

TERMINAL STEPS OF HEME BIOSYNTHESIS IN PROKARYOTES:
IDENTIFICATION AND CHARACTERIZATION OF NOVEL ENZYMES

by

TYE O'HARA BOYNTON

(Under the Direction of Harry A. Dailey)

ABSTRACT

Heme is an intrinsic and indispensable molecule in most forms of life, performing a myriad of functions such as gas transport, electron transport, and cellular signaling. In eukaryotes, heme is synthesized by a tightly regulated pathway of eight enzymes that is both well understood and fully characterized. In prokaryotes, however, there is a notable absence of identifiable biosynthetic enzymes in the terminal portion of the pathway. Recognized forms of protoporphyrinogen oxidase (PPO, HemY), the penultimate step in heme biosynthesis, are lacking in 50% of heme synthesizing genomes to date. An identifiable form of coproporphyrinogen oxidase (CPO), the antepenultimate step, is likewise missing in a subset of bacteria. Herein we describe the presence and characterization of three variant enzymes to fill these steps. The first, HemG, was shown to be a flavodoxin-like enzyme that has evolved PPO activity via a long chain insert loop specific to this function. Through the use of quinones, HemG functions during both aerobic and anaerobic respiration and is found in the enterics. Using comparative genomics, we have further identified another variant PPO, HemJ, which was proven to contain catalytic function through characterization and

complementation of a $\Delta hemJ$ *Acinetobacter baylyi* ADP1 mutant. Finally, we have proven the existence of novel CPO present in the Firmicutes that requires a cobalt cofactor and utilizes oxygen. This is distinct from the two currently known forms, HemF and HemN, which use only oxygen or a radical-SAM mechanism respectively. In addition to these variant enzymes, we show evidence of novel behavior for HemY in the Firmicutes and Actinobacteria, which requires the presence of the following pathway enzyme HemH and a newly identified enzyme HemQ to function correctly. This data indicates a novel mechanism of heme formation, despite overall HemY sequence homology with eukaryotes. It is intriguing that, while highly conserved elsewhere, these steps exhibit this degree of diversity in a pathway where the final intermediates are highly toxic. With our findings, enzymes with protoporphyrinogen oxidase activity can now be asserted in all heme-synthesizing bacteria. With regards to CPO activity, the Firmicutes have been satisfied, leaving the only Actinobacteria without a recognized form.

INDEX WORDS: Heme, Prokaryotic Heme Biosynthesis, Protoporphyrinogen oxidase, Coproporphyrinogen oxidase

TERMINAL STEPS OF HEME BIOSYNTHESIS IN PROKARYOTES:
IDENTIFICATION AND CHARACTERIZATION OF NOVEL ENZYMES

by

TYE O'HARA BOYNTON

B. S., The University of Georgia, 2002

M.S., The University of Georgia, 2005

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial
Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2011

© 2011

Tye O'hara Boynton

All Rights Reserved

TERMINAL STEPS OF HEME BIOSYNTHESIS IN PROKARYOTES:
IDENTIFICATION AND CHARACTERIZATION OF NOVEL ENZYMES

by

TYE O'HARA BOYNTON

Major Professor: Harry A. Dailey

Committee: William B. Whitman
Jan Mrázek
William N. Lanzilotta

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
May 2011

DEDICATION

This dissertation is dedicated to my fiancé, Jillian, my entire family, and most especially my mentor, Harry Dailey. Without them, I could not have hoped to achieve the enormous task this dissertation encompassed.

“Whether our efforts are, or not, favored by life, let us be able to say, when we come near the great goal, "I have done what I could.” – Louis Pasteur

ACKNOWLEDGEMENTS

I would like to acknowledge in no particular order the following people for the help or encouragement they provided during my current academic career:

Harry A. Dailey and my entire committee, Tamara A. Dailey, Svetlana Gerdes, Ellen Neidle, Sarah Craven, John D. Phillips, Amy E. Medlock, Wendy Dustman, Joseph Burch, Lizzie May, Lauren Daugherty, Aisha Mahmood, Daniel Thames, and Colton Wood.

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS.....	v
CHAPTER	
1 INTRODUCTION AND LITERATURE REVIEW	1
References.....	42
Figures	71
2 IDENTIFICATION OF <i>E. COLI</i> HEMG AS A NOVEL, MENADIONE- DEPENDENT FLAVODOXIN WITH PROTOPORPHYRINOGEN OXIDASE ACTIVITY	83
References.....	100
Figures	106
Tables	112
3 REQUIREMENT OF HEMQ FOR PROPER FUNCTION OF HEMY AND HEMH IN THE FIRMICUTES AND ACTINOBACTERIA.....	113
References.....	123
Figures	126
Tables	129
4 DISCOVERY OF A GENE INVOLVED IN A THIRD BACTERIAL PROTOPORPHYRINOGEN OXIDASE ACTIVITY THROUGH	

COMPARATIVE GENOMIC ANALYSIS AND FUNCTIONAL COMPLEMENTATION	130
References.....	150
Figures	157
Tables	163
5 DISCOVERY AND CHARACTERIZATION OF A COBALT- DEPENDENT, OXYGEN UTILIZING BACTERIAL COPROPORPHYRINOGEN DECARBOXYLASE	166
References.....	188
Figures	195
Tables	202
6 CONCLUSIONS	208
References.....	216

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

Importance of Tetrapyrroles and Heme

At its most basic molecular form, life can be defined as a system of biochemical reactions out of equilibrium with its environment. While this "bag of enzymes" theory is overly simplistic, it does highlight the importance of such cellular processes as genetic replication, photosynthesis, respiration, cell communication, signal transduction, and response to external stimuli.

Porphyryns, one of the oldest families of cofactors known, are intimately entangled in all of these processes. The chemical nature of these molecules allows this, enabling a wide variety of functionality based on the exact metal ion chelated into the center and the side chains present. Dubbed the "pigments of life", porphyryns have been attributed to the origin of life as it currently exists. Without the influx of free solar energy generated by photosynthesis, complex life would never have formed (115).

The various types of biological porphyryns are shown in figure 1.1. The basic backbone of all porphyryns enable their special properties. Their structure consists of a macrocycle of four pyrrole rings connected by methylene bridges. These bridges generate a stable and aromatic system with the nitrogen atom of each pyrrole ring facing center, allowing chelation of octahedral metals on the

plane of their octahedron and leaving axial positions open as electron sources or sinks. These positions are the crux of all porphyrin reactions, through the binding of various ligands. For example, the transition metal iron can function as a cofactor in biological processes due to its ability to cycle between redox states. When iron is inserted into porphyrin, the catalytic capacity increases as great as 10 orders of magnitude (199).

The most prevalent tetrapyrroles in nature are heme and chlorophyll, which contain iron and magnesium respectively. The most versatile of these is heme, which affects the majority of cellular processes previously mentioned through distinct, separate mechanisms as a cofactor of various hemoproteins. Among hemoproteins there is no common fold devoted to heme binding. It is the protein environment that gives the heme moiety its exact mode of action (125). Various forms of heme exist (figure 1.2), but each is synthesized through modification of the principle form heme B, an iron-containing porphyrin that bears vinyl groups attached to the A and B rings and propionate groups attached to the C and D rings. Heme A and heme O both contain isoprenoid groups attached to the A ring vinyl moiety, with a formyl-substitution at ring D. Both are cofactors of their respective cytochrome oxidases (102, 155). Heme C is covalently attached rather than merely bound to an apoprotein. This is accomplished through the formation of thioether bonds at the position of vinyl groups to a specific amino acid sequence, usually CXXCH. The two cysteines form the covalent attachment and the histidine serves as an axial ligand to the iron (198). The most extreme modification is found in heme D, containing a fifth ring, a lactone, off of ring C.

Heme D is important in terminal respiratory oxidases and some catalases (207). Other less common types include heme L in lactoperoxidase (158), heme M in myeloperoxidase (39), and heme S found in the hemoglobin of marine worms.

Heme has long been known for several functions. The first is as a diatomic gas transporter in hemoglobin and myoglobin, where it binds oxygen through an axial ligand of iron (145). Cytochromes bind various types of heme to transfer single electrons throughout electron transport chains (72). Finally, heme has been well-studied as a cofactor in various oxidases, the most familiar being catalases and peroxidases (221), which catalyze the breakdown of hydrogen peroxide, as well as cytochrome p450s that activate molecular oxygen and oxidize a wide variety of organic substrates including steroid and lipid biosynthesis (53).

More recently, research has revealed that heme is much more widespread in nature's design. In addition to diatomic gas transport, the same properties allow hemoproteins to act as gas sensors, enabling catalytic activity in the presence or absence of a specific substrate (164). The electron transfer capability of heme within cytochromes also gives heme the ability to detect changes in cellular redox potential, which has given rise to many cytoprotective hemoproteins that are switched on or off during variations in environmental conditions. *E. coli* switches between aerobic and anaerobic respiration through one such protein, *Ec* DOS (90). Still other hemoproteins contain a heme regulatory motif (HRM) and are able to affect DNA transcription (47) and have even been found to exert control over circadian rhythm, integral to coordinating

biological processes (159). Heme has further been found to control processing of microRNAs, which regulate gene expression (37). In a wide variety of prokaryotes including *Staphylococcus*, *Bacillus*, *Haemophilus* and *Pseudomonas* species, hemoproteins have been shown to act as virulence switches that sense the presence or absence of heme in addition to many of the afore mentioned functions (52, 122, 140).

Heme biosynthesis in Prokaryotes

The ubiquity of porphyrins extends to virtually all prokaryotic organisms. Among these microorganisms can be found all known types of organic tetrapyrroles, indicating their importance in the ancient lineage of life. Heme biosynthesis occurred very early in evolution, being present in *Aquifex* species, the deepest branching of the bacteria. The majority of organisms that require heme but lack the ability to produce it, such as *Haemophilus* and *Streptococcus* species, have compensated by having heme uptake systems. These are present in many pathogenic microorganisms (119, 162).

Heme is synthesized through an eight step process (scheme 1.1). Outside of the α -proteobacteria, the first four steps are common to all tetrapyrrole synthetic pathways. The next three are further shared in chlorophyll biosynthesis, where heme and chlorophyll deviate based on metal chelation. Synthesis starts with the formation of the precursor compound d-aminolevulinic acid (ALA). Two ALAs are condensed to form the monopyrrole porphobilinogen (PBG). Four PBG molecules are required for the formation of

hydroxymethylbilane, an unstable linear tetrapyrrole that is quickly cyclized. During this cyclization, the D-ring of the porphyrin is also enzymatically inverted. The resulting uroporphyrinogen III is decarboxylated to form coproporphyrinogen III, which is then oxidatively decarboxylated to form protoporphyrinogen IX. Protoporphyrinogen IX is reduced to form the fully aromatic protoporphyrin IX and finally iron is enzymatically inserted to produce heme. In the following pages, the mechanisms and enzymes involved in the process will be reviewed in more detail.

Synthesis of δ -Aminolevulinic Acid

The first committed step to tetrapyrrole biosynthesis is the formation of the precursory compound δ -aminolevulinic acid (ALA). Two distinct systems exist for synthesizing ALA, the C4 (classically known as the Shemin pathway) and the C5 pathway, named in regards to the initial carbon skeleton of succinate and glutamate, respectively. In higher organisms, the former is found in animals and fungi, while plants and algae contain the latter. In prokaryotes, the C4 pathway is found exclusive to the α -proteobacteria, a biologically diverse group of organisms including photosynthesizers and endosymbionts. Here, succinyl-CoA and glycine are condensed via the enzyme ALA synthase (HemA) to form ALA with the release of carbon dioxide. The C5 pathway, found in archaea and all remaining bacteria, employs two enzymes, glutamyl-tRNA reductase (GTR) and glutamate semialdehyde aminotransferase (HemL) to convert glutamate in the form of glutamyl-tRNA into ALA. The enzyme glutamyl-tRNA synthetase is

required for this, but is not a committed step and is thus considered separate from heme biosynthesis.

The C4 Pathway

Early characterization of the C4 pathway occurred in the late 1940's by following the route of [¹⁵N]glycine in the human body. Following ingestion of labeled glycine, David Shemin took his own blood samples and traced the atoms to blood porphyrins (157, 185). Further studies with labeled carbon proved that glycine was responsible for eight of the carbons within the macrocycle with the rest coming from acetate. This activity was eventually linked to the tricarboxylic acid cycle, and Shemin went on to postulate the presence of a succinyl coenzyme compound required during this catalysis, years before succinyl-CoA was actually discovered (136, 184). Finally, Shemin found that α -aminolevulinic acid could effectively replace glycine and succinate, and was responsible for all atoms within the final tetrapyrrole (186). A reaction mechanism was proposed (figure 1.3A), and later confirmed, whereby glycine forms a linkage with the cofactor pyridoxal-5'-phosphate (PLP) and the resulting radical condenses with succinyl-CoA, carbon dioxide and the coenzyme are released, and finally a proton replaces the carboxyl group lost and ALA is released (86, 132). A single lysine residue in murine ALAS has been proposed to help form the quinonoid reaction intermediate necessary for catalysis to proceed (61).

After being mapped via a mutation, purification of the enzyme confirmed PLP as a cofactor (127), and through pre-steady kinetics it was determined that the rate-limiting step during catalysis was product release (128). This led to the

notion that conformational changes occur within ALAS during substrate binding and product release (224). This was confirmed in 2005 when crystal structures of ALAS were obtained from *Rhodobacter sphaeroides* with both glycine and succinyl-CoA substrates bound respectively (6) (figure 1.3B). ALAS was shown to be a homodimer of a three-domain enzyme, with the active site and PLP found in the center domain, deeply buried from the surface. The monomer was seen to adopt both an open and closed formation, whereby the open conformation allows for binding of substrates, but residues are not properly aligned for catalysis. Functionally, ALAS belongs to a subfamily of PLP-dependent enzymes that catalyze the decarboxylative condensation of a carboxyl acid CoA molecule with an amino acid (188, 204).

Overexpression of HemA or addition of exogenous ALA leads to an upregulation of tetrapyrrole biosynthesis. This same overexpression of HemA in organisms normally containing the C5 pathway also exhibit this same increase due to the increased internal ALA levels (104, 152). Exogenous heme, on the other hand, initiates feedback inhibition at the point of ALAS within prokaryotes (215). Not surprisingly, regulation of HemA has been extensively studied in an effort to understand how cells regulate heme synthesis in general. In *Rhodobacter sphaeroides*, ALAS exists as two separate isozymes, HemA and HemT. HemT is not expressed under aerobic or photosynthetic growth, but the encoded enzyme is fully functional (135). *hemA* is transcribed in response to changes in oxygen tension, and is attributed to internal redox levels created by the changes in oxygen rather than the molecule itself. This is mediated by the

global regulatory proteins PrrA and FnrL (36, 161). Regulation in this system has not been fully elucidated, however, as heme and vitamin B12 levels must be maintained despite changes in bacterial chlorophyll production (222). In *Rhodobacter capsulatus*, feedback inhibition of heme has been attributed to the heme-binding regulatory protein HbrL that activates expression in the absence of bound heme (194).

The C5 Pathway

The so-called “alternate route” of ALA synthesis, the C5 pathway, was first discovered when researchers found that radioactively-labeled succinate and glycine did not lead to labeled porphyrins in plant extracts (9). Instead, the atoms were seen to originate predominantly from glutamate (10). Though not understood, it was assumed that this pathway was utilized for plastid porphyrins only, until the determination that glutamate led to both plastid and mitochondrial porphyrins. The same was found to be true in algae and eventually the majority of prokaryotic organisms (163, 211). Further research revealed that glutamate must be present in the form of glutamyl-tRNA and that multiple enzymes fractions were required (212, 213). These findings resulted in the identification of the two enzymes previously mentioned: Glutamyl-tRNA reductase, and glutamate-1-semialdehyde (GSA) aminotransferase. These enzymes catalyze the formation of ALA from glutamyl tRNA (figure 1.4A).

GTR is a NADPH-dependant enzyme that cleaves the nucleic acid backbone from the glutamyl-tRNA leaving GSA (70). Only the tRNA bearing the UUC anticodon is recognizable by the enzyme, and a single cysteine and

histidine are thought to be solely responsible for catalysis; a substitution in either abolishes activity (123, 177). The current accepted mechanism involves a nucleophilic attack by the cysteine on the α -carbon of glutamate, releasing the tRNA. Hydride transfer from NADPH, facilitated by the histidine, causes release of the newly created semialdehyde (108). GTR is also known to function as an esterase, whereby water is used to instead generate glutamate as a final product, though this is most likely an *in vitro* artifact and not an actual physiological function (176).

Structurally, GTR is a V-shaped dimer with subunits containing an NADPH-binding, catalytic, and dimerization domain (figure 1.4B). The catalytic domain itself has two separate functions, a tRNA tertiary core and a glutamate recognition site that properly align the substrate for the cleaving the tRNA and hydride transfer respectively (124, 160). The tertiary structure also bears a large void between catalytic sites that was proposed to contain its partner HemL *in vivo*. Studies using co-immunoprecipitation and gel chromatography later confirmed this tightly-bound complex, most likely developed as a means of protecting the highly reactive GSA intermediate through substrate channeling . *In vitro*, complex formation also prevents esterase activity of GTR (107).

HemL is the enzyme responsible for converting GSA into ALA in the C5 pathway and is sometimes referred to as GSA aminomutase or aminotransferase. During catalysis, an amine group is translocated from 2-carbon to the 1-carbon position of GSA via a cofactor of the enzyme, PLP. Because of the early observations that both PLP and pyridoxamine-5'-phosphate

(PMP) can both bind as cofactors, two mechanisms have been proposed for the reaction. In the first, PLP transfers the amine of GSA onto itself, generating PMP and dioxoalate (DOVA). PMP is then regenerated to PLP by transferring the amine back to GSA leaving ALA (206). The opposite occurs in the second mechanism, whereby PMP transfers its amine group to the 1-carbon position leaving PLP and diaminoalate (DAVA). PLP then regenerates to PMP by removing the 2-carbon amine leaving ALA (182). Studies using both of these intermediates as substrates as well as both cofactors showed that HemL is capable of either mechanism, though the preferred route is to use PLP as an initial cofactor (21). In this conformation a salt bridge ensures that PLP is always protonated, increasing the ability to accept electrons.

Despite sharing only 14% sequence homology, the overall tertiary structure of HemL is very similar to HemA (figure 1.4C). Within the family of aminotransferases, they are structurally the closest neighbors. Most key differences between the two are restricted to the substrate recognition residues within the active sites, and the overall similarity has led to the hypothesis that HemA evolved from HemL (182). The appearance of the citric acid cycle and consequently succinyl CoA is relatively young with respect to evolution, and the tRNA requirement of GTR may be a remnant of the theoretical ancient RNA world. Heme biosynthesis eventually evolved to be uncoupled from protein synthesis and instead be linked to metabolism. According to the theory of endosymbiosis, mitochondria arose from an ancient α -proteobacterium, which would have contained this newly created system (141). The gene may have

been passed to the nuclear genome, giving rise to the C4 pathway in animals and fungi. Acquisition of ALA synthase has been observed in certain species of *Streptomyces*, where both systems are present and functional. Here, the C5 pathway is utilized for heme production, while ALA from HemA is exclusively used in product of the antibiotic Asukamycin (146).

ALA to Porphobilinogen

The next step in the synthesis of all tetrapyrroles is the formation of the initial monopyrrole, porphobilinogen (PBG) through the enzymatic condensation of two molecules of ALA. This condensation is uniquely asymmetric, requiring precise binding and orientation of the separate acids to complete the reaction (figure 1.5A). In literature the enzyme responsible is referred to as ALA dehydratase (ALAD), PBG synthase (PBGS), and HemB, the latter being ascribed to the enzyme in prokaryotes. HemB is a well-conserved enzyme in all forms of life, even the most phylogenetically distant organisms share 35% sequence similarity. Despite this conformity, HemB is subdivided into distinct groups based on divalent cation binding.

Early experiments using PBG as a precursor to heme led to an assumption that the molecule was the fundamental building block to tetrapyrrole synthesis (48). Once the Shemin pathway was proposed, it became clear that two molecules of ALA could be used to synthesize PBG with the loss of two molecules of water, thus the original name of “dehydratase”. This activity was found linked to a single purified enzyme (129). In these early studies, enzymatic

activity of *Rhodobacter sphaeroides* was stimulated by the addition of monovalent cations (131). The actual function of these ions has yet to be attributed and is not explained by current structural models. An initial reaction mechanism was proposed whereby the first molecule of ALA formed a Schiff base with the enzyme, a carbon-carbon bond was formed with the second, and the original Schiff base was replaced by the second ALA's amine group. This was not unlike reactions proposed of known aldolases. The attachment via a Schiff base was proven by the reduction of the enzyme in the presence of ^{14}C -labeled ALA, which stabilized the bond and left the protein radioactive (130). Though this was accepted, debate occurred over which molecule of ALA bound first. PBG is an asymmetric compound, with one acid providing the acetyl group (the A-side) and the second the propionyl group (the P-side) of PBG. A unique approach was required to deduce the order, as both substrates were identical and could not be determined using standard kinetic studies. This was solved by an elegant set of experiments by P. M. Jordan using ^{13}C and ^{14}C labeled ALA. Using single turnover experiments, a stoichiometrical amount of labeled ALA was added to the enzyme and allowed to bind, presumably under high affinity. Unlabelled substrate was then added and the product was examined for radioactivity. From this, it was seen that the P-side acid always bound to the enzyme first (76).

Following these findings, molecular advances revealed that at least two separate types of HemB existed based on the necessity for either zinc or magnesium cofactors (42, 120). Once crystal structures of these enzymes were

obtained, five separate types were revealed based on the amount of Zn or Mg atoms bound to the enzyme (43). HemB monomers exist as a TIM barrel with a buried central active site with a N-terminal arm of approximately 30 residues (figure 1.5B). A small loop with a single alpha helix also sits directly above the active site acting as a lid. Functional HemB is found to adopt an tetramer of asymmetric dimers, held together by interactions of the N-terminal extensions. Crystal structures of these octamers revealed three distinct metal binding sites, a Zn site found only in animals and fungi, a second Zn site within the active site, and a Mg site. In prokaryotes, these can be functionally divided into enzymes that use zinc and magnesium as in *E. coli*, those that use magnesium as in *B. japonicum* and *P. aeruginosa*, and enzymes requiring no divalent cations as in *Rhodobacter* species (34, 35, 58).

Mechanistically, it was discovered that the Mg binding site is used to control the active site flap location, and allow for allosteric activation (69). Enzymes requiring Zn, on the other hand, bind the metal with three cysteine ligands and properly position the A-site substrate before catalysis occurs (121). Those without Zn are now thought to bind the second ALA via a second lysine in the active site directly across from the lysine responsible for the Schiff base with the P-side ALA (190). The *P. aeruginosa* Zn-independent HemB can be made Zn-dependent by mutating two amino acids within the active site, A129C and D139C (44). From this it has been conjectured that Zn-dependency was evolutionarily lost in some enzymes. This could have occurred through gene duplication followed by mutation as supported by *Nostoc* species that contain

both types. Another possibility is that a mutation in a HemB gene containing only two encoded cysteine ligands, as seen in *M. tuberculosis*, could easily lead to a loss of Zn binding.

