directly (Skaar, Humayun, et al., 2004). I hypothesized the reason that heme is more beneficial may be due to the activity of the heme degradation products in *S. aureus*. This hypothesis led us to try to identify the impact that these heme catabolites have on *S. aureus* by RNA-Seq analysis. However, because I utilized a constitutively expressed form of *isdI* in order to ensure high levels of the degradation products, this led to unregulated heme degradation and significant changes in the physiology of the cell. The physiological changes from a respiration competent to a fermentative energy production state are vast, likely obscuring the impact of the heme catabolites on the transcriptome.

However, the question still stands as to what the importance of these products are. Since modifying the regulation of the heme oxygenases had unintended side effects on the cell, I would propose repeating the RNA-Seq experiment with *isdI* under its native promoter. This would allow for the identification of any transcripts that were differentially abundant in the presence of the products. However, I would then need to determine what the role of each product is within the

context of the cell. Since the physiological role of iron and the Fur regulon have been well defined within *S. aureus*, any transcripts conserved between the transcripts identified in the presence of the heme degradation products and the Fur regulon are likely due to an increase in free iron in the presence of the heme oxygenase, leading to Fur repression. Any transcripts identified that are not known to be regulated by iron are likely related either to the presence of heme within the cell or the heme catabolites.

To confirm the RNA-Seq results, a combination of iron, staphylobilin, and/or formaldehyde can be added in different concentrations and growth can be monitored to see if the addition of the catabolites complement a strain lacking heme oxygenase expression when grown under iron limiting conditions. While heme is imported into *S. aureus* under iron limitation, it is unknown whether staphylobilin would be readily imported into the cell in the same fashion. An HPLC based methodology, similar to that developed for the identification of the heme biosynthesis products in Surdel *et. al*, would be important to ensure that staphylobilin is entering the cell (Surdel et al., 2017). Formaldehyde is readily oxidized to formic acid in the presence of oxygen. In addition, formaldehyde is a good electrophile meaning that it can readily react with amines, such as those found in proteins, which would restrict how much formaldehyde is entering the cytoplasm. Using a formaldehyde inducible promoter system driving GFP, one could monitor the presence of formaldehyde in the cytoplasm, to ensure that formaldehyde is getting into the cell and is accessible (Rohlhill, Sandoval, & Papoutsakis, 2017). However, if the exogenous addition of the heme catabolites does not complement the growth defect in the heme oxygenase lacking strain, then qRT-PCR could be used to monitor any changes in transcription.

Investigate the effect of unregulated heme oxygenase production on *S. aureus* pathogenesis

Chapter 3 of this thesis discussed the physiological changes that occur to *S. aureus* when the heme oxygenases are dysregulated. However, I have not yet addressed whether these physiological changes affect the bacteria's ability to infect and thrive in a host. Many of the physiological changes seen with increased heme oxygenase expression are similar to those seen with small colony variants (SCVs) (von Eiff et al., 1997). In the context of infection, *S. aureus* strains that are unable to biosynthesize heme and are SCVs are attenuated and have significantly lower colony forming units (CFUs) in the hearts and the livers of systemically infected mice, when compared to mice infected with wildtype *S. aureus* (Hammer et al., 2013). However, since the *S. aureus* strains expressing the heme oxygenases *in trans* do have low levels of heme under iron replete conditions, I would hypothesize a more modest defect in colonization, especially in organs that are more iron replete.

In order to begin to probe the effect of having reduced heme levels within the cell on *S. aureus* pathogenicity, creating a strain of *S. aureus* which has the heme oxygenases under either endogenous or constitutive promoters incorporated into the chromosome or expressed on a stable plasmid without an antibiotic selection required would need to be done. Currently, the differentially expressed heme oxygenases are encoded within a high copy number plasmid (pOS1), which contains a chloramphenicol resistance marker for selection. However, if these strains were utilized in an infection model, the plasmids have the potential to be lost without chloramphenicol selection, and the plasmid could potentially also be lost in the mice being treated with chloramphenicol since complete penetrance is uncertain. To ensure that the plasmids are not lost

under conditions where they might be selected against, encoding the plasmids within the chromosome with either promoter would ensure homogeneity of the bacterial population during the infection. After the creation of these strains, growth experiments would neet to be performed in the low iron conditions to see if the same growth phenotypes are retained if the heme oxygenases are not expressed off the high copy number plasmid. Finally, a septic model of infection should be used comparing the heme oxygenase lacking strain, strains expressing the heme oxygenases endogenously and constitutively, and a strain lacking the ability to biosynthesize heme. This would allow for a head-to-head comparison of not only strains containing differentially expressed heme oxygenases, but also of cells with varying intracellular heme levels to determine the role of heme at different infection sites.

Determine the most important heme utilizing enzymes within the cell

One of the interesting questions raised during the work described in Chapter 3 was whether there is a specific order to which heme binding proteins are populated by heme within the cell. Despite how much is known about the regulation of heme biosynthesis in other organisms, regulation of heme biosynthesis in *S. aureus* remains unknown. Because of this, the creation of the constitutively expressed heme oxygenase containing strains are the first strains that toggle the amount in the heme in the staphylococcal cell. Using these strains, one could interrogate which heme binding proteins contain heme under conditions where there are different amounts of heme within the cell. This would allow for the determination of which heme binding proteins are most essential under various conditions, since they would be hypothesized to be prioritized for heme

insertion. This would expand our knowledge of heme homeostasis within the cell and how *S. aureus* prioritizes the use of a valuable cellular resource.

APPENDIX A

SUPPLEMENTARY TABLE ASSOCIATE WITH CHAPTER II

Sequences identified to be homologous to characterized IsdG family proteins.

sequence id	domain	lineage	taxon	taxonomy id	catalytic triad	functionally characterized	notes
MHUD_MYCTU	Bacteria	Actinobacteria	Mycobacterium_tuberculosis	83332	3(NWH)	MhuD	
X8F931_MYCUL	Bacteria	Actinobacteria	Mycobacterium_ulcerans	1809	0(---)		
A0A0K9EV67_9ACTO	Bacteria	Actinobacteria	Actinomycetales	2037	0(CYP)		
D2NSQ3_ROTMD	Bacteria	Actinobacteria	Rothia_mucilaginosa	680646	1(HYH)		
B5HBJ5_STRPR	Bacteria	Actinobacteria	Streptomyces_pristinaespiralis	38300	1(N--)		
A0A0M9ZLZ8_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	2(-WH)		
A0A0M9YDN3_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	2(-WH)		
A0A0M8Y9Z8_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	2(-WH)		
G0HEU7_CORVD	Bacteria	Actinobacteria	Corynebacterium_variabile	858619	2(-WH)		
J9SBJ3_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	2(-WH)		

C7MAB6_ BRAFD	Bacteria	Actinobacteria	Brachybacterium_ faecium	446465	2(NWP)		
A0A0D1L2H3 – 9MYCO	Bacteria	Actinobacteria	Mycobacteriaceae	1762	2(NWP)		
U7MIJ1_ 9CORY	Bacteria	Actinobacteria	Corynebacteriaceae	1653	2(SWH)		
G7GLA8_ 9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	2(SWH)		
U7L1T3_ 9CORY	Bacteria	Actinobacteria	Corynebacteriaceae	1653	2(TWH)		
I6WRW4_ PSEPQ	Bacteria	Actinobacteria	Pseudopropionibacterium_ propionicum	767029	2(TWH)		
A0A0M4GQF6 – 9ACTO	Bacteria	Actinobacteria	Actinomycetales	2037	2(TWH)		
U1RK52_ 9ACTO	Bacteria	Actinobacteria	Actinomycetales	2037	2(TWH)		
E7NDT5_ 9ACTO	Bacteria	Actinobacteria	Actinomycetales	2037	2(TWH)		
F9PJ40_ 9ACTO	Bacteria	Actinobacteria	Actinomycetales	2037	2(TWH)		
F9EHX2_ 9ACTO	Bacteria	Actinobacteria	Actinomycetales	2037	2(TWH)		
J0N3K2_ 9ACTO	Bacteria	Actinobacteria	Actinomycetales	2037	2(TWH)		
F5XN95_ MICPN	Bacteria	Actinobacteria	Microlunatus_ phosphovorus	1032480	2(TWH)		
A0A096AGV7 _9MICC	Bacteria	Actinobacteria	Micrococcaceae	1268	2(TWH)		
Q8NMA1_	Bacteria	Actinobacteria	Corynebacterium_	196627	2(TWH)		

CORGL			glutamicum				
A0A0X8JD09_9ACTO	Bacteria	Actinobacteria	Actinomycetales	2037	2(TWH)		
A0A023X207_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
A0A151C0Z2_9MICO	Bacteria	Actinobacteria	Micrococcales	85006	3(NWH)		
A0A0Q7KDM5_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
E2SE28_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
A0A0Q6FJW8_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
R7XTL6_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
C8XE02_NAKMY	Bacteria	Actinobacteria	Nakamurella_multipartita	479431	3(NWH)		
A8KZJ6_FRASN	Bacteria	Actinobacteria	Frankia_sp	298653	3(NWH)		
A0A099J2D7_9MICO	Bacteria	Actinobacteria	Micrococcales	85006	3(NWH)		
N1V167_9MICC	Bacteria	Actinobacteria	Micrococcaceae	1268	3(NWH)		
A0A0K1JL49_9MICO	Bacteria	Actinobacteria	Micrococcales	85006	3(NWH)		
A0A0U2WW81_9MICC	Bacteria	Actinobacteria	Micrococcaceae	1268	3(NWH)		
A0A0M2CTM9_9MICC	Bacteria	Actinobacteria	Micrococcaceae	1268	3(NWH)		
A0A010NNJ8_9MICC	Bacteria	Actinobacteria	Micrococcaceae	1268	3(NWH)		
A0A1E8F3N9_9MICC	Bacteria	Actinobacteria	Micrococcaceae	1268	3(NWH)		
A0A0C1RCG9_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		

A0A0C1UUQ3_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
A0A0G3V296_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
D1ABQ2_THECD	Bacteria	Actinobacteria	Thermomonospora_curvata	471852	3(NWH)		
D6Y690_THEBD	Bacteria	Actinobacteria	Thermobispora_bispora	469371	3(NWH)		
A0A0Q5VMD4_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
A0A101NH48_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
H2K7F8_STRHJ	Bacteria	Actinobacteria	Streptomyces_hygroscopicus subsp	1133850	3(NWH)		
S5VJM8_STRC3	Bacteria	Actinobacteria	Streptomyces_collinus	1214242	3(NWH)		
A0A0M8XI27_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
S4N0J3_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
A0A0G3ADB2_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
A0A089XCB7_STRGA	Bacteria	Actinobacteria	Streptomyces_glaucescens	1907	3(NWH)		
A0A0D4DU44_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
A0A0X3XUH2_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
A0A101J6R4_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
K4QXE6_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
A0A101PFB4_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		

A0A176L9W5_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
A0A0N0N865_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
A0A086MSV8_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
A0A177HVH0_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
A0A0B5D7R1_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
A0A0X3SKW4_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
A0A0C5GDZ5_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
A0A0N1GRV6_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
A0A1K2FHG3_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
A0A1J4Q6M5_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
A0A1D7VH98_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
J1ZT77_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
A0A059W6D6_STRA9	Bacteria	Actinobacteria	Streptomyces_albulus	68570	3(NWH)		
A0A0U3NP43_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
D9X4M1_STRVT	Bacteria	Actinobacteria	Streptomyces_viridochromogenes	591159	3(NWH)		
A0A0M8VNK7_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
V6KE10_STRRC	Bacteria	Actinobacteria	Streptomyces_roseochromogenus subsp	149682	3(NWH)		
A0A124I2D5_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		

A0A0M3QL09_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
B1VNT2_STRGG	Bacteria	Actinobacteria	Streptomyces_ griseus subsp	455632	3(NWH)		
A0A1J4P5W0_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
A0A022MCA4_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
G2GFM1_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
A0A171BH93_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
A0A101T4I1_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
A0A0Q9AQ68_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
D6K8V7_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
A0A0L0JWV0_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
A0A0F4K7B4_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
A0A0F4JK55_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
A0A0M8TCY0_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
D9VU91_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
A0A1E5Q1S4_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
A0A0C1SA86_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
A0A0F4J8X8_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		

A0A0F0H9E0_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
A0A0F2T8A5_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
D9WW83_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
G2P6W6_STRVO	Bacteria	Actinobacteria	Streptomyces_violaceusniger	68280	3(NWH)		
D7CAY2_STRBB	Bacteria	Actinobacteria	Streptomyces_bingchenggensis	749414	3(NWH)		
A0A0K9XMS6_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
A0A0W7XAR1_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
A0A0D0PHL4_KITGR	Bacteria	Actinobacteria	Kitasatospora_griseola	2064	3(NWH)		
E4NAL3_KITSK	Bacteria	Actinobacteria	Kitasatospora_setae	452652	3(NWH)		
A0A1E7JKE4_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
C9ZFN2_STRSW	Bacteria	Actinobacteria	Streptomyces_scabiei	680198	3(NWH)		
A0A0X3W474_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
A0A0A8ESE7_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
A0A0X3UXH7_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
A0A1B6ATY7_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
L8ELN5_STRRM	Bacteria	Actinobacteria	Streptomyces_rimosus	1927	3(NWH)		
A0A0M8YKA7_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
F2RKU8_	Bacteria	Actinobacteria	Streptomyces_venezuelae	953739	3(NWH)		

STRVP							
A0A1E5PC13_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
S2XVA0_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
L1KME9_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
S3ZP62_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
A0A081XLA8_STRTO	Bacteria	Actinobacteria	Streptomyces_toyocaensis	55952	3(NWH)		
A0A1D2IA20_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
A0A101V2J2_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
A0A1B1MDL0_STRLN	Bacteria	Actinobacteria	Streptomyces_lincolnensis	1915	3(NWH)		
D5ZPK1_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
A0A1A5P3K7_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
A0A0M8QDV9_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
A0A072SDE8_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
A0A0M8VPX9_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
L7FIY7_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
A0A101R434_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
V4IJQ8_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
B5HTJ5_	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		

9ACTN							
I2MXP1_ 9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
E2Q217_ STRC2	Bacteria	Actinobacteria	Streptomyces_ clavuligerus	443255	3(NWH)		
A0A066YW54 _9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
A0A0Q8NNM 3_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
A0A0F0HPE0_ 9PSEU	Bacteria	Actinobacteria	Pseudonocardiales	85010	3(NWH)		
A0A0F4JZP8_ 9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
V6L0J0_ 9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
A0A1E7K201_ 9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
A0A0N0YQ55 _9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
A0A0F7FXV5 _9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
A0A0F5VPB0 _9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
A0A0T6LL15_ 9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
A0A132MYA9 _9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
Q47SB0_ THEFY	Bacteria	Actinobacteria	Thermobifida_fusca	269800	3(NWH)		
D7B4D0_NOC DD	Bacteria	Actinobacteria	Nocardiopsis_dassonvillei	446468	3(NWH)		
J7L2N7_ NOCAA	Bacteria	Actinobacteria	Nocardiopsis_alba	1205910	3(NWH)		
D2SDV9_	Bacteria	Actinobacteria	Geodermatophilus_	526225	3(NWH)		

GEOOG			obscurus				
I4F289_ 9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
A0A098Y899_ 9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
H6RVU3_ BLASD	Bacteria	Actinobacteria	Blastococcus_saxobsidens	1146883	3(NWH)		
W5XRS5_9CORY	Bacteria	Actinobacteria	Corynebacteriaceae	1653	3(NWH)		
K9AJZ4_ 9MICO	Bacteria	Actinobacteria	Micrococcales	85006	3(NWH)		
A0A0B8ZWM3_BRELN	Bacteria	Actinobacteria	Brevibacterium_linens	1703	3(NWH)		
A0A0H0ZQ18_9MICC	Bacteria	Actinobacteria	Micrococcaceae	1268	3(NWH)		
E1W0L8_ GLUAR	Bacteria	Actinobacteria	Glutamicibacter_ arilaitensis	861360	3(NWH)		
M7N521_9MICC	Bacteria	Actinobacteria	Micrococcaceae	1268	3(NWH)		
A0A0M4QYZ1_9MICC	Bacteria	Actinobacteria	Micrococcaceae	1268	3(NWH)		
A0A0Q4V4V7_9MICO	Bacteria	Actinobacteria	Micrococcales	85006	3(NWH)		
A0A0Q5MVP2_9MICO	Bacteria	Actinobacteria	Micrococcales	85006	3(NWH)		
A0A0N0KTP0_9MICO	Bacteria	Actinobacteria	Micrococcales	85006	3(NWH)		
A0A0Q5L8K0_9MICO	Bacteria	Actinobacteria	Micrococcales	85006	3(NWH)		
A0A0Q5CFL3_9MICO	Bacteria	Actinobacteria	Micrococcales	85006	3(NWH)		
X5DT36_ 9CORY	Bacteria	Actinobacteria	Corynebacteriaceae	1653	3(NWH)		
C5CAU5_	Bacteria	Actinobacteria	Micrococcus_luteus	465515	3(NWH)		

MICLC							
B2GLK5_ KOCRD	Bacteria	Actinobacteria	Kocuria_rhizophila	378753	3(NWH)		
A0A0P0EEI4_ 9MICO	Bacteria	Actinobacteria	Micrococcales	85006	3(NWH)		
A0A061LX03_ 9MICO	Bacteria	Actinobacteria	Micrococcales	85006	3(NWH)		
A4F6V2_ SACEN	Bacteria	Actinobacteria	Saccharopolyspora_ erythraea	405948	3(NWH)		
A0A0A6UJA6 _ACTUT	Bacteria	Actinobacteria	Actinoplanes_utahensis	1869	3(NWH)		
I0GZU2_ ACTM4	Bacteria	Actinobacteria	Actinoplanes_ missouriensis	512565	3(NWH)		
A0A124G9Z0_ 9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
G8S1R2_ ACTS5	Bacteria	Actinobacteria	Actinoplanes_sp	134676	3(NWH)		
A0A073AWL1 _9PSEU	Bacteria	Actinobacteria	Pseudonocardiales	85010	3(NWH)		
A0A1D8BQV5 _9PSEU	Bacteria	Actinobacteria	Pseudonocardiales	85010	3(NWH)		
U5VVI0_ L9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
A0A0C1KYM 8_9PSEU	Bacteria	Actinobacteria	Pseudonocardiales	85010	3(NWH)		
H5X206_ 9PSEU	Bacteria	Actinobacteria	Pseudonocardiales	85010	3(NWH)		
R4TIK2_ AMYOR	Bacteria	Actinobacteria	Amycolatopsis_orientalis	31958	3(NWH)		
D9V4Y4_ 9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
I1D750_	Bacteria	Actinobacteria	Pseudonocardiales	85010	3(NWH)		

9PSEU							
C7MQ74_ SACVD	Bacteria	Actinobacteria	Saccharomonospora_ viridis	471857	3(NWH)		
A0A076N9D7_ AMYME	Bacteria	Actinobacteria	Amycolatopsis_ methanolica	1814	3(NWH)		
G0G6K5_ AMYMS	Bacteria	Actinobacteria	Amycolatopsis_ mediterranei	713604	3(NWH)		
W7IMM4_9PSEU	Bacteria	Actinobacteria	Pseudonocardiales	85010	3(NWH)		
A0A099D4K4_ 9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
E3J591_FRAIE	Bacteria	Actinobacteria	Frankia_inefficax	298654	3(NWH)		
Q0RQ12_ FRAAA	Bacteria	Actinobacteria	Frankia_alni	326424	3(NWH)		
A0A166S207_ 9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
D3CVB5_ 9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
A0A0P0SJR7_ 9PSEU	Bacteria	Actinobacteria	Pseudonocardiales	85010	3(NWH)		
A0A0M4RCM5_9PSEU	Bacteria	Actinobacteria	Pseudonocardiales	85010	3(NWH)		
D6Z9W1_ SEGRD	Bacteria	Actinobacteria	Segniliparus_rotundus	640132	3(NWH)		
G7GUC8_ 9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
H0R0A9_ 9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
D0L447_ GORB4	Bacteria	Actinobacteria	Gordonia_bronchialis	526226	3(NWH)		
H5TYZ1_	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		