Formation of a Linear Tetrapyrrole

The next step in the tetrapyrrole biosynthetic pathway is polymerization of four PBG molecules by the enzyme hydroxymethylbilane synthase (HemC) to form the linear tetrapyrrole 1-hydroxymethylbilane (figure 1.6A). As with other enzymes in the pathway, literature is complicated by the use of additional nomenclature, with the enzyme often being referred to as PBG deaminase and occasionally uroporphyrinogen I synthase. The latter is a misnomer brought about by the early observation that uroporphyrinogen I was produced by HemC when given PBG (167). In reality, this compound arose from the non-enzymatic, spontaneous cyclization of the highly reactive linear tetrapyrrole created. *In vivo*, this spontaneous reaction does not occur due to the following enzyme in the pathway, uroporphyrinogen III synthase (HemD). Only one type of HemC exists across all forms of life, with 45% sequence identity among the most distant proteins (49).

Early investigations into HemC led to the discovery that first pyrrole was covalently linked to the enzyme and each additional molecule was added in a stepwise fashion (78). The released product was initially proposed to be aminomethylbilane. Whether the inversion of the D ring was a product of the deaminase was a topic of debate. Once the difficult chemical synthesis of

aminomethylbilane and other variations was achieved, it was determined that the best possible substrate for a HemC/HemD coenzyme system was a aminomethylbilane of the order AP-AP-AP-AP (where A and P indicate the acetyl and propionyl groups respectively), indicating that the inversion of the D ring most likely occurred by HemD rather than the deaminase itself (167). Eventually experiments were conducted with HemC and HemD separately, leading to the realization that the aminomethylbilane could not be used as a substrate for HemD (15). A bona fide intermediate could be purified by cooling a short reaction with HemC and PBG, and this unknown compound could serve as a substrate for HemD (75). Chemical synthesis of a variety of methylbilane variations eventually led to the findings that 1-hydroxymethylbilane had identical NMR signals to that of the previously unidentified intermediate (7). Furthermore, this bilane was an excellent substrate for HemD.

Once identification, cloning, and sequencing of the *E. coli hemC* gene was complete (203), work was done to identify how the substrate was covalently attached. While these experiments did not provide the intended results, proteolytic digests revealed the presence of a chromophore with properties matching dipyrromethane, a dipyrrole made during the first two steps of the reaction (74). It was initially assumed that this dipyrromethane was an artifact left from incomplete turnover, and *E. coli* cells overexpressing HemC were grown in the presence of ¹⁴C-labelled PBG to follow incorporation. Surprisingly, when purified enzyme was incubated with unlabelled PBG, no radioactivity could be seen in the final product. In fact, the signal remained exclusive to the enzyme.

This was the first indication that HemC used its own substrate as a cofactor in catalysis (79). This was independently confirmed by other groups as well, all leading to the same conclusion (56). The requirement of two substrate molecules to form a mature, active enzyme is quite remarkable, and perhaps the most interesting feature of the HemC enzyme. Apoenzyme is unstable without this cofactor and conversion to holoenzyme produces a highly stable tertiary complex capable of catalysis. The cofactor itself serves as a primer for attachment and assembly of four more molecules of PBG into product, but it is never catalytically turned over.

To solve the problem of cofactor attachment, a combination of protein chemistry and site-directed mutagenesis was employed. Using Ellman's reagent, only three of the four cysteines present in the sequence could be detected spectrophotometrically. One of these residues was presumably being masked by attachment of the cofactor, and cysteine 242 of the *E. coli* HemC was implicated by individually mutating these residues (57). Other residues important for dipyrromethane binding are two lysines conserved across sequences that were presumed to lie within the catalytic core and bind the carboxyl side chains of the cofactor (54). Many other integral residues have been identified through the sequencing of human mutants described for nonfunctional PBG deaminase, and by the distribution it appears that the vast majority simply lead to instability of the tertiary structure (189).

The structure of *E. coli* HemC (figure 1.6B) reveals a three-domain globular protein, the first two consisting of wound parallel β -strands surrounded

by α -helices. These two portions are also connected by two short loop regions proposed to act as a hinge region. Domain 3 consists of an anti-parallel β -sheet with surrounding α -helices and contains the cysteine responsible for cofactor attachment. While covalently bound to domain 3, the dipyrromethane lies in a deep cleft at the active site between domains 1 and 2 with a large internal cavity behind it. This cavity may allow for extension of the polymerization, as the enzyme would need to occupy single, di, tri, and tetrapyrrole products but maintain a single catalytic site. One possibility is that the active site cavity can be modulated by movement in the hinge region, but the mechanisms of this remain unclear (55, 80, 106). Little is known as well about the catalytic mechanism, but currently one model is proposed. PBG is first deaminated by nearby basic residues to form a methylene intermediate. A nucleophilic attack between this and the α -carbon of the attached cofactor links the two, and deprotonation at this position completes the reaction (196).

Inversion of the D ring and Cyclization

As previously mentioned, it has been well established that uroporphyrinogen III synthase (HemD) is responsible for cyclization in tetrapyrrole biosynthesis and that inversion of the D ring occurs during this step (figure 1.7A). HemD is also the last step in the common pathways of all tetrapyrrolic molecules, leading into the creation of heme, siroheme, chlorophyll, vitamin B12, and a variety of other biologically important cofactors. Whereas generation of uroporphyrinogen I from HMB is spontaneous, the remarkable

transformation into uroporphyrinogen III requires a complex enzymatic process whereby the entire D ring of the molecule is inverted with respect to the other three. As such, most research into this amazing enzyme has involved mechanistic investigations and structure determination.

The first mechanism for this intricate process was proposed in 1961, a few years after the discovery of the enzyme itself (112). Mathewson proposed that the hydroxyl group was lost and formed a methylene bridge with the α -carbon of ring D. The resulting spirointermediate could then be cleaved between the C and D rings and give rise to uroporphyrinogen III. This mechanism was supported by previous work inexorably linking cyclization with inversion. Substrate analogs in which the D-ring was already inverted led to the enzymatic product uroporphyrinogen I, indicating that inversion occurred regardless of side chain orientation (8). Indeed, side chain substitutions of the C and D rings had little to no effect on catalysis, indicating that only the A and B rings were required for substrate recognition (153). It has been proposed that catalysis is instigated by a nucleophilic attack on the hydroxyl group by the enzyme, though this has yet to be proven (181). Furthermore, a spirocyclic compound closely resembling the proposed intermediate was shown to act as a potent inhibitor of the enzyme, strengthening this mechanism (101).

Early investigations of heme-deficient mutants led to the discovery and sequencing of the *hemD* gene in *Salmonella* and eventually *E. coli* (77, 173, 175). Those of *Bacillus subtilis*, *Synechococcus* PCC 7942, human, murine, and *Saccharomyces cerevisiae* followed shortly (197). Sequence comparison

revealed 20% similarity between *E. coli* and human, and consequently this has hindered identification of *hemD* in some organisms. Recently, advances in functional complementation have allowed identification of the enzyme in *Arabidopsis thaliana* (202). Examination of sequences reveals 7 invariable residues, only one tyrosine of which is absolutely required for catalysis to occur (165). Recombinant human and *E. coli* enzymes were eventually expressed and purified, opening the door for structure determination (3, 143).

The structure of human uroporphyrinogen III synthase was first solved in 2001 by the Phillips group and revealed a small, monomeric protein containing two similar α/β globular domains linked by a short two β -strand sheet. Though very similar, there is a 13.5° rotation of the domains with respect to each other (111). The linker between them creates a cleft that was proposed to be the active site and bind product. Conserved residues among enzymes were found to line the active site cleft, presumably interacting with substrate in some fashion. A series of structures from *Thermus thermophilus* with and without product bound led to the discovery that HemD has a huge range of conformational flexibility around the linker domain (figure 1.7B). With product bound, the enzyme adopts a “closed” state, where each domain closes in around the tetrapyrrole. These structures further proved the importance of the conserved tyrosine residue, which was poised hydrogen bond with the hydroxyl group of HMB and facilitate its loss (180). Comparative density studies using quantum mechanical calculations suggest that the D-ring initially binds in a manner that shields its terminal portion, thereby forcing an alignment that would favor the attack on the α -carbon as

predicted (192). All in all, these experiments fully support the initial mechanism presented in 1961.

Decarboxylation of Uroporphyrinogen

The next step in heme biosynthesis proceeds with the step-wise decarboxylation of uroporphyrinogen III to coproporphyrinogen III through the enzyme uroporphyrinogen decarboxylase (UROD, HemE), with the decarboxylation occurring at each of the four acetate side chains of the pyrrole rings (figure 1.8A). The enzymatic step was first described in the pioneering work of porphyrin biosynthesis by the Granick group in 1957, where it was attributed to an as yet unidentified protein (114). Further studies of accumulated intermediates revealed that one enzyme was responsible for all four steps, and that each occurred separately (31). The enzyme was subsequently identified and characterized from a wide variety of organisms including human (150), murine (195), *S. cerevisiae* (46), *Chlorella kessleri* (81), *E. coli* (172), and *Synechococcus elongatus* (210). Despite very low sequence homology (~10%), even the most distant types are capable of replacing each other *in vivo*, suggesting that very few conserved residues are required for catalysis and recognition (219). These studies have also revealed a rather unique feature of HemE, the absence of any cofactor or prosthetic group required for enzymatic activity as would be expected.

Most early work on HemE was aimed at elucidating the order of decarboxylation. By examining the structures of intermediates accumulating in

rats, the Jackson group first proposed the “clockwise hypothesis” where the D ring of uroporphyrinogen is decarboxylated first followed by the A, B, and C rings respectively (63). The initial recognition of the D ring is appealing, as only the III-isomer is recognized by the enzyme efficiently (114). In opposition, other data existed suggesting that decarboxylation might be random, as all versions of hepta-, hexa-, and pentacarboxylic porphyrinogens could be used effectively as substrates (110). In recent years, it has become widely accepted that the latter is an artifact of high, non-physiological substrate concentrations, and HemE most likely follows the clockwise pattern *in vivo* (109).

The number of active sites within HemE was a topic of debate among early researchers. The lack of conserved repeats within sequence implies either one active site or the existence of completely separate domains responsible for the same step (214). While the latter seems unlikely, chemical modification of purified HemE resulted in an enzyme that could only intermediate forms of the substrate such as heptacarboxylic porphyrinogen. This implied more than one catalytic location with a mutation in the first decarboxylative site (84). With the solving of the crystal structure of the human enzyme, the consensus shifted back to favoring one active site, and a more recent structure with product bound appears to confirm this (151). The results of the chemical modification are more easily explained by the idea that initial binding of the D ring is sterically hindered, a hindrance overcome by heptacarboxylic porphyrinogen.

The structure of human uroporphyrinogen decarboxylase adopts a $(\beta/\alpha)_8$ barrel formation with an active site cleft centered on the C-terminal end of the

beta-strands in a homodimer with active sites facing one another (214) (figure 1.8B). This has promoted the idea of substrate shuttling between each monomer in order to rearrange the porphyrinogen for the next decarboxylation, but this was disproved by artificially connecting two HemE ORFs by a small linker with one subunit inactivated. This heterodimeric complex displays identical activity to wildtype, therefore substrate shuttling may occur, but is not required (149). HemE has also been shown to be very tolerant of missense mutations within its sequence, and minor structural changes have little to no impact on activity (209). In contrast, substrate recognition is extremely stringent, and even the smallest changes uroporphyrinogen III inhibit the enzyme completely. One such molecule is porphomethene, identical to the natural substrate except for the methyl bridge between rings A and B being reduced (148). Mechanistically, HemE is thought to function by protonating each pyrrole ring by an as yet unidentified residue. The resulting positively charged intermediate is stabilized by an aspartate residue within the active cleft and serves as an electron sink for the release of the carboxylate as carbon dioxide (151). In this model, the protonated pyrrole ring acts analogously to the pyridine ring of pyridoxyl phosphate, making the substrate function act as its own cofactor during catalysis (214).

In prokaryotes, little work has been done on HemE, and consequently little specific information is known about these enzymes. Of the few that have been examined, a crystal structure exists for the enzyme in *Bacillus subtilis* that exhibits remarkable three dimensional similarities to its human counterpart. The active site cleft contains the same invariant aspartate, as well as other conserved

polar residues that may play a role in protonation (38). Structural aspects aside, the prokaryotic *hemE* is also known to bind a regulator in the Hsp90 family, HtpG in *E. coli* and *S. elongatus*. HtpG is highly conserved in cyanobacteria and proteobacteria, but is not found in the Actinobacteria or Firmicutes. The significance of this has not been established yet, but it is currently held that HtpG regulates heme synthesis by suppressing HemE activity (168, 210).

Oxidative Decarboxylation of Coproporphyrinogen

The antepenultimate step in heme biosynthesis uses the enzyme coproporphyrinogen oxidase (CPO) to concomitantly decarboxylate and oxidize the propionates attached to the A and B rings of the porphyrin moiety into vinyl groups of the product protoporphyrinogen IX (figure 1.9A). In higher organisms and some prokaryotes this reaction is mediated by oxygen dependant CPO (HemF) and requires no cofactors other than diatomic oxygen to occur. In other prokaryotes a second form of this enzyme exists, the oxygen-independent CPO (HemN), a member of the “Radical-SAM” protein family (97). These enzymes utilize a 4Fe-4S cluster to transfer a single electron to S-adenosylmethionine, cleaving methionine and generating the newly formed 5'-adenosyl radical. Here, the subsequent, highly oxidizing radical is used to carry out catalysis in the absence of oxygen (95). The majority of bacteria with an identifiable CPO contain either HemN alone or concurrent alongside HemF (183, 217) with the exception of a few obligate aerobes such as *Rickettsia* that may have lost the unneeded anaerobic version. This suggests a scenario whereby aerobic CPO

developed as a secondary means of protoporphyrinogen IX generation and anaerobic CPO was no longer required. Unfortunately, certain classes of radical SAM enzymes are often misannotated as *hemN* in published genomes, complicating the global CPO perspective in prokaryotes (50).

The oxygen-dependent HemF was initially characterized in eukaryotes, and most insight into the enzyme derives from these models. The enzyme was initially purified from bovine and murine liver independently and found to have an absolute requirement for oxygen but not reducing agents (32, 220). Isolation of intermediate porphyrinogens from rats also revealed that the initiatory decarboxylation occurs at the A ring (85). As the enzyme was successfully purified from other organisms, including *S. typhimurium*, it became clear that no cofactor was required for the reaction to proceed (16, 126, 218). This is perhaps the most striking feature of HemF, as a catalyst is required in order for dioxygen to break its activation barrier and react with a organic molecule. The majority of oxidases surmount this obstacle through transition metals or electron accepting cofactors such as flavins, unlike HemF (41). While uncharacterized, *R. sphaeroides hemF* is known to be regulated by the regulatory protein PrrA. PrrA responds to oxygen tensions, negatively repressing *hemF* transcription in low or apoxic oxygen concentrations (223). At the turn of the millennium, it was reported that *E. coli* HemF was responsible for oxidation of protoporphyrinogen IX, the proceeding step in heme production (133). This role is most likely an artifact of overexpression and not physiological, as the protoporphyrinogen oxidase in *E. coli*, HemG, functions during aerobic respiration.

The reduction/decarboxylation reaction of an enzyme lacking any known cofactors has presented a unique challenge to researchers over the years, and various mechanisms have been proposed with varying degrees of evidential support (100, 192). The currently accepted mechanism involves an initial deprotonation of the pyrrole NH by a basic residue sitting below the active site of the enzyme. The resulting radical can then react with oxygen to form a peroxide anion. A deprotonation of the carboxyl side chain by the anion would then lead to reordering of the entire ring structure, with the final formation of product, carbon dioxide, and hydrogen peroxide (93). Studies using radiolabeled substrate previously confirmed this loss of carbon dioxide (33). More recently evidence of hydrogen peroxide generation was found in *in vitro* enzyme assays (12). Proton abstraction by a nearby residue is appealing, as a single side chain could easily be poised to catalyze both the A and B ring reactions if the substrate adopts a structure where the center of the tetrapyrrole faces away from the direction of the hydrophobic side chains. Coproporphyrinogen III is known to adopt this conformation in the crystal structure of HemE, the previous enzyme in the pathway, where each pyrrole NH is anchored by a single conserved aspartate residue (151). Finally, this mechanism mimics that of the early steps of urate decarboxylase, an enzyme that uses molecular oxygen in a similar fashion without cofactors (62).

In contrast to HemF, HemN is an oxygen-independent coproporphyrinogen oxidase and is exclusive to bacterial species. Since HemN functions without the need for oxygen, the term “oxidase” does not apply, but the

nomenclature has persisted despite this. Anaerobic activity was initially reported early on from *R. sphaeroides* and shown to require Mg^{2+} , NADP, NADH, ATP, and methionine to stabilize in crude cell extracts (201). In the same study, anaerobic activity *Chromatium* extracts was seen in the presence of S-adenosylmethionine (SAM), an early hint at the Radical-SAM nature of the enzyme. Experiments with these extracts proved that the reaction proceeded in a similar fashion to HemF, with hydrogen being lost from the B-position of each ring (183). The gene responsible was eventually identified as a single enzyme using transposon mutagenesis and sequencing (18). The same gene was eventually identified in *S. typhimurium* and *E. coli*. Under anaerobic conditions, *E. coli hemN* expression is three-fold greater than during aerobic growth, and expression was shown to be dependent on iron as well (205). In *Bacillus subtilis*, *hemN* was identified by functional complementation of a *hemF/hemN* double knockout in *S. typhimurium*. A second gene was found to encode a protein with high amino acid similarity to HemN and was termed *hemZ* (166). Anaerobic induction of these proteins is regulated by a number of distinct systems, including ANR and DNR in *Pseudomonas* and resDE and FNR in *Bacillus* (60, 103). Knockouts of *hemN* and *hemZ* in *Bacillus* has no abnormal growth phenotype, implying the existence of another CPO-like gene in this species (59).

Until biochemical purification of HemN in *E. coli*, the exact nature of the enzyme remained undetermined. Spectroscopic analysis revealed a pattern typical of a 4Fe-4S cluster that was found to be oxygen sensitive and chelation of

this cluster led to complete inhibition of the enzyme. Furthermore, a cysteine motif (CXXXCXXC) essential for cluster formation was found to be common to all known HemN and HemZ sequences. This put these proteins in a superfamily of radical SAM enzymes and explained the early observed stimulation exhibited by exogenous SAM. Site-directed mutagenesis of conserved cysteines residues abolished activity, as did removal of the putative SAM binding tyrosine and histidine residues (99). A mechanism was proposed whereby flavodoxin oxidoreductase reduces the iron-sulfur cluster using NAD(P)H, leading to the formation of the desoxyadenosine radical. This portion of the mechanism, common to all radical SAM enzymes, accounts for the observed stimulation by nicotinamide cofactors. The radical then abstracts the B-position proton that finally condenses as a vinyl group with the help of an unknown electron acceptor (94).

Structure determination of *E. coli* HemN revealed a monomeric protein with two dissimilar domains (figure 1.9B). The N-terminal end consists of a parallel β -sheet with alpha helices on the surface. This domain resembles that of the TIM barrel of HemE except that one quarter of the barrel is missing. This open lateral curve creates a substrate-binding pocket perpendicular to the strand axis. The C-terminal domain functions to plug the open void of the central cavity and creates a tunnel wide enough to accommodate cofactors and the coproporphyrinogen molecule (97). The geometry of the 4Fe-4S cluster and bound SAM is appropriate for radical SAM function. Surprisingly, a second SAM molecule was found bound in a secondary position near the cluster and other

SAM binding site. This secondary motif is conserved in all known sequences of bona fide HemN, and mutagenesis studies have shown it is essential for catalysis (94). This cofactor has been proposed to act as the electron acceptor during the final steps of the reaction mechanism, but experimental evidence has not been shown (98). Comparison of HemN structure with that of other radical SAM enzymes denotes a strong ancestral lineage of enzymes that have evolved distinct functions with underlying structural changes (96).

Generation of the Planar Macrocycle

Protoporphyrinogen oxidase (PPO) catalyzes the six-electron oxidation of protoporphyrinogen IX into protoporphyrin IX during the penultimate step of heme biosynthesis. The oxidation reaction results in a planar tetrapyrrole bearing methene bridges joining each ring rather than the methylene bridges of its porphyrinogen counterpart (figure 1.10A). The resulting structure affords heme its chromophoric characteristics. PPO is the last shared step between heme and chlorophyll synthesis, with the generated porphyrin going on to have either iron or magnesium inserted, respectively. Formation of can occur spontaneously, and for a while two camps of thought existed. Either this step occurred spontaneously through the delivery of protoporphyrinogen to the final enzyme ferrochelatase or another enzyme existed. \

Evidence that a catalytic step existed was shown when the Sano group found that rat liver extracts significantly enhanced the rate of protoporphyrin accumulation (170). This enhancement showed a reliance on oxygen, leading to

the given name oxidase. The enzyme was independently purified to homogeneity from bovine and murine livers and shown to contain the cofactor flavin adenine dinucleotide (FAD) (26, 191). Eventually, the Dailey group observed that during catalysis, 3 mol of diatomic oxygen was consumed per mol of protoporphyrinogen and that hydrogen peroxide was generated during catalysis. Though stoichiometric amounts of hydrogen peroxide could not be determined, it is generally accepted that the six electron oxidation would yield 3 hydrogen peroxide molecules (40). In plants, the enzyme is the target of diphenyl ether herbicides such as acifluorfen that causes the build up of protoporphyrinogen. These compounds diffuse through the plant cell and lead to reactive oxygen species and eventually cell death (113).

In prokaryotes, PPO is designated HemY and has been cloned, expressed, and characterized from *Myxococcus xanthus* (23), *Aquifex aeolicus* (208), and *Bacillus subtilis* (30). The former two encode membrane-bound enzymes identical to eukaryotic PPO and are inhibited by acifluorfen. These enzymes are a member of an FAD superfamily also containing monoamine oxidases and phytoene desaturases (28). In contrast, *Bacillus subtilis hemY* is unique and encodes a soluble protein that is not inhibited by herbicides and exhibits broad substrate specificity, oxidizing copro- as well as protoporphyrinogen (20). This soluble form of HemY is present in all Firmicutes and Actinobacteria capable of synthesizing heme.

Structure determinations of both the membrane bound and soluble forms of HemY exist (19, 156). Both have a three-domain structure sharing the FAD-

binding and substrate binding domains. The main difference lies in the third domain, the membrane-binding portion. In *M. xanthus*, this domain consists of a hydrophobic, α -helical core that is poised to bury itself within a membrane (figure 1.10B). In *B. subtilis*, the domain is missing this hydrophobic core, and is instead quite soluble. Both structures were solved in the presence of acifluorfen, with the inhibitors bound to a pocket separate from the active site.