9ACTN							
V8CUM0_ 9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
H6MUR3_ GORPV	Bacteria	Actinobacteria	Gordonia_ polyisoprenivorans	1112204	3(NWH)		
K6WHQ8_ 9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
M0QKW1_ 9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
M3TMN0_ 9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
M7A2Z5_ 9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
A0A0Q5QS10 _9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
U5WSR9_ MYCKA	Bacteria	Actinobacteria	Mycobacterium_kansasii	1768	3(NWH)		
X7XTP3_ MYCKA	Bacteria	Actinobacteria	Mycobacterium_kansasii	1768	3(NWH)		
B2HJ32_ MYCMM	Bacteria	Actinobacteria	Mycobacterium_marinum	216594	3(NWH)		
A0A0I9TCN2_ 9MYCO	Bacteria	Actinobacteria	Mycobacteriaceae	1762	3(NWH)		
A0A1A0MGE0 _9MYCO	Bacteria	Actinobacteria	Mycobacteriaceae	1762	3(NWH)		
A0A1A2Z5J0_ 9MYCO	Bacteria	Actinobacteria	Mycobacteriaceae	1762	3(NWH)		
A0A1E3TJJ4_ MYCSH	Bacteria	Actinobacteria	Mycobacterium_shimoidei	29313	3(NWH)		
A0A1A3PGN1 _9MYCO	Bacteria	Actinobacteria	Mycobacteriaceae	1762	3(NWH)		
I0RMF7_	Bacteria	Actinobacteria	Mycobacterium_xenopi	1789	3(NWH)		

MYCXE							
X8C9K0_ MYCXE	Bacteria	Actinobacteria	Mycobacterium_xenopi	1789	3(NWH)		
Q743X4_ MYCPA	Bacteria	Actinobacteria	Mycobacterium_ paratuberculosis	262316	3(NWH)		
X8CLW5_ MYCIT	Bacteria	Actinobacteria	Mycobacterium_ intracellulare	1767	3(NWH)		
Q9CBI8_ MYCLE	Bacteria	Actinobacteria	Mycobacterium_leprae	272631	3(NWH)		
A0A1A3CA02 _MYCAS	Bacteria	Actinobacteria	Mycobacterium_asiaticum	1790	3(NWH)		
A0A024K6I8_ 9MYCO	Bacteria	Actinobacteria	Mycobacteriaceae	1762	3(NWH)		
A0A1E3SEN8 _9MYCO	Bacteria	Actinobacteria	Mycobacteriaceae	1762	3(NWH)		
D5P8M8_9MY CO	Bacteria	Actinobacteria	Mycobacteriaceae	1762	3(NWH)		
A0A1A2ZUN1 _9MYCO	Bacteria	Actinobacteria	Mycobacteriaceae	1762	3(NWH)		
A0A172UIY4_ 9MYCO	Bacteria	Actinobacteria	Mycobacteriaceae	1762	3(NWH)		
A0A1A3T960_ 9MYCO	Bacteria	Actinobacteria	Mycobacteriaceae	1762	3(NWH)		
A0A132T9S5_ 9MYCO	Bacteria	Actinobacteria	Mycobacteriaceae	1762	3(NWH)		
I4BQ47_ MYCCN	Bacteria	Actinobacteria	Mycobacterium_ chubuense	710421	3(NWH)		
A0A1A1ZB52 _9MYCO	Bacteria	Actinobacteria	Mycobacteriaceae	1762	3(NWH)		
L0J627_ 9MYCO	Bacteria	Actinobacteria	Mycobacteriaceae	1762	3(NWH)		
A1UMG9_ MYCSK	Bacteria	Actinobacteria	Mycobacterium_sp	189918	3(NWH)		

G8RSZ0_ MYCRN	Bacteria	Actinobacteria	Mycobacterium_rhodesiae	710685	3(NWH)		
A0A1E3RSU3_9MYCO	Bacteria	Actinobacteria	Mycobacteriaceae	1762	3(NWH)		
A0A0M2ZEE0_9MYCO	Bacteria	Actinobacteria	Mycobacteriaceae	1762	3(NWH)		
W9BKA3_9MYCO	Bacteria	Actinobacteria	Mycobacteriaceae	1762	3(NWH)		
A0A132PM29_9MYCO	Bacteria	Actinobacteria	Mycobacteriaceae	1762	3(NWH)		
A0R569_ MYCS2	Bacteria	Actinobacteria	Mycobacterium_ smegmatis	246196	3(NWH)		
A0A100VXA0_9MYCO	Bacteria	Actinobacteria	Mycobacteriaceae	1762	3(NWH)		
G7CF04_ MYCT3	Bacteria	Actinobacteria	Mycobacterium_ thermoresistibile	1078020	3(NWH)		
A0A139VMN4_MYCPH	Bacteria	Actinobacteria	Mycobacterium_phlei	1771	3(NWH)		
X5LDQ8_9MYCO	Bacteria	Actinobacteria	Mycobacteriaceae	1762	3(NWH)		
A0A0A1FY33_9MYCO	Bacteria	Actinobacteria	Mycobacteriaceae	1762	3(NWH)		
A1TG23_ MYCVP	Bacteria	Actinobacteria	Mycobacterium_ vanbaalenii	350058	3(NWH)		
A0A178M2X9_9MYCO	Bacteria	Actinobacteria	Mycobacteriaceae	1762	3(NWH)		
F5YS85_ MYCSD	Bacteria	Actinobacteria	Mycobacterium_sinense	875328	3(NWH)		
A0A1B1WC23_9MYCO	Bacteria	Actinobacteria	Mycobacteriaceae	1762	3(NWH)		
A0A0D1KY19_9MYCO	Bacteria	Actinobacteria	Mycobacteriaceae	1762	3(NWH)		
A0A0T1WFM3_9MYCO	Bacteria	Actinobacteria	Mycobacteriaceae	1762	3(NWH)		

A0A0Q2R5J7_MYCGO	Bacteria	Actinobacteria	Mycobacterium_gordonae	1778	3(NWH)		
A0A1A1XBE9_9MYCO	Bacteria	Actinobacteria	Mycobacteriaceae	1762	3(NWH)		
K5BIW2_MYCHD	Bacteria	Actinobacteria	Mycobacterium_hassiacum	1122247	3(NWH)		
A0A0J6WCR5_9MYCO	Bacteria	Actinobacteria	Mycobacteriaceae	1762	3(NWH)		
A0A0J6WET7_9MYCO	Bacteria	Actinobacteria	Mycobacteriaceae	1762	3(NWH)		
A0A0G3IUB8_9MYCO	Bacteria	Actinobacteria	Mycobacteriaceae	1762	3(NWH)		
X8E0Y9_9MYCO	Bacteria	Actinobacteria	Mycobacteriaceae	1762	3(NWH)		
E9T1H5_RHOHA	Bacteria	Actinobacteria	Rhodococcus_hoagii	43767	3(NWH)		
A0A177YK21_9NOCA	Bacteria	Actinobacteria	Nocardiaceae	85025	3(NWH)		
A0A143QD21_9NOCA	Bacteria	Actinobacteria	Nocardiaceae	85025	3(NWH)		
A0A164KY00_9NOCA	Bacteria	Actinobacteria	Nocardiaceae	85025	3(NWH)		
A0A0D0LDR0_9NOCA	Bacteria	Actinobacteria	Nocardiaceae	85025	3(NWH)		
A0A161X6Q5_9NOCA	Bacteria	Actinobacteria	Nocardiaceae	85025	3(NWH)		
A0A1A2BIL8_9MYCO	Bacteria	Actinobacteria	Mycobacteriaceae	1762	3(NWH)		
W5T8G0_9NOCA	Bacteria	Actinobacteria	Nocardiaceae	85025	3(NWH)		
A0A0B8N9M9_9NOCA	Bacteria	Actinobacteria	Nocardiaceae	85025	3(NWH)		
K0ERV2_9NOCA	Bacteria	Actinobacteria	Nocardiaceae	85025	3(NWH)		
H6RB15_	Bacteria	Actinobacteria	Nocardia_cyriacigeorgica	1127134	3(NWH)		

NOCCG							
Q0S8B4_RHOJR	Bacteria	Actinobacteria	Rhodococcus_jostii	101510	3(NWH)		
C0ZPP4_RHOE4	Bacteria	Actinobacteria	Rhodococcus_erythropolis	234621	3(NWH)		
A0A059MU50_9NOCA	Bacteria	Actinobacteria	Nocardiaceae	85025	3(NWH)		
Q5Z2S2_NOCFA	Bacteria	Actinobacteria	Nocardia_farcinica	247156	3(NWH)		
V9XGE0_9NOCA	Bacteria	Actinobacteria	Nocardiaceae	85025	3(NWH)		
A0A0K2YHJ0_9NOCA	Bacteria	Actinobacteria	Nocardiaceae	85025	3(NWH)		
R7WML8_9NOCA	Bacteria	Actinobacteria	Nocardiaceae	85025	3(NWH)		
U5EBD1_NOCAS	Bacteria	Actinobacteria	Nocardia_asteroides	1824	3(NWH)		
A0A137ZYV7_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
F6EGX9_HOYSD	Bacteria	Actinobacteria	Hoyosella_subflava	443218	3(NWH)		
G7H1I7_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
M3VAS8_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
L7LGA0_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
A0A0N9N7Q2_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
L7LCV8_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		

A0A0Q7B861_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
A0A1A9GRK8_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
A0A0Q9RZY9_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
A0A0Q7YYH5_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
E6J7G2_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
A0A1E3RTM1_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
A0A1J4N8D6_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
A0A0Q8VHZ5_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
A0A0A1DR65_NOCSI	Bacteria	Actinobacteria	Nocardioides_simplex	2045	3(NWH)		
A0A0Q8PZE0_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
A0A1A3JCL4_9MYCO	Bacteria	Actinobacteria	Mycobacteriaceae	1762	3(NWH)		
A0A1E3SVF4_9MYCO	Bacteria	Actinobacteria	Mycobacteriaceae	1762	3(NWH)		
K0JQ72_SACES	Bacteria	Actinobacteria	Saccharothrix_espanaensis	1179773	3(NWH)		
A0A0M8YK53_9PSEU	Bacteria	Actinobacteria	Pseudonocardiales	85010	3(NWH)		
C6WGJ2_ACTMD	Bacteria	Actinobacteria	Actinosynnema_mirum	446462	3(NWH)		
A0A0F0GM07_NOCAE	Bacteria	Actinobacteria	Lechevalieria_aerocolonigenes	68170	3(NWH)		
A0A0M8W999_9NOCA	Bacteria	Actinobacteria	Nocardiaceae	85025	3(NWH)		
D5UST7_	Bacteria	Actinobacteria	Tsukamurella_	521096	3(NWH)		

TSUPD			paurometabola				
D9T4K5_ MICAI	Bacteria	Actinobacteria	Micromonospora_ aurantiaca	644283	3(NWH)		
F4F197_ VERMA	Bacteria	Actinobacteria	Verrucosispora_maris	263358	3(NWH)		
F4CLT0_ PSEUX	Bacteria	Actinobacteria	Pseudonocardia_ dioxanivorans	675635	3(NWH)		
A0A1E4P0F2_ 9PSEU	Bacteria	Actinobacteria	Pseudonocardiales	85010	3(NWH)		
D9QC27_ CORP2	Bacteria	Actinobacteria	Corynebacterium_ pseudotuberculosis	681645	3(NWH)		
A0A0B6TK15 _9CORY	Bacteria	Actinobacteria	Corynebacteriaceae	1653	3(NWH)		
A0A0G3H625_ 9CORY	Bacteria	Actinobacteria	Corynebacteriaceae	1653	3(NWH)		
M1NWI6_ 9CORY	Bacteria	Actinobacteria	Corynebacteriaceae	1653	3(NWH)		
A0A096AGB5 _9CORY	Bacteria	Actinobacteria	Corynebacteriaceae	1653	3(NWH)		
S5THU3_9CO RY	Bacteria	Actinobacteria	Corynebacteriaceae	1653	3(NWH)		
A0A076NMS4 _9CORY	Bacteria	Actinobacteria	Corynebacteriaceae	1653	3(NWH)		
A0A077HJN0_ 9CORY	Bacteria	Actinobacteria	Corynebacteriaceae	1653	3(NWH)		
A0A0G3HBB7 _9CORY	Bacteria	Actinobacteria	Corynebacteriaceae	1653	3(NWH)		
C0XQ83_ 9CORY	Bacteria	Actinobacteria	Corynebacteriaceae	1653	3(NWH)		
D7WEQ1_ 9CORY	Bacteria	Actinobacteria	Corynebacteriaceae	1653	3(NWH)		
A0A097IEH8_ 9CORY	Bacteria	Actinobacteria	Corynebacteriaceae	1653	3(NWH)		

A0A0K1RER0_9CORY	Bacteria	Actinobacteria	Corynebacteriaceae	1653	3(NWH)		
Q8FLN9_COREF	Bacteria	Actinobacteria	Corynebacterium_efficiens	196164	3(NWH)		
Q8NLF1_CORGL	Bacteria	Actinobacteria	Corynebacterium_glutamicum	196627	3(NWH)		
K1E2A9_9MICO	Bacteria	Actinobacteria	Micrococcales	85006	3(NWH)		
A0A0Q9K7T9_9MICO	Bacteria	Actinobacteria	Micrococcales	85006	3(NWH)		
Q6NFP6_CORDI	Bacteria	Actinobacteria	Corynebacterium_diphtheriae	257309	3(NWH)		
L1MC46_9CORY	Bacteria	Actinobacteria	Corynebacteriaceae	1653	3(NWH)		
C7QYN2_JONDD	Bacteria	Actinobacteria	Jonesia_denitrificans	471856	3(NWH)		
A0A0S9Q485_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
A0A0N9HV33_9PSEU	Bacteria	Actinobacteria	Pseudonocardiales	85010	3(NWH)		
A0A0Q0UFT1_9CORY	Bacteria	Actinobacteria	Corynebacteriaceae	1653	3(NWH)		
A0A172X5D3_9MICO	Bacteria	Actinobacteria	Micrococcales	85006	3(NWH)		
A0A175RF64_9MICO	Bacteria	Actinobacteria	Micrococcales	85006	3(NWH)		
M2YBQ4_9MICC	Bacteria	Actinobacteria	Micrococcaceae	1268	3(NWH)		
D4YL80_9MICO	Bacteria	Actinobacteria	Micrococcales	85006	3(NWH)		
A0A0G3H6E2_9CORY	Bacteria	Actinobacteria	Corynebacteriaceae	1653	3(NWH)		
W5WP97_	Bacteria	Actinobacteria	Corynebacteriaceae	1653	3(NWH)		

9CORY							
A0A1B7LXS4_9MICC	Bacteria	Actinobacteria	Micrococcaceae	1268	3(NWH)		
I7IXM5_9CORY	Bacteria	Actinobacteria	Corynebacteriaceae	1653	3(NWH)		
A0A088QGL6_9CORY	Bacteria	Actinobacteria	Corynebacteriaceae	1653	3(NWH)		
W9GLG3_9MICO	Bacteria	Actinobacteria	Micrococcales	85006	3(NWH)		
W7T2I9_9PSEU	Bacteria	Actinobacteria	Pseudonocardiales	85010	3(NWH)		
W5WCM7_9PSEU	Bacteria	Actinobacteria	Pseudonocardiales	85010	3(NWH)		
M5AGN8_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
H0EBN1_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	3(NWH)		
E0DGK3_9CORY	Bacteria	Actinobacteria	Corynebacteriaceae	1653	3(NWH)		
A0A075TX84_9CORY	Bacteria	Actinobacteria	Corynebacteriaceae	1653	3(NWH)		
E3H3C9_ROTDC	Bacteria	Actinobacteria	Rothia_dentocariosa	762948	3(NWH)		
A4X1G9_SALTO	Bacteria	Actinobacteria	Salinispora_tropica	369723	3(NWH)		
A0A0F0GYW8_NOCAE	Bacteria	Actinobacteria	Lechevalieria_aerocolonigenes	68170	3(NWH)		
A0A0N0A1A8_9NOCA	Bacteria	Actinobacteria	Nocardiaceae	85025	3(NWH)		
Q1AWN0_RUBXD	Bacteria	Actinobacteria	Rubrobacter_xylanophilus	266117	na		failed iqtree; removed from

							phylogenetic analysis
A0A0Q9TGA2_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	na		failed iqtree; removed from phylogenetic analysis
A0A0S9R9W1_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	na		failed iqtree; removed from phylogenetic analysis
A0A0B2B830_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	na		failed iqtree; removed from phylogenetic analysis
A0A0A0C0N4_9CELL	Bacteria	Actinobacteria	Cellulomonadaceae	85016	na		failed iqtree; removed from phylogenetic analysis
A0A021VYC4_9CELL	Bacteria	Actinobacteria	Cellulomonadaceae	85016	na		failed iqtree; removed from phylogenetic analysis
C7NJ91_KYTSD	Bacteria	Actinobacteria	Kytococcus_sedentarius	478801	na		failed iqtree; removed from phylogenetic analysis
R4YWA3_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	na		failed iqtree; removed from phylogenetic analysis
K6X673_9MICO	Bacteria	Actinobacteria	Micrococcales	85006	na		failed iqtree; removed from phylogenetic analysis
H5USM1_9MICO	Bacteria	Actinobacteria	Micrococcales	85006	na		failed iqtree; removed from phylogenetic analysis

A0A127A6J9_9MICC	Bacteria	Actinobacteria	Micrococcaceae	1268	na		failed iqtree; removed from phylogenetic analysis
A0A0B2AQ08_9MICC	Bacteria	Actinobacteria	Micrococcaceae	1268	na		failed iqtree; removed from phylogenetic analysis
D2AVP2_STRRD	Bacteria	Actinobacteria	Streptosporangium_roseum	479432	na		failed iqtree; removed from phylogenetic analysis
A0A161LNP0_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	na		failed iqtree; removed from phylogenetic analysis
W2EJN4_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	na		failed iqtree; removed from phylogenetic analysis
B4V3Z5_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	na		failed iqtree; removed from phylogenetic analysis
A0A0N0N657_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	na		failed iqtree; removed from phylogenetic analysis
A0A0X3S049_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	na		failed iqtree; removed from phylogenetic analysis
Q2J6D1_FRACC	Bacteria	Actinobacteria	Frankia_casuarinae	106370	na		failed iqtree; removed from phylogenetic analysis
Q0RCY6_FRAAA	Bacteria	Actinobacteria	Frankia_alni	326424	na		failed iqtree; removed from

							phylogenetic analysis
A0A166Q9K2_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	na		failed iqtree; removed from phylogenetic analysis
D3D6D5_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	na		failed iqtree; removed from phylogenetic analysis
A8L2H5_FRASN	Bacteria	Actinobacteria	Frankia_sp	298653	na		failed iqtree; removed from phylogenetic analysis
F8AX51_FRADG	Bacteria	Actinobacteria	Frankia_symbiont subsp	656024	na		failed iqtree; removed from phylogenetic analysis
A0A126Z9B6_9MICO	Bacteria	Actinobacteria	Micrococcales	85006	na		failed iqtree; removed from phylogenetic analysis
A0A0H5CT56_9PSEU	Bacteria	Actinobacteria	Pseudonocardiales	85010	na		failed iqtree; removed from phylogenetic analysis
A6W7P6_KINRD	Bacteria	Actinobacteria	Kineococcus_ radiotolerans	266940	na		failed iqtree; removed from phylogenetic analysis
A0A0Q8E4B7_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	na		failed iqtree; removed from phylogenetic analysis
A0A125Q117_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	na		failed iqtree; removed from phylogenetic analysis