The precise mechanism of the enzyme remains unclear, though two separate models have been proposed. Initially, a basic model was presented based on stereochemical studies that suggested three desaturation reactions and a final tautomeric loss of hydrogen resulting from final rearrangement (19). The second mechanism was proposed by Koch *et al* based on their published structure of tobacco PPO. Their mechanism envisioned hydride transfer at a single C20 position which reduces the FAD cofactor. FAD is regenerated through the production of hydrogen peroxide, with the tetrapyrrole undergoing rearrangement and loss of another hydrogen. This is repeated two times in the same fashion, with the enzyme still acting on the same C20 position (89). This proposed mechanism is consistent with the active site of the enzyme, which sterically disfavors rotation of the substrate. This hindrance is also seen in *M. xanthus* HemY. The possibility of multiple sites within the active tunnel has not been ruled out.

Protoporphyrinogen oxidase activity has also been described in *E. coli*, though a homolog of *hemY* is not present in the genome. This is not surprising, as an enzyme reliant on oxygen could not function during anaerobic conditions.

Experiments using *E. coli* cell extracts were carried out in the 1970s by the Jacobs group and revealed the following information: First, fumarate was found to enhance protoporphyrin generation and presumed to function as an electron acceptor in anaerobic assays (65). Delving further into this, they determined that nitrate and oxygen had the same effect as potential electron acceptors (66). Next, enzymatic activity can be blocked by the electron transport inhibitor 2-heptyl-4-hydroxy quinoline-N-oxide, suggesting quinones play a vital role (64). This led to the hypothesis that this step in heme biosynthesis is coupled to the electron transport chain itself, rather than directly to single electron acceptors such as oxygen in HemY (67). Finally, the reliance on quinones was examined in *E. coli* cells devoid of menaquinone or ubiquinone and only those containing the former contained activity. Additionally, this same study determined activity was limited solely to the membrane fraction (68).

Shortly after the previous findings, researchers working with heme-deficient *E. coli* mutants found that one strain accumulated porphyrins associated with terminal enzymes (uro- copro- and proto-) but remained functional in the final step of heme biosynthesis, ferrochelatase. Porphyrin accumulation does not occur when a functional biosynthetic pathway is present, and this strain was deemed deficient in PPO activity (171). The mutation was mapped to a gene termed *hemG* (174) and is present only in the enterics, a subclass within the α -proteobacteria. The *E. coli* mutant strain has since been used as a complementation tool for testing other bona fide PPOs including the human enzyme, validating a lack of this activity within the cells(91, 134, 138). The

protein encoded by *hemG* was never purified but bears striking homology with flavodoxins, small non-catalytic proteins involved in electron transport, and its exact role in heme biosynthesis was not determined.

PPO activity has been studied to a lesser extent in other prokaryotes as well. In *Desulfovibrio gigas*, the membrane fraction of crude cell extracts was capable of oxidizing protoporphyrinogen IX and reported to function best with sulfite and nicotinamide electron acceptors (87). Later, extracts were fractionated using size exclusion chromatography, and the fraction containing PPO activity was composed of three distinct proteins (88). To date, no additional research has been carried out to identify the proteins involved. Recently it has been suggested that heme is synthesized in these organisms through a novel pathway branching from siroheme biosynthesis, using sirohydrochlorin as a substrate (105). This represents an appealing system for organisms that lack the terminal enzymes of heme biosynthesis but still require heme, and may be the predominant route of heme synthesis in archaea.

Iron Insertion and the Formation of Heme

The terminal step in heme biosynthesis is the insertion of iron into the center of the porphyrin macrocycle by the enzyme ferrochelatase (FC, HemH) (figure 1.11A), referred sometimes as heme synthase or protoheme ferrolyase. Each of the tetrapyrrole metabolic pathways contain a specific, distinct metal ion chelatase to ensure the proper metal is inserted. Thus, the specificity of ferrochelatase must be precise *in vivo*. Surprisingly in contrast, ferrochelatases

of various organisms have been reported to chelate a wide variety of metals *in vitro* (117). The enzyme's observed stringency *in vivo* has been attributed to compartmentalization and theoretical metal chaperones. Evolutionarily, ferrochelatase appears to have arisen from a fusion of an ancient ancestral chelatase, the small cobalt chelatase recently discovered in archaea (13). This enzyme is much less specific, and can accommodate the insertion of different metals into different tetrapyrrole substrates.

Much like protoporphyrinogen oxidase, ferrochelatase exists in two forms known forms, a monomeric, water soluble form in the Actinobacteria and Firmicutes and a homodimeric, membrane-associated form in other heme-synthesizers. Both are designated HemH and contain very low sequence homology among each other and with higher organisms. HemH can be further subdivided by those that contain a 2Fe-2S cluster and those that do not (29, 187). Interestingly, this division is separate from the soluble/membrane-bound division, as both groups contain organisms with and without clusters. Furthermore, the carboxy terminal region that binds this cluster bears no similarity to their counterparts in eukaryotes (24). However the clusters evolved, they are not involved in catalysis, but removal has an impact on enzyme activity. This suggests a role in further linking internal iron levels to heme biosynthesis.

The microbial enzyme was first reported from *R. sphaeroides* crude extracts (73), followed by reports that heme could be produced under non-enzymatic conditions. This was later settled by the characterization of a ferrochelatase-deficient mutant of *Spirillum itersnoii*, as well as detection of the

enzyme in *E. coli*, the alga *Cyanidium caldarium*, and a few other bacterial species (14, 25). The *hemH* gene was eventually mapped, cloned, and expressed from a wide variety of organisms (45, 82), but most of our knowledge comes from the *Bacillus subtilis* and human enzymes, both of which various structures have been solved (figure 1.11B). The bacillus enzyme contains two domains of similar core structures bearing the Rossman fold (parallel β sheet surrounded by α -helices) with a metal binding histidine residue located inside the predicted active site cleft (2). It has been proposed that substrate positioning inside the active cleft of the enzyme is crucial for proper distortion of the macrocycle (83), but this data is based on complexes with substrate analogues bound outside the actual active site. Thus, the conclusions may not be physiologically relevant. Several conserved surface residues are present in HemH that seemingly play no role in structure or catalysis, yet when mutated the resulting enzyme is inactive. From this, researchers have postulated these residues play a role in either substrate or product channeling (142).

Most of the data concerning the mechanism of ferrochelatase is a result of various human ferrochelatase structures solved to date. Much of this is relevant to the overall nature of the enzyme in all organisms, and thus will be discussed here. The general accepted mechanism prior to the solution of these structures involved enzymatic distortion of the planar protoporphyrin IX macrocycle, followed by insertion of bound iron and product release (22). While seemingly simplistic, this mechanism requires the enzyme to differentiate between very similar substrates and products during this catalysis (24). Insight came with the

structure of an E343K variant that was found to crystallize with substrate bound. The porphyrin displayed a 11.5° bend, consistent with previous theoretical configurations that would sufficiently allow iron insertion (116). Wild-type and heme-bound structures soon followed this, and a comparison of all three structures revealed a large clue as to the enzyme's mechanism. The residue H263 anchors a complex hydrogen bond network within the active site (27). Once protoporphyrin binds, this hydrogen bond network is altered dramatically and the active site is closed. Together, this facilitates iron insertion. Finally, once the product is formed, a conserved π -helix partially unwinds creating an alternate exit route for heme (118). Furthermore, it was also shown that only insertion of the proper metal into the macrocycle can initiate this release. Inhibitory metals such as Pb, Hg, Cd, or Ni are chelated into the macrocycle, but the enzyme will not release the product. In the case of Mn, the helix does not unwind at all and maintains a closed active site (117).

Regulation and Conservation of Heme Biosynthesis, and a Genomic Perspective

Heme biosynthesis represents a paradox within the cell. On one side, heme is vital to cellular operations and is required in organisms that contain hemoproteins. On the other side, heme itself is a highly toxic molecule to living systems. This is due in part to the nature of heme to incorporate into cell walls, increase permeability, and potentially cause lysis (17), as well as the observation that heme exhibits mono-oxygenase-like activity resulting in DNA, protein, and

lipid damage (1, 11, 137). The breakdown of heme within a cell can also lead to uncontrolled iron release and Fenton chemistry resulting in reactive oxygen species formation, though a direct correlation within cells has yet to be noted in microbial systems (5). In addition to heme toxicity, the intermediates of the biosynthetic pathway are also quite harmful to cells. Uncoupled to proteins, the intermediates uro-, copro-, and protoporphyrinogen are rapidly oxidized to their porphyrin counterparts. These porphyrins are highly photoreactive, and in the presence of light and oxygen give rise to singlet oxygen species and cellular damage (92). This feature has been well-documented and cleverly used in anti-cancer treatments where porphyrins are targeted to tumors. The porphyrins are then excited with light to induce tumor necrosis (139). The photoreactiveness is also responsible for light sensitivity seen in types of human porphyrias where mutations in heme biosynthetic enzymes lead to the build up of these latter intermediates (200).

In light of the adverse effects heme and its intermediates can cause, heme biosynthesis must be highly regulated and should be highly conserved. This conservation is reinforced by the idea that heme synthesis occurred very early in the evolution of life. During synthesis, all intermediates are thought to be tightly shuttled through the pathway, with no free substrate available to the cell. This creates a regulatory control point at the synthesis of ALA. Addition of exogenous ALA to a bacterial culture causes an unregulated acute increase in internal porphyrin and heme levels, suggesting that the production of the precursory compound is indeed the major regulatory route of microorganisms (147). In *E.*

coli and *Salmonella typhimurium*, the enzyme GTR is regulated *in vivo* by the presence of heme (71). This is due to a conformation change that occurs within GTR when heme is bound that allows proteolytic cleavage of a domain that would otherwise be unexposed (179). For the minority of bacteria that contain HemA, regulatory control is exhibited through direct feedback inhibition of HemA by heme (36).

For the most part conservation of the heme biosynthetic pathway also holds true. The enzymes that catalyze the conversion of ALA into uroporphyrinogen are fully conserved across all taxa and are used for the synthesis of all forms of tetrapyrroles. The only real variation is the appearance of a gene fusion of HemD with an enzyme in siroheme biosynthesis in a small number of genomes (4). The Granick hypothesis states that enzymatic pathways recapitulate themselves, and that the pathway intermediates reflect the evolutionary history of porphyrins, each with their own use (51). This further implies that uroporphyrin may have been a universal cofactor before specialization occurred. The next step in heme biosynthesis, HemE is also as conserved. Finally, the terminal enzyme, HemH, is conserved in all heme-producing species. Unfortunately, the same cannot be said of the antepenultimate and penultimate enzymes, coproporphyrinogen oxidase and protoporphyrinogen oxidase. Here, a great deal of diversity exists among these steps, and neither is limited to a specific taxonomic group (144). This is surprising, given that the final enzyme remains the same.

The antepenultimate step in heme biosynthesis is the conversion of coproporphyrinogen III to protoporphyrin IX. As stated earlier, two separate enzymes are known to catalyze this reaction, the oxygen-dependent HemF and the oxygen-independent HemN. The HemF version is homologous to the eukaryotic enzyme and represents the version adapted for life in an oxygen-rich environment. In prokaryotes, it is only present in subclasses of proteobacteria, the Bacteroidetes, the green sulfur bacteria, green non-sulfur bacteria, and the cyanobacteria. Any bacterium capable of anaerobic respiration could not contain *hemF* only, and must contain a separate form that does not require oxygen. An examination of available genomic sequence data would suggest that HemN is present in all heme-synthesizers annotated as “probable oxygen-independent coproporphyrinogen III oxidases”. Unfortunately, recent examinations and publications have revealed that these enzymes are often misannotated. It has been estimated that up to 49% of all computationally annotated genes are annotated incorrectly (178). HutW is one example, where an enzyme thought to be involved in heme uptake is annotated in databases as HemN, despite being more closely homologous with a separate family of proteins (216). When the sequence of these genes are examined, the only regions homologous are the residues required for iron-sulfur cluster formation and S-adenosyl methionine binding, indicating they are all Radical-SAM enzymes. There is no biochemical or biological data to support many of the annotated genes’ involvement in heme biosynthesis. A closer look, as will be detailed later, reveals distinct protein families with wide taxonomic spread. When searching for those that resemble

the bona fide, experimentally-verified HemN of *E. coli* and *Salmonella*, a clearer picture emerges that leaves the Firmicutes and Actinobacteria without an identifiable CPO. This indicates that at least a third uncharacterized variant of the enzyme exists, possibly more.

Even more surprising, however, is the case of the penultimate step. Here, at the six electron-oxidation of protoporphyrinogen IX to protoporphyrin IX, the porphyrin becomes planar and biologically important for both the heme and chlorophyll biosynthetic pathways. Unfortunately, it is also here that the biggest gap in our knowledge of the pathways exists. Like the previous step, the HemY enzyme homologous to the eukaryotic protoporphyrinogen oxidase is oxygen-dependent, and therefore cannot be solely responsible in obligate or facultative anaerobes. In fact, this enzyme exists only in the Firmicutes, Actinobacteria, and a few Gram-negative organisms such as *Mxyococcus* and *Aquifex* species. The only clues to activity outside of these organisms come from experiments of crude cell extracts carried out in the mid 1970s. Jacobs and Jacobs found that the enzymatic reaction required the electron transport chain of *E. coli* to be intact and that the electron acceptor used for growing the culture be present (nitrate or fumarate for anaerobic growth, oxygen for aerobic growth) (64-67). The activity was only present when cellular membrane fractions were included (68). Separate from these experiments, the Sasarman group isolated an *E. coli* mutant that exhibited the features of protoporphyrinogen oxidase loss, an acquired auxotrophy for heme and a buildup of the pathway intermediates (171). This mutation was mapped to a specific gene, now called *hemG*, sequenced, and

nothing more done (174). Prior to the research presented herein, this gene was never expressed and the encoded protein never characterized. As stated, *hemG* contains a distinct flavodoxin motif. Flavodoxins are small flavin-containing proteins that act as electron carriers between two systems and are not known for catalytic activity themselves (169). This led most researchers to conclude the encoded protein merely shuttled electrons from an as yet unidentified enzyme to the electron transport chain. Even still, *hemG* is only present in the enteric subset of α -proteobacteria. Further complicating matters is the presence of a gene annotated as *hemY* in *E. coli* and related organisms that is unrelated to bona fide protoporphyrinogen oxidase. Prior to the current studies, the vast remaining heme-synthesizing bacteria contained no identifiable enzyme, accounting for nearly 50% of sequenced genomes. This includes the well-studied species *Pseudomonas aeruginosa*, *Caulobacter crescentus*, and *Synechocystis* PCC 6803.

Together, the features outlined above represent a major flaw in our understanding of prokaryotic terminal heme biosynthesis. It is not surprising to find this amongst pathways in microorganisms, considering the outstanding diversity they display. In fact, the overall conservation of the pathway outside of these steps is quite extraordinary. It is perplexing though that the final chelation steps do not display the same diversity. This could be due to a variety of reasons. First, it is possible that chelation of specific metals first evolved before modifications of the tetrapyrrole side chains occurred, as uroporphyrin. Patchwork enzyme recruitment could then allow for further modifications that

were 'inserted' before chelation. Since the final product remained the same throughout the diversity, it is more likely that one version of each step was the original ancestral enzyme and then replaced as oxygen-requiring enzymes were adapted. Identification of the novel, missing enzymes will provide insight into this topic, as well as deliver new candidate drug targets for a variety of pathogenic microorganisms. This has been previously proposed elsewhere, including as an antifilarial target for nematodes that maintain a *Wolbachia* endosymbiont to provide them heme (193). Currently, a diphenyl ether inhibitor of plant protoporphyrinogen oxidase, acifluorfen, is used as a potent herbicide, but also affects human PPO (154). The novel genes in heme biosynthesis would not be present in humans, limiting their adverse effects. It is the author's hope that after examining the evidence presented in the following work, the reader will agree that the vast majority of these unidentified steps in heme biosynthesis have now been filled with forms unique to the microbial world.

References

1. **Aft RL & Mueller GC.** 1984. Hemin-mediated oxidative degradation of proteins. *J Biol Chem.* 259:301-305.
2. **Al-Karadaghi S, Hansson M, Nikonov S, Jonsson B, & Hederstedt L.** 1997. Crystal structure of ferrochelatase: the terminal enzyme in heme biosynthesis. *Structure.* 5:1501-1510.
3. **Alwan AF, Mgbeje BI, & Jordan PM.** 1989. Purification and properties of uroporphyrinogen III synthase (co-synthase) from an overproducing recombinant strain of *Escherichia coli* K-12. *Biochem J.* 264:397-402.
4. **Anderson PJ, Entsch B, & McKay DB.** 2001. A gene, *cobA* + *hemD*, from *Selenomonas ruminantium* encodes a bifunctional enzyme involved in the synthesis of vitamin B12. *Gene.* 281:63-70.
5. **Anzaldi LL & Skaar EP.** 2010. Overcoming the heme paradox: heme toxicity and tolerance in bacterial pathogens. *Infect Immun.* 78:4977-4989.
6. **Astner I, et al.** 2005. Crystal structure of 5-aminolevulinate synthase, the first enzyme of heme biosynthesis, and its link to XLSA in humans. *EMBO J.* 24:3166-3177.
7. **Battersby A, Fookes CJR, Gustafson-Potter KE, Matcham GWJ, & McDonald E.** 1979. Proof by synthesis that unrearranged hydroxymethylbilane is the product from deaminase and the substrate for cosynthetase in the biosynthesis of uro'gen-III. *J Chem Soc Chem Commun.* 1155-1158.

8. **Battersby AR, Fookes CJR, & Pandey PS.** 1983. Linear tetrapyrrolic intermediates for biosynthesis of the natural porphyrins. *Tetrahedron*. 39:1919-1926.
9. **Beale SI & Castelfranco PA.** 1974. The Biosynthesis of delta-Aminolevulinic Acid in Higher Plants: II. Formation of C-delta-Aminolevulinic Acid from Labeled Precursors in Greening Plant Tissues. *Plant Physiol*. 53:297-303.
10. **Beale SI, Gough SP, & Granick S.** 1975. Biosynthesis of delta-aminolevulinic acid from the intact carbon skeleton of glutamic acid in greening barley. *Proc Natl Acad Sci USA*. 72:2719-2723.
11. **Beppu M, Nagoya M, & Kikugawa K.** 1986. Role of heme compounds in the erythrocyte membrane damage induced by lipid hydroperoxide. *Chem Pharm Bull (Tokyo)*. 34:5063-5070.
12. **Breckau D, Mahlitz E, Sauerwald A, Layer G, & Jahn D.** 2003. Oxygen-dependent coproporphyrinogen III oxidase (HemF) from *Escherichia coli* is stimulated by manganese. *J Biol Chem*. 278:46625-46631.
13. **Brindley AA, Raux E, Leech HK, Schubert HL, & Warren MJ.** 2003. A story of chelatase evolution: identification and characterization of a small 13-15-kDa "ancestral" cobaltochelatase (CbiXS) in the archaea. *J Biol Chem*. 278:22388-22395.
14. **Brown SB, Holroyd JA, Vernon DI, & Jones OT.** 1984. Ferrochelatase activity in the photosynthetic alga *Cyanidium caldarium*. Development of the enzyme during biosynthesis of photosynthetic pigments. *Biochem J*. 220:861-863.

15. **Burton G, Fagerness PE, Hosozawa S, Jordan PM, & Scott AI.** 1979. 13 C n.m.r. evidence for a new intermediate, pre-uroporphyrinogen, in the enzymic transformation of porphobilinogen into uroporphyrinogens I and III. *J Chem Soc Chem Commun.* 5:202-204.
16. **Camadro JM, Chambon H, Jolles J, & Labbe P.** 1986. Purification and properties of coproporphyrinogen oxidase from the yeast *Saccharomyces cerevisiae*. *Eur J Biochem.* 156:579-587.
17. **Chou AC & Fitch CD.** 1981. Mechanism of hemolysis induced by ferriprotoporphyrin IX. *J Clin Invest.* 68:672-677.
18. **Coomber SA, Jones RM, Jordan PM, & Hunter CN.** 1992. A putative anaerobic coproporphyrinogen III oxidase in *Rhodobacter sphaeroides*. I. Molecular cloning, transposon mutagenesis and sequence analysis of the gene. *Mol Microbiol.* 6:3159-3169.
19. **Corradi HR, et al.** 2006. Crystal structure of protoporphyrinogen oxidase from *Myxococcus xanthus* and its complex with the inhibitor acifluorfen. *J Biol Chem.* 281:38625-38633.
20. **Corrigall AV, et al.** 1998. Purification of and kinetic studies on a cloned protoporphyrinogen oxidase from the aerobic bacterium *Bacillus subtilis*. *Arch Biochem Biophys.* 358:251-256.
21. **D'Aguanno S, Gonzales IN, Simmaco M, Contestabile R, & John RA.** 2003. Stereochemistry of the reactions of glutamate-1-semialdehyde aminomutase with 4,5-diaminovalerate. *J Biol Chem.* 278:40521-40526.

22. **Dailey HA.** 1996. Ferrochelatase. *Mechanisms of Metallocenter Assembly*, eds Hausinger RP, Eichorn GL, & G. ML (VCH, New York, New York), pp 77-98.
23. **Dailey HA & Dailey TA.** 1996. Protoporphyrinogen oxidase of *Myxococcus xanthus*. Expression, purification, and characterization of the cloned enzyme. *J Biol Chem.* 271:8714-8718.
24. **Dailey HA, et al.** 2000. Ferrochelatase at the millennium: structures, mechanisms and [2Fe-2S] clusters. *Cell Mol Life Sci.* 57:1909-1926.
25. **Dailey HA, Jr. & Lascelles J.** 1977. Reduction of iron and synthesis of protoheme by *Spirillum itersonii* and other organisms. *J Bacteriol.* 129:815-820.
26. **Dailey HA & Karr SW.** 1987. Purification and characterization of murine protoporphyrinogen oxidase. *Biochemistry.* 26:2697-2701.
27. **Dailey HA, et al.** 2007. Altered orientation of active site residues in variants of human ferrochelatase. Evidence for a hydrogen bond network involved in catalysis. *Biochemistry.* 46:7973-7979.
28. **Dailey TA & Dailey HA.** 1998. Identification of an FAD superfamily containing protoporphyrinogen oxidases, monoamine oxidases, and phytoene desaturase. Expression and characterization of phytoene desaturase of *Myxococcus xanthus*. *J Biol Chem.* 273:13658-13662.
29. **Dailey TA & Dailey HA.** 2002. Identification of [2Fe-2S] clusters in microbial ferrochelatases. *J Bacteriol.* 184:2460-2464.