A0A0D0VZH7_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	na		failed iqtree; removed from phylogenetic analysis
C4RGJ4_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	na		failed iqtree; removed from phylogenetic analysis
A0A0M2RLR7_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	na		failed iqtree; removed from phylogenetic analysis
I0L0M5_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	na		failed iqtree; removed from phylogenetic analysis
A0A136PPU2_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	na		failed iqtree; removed from phylogenetic analysis
D1BDF9_SANKS	Bacteria	Actinobacteria	Sanguibacter_keddieii	446469	na		failed iqtree; removed from phylogenetic analysis
A0A0J0V0W2_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	na		failed iqtree; removed from phylogenetic analysis
A0A0Q6FGB3_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	na		failed iqtree; removed from phylogenetic analysis
A0A0Q7JLE7_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	na		failed iqtree; removed from phylogenetic analysis
A0A101SMY0_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	na		failed iqtree; removed from

							phylogenetic analysis
A0A0U3NJK9_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	na		failed iqtree; removed from phylogenetic analysis
O54100_STRCO	Bacteria	Actinobacteria	Streptomyces_coelicolor	100226	na		failed iqtree; removed from phylogenetic analysis
A0A0M8WFP2_9NOCA	Bacteria	Actinobacteria	Nocardiaceae	85025	na		failed iqtree; removed from phylogenetic analysis
A0A0Q4HC32_9MICO	Bacteria	Actinobacteria	Micrococcales	85006	na		failed iqtree; removed from phylogenetic analysis
A0A0X8E3N2_9MICO	Bacteria	Actinobacteria	Micrococcales	85006	na		failed iqtree; removed from phylogenetic analysis
A0A089XKJ5_STRGA	Bacteria	Actinobacteria	Streptomyces_glaucescens	1907	na		failed iqtree; removed from phylogenetic analysis
D2AY66_STRRD	Bacteria	Actinobacteria	Streptosporangium_roseum	479432	na		failed iqtree; removed from phylogenetic analysis
A0A0K9XA78_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	na		failed iqtree; removed from phylogenetic analysis
D2AY38_STRRD	Bacteria	Actinobacteria	Streptosporangium_roseum	479432	na		failed iqtree; removed from phylogenetic analysis

D2AY67_ STRRD	Bacteria	Actinobacteria	Streptosporangium_ roseum	479432	na		failed iqtree; removed from phylogenetic analysis
W5T828_ 9NOCA	Bacteria	Actinobacteria	Nocardiaceae	85025	na		failed iqtree; removed from phylogenetic analysis
A0A0K9XL25 _9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	na		failed iqtree; removed from phylogenetic analysis
U5EK13_ NOCAS	Bacteria	Actinobacteria	Nocardia_asteroides	1824	na		failed iqtree; removed from phylogenetic analysis
A0A1E5P2Y9_ 9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	na		failed iqtree; removed from phylogenetic analysis
G2G7I6_ 9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	na		failed iqtree; removed from phylogenetic analysis
A8L408_ FRASN	Bacteria	Actinobacteria	Frankia_sp	298653	na		failed iqtree; removed from phylogenetic analysis
C9ZEK2_ STRSW	Bacteria	Actinobacteria	Streptomyces_scabiei	680198	na		failed iqtree; removed from phylogenetic analysis
A0A1A0TF16_ 9MYCO	Bacteria	Actinobacteria	Mycobacteriaceae	1762	na		failed iqtree; removed from phylogenetic analysis
D3EZA8_ CONWI	Bacteria	Actinobacteria	Conexibacter_woesei	469383	na		long branch; removed from final

							phylogenetic analysis
C4LLL6_ CORK4	Bacteria	Actinobacteria	Corynebacterium_ kroppenstedtii	645127	na		long branch; removed from final phylogenetic analysis
K0YNF4_ 9ACTO	Bacteria	Actinobacteria	Actinomycetales	2037	na		long branch; removed from final phylogenetic analysis
A0A0F0GP16_ NOCAE	Bacteria	Actinobacteria	Lechevalieria_ aerocolonigenes	68170	na		long branch; removed from final phylogenetic analysis
A0A0Q9J7K0_ 9MICO	Bacteria	Actinobacteria	Micrococcales	85006	na		long branch; removed from final phylogenetic analysis
A0A0Q8CKB6 _9MICO	Bacteria	Actinobacteria	Micrococcales	85006	na		long branch; removed from final phylogenetic analysis
A0A1E8F4W7 _9MICC	Bacteria	Actinobacteria	Micrococcaceae	1268	na		long branch; removed from final phylogenetic analysis
W7T0M6_ 9PSEU	Bacteria	Actinobacteria	Pseudonocardiales	85010	na		long branch; removed from final phylogenetic analysis
F8B0H0_ FRADG	Bacteria	Actinobacteria	Frankia_symbiont subsp	656024	na		long branch; removed from final phylogenetic analysis
E3IZH0_ FRAIE	Bacteria	Actinobacteria	Frankia_inefficax	298654	na		long branch; removed from final phylogenetic analysis

L7F1N0_ 9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	na		long branch; removed from final phylogenetic analysis
A0A124G6X0_ 9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	na		long branch; removed from final phylogenetic analysis
A0A0M8WNX2_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	na		long branch; removed from final phylogenetic analysis
A0A0N0HZV1_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	na		long branch; removed from final phylogenetic analysis
G0FY96_ AMYMS	Bacteria	Actinobacteria	Amycolatopsis_ mediterranei	713604	na		long branch; removed from final phylogenetic analysis
R4SW22_ AMYOR	Bacteria	Actinobacteria	Amycolatopsis_orientalis	31958	na		long branch; removed from final phylogenetic analysis
B1VTB6_ STRGG	Bacteria	Actinobacteria	Streptomyces_ griseus subsp	455632	na		long branch; removed from final phylogenetic analysis
V6UH64_ 9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	na		long branch; removed from final phylogenetic analysis
I2N5E4_ 9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	na		long branch; removed from final phylogenetic analysis
G2GGS3_ 9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	na		long branch; removed from final

							phylogenetic analysis
A0A197SI33_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	na		long branch; removed from final phylogenetic analysis
A0A0A6US84_ACTUT	Bacteria	Actinobacteria	Actinoplanes_utahensis	1869	na		long branch; removed from final phylogenetic analysis
I0L7D6_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	na		long branch; removed from final phylogenetic analysis
A0A0M8YDH2_9ACTN	Bacteria	Actinobacteria	Actinobacteria	201174	na		long branch; removed from final phylogenetic analysis
C1DU99_SULAA	Bacteria	Aquificae	Sulfurihydrogenibium_azorense	204536	1(TYH)		
B2V725_SULSY	Bacteria	Aquificae	Sulfurihydrogenibium_sp	436114	na		failed iqtree; removed from phylogenetic analysis
B8LE39_THAPS	Eukaryote	Bacillariophyta	Thalassiosira_pseudonana	35128	3(NWH)		N terminal ABM domain
K0TJM3_THAOC	Eukaryote	Bacillariophyta	Thalassiosira_oceanica	159749	3(NWH)		N terminal ABM domain
K0TJM3_THAOC	Eukaryote	Bacillariophyta	Thalassiosira_oceanica	159749	3(NWH)		C terminal ABM domain
B8LE39_THAPS	Eukaryote	Bacillariophyta	Thalassiosira_pseudonana	35128	3(NWH)		C terminal ABM domain
A0A1E7F1Z8_9STRA	Eukaryote	Bacillariophyta	Fragilariopsis_cylindrus	635003	3(NWH)		N terminal ABM domain

A0A1E7F1Z8_9STRA	Eukaryote	Bacillariophyta	Fragilariopsis_cylindrus	635003	3(NWH)		C terminal ABM domain
A0A098C3W9_9PORP	Bacteria	Bacteroidetes Chlorobi	Porphyromonadaceae	171551	na		long branch; removed from final phylogenetic analysis
D6THY3_9CHLR	Bacteria	Chloroflexi	Chloroflexi	200795	2(NYH)		
A0A0M9UCJ3_9CHLR	Bacteria	Chloroflexi	Chloroflexi	200795	3(NWH)		
A9WHC7_CHLAA	Bacteria	Chloroflexi	Chloroflexus_aurantiacus	324602	3(NWH)		
A5UZ74_ROSS1	Bacteria	Chloroflexi	Roseiflexus_sp	357808	3(NWH)		
A0A136LHU6_9CHLR	Bacteria	Chloroflexi	Chloroflexi	200795	3(NWH)		
A8I5W8_CHLRE	Eukaryote	Chlorophyta	Chlamydomonas_reinhardtii	3055	3(NWH)	LFO1	failed iqtree; but retained eukaryotic sequence
A0A090N3X1_OSTTA	Eukaryote	Chlorophyta	Ostreococcus_tauri	70448	2(NWR)		C terminal ABM domain
A4S0P7_OSTLU	Eukaryote	Chlorophyta	Ostreococcus_lucimarinus	436017	2(NWR)		C terminal ABM domain
C1MUT0_MICPC	Eukaryote	Chlorophyta	Micromonas_pusilla	564608	3(NWH)		C terminal ABM domain
D8U6C7_VOLCA	Eukaryote	Chlorophyta	Volvox_carteri	3067	3(NWH)		
C1E6M3_MICCC	Eukaryote	Chlorophyta	Micromonas_commoda	296587	3(NWH)		N terminal ABM domain

C1MUT0_ MICPC	Eukaryote	Chlorophyta	Micromonas_pusilla	564608	3(NWH)		N terminal ABM domain
A0A090N3X1_ OSTTA	Eukaryote	Chlorophyta	Ostreococcus_tauri	70448	3(NWH)		N terminal ABM domain
A4S0P7_ OSTLU	Eukaryote	Chlorophyta	Ostreococcus_lucimarinus	436017	3(NWH)		N terminal ABM domain
C1E6M3_ MICCC	Eukaryote	Chlorophyta	Micromonas_commoda	296587	3(NWH)		C terminal ABM domain
Q9YCC3_ AERPE	Archaea	Crenarchaeota	Aeropyrum_pernix	272557	3(NWH)		failed iqtree; but retained archaeal sequence
A0A166TTZ3_ 9CYAN	Bacteria	Cyanobacteria	Cyanobacteria	1117	3(NWH)		
A0A0C1VFH2 _9CYAN	Bacteria	Cyanobacteria	Cyanobacteria	1117	3(NWH)		
B0C735_ ACAM1	Bacteria	Cyanobacteria	Acaryochloris_marina	329726	na		failed iqtree; removed from phylogenetic analysis
A0A0C1Y524_ 9CYAN	Bacteria	Cyanobacteria	Cyanobacteria	1117	na		long branch; removed from final phylogenetic analysis
A0A0C2M485 _9CYAN	Bacteria	Cyanobacteria	Cyanobacteria	1117	na		long branch; removed from final phylogenetic analysis
A0A0C2QMC1 _9CYAN	Bacteria	Cyanobacteria	Cyanobacteria	1117	na		long branch; removed from final phylogenetic analysis
L0A398_ DEIPD	Bacteria	Deinococcus Thermus	Deinococcus_ peraridilitoris	937777	3(NWH)		
H8GT32_	Bacteria	Deinococcus	Deinococcus_gobiensis	745776	3(NWH)		

DEIGI		Thermus					
A0A0Q1AZ35_9DEIO	Bacteria	Deinococcus Thermus	Deinococcaceae	183710	3(NWH)		
A0A0F7JTH1_9DEIO	Bacteria	Deinococcus Thermus	Deinococcaceae	183710	3(NWH)		
Q1IZZ8_DEIGD	Bacteria	Deinococcus Thermus	Deinococcus_geothermalis	319795	3(NWH)		
F0RN07_DEIPM	Bacteria	Deinococcus Thermus	Deinococcus_proteolyticus	693977	3(NWH)		
A0A0A7KGV0_9DEIO	Bacteria	Deinococcus Thermus	Deinococcaceae	183710	3(NWH)		
Q1IW94_DEIGD	Bacteria	Deinococcus Thermus	Deinococcus_geothermalis	319795	3(NWH)		
E8U471_DEIML	Bacteria	Deinococcus Thermus	Deinococcus_maricopensis	709986	3(NWH)		
D7CVU1_TRURR	Bacteria	Deinococcus Thermus	Truepera_radiovictrix	649638	3(NWH)		
Q53VV7_THET8	Bacteria	Deinococcus Thermus	Thermus_thermophilus	300852	3(NWH)		
D3PM89_MEIRD	Bacteria	Deinococcus Thermus	Meiothermus_ruber	504728	3(NWH)		
D7BHV0_MEISD	Bacteria	Deinococcus Thermus	Meiothermus_silvanus	526227	3(NWH)		
E4U5B9_OCEP5	Bacteria	Deinococcus Thermus	Oceanithermus_profundus	670487	3(NWH)		
F2NKW3_MARHT	Bacteria	Deinococcus Thermus	Marinithermus_hydrothermalis	869210	3(NWH)		
Q9RX69_DEIRA	Bacteria	Deinococcus Thermus	Deinococcus_radiodurans	243230	3(NWH)		

F0RKB4_ DEIPM	Bacteria	Deinococcus Thermus	Deinococcus_ proteolyticus	693977	3(NWH)		
Q9RXN8_ DEIRA	Bacteria	Deinococcus Thermus	Deinococcus_radiodurans	243230	3(NWH)		
A0A0F7JLG3_ 9DEIO	Bacteria	Deinococcus Thermus	Deinococcaceae	183710	na		failed iqtree; removed from phylogenetic analysis
T0LQ28_ 9EURY	Archaea	Euryarchaeota	Euryarchaeota	28890	3(NWH)		
A0A166TA40_ 9EURY	Archaea	Euryarchaeota	Euryarchaeota	28890	3(NWH)		
T0N5V9_ 9EURY	Archaea	Euryarchaeota	Euryarchaeota	28890	3(NWH)		
A0A1M4Q4X2 _STAAU	Bacteria	Firmicutes	Staphylococcus_aureus	1280	3(NWH)	IsdG	
A0A133QBY0 _STALU	Bacteria	Firmicutes	Staphylococcus_ lugdunensis	28035	3(NWH)	IsdG	
HDOX_ BACAN	Bacteria	Firmicutes	Bacillus_anthracis	1392	3(NWH)	IsdG	
HDOX_ LISMO	Bacteria	Firmicutes	Listeria_ monocytogenes serovar 1	169963	3(NWH)	IsdG	
A0A113JIK9_ STAAU	Bacteria	Firmicutes	Staphylococcus_aureus	1280	3(NWH)	IsdI	
A8FBM3_ BACP2	Bacteria	Firmicutes	Bacillus_pumilus	315750	1(NFN)		
A0A0A3IHC6_ 9BACI	Bacteria	Firmicutes	Bacillaceae	186817	1(QWL)		
F5SD85_ 9BACL	Bacteria	Firmicutes	Bacillales	1385	2(-WH)		

A0A1B9AQP0_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	2(-WH)		
X0R6N6_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	2(KWH)		
A0A094YSI1_BACAO	Bacteria	Firmicutes	Bacillus_alcalophilus	1445	2(KWH)		
A0A081PAW1_9BACL	Bacteria	Firmicutes	Bacillales	1385	2(KWH)		
H3SCR7_9BACL	Bacteria	Firmicutes	Bacillales	1385	2(KWH)		
C3BLJ9_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	2(KWH)		
Q81P32_BACAN	Bacteria	Firmicutes	Bacillus_anthracis	1392	2(KWH)		
Q3EP94_BACTI	Bacteria	Firmicutes	Bacillus_ thuringiensis subsp	1430	2(KWH)		
Q81C15_BACCR	Bacteria	Firmicutes	Bacillus_cereus	226900	2(KWH)		
A0A073JX78_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	2(KWH)		
M8DI09_9BACL	Bacteria	Firmicutes	Bacillales	1385	2(KWH)		
V9W098_9BACL	Bacteria	Firmicutes	Bacillales	1385	2(KWH)		
A0A0M2VWE2_9BACL	Bacteria	Firmicutes	Bacillales	1385	2(KWH)		
D3E9W0_GEOS4	Bacteria	Firmicutes	Geobacillus_sp	481743	2(KWH)		
G4HG90_9BACL	Bacteria	Firmicutes	Bacillales	1385	2(KWH)		
V9G998_	Bacteria	Firmicutes	Bacillales	1385	2(KWH)		

9BACL							
A0A165Y0L0_ 9BACI	Bacteria	Firmicutes	Bacillaceae	186817	2(KWH)		
A0A0B0IGT8_ 9BACI	Bacteria	Firmicutes	Bacillaceae	186817	2(KWH)		
Q9K649_ BACHD	Bacteria	Firmicutes	Bacillus_halodurans	272558	2(KWH)		
E5WR99_ 9BACI	Bacteria	Firmicutes	Bacillaceae	186817	2(KWH)		
D3FRS6_ BACPE	Bacteria	Firmicutes	Bacillus_pseudofirmus	398511	2(KWH)		
Q5WGM4_ BACSK	Bacteria	Firmicutes	Bacillus_clausii	66692	2(KWH)		
A0A0M0KZX0 _9BACI	Bacteria	Firmicutes	Bacillaceae	186817	2(KWH)		
A0A147KBN1 _9BACI	Bacteria	Firmicutes	Bacillaceae	186817	2(KWH)		
W4ERL6_ 9BACL	Bacteria	Firmicutes	Bacillales	1385	2(KWH)		
A0A077J893_ 9BACI	Bacteria	Firmicutes	Bacillaceae	186817	2(KWH)		
A0A0W7Y8E0 _9BACI	Bacteria	Firmicutes	Bacillaceae	186817	2(KWH)		
X2GXF4_ 9BACI	Bacteria	Firmicutes	Bacillaceae	186817	2(KWH)		
A0A0N0S3J2_ 9BACI	Bacteria	Firmicutes	Bacillaceae	186817	2(KWH)		
A0A0M9DGZ9 _9BACI	Bacteria	Firmicutes	Bacillaceae	186817	2(KWH)		
A0A098EP32_ 9BACL	Bacteria	Firmicutes	Bacillales	1385	2(KWH)		
C0Z890_ BREBN	Bacteria	Firmicutes	Brevibacillus_brevis	358681	2(KWH)		

J3BB78_ 9BACL	Bacteria	Firmicutes	Bacillales	1385	2(KWH)		
W4EQI4_ 9BACL	Bacteria	Firmicutes	Bacillales	1385	2(NWF)		
A0A0Q3WYY5_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	2(NWQ)		
F5LH48_ 9BACL	Bacteria	Firmicutes	Bacillales	1385	2(NYH)		
A0A0P9CY60_9BACL	Bacteria	Firmicutes	Bacillales	1385	2(NYH)		
A0A0D3VFD1_9BACL	Bacteria	Firmicutes	Bacillales	1385	2(RWH)		
C6J0B9_ 9BACL	Bacteria	Firmicutes	Bacillales	1385	2(RWH)		
A0A090ZAP8_PAEMA	Bacteria	Firmicutes	Paenibacillus_macerans	44252	2(RWH)		
A0A098MDH7_9BACL	Bacteria	Firmicutes	Bacillales	1385	2(RWH)		
A0A0U2W9R9_9BACL	Bacteria	Firmicutes	Bacillales	1385	2(RWH)		
H6NBK8_ 9BACL	Bacteria	Firmicutes	Bacillales	1385	2(RWH)		
A0A1C1A6W3_9BACL	Bacteria	Firmicutes	Bacillales	1385	2(RWH)		
A0A0Q9M2T8_9BACL	Bacteria	Firmicutes	Bacillales	1385	2(RWH)		
A0A0A2TXC3_9BACL	Bacteria	Firmicutes	Bacillales	1385	2(RWH)		
A0A089KXF0_9BACL	Bacteria	Firmicutes	Bacillales	1385	2(RWH)		
A0A101XXE6_9BACL	Bacteria	Firmicutes	Bacillales	1385	2(RWH)		
A0A0B0HT13_9BACL	Bacteria	Firmicutes	Bacillales	1385	2(RWH)		