30. **Dailey TA, Meissner P, & Dailey HA.** 1994. Expression of a cloned protoporphyrinogen oxidase. *J Biol Chem.* 269:813-815.
31. **de Viale LC, Garcia RC, De Pisarev DK, Tomio JM, & Grinstein M.** 1969. Studies on uroporphyrinogen decarboxylase from chicken erythrocytes. *FEBS Lett.* 5:149-152.
32. **del Batlle AM, Benson A, & Rimington C.** 1965. Purification and properties of coproporphyrinogenase. *Biochem J.* 97:731-740.
33. **Elder GH, Evans JO, Jackson JR, & Jackson AH.** 1978. Factors determining the sequence of oxidative decarboxylation of the 2- and 4-propionate substituents of coproporphyrinogen III by coproporphyrinogen oxidase in rat liver. *Biochem J.* 169:215-223.
34. **Erskine PT, et al.** 2003. X-ray structure of a putative reaction intermediate of 5-aminolaevulinic acid dehydratase. *Biochem J.* 373:733-738.
35. **Erskine PT, et al.** 1999. X-ray structure of 5-aminolevulinic acid dehydratase from *Escherichia coli* complexed with the inhibitor levulinic acid at 2.0 Å resolution. *Biochemistry.* 38:4266-4276.
36. **Fales L, Kryszak L, & Zeilstra-Ryalls J.** 2001. Control of hemA expression in *Rhodobacter sphaeroides* 2.4.1: effect of a transposon insertion in the hbdA gene. *J Bacteriol.* 183:1568-1576.
37. **Faller M, Matsunaga M, Yin S, Loo JA, & Guo F.** 2007. Heme is involved in microRNA processing. *Nat Struct Mol Biol.* 14:23-29.

38. **Fan J, Liu Q, Hao Q, Teng M, & Niu L.** 2007. Crystal structure of uroporphyrinogen decarboxylase from *Bacillus subtilis*. *J Bacteriol.* 189:3573-3580.
39. **Fenna R, Zeng J, & Davey C.** 1995. Structure of the green heme in myeloperoxidase. *Arch Biochem Biophys.* 316:653-656.
40. **Ferreira GC & Dailey HA.** 1988. Mouse protoporphyrinogen oxidase. Kinetic parameters and demonstration of inhibition by bilirubin. *Biochem J.* 250:597-603.
41. **Fetzner S & Steiner RA.** Cofactor-independent oxidases and oxygenases. *Appl Microbiol Biotechnol.* 86:791-804.
42. **Frankenberg N, Heinz DW, & Jahn D.** 1999. Production, purification, and characterization of a Mg²⁺-responsive porphobilinogen synthase from *Pseudomonas aeruginosa*. *Biochemistry.* 38:13968-13975.
43. **Frankenberg N, Jahn D, & Jaffe EK.** 1999. *Pseudomonas aeruginosa* contains a novel type V porphobilinogen synthase with no required catalytic metal ions. *Biochemistry.* 38:13976-13982.
44. **Frere F, Reents H, Schubert WD, Heinz DW, & Jahn D.** 2005. Tracking the evolution of porphobilinogen synthase metal dependence in vitro. *J Mol Biol.* 345:1059-1070.
45. **Frustaci JM & O'Brian MR.** 1993. The *Escherichia coli* *visA* gene encodes ferrochelatase, the final enzyme of the heme biosynthetic pathway. *J Bacteriol.* 175:2154-2156.

46. **Garey JR, et al.** 1992. Uroporphyrinogen decarboxylase in *Saccharomyces cerevisiae*. HEM12 gene sequence and evidence for two conserved glycines essential for enzymatic activity. *Eur J Biochem.* 205:1011-1016.
47. **Giardina G, et al.** 2008. NO sensing in *Pseudomonas aeruginosa*: structure of the transcriptional regulator DNR. *J Mol Biol.* 378:1002-1015.
48. **Gibson KD, Neuberger A, & Scott JJ.** 1955. The purification and properties of delta-aminolaevulinic acid dehydrase. *Biochem J.* 61:618-629.
49. **Gill R, et al.** 2009. Structure of human porphobilinogen deaminase at 2.8 Å: the molecular basis of acute intermittent porphyria. *Biochem J.* 420:17-25.
50. **Goto T, Aoki R, Minamizaki K, & Fujita Y.** 2010. Functional differentiation of two analogous coproporphyrinogen III oxidases for heme and chlorophyll biosynthesis pathways in the cyanobacterium *Synechocystis* sp. PCC 6803. *Plant Cell Physiol.* 51:650-663.
51. **Granick S.** 1957. Speculations on the origins and evolution of photosynthesis. *Ann N Y Acad Sci.* 69:292-308.
52. **Grigg JC, Ukpabi G, Gaudin CF, & Murphy ME.** 2010. Structural biology of heme binding in the *Staphylococcus aureus* Isd system. *J Inorg Biochem.* 104:341-348.
53. **Guengerich FP.** 2007. Mechanisms of cytochrome P450 substrate oxidation: MiniReview. *J Biochem Mol Toxicol.* 21:163-168.
54. **Hadener A, Alefounder PR, Hart GJ, Abell C, & Battersby AR.** 1990. Investigation of putative active-site lysine residues in hydroxymethylbilane synthase. Preparation and characterization of mutants in which (a) Lys-55,

- (b) Lys-59 and (c) both Lys-55 and Lys-59 have been replaced by glutamine. *Biochem J.* 271:487-491.
55. **Hadener A, et al.** 1993. Purification, characterization, crystallisation and X-ray analysis of selenomethionine-labelled hydroxymethylbilane synthase from *Escherichia coli*. *Eur J Biochem.* 211:615-624.
56. **Hart GH, Miller AD, Leeper FJ, & Battersby A.** 1987. Biosynthesis of the natural porphyrins: proof that hydroxymethylbilane synthase (porphobilinogen deaminase) uses a novel binding group in its catalytic action. *J Chem Soc Chem Commun.* 1762-1765.
57. **Hart GJ, Miller AD, & Battersby AR.** 1988. Evidence that the pyrromethane cofactor of hydroxymethylbilane synthase (porphobilinogen deaminase) is bound through the sulphur atom of a cysteine residue. *Biochem J.* 252:909-912.
58. **Heinemann IU, et al.** 2010. Structure of the heme biosynthetic *Pseudomonas aeruginosa* porphobilinogen synthase in complex with the antibiotic alaremycin. *Antimicrob Agents Chemother.* 54:267-272.
59. **Hippler B, et al.** 1997. Characterization of *Bacillus subtilis* hemN. *J Bacteriol.* 179:7181-7185.
60. **Homuth G, Rompf A, Schumann W, & Jahn D.** 1999. Transcriptional control of *Bacillus subtilis* hemN and hemZ. *J Bacteriol.* 181:5922-5929.
61. **Hunter GA & Ferreira GC.** 1999. Lysine-313 of 5-Aminolevulinate Synthase Acts as a General Base during Formation of the Quinonoid Reaction Intermediates. *Biochemistry.* 38:12526.

62. **Imhoff RD, Power NP, Borrok MJ, & Tipton PA.** 2003. General base catalysis in the urate oxidase reaction: evidence for a novel Thr-Lys catalytic diad. *Biochemistry.* 42:4094-4100.
63. **Jackson AH, et al.** 1976. Macrocyclic intermediates in the biosynthesis of porphyrins. *Philos Trans R Soc Lond B Biol Sci.* 273:191-206.
64. **Jacobs JM & Jacobs NJ.** 1977. The late steps of anaerobic heme biosynthesis in *E. coli*: role for quinones in protoporphyrinogen oxidation. *Biochem Biophys Res Commun.* 78:429-433.
65. **Jacobs NJ & Jacobs JM.** 1975. Fumarate as alternate electron acceptor for the late steps of anaerobic heme synthesis in *Escherichia coli*. *Biochem Biophys Res Commun.* 65:435-441.
66. **Jacobs NJ & Jacobs JM.** 1976. Nitrate, fumarate, and oxygen as electron acceptors for a late step in microbial heme synthesis. *Biochim Biophys Acta.* 449:1-9.
67. **Jacobs NJ & Jacobs JM.** 1977. Evidence for involvement of the electron transport system at a late step of anaerobic microbial heme synthesis. *Biochim Biophys Acta.* 459:141-144.
68. **Jacobs NJ & Jacobs JM.** 1978. Quinones as hydrogen carriers for a late step in anaerobic heme biosynthesis in *Escherichia coli*. *Biochim Biophys Acta.* 544:540-546.
69. **Jaffe EK.** 2004. The porphobilinogen synthase catalyzed reaction mechanism. *Bioorg Chem.* 32:316-325.

70. **Jahn D, Verkamp E, & Soll D.** 1992. Glutamyl-transfer RNA: a precursor of heme and chlorophyll biosynthesis. *Trends Biochem Sci.* 17:215-218.
71. **Javor GT & Febre EF.** 1992. Enzymatic basis of thiol-stimulated secretion of porphyrins by *Escherichia coli*. *J Bacteriol.* 174:1072-1075.
72. **Jeuken LJ, Jones AK, Chapman SK, Cecchini G, & Armstrong FA.** 2002. Electron-transfer mechanisms through biological redox chains in multicenter enzymes. *J Am Chem Soc.* 124:5702-5713.
73. **Jones MS & Jones OT.** 1970. Ferrochelatase of *Rhodopseudomonas spheroides*. *Biochem J.* 119:453-462.
74. **Jordan PM.** 1991. The biosynthesis of 5-aminolaevulinic acid and its transformation into uroporphyrinogen III. *Biosynthesis of Tetrapyrroles*, ed Jordan PM (Elsevier, Amsterdam).
75. **Jordan PM, et al.** 1979. Pre-uroporphyrinogen: a substrate for uroporphyrinogen III cosynthetase. *J Chem Soc Chem Commun.* 5:204-205.
76. **Jordan PM & Gibbs PN.** 1985. Mechanism of action of 5-aminolaevulinate dehydratase from human erythrocytes. *Biochem J.* 227:1015-1020.
77. **Jordan PM, Mgbeje BI, Thomas SD, & Alwan AF.** 1988. Nucleotide sequence for the hemD gene of *Escherichia coli* encoding uroporphyrinogen III synthase and initial evidence for a hem operon. *Biochem J.* 249:613-616.

78. **Jordan PM & Seehra JS.** 1979. The biosynthesis of uroporphyrinogen III: order of assembly of the four porphobilinogen molecules in the formation of the tetrapyrrole ring. *FEBS Lett.* 104:364-366.
79. **Jordan PM & Warren MJ.** 1987. Evidence for a dipyrromethane cofactor at the catalytic site of *E. coli* porphobilinogen deaminase. *FEBS Lett.* 225:87-92.
80. **Jordan PM, et al.** 1992. Crystallization and preliminary X-ray investigation of *Escherichia coli* porphobilinogen deaminase. *J Mol Biol.* 224:269-271.
81. **Juarez AB, Aldonatti C, Vigna MS, & Rios de Molina Mdel C.** 2007. Studies on uroporphyrinogen decarboxylase from *Chlorella kessleri* (*Trebouxiophyceae, Chlorophyta*). *Can J Microbiol.* 53:303-312.
82. **Kanazireva E & Biel AJ.** 1996. Nucleotide sequence of the *Rhodobacter capsulatus* hemH gene. *Gene.* 170:149-150.
83. **Karlberg T, et al.** 2008. Porphyrin binding and distortion and substrate specificity in the ferrochelatase reaction: the role of active site residues. *J Mol Biol.* 378:1074-1083.
84. **Kawanishi S, Seki Y, & Sano S.** 1983. Uroporphyrinogen decarboxylase. Purification, properties, and inhibition by polychlorinated biphenyl isomers. *J Biol Chem.* 258:4285-4292.
85. **Kennedy GY, Jackson AH, Kenner GW, & Suckling CJ.** 1970. Isolation, structure and synthesis of a tricarboxylic porphyrin from the harderian glands of the rat. *FEBS Lett.* 6:9-12.

86. **Kikuchi G, Kumar A, Talmage P, & Shemin D.** 1958. The enzymatic synthesis of delta-aminolevulinic acid. *J Biol Chem.* 233:1214-1219.
87. **Klemm DJ & Barton LL.** 1985. Oxidation of protoporphyrinogen in the obligate anaerobe *Desulfovibrio gigas*. *J Bacteriol.* 164:316-320.
88. **Klemm DJ & Barton LL.** 1987. Purification and properties of protoporphyrinogen oxidase from an anaerobic bacterium, *Desulfovibrio gigas*. *J Bacteriol.* 169:5209-5215.
89. **Koch M, et al.** 2004. Crystal structure of protoporphyrinogen IX oxidase: a key enzyme in haem and chlorophyll biosynthesis. *EMBO J.* 23:1720-1728.
90. **Kurokawa H, et al.** 2004. A redox-controlled molecular switch revealed by the crystal structure of a bacterial heme PAS sensor. *J Biol Chem.* 279:20186-20193.
91. **Kusaba A, et al.** 2002. Cloning and expression of a *Porphyromonas gingivalis* gene for protoporphyrinogen oxidase by complementation of a hemG mutant of *Escherichia coli*. *Oral Microbiol Immunol.* 17:290-295.
92. **Kusama Y, Bernier M, & Hearse DJ.** 1990. Photoactivation of porphyrins: studies of reactive oxygen intermediates and arrhythmogenesis in the aerobic rat heart. *Cardiovasc Res.* 24:676-682.
93. **Lash TD.** 2005. The enigma of coproporphyrinogen oxidase: how does this unusual enzyme carry out oxidative decarboxylations to afford vinyl groups? *Bioorg Med Chem Lett.* 15:4506-4509.

94. **Layer G, et al.** 2005. Radical S-adenosylmethionine enzyme coproporphyrinogen III oxidase HemN: functional features of the [4Fe-4S] cluster and the two bound S-adenosyl-L-methionines. *J Biol Chem.* 280:29038-29046.
95. **Layer G, Heinz DW, Jahn D, & Schubert WD.** 2004. Structure and function of radical SAM enzymes. *Curr Opin Chem Biol.* 8:468-476.
96. **Layer G, et al.** 2005. Structural and functional comparison of HemN to other radical SAM enzymes. *Biol Chem.* 386:971-980.
97. **Layer G, Moser J, Heinz DW, Jahn D, & Schubert WD.** 2003. Crystal structure of coproporphyrinogen III oxidase reveals cofactor geometry of Radical SAM enzymes. *EMBO J.* 22:6214-6224.
98. **Layer G, et al.** 2006. The substrate radical of *Escherichia coli* oxygen-independent coproporphyrinogen III oxidase HemN. *J Biol Chem.* 281:15727-15734.
99. **Layer G, Verfurth K, Mahlitz E, & Jahn D.** 2002. Oxygen-independent coproporphyrinogen-III oxidase HemN from *Escherichia coli*. *J Biol Chem.* 277:34136-34142.
100. **Lee DS, et al.** 2005. Structural basis of hereditary coproporphyrinemia. *Proc Natl Acad Sci USA.* 102:14232-14237.
101. **Leeper FJ.** 1994. The evidence for a spirocyclic intermediate in the formation of uroporphyrinogen III by cosynthase. *Ciba Found Symp.* 180:111-123; discussion 124-130.

102. **Lewin A & Hederstedt L.** 2006. Compact archaeal variant of heme A synthase. *FEBS Lett.* 580:5351-5356.
103. **Lieb C, Siddiqui RA, Hippler B, Jahn D, & Friedrich B.** 1998. The *Alcaligenes eutrophus* hemN gene encoding the oxygen-independent coproporphyrinogen III oxidase, is required for heme biosynthesis during anaerobic growth. *Arch Microbiol.* 169:52-60.
104. **Lin J, Fu W, & Cen P.** 2009. Characterization of 5-aminolevulinate synthase from *Agrobacterium radiobacter*, screening new inhibitors for 5-aminolevulinate dehydratase from *Escherichia coli* and their potential use for high 5-aminolevulinate production. *Bioresour Technol.* 100:2293-2297.
105. **Lobo SA, Brindley A, Warren MJ, & Saraiva LM.** 2009. Functional characterization of the early steps of tetrapyrrole biosynthesis and modification in *Desulfovibrio vulgaris* Hildenborough. *Biochem J.* 420:317-325.
106. **Louie GV, et al.** 1992. Structure of porphobilinogen deaminase reveals a flexible multidomain polymerase with a single catalytic site. *Nature.* 359:33-39.
107. **Luer C, et al.** 2005. Complex formation between glutamyl-tRNA reductase and glutamate-1-semialdehyde 2,1-aminomutase in *Escherichia coli* during the initial reactions of porphyrin biosynthesis. *J Biol Chem.* 280:18568-18572.
108. **Luer C, et al.** 2007. Glutamate recognition and hydride transfer by *Escherichia coli* glutamyl-tRNA reductase. *FEBS J.* 274:4609-4614.

109. **Luo J & Lim CK.** 1993. Order of uroporphyrinogen III decarboxylation on incubation of porphobilinogen and uroporphyrinogen III with erythrocyte uroporphyrinogen decarboxylase. *Biochem J.* 289 (Pt 2):529-532.
110. **Luo JL & Lim CK.** 1991. Random decarboxylation of uroporphyrinogen III by human hepatic uroporphyrinogen decarboxylase. *J Chromatogr.* 566:409-413.
111. **Mathews MA, et al.** 2001. Crystal structure of human uroporphyrinogen III synthase. *EMBO J.* 20:5832-5839.
112. **Mathewson JH & Corwin AH.** 1961. Biosynthesis of pyrrole pigments: A mechanism for porphobilinogen polymerization. *J Am Chem Soc.* 83:135-137.
113. **Matringe M, Camadro JM, Labbe P, & Scalla R.** 1989. Protoporphyrinogen oxidase as a molecular target for diphenyl ether herbicides. *Biochem J.* 260:231-235.
114. **Mauzerall D & Granick S.** 1958. Porphyrin biosynthesis in erythrocytes. III. Uroporphyrinogen and its decarboxylase. *J Biol Chem.* 232:1141-1162.
115. **Mauzerall DC.** 1998. Evolution of porphyrins. *Clin Dermatol.* 16:195-201.
116. **Medlock A, Swartz L, Dailey TA, Dailey HA, & Lanzilotta WN.** 2007. Substrate interactions with human ferrochelatase. *Proc Natl Acad Sci USA.* 104:1789-1793.
117. **Medlock AE, Carter M, Dailey TA, Dailey HA, & Lanzilotta WN.** 2009. Product release rather than chelation determines metal specificity for ferrochelatase. *J Mol Biol.* 393:308-319.

118. **Medlock AE, Dailey TA, Ross TA, Dailey HA, & Lanzilotta WN.** 2007. A pi-helix switch selective for porphyrin deprotonation and product release in human ferrochelatase. *J Mol Biol.* 373:1006-1016.
119. **Meehan M, Burke FM, Macken S, & Owen P.** 2010. Characterization of the haem-uptake system of the equine pathogen *Streptococcus equi* subsp. *equi*. *Microbiology.* 156:1824-1835.
120. **Mitchell LW & Jaffe EK.** 1993. Porphobilinogen synthase from *Escherichia coli* is a Zn(II) metalloenzyme stimulated by Mg(II). *Arch Biochem Biophys.* 300:169-177.
121. **Mitchell LW, Volin M, & Jaffe EK.** 1995. The phylogenetically conserved histidines of *Escherichia coli* porphobilinogen synthase are not required for catalysis. *J Biol Chem.* 270:24054-24059.
122. **Morton DJ, et al.** 2009. The heme-binding protein (HbpA) of *Haemophilus influenzae* as a virulence determinant. *Int J Med Microbiol.* 299:479-488.
123. **Moser J, Lorenz S, Hubschwerlen C, Rompf A, & Jahn D.** 1999. *Methanopyrus kandleri* glutamyl-tRNA reductase. *J Biol Chem.* 274:30679-30685.
124. **Moser J, et al.** 2001. V-shaped structure of glutamyl-tRNA reductase, the first enzyme of tRNA-dependent tetrapyrrole biosynthesis. *EMBO J.* 20:6583-6590.
125. **Munro AW, Girvan HG, McLean KJ, Cheesman MR, & Leys D.** 2009. Heme and Hemoproteins. *Tetrapyrroles: Birth, Life and Death*, eds Warren MJ & Smith AG (Springer Science, New York, New York), pp 160-183.

126. **Nagaraj VA, Prasad D, Arumugam R, Rangarajan PN, & Padmanaban G.** Characterization of coproporphyrinogen III oxidase in *Plasmodium falciparum* cytosol. *Parasitol Int.* 59:121-127.
127. **Nandi DL.** 1978. Delta-aminolevulinic acid synthase of *Rhodopseudomonas spheroides*. Binding of pyridoxal phosphate to the enzyme. *Arch Biochem Biophys.* 188:266-271.
128. **Nandi DL.** 1978. Studies on delta-aminolevulinic acid synthase of *Rhodopseudomonas spheroides*. Reversibility of the reaction, kinetic, spectral, and other studies related to the mechanism of action. *J Biol Chem.* 253:8872-8877.
129. **Nandi DL, Baker-Cohen KF, & Shemin D.** 1968. Delta-aminolevulinic acid dehydratase of *Rhodopseudomonas spheroides*. *J Biol Chem.* 243:1224-1230.
130. **Nandi DL & Shemin D.** 1968. Delta-aminolevulinic acid dehydratase of *Rhodopseudomonas spheroides*. 3. Mechanism of porphobilinogen synthesis. *J Biol Chem.* 243:1236-1242.
131. **Nandi DL & Shemin D.** 1968. Delta-aminolevulinic acid dehydratase of *Rhodopseudomonas spheroides*. II. Association to polymers and dissociation to subunits. *J Biol Chem.* 243:1231-1235.
132. **Nandi DL & Shemin D.** 1977. Quaternary structure of delta-aminolevulinic acid synthase from *Rhodopseudomonas spheroides*. *J Biol Chem.* 252:2278-2280.

133. **Narita S, Taketani S, & Inokuchi H.** 1999. Oxidation of protoporphyrinogen IX in *Escherichia coli* is mediated by the aerobic coproporphyrinogen oxidase. *Mol Gen Genet.* 261:1012-1020.
134. **Narita S, et al.** 1996. Molecular cloning and characterization of a cDNA that encodes protoporphyrinogen oxidase of *Arabidopsis thaliana*. *Gene.* 182:169-175.
135. **Neidle EL & Kaplan S.** 1993. Expression of the *Rhodobacter sphaeroides* hemA and hemT genes, encoding two 5-aminolevulinic acid synthase isozymes. *J Bacteriol.* 175:2292-2303.
136. **Nemeth AM, Russell CS, & Shemin D.** 1957. The succinate-glycine cycle. II. Metabolism of delta-aminolevulinic acid. *J Biol Chem.* 229:415-422.
137. **Nir U, Ladan H, Malik Z, & Nitzan Y.** 1991. In vivo effects of porphyrins on bacterial DNA. *J Photochem Photobiol B.* 11:295-306.
138. **Nishimura K, Taketani S, & Inokuchi H.** 1995. Cloning of a human cDNA for protoporphyrinogen oxidase by complementation in vivo of a hemG mutant of *Escherichia coli*. *J Biol Chem.* 270:8076-8080.
139. **O'Connor AE, Gallagher WM, & Byrne AT.** 2009. Porphyrin and nonporphyrin photosensitizers in oncology: preclinical and clinical advances in photodynamic therapy. *Photochem Photobiol.* 85:1053-1074.
140. **Oglesby-Sherrouse AG & Vasil ML.** 2010. Characterization of a heme-regulated non-coding RNA encoded by the prrF locus of *Pseudomonas aeruginosa*. *PLoS One.* 5:e9930.

141. **Oh-hama T.** 1997. Evolutionary consideration on 5-aminolevulinate synthase in nature. *Orig Life Evol Biosph.* 27:405-412.
142. **Olsson U, Billberg A, Sjoval S, Al-Karadaghi S, & Hansson M.** 2002. In vivo and in vitro studies of *Bacillus subtilis* ferrochelatase mutants suggest substrate channeling in the heme biosynthesis pathway. *J Bacteriol.* 184:4018-4024.
143. **Omata Y, Sakamoto H, Higashimoto Y, Hayashi S, & Noguchi M.** 2004. Purification and characterization of human uroporphyrinogen III synthase expressed in *Escherichia coli*. *J Biochem.* 136:211-220.
144. **Panek H & O'Brian MR.** 2002. A whole genome view of prokaryotic haem biosynthesis. *Microbiology.* 148:2273-2282.
145. **Perutz MF.** 1960. Structure of hemoglobin. *Brookhaven Symp Biol.* 13:165-183.
146. **Petricek M, Petrickova K, Havlicek L, & Felsberg J.** 2006. Occurrence of two 5-aminolevulinate biosynthetic pathways in *Streptomyces nodosus* subsp. *asukaensis* is linked with the production of asukamycin. *J Bacteriol.* 188:5113-5123.
147. **Philipp-Dormston WK & Doss M.** 1973. Comparison of porphyrin and heme biosynthesis in various heterotrophic bacteria. *Enzyme.* 16:57-64.
148. **Phillips JD, Bergonia HA, Reilly CA, Franklin MR, & Kushner JP.** 2007. A porphomethene inhibitor of uroporphyrinogen decarboxylase causes porphyria cutanea tarda. *Proc Natl Acad Sci USA.* 104:5079-5084.

149. **Phillips JD, Warby CA, Whitby FG, Kushner JP, & Hill CP.** 2009. Substrate shuttling between active sites of uroporphyrinogen decarboxylase is not required to generate coproporphyrinogen. *J Mol Biol.* 389:306-314.
150. **Phillips JD, Whitby FG, Kushner JP, & Hill CP.** 1997. Characterization and crystallization of human uroporphyrinogen decarboxylase. *Protein Sci.* 6:1343-1346.
151. **Phillips JD, Whitby FG, Kushner JP, & Hill CP.** 2003. Structural basis for tetrapyrrole coordination by uroporphyrinogen decarboxylase. *EMBO J.* 22:6225-6233.
152. **Piao Y, Kiatpapan P, Yamashita M, & Murooka Y.** 2004. Effects of expression of hemA and hemB genes on production of porphyrin in *Propionibacterium freudenreichii*. *Appl Environ Microbiol.* 70:7561-7566.
153. **Pichon C, Atshaves BP, Xue T, Stolowich NJ, & Scott AI.** 1994. Studies on uro'gen III synthase with modified bilanes. *Bioorg Med Chem Lett.* 4:1105-1110.
154. **Prasad AR & Dailey HA.** 1995. Generation of resistance to the diphenyl ether herbicide acifluorfen by MEL cells. *Biochem Biophys Res Commun.* 215:186-191.
155. **Puustinen A & Wikstrom M.** 1991. The heme groups of cytochrome o from *Escherichia coli*. *Proc Natl Acad Sci USA.* 88:6122-6126.
156. **Qin X, et al.** Structural insight into unique properties of protoporphyrinogen oxidase from *Bacillus subtilis*. *J Struct Biol.* 170:76-82.

157. **Radin NS, Rittenberg D, & Shemin D.** 1950. The role of glycine in the biosynthesis of heme. *J Biol Chem.* 184:745-753.
158. **Rae TD & Goff HM.** 1998. The heme prosthetic group of lactoperoxidase. Structural characteristics of heme I and heme I-peptides. *J Biol Chem.* 273:27968-27977.
159. **Ramakrishnan SN & Muscat GE.** 2006. The orphan Rev-erb nuclear receptors: a link between metabolism, circadian rhythm and inflammation? *Nucl Recept Signal.* 4:e009.
160. **Randau L, et al.** 2004. tRNA recognition by glutamyl-tRNA reductase. *J Biol Chem.* 279:34931-34937.
161. **Ranson-Olson B & Zeilstra-Ryalls JH.** 2008. Regulation of the *Rhodobacter sphaeroides* 2.4.1 hemA gene by PrrA and FnrL. *J Bacteriol.* 190:6769-6778.
162. **Reilly TJ, Green BA, Zlotnick GW, & Smith AL.** 2001. Contribution of the DDDD motif of *H. influenzae* e (P4) to phosphomonoesterase activity and heme transport. *FEBS Lett.* 494:19-23.
163. **Rieble S & Beale SI.** 1988. Transformation of glutamate to delta-aminolevulinic acid by soluble extracts of *Synechocystis* sp. PCC 6803 and other oxygenic prokaryotes. *J Biol Chem.* 263:8864-8871.
164. **Roberts GP, Kerby RL, Youn H, & Conrad M.** 2005. CooA, a paradigm for gas sensing regulatory proteins. *J Inorg Biochem.* 99:280-292.
165. **Roessner CA, Ponnampereuma K, & Scott AI.** 2002. Mutagenesis identifies a conserved tyrosine residue important for the activity of

- uroporphyrinogen III synthase from *Anacystis nidulans*. FEBS Lett. 525:25-28.
166. **Rompf A, et al.** 1998. Regulation of *Pseudomonas aeruginosa* hemF and hemN by the dual action of the redox response regulators Anr and Dnr. Mol Microbiol. 29:985-997.
167. **Rossetti MV, Jukbat de Geralnik AA, & Batlle AMC.** 1977. Porphyrin biosynthesis in *Euglena gracilis*—II. Pyrrolymethane intermediates in the enzymic cyclotetramerization of porphobilinogen Int J Biochem. 8:781-787.
168. **Saito M, Watanabe S, Yoshikawa H, & Nakamoto H.** 2008. Interaction of the molecular chaperone HtpG with uroporphyrinogen decarboxylase in the cyanobacterium *Synechococcus elongatus* PCC 7942. Biosci Biotechnol Biochem. 72:1394-1397.
169. **Sancho J.** 2006. Flavodoxins: sequence, folding, binding, function and beyond. Cell Mol Life Sci. 63:855-864.
170. **Sano S & Granick S.** 1961. Mitochondrial coproporphyrinogen oxidase and protoporphyrin formation. J Biol Chem. 236:1173-1180.
171. **Sasarman A, et al.** 1979. Mapping of a new hem gene in *Escherichia coli* K12. J Gen Microbiol. 113:297-303.
172. **Sasarman A, et al.** 1975. Uroporphyrin-accumulating mutant of *Escherichia coli* K-12. J Bacteriol. 124:1205-1212.
173. **Sasarman A & Desrochers M.** 1976. Uroporphyrinogen III cosynthase-deficient mutant of *Salmonella typhimurium* LT2. J Bacteriol. 128:717-721.

174. **Sasarman A, et al.** 1993. Nucleotide sequence of the hemG gene involved in the protoporphyrinogen oxidase activity of *Escherichia coli* K12. *Can J Microbiol.* 39:1155-1161.
175. **Sasarman A, et al.** 1987. Molecular cloning and sequencing of the hemD gene of *Escherichia coli* K-12 and preliminary data on the Uro operon. *J Bacteriol.* 169:4257-4262.
176. **Schauer S, et al.** 2002. *Escherichia coli* glutamyl-tRNA reductase. Trapping the thioester intermediate. *J Biol Chem.* 277:48657-48663.
177. **Schneegurt MA & Beale SI.** 1988. Characterization of the RNA Required for Biosynthesis of delta-Aminolevulinic Acid from Glutamate : Purification by Anticodon-Based Affinity Chromatography and Determination That the UUC Glutamate Anticodon Is a General Requirement for Function in ALA Biosynthesis. *Plant Physiol.* 86:497-504.
178. **Schoes AM, Brown SD, Dodevski I, & Babbitt PC.** 2009. Annotation error in public databases: misannotation of molecular function in enzyme superfamilies. *PLoS Comput Biol.* 5:e1000605.
179. **Schobert M & Jahn D.** 2002. Regulation of heme biosynthesis in non-phototrophic bacteria. *J Mol Microbiol Biotechnol.* 4:287-294.
180. **Schubert HL, Phillips JD, Heroux A, & Hill CP.** 2008. Structure and mechanistic implications of a uroporphyrinogen III synthase-product complex. *Biochemistry.* 47:8648-8655.

181. **Schubert HL, et al.** 2002. Structural diversity in metal ion chelation and the structure of uroporphyrinogen III synthase. *Biochem Soc Trans.* 30:595-600.
182. **Schulze JO, Schubert WD, Moser J, Jahn D, & Heinz DW.** 2006. Evolutionary relationship between initial enzymes of tetrapyrrole biosynthesis. *J Mol Biol.* 358:1212-1220.
183. **Seehra JS, Jordan PM, & Akhtar M.** 1983. Anaerobic and aerobic coproporphyrinogen III oxidases of *Rhodopseudomonas spheroides*. Mechanism and stereochemistry of vinyl group formation. *Biochem J.* 209:709-718.
184. **Shemin D & Kumin S.** 1952. The mechanism of porphyrin formation; the formation of a succinyl intermediate from succinate. *J Biol Chem.* 198:827-837.
185. **Shemin D & Rittenberg D.** 1945. The utilization of glycine for the synthesis of a porphyrin. *J Biol Chem.* 159:567-568.
186. **Shemin D & Russell CS.** 1953. delta-Aminolevulinic acid, its role in the biosynthesis of porphyrins and purines. *J Am Chem Soc.* 75:4873-4874.
187. **Shepherd M, Dailey TA, & Dailey HA.** 2006. A new class of [2Fe-2S]-cluster-containing protoporphyrin (IX) ferrochelatases. *Biochem J.* 397:47-52.
188. **Shoolingin-Jordan PM, et al.** 2003. 5-Aminolevulinic acid synthase: mechanism, mutations and medicine. *Biochim Biophys Acta.* 1647:361-366.

189. **Shoolingin-Jordan PM, Al-Dbass A, McNeill LA, Sarwar M, & Butler D.** 2003. Human porphobilinogen deaminase mutations in the investigation of the mechanism of dipyrromethane cofactor assembly and tetrapyrrole formation. *Biochem Soc Trans.* 31:731-735.
190. **Shoolingin-Jordan PM, et al.** 2002. 5-Aminolaevulinic acid dehydratase: metals, mutants and mechanism. *Biochem Soc Trans.* 30:584-590.
191. **Siepkner LJ, Ford M, de Kock R, & Kramer S.** 1987. Purification of bovine protoporphyrinogen oxidase: immunological cross-reactivity and structural relationship to ferrochelatase. *Biochim Biophys Acta.* 913:349-358.
192. **Silva PJ & Ramos MJ.** 2008. A comparative density-functional study of the reaction mechanism of the O₂-dependent coproporphyrinogen III oxidase. *Bioorg Med Chem.* 16:2726-2733.
193. **Slatko BE, Taylor MJ, & Foster JM.** 2010. The *Wolbachia* endosymbiont as an anti-filarial nematode target. *Symbiosis.* 51:55-65.
194. **Smart JL & Bauer CE.** 2006. Tetrapyrrole biosynthesis in *Rhodobacter capsulatus* is transcriptionally regulated by the heme-binding regulatory protein, HbrL. *J Bacteriol.* 188:1567-1576.
195. **Smith AG & Francis JE.** 1979. Decarboxylation of porphyrinogens by rat liver uroporphyrinogen decarboxylase. *Biochem J.* 183:455-458.
196. **Song G, et al.** 2009. Structural insight into acute intermittent porphyria. *FASEB J.* 23:396-404.

197. **Stamford NP, Capretta A, & Battersby AR.** 1995. Expression, purification and characterisation of the product from the *Bacillus subtilis* hemD gene, uroporphyrinogen III synthase. *Eur J Biochem.* 231:236-241.
198. **Stevens JM, Uchida T, Daltrop O, & Ferguson SJ.** 2005. Covalent cofactor attachment to proteins: cytochrome c biogenesis. *Biochem Soc Trans.* 33:792-795.
199. **Stojiljkovic I, Evavold BD, & Kumar V.** 2001. Antimicrobial properties of porphyrins. *Expert Opin Investig Drugs.* 10:309-320.
200. **Stolzel U, Stauch T, & Doss MO.** 2010. [Porphyrias]. *Internist (Berl).* 51:1525-1533; quiz 1534.
201. **Tait GH.** 1972. Coproporphyrinogenase activities in extracts of *Rhodopseudomonas spheroides* and *Chromatium* strain D. *Biochem J.* 128:1159-1169.
202. **Tan FC, et al.** 2008. Identification and characterization of the *Arabidopsis* gene encoding the tetrapyrrole biosynthesis enzyme uroporphyrinogen III synthase. *Biochem J.* 410:291-299.
203. **Thomas SD & Jordan PM.** 1986. Nucleotide sequence of the hemC locus encoding porphobilinogen deaminase of *Escherichia coli* K12. *Nucleic Acids Res.* 14:6215-6226.
204. **Toney MD.** 2005. Reaction specificity in pyridoxal phosphate enzymes. *Arch Biochem Biophys.* 433:279-287.

205. **Troup B, Hungerer C, & Jahn D.** 1995. Cloning and characterization of the *Escherichia coli* hemN gene encoding the oxygen-independent coproporphyrinogen III oxidase. *J Bacteriol.* 177:3326-3331.
206. **Tyacke RJ, Harwood JL, & John RA.** 1993. Properties of the pyridoxalimine form of glutamate semialdehyde aminotransferase (glutamate-1-semialdehyde 2,1-aminomutase) and analysis of its role as an intermediate in the formation of aminolaevulinate. *Biochem J.* 293 (Pt 3):697-701.
207. **Vos MH, Borisov VB, Liebl U, Martin JL, & Konstantinov AA.** 2000. Femtosecond resolution of ligand-heme interactions in the high-affinity quinol oxidase bd: A di-heme active site? *Proc Natl Acad Sci USA.* 97:1554-1559.
208. **Wang KF, Dailey TA, & Dailey HA.** 2001. Expression and characterization of the terminal heme synthetic enzymes from the hyperthermophile *Aquifex aeolicus*. *FEMS Microbiol Lett.* 202:115-119.
209. **Warby CA, et al.** 2009. Structural and kinetic characterization of mutant human uroporphyrinogen decarboxylases. *Cell Mol Biol (Noisy-le-grand).* 55:40-45.
210. **Watanabe S, et al.** 2007. Studies on the role of HtpG in the tetrapyrrole biosynthesis pathway of the cyanobacterium *Synechococcus elongatus* PCC 7942. *Biochem Biophys Res Commun.* 352:36-41.

211. **Weinstein JD & Beale SI.** 1984. Biosynthesis of Protoheme and Heme a Precursors Solely from Glutamate in the Unicellular Red Alga *Cyanidium caldarium*. *Plant Physiol.* 74:146-151.
212. **Weinstein JD & Beale SI.** 1985. RNA is required for enzymatic conversion of glutamate to delta-aminolevulinate by extracts of *Chlorella vulgaris*. *Arch Biochem Biophys.* 239:87-93.
213. **Weinstein JD, Mayer SM, & Beale SI.** 1987. Formation of delta-Aminolevulinic Acid from Glutamic Acid in Algal Extracts : Separation into an RNA and Three Required Enzyme Components by Serial Affinity Chromatography. *Plant Physiol.* 84:244-250.
214. **Whitby FG, Phillips JD, Kushner JP, & Hill CP.** 1998. Crystal structure of human uroporphyrinogen decarboxylase. *EMBO J.* 17:2463-2471.
215. **Woodard SI & Dailey HA.** 1995. Regulation of heme biosynthesis in *Escherichia coli*. *Arch Biochem Biophys.* 316:110-115.
216. **Wyckoff EE, Schmitt M, Wilks A, & Payne SM.** 2004. HutZ is required for efficient heme utilization in *Vibrio cholerae*. *J Bacteriol.* 186:4142-4151.
217. **Xu K, Delling J, & Elliott T.** 1992. The genes required for heme synthesis in *Salmonella typhimurium* include those encoding alternative functions for aerobic and anaerobic coproporphyrinogen oxidation. *J Bacteriol.* 174:3953-3963.
218. **Xu K & Elliott T.** 1993. An oxygen-dependent coproporphyrinogen oxidase encoded by the hemF gene of *Salmonella typhimurium*. *J Bacteriol.* 175:4990-4999.

219. **Xu K, Ndivo V, & Majumdar D.** 2002. Cloning and expression of Chlorobium vibrioforme uroporphyrinogen decarboxylase gene in a *Salmonella* heme auxotroph. *Curr Microbiol.* 45:456-458.
220. **Yoshinaga T & Sano S.** 1980. Coproporphyrinogen oxidase. I. Purification, properties, and activation by phospholipids. *J Biol Chem.* 255:4722-4726.
221. **Zamocky M, Furtmuller PG, & Obinger C.** 2010. Evolution of structure and function of Class I peroxidases. *Arch Biochem Biophys.* 500:45-57.
222. **Zeilstra-Ryalls JH & Kaplan S.** 1995. Regulation of 5-aminolevulinic acid synthesis in *Rhodobacter sphaeroides* 2.4.1: the genetic basis of mutant H-5 auxotrophy. *J Bacteriol.* 177:2760-2768.
223. **Zeilstra-Ryalls JH & Schornberg KL.** 2006. Analysis of hemF gene function and expression in *Rhodobacter sphaeroides* 2.4.1. *J Bacteriol.* 188:801-804.
224. **Zhang J & Ferreira GC.** 2002. Transient state kinetic investigation of 5-aminolevulinic acid synthase reaction mechanism. *J Biol Chem.* 277:44660-44669.

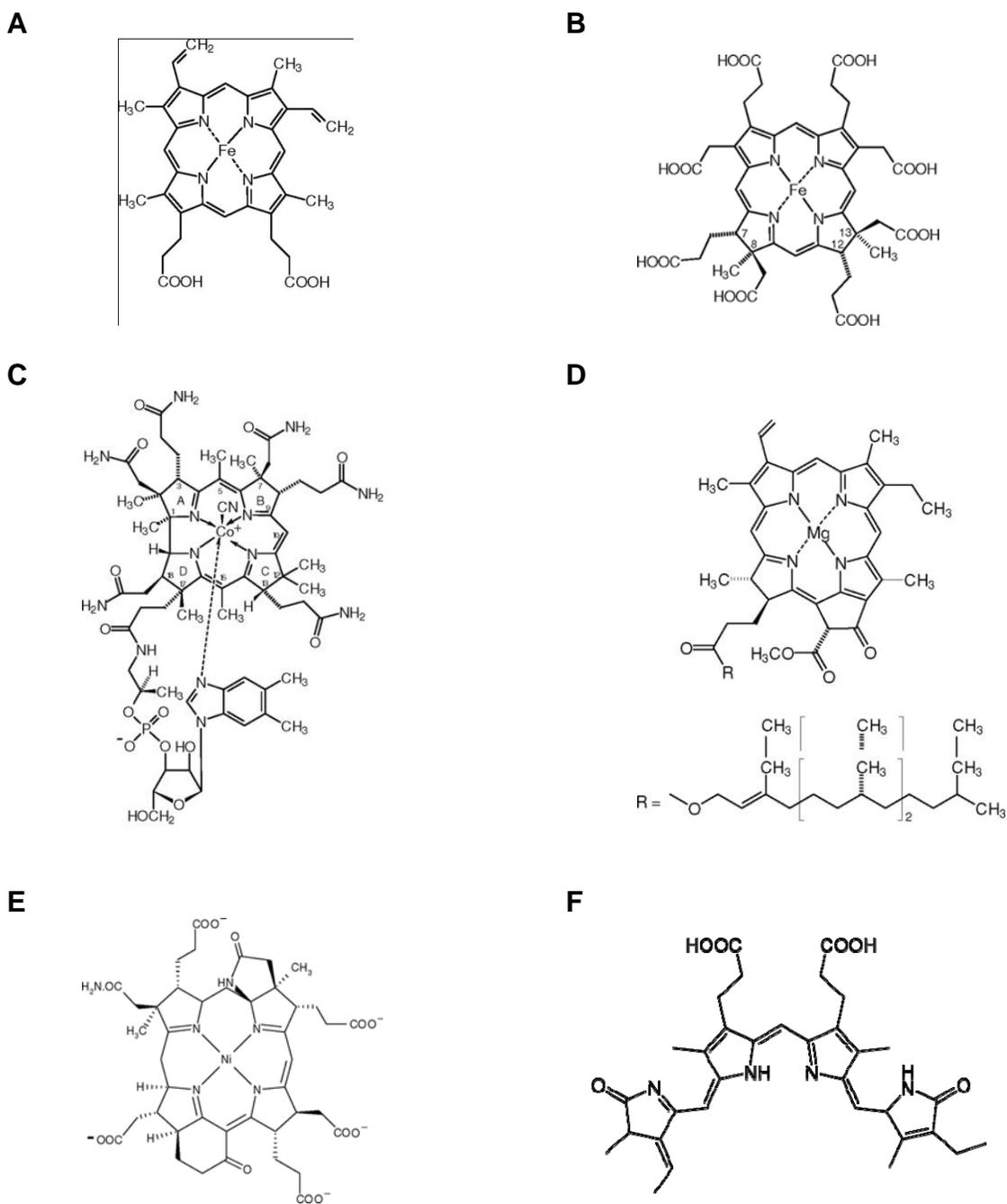


Figure 1.1. Examples of biological tetrapyrroles: **A)** Heme, a cofactor various proteins including hemoglobin, cytochromes, and peroxidases. **B)** Siroheme, used in sulfur and nitrogen reduction. **C)** Cobalamin, also known as vitamin B12, used by enzymes such as isomerases and methyltransferases. **D)** Chlorophyll, the light absorbing pigment used in photosynthesis. **E)** Coenzyme F₄₃₀, a cofactor in archaeal methanogenesis. **F)** Phycocyanobilin, a linear tetrapyrrole chromophore in light harvesting proteins.