A0A0A0E7R2_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	2(RWH)		
HMOA_BACSU	Bacteria	Firmicutes	Bacillus_subtilis	224308	2(RWH)		
V6M971_9BACL	Bacteria	Firmicutes	Bacillales	1385	2(RWH)		
A0A0M4FU31_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	2(RWH)		
A0A0K9H1J9_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	2(RWH)		
A0A1E3L0E2_9BACL	Bacteria	Firmicutes	Bacillales	1385	2(RWH)		
A0A172ZIK3_9BACL	Bacteria	Firmicutes	Bacillales	1385	2(RWH)		
A0A0X8CQU1_9BACL	Bacteria	Firmicutes	Bacillales	1385	2(RWH)		
A0A069DA78_9BACL	Bacteria	Firmicutes	Bacillales	1385	2(RWH)		
A0A0A3J467_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	2(RWH)		
A0A0M9X0H6_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	2(RWH)		
F2F6A5_SOLSS	Bacteria	Firmicutes	Solibacillus_silvestris	1002809	2(RWH)		
U5L5Z8_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	2(RWH)		
A0A143HAZ2_9BACL	Bacteria	Firmicutes	Bacillales	1385	2(RWH)		
A0A1E7DKB7_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	2(SWH)		
A0A0V8HGS6_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	2(SWH)		
A0A0N8GH93_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	2(SWH)		
A6CK76_	Bacteria	Firmicutes	Bacillaceae	186817	2(SWH)		

9BACI							
A0A0M0G691_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	2(SWH)		
A0A0M9GT07_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	2(SWH)		
A0A075JNZ2_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	2(SWH)		
W4Q1K6_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	2(SWH)		
W4QU03_BACA3	Bacteria	Firmicutes	Bacillus_akibai	1236973	2(SWH)		
D3FSS8_BACPE	Bacteria	Firmicutes	Bacillus_pseudofirmus	398511	2(SWH)		
U5LEG5_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	2(SWH)		
A0A0A2TAH1_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	2(SWH)		
A0A0A5FWU8_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	2(SWH)		
I0JQ76_HALH3	Bacteria	Firmicutes	Halobacillus_halophilus	866895	2(SWH)		
L5N9C9_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	2(SWH)		
A0A0B0D545_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	2(SWH)		
K2GDC9_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	2(SWH)		
I8J6H2_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	2(SWH)		
D5DX46_BACMQ	Bacteria	Firmicutes	Bacillus_megaterium	545693	2(SWH)		
A0A060M2B1_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	2(SWH)		

A0A0K9HAY5_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	2(SWH)		
A0A0J5QS38_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	2(SWH)		
A0A061PCL7_9BACL	Bacteria	Firmicutes	Bacillales	1385	2(SWH)		
A0A061P2I0_9BACL	Bacteria	Firmicutes	Bacillales	1385	2(SWH)		
A0A0F5HMR0_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	2(SWH)		
H6NTN3_9BACL	Bacteria	Firmicutes	Bacillales	1385	2(SWH)		
A0A172ZGC6_9BACL	Bacteria	Firmicutes	Bacillales	1385	2(SWH)		
W9B787_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	2(SWH)		
A0A0U2XTX3_9BACL	Bacteria	Firmicutes	Bacillales	1385	2(SWH)		
A0A098ELL9_9BACL	Bacteria	Firmicutes	Bacillales	1385	2(SWH)		
Q3EZT5_BACTI	Bacteria	Firmicutes	Bacillus_thuringiensis subsp	1430	3(NWH)		
Q81GW8_BACCR	Bacteria	Firmicutes	Bacillus_cereus	226900	3(NWH)		
Q81U25_BACAN	Bacteria	Firmicutes	Bacillus_anthracis	1392	3(NWH)		
C3BGZ8_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
A0A073K0P6_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
Q65LR6_BACLD	Bacteria	Firmicutes	Bacillus_licheniformis	279010	3(NWH)		
HMOB_	Bacteria	Firmicutes	Bacillus_subtilis	224308	3(NWH)		

BACSU							
A0A084GJ82_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
A0A176J6F3_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
C5D6M3_GEOSW	Bacteria	Firmicutes	Geobacillus_sp	471223	3(NWH)		
A0A160FDJ3_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
A0A023DET4_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
A0A160F3A4_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
A0A094LBK6_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
B7GF20_ANOFW	Bacteria	Firmicutes	Anoxybacillus_flavithermus	491915	3(NWH)		
A0A0M0LGX8_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
A0A0V8JPG2_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
D5DZ10_BACMQ	Bacteria	Firmicutes	Bacillus_megaterium	545693	3(NWH)		
A0A0M9GT57_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
A0A0H4KSL1_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
A0A147KAF9_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
U5L681_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
A0A1E7DLK5_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
A0A177LCL9_9BACL	Bacteria	Firmicutes	Bacillales	1385	3(NWH)		

A0A0D6ZCH9_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
S2XZX7_9BACL	Bacteria	Firmicutes	Bacillales	1385	3(NWH)		
A6CQP3_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
A0A0P6W9N5_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
A0A0V8HI78_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
A0A0M0GRK0_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
A0A0C2V9I6_9BACL	Bacteria	Firmicutes	Bacillales	1385	3(NWH)		
A0A0B5APR3_9BACL	Bacteria	Firmicutes	Bacillales	1385	3(NWH)		
A0A0C2RZU0_9BACL	Bacteria	Firmicutes	Bacillales	1385	3(NWH)		
A0A0K9F0X1_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
X2H0M4_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
A0A0W7YQ34_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
A0A0M9DKG4_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
A0A0M8QEV7_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
A0A0A3HTA3_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
A0A0A3IAZ4_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
A0A150YIB3_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
A0A087N393_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		

A0A0B4RA39_9BACL	Bacteria	Firmicutes	Bacillales	1385	3(NWH)		
A0A0U2Q7B3_9BACL	Bacteria	Firmicutes	Bacillales	1385	3(NWH)		
A0A098EMF6_9BACL	Bacteria	Firmicutes	Bacillales	1385	3(NWH)		
A0A0F5I5J4_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
A0A0K9H2T6_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
A0A0A8JI35_BACSX	Bacteria	Firmicutes	Bacillus_sp	98228	3(NWH)		
A0A135WH95_9BACL	Bacteria	Firmicutes	Bacillales	1385	3(NWH)		
M7NDU8_9BACL	Bacteria	Firmicutes	Bacillales	1385	3(NWH)		
A0A098F640_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
A0A0M4GFR5_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
A0A0M2SPC9_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
N0APC6_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
A0A0M1NQC8_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
A0A0A2TAW6_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
A0A0A2UQ22_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
A0A0A5FY70_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
A0A0A5GN86_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
I0JKB1_HALH3	Bacteria	Firmicutes	Halobacillus_halophilus	866895	3(NWH)		

A0A0B0D9E7_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
A0A059NZ95_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
L5NA23_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
A0A024QEL8_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
A0A0L0QTY2_VIRPA	Bacteria	Firmicutes	Virgibacillus_pantothenticus	1473	3(NWH)		
A0A160IIV5_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
A0A0V8JAY9_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
I8AJW9_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
W4Q1R3_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
A0A0B0IMC5_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
W4QUC6_BACA3	Bacteria	Firmicutes	Bacillus_akibai	1236973	3(NWH)		
D3FVT5_BACPE	Bacteria	Firmicutes	Bacillus_pseudofirmus	398511	3(NWH)		
A0A0U2PDV0_9BACL	Bacteria	Firmicutes	Bacillales	1385	3(NWH)		
A0A098EHV5_9BACL	Bacteria	Firmicutes	Bacillales	1385	3(NWH)		
A0A0B4REM9_9BACL	Bacteria	Firmicutes	Bacillales	1385	3(NWH)		
A0A172ZLV9_9BACL	Bacteria	Firmicutes	Bacillales	1385	3(NWH)		
A0A1E3L037_9BACL	Bacteria	Firmicutes	Bacillales	1385	3(NWH)		

A0A0A2U6Y8_9BACL	Bacteria	Firmicutes	Bacillales	1385	3(NWH)		
A0A069DIB5_9BACL	Bacteria	Firmicutes	Bacillales	1385	3(NWH)		
A0A075LJD0_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
A0A075JN73_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
W9AAU2_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
A0A061P5Z4_9BACL	Bacteria	Firmicutes	Bacillales	1385	3(NWH)		
D3FV58_BACPE	Bacteria	Firmicutes	Bacillus_pseudofirmus	398511	3(NWH)		
A0A0J6CT63_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
A0A1E5LEK5_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
W7YU30_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
A0A060M1C3_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
Q5WHZ8_BACSK	Bacteria	Firmicutes	Bacillus_clausii	66692	3(NWH)		
E6TTI9_BACCJ	Bacteria	Firmicutes	Bacillus_cellulosilyticus	649639	3(NWH)		
A0A1D7QWV8_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
D6XXL8_BACIE	Bacteria	Firmicutes	Bacillus_selenitireducens	439292	3(NWH)		
A0A0D1XEQ0_ANEMI	Bacteria	Firmicutes	Aneurinibacillus_migulanus	47500	3(NWH)		
A0A0X8D6B5_9BACL	Bacteria	Firmicutes	Bacillales	1385	3(NWH)		

A0A0Q3WG02_BRECH	Bacteria	Firmicutes	Brevibacillus_ choshinensis	54911	3(NWH)		
C0ZEI8_ BREBN	Bacteria	Firmicutes	Brevibacillus_brevis	358681	3(NWH)		
J2QRN4_ 9BACL	Bacteria	Firmicutes	Bacillales	1385	3(NWH)		
V6M161_ 9BACL	Bacteria	Firmicutes	Bacillales	1385	3(NWH)		
A0A1A5XSH9_9BACL	Bacteria	Firmicutes	Bacillales	1385	3(NWH)		
M8DCD8_ 9BACL	Bacteria	Firmicutes	Bacillales	1385	3(NWH)		
A0A074LTF9_9BACL	Bacteria	Firmicutes	Bacillales	1385	3(NWH)		
HDOX1_ STAA8	Bacteria	Firmicutes	Staphylococcus_aureus	93061	3(NWH)		
B9DIC2_ STACT	Bacteria	Firmicutes	Staphylococcus_carnosus	396513	3(NWH)		
K9ARK7_ 9STAP	Bacteria	Firmicutes	Staphylococcaceae	90964	3(NWH)		
A0A033UIU7_STAAU	Bacteria	Firmicutes	Staphylococcus_aureus	1280	3(NWH)		
H0DF41_ 9STAP	Bacteria	Firmicutes	Staphylococcaceae	90964	3(NWH)		
HDOX2_ STAA8	Bacteria	Firmicutes	Staphylococcus_aureus	93061	3(NWH)		
A0A0A8HSN4_STAHY	Bacteria	Firmicutes	Staphylococcus_hyicus	1284	3(NWH)		
E8SGB4_ STAPH	Bacteria	Firmicutes	Staphylococcus_ pseudintermedius	937773	3(NWH)		

A0A0A5G003_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
A0A0X1RW94_9BACL	Bacteria	Firmicutes	Bacillales	1385	3(NWH)		
B7GLW7_ANOFW	Bacteria	Firmicutes	Anoxybacillus_flavithermus	491915	3(NWH)		
A0A0B0HN95_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
A0A165Y103_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
C6D4S2_PAESJ	Bacteria	Firmicutes	Paenibacillus_sp	324057	3(NWH)		
A0A1B8WF14_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
A0A0Q9L3G6_9BACL	Bacteria	Firmicutes	Bacillales	1385	3(NWH)		
A0A0Q4QZ50_9BACL	Bacteria	Firmicutes	Bacillales	1385	3(NWH)		
A0A172TIU4_9BACL	Bacteria	Firmicutes	Bacillales	1385	3(NWH)		
A0A098ME35_9BACL	Bacteria	Firmicutes	Bacillales	1385	3(NWH)		
A0A0E4HII4_9BACL	Bacteria	Firmicutes	Bacillales	1385	3(NWH)		
A0A089ISN0_9BACL	Bacteria	Firmicutes	Bacillales	1385	3(NWH)		
A0A1B8VUE2_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
X5A6L5_9BACL	Bacteria	Firmicutes	Bacillales	1385	3(NWH)		
A0A101XVJ3_9BACL	Bacteria	Firmicutes	Bacillales	1385	3(NWH)		
A0A089LYL2_9BACL	Bacteria	Firmicutes	Bacillales	1385	3(NWH)		
V9GA49_9BACL	Bacteria	Firmicutes	Bacillales	1385	3(NWH)		

D3EG83_ GEOS4	Bacteria	Firmicutes	Geobacillus_sp	481743	3(NWH)		
G4HCL9_ 9BACL	Bacteria	Firmicutes	Bacillales	1385	3(NWH)		
A0A0M2VSB6 _9BACL	Bacteria	Firmicutes	Bacillales	1385	3(NWH)		
K4ZNN3_ PAEAL	Bacteria	Firmicutes	Paenibacillus_alvei	44250	3(NWH)		
H3SA54_ 9BACL	Bacteria	Firmicutes	Bacillales	1385	3(NWH)		
W7YNV8_ 9BACL	Bacteria	Firmicutes	Bacillales	1385	3(NWH)		
V9W8P9_ 9BACL	Bacteria	Firmicutes	Bacillales	1385	3(NWH)		
A0A087N2X2_ 9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
A0A127W3I6_ SPOPS	Bacteria	Firmicutes	Sporosarcina_ psychrophila	1476	3(NWH)		
A0A0F7HHW9 _9STAP	Bacteria	Firmicutes	Staphylococcaceae	90964	3(NWH)		
A0A0A1MRH 4_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
A0A075JMY4 _9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
A0A0A8JKG5 _BACSX	Bacteria	Firmicutes	Bacillus_sp	98228	3(NWH)		
W4QJZ4_ 9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
A0A094XEH3 _BACAO	Bacteria	Firmicutes	Bacillus_alcalophilus	1445	3(NWH)		
A0A0J5R0C6_ 9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
HDOX_	Bacteria	Firmicutes	Bacillus_halodurans	272558	3(NWH)		

BACHD							
HDOX_ BACCR	Bacteria	Firmicutes	Bacillus_cereus	226900	3(NWH)		
A0A1A5YRT1 _9BACL	Bacteria	Firmicutes	Bacillales	1385	3(NWH)		
C6CXI8_ PAESJ	Bacteria	Firmicutes	Paenibacillus_sp	324057	3(NWH)		
A0A0U2W3V3 _9BACL	Bacteria	Firmicutes	Bacillales	1385	3(NWH)		
A0A081NTU9 _9BACL	Bacteria	Firmicutes	Bacillales	1385	3(NWH)		
E3E6U8_ PAEPS	Bacteria	Firmicutes	Paenibacillus_polymyxa	886882	3(NWH)		
A0A090Z770_ PAEMA	Bacteria	Firmicutes	Paenibacillus_macerans	44252	3(NWH)		
A0A172ZKW0 _9BACL	Bacteria	Firmicutes	Bacillales	1385	3(NWH)		
A0A1E3L9N0_ 9BACL	Bacteria	Firmicutes	Bacillales	1385	3(NWH)		
A0A089KXY8 _9BACL	Bacteria	Firmicutes	Bacillales	1385	3(NWH)		
W7ZG88_ 9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
HDOX_ BACSK	Bacteria	Firmicutes	Bacillus_clausii	66692	3(NWH)		
A0A060LXF3_ 9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
A0A061PCH7_ 9BACL	Bacteria	Firmicutes	Bacillales	1385	3(NWH)		
A0A0K9GXP1 _9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
A0A0D0SMA5 _STAGA	Bacteria	Firmicutes	Staphylococcus_ gallinarum	1293	3(NWH)		
H0DFD1_	Bacteria	Firmicutes	Staphylococcaceae	90964	3(NWH)		

9STAP							
HDOX_ STAEQ	Bacteria	Firmicutes	Staphylococcus_ epidermidis	176279	3(NWH)		
A0A0M0L5X0 _9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
A0A0U2Z734_ 9BACL	Bacteria	Firmicutes	Bacillales	1385	3(NWH)		
A0A098EHK9 _9BACL	Bacteria	Firmicutes	Bacillales	1385	3(NWH)		
A0A094WRD1 _BACAO	Bacteria	Firmicutes	Bacillus_alcalophilus	1445	3(NWH)		
A0A143HGY7 _9BACL	Bacteria	Firmicutes	Bacillales	1385	3(NWH)		
W4F3U8_ 9BACL	Bacteria	Firmicutes	Bacillales	1385	3(NWH)		
A0A109Q904_ 9BACL	Bacteria	Firmicutes	Bacillales	1385	3(NWH)		
A0A0A0E503_ 9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
B9E9V0_ MACCJ	Bacteria	Firmicutes	Macrococcus_caseolyticus	458233	3(NWH)		
A0A033UDM6 _STAAU	Bacteria	Firmicutes	Staphylococcus_aureus	1280	3(NWH)		
A0A0N0CUB3 _9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
A0A0W7Y2H8 _9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
A0A0K9FHD9 _9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
X2GY49_ 9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
A0A0A3IG79_ 9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
F2FAT2_ SOLSS	Bacteria	Firmicutes	Solibacillus_silvestris	1002809	3(NWH)		

A0A168Q1M1_9BACL	Bacteria	Firmicutes	Bacillales	1385	3(NWH)		
A0A167BAI0_9BACL	Bacteria	Firmicutes	Bacillales	1385	3(NWH)		
A0A0M1NS72_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
B1YG40_EXIS2	Bacteria	Firmicutes	Exiguobacterium_sibiricum	262543	3(NWH)		
A0A1C0YLB4_9BACL	Bacteria	Firmicutes	Bacillales	1385	3(NWH)		
A0A1B9AIK4_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
A0A0K9GCA9_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
B9DM71_STACT	Bacteria	Firmicutes	Staphylococcus_carnosus	396513	3(NWH)		
A0A0M8Q3Q9_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
A0A1C0YDF8_9BACL	Bacteria	Firmicutes	Bacillales	1385	3(NWH)		
A0A150YFW9_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	3(NWH)		
A0A0R2HW60_CARDV	Bacteria	Firmicutes	Carnobacterium_divergens	2748	3(NWH)		
W7C1E8_9LIST	Bacteria	Firmicutes	Listeriaceae	186820	3(NWH)		
W7B0I4_9LIST	Bacteria	Firmicutes	Listeriaceae	186820	3(NWH)		
W7CSV2_BROTH	Bacteria	Firmicutes	Brochothrix_thermosphacta	2756	3(NWH)		
K8ELZ5_CARML	Bacteria	Firmicutes	Carnobacterium_maltaromaticum	2751	3(NWH)		
E6LHR8_	Bacteria	Firmicutes	Enterococcaceae	81852	3(NWH)		

9ENTE							
S0KV52_ 9ENTE	Bacteria	Firmicutes	Enterococcaceae	81852	3(NWH)		
C4L0B6_ EXISA	Bacteria	Firmicutes	Exiguobacterium_sp	360911	3(NWH)		
B1YGM6_ EXIS2	Bacteria	Firmicutes	Exiguobacterium_ sibiricum	262543	3(NWH)		
A0A0N0HFV4 _THEVU	Bacteria	Firmicutes	Thermoactinomyces_ vulgaris	2026	3(NWH)		
F5SIK2_ 9BACL	Bacteria	Firmicutes	Bacillales	1385	3(NWH)		
D7UWG3_ LISGR	Bacteria	Firmicutes	Listeria_grayi	1641	3(NWH)		
A0A0A8X505_ 9BACI	Bacteria	Firmicutes	Bacillaceae	186817	na		failed iqtree; removed from phylogenetic analysis
A0A127W1Z9 _SPOPS	Bacteria	Firmicutes	Sporosarcina_ psychrophila	1476	na		failed iqtree; removed from phylogenetic analysis
A0A0A3J6N7_ 9BACI	Bacteria	Firmicutes	Bacillaceae	186817	na		failed iqtree; removed from phylogenetic analysis
A0A0K9GS06 _9BACI	Bacteria	Firmicutes	Bacillaceae	186817	na		failed iqtree; removed from phylogenetic analysis
A0A0H4PEM5 _9BACI	Bacteria	Firmicutes	Bacillaceae	186817	na		failed iqtree; removed from phylogenetic analysis

A0A0Q3TLX0_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	na		failed iqtree; removed from phylogenetic analysis
K2FL47_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	na		failed iqtree; removed from phylogenetic analysis
A0A179SRN8_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	na		failed iqtree; removed from phylogenetic analysis
A0A0M2PHP3_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	na		failed iqtree; removed from phylogenetic analysis
W9AII1_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	na		failed iqtree; removed from phylogenetic analysis
A0A0M4FVC5_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	na		failed iqtree; removed from phylogenetic analysis
A0A0A1MWT7_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	na		failed iqtree; removed from phylogenetic analysis
Q8ERY1_OCEIH	Bacteria	Firmicutes	Oceanobacillus_iheyensis	221109	na		failed iqtree; removed from phylogenetic analysis
W1SNN8_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	na		failed iqtree; removed from phylogenetic analysis
G2TP25_BACCO	Bacteria	Firmicutes	Bacillus_coagulans	1398	na		failed iqtree; removed from

							phylogenetic analysis
A0A0A6VGY2_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	na		failed iqtree; removed from phylogenetic analysis
A0A0M0X659_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	na		failed iqtree; removed from phylogenetic analysis
A0A165YGU6_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	na		failed iqtree; removed from phylogenetic analysis
A0A0L0QMT2_VIRPA	Bacteria	Firmicutes	Virgibacillus_ pantothenticus	1473	na		failed iqtree; removed from phylogenetic analysis
A0A061PEL7_9BACL	Bacteria	Firmicutes	Bacillales	1385	na		failed iqtree; removed from phylogenetic analysis
A0A0D5NJZ5_9BACL	Bacteria	Firmicutes	Bacillales	1385	na		failed iqtree; removed from phylogenetic analysis
A0A0F5R950_9BACL	Bacteria	Firmicutes	Bacillales	1385	na		failed iqtree; removed from phylogenetic analysis
A0A147KBP7_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	na		failed iqtree; removed from phylogenetic analysis
A0A0U2MCR9_9BACL	Bacteria	Firmicutes	Bacillales	1385	na		failed iqtree; removed from phylogenetic analysis

W7YBT6_ 9BACL	Bacteria	Firmicutes	Bacillales	1385	na		failed iqtree; removed from phylogenetic analysis
X0RKQ6_ 9BACI	Bacteria	Firmicutes	Bacillaceae	186817	na		failed iqtree; removed from phylogenetic analysis
A0A0Q9Y6W1 _9BACI	Bacteria	Firmicutes	Bacillaceae	186817	na		failed iqtree; removed from phylogenetic analysis
A0A0Q9YJS8_ 9BACI	Bacteria	Firmicutes	Bacillaceae	186817	na		failed iqtree; removed from phylogenetic analysis
A0A075LN66_ 9BACI	Bacteria	Firmicutes	Bacillaceae	186817	na		failed iqtree; removed from phylogenetic analysis
HDOX_STAS1	Bacteria	Firmicutes	Staphylococcus_ saprophyticus subsp	342451	na		failed iqtree; removed from phylogenetic analysis
A0A0Q3W0P3 _9BACI	Bacteria	Firmicutes	Bacillaceae	186817	na		failed iqtree; removed from phylogenetic analysis
A0A0J5QVM0 _9BACI	Bacteria	Firmicutes	Bacillaceae	186817	na		failed iqtree; removed from phylogenetic analysis
H7F935_9LIST	Bacteria	Firmicutes	Listeriaceae	186820	na		failed iqtree; removed from phylogenetic analysis
D7UX71_ LISGR	Bacteria	Firmicutes	Listeria_grayi	1641	na		failed iqtree; removed from

							phylogenetic analysis
R2T3V6_9ENTE	Bacteria	Firmicutes	Enterococcaceae	81852	na		failed iqtree; removed from phylogenetic analysis
A0A1E5GHP0_9ENTE	Bacteria	Firmicutes	Enterococcaceae	81852	na		failed iqtree; removed from phylogenetic analysis
W4QUC0_BACA3	Bacteria	Firmicutes	Bacillus_akibai	1236973	na		failed iqtree; removed from phylogenetic analysis
A0A167G471_9BACL	Bacteria	Firmicutes	Bacillales	1385	na		failed iqtree; removed from phylogenetic analysis
A0A168J641_9BACL	Bacteria	Firmicutes	Bacillales	1385	na		failed iqtree; removed from phylogenetic analysis
E6TVW7_BACCJ	Bacteria	Firmicutes	Bacillus_cellulosilyticus	649639	na		failed iqtree; removed from phylogenetic analysis
F5LEQ5_9BACL	Bacteria	Firmicutes	Bacillales	1385	na		failed iqtree; removed from phylogenetic analysis
A0A0D1XUU2_ANEMI	Bacteria	Firmicutes	Aneurinibacillus_migulanus	47500	na		failed iqtree; removed from phylogenetic analysis
K4ZMN9_PAEAL	Bacteria	Firmicutes	Paenibacillus_alvei	44250	na		failed iqtree; removed from phylogenetic analysis

A0A1A5XRX8_9BACL	Bacteria	Firmicutes	Bacillales	1385	na		failed iqtree; removed from phylogenetic analysis
A0A075RBM5_BRELA	Bacteria	Firmicutes	Brevibacillus_laterosporus	1465	na		failed iqtree; removed from phylogenetic analysis
A0A0Q3S0D5_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	na		failed iqtree; removed from phylogenetic analysis
A0A0K9FI75_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	na		failed iqtree; removed from phylogenetic analysis
A0A0A8X2K0_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	na		failed iqtree; removed from phylogenetic analysis
W4QEP1_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	na		failed iqtree; removed from phylogenetic analysis
A0A0M2SW87_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	na		failed iqtree; removed from phylogenetic analysis
A0A150L1D3_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	na		failed iqtree; removed from phylogenetic analysis
A0A0M1NRB8_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	na		failed iqtree; removed from phylogenetic analysis
A0A0M0XBX0_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	na		failed iqtree; removed from

							phylogenetic analysis
A0A0Q3SZK8_BRECH	Bacteria	Firmicutes	Brevibacillus_ choshinensis	54911	na		failed iqtree; removed from phylogenetic analysis
W4Q784_ 9BACI	Bacteria	Firmicutes	Bacillaceae	186817	na		failed iqtree; removed from phylogenetic analysis
A0A0F5R109_9BACL	Bacteria	Firmicutes	Bacillales	1385	na		failed iqtree; removed from phylogenetic analysis
A0A0K9G9J4_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	na		failed iqtree; removed from phylogenetic analysis
A0A1B9AMJ8_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	na		failed iqtree; removed from phylogenetic analysis
A0A0A3IMY4_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	na		failed iqtree; removed from phylogenetic analysis
A0A0A3HSG0_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	na		failed iqtree; removed from phylogenetic analysis
A0A0K9GYI4_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	na		failed iqtree; removed from phylogenetic analysis
A0A1E7DVL1_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	na		failed iqtree; removed from phylogenetic analysis

E6SMP9_ THEM7	Bacteria	Firmicutes	Thermaerobacter_ marianensis	644966	na		failed iqtree; removed from phylogenetic analysis
A0A084GW15_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	na		failed iqtree; removed from phylogenetic analysis
X0RBC2_ 9BACI	Bacteria	Firmicutes	Bacillaceae	186817	na		failed iqtree; removed from phylogenetic analysis
A0A0D0SQW0_STAGA	Bacteria	Firmicutes	Staphylococcus_ gallinarum	1293	na		failed iqtree; removed from phylogenetic analysis
Q2FV13_ STAA8	Bacteria	Firmicutes	Staphylococcus_aureus	93061	na		failed iqtree; removed from phylogenetic analysis
Q5HL14_ STAEQ	Bacteria	Firmicutes	Staphylococcus_ epidermidis	176279	na		failed iqtree; removed from phylogenetic analysis
A0A133Q6N2_STALU	Bacteria	Firmicutes	Staphylococcus_ lugdunensis	28035	na		failed iqtree; removed from phylogenetic analysis
A0A0D1VE95_ANEMI	Bacteria	Firmicutes	Aneurinibacillus_ migulanus	47500	na		failed iqtree; removed from phylogenetic analysis
A0A179T016_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	na		failed iqtree; removed from phylogenetic analysis
A0A0M2PMN6_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	na		failed iqtree; removed from

							phylogenetic analysis
R2SIR2_ 9ENTE	Bacteria	Firmicutes	Enterococcaceae	81852	na		failed iqtree; removed from phylogenetic analysis
A0A073JXS6_ 9BACI	Bacteria	Firmicutes	Bacillaceae	186817	na		failed iqtree; removed from phylogenetic analysis
E5WEM3_ 9BACI	Bacteria	Firmicutes	Bacillaceae	186817	na		failed iqtree; removed from phylogenetic analysis
I3E7G7_ BACMT	Bacteria	Firmicutes	Bacillus_methanolicus	1471	na		failed iqtree; removed from phylogenetic analysis
A0A0A5HM81 _9BACI	Bacteria	Firmicutes	Bacillaceae	186817	na		failed iqtree; removed from phylogenetic analysis
A0A075JRI1_9 BACI	Bacteria	Firmicutes	Bacillaceae	186817	na		failed iqtree; removed from phylogenetic analysis
A0A0U3W677 _9BACI	Bacteria	Firmicutes	Bacillaceae	186817	na		failed iqtree; removed from phylogenetic analysis
A0A0J6CVW4 _9BACI	Bacteria	Firmicutes	Bacillaceae	186817	na		failed iqtree; removed from phylogenetic analysis
N4W5X2_ 9BACI	Bacteria	Firmicutes	Bacillaceae	186817	na		failed iqtree; removed from phylogenetic analysis

A0A0W7YP89_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	na		failed iqtree; removed from phylogenetic analysis
H6NI20_9BACL	Bacteria	Firmicutes	Bacillales	1385	na		failed iqtree; removed from phylogenetic analysis
G2TKN9_BACCO	Bacteria	Firmicutes	Bacillus_coagulans	1398	na		failed iqtree; removed from phylogenetic analysis
A0A176J7V1_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	na		failed iqtree; removed from phylogenetic analysis
A0A0M2STW3_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	na		failed iqtree; removed from phylogenetic analysis
E8SI76_STAPH	Bacteria	Firmicutes	Staphylococcus_pseudintermedius	937773	na		failed iqtree; removed from phylogenetic analysis
A0A078M7T3_9STAP	Bacteria	Firmicutes	Staphylococcaceae	90964	na		failed iqtree; removed from phylogenetic analysis
A0A090IX75_9BACI	Bacteria	Firmicutes	Bacillaceae	186817	na		failed iqtree; removed from phylogenetic analysis
F5WRX8_ERYRF	Bacteria	Firmicutes	Erysipelothrix_rhusiopathiae	650150	na		failed iqtree; removed from phylogenetic analysis
I0ILF5_LEPFC	Bacteria	Nitrospirae	Leptospirillum_ferrooxidans	1162668	3(NWH)		

J9ZA06_ LEPFM	Bacteria	Nitrospirae	Leptospirillum_ ferriphilum	1048260	3(NWH)		
A0A0G4GEP9 _VITBC	Eukaryote	other_Alveolata	Vitrella_brassicaformis	1169540	1(TFH)		N terminal ABM domain
A0A0G4GEP9 _VITBC	Eukaryote	other_Alveolata	Vitrella_brassicaformis	1169540	3(NWH)		C terminal ABM domain
A0A136P287_ 9BACT	Bacteria	other_Bacteria	Bacteria	2	2(-WH)		
A0A0E3ZHP3 _9BACT	Bacteria	other_Bacteria	Bacteria	2	2(SWH)		
J1FFB7_ 9BACT	Bacteria	other_Bacteria	Bacteria	2	2(SWH)		
A0A0P0C6U1_ 9BACT	Bacteria	other_Bacteria	Bacteria	2	2(SWH)		
A0A127B3Z1_ 9BACT	Bacteria	other_Bacteria	Bacteria	2	2(SWH)		
W4LDS2_ 9BACT	Bacteria	other_Bacteria	Bacteria	2	3(NWH)		
W4L9X6_ 9BACT	Bacteria	other_Bacteria	Bacteria	2	3(NWH)		
T0XVM2_ 9BACT	Bacteria	other_Bacteria	Bacteria	2	3(NWH)		
C6HXU1_ 9BACT	Bacteria	other_Bacteria	Bacteria	2	3(NWH)		
W4LYY6_ 9BACT	Bacteria	other_Bacteria	Bacteria	2	3(NWH)		
W4LTU4_ 9BACT	Bacteria	other_Bacteria	Bacteria	2	3(NWH)		
D5MMH7_ 9BACT	Bacteria	other_Bacteria	Bacteria	2	3(NWH)		
D5MJI1_	Bacteria	other_Bacteria	Bacteria	2	3(NWH)		

9BACT							
A0A142L7F6_ 9BACT	Bacteria	other_Bacteria	Bacteria	2	3(NWH)		
A0A136L2D3_ 9BACT	Bacteria	other_Bacteria	Bacteria	2	3(NWH)		
A0A191TLC1_ 9BACT	Bacteria	other_Bacteria	Bacteria	2	na		failed iqtree; removed from phylogenetic analysis
W0RBW4_ 9BACT	Bacteria	other_Bacteria	Bacteria	2	na		failed iqtree; removed from phylogenetic analysis
F0XXQ6_ AURAN	Eukaryote	other_ Stramenopiles	Aureococcus_ anophagefferens	44056	2(NWQ)		failed iqtree; but retained eukaryotic sequence
Q09CQ8_ STIAD	Bacteria	Proteobacteria	Stigmatella_aurantiaca	378806	0(SYK)		
A0A0G2ZMH8 _9DELT	Bacteria	Proteobacteria	Deltaproteobacteria	28221	1(NYK)		
A0A0F6W4E3 _9DELT	Bacteria	Proteobacteria	Deltaproteobacteria	28221	1(SYH)		
F2J5R3_ POLGS	Bacteria	Proteobacteria	Polymorphum_gilvum	991905	2(-WH)		
A0A080KCX7 _9GAMM	Bacteria	Proteobacteria	Gammaproteobacteria	1236	2(-WH)		
A7BUJ5_ 9GAMM	Bacteria	Proteobacteria	Gammaproteobacteria	1236	2(SWH)		
W6KB90_ 9PROT	Bacteria	Proteobacteria	Proteobacteria	1224	2(SWH)		
A0A0C1ZMS9 _9DELT	Bacteria	Proteobacteria	Deltaproteobacteria	28221	2(SWH)		
L0DUM1_ THIND	Bacteria	Proteobacteria	Thioalkalivibrio_ nitratireducens	1255043	2(SWH)		

A0A0S2TDM1_9GAMM	Bacteria	Proteobacteria	Gammaproteobacteria	1236	3(NWH)		
Q1ZR24_PHOAS	Bacteria	Proteobacteria	Photobacterium_angustum	314292	3(NWH)		
A0A0D8PM30_9GAMM	Bacteria	Proteobacteria	Gammaproteobacteria	1236	3(NWH)		
C5BP80_TERTT	Bacteria	Proteobacteria	Teredinibacter_turnerae	377629	3(NWH)		
Q0VNS0_ALCBS	Bacteria	Proteobacteria	Alcanivorax_borkumensis	393595	3(NWH)		
A0A1A8TF63_9GAMM	Bacteria	Proteobacteria	Gammaproteobacteria	1236	3(NWH)		
F2JTU2_MARM1	Bacteria	Proteobacteria	Marinomonas_mediterranea	717774	3(NWH)		
A0A0F4QLW5_9GAMM	Bacteria	Proteobacteria	Gammaproteobacteria	1236	3(NWH)		
A0A0P8AS74_9GAMM	Bacteria	Proteobacteria	Gammaproteobacteria	1236	3(NWH)		
A0A085DVL3_9GAMM	Bacteria	Proteobacteria	Gammaproteobacteria	1236	3(NWH)		
A0A0D7UV92_9GAMM	Bacteria	Proteobacteria	Gammaproteobacteria	1236	3(NWH)		
A0A060B440_9GAMM	Bacteria	Proteobacteria	Gammaproteobacteria	1236	3(NWH)		
A0A101D125_9GAMM	Bacteria	Proteobacteria	Gammaproteobacteria	1236	3(NWH)		
A0A081G4E7_9GAMM	Bacteria	Proteobacteria	Gammaproteobacteria	1236	3(NWH)		
N6WYY9_9ALTE	Bacteria	Proteobacteria	Alteromonadaceae	72275	3(NWH)		
Q2BJL2_NEPCE	Bacteria	Proteobacteria	Neptuniibacter_caesariensis	207954	3(NWH)		
A0A136HG02_9GAMM	Bacteria	Proteobacteria	Gammaproteobacteria	1236	3(NWH)		

A0A191ZF22_9GAMM	Bacteria	Proteobacteria	Gammaproteobacteria	1236	3(NWH)		
A0A0S8A6W6_9GAMM	Bacteria	Proteobacteria	Gammaproteobacteria	1236	3(NWH)		
A0A1E2V8Y3_9GAMM	Bacteria	Proteobacteria	Gammaproteobacteria	1236	3(NWH)		
A0A084IHJ3_9GAMM	Bacteria	Proteobacteria	Gammaproteobacteria	1236	3(NWH)		
U2EQX8_9GAMM	Bacteria	Proteobacteria	Gammaproteobacteria	1236	3(NWH)		
A0A1E4L4P8_9PROT	Bacteria	Proteobacteria	Proteobacteria	1224	3(NWH)		
A0A0Q5MDC5_9BURK	Bacteria	Proteobacteria	Burkholderiales	80840	3(NWH)		
A1TPJ7_ACIAC	Bacteria	Proteobacteria	Acidovorax_citrulli	397945	3(NWH)		
A0A1A8T779_9GAMM	Bacteria	Proteobacteria	Gammaproteobacteria	1236	3(NWH)		
A0A0C5V1Q7_9GAMM	Bacteria	Proteobacteria	Gammaproteobacteria	1236	3(NWH)		
A0A1B1YPY0_9GAMM	Bacteria	Proteobacteria	Gammaproteobacteria	1236	3(NWH)		
G2DZS7_9GAMM	Bacteria	Proteobacteria	Gammaproteobacteria	1236	3(NWH)		
A1KBW1_AZOSB	Bacteria	Proteobacteria	Azoarcus_sp	62928	3(NWH)		
C1DFU6_AZOVD	Bacteria	Proteobacteria	Azotobacter_vinelandii	322710	3(NWH)		
I3YF20_THIV6	Bacteria	Proteobacteria	Thiocystis_violascens	765911	3(NWH)		
A0A0U1PZN7_9BURK	Bacteria	Proteobacteria	Burkholderiales	80840	3(NWH)		
A0A133XIQ8_9RHOO	Bacteria	Proteobacteria	Rhodocyclales	206389	3(NWH)		

Q47G09_ DECAR	Bacteria	Proteobacteria	Dechloromonas_aromatica	159087	3(NWH)		
G8QQ13_ DECSP	Bacteria	Proteobacteria	Dechlorosoma_suillum	640081	3(NWH)		
A0A0F7KP90_ 9SPHN	Bacteria	Proteobacteria	Sphingomonadales	204457	3(NWH)		
V4TC33_ 9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		
A8HW19_ AZOC5	Bacteria	Proteobacteria	Azorhizobium_ caulinodans	438753	3(NWH)		
A7IN18_ XANP2	Bacteria	Proteobacteria	Xanthobacter_ autotrophicus	78245	3(NWH)		
A0A0K2DKR4 _9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		
D7AAL6_ STAND	Bacteria	Proteobacteria	Starkeya_novella	639283	3(NWH)		
W3RIQ4_ 9BRAD	Bacteria	Proteobacteria	Bradyrhizobiaceae	41294	3(NWH)		
A0A0Q7TDY5 _9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		
J6LJE2_ 9RHOB	Bacteria	Proteobacteria	Rhodobacterales	204455	3(NWH)		
A0A0D1NJE3_ BRAEL	Bacteria	Proteobacteria	Bradyrhizobium_elkanii	29448	3(NWH)		
A0A120FLP4_ 9BRAD	Bacteria	Proteobacteria	Bradyrhizobiaceae	41294	3(NWH)		
F7QKR5_ 9BRAD	Bacteria	Proteobacteria	Bradyrhizobiaceae	41294	3(NWH)		
B6JAV4_ OLICO	Bacteria	Proteobacteria	Oligotropha_ carboxidovorans	504832	3(NWH)		
Q3SQ62_	Bacteria	Proteobacteria	Nitrobacter_winogradskyi	323098	3(NWH)		