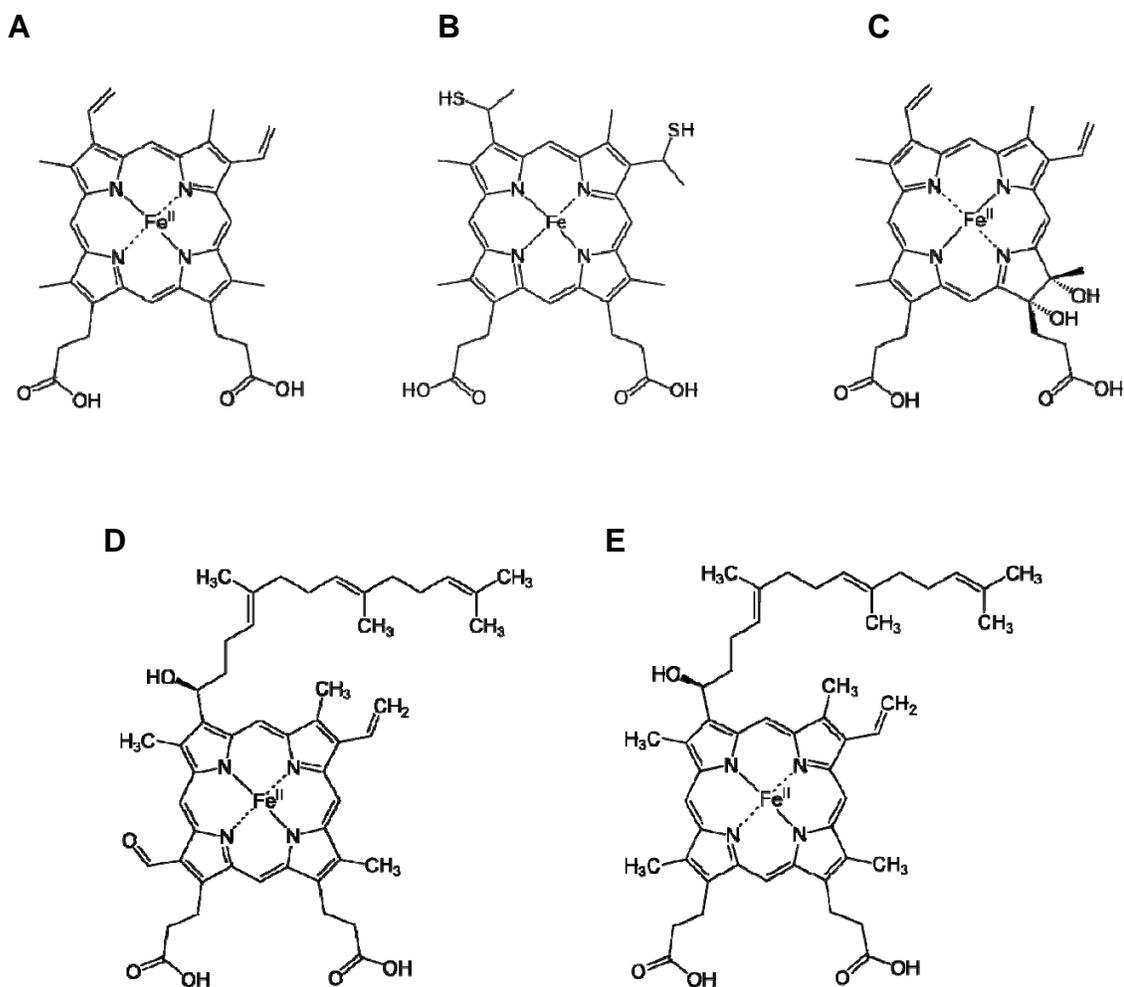
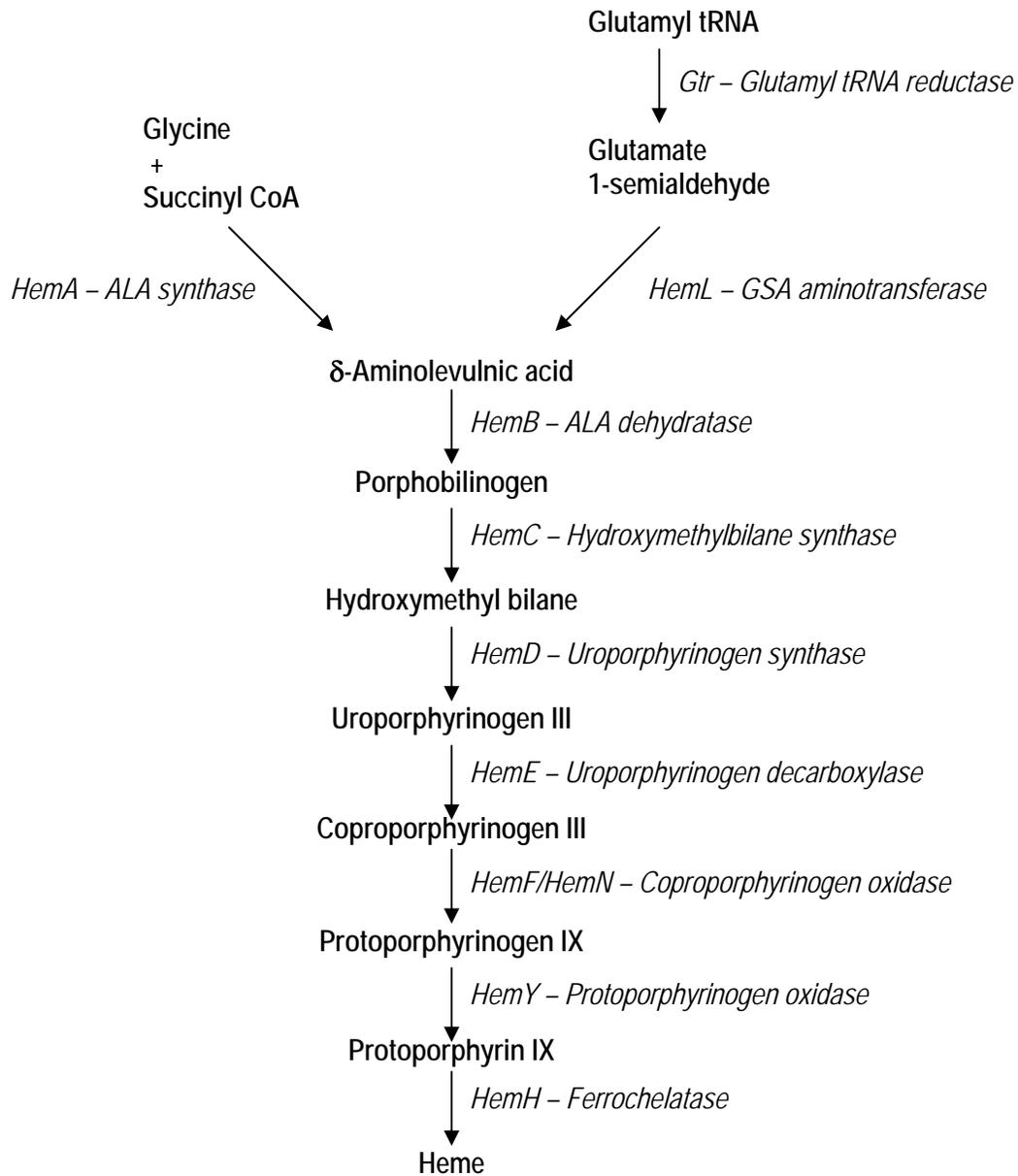


Figure 1.2. Biological forms of Heme. **A)** Protoheme, or heme B, the principle form generated by the biosynthetic pathway. **B)** Heme C, modified to contain thiol groups for covalent attachment to proteins, **C)** Heme D, shown in the trans-formation without the added lactone ring. **D)** Heme A and **E)** Heme O, two forms containing a farnesyl side chain used as a lipophilic anchor in bacterial oxidases.



Scheme 1.1. The heme biosynthetic pathway. Intermediates are shown in bold; enzymes involved in catalysis are shown in italics.

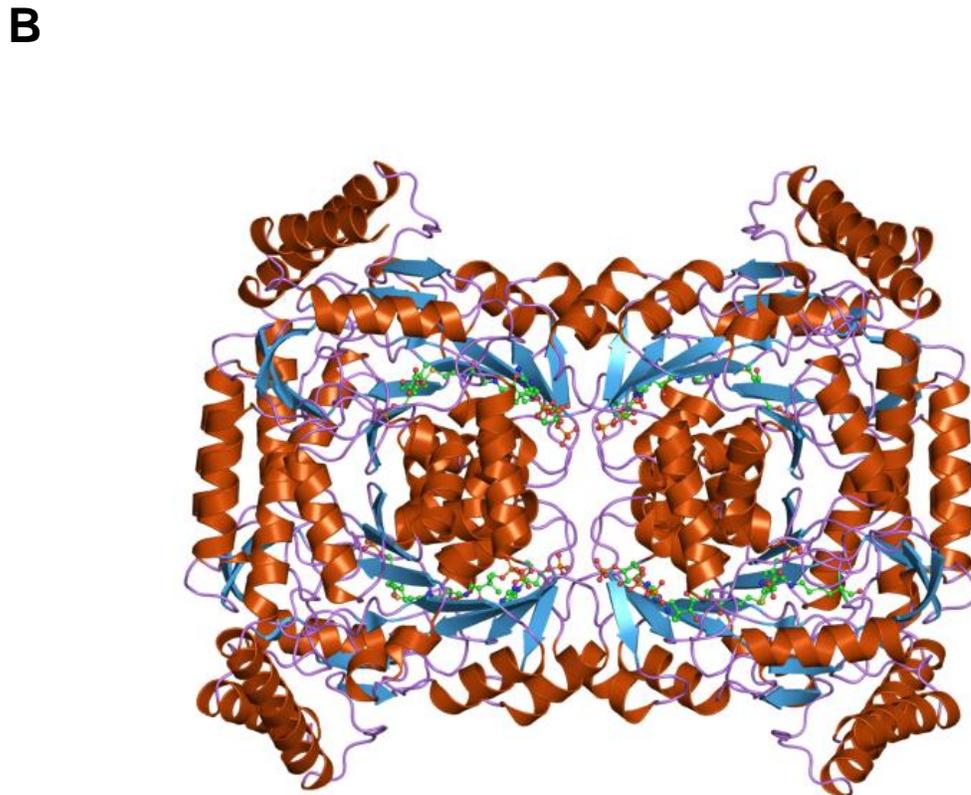
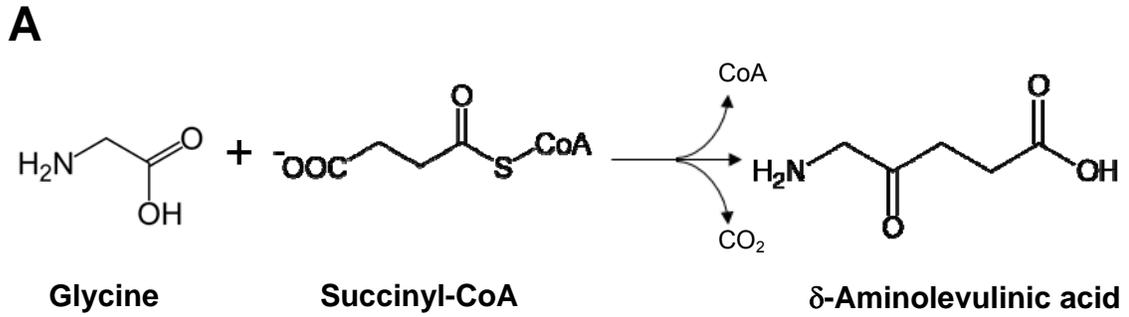


Figure 1.3. The C₄ pathway. **A)** δ -Aminolevulinic acid is generated through a condensation of glycine and succinyl-CoA. This reaction is catalyzed by the enzyme ALA synthase. The reaction also requires the turnover of a PLP molecule. **B)** Crystal structure of ALAS from *Rhodobacter capsulatus*. The enzyme is presented as a dimer of dimers.

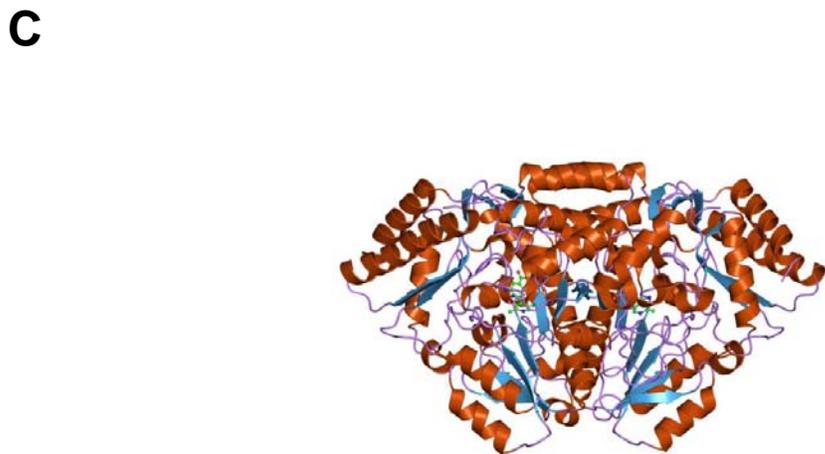
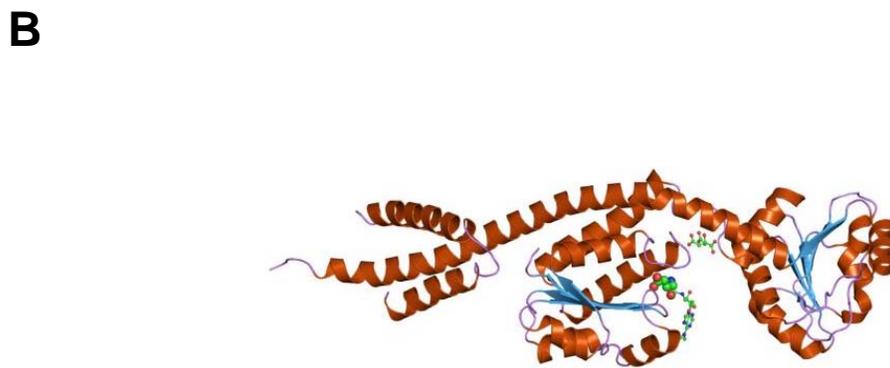
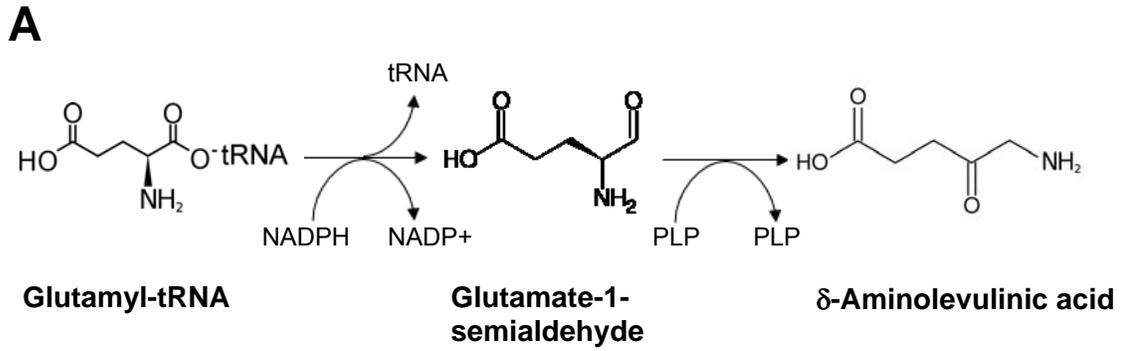


Figure 1.4. The C5 pathway **A)** The creation of ALA requires two steps, the conversion of glutamyl-tRNA to glutamate semialdehyde followed by an isomerization step to form the final product by the enzymes **B)** glutamyl tRNA reductase and **C)** GSA aminotransferase. The former is shown as a monomer, but physiologically exists as a dimer that interacts with the latter dimer *in vivo*. The structures are from *Methanopyrus kandleri* and *Synechococcus*, respectively.

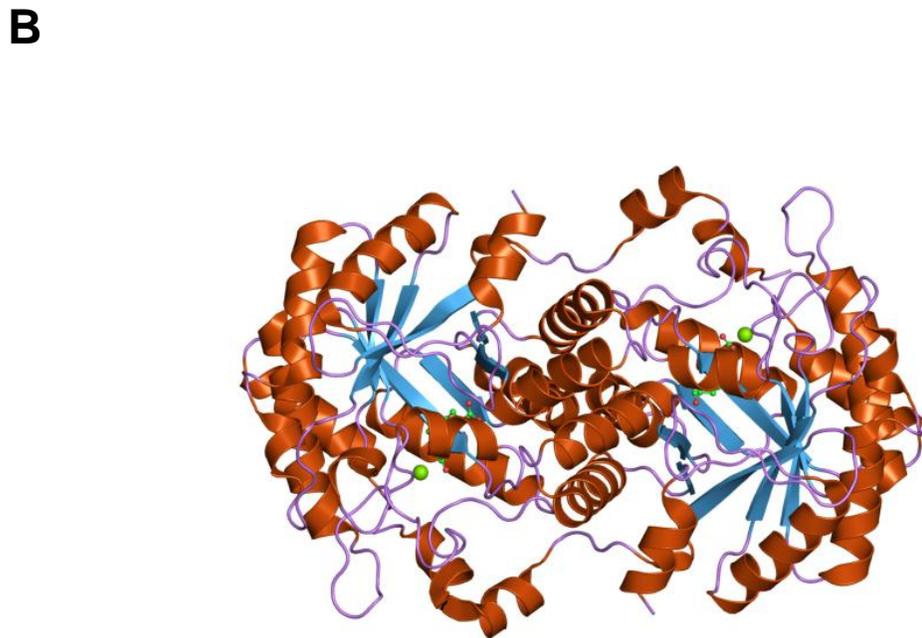
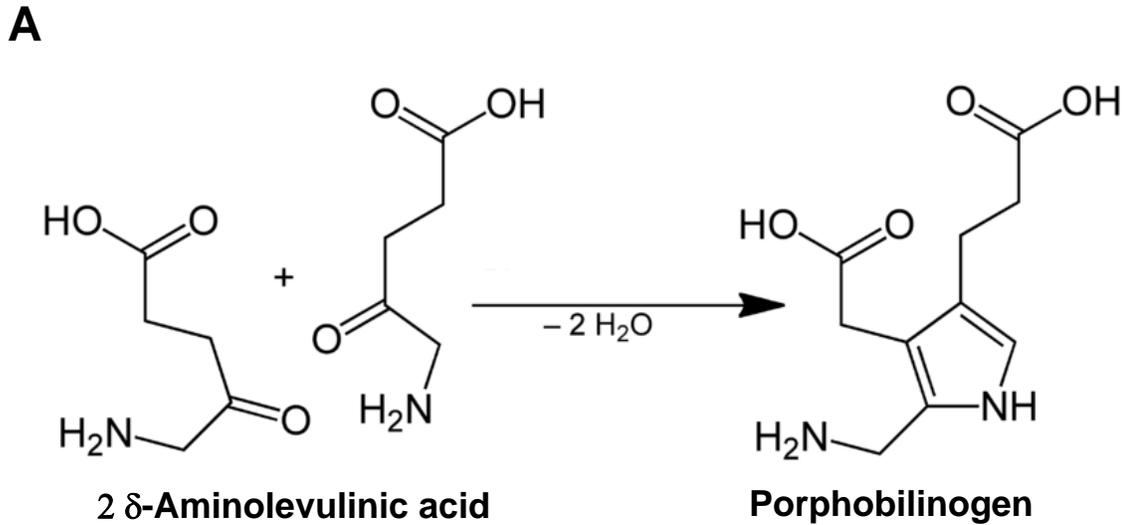


Figure 1.5. The creation of the monopyrrole porphobilinogen. **A)** Two molecules of ALA are condensed together to form porphobilinogen with the loss of two waters. This reaction is catalyzed by ALA dehydratase. Each ALA is aligned to show its position in the final product. **B)** Crystal structure of ALA dehydratase from *E. coli*, shown as a dimer. *In vivo*, four dimers adopt a homooctamer configuration.

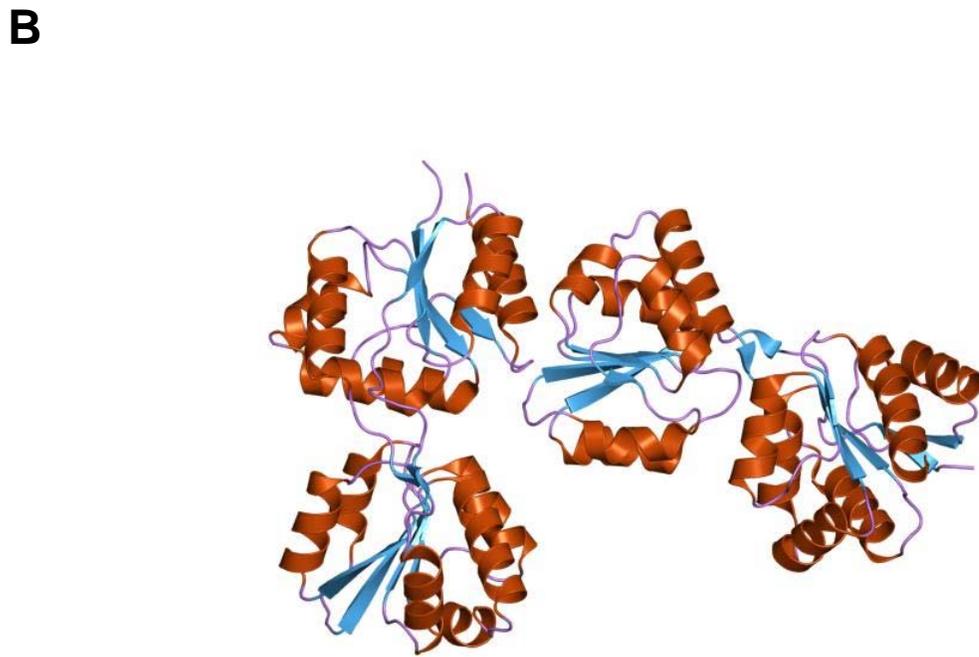
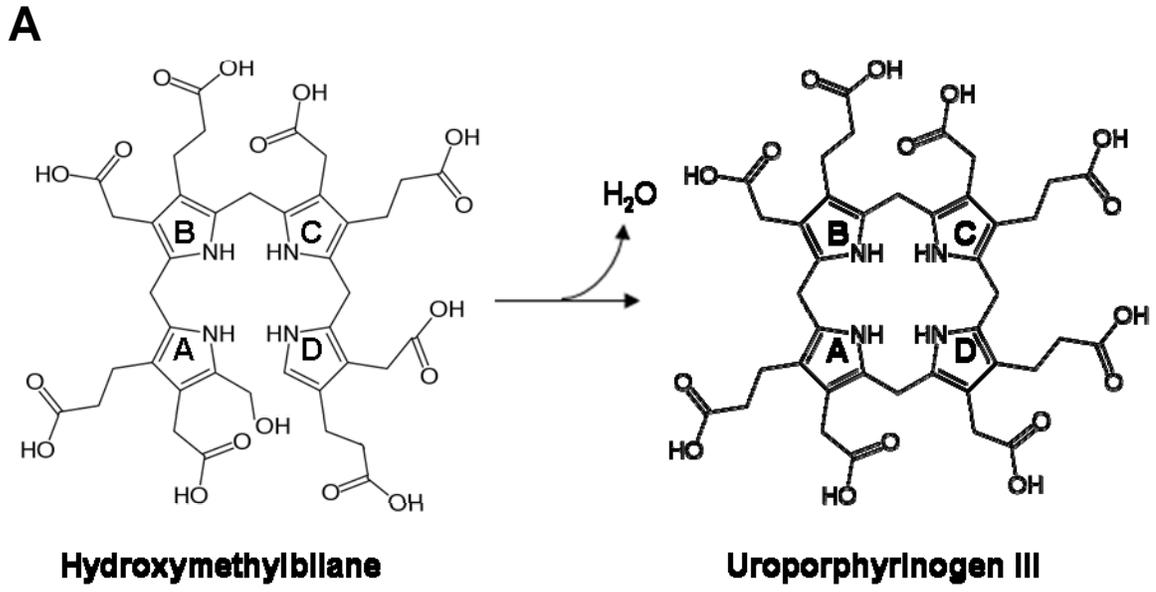


Figure 1.7. Cyclization of the tetrapyrrole ring. **A)** Hydroxymethylbilane will spontaneously cyclized to the non-biological form uroporphyrinogen I. The enzyme uroporphyrinogen synthase functions to invert the D ring leading to the correct form, uroporphyrinogen III. **B)** Crystal structure of *Thermus thermophilus* uroporphyrinogen synthase. The structure is shown as an asymmetric dimer.