NITWN							
A0A163YS83_9BRAD	Bacteria	Proteobacteria	Bradyrhizobiaceae	41294	3(NWH)		
A0A0D7ET67_RHOPL	Bacteria	Proteobacteria	Rhodopseudomonas_palustris	1076	3(NWH)		
Q217A5_RHOPB	Bacteria	Proteobacteria	Rhodopseudomonas_palustris	316056	3(NWH)		
A0A0D6JFD2_9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		
A0A109BCB0_HYPSL	Bacteria	Proteobacteria	Hyphomicrobium_sulfonivorans	121290	3(NWH)		
V5SF37_9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		
A0A0A8K4K4_9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		
A0A1E3VZ41_9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		
A0A1E3WAY3_9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		
A0A1E3VNC2_9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		
A0A1E2S159_9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		
A0A0S2ELA0_9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		
A0A175RL64_9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		
A0A0Q5GU57_9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		
A0A0B1Q3Z9_9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		
A0A0D2W186_9PROT	Bacteria	Proteobacteria	Proteobacteria	1224	3(NWH)		
W9HBV7_9PROT	Bacteria	Proteobacteria	Proteobacteria	1224	3(NWH)		
Q0G7X1_	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		

9RHIZ							
B2JSL5_ PARP8	Bacteria	Proteobacteria	Paraburkholderia_ phymatum	391038	3(NWH)		
A0A0S1Y6U5 _9BORD	Bacteria	Proteobacteria	Bordetella	517	3(NWH)		
A0A069PU02_ 9BURK	Bacteria	Proteobacteria	Burkholderiales	80840	3(NWH)		
Q13QE3_ PARXL	Bacteria	Proteobacteria	Paraburkholderia_ xenovorans	266265	3(NWH)		
A0A0Q5P8X5 _9BURK	Bacteria	Proteobacteria	Burkholderiales	80840	3(NWH)		
B5WNN4_ 9BURK	Bacteria	Proteobacteria	Burkholderiales	80840	3(NWH)		
G8MEM6_ 9BURK	Bacteria	Proteobacteria	Burkholderiales	80840	3(NWH)		
A0A0L0LZP1_ 9BURK	Bacteria	Proteobacteria	Burkholderiales	80840	3(NWH)		
A0A0K9JRN4 _9BURK	Bacteria	Proteobacteria	Burkholderiales	80840	3(NWH)		
A0A0L1L1W5 _9BURK	Bacteria	Proteobacteria	Burkholderiales	80840	3(NWH)		
A0A0K9JXR9 _9BURK	Bacteria	Proteobacteria	Burkholderiales	80840	3(NWH)		
A0A0L0M8L7 _9BURK	Bacteria	Proteobacteria	Burkholderiales	80840	3(NWH)		
G4MEL6_ 9BURK	Bacteria	Proteobacteria	Burkholderiales	80840	3(NWH)		
A0A0C1ZBA1 _9BURK	Bacteria	Proteobacteria	Burkholderiales	80840	3(NWH)		
A0A1A5X484_ 9BURK	Bacteria	Proteobacteria	Burkholderiales	80840	3(NWH)		
C5ABZ0_ BURGB	Bacteria	Proteobacteria	Burkholderia_glumae	626418	3(NWH)		
Q1YJB4_	Bacteria	Proteobacteria	Aurantimonas_	287752	3(NWH)		

AURMS			manganoxydans				
A0A0Q6D2Z4_9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		
A0A0J6UA14_9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		
A0A0Q4ZCQ3_9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		
C5AWA2_METEA	Bacteria	Proteobacteria	Methylobacterium_extorquens	272630	3(NWH)		
A0A0Q4XTJ4_9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		
A0A0Q6A2F1_9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		
B8IUS6_METNO	Bacteria	Proteobacteria	Methylobacterium_nodulans	460265	3(NWH)		
B0UI59_METS4	Bacteria	Proteobacteria	Methylobacterium_sp	426117	3(NWH)		
Q0BUX7_GRABC	Bacteria	Proteobacteria	Granulibacter_bethesdensis	391165	3(NWH)		
F6IIG8_9SPHN	Bacteria	Proteobacteria	Sphingomonadales	204457	3(NWH)		
G6EA58_9SPHN	Bacteria	Proteobacteria	Sphingomonadales	204457	3(NWH)		
H0TIW0_9BRAD	Bacteria	Proteobacteria	Bradyrhizobiaceae	41294	3(NWH)		
A4Z0G0_BRASO	Bacteria	Proteobacteria	Bradyrhizobium_sp	114615	3(NWH)		
A0A0Q6ABK6_9BRAD	Bacteria	Proteobacteria	Bradyrhizobiaceae	41294	3(NWH)		
Q89DL6_BRADU	Bacteria	Proteobacteria	Bradyrhizobium_diazoefficiens	224911	3(NWH)		

A0A0R3E4X8_9BRAD	Bacteria	Proteobacteria	Bradyrhizobiaceae	41294	3(NWH)		
U2Y363_9SPHN	Bacteria	Proteobacteria	Sphingomonadales	204457	3(NWH)		
D5RKZ4_9PROT	Bacteria	Proteobacteria	Proteobacteria	1224	3(NWH)		
H0A543_9PROT	Bacteria	Proteobacteria	Proteobacteria	1224	3(NWH)		
A0A0Q8QVI2_9SPHN	Bacteria	Proteobacteria	Sphingomonadales	204457	3(NWH)		
A0A0B1ZKY4_9SPHN	Bacteria	Proteobacteria	Sphingomonadales	204457	3(NWH)		
G8ANC8_AZOBR	Bacteria	Proteobacteria	Azospirillum_brasilense	192	3(NWH)		
A0A0Q7CZN9_9CAUL	Bacteria	Proteobacteria	Caulobacteraceae	76892	3(NWH)		
G7ZF76_AZOL4	Bacteria	Proteobacteria	Azospirillum_lipoferum	862719	3(NWH)		
Q1YLW7_AURMS	Bacteria	Proteobacteria	Aurantimonas_manganoxydans	287752	3(NWH)		
A0A0B3RUT5_9RHOB	Bacteria	Proteobacteria	Rhodobacterales	204455	3(NWH)		
A3K8B1_9RHOB	Bacteria	Proteobacteria	Rhodobacterales	204455	3(NWH)		
A0A0W7WL00_9RHOB	Bacteria	Proteobacteria	Rhodobacterales	204455	3(NWH)		
A0A1B1PSZ2_9RHOB	Bacteria	Proteobacteria	Rhodobacterales	204455	3(NWH)		
S9RDS0_9RHOB	Bacteria	Proteobacteria	Rhodobacterales	204455	3(NWH)		
Q0FGS4_PELBH	Bacteria	Proteobacteria	Pelagibaca_bermudensis	314265	3(NWH)		

W6KAE6_9PROT	Bacteria	Proteobacteria	Proteobacteria	1224	3(NWH)		
K9GMS6_9PROT	Bacteria	Proteobacteria	Proteobacteria	1224	3(NWH)		
A0A0D6P616_9PROT	Bacteria	Proteobacteria	Proteobacteria	1224	3(NWH)		
A0A0W0A5G2_9PROT	Bacteria	Proteobacteria	Proteobacteria	1224	3(NWH)		
A0A0H5BC55_BLAVI	Bacteria	Proteobacteria	Blastochloris_viridis	1079	3(NWH)		
A0A192IMR2_9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		
A0A0Q6EX43_9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		
A0A0Q6CN13_9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		
A0A0Q6DQF3_9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		
U7P0U6_9GAMM	Bacteria	Proteobacteria	Gammaproteobacteria	1236	3(NWH)		
A0A0X8HBB1_9GAMM	Bacteria	Proteobacteria	Gammaproteobacteria	1236	3(NWH)		
A0A139CBL4_9GAMM	Bacteria	Proteobacteria	Gammaproteobacteria	1236	3(NWH)		
W7QPS6_9GAMM	Bacteria	Proteobacteria	Gammaproteobacteria	1236	3(NWH)		
A0A086D693_9GAMM	Bacteria	Proteobacteria	Gammaproteobacteria	1236	3(NWH)		
A0A172YB90_9GAMM	Bacteria	Proteobacteria	Gammaproteobacteria	1236	3(NWH)		
A0A1E4C1S4_9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		
A0A101W7I8_9PROT	Bacteria	Proteobacteria	Proteobacteria	1224	3(NWH)		
A0A086MGR5_9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		

A0A1B2EFT8_9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		
B8ER95_METSB	Bacteria	Proteobacteria	Methylocella_silvestris	395965	3(NWH)		
A0A0P6VNS7_9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		
A0A0Q9JH66_9BRAD	Bacteria	Proteobacteria	Bradyrhizobiaceae	41294	3(NWH)		
A0A085EW32_9BRAD	Bacteria	Proteobacteria	Bradyrhizobiaceae	41294	3(NWH)		
A0A126NYM5_9BRAD	Bacteria	Proteobacteria	Bradyrhizobiaceae	41294	3(NWH)		
A0A0Q3KJD4_9BRAD	Bacteria	Proteobacteria	Bradyrhizobiaceae	41294	3(NWH)		
A0A117MSU7_9BRAD	Bacteria	Proteobacteria	Bradyrhizobiaceae	41294	3(NWH)		
A0A1B3NGM2_9BRAD	Bacteria	Proteobacteria	Bradyrhizobiaceae	41294	3(NWH)		
A0A0Q9HCN8_9BRAD	Bacteria	Proteobacteria	Bradyrhizobiaceae	41294	3(NWH)		
A0A0N1B5K0_9BRAD	Bacteria	Proteobacteria	Bradyrhizobiaceae	41294	3(NWH)		
A0A0Q6KBM8_9BRAD	Bacteria	Proteobacteria	Bradyrhizobiaceae	41294	3(NWH)		
A0A177PK25_9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		
A0A1A6FPJ2_9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		
J7QVB4_METSZ	Bacteria	Proteobacteria	Methylocystis_sp	187303	3(NWH)		
B2IJU2_BEII9	Bacteria	Proteobacteria	Beijerinckia_indica subsp	395963	3(NWH)		
B8ENU3_METSB	Bacteria	Proteobacteria	Methylocella_silvestris	395965	3(NWH)		
K0C859_CYCSP	Bacteria	Proteobacteria	Cycloclasticus_sp	385025	3(NWH)		

E3I0H5_ RHOVT	Bacteria	Proteobacteria	Rhodomicrobium_ vannielii	648757	3(NWH)		
W8SR89_ 9RHOB	Bacteria	Proteobacteria	Rhodobacterales	204455	3(NWH)		
A8LKX8_ DINSH	Bacteria	Proteobacteria	Dinoroseobacter_shibae	398580	3(NWH)		
A0A0D6AZS6 _RHOSU	Bacteria	Proteobacteria	Rhodovulum_ sulfidophilum	35806	3(NWH)		
A0A0N8KDG0 _9RHOB	Bacteria	Proteobacteria	Rhodobacterales	204455	3(NWH)		
Q3IZM9_ RHOS4	Bacteria	Proteobacteria	Rhodobacter_sphaeroides	272943	3(NWH)		
A0A161GJB1_ 9RHOB	Bacteria	Proteobacteria	Rhodobacterales	204455	3(NWH)		
A0A086YAB1 _9RHOB	Bacteria	Proteobacteria	Rhodobacterales	204455	3(NWH)		
A0A095WWI7 _9GAMM	Bacteria	Proteobacteria	Gammaproteobacteria	1236	3(NWH)		
W6W8W8_ 9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		
A0A095VFI2_ 9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		
A0A081MJY8 _9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		
L0NHG4_ 9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		
A0A0M3BFP6 _9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		
A0A0Q5CUX7 _9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		
A0A0Q8BJK4 _9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		
A0A081CQF9_ 9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		

A9CHT4_ AGRFC	Bacteria	Proteobacteria	Agrobacterium_fabrum	176299	3(NWH)		
B9JXJ0_ AGRVS	Bacteria	Proteobacteria	Agrobacterium_vitis	311402	3(NWH)		
K2QW79_ 9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		
A0A0Q7XI56_ 9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		
A0A1E3H442_ 9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		
G8PKB1_ PSEUV	Bacteria	Proteobacteria	Pseudovibrio_sp	911045	3(NWH)		
W6REQ1_ 9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		
K0PXY0_ 9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		
A0A0Q6WB77 _9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		
K0VJ16_ 9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		
Q2K560_ RHIEC	Bacteria	Proteobacteria	Rhizobium_etli	347834	3(NWH)		
F2AEI7_ RHIET	Bacteria	Proteobacteria	Rhizobium_etli	29449	3(NWH)		
A0A0B4X618_ 9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		
J2B2J0_9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		
A0A0Q8NKF1 _9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		
B9J859_ AGRRK	Bacteria	Proteobacteria	Agrobacterium_ radiobacter	311403	3(NWH)		

N6V9E9_9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		
A0A1B9RQ47_9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		
U4VB73_9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		
A6X526_OCHA4	Bacteria	Proteobacteria	Ochrobactrum_anthropi	439375	3(NWH)		
Q2YKI2_BRUA2	Bacteria	Proteobacteria	Brucella_abortus	359391	3(NWH)		
A0A1A9FNH1_9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		
X6FLI6_9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		
G6YCW7_9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		
M5F6F4_9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		
A0A090FS72_9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		
A0A117N4F1_RHILI	Bacteria	Proteobacteria	Rhizobium_loti	381	3(NWH)		
E8TJB6_MESCW	Bacteria	Proteobacteria	Mesorhizobium_ciceri biovar biserrulae	765698	3(NWH)		
A0A1C2DND7_9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		
A0A0Q6MK95_9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		
Q98L66_RHILO	Bacteria	Proteobacteria	Rhizobium_loti	266835	3(NWH)		
A0A0Q7WGT6_9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		
H0HST2_	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		

9RHIZ							
A0A0H1ADK4_9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		
A0A011VN94_9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		
A0A135HW59_9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		
Q11IE6_CHESB	Bacteria	Proteobacteria	Chelativorans_sp	266779	3(NWH)		
A0A0Q8BA79_9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		
B9QYS4_LABAD	Bacteria	Proteobacteria	Labrenzia_alexandrii	244592	3(NWH)		
A0A0U3ET72_9RHOB	Bacteria	Proteobacteria	Rhodobacterales	204455	3(NWH)		
A0A1E3VDN3_9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		
C3MGD7_SINFN	Bacteria	Proteobacteria	Sinorhizobium_fredii	394	3(NWH)		
A0A0Q6FMS2_9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		
A0A1C7NU50_9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		
A0A0Q7NZ70_9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		
A0A0Q5WR82_9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		
A0A0Q6SMM3_9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		
A0A1B3M7W3_9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		
A0A0D5LPH0_9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		
A0A127CFR9_9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		

A0A0F2PMI4_9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		
A0A0N0VMF1_9RHOB	Bacteria	Proteobacteria	Rhodobacterales	204455	3(NWH)		
A0A0M7A3B9_9RHOB	Bacteria	Proteobacteria	Rhodobacterales	204455	3(NWH)		
A0A0F2RGF8_9PROT	Bacteria	Proteobacteria	Proteobacteria	1224	3(NWH)		
A0A0P1H409_9RHOB	Bacteria	Proteobacteria	Rhodobacterales	204455	3(NWH)		
A4EWR3_9RHOB	Bacteria	Proteobacteria	Rhodobacterales	204455	3(NWH)		
A3X3D9_9RHOB	Bacteria	Proteobacteria	Rhodobacterales	204455	3(NWH)		
A0A0N7LS90_9RHOB	Bacteria	Proteobacteria	Rhodobacterales	204455	3(NWH)		
A0A0P7YAL4_9RHOB	Bacteria	Proteobacteria	Rhodobacterales	204455	3(NWH)		
A0A058ZLB0_9RHOB	Bacteria	Proteobacteria	Rhodobacterales	204455	3(NWH)		
A0A176FFY8_9RHOB	Bacteria	Proteobacteria	Rhodobacterales	204455	3(NWH)		
A0A0T5P6I3_9RHOB	Bacteria	Proteobacteria	Rhodobacterales	204455	3(NWH)		
B9NQB2_9RHOB	Bacteria	Proteobacteria	Rhodobacterales	204455	3(NWH)		
A0A0L6CVN9_9RHOB	Bacteria	Proteobacteria	Rhodobacterales	204455	3(NWH)		
A0A0P1GNZ1_9RHOB	Bacteria	Proteobacteria	Rhodobacterales	204455	3(NWH)		
A0A0B5DZA8_9RHOB	Bacteria	Proteobacteria	Rhodobacterales	204455	3(NWH)		
A0A0S6WWY2_9SPHN	Bacteria	Proteobacteria	Sphingomonadales	204457	3(NWH)		
A0A1E4BUG8_9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		
W4HQL7_	Bacteria	Proteobacteria	Rhodobacterales	204455	3(NWH)		

9RHOB							
Q0FJ41_ PELBH	Bacteria	Proteobacteria	Pelagibaca_bermudensis	314265	3(NWH)		
D0D788_ 9RHOB	Bacteria	Proteobacteria	Rhodobacterales	204455	3(NWH)		
X6KW59_ 9RHOB	Bacteria	Proteobacteria	Rhodobacterales	204455	3(NWH)		
A0A0T5NU08 _9RHOB	Bacteria	Proteobacteria	Rhodobacterales	204455	3(NWH)		
A9DQF6_ 9RHOB	Bacteria	Proteobacteria	Rhodobacterales	204455	3(NWH)		
X7F5U0_ 9RHOB	Bacteria	Proteobacteria	Rhodobacterales	204455	3(NWH)		
A0A073IHA6_ 9RHOB	Bacteria	Proteobacteria	Rhodobacterales	204455	3(NWH)		
A0A196NZN4 _9RHOB	Bacteria	Proteobacteria	Rhodobacterales	204455	3(NWH)		
Q160G3_ ROSDO	Bacteria	Proteobacteria	Roseobacter_denitrificans	375451	3(NWH)		
A0A0B4BF96_ 9RHOB	Bacteria	Proteobacteria	Rhodobacterales	204455	3(NWH)		
X7EMM6_ 9RHOB	Bacteria	Proteobacteria	Rhodobacterales	204455	3(NWH)		
Q28QF8_ JANSC	Bacteria	Proteobacteria	Jannaschia_sp	290400	3(NWH)		
W5Y4H9_ KOMXY	Bacteria	Proteobacteria	Komagataeibacter_xylinus	28448	3(NWH)		
F3S623_ 9PROT	Bacteria	Proteobacteria	Proteobacteria	1224	3(NWH)		
Q5FP51_	Bacteria	Proteobacteria	Gluconobacter_oxydans	290633	3(NWH)		

GLUOX							
K7TDX8_ GLUOY	Bacteria	Proteobacteria	Gluconobacter_oxydans	442	3(NWH)		
G6XH39_ 9PROT	Bacteria	Proteobacteria	Proteobacteria	1224	3(NWH)		
F7VEH9_ 9PROT	Bacteria	Proteobacteria	Proteobacteria	1224	3(NWH)		
A0A023D3E4_ ACIMT	Bacteria	Proteobacteria	Acidomonas_methanolica	437	3(NWH)		
W5YDT4_ KOMXY	Bacteria	Proteobacteria	Komagataeibacter_xylinus	28448	3(NWH)		
D5QHZ6_ KOMHA	Bacteria	Proteobacteria	Komagataeibacter_ hansenii	436	3(NWH)		
A0A060QFM4 _9PROT	Bacteria	Proteobacteria	Proteobacteria	1224	3(NWH)		
F9U5X0_ 9GAMM	Bacteria	Proteobacteria	Gammaproteobacteria	1236	3(NWH)		
H6SMU7_ RHOPH	Bacteria	Proteobacteria	Rhodospirillum_ photometricum	1084	3(NWH)		
W8KL48_ HALHR	Bacteria	Proteobacteria	Halorhodospira_ halochloris	1052	3(NWH)		
J1JZQ3_ 9RHIZ	Bacteria	Proteobacteria	Rhizobiales	356	3(NWH)		
A0A0S8A0C2_ 9GAMM	Bacteria	Proteobacteria	Gammaproteobacteria	1236	3(NWH)		
A0A0S2JEB6_ 9GAMM	Bacteria	Proteobacteria	Gammaproteobacteria	1236	3(NWH)		
D4ZKN8_ SHEVD	Bacteria	Proteobacteria	Shewanella_violacea	637905	3(NWH)		
A0A191ZG87_ 9GAMM	Bacteria	Proteobacteria	Gammaproteobacteria	1236	3(NWH)		

D0L0J1_ HALNC	Bacteria	Proteobacteria	Halothiobacillus_ neapolitanus	555778	3(NWH)		
H8YVF6_ 9GAMM	Bacteria	Proteobacteria	Gammaproteobacteria	1236	3(NWH)		
A0A1E2UYF3 _9GAMM	Bacteria	Proteobacteria	Gammaproteobacteria	1236	3(NWH)		
F9ZRI1_ ACICS	Bacteria	Proteobacteria	Acidithiobacillus_caldus	990288	3(NWH)		
A0A179BJY2_ ACIFR	Bacteria	Proteobacteria	Acidithiobacillus_ ferrooxidans	920	3(NWH)		
A0A060ULT0_ 9PROT	Bacteria	Proteobacteria	Proteobacteria	1224	3(NWH)		
B7J7X6_ ACIF2	Bacteria	Proteobacteria	Acidithiobacillus_ ferrooxidans	243159	3(NWH)		
A0A1C2I1R0_ ACITH	Bacteria	Proteobacteria	Acidithiobacillus_ thiooxidans	930	3(NWH)		
A0A1A6C5M2 _9GAMM	Bacteria	Proteobacteria	Gammaproteobacteria	1236	3(NWH)		
A7BZR6_ 9GAMM	Bacteria	Proteobacteria	Gammaproteobacteria	1236	3(NWH)		
Q0A9W5_ ALKEH	Bacteria	Proteobacteria	Alkalilimnicola_ehrlichii	187272	3(NWH)		
S2L192_ 9GAMM	Bacteria	Proteobacteria	Gammaproteobacteria	1236	3(NWH)		
E1V6S4_HAL ED	Bacteria	Proteobacteria	Halomonas_elongata	768066	3(NWH)		
A0A098RKA2 _9GAMM	Bacteria	Proteobacteria	Gammaproteobacteria	1236	3(NWH)		
A0A0F4RC52_ 9GAMM	Bacteria	Proteobacteria	Gammaproteobacteria	1236	3(NWH)		
A0A0X8HC85 _9GAMM	Bacteria	Proteobacteria	Gammaproteobacteria	1236	3(NWH)		

Q1QUM3_ CHRSD	Bacteria	Proteobacteria	Chromohalobacter_ salexigens	290398	3(NWH)		
A0A0P9C414_ 9GAMM	Bacteria	Proteobacteria	Gammaproteobacteria	1236	3(NWH)		
A0A017STS4_ 9DELT	Bacteria	Proteobacteria	Deltaproteobacteria	28221	3(NWH)		
S4Y799_ SORCE	Bacteria	Proteobacteria	Sorangium_cellulosum	56	3(NWH)		
A9GTF2_ SORC5	Bacteria	Proteobacteria	Sorangium_cellulosum	448385	3(NWH)		
A0A0S8BYZ4 _9DELT	Bacteria	Proteobacteria	Deltaproteobacteria	28221	3(NWH)		
A6GFD3_ 9DELT	Bacteria	Proteobacteria	Deltaproteobacteria	28221	3(NWH)		
A0LA75_ MAGMM	Bacteria	Proteobacteria	Magnetococcus_marinus	156889	3(NWH)		
X2HBR6_ 9GAMM	Bacteria	Proteobacteria	Gammaproteobacteria	1236	na		failed iqtree; removed from phylogenetic analysis
A7I3R4_ CAMHC	Bacteria	Proteobacteria	Campylobacter_hominis	360107	na		failed iqtree; removed from phylogenetic analysis
A0A085WJ43_ 9DELT	Bacteria	Proteobacteria	Deltaproteobacteria	28221	na		failed iqtree; removed from phylogenetic analysis
G8AFH9_ AZOBR	Bacteria	Proteobacteria	Azospirillum_brasilense	192	na		failed iqtree; removed from phylogenetic analysis
A0A0Q9EX93 _9GAMM	Bacteria	Proteobacteria	Gammaproteobacteria	1236	na		failed iqtree; removed from

							phylogenetic analysis
A0A176I893_ 9GAMM	Bacteria	Proteobacteria	Gammaproteobacteria	1236	na		failed iqtree; removed from phylogenetic analysis
A0A080KBG6 _9GAMM	Bacteria	Proteobacteria	Gammaproteobacteria	1236	na		failed iqtree; removed from phylogenetic analysis
A0A080K8U9_ 9GAMM	Bacteria	Proteobacteria	Gammaproteobacteria	1236	na		failed iqtree; removed from phylogenetic analysis
E8UEJ4_ TAYEM	Bacteria	Proteobacteria	Taylorella_equigenitalis	937774	na		failed iqtree; removed from phylogenetic analysis
R7I975_ 9BURK	Bacteria	Proteobacteria	Burkholderiales	80840	na		failed iqtree; removed from phylogenetic analysis
A0A0U2QRE8 _9ALTE	Bacteria	Proteobacteria	Alteromonadaceae	72275	na		long branch; removed from final phylogenetic analysis
A0A091AC91_ 9DELT	Bacteria	Proteobacteria	Deltaproteobacteria	28221	na		long branch; removed from final phylogenetic analysis
Q39WX1_ GEOMG	Bacteria	Proteobacteria	Geobacter_ metallireducens	269799	na		long branch; removed from final phylogenetic analysis
Q82PL7_ STRAW	Bacteria	Tenericutes	Phytoplasma_sp	59889	3(NWH)		
U4KRG7_	Bacteria	Tenericutes	Acholeplasma_palmae	1318466	na		long branch; removed from final

ACHPJ							phylogenetic analysis
A0A061A9P6_9MOLU	Bacteria	Tenericutes	Mollicutes	31969	na		long branch; removed from final phylogenetic analysis
A9NGT9_ACHLI	Bacteria	Tenericutes	Acholeplasma_laidlawii	441768	na		long branch; removed from final phylogenetic analysis
U4KPB2_9MOLU	Bacteria	Tenericutes	Mollicutes	31969	na		long branch; removed from final phylogenetic analysis

BIBLIOGRAPHY

Alberts, B., Johnson, A., Lewis, J., & al., e. (2002). Molecular Biology of the Cell. 4th Edition. New York: Garland Science.

Alonzo, F., 3rd, & Torres, V. J. (2014). The bicomponent pore-forming leucocidins of *Staphylococcus aureus*. *Microbiol Mol Biol Rev, 78*(2), 199-230. doi:10.1128/MMBR.00055-13

Bagg, A., & Neilands, J. B. (1987). Ferric uptake regulation protein acts as a repressor, emloying iron(II) as a cofactor to bing the operator of an iron transport operon in *Escherichia coli*. *Biochemistry, 26*, 5471-5477. doi:10.1021/bi00391a039,

Barker, K. D., Barkovits, K., & Wilks, A. (2012). Metabolic flux of extracellular heme uptake in *Pseudomonas aeruginosa* is driven by the iron-regulated heme oxygenase (HemO). *J Biol Chem, 287*(22), 18342-18350. doi:10.1074/jbc.M112.359265

Bartlett, M. G., Busch, K. L., Wells, C. A., & Schey, K. L. (1996). Use of 2-hydroxy-1-naphthoic acid as a matrix for matrix-assisted laser desorption/ionization mass spectrometry of low molecular weight porphyrins and peptides. *J Mass Spectrom, 31*(3), 275-279.

Bateman, A., Bimey, E., Cerruti, L., Durbin, R., Etwiller, L., Eddy, S. R., . . . Sonnhammer, E. L. L. (2002). The Pfam protein families database *Nucleic Acids Res, 30*(1), 276-280.

Bates, D., von Eiff, C., McNamara, P., Peters, G., Yeaman, M. R., Bayer, A. S., & Proctor, R. A. (2003). *Staphylococcus aureus menD* and *hemB* mutants are as infective as the parental strains, but the menadione biosynthesic mutant persists within the kidney. *J Infect Dis, 187*, 1654-1661.

Beall, B. W., & Sanden, G. N. (1995). Cloning and initial characterization of *Bordetella pertussis fur* gene. *Curr Microbiol, 30*, 223-226.

Blaby-Haas, C. E., & Merchant, S. S. (2013). Iron sparing and recycling in a compartmentalized cell. *Curr Opin Microbiol, 16*(6), 677-685. doi:10.1016/j.mib.2013.07.019

Bryan, L. E., & Kwan, S. (1983). Roles of ribosomal binding, membrane potential, and electron transport in bacterial uptake of streptomycin and gentamicin. *Antimicrob Agents Chemother, 23*(6), 835-845.

Brzezowski, P., Schlicke, H., Richter, A., Dent, R. M., Niyogi, K. K., & Grimm, B. (2014). The GUN4 protein plays a regulatory role in tetrapyrrole biosynthesis and chloroplast-to-nucleus signalling in *Chlamydomonas reinhardtii*. *Plant J, 79*(2), 285-298. doi:10.1111/tpj.12560

Chang, T., Wegmann, B., & Wang, W. (1990). Purification and characterization of glutamyl-tRNA synthetase. *Plant Physiol, 93*, 1641-1649.

Chim, N., Iniguez, A., Nguyen, T. Q., & Goulding, C. W. (2010). Unusual diheme conformation of the heme-degrading protein from *Mycobacterium tuberculosis*. *J Mol Biol, 395*(3), 595-608. doi:10.1016/j.jmb.2009.11.025

Choby, J. E., & Skaar, E. P. (2016). Heme Synthesis and Acquisition in Bacterial Pathogens. *J Mol Biol, 428*(17), 3408-3428. doi:10.1016/j.jmb.2016.03.018

Dailey, H. A., Dailey, T. A., Gerdes, S., Jahn, D., Jahn, M., O'Brian, M. R., & Warren, M. J. (2017). Prokaryotic Heme Biosynthesis: Multiple Pathways to a Common Essential Product. *Microbiol Mol Biol Rev, 81*(1). doi:10.1128/MMBR.00048-16

Dailey, H. A., Gerdes, S., Dailey, T. A., Burch, J. S., & Phillips, J. D. (2015). Noncanonical coproporphyrin-dependent bacterial heme biosynthesis pathway that does not use protoporphyrin. *Proc Natl Acad Sci U S A, 112*(7), 2210-2215. doi:10.1073/pnas.1416285112

Dammeyer, T., Bagby, S. C., Sullivan, M. B., Chisholm, S. W., & Frankenberg-Dinkel, N. (2008). Efficient phage-mediated pigment biosynthesis in oceanic cyanobacteria. *Curr Biol, 18*(6), 442-448. doi:10.1016/j.cub.2008.02.067

DeLeo, F. R., & Chambers, H. F. (2009). Reemergence of antibiotic-resistant *Staphylococcus aureus* in the genomics era. *J Clin Invest, 119*(9), 2464-2474. doi:10.1172/JCI38226

Diaz, A., Loewen, P. C., Fita, I., & Carpena, X. (2012). Thirty years of heme catalases structural biology. *Arch Biochem Biophys, 525*(2), 102-110. doi:10.1016/j.abb.2011.12.011

Duanmu, D., Casero, D., Dent, R. M., Gallaher, S., Yang, W., Rockwell, N. C., . . . Lagarias, J. C. (2012). Retrograde bilin signaling enables *Chlamydomonas* greening and phototrophic survival. *Proc Natl Acad Sci U S A, 110*(9), 3621-3626.

Duong, T., Park, K., Kim, T., Kang, S. W., Hahn, M. J., Hwang, H. Y., . . . Kim, K. K. (2014). Structural and functional characterization of an Isd-type haem-degradation enzyme from *Listeria monocytogenes*. *Acta Crystallogr D Biol Crystallogr, 70*(Pt 3), 615-626. doi:10.1107/S1399004713030794

Duthie, E. S., & Lorenz, L. L. (1952). Staphylococcal coagulase: mode of action and antigenicity. *J Gen Microbiol, 6*(1-2), 95-107.

Duthie, E. S., & Lorenz, L. L. (1952). Staphylococcal coagulase: mode of action and antigenicity. *J Gen Microbiol, 6*, 95-107.

Elbaz, A., Wei, Y. Y., Meng, Q., Zheng, Q., & Yang, Z. M. (2010). Mercury-induced oxidative stress and impact on antioxidant enzymes in *Chlamydomonas reinhardtii*. *Ecotoxicology, 19*(7), 1285-1293. doi:10.1007/s10646-010-0514-z

Farrand, A. J., Haley, K. P., Lareau, N. M., Heilbronner, S., McLean, J. A., Foster, T., & Skaar, E. P. (2015). An Iron-Regulated Autolysin Remodels the Cell Wall To Facilitate Heme Acquisition in *Staphylococcus lugdunensis*. *Infect Immun, 83*(9), 3578-3589. doi:10.1128/IAI.00397-15

Finazzi, G. (2005). The central role of the green alga *Chlamydomonas reinhardtii* in revealing the mechanism of state transitions. *J Exp Bot, 56*(411), 383-388. doi:10.1093/jxb/erh230

Finn, R. D., Clements, J., & Eddy, S. R. (2011). HMMER web server: interactive sequence similarity searching. *Nucleic Acids Res, 39*(Web Server issue), W29-37. doi:10.1093/nar/gkr367

George, P. (1948). A comparison of the decomposition of hydrogen peroxide by catalase, ferrous and ferric ions, haemin and ferrous phthalocyanine. *Biochem J, 43*, 287-295.

Glaesener, A. G., Merchant, S. S., & Blaby-Haas, C. E. (2013). Iron economy in *Chlamydomonas reinhardtii*. *Front Plant Sci, 4*, 337. doi:10.3389/fpls.2013.00337

Goldman, B. S., & Roth, J. R. (1993). Genetic structure and regulation of the *cysG* gene in *Salmonella typhimurium*. *J Bacteriol, 175*(5), 1457-1466.

Graves, A. B., Morse, R. P., Chao, A., Iniguez, A., Goulding, C. W., & Liptak, M. D. (2014). Crystallographic and spectroscopic insights into heme degradation by *Mycobacterium tuberculosis* MhuD. *Inorg Chem, 53*(12), 5931-5940. doi:10.1021/ic500033b

Green, M. R., & Sambrook, J. (2012). *Molecular Cloning. A Laboratory Manual.* Cold Spring Harbor, NY: John Inglis.

Grossman, A. R. (2000). *Chlamydomonas reinhardtii* and photosynthesis: genetics to genomics. *Curr Opin Plant Biol, 3*, 132-137.

Haley, K. P., Janson, E. M., Heilbronner, S., Foster, T. J., & Skaar, E. P. (2011). *Staphylococcus lugdunensis* IsdG liberates iron from host heme. *J Bacteriol, 193*(18), 4749-4757. doi:10.1128/JB.00436-11

Hammer, N. D., Reniere, M. L., Cassat, J. E., Zhang, Y., Hirsch, A. O., Indriati Hood, M., & Skaar, E. P. (2013). Two heme-dependent terminal oxidases power *Staphylococcus aureus* organ-specific colonization of the vertebrate host. *MBio, 4*(4). doi:10.1128/mBio.00241-13

Hirotsu, S., Chu, G. C., Unno, M., Lee, D. S., Yoshida, T., Park, S. Y., . . . Ikeda-Saito, M. (2004). The crystal structures of the ferric and ferrous forms of the heme complex of HmuO, a

heme oxygenase of *Corynebacterium diphtheriae*. *J Biol Chem, 279*(12), 11937-11947. doi:10.1074/jbc.M311631200

Hoober, J. K. (1989). *The Chlamydomonas Sourcebook. A Comprehensive Guide to Biology and Laboratory Use* (Vol. 780). San Diego, CA: Academic Press.

Howe, G., & Merchant, S. S. (1994). Role of heme in the biosynthesis of cytochrome *c*6. *J Biol Chem, 269*(8), 5824-5832.

Hurt, E., & Hauska, G. (1981). A cytochrome f/b6 complex of five polypeptides with plastoquinol-plastocyanin-oxidoreductase activity from spinich chloroplasts. *Eur J Biochem, 117*, 591-599.

Igo, R. P., Mackler, B., & Duncan, H. (1961). Liver aldehyde oxidase: The nature of hematin component. *Arch Biochem Biophys, 93*, 435-439.

Jayaraman, P. S., Peakman, T. C., Busby, S. J. W., Quincey, R. V., & Cole, J. A. (1987). Location and sequence of the promoter of the gene for the NADH-dependent nitrite reductase of *Escherichia coli* and its regulation by oxygen, the Fnr protein and nitrite. *J Mol Biol, 196*, 781-788.

Jessome, L. L., & Volmer, D. A. (2006). Ion suppression: A major concern in mass spectrometry. *LC GC N Am, 24*(5), 498-510.

Johansson, P., & Hederstedt, L. (1999). Organization of genes for tetrapyrrole biosynthesis in Gram-positive bacteria. *Microbiology, 145*, 529-538.

Jurtchuk, P. J. (1996). Bacterial Metabolism. In S. Baron (Ed.), *Medical Microbiology. 4th edition.* Galveston, TX: University of Texas Medical Branch at Gelveston.

Kahl, B., Herrmann, M., Everding, A. S., Koch, H. G., Becker, K., Harms, E., . . . Peters, G. (1998). Persistent infection with small colony variant strains of *Staphylococcus aureus* in patients with cystic fibrosis. *J Infect Dis, 177*(4), 1023-1029.

Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol, 30*(4), 772-780. doi:10.1093/molbev/mst010

Kendrew, S. G., Hopwood, D. A., & Marsh, E. N. G. (1997). Identification of a monooxygenase from *Streptomyces coelicolor* A3(2) involved in biosynthesis of actinorhodin: purification and characterization of the recombinant enzyme. *J Bacteriol, 179*(13), 4305-4310.

Klevens, R. M., Morrison, M. A., Nadle, J., Petit, S., Gershman, K., Ray, S., . . . Investigators, f. t. A. B. C. s. A. s. M. (2007). Invasive methicillin-resistant *Staphylococcus aureus* infections in the United States. *J Am Med Assoc, 298*(15), 1763-1771.

Krishna Kumar, K., Jacques, D. A., Pishchany, G., Caradoc-Davies, T., Spirig, T., Malmirchegini, G. R., . . . Gell, D. A. (2011). Structural basis for hemoglobin capture by *Staphylococcus aureus* cell-surface protein, IsdH. *J Biol Chem, 286*(44), 38439-38447. doi:10.1074/jbc.M111.287300

LaMattina, J. W., Nix, D. B., & Lanzilotta, W. N. (2016). Radical new paradigm for heme degradation in *Escherichia coli* O157:H7. *Proc Natl Acad Sci U S A, 113*(43), 12138-12143. doi:10.1073/pnas.1603209113

Lee, W. C., Reniere, M. L., Skaar, E. P., & Murphy, M. E. (2008). Ruffling of metalloporphyrins bound to IsdG and IsdI, two heme-degrading enzymes in *Staphylococcus aureus*. *J Biol Chem, 283*(45), 30957-30963. doi:10.1074/jbc.M709486200

Letunic, I., & Bork, P. (2016). Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res, 44*(W1), W242-245. doi:10.1093/nar/gkw290

Li, H., & Poulos, T. L. (1994). Structural variation in heme enzymes: a comparative analysis of peroxidase and P450 crystal structures. *Curr Biol, 15*(2), 461-464.

Liping, Z., Hongbo, S., Xiaohua, L., & Zhaopu, L. (2013). Gene regulation of iron-deficiency responses is associated with carbon monoxide and heme oxydase 1 in *Chlamydomonas reinhardtii*. *PLoS One, 8*(1), e53835. doi:10.1371/journal.pone.0053835

Litwin, C. M., Boyko, S. A., & Calderwood, S. B. (1992). Cloning, sequencing, and transcriptional regulation of the *Vibrio cholerae fur* gene. *J Bacteriol, 174*(6), 1879-1903.

Maehly, A. C. (1952). Splitting of horseradish peroxidase into prosthetic group and protein as a means of studying the linkages between hemin and protein. *Biochim Biophys Acts, 8*, 1-17.

Maines, M. D. (1997). The Heme Oxygenase System: A Regulator of Second Messenger Gases. *Annu Rev Pharmacol Toxicol, 317*, 517-554.