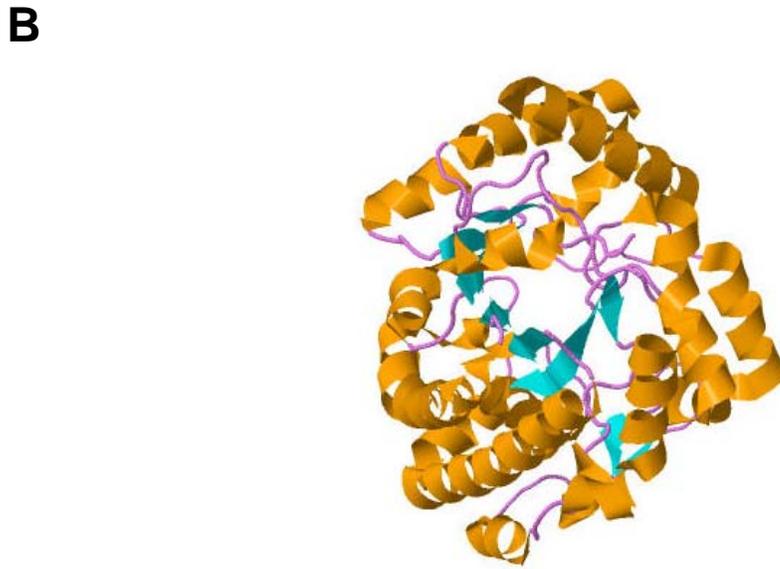
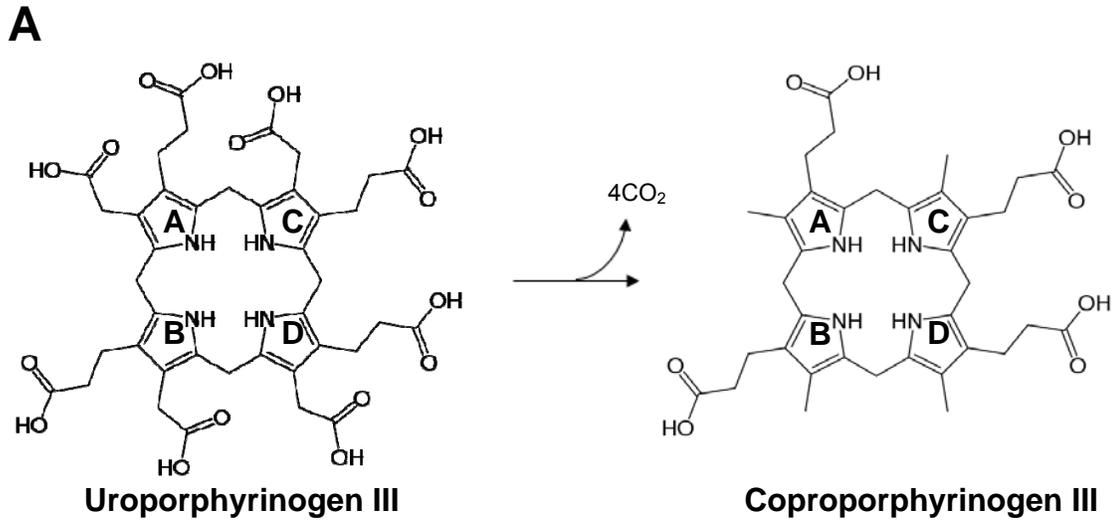


Figure 1.8. Decarboxylation of the acetate side-chains. **A)** Each of the side chains, starting at the D-ring, is decarboxylated by the enzyme uroporphyrinogen decarboxylase. **B)** The crystal structure of the *Bacillus subtilis* enzyme, as viewed down the $(a/b)_8$ barrel. This enzyme is one of the few decarboxylases that contain no cofactors.

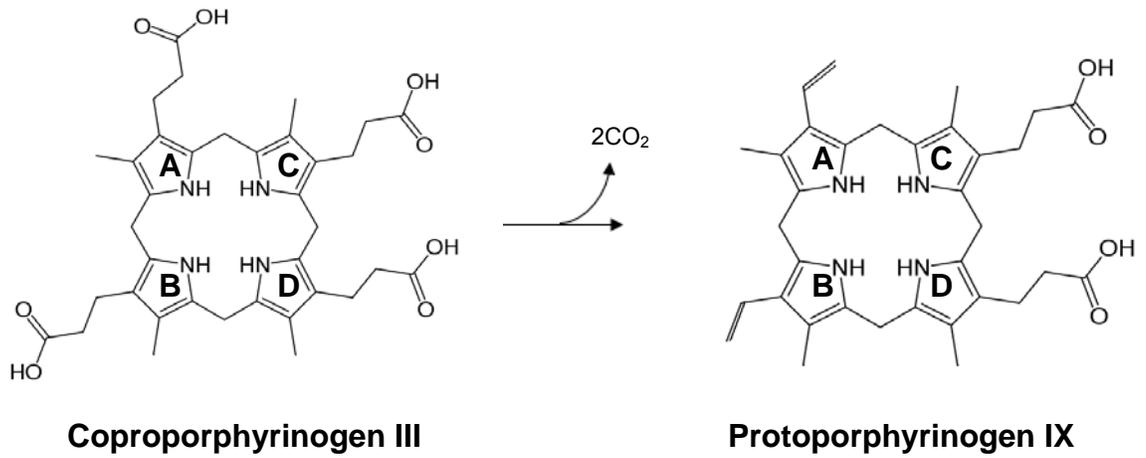
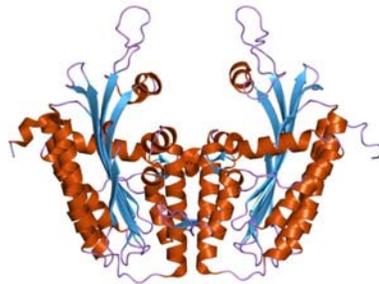
A**B****C**

Figure 1.9. The antepenultimate step of heme biosynthesis. A) This step requires the oxidative decarboxylation of the propionate groups of rings A and B to form vinyl groups. In prokaryotes, this is catalyzed by two known enzymes, the oxygen-dependent HemF (**B**), and the oxygen-independent Radical-SAM enzyme HemN (**C**). The structures are from *Saccharomyces cerevisiae* coproporphyrinogen oxidase (homologous to HemF) and *E. coli* HemN respectively.

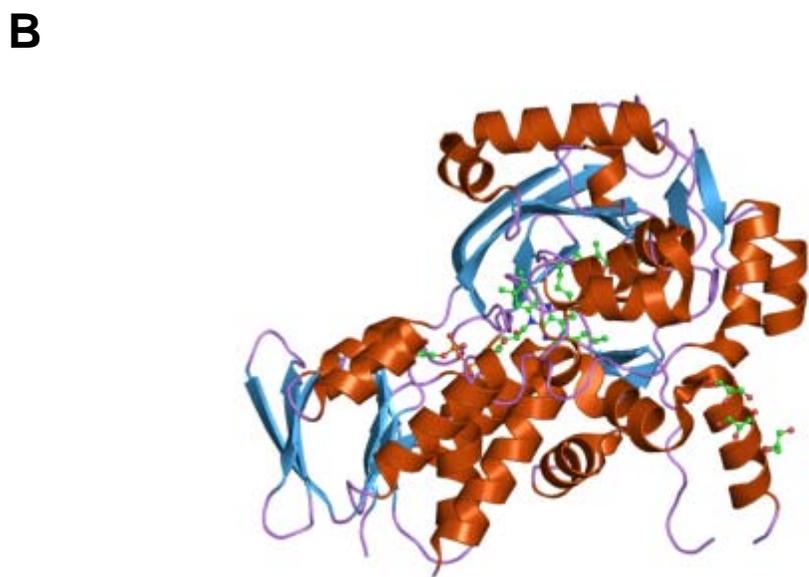
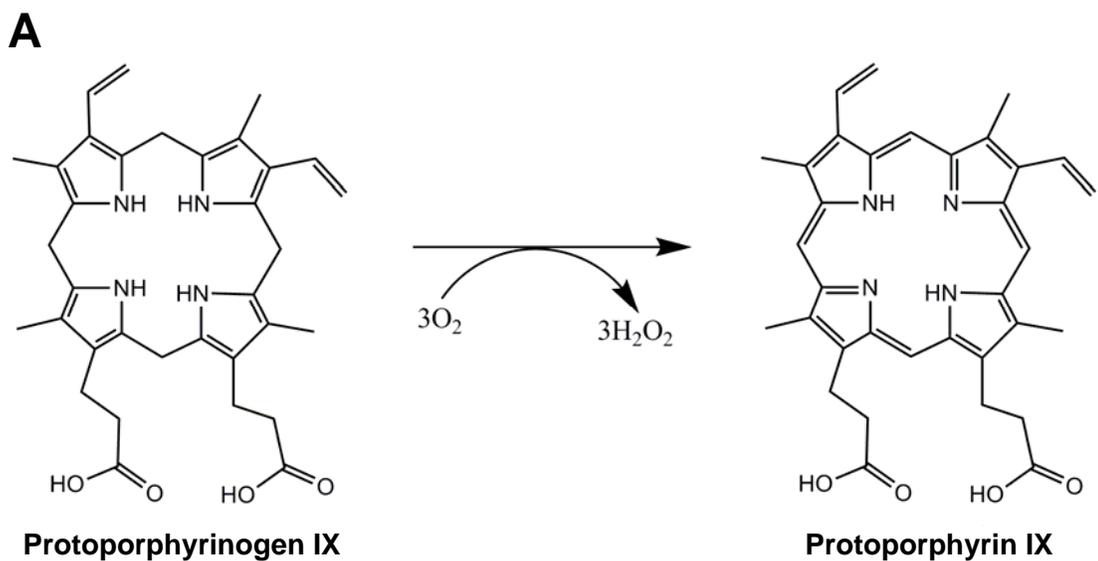


Figure 1.10. Porphyrin formation. **A)** Protoporphyrinogen undergoes a six-electron oxidation to generate the planar macrocycle of protoporphyrin. In organisms containing HemY, this mechanism requires molecular oxygen as the electron acceptor. **B)** Crystal structure of *Myxococcus xanthus* protoporphyrinogen oxidase (HemY) showing the FAD cofactor in the center.

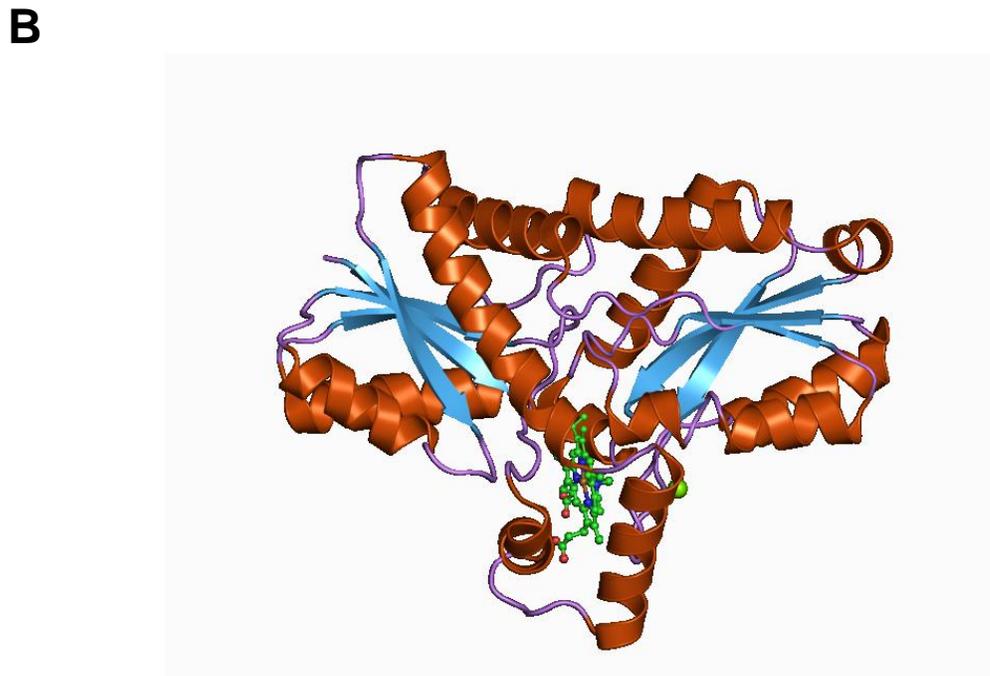
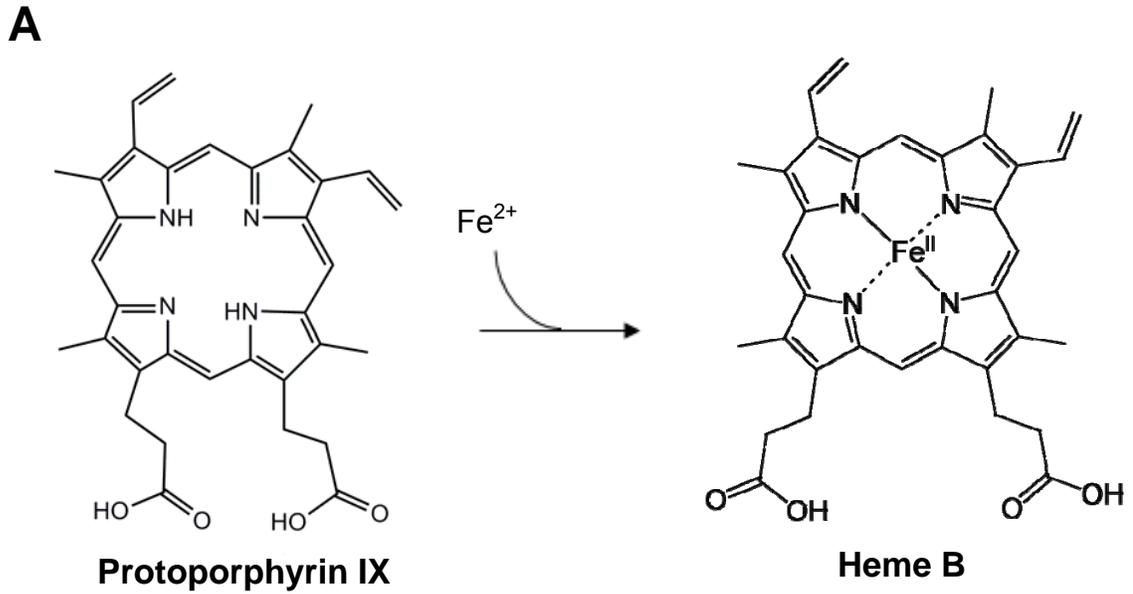


Figure 1.11. Final step in heme synthesis, the chelation of iron. **A)** Ferrous iron is inserted by the enzyme ferrochelatase, producing the principle form of heme. **B)** Crystal structure of *Bacillus subtilis* ferrochelatase showing the position of a substrate analog in wireframe.

CHAPTER 2

IDENTIFICATION OF *E. COLI* HEMG AS A NOVEL, MENADIONE- DEPENDENT FLAVODOXIN WITH PROTOPORPHYRINOGEN OXIDASE ACTIVITY ¹

¹ Boynton TO, Daugherty LE, Dailey TA, Dailey HA 2009. *Biochemistry*. 48:6705-6711.
- Reprinted here with permission of the publisher.

Abstract

Protoporphyrinogen oxidase (PPO, EC 1.3.3.4) catalyzes the six electron oxidation of protoporphyrinogen IX to the fully conjugated protoporphyrin IX. Eukaryotes and Gram-positive bacteria possess an oxygen-dependent, FAD-containing enzyme for this step while the majority of Gram-negative bacteria lack this oxygen-dependent PPO. In *E. coli*, PPO activity is known to be linked to respiration and the quinone pool. In *E. coli* SASX38 the knockout of *hemG* causes loss of measurable PPO activity. HemG is a small soluble protein typical of long chain flavodoxins. Herein purified recombinant HemG was shown to be capable of a menadione-dependent conversion of protoporphyrinogen IX to protoporphyrin IX. Electrochemical analysis of HemG revealed similarities to other flavodoxins. Interestingly, HemG, a member of a class of the long chain flavodoxin family that is unique to the γ -proteobacteria, possesses a 22 residue sequence that, when transferred into *E. coli* flavodoxin A, produces a chimera that will complement an *E. coli* *hemG* mutant, indicating that this region confers PPO activity on the flavodoxin. These findings reveal a previously unidentified class of PPO enzymes that do not utilize oxygen as an electron acceptor thereby allowing γ -proteobacteria to synthesize heme in both aerobic and anaerobic environments.

Introduction

Heme is a highly versatile small organic compound that is a key component of most living organisms. As a cofactor it functions in diverse reactions ranging from a gas sensor and transporter of diatomic gasses to serving as an electron carrier in redox-linked reactions such as is found in mixed-function oxidases, catalases or electron transport chains (29, 31, 32). In addition, it is now recognized that it is also a regulatory molecule. It has been shown to manage protein expression at the transcriptional (41, 42) and translational levels (26) as well as through protein targeting and degradation (19), it participates in regulating circadian rhythm, lipid metabolism, and gluconeogenesis (35, 36, 45) and in prokaryotic organisms heme may even serve to alter pathogenesis by triggering decreased virulence levels and evasion of the host immune response (4, 43).

The biosynthesis of heme has been studied in both eukaryotes and prokaryotes. It is a tightly regulated metabolic pathway consisting of seven enzymatic steps from a precursor compound 5-aminolevulinic acid to the end product, protoheme (30). In humans and animals defects in any of these seven enzymes result in clinically distinct diseases, porphyrias, due to an accumulation of pathway intermediates (18). While the system has been extensively studied and is relatively well understood in eukaryotes (1), the pathway is less well-defined in bacteria and archaea (33). Two steps in particular, the penultimate and antepenultimate, have characterized enzymes that require molecular oxygen as an electron acceptor in eukaryotes. In higher organisms and a handful of

Gram-negative organisms such as *Myxococcus xanthus*, the penultimate enzyme, protoporphyrinogen oxidase (PPO), exists as a three domain, membrane-associated enzyme that utilizes an FAD cofactor (7, 25) and converts oxygen to hydrogen peroxide during the reaction (scheme 2.1). This enzyme belongs to an FAD superfamily of proteins that also contains monoamine oxidases and phytoene desaturases (11). A soluble form of the enzyme named HemY is found in Gram-positive bacteria that produce heme and apart from its solubility, it is virtually the same as the eukaryotic enzyme (12). Not surprising, these oxygen-dependent enzymes have not been found in facultative or anaerobic bacteria although it is clear that some mechanism to catalyze this reaction must exist. The current work focuses on this penultimate step in the facultative bacterium *E. coli*.

Investigations by the Jacobs group in the 1970's found that PPO enzymatic activity was found in crude cellular extracts of *E. coli* and was reported to be membrane-associated (23). The activity was intimately linked to cellular respiration and increased in the presence of terminal electron acceptors such as fumarate, nitrate, or oxygen (20-22). Respiratory enzymes using these substrates all share one thing in common: the usage of quinones, specifically menaquinone-8, as an electron donor (17, 44). These same studies also showed that menaquinone-deficient *E. coli* were deficient in PPO activity (23).

Mutagenesis of *E. coli* by the Sasarman group resulted in the production of one mutant, SASX38, that was deficient in PPO activity (37). The gene, named *hemG*, was mapped and later annotated as being responsible for the penultimate

step in heme biosynthesis (38). Sequence analysis of the encoded protein reveals it to be a member of the protein family known as long chain flavodoxins. These are small electron transfer proteins containing an FMN cofactor. They are distinct from the small chain flavodoxins in that they possess an additional loop inserted in the fifth beta strand. This insert has previously been implicated in specificity and function between redox partners (27, 28). To date, however, there are no published data for the expression, purification and characterization of HemG. Of particular note, because of HemG's resemblance to flavodoxins, it has been generally assumed that it may simply function as an electron carrier for a bonafide PPO enzyme, and PPO enzymatic activity has never been reported for HemG (9).

In the current study the HemG protein was expressed, purified and characterized. In addition the nature of the *hemG* mutation in *E. coli* SASX38 was identified. EPR-monitored redox titrations of HemG were performed to examine the FMN cofactor's role in electron transfer and a menadione-dependent PPO activity was identified and characterized. The possible function for the long chain insert loop, which is unique to HemG of the γ -proteobacteria, was also examined. The data presented unequivocally demonstrates that HemG functions as a protoporphyrinogen oxidase and quinone reductase and that the PPO activity is conveyed by the long chain insert loop.

Materials and Methods

Bacterial strains and constructed plasmids.

The *hemG* gene was amplified from *E. coli* genomic DNA and cloned into the *NheI* and *HindIII* sites of the 6x-His tag vector pTrchisA (Invitrogen, Carlsbad, CA, USA). The resulting plasmid pTHHemG was then tested for functionality by transforming into the SASX38 cell line that contains a knockout of *hemG*.

Overexpression of recombinant protein was carried out in JM109 cells.

Bdellovibrio bacteriovorus Bd2899 was cloned identically to *E. coli hemG*. The chimeric FldG plasmid was constructed by first cloning the *fldA* gene of *E. coli* into pTrchisA and then Quickchange (Stratagene, Jolla, CA, USA) mutagenesis was used to swap the 66 nucleotide insert region to that of *hemG*.

Protein expression, purification, and characterization.

Cells containing pTHHemG were grown for 7 hours at 30 °C in 100 mL Circlegrow (MP Biomedical, Solan, OH, USA) containing 50 µg/mL ampicillin and then transferred to 1 L of the same media and grown for 22 hours at 30° C. Six hours prior to harvesting, the cells were supplemented with riboflavin at a final concentration of 0.75 mg/mL. Cells were then collected by centrifugation and resuspended in a solubilization buffer containing 50 mM Tris-MOPS pH 8.1, 100 mM KCl, 1% sodium cholate, and 10 µg/ml PMSF. The resulting suspension was sonicated three times for 30 s each and centrifuged at 100,000 x g for 30 minutes after which the supernatant was applied to a column containing HisPur Cobalt affinity resin (Thermo Sci., Rockford, IL, USA). The protein was then washed

with solubilization buffer containing 15 mM imidazole, and eluted in solubilization buffer containing 300 mM imidazole and 10% glycerol. Concentration determination and spectrophotometric analysis was carried out using a Cary 1G scanning spectrophotometer (Varian, Palo Alto, CA, USA). SDS-page was done using Bio-Rad 4-20% Tris HCl ready made gels (Bio-Rad, Hercules, CA, USA). FPLC was done using an AktaPrime machine equipped with a Hi-prep Sephacryl S-300 column (GE healthcare, Piscataway, NJ, USA) using solubilization buffer.

EPR-monitored Redox Titrations.