Mansuy, D., & Battioni, P. (2013). *Handbook of Porphyrin Science (Volume 26): With Applications to Chemistry, Physics, Material Science, Engineering, Biology and Medicine*: World Scientific.

Matsui, T., Nambu, S., Goulding, C. W., Takahashi, S., Fujii, H., & Ikeda-Saito, M. (2016). Unique coupling of mono- and dioxygenase chemistries in a single active site promotes heme degradation. *Proc Natl Acad Sci U S A, 113*(14), 3779-3784. doi:10.1073/pnas.1523333113

Matsui, T., Nambu, S., Ono, Y., Goulding, C. W., Tsumoto, K., & Ikeda-Saito, M. (2013). Heme degradation by *Staphylococcus aureus* IsdG and IsdI liberates formaldehyde rather than carbon monoxide. *Biochemistry, 52*(18), 3025-3027. doi:10.1021/bi400382p

Mazmanian, S. K., Skaar, E. P., Gaspar, A. H., Humayun, M., Gornicki, P., Jelenska, J., . . . Schneewind, O. (2003). Passage of heme-iron across the envelope of *Staphylococcus aureus*. *Science, 299*(5608), 906-909. doi:10.1126/science.1081147

Mazmanian, S. K., Ton-That, H., Su, K., & Schneewind, O. (2002). An iron-regulated sortase anchors a class of surface protein during *Staphylococcus aureus* pathogenesis. *Proc Natl Acad Sci U S A, 99*(4), 2293-2298.

McClure, R., Balasubramanian, D., Sun, Y., Bobrovskyy, M., Sumby, P., Genco, C. A., . . . Tjaden, B. (2013). Computational analysis of bacterial RNA-Seq data. *Nucleic Acids Res, 41*(14), e140. doi:10.1093/nar/gkt444

Merchant, S. S., Prochnik, S. E., Vallon, O., Harris, E. H., Karpowicz, S. J., Witman, G. B., . . . Grossman, A. R. (2007). The *Chlamydomonas* genome reveals the evolution of key animal and plant functions. *Science, 318*(5848), 245-250. doi:10.1126/science.1143609

Michel, B. W., Lippert, A. R., & Chang, C. J. (2012). A reaction-based fluorescent probe for selective imaging of carbon monoxide in living cells using a palladium-mediated carbonylation. *J Am Chem Soc, 134*(38), 15668-15671. doi:10.1021/ja307017b

Mike, L. A., Dutter, B. F., Stauff, D. L., Moore, J. L., Vitko, N. P., Aranmolate, O., . . . Skaar, E. P. (2013). Activation of heme biosynthesis by a small molecule that is toxic to fermenting *Staphylococcus aureus*. *Proc Natl Acad Sci U S A, 110*(20), 8206-8211.

Moriwaki, Y., Terada, T., Caaveiro, J. M., Takaoka, Y., Hamachi, I., Tsumoto, K., & Shimizu, K. (2013). Heme binding mechanism of structurally similar iron-regulated surface determinant near transporter domains of *Staphylococcus aureus* exhibiting different affinities for heme. *Biochemistry, 52*(49), 8866-8877. doi:10.1021/bi4008325

Morrison, M., & Stotz, E. (1955). Partition chromatography of hemins. Separation of the prosthetic groups of cytochromes a and a3. *J Biol Chem, 213*, 373-378.

Mortensen, B. L., Rathi, S., Chazin, W. J., & Skaar, E. P. (2014). *Acinetobacteri baumannii* response to host-mediated zinc limitation requires the transcriptional regulator Zur. *J Bacteriol, 196*(14), 2616-2626.

Moseley, J. L., Allinger, T., Herzog, S., Hoerth, P., Wehinger, E., Merchant, S. S., & Hippler, M. (2002). Adaptation to Fe-deficiency requires remodeling of the photosynthetic apparatus. *EMBO J, 21*(24), 6709-6720.

Musher, D. M., Baughn, R. E., & Merrell, G. L. (1979). Selection of small-colony variants of *Enterobacteriaceae* by in vitro exposure to aminoglycosides: Pathogenicity for experimental animals. *J Infect Dis, 140*(2), 209-214.

Musser, S. M., & Chan, S. I. (1998). Evolution of the Cytochrome c Oxidase Proton Pump. *J Mol Evol, 46*(5), 508-520. doi:10.1007/pl00006332

Nakajima, H., Takemura, T., Nakajima, O., Yamaoka, K. (1963). Studies on heme a-methenyl oxygenase: I. The enzymatic conversion of pyridin-hemichromogen and hemoglobin-haptoglobin into a possible precursor of biliverdin. *Journal of Biochemistry, 238*(3784-3796).

Nambu, S., Matsui, T., Goulding, C. W., Takahashi, S., & Ikeda-Saito, M. (2013). A new way to degrade heme: the *Mycobacterium tuberculosis* enzyme MhuD catalyzes heme degradation without generating CO. *J Biol Chem, 288*(14), 10101-10109. doi:10.1074/jbc.M112.448399

Neubauer, H., Pantel, I., & Gotz, F. (1999). Molecular characterization of the nitrite-reducing system of *Staphylococcus carnosus*. *J Bacteriol, 181*(5), 1481-1488.

Nguyen, L. T., Schmidt, H. A., von Haeseler, A., & Minh, B. Q. (2015). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol, 32*(1), 268-274. doi:10.1093/molbev/msu300

Nierth, A., & Marletta, M. A. (2014). Direct meso-alkynylation of metalloporphyrins through gold catalysis for hemoprotein engineering. *Angew Chem Int Ed Engl, 53*(10), 2611-2614. doi:10.1002/anie.201310145

Page, M. D., Allen, M. D., Kropat, J., Urzica, E. I., Karpowicz, S. J., Hsieh, S. I., . . . Merchant, S. S. (2012). Fe sparing and Fe recycling contribute to increased superoxide dismutase capacity in iron-starved *Chlamydomonas reinhardtii*. *Plant Cell, 24*(6), 2649-2665. doi:10.1105/tpc.112.098962

Park, S. H., Choi, S. R., & Choe, J. W. (2012). *Bacillus subtilis* HmoB is a heme oxygenase with a novel structure. *BMB Reports, 45*(4), 239-241. doi:10.5483/BMBRep.2012.45.4.239

Pfeiffer, T., Schuster, S., & Bonhoeffer, S. (2001). Cooperation and competition in the evolution of ATP-production pathways. *Science, 292*(5516), 504-507.

Purcell, E. B., & Crosson, S. (2008). Photoregulation in prokaryotes. *Curr Opin Microbiol, 11*(2), 168-178. doi:10.1016/j.mib.2008.02.014

Puri, S., & O'Brian, M. R. (2006). The *hmuQ* and *hmuD* genes from *Bradyrhizobium japonicum* encode heme-degrading enzymes. *J Bacteriol, 188*(18), 6476-6482. doi:10.1128/JB.00737-06

Ratliff, M., Zhu, W., Deshmukh, R., Wilks, A., & Stojiljkovic, I. (2001). Homologues of neisserial heme oxygenase in gram-negative bacteria: degradation of heme by the product of the *pigA*

gene of *Pseudomonas aeruginosa*. *J Bacteriol, 183*(21), 6394-6403. doi:10.1128/JB.183.21.6394-6403.2001

Reniere, M. L., Haley, K. P., & Skaar, E. P. (2011). The flexible loop of *Staphylococcus aureus* IsdG is required for its degradation in the absence of heme. *Biochemistry, 50*(31), 6730-6737. doi:10.1021/bi200999q

Reniere, M. L., & Skaar, E. P. (2008). *Staphylococcus aureus* haem oxygenases are differentially regulated by iron and haem. *Mol Microbiol, 69*(5), 1304-1315. doi:10.1111/j.1365-2958.2008.06363.x

Reniere, M. L., Ukpabi, G. N., Harry, S. R., Stec, D. F., Krull, R., Wright, D. W., . . . Skaar, E. P. (2010). The IsdG-family of haem oxygenases degrades haem to a novel chromophore. *Mol Microbiol, 75*(6), 1529-1538. doi:10.1111/j.1365-2958.2010.07076.x

Rochaix, J. D. (1995). *Chlamydomonas reinhardtii* as the photosynthetic yeast. *Annu Rev Genet, 29*, 209-230.

Rohlhill, J., Sandoval, N. R., & Papoutsakis, E. T. (2017). Sort-Seq Approach to Engineering a Formaldehyde-Inducible Promoter for Dynamically Regulated *Escherichia coli* Growth on Methanol. *ACS Synth Biol, 6*(8), 1584-1595. doi:10.1021/acssynbio.7b00114

Schafer, G., Purschke, W., & Schmidt, C. L. (1996). On the origin of respiration: electron transport proteins from archaea to man. *FEMS Microbiol Rev, 18*, 173-188.

Schmitt, M. P. (1997). Transciption of the *Corynebacterium diphtheriae hmuO* gene is regulated by iron and heme. *Infect Immun, 65*(11), 4634-4641.

Schroder, I., Johansson, P., Rutberg, L., & Hederstedt, L. (1994). The *hemX* gene of the *Bacillus subtilis hemAXCDBL* operon encodes a membrane protein, negatively affecting the steady-state cellular concentration of HemA (glutamyl-tRNA reductase). *Microbiology, 140*, 731-741.

Schuller, D. J., Wilks, A., Ortiz de Montellano, P. R., & Poulos, T. L. (1999). Crystal structure of human heme oxygenase-1. *Nat Struct Biol, 6*(9), 860-867.

Sciara, G., Kendrew, S. G., Miele, A. E., Marsh, N. G., Federici, L., Malatesta, F., . . . Vallone, B. (2003). The structure of ActVA-Orf6, a novel type of monooxygenase involved in actinorhodin biosynthesis. *EMBO J, 22*(2), 205-215.

Selvan, A. T., & Sankaran, K. (2008). Localization and characterization of prolipoprotein diacylglyceryl transferase (Lgt) critical in bacterial lipoprotein biosynthesis. *Biochimie, 90*(11-12), 1647-1655. doi:10.1016/j.biochi.2008.06.005

Skaar, E. P., Gaspar, A. H., & Schneewind, O. (2004). IsdG and IsdI, heme-degrading enzymes in the cytoplasm of *Staphylococcus aureus*. *J Biol Chem, 279*(1), 436-443. doi:10.1074/jbc.M307952200

Skaar, E. P., Gaspar, A. H., & Schneewind, O. (2006). *Bacillus anthracis* IsdG, a heme-degrading monooxygenase. *J Bacteriol, 188*(3), 1071-1080. doi:10.1128/JB.188.3.1071-1080.2006

Skaar, E. P., Humayun, M., Bae, T., DeBord, K. L., & Schneewind, O. (2004). Iron-source preference of *Staphylococcus aureus* infections. *Science, 305*, 1626-1628.

Smith, A. G., Marsh, O., & Elder, G. H. (1993). Investigation of the subcellular location of the tetrapyrrole-biosynthesis enzyme coproporphyrinogen oxidase in high plants. *Biochem J, 292*(2), 503-508.

Streit, B. R., Kant, R., Tokmina-Lukaszewska, M., Celis, A. I., Machovina, M. M., Skaar, E. P., . . . DuBois, J. L. (2016). Time-resolved Studies of IsdG Protein Identify Molecular Signposts along the Non-canonical Heme Oxygenase Pathway. *J Biol Chem, 291*(2), 862-871. doi:10.1074/jbc.M115.666560

Stryer, L. (1961). A conformation-dependent Cotton effect in the Soret band of hemin:poly-L-lysine. *Biochim Biophys Acta, 54*, 395-397.

Sugishima, M., Migita, C. T., Zhang, X., Yoshida, T., & Fukuyama, K. (2004). Crystal structure of heme oxygenase-1 from cyanobacterium *Synechocystis* sp. PCC 6803 in complex with heme. *Eur J Biochem, 271*(22), 4517-4525. doi:10.1111/j.1432-1033.2004.04411.x

Sugishima, M., Sakamoto, H., Higashimoto, Y., Noguchi, M., & Fukuyama, K. (2003). Crystal structure of rat heme oxygenase-1 in complex with biliverdin-iron chelate. Conformational change of the distal helix during the heme cleavage reaction. *J Biol Chem, 278*(34), 32352-32358. doi:10.1074/jbc.M303682200

Surdel, M. C., Horvath, D. J., Jr., Lojek, L. J., Fullen, A. R., Simpson, J., Dutter, B. F., . . . Skaar, E. P. (2017). Antibacterial photosensitization through activation of coproporphyrinogen oxidase. *Proc Natl Acad Sci U S A, 114*(32), E6652-E6659. doi:10.1073/pnas.1700469114

Tanui, C. K., Shyntum, D. Y., Priem, S. L., Theron, J., & Moleleki, L. N. (2017). Influence of the ferric uptake regulator (Fur) protein on pathogenicity in *Pectobacterium carotovorum subsp. brasiliense*. *PLoS One, 12*(5), e0177647. doi:10.1371/journal.pone.0177647

Tarahi Tabrizi, S., Sawicki, A., Zhou, S., Luo, M., & Willows, R. D. (2016). GUN4-Protoporphyrin IX Is a Singlet Oxygen Generator with Consequences for Plastid Retrograde Signaling. *J Biol Chem, 291*(17), 8978-8984. doi:10.1074/jbc.C116.719989

Tardif, M., Atteia, A., Specht, M., Cogne, G., Rolland, N., Brugiere, S., . . . Cournac, L. (2012). PredAlgo: a new subcellular localization prediction tool dedicated to green algae. *Mol Biol Evol, 29*(12), 3625-3639. doi:10.1093/molbev/mss178

Tenhunen, R., Marver, H. S., & Schmid, R. (1968). The enzymatic conversion of heme to bilirubin by microsomal heme oxygenase *Biochem J, 61*, 748-755.

Terauchi, A. M., Peers, G., Kobayashi, M. C., Niyogi, K. K., & Merchant, S. S. (2010). Trophic status of *Chlamydomonas reinhardtii* influences the impact of iron deficiency on photosynthesis. *Photosynth Res, 105*(1), 39-49. doi:10.1007/s11120-010-9562-8

Tien, W., & White, D. C. (1968). Linear sequential arrangment of genes for the biosynthetic pathway of protoheme in *Staphylococcus aureus*. *Biochem J, 61*, 1392-1398.

Torres, V. J., Attia, A. S., Mason, W. J., Hood, M. I., Corbin, B. D., Beasley, F. C., . . . Skaar, E. P. (2010). *Staphylococcus aureus* fur regulates the expression of virulence factors that contribute to the pathogenesis of pneumonia. *Infect Immun, 78*(4), 1618-1628. doi:10.1128/IAI.01423-09

Torres, V. J., Pishchany, G., Humayun, M., Schneewind, O., & Skaar, E. P. (2006). *Staphylococcus aureus* IsdB is a hemoglobin receptor required for heme iron utilization. *J Bacteriol, 188*(24), 8421-8429. doi:10.1128/JB.01335-06

Torres, V. J., Stauff, D. L., Pishchany, G., Bezbradica, J. S., Gordy, L. E., Iturregui, J., . . . Skaar, E. P. (2007). A *Staphylococcus aureus* regulatory system that responds to host heme and modulated virulence. *Cell Host Microbe, 1*(2), 109-119.

Trakshel, G. M., Kutty, R. K., & Maines, M. D. (1986). Purification and characterization of the major constitutive form of testicular heme oxygenase. The noninducible isoform. *J Biol Chem, 261*(24), 11131-11137.

Tsolis, R. M., Baumler, A. J., Stojiljkovic, I., & Heffron, F. (1995). Fur regulon of *Salmonella typhimurium*: identification of new iron regulated genes. *J Bacteriol, 177*(16), 4628-4637.

Urzica, E. I., Casero, D., Yamasaki, H., Hsieh, S. I., Adler, L. N., Karpowicz, S. J., . . . Merchant, S. S. (2012). Systems and trans-system level analysis identifies conserved iron deficiency responses in the plant lineage. *Plant Cell, 24*(10), 3921-3948. doi:10.1105/tpc.112.102491

van Lis, R., Atteia, A., Nogaj, L. A., & Beale, S. I. (2005). Subcellular localization and light-regulated expression of protoporphyrinogen IX oxidase and ferrochelatase in *Chlamydomonas reinhardtii*. *Plant Physiol, 139*(4), 1946-1958. doi:10.1104/pp.105.069732

von Eiff, C., Heilmann, C., Proctor, R. A., Woltz, C., Peters, G., & Gotz, F. (1997). A Site-Directed *Staphylococcus aureus hemB* Mutant is a Small-Colony Variant which Persists Intracellularly. *J Bacteriol, 179*(15), 4706-4712.

von Eiff, C., McNamara, P., Becker, K., Bates, D., Lei, X. H., Ziman, M., . . . Proctor, R. A. (2006). Phenotype microarray profiling of *Staphylococcus aureus menD* and *hemB* mutants with the small-colony-variant phenotype. *J Bacteriol, 188*(2), 687-693. doi:10.1128/JB.188.2.687-693.2006

von Gromoff, E. D., Alawady, A., Meinecke, L., Grimm, B., & Beck, C. F. (2008). Heme, a plastid-derived regulator of nuclear gene expression in *Chlamydomonas*. *Plant Cell, 20*(3), 552-567. doi:10.1105/tpc.107.054650

Voss, B., Meinecke, L., Kurz, T., Al-Babili, S., Beck, C. F., & Hess, W. R. (2011). Hemin and magnesium-protoporphyrin IX induce global changes in gene expression in *Chlamydomonas reinhardtii*. *Plant Physiol, 155*(2), 892-905. doi:10.1104/pp.110.158683

Watanabe, T., & Horie, S. (1976). Studies on P-450. X. On the coordination structure of hemoprotein P-450. *J Biochem, 79*, 829-840.

Wilks, A. (2002). Heme Oxygenase: Evolution, Structure, and Mechanism. *Antioxid Redox Signal, 4*(4), 603-614.

Wilks, A., & Heinzl, G. (2014). Heme oxygenation and the widening paradigm of heme degradation. *Arch Biochem Biophys, 544*, 87-95. doi:10.1016/j.abb.2013.10.013

Wilks, A., Ortiz de Montellano, P. R., Sun, J., & Loehr, T. M. (1996). Heme oxygenase (HO-1): His-132 stabilizes a distal water ligand and assists catalysis. *Biochemistry, 35*, 930-936.

Wilks, A., & Schmitt, M. P. (1998). Expression and Characterization of a Heme Oxygenase (HmuO) from *Corynebacterium diptheriae*: Iron Acquisition Requires Oxidative Cleavage of the Heme Macrocycle. *J Biol Chem, 273*(2), 837-841.

Wilks, A., Torpey, J., & Ortiz de Montellano, P. R. (1994). Heme oxygenase (HO-1). Evidence for electrophilic oxygen addition to the porphyrin ring in the formation of a-meso-hydroxyheme. *J Biol Chem, 269*(47), 29553-29556.

Wu, R., Skaar, E. P., Zhang, R., Joachimiak, G., Gornicki, P., Schneewind, O., & Joachimiak, A. (2005). *Staphylococcus aureus* IsdG and IsdI, heme-degrading enzymes with structural similarity to monooxygenases. *J Biol Chem, 280*(4), 2840-2846. doi:10.1074/jbc.M409526200

Xiong, A., Singh, V. S., Cabrera, G., & Jayaswal, R. K. (2000). Molecular Characterization of the Ferric-Uptake Regulator, Fur, from *Staphylococcus aureus*. *Microbiology, 146*, 659-668.

Zhu, H., Xie, G., Liu, M., Olson, J. S., Fabian, M., Dooley, D. M., & Lei, B. (2008). Pathway for heme uptake from human methemoglobin by the iron-regulated surface *determinants system of Staphylococcus aureus*. *J Biol Chem, 283*(26), 18450-18460. doi:10.1074/jbc.M801466200

Zhu, W., Wilks, A., & Stojiljkovic, I. (2000). Degradation of heme in Gram-negative bacteria: the product of the *hemO* gene of *Neisseriae* is a heme oxygenase. *J Bacteriol, 182*(23), 6783-6790.