Redox titrations were carried out at 25° C using a solution containing 100 μ M HemG in solubilization buffer (pH 7.0) and the following redox mediator dyes at 50 μ M each: Methyl viologen, benzyl viologen, neutral red, safranin O, anthraquinone-2-sulphonate, phenosafranin, anthraquinone-1,5-disulphonate, 2-OH-1,4-naphthoquinone, indigo-disulphonate, and methylene blue. The reductant and oxidant used were sodium dithionite (20 mg/ml) and potassium ferrocyanide (50 mg/ml) respectively. HemG was fully reduced anaerobically and titration points were taken from potentials ranging from \sim -450 to -150 mV by adding small amounts of oxidant. After each addition of oxidant, the sample was stirred and allowed to come to a stable potential, and 250 μ l of sample was transferred to an anaerobic EPR tube and frozen in liquid nitrogen. Potentials were obtained using a Ag/AgCl electrode. Reported potentials were then recalculated with respect to NHE values. EPR data was then collected at 9.18

gHz with a microwave power of 20 μ W and the intensity of the semiquinone signal was quantified using double integration.

PPO assay.

PPO activity was monitored as previously described (39). Briefly, substrate was prepared by solubilizing 6 mg protoporphyrin IX (Porphyrin Products, Logan, UT, USA) with ~5 drops of 30% ammonium hydroxide and then diluted with 5 mL 10 mM KOH containing 20% ethanol followed by 5 mL 10 mM KOH. The resulting solution was then reduced in the dark by addition of 30g sodium amalgam made from 3.5g sodium and 75g mercury under nitrogen. Once reduced, removal of excess protoporphyrin was achieved by filtration through glass wool. The resulting protoporphyrinogen IX was used immediately. Concentration was determined by allowing unused substrate to oxidize in light and quantifying porphyrin absorbance. Reaction mixtures consisted of 50 mM NaH_2PO_4 (pH 8.0), 0.2% (w/v) Tween 20, 2.5 mM glutathione, 100 nM HemG and varying amounts of protoporphyrinogen IX and menadione. Experiments were also done using mesoporphyrinogen and coproporphyrinogen prepared as for protoporphyrinogen. Activity was monitored at 37° C by detecting the accumulation of porphyrin fluorescence at ~545nm using Synergy HTI plate reader (BioTek, Winooski, VT, USA) with 528 nm and 545nm bandpass filters and a 550 nm highpass filter on the excitation light and 635 nm bandpass filter on the emission light. Data (v vs. $[S]$) was then fitted to equation 1,

$$v = V_{\text{app}}[S]/(K_{\text{app}}+[S]),$$

where V_{app} is the apparent maximum rate and K_{app} is the apparent Michaelis-Menton constant.

Results

Protein expression and characterization.

Sequence analysis of HemG from *E. coli* revealed that it is a member of a protein family found in the γ -proteobacteria that is part of a larger COG known as long chain flavodoxins (figure 2.1) (5). The *hemG* gene of *E. coli* was cloned, then expressed and purified as a six histidine-tagged recombinant protein, and the yield was approximately 20 mg protein/L culture. A band corresponding to approximately 22 kDa was observed on SDS-PAGE (figure 2.2A), which is in agreement with the theoretical value of 22.5 kDa calculated from the amino acid sequence. Maldi TOF/TOF mass spectroscopy identified this band as HemG. A faint band located at approximately 40 kDa was identified as HemG and may represent a dimer of the protein. The cofactor was determined to be similar in size to FMN by ESI MS. The recombinant expression plasmid rescued the HemG-deficient strain of *E. coli* SAS38X, which alone cannot grow without the presence of exogenous heme.

The UV-visible spectrum of recombinant HemG possesses the expected features associated with protein and the FMN cofactor (figure 2.2B) along with a feature at 416 nm corresponding to the Soret band of porphyrins, but it is yet to be determined if this is due to bound protoporphyrin IX. This last feature is distinct between the spectra of HemG and that of its closest relative *E. coli*

flavodoxin A, where no band is observed. Determination of subunit composition by FPLC identified a peak corresponding to ~108 kDa. Based on this, HemG would then exist in solution as a homotetramer.

EPR-monitored redox titrations.

The EPR spectra of HemG revealed a signal typical of the FMN semiquinone radical centered at $g = 2.00$. Using sodium dithionite (potential of approximately -450 mV at pH 7.0) it was possible to fully reduce the FMN. Titrations of HemG included methyl and benzyl viologens, which possess EPR signals that overlap with the FMN semiquinone radical, so the data were corrected by subtracting out the signal of the viologens alone. Data obtained at stable and reliable potentials from the double integration of interpretable spectra are shown in figure 2.2C. The data was fitted to the Nernst equation (equation 2) for two consecutive one electron redox steps:

$$[Fidsq]/[FIdtotal] = 1/\{1 + \exp(F(E-E2)/RT) + \exp(-F(E-E1)/RT)\},$$

and normalized to the fitted maximum. E2 and E1 represent the first and second reduction steps of HemG and were determined to be -241 mV and -412 mV respectively. These values (also shown in figure 2.2C) are comparable to those of standard flavodoxins (40). In typical flavodoxin reactions, the semiquinone/hydroxyquinone couple is stabilized so that flavodoxins effectively function as one electron donor/acceptors. Whether HemG functions as a one or two electron redox partner in the PPO reaction is not known. Accurate

determination of this is complicated by spectral overlap of porphyrin product with the flavin spectra.

PPO assays.

To investigate the role of HemG in the conversion of protoporphyrinogen IX to protoporphyrin IX, HemG was assayed in the presence and absence of the potential electron acceptor, menadione, a soluble analog of menaquinone-8. The results of this assay are shown in figure 2.3A. Upon addition of excess menadione the activity of HemG increased reaching a rate of roughly half of that found with human PPO (3.71 and 8.54 min⁻¹, respectively). A small amount of light-induced autooxidation occurred as indicated by the assay containing protoporphyrinogen IX only.

To provide additional evidence that HemG is a bonafide PPO, we carried out the menadione-dependent assay with the closest COG relative of HemG in *E. coli*, flavodoxin A, which cannot complement SASX38 cells, and FMN alone. Both yielded data equivalent to the rate of autooxidation seen from substrate only (results not shown). Given the close similarity of HemG and FldA, these data indicate that the PPO activity of HemG is not an artifact or side reaction catalyzed by flavodoxins.

As an additional test, we cloned, expressed and purified the long chain flavodoxin BD2899 from *Bdellovibrio bacteriovorus* that is currently annotated as a PPO. This protein has high homology to HemG including the long chain insert region (fig. 2B). However, this organism possesses the typical HemY PPO in an

operon with other heme synthesizing enzyme genes and there exist no experimental data to support a role for this HemG-related flavodoxin in heme synthesis. Interestingly, while the insert region of the *Bd. bacteriovorus* protein is more similar to that found in HemG than in FldA, it lacks an acidic residue at position 136 and a methionine at position 144 that are conserved in all other HemGs. When expressed in the *E. coli* SASX38, this protein did not complement the *hemG* deletion indicating that it is not a bonafide PPO.

HemG PPO activity was found to be specific for protoporphyrinogen. Neither mesoporphyrinogen nor coproporphyrinogen were oxidized by HemG. Kinetic parameters for HemG were determined for both protoporphyrinogen IX and menadione. These data are shown in figure 2.3B, and the curves were fitted to the Michaelis equation (eq. 2). The K_m and K_{cat} for protoporphyrinogen IX were found to be 7.0 μM and 17.52 min^{-1} respectively, and the K_m and K_{cat} for menadione were 3.76 μM and 16.87 min^{-1} . The K_{cat}/K_m for HemG was 2.5. The kinetic parameters, along with those of other well-classified PPOs are shown in table 2.1 (8, 10, 14, 24). The K_{cat} of HemG is slightly higher than most PPO enzymes, with the exception of *Desulfovibrio gigas* PPO, and the K_m for substrate falls within the range of other PPOs. HemG also has a high affinity for its electron acceptor, menadione. By comparison, the reported K_m of mouse PPO for oxygen is 125 μM (14). HemG's affinity for its electron acceptor could allow function when only a small amount of the menaquinone pool is oxidized, as in the case of aerobic respiration, where ubiquinone is preferentially oxidized by terminal oxidases.

Role of the long chain insert loop.

The *E. coli* strain SASX38 lacks functional PPO activity and was reported to be deficient in HemG. However, the nature of this deletion was never identified. Sequence analysis of the *hemG* gene in SASX38 cells revealed a deletion from nucleotides 152 to 470. This results in a predicted HemG protein that remains in frame but is lacking 106 internal amino acids, including the long chain insert loop. This long chain loop has been previously reported as being responsible for activity, specificity, and/or recognition of redox partners (27, 28). Unique to HemG, this region is predicted to contain an alpha helix (figure 2.1). Since long chain insert loops have previously been implicated in specificity and function of flavodoxins, this region of HemG was examined by mutating the long chain insert loop of *E. coli* flavodoxin A to that of HemG. The resulting chimeric clone, FldG, was expressed in SASX38 cells deficient in HemG activity and complementation was obtained indicating that *in vivo* the chimera can contain PPO activity. Expression and purification of the FldG protein was performed as described for HemG. Spectral analysis revealed no feature at 416 nm, and PPO assays of purified FldG showed no activity in the presence or absence of menadione.

Discussion

The lack of an identifiable protoporphyrinogen oxidase in most Gram-negative bacteria represents a significant gap in our current understanding of prokaryotic heme synthesis. These organisms require heme and contain all

other enzymes within the pathway so clearly an enzyme with PPO activity must exist. Dependence upon non-enzymatic conversion of protoporphyrinogen IX *in vivo* would be untenable since it would be unregulated and the accumulated highly reactive excess product would be toxic to the cell. Data presented herein clearly demonstrates that HemG is a previously uncharacterized form of PPO that is distinct from currently identified oxygen-dependent bacterial HemY PPOs. Unlike other known PPOs, HemG is a flavodoxin-based enzyme that employs menaquinone, rather than oxygen, as an electron acceptor. This is appropriate since *E. coli* is a facultative anaerobe and would be unable to synthesize heme in the absence of oxygen if it possessed a HemY type of PPO.

Due to the hydrophobicity of menaquinone and the nature of these assays, the soluble form menadione was used, which lacks the long aliphatic side chain not involved in redox properties. Our finding that HemG utilizes menadione *in vitro* is consistent with previous research reported by others that was done with cell extracts of *E. coli* (20-23). These authors found that PPO activity was respiratory chain associated (with fumarate, nitrate or oxygen as terminal electron acceptors) and the activity was localized to the membrane fraction. This location would be anticipated since the respiratory chain components, including the menaquinone-8 pool, are membrane localized. Thus, the coupled reaction will be one where the soluble HemG interacts with menaquinone at the membrane.

The proposition that a flavodoxin has evolved to gain function is not without precedent. WrbA, another protein similar to *E. coli* flavodoxin, possesses

a quinone reductase activity much like HemG (16, 34). It appears that this protein represents an evolutionary link between flavodoxins and eukaryotic NAD(P)H/quinone oxidoreductases (3). Like HemG, it also exists as a homotetramer and the regions adjacent to the four FMN binding sites form an active site complex in the center of the quaternary structure (2, 15). Another interesting property of WrbA and other quinone reductases in general is their ability to transfer two electrons in a ping-pong mechanism. This avoids the generation of semiquinone radicals that can lead to harmful reactive oxygen species (13). It will be of interest to determine if HemG also has the ability to function as a two electron transfer protein. Such an ability would decrease by half the number of turnovers required to oxidize the porphyrinogen.

The data obtained from the chimeric FldG in which the long chain insert of HemG is substituted for the analogous position of flavodoxin A presents possibly the most intriguing aspect of HemG. By swapping the long chain insert loop out of HemG into *E. coli* flavodoxin A, we were able to complement a cell line that must have PPO activity to grow under respiring conditions. This clearly indicates a transfer of function from HemG to a flavodoxin normally only capable of transferring electrons between redox partners. No *in vitro* activity with the chimera could be seen in the presence of menadione, but *in vivo* complementation can only occur if function is present. Lack of *in vitro* activity probably results from the absence of a normal flavodoxin A redox partner to accept electrons during the reaction. The most compelling reason for this is that the loop region is responsible for binding and oxidation of the tetrapyrrole but

does not confer reoxidation of the cofactor by quinones. From the complementation results, it is clear that FldG is still able to interact with a physiological redox partner of FldA in vivo so that it can turn over.

Computational modeling of HemG results in a predicted structure that is similar to the typical flavodoxin motif. The predicted structure generated using 3D-JIGSAW is shown in figure 2.4 (6). The FMN binding pocket is formed by the 50's and 90's loops, and the long chain insert loop extends behind this pocket presenting the possible interaction site for FMN and substrate, further supporting the role of the long chain insert loop.

Interestingly, the conversion of protoporphyrinogen IX to protoporphyrin IX appears to be an activity where Gram-negative bacteria have scavenged and adapted other proteins to do the job. *Myxococcus* species contain an FAD-containing, oxygen dependent enzyme essentially identical to eukaryotic PPOs (7) whereas *Desulfovibrio gigas* is reported to carry out catalysis using a heterotrimer complex (24) that remains to be characterized. HemG itself is only present in α -proteobacteria and so the identity of PPO in the remaining Gram negative organisms remains unknown. Thus, HemG's role as a protoporphyrinogen oxidase fills only a portion of the gap in our understanding of this point of the pathway.

There are many unanswered questions concerning HemG, its function, and the role of the long chain insert. Since it is a relatively small protein, it will be of interest to determine how it binds the substrates protoporphyrinogen and

menaquinone and catalyzes the reaction. Ongoing studies aimed at determining the X-ray crystal structure of HemG will help to answer these questions.

Supporting Information Available

SAS38X complementation was identified by the appearance of colony growth after transformation of appropriate plasmid. In an effort to give some quantitation to this, the following growth curves were generated: SAS38X, SASX38 with heme, SAS38X transformed with pTHHemG, and SAS38x transformed with FldG (figure S2.1).

References

1. **Ajioka RS, Phillips JD, & Kushner JP.** 2006. Biosynthesis of heme in mammals. *Biochim Biophys Acta.* 1763:723-736.
2. **Andrade SL, Patridge EV, Ferry JG, & Einsle O.** 2007. Crystal structure of the NADH:quinone oxidoreductase WrbA from *Escherichia coli*. *J Bacteriol.* 189:9101-9107.
3. **Carey J, et al.** 2007. WrbA bridges bacterial flavodoxins and eukaryotic NAD(P)H:quinone oxidoreductases. *Protein Sci.* 16:2301-2305.
4. **Cho HY, Cho HC, Kim YM, Oh JI, & Kang BS.** 2009. Structural insight into the heme-based redox sensing by DosS from *Mycobacterium tuberculosis*. *J Biol Chem.*
5. **Cole C, Barber JD, & Barton GJ.** 2008. The Jpred 3 secondary structure prediction server. *Nucleic Acids Res.* 36:W197-201.
6. **Contreras-Moreira B & Bates PA.** 2002. Domain fishing: a first step in protein comparative modelling. *Bioinformatics.* 18:1141-1142.
7. **Corradi HR, et al.** 2006. Crystal structure of protoporphyrinogen oxidase from *Myxococcus xanthus* and its complex with the inhibitor acifluorfen. *J Biol Chem.* 281:38625-38633.
8. **Corrigall AV, et al.** 1998. Purification of and kinetic studies on a cloned protoporphyrinogen oxidase from the aerobic bacterium *Bacillus subtilis*. *Arch Biochem Biophys.* 358:251-256.
9. **Dailey HA.** 2002. Terminal steps of haem biosynthesis. *Biochem Soc Trans.* 30:590-595.

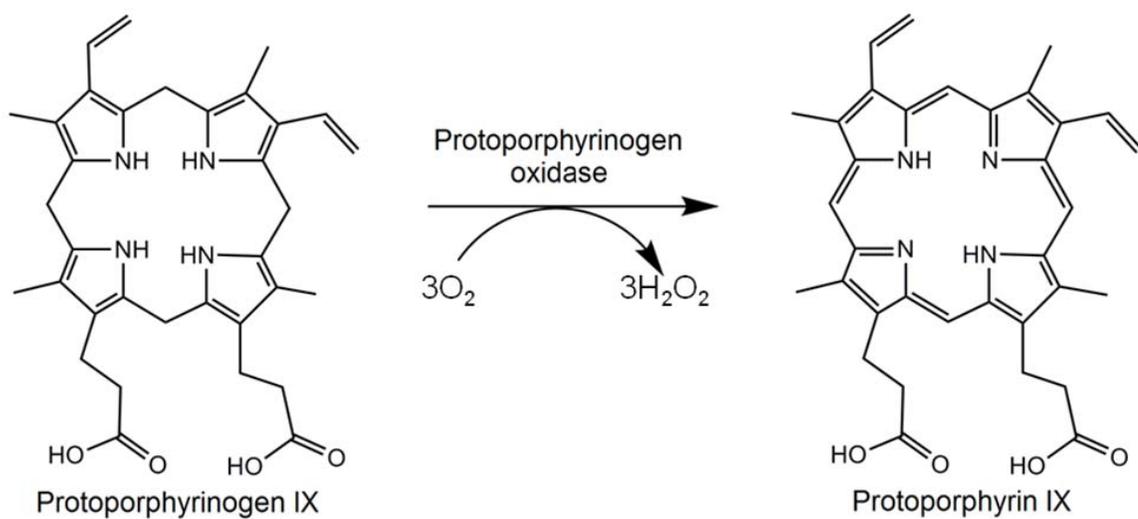
10. **Dailey TA & Dailey HA.** 1996. Human protoporphyrinogen oxidase: expression, purification, and characterization of the cloned enzyme. *Protein Sci.* 5:98-105.
11. **Dailey TA & Dailey HA.** 1998. Identification of an FAD superfamily containing protoporphyrinogen oxidases, monoamine oxidases, and phytoene desaturase. Expression and characterization of phytoene desaturase of *Myxococcus xanthus*. *J Biol Chem.* 273:13658-13662.
12. **Dailey TA, Meissner P, & Dailey HA.** 1994. Expression of a cloned protoporphyrinogen oxidase. *J Biol Chem.* 269:813-815.
13. **Deller S, Macheroux P, & Sollner S.** 2008. Flavin-dependent quinone reductases. *Cell Mol Life Sci.* 65:141-160.
14. **Ferreira GC & Dailey HA.** 1988. Mouse protoporphyrinogen oxidase. Kinetic parameters and demonstration of inhibition by bilirubin. *Biochem J.* 250:597-603.
15. **Gorman J & Shapiro L.** 2005. Crystal structures of the tryptophan repressor binding protein WrbA and complexes with flavin mononucleotide. *Protein Sci.* 14:3004-3012.
16. **Grandori R, et al.** 1998. Biochemical characterization of WrbA, founding member of a new family of multimeric flavodoxin-like proteins. *J Biol Chem.* 273:20960-20966.
17. **Hastings SF, et al.** 1998. Identification of a stable semiquinone intermediate in the purified and membrane bound ubiquinol oxidase-cytochrome bd from *Escherichia coli*. *Eur J Biochem.* 255:317-323.

18. **Hift RJ, et al.** 1997. Variegate porphyria in South Africa, 1688-1996--new developments in an old disease. *S Afr Med J.* 87:722-731.
19. **Iwai K, et al.** 1998. Iron-dependent oxidation, ubiquitination, and degradation of iron regulatory protein 2: implications for degradation of oxidized proteins. *Proc Natl Acad Sci U S A.* 95:4924-4928.
20. **Jacobs JM & Jacobs NJ.** 1977. The late steps of anaerobic heme biosynthesis in *E. coli*: role for quinones in protoporphyrinogen oxidation. *Biochem Biophys Res Commun.* 78:429-433.
21. **Jacobs NJ & Jacobs JM.** 1975. Fumarate as alternate electron acceptor for the late steps of anaerobic heme synthesis in *Escherichia coli*. *Biochem Biophys Res Commun.* 65:435-441.
22. **Jacobs NJ & Jacobs JM.** 1976. Nitrate, fumarate, and oxygen as electron acceptors for a late step in microbial heme synthesis. *Biochim Biophys Acta.* 449:1-9.
23. **Jacobs NJ & Jacobs JM.** 1978. Quinones as hydrogen carriers for a late step in anaerobic heme biosynthesis in *Escherichia coli*. *Biochim Biophys Acta.* 544:540-546.
24. **Klemm DJ & Barton LL.** 1987. Purification and properties of protoporphyrinogen oxidase from an anaerobic bacterium, *Desulfovibrio gigas*. *J Bacteriol.* 169:5209-5215.
25. **Koch M, et al.** 2004. Crystal structure of protoporphyrinogen IX oxidase: a key enzyme in haem and chlorophyll biosynthesis. *EMBO J.* 23:1720-1728.

26. **Kramer G, Cimadevilla JM, & Hardesty B.** 1976. Specificity of the protein kinase activity associated with the hemin-controlled repressor of rabbit reticulocyte. *Proc Natl Acad Sci U S A.* 73:3078-3082.
27. **Lopez-Llano J, et al.** 2004. The long and short flavodoxins: I. The role of the differentiating loop in apoflavodoxin structure and FMN binding. *J Biol Chem.* 279:47177-47183.
28. **Lopez-Llano J, et al.** 2004. The long and short flavodoxins: II. The role of the differentiating loop in apoflavodoxin stability and folding mechanism. *J Biol Chem.* 279:47184-47191.
29. **Lukin JA & Ho C.** 2004. The structure--function relationship of hemoglobin in solution at atomic resolution. *Chem Rev.* 104:1219-1230.
30. **Medlock AE & Dailey HA.** 2009. Regulation of mammalian heme biosynthesis. *Tetrapyrroles: Birth, Life and Death*, eds Warren MJ & Smith AG (Landes Bioscience), pp 116-127.
31. **Michel H, Behr J, Harrenga A, & Kannt A.** 1998. Cytochrome c oxidase: structure and spectroscopy. *Annu Rev Biophys Biomol Struct.* 27:329-356.
32. **Munro AW, Girvan HM, McLean KJ, Cheesman MR, & Leys D.** 2009. Heme and Hemoproteins. *Tetrapyrroles: Birth, Life and Death*, eds Warren MJ & Smith AG (Landes Bioscience), pp 160-183.
33. **Panek H & O'Brian MR.** 2002. A whole genome view of prokaryotic haem biosynthesis. *Microbiology.* 148:2273-2282.

34. **Patridge EV & Ferry JG.** 2006. WrbA from *Escherichia coli* and *Archaeoglobus fulgidus* is an NAD(P)H:quinone oxidoreductase. *J Bacteriol.* 188:3498-3506.
35. **Raghuram S, et al.** 2007. Identification of heme as the ligand for the orphan nuclear receptors REV-ERBalpha and REV-ERBbeta. *Nat Struct Mol Biol.* 14:1207-1213.
36. **Raspe E, et al.** 2002. Identification of Rev-erbalpha as a physiological repressor of apoC-III gene transcription. *J Lipid Res.* 43:2172-2179.
37. **Sasarman A, et al.** 1979. Mapping of a new hem gene in *Escherichia coli* K12. *J Gen Microbiol.* 113:297-303.
38. **Sasarman A, et al.** 1993. Nucleotide sequence of the hemG gene involved in the protoporphyrinogen oxidase activity of *Escherichia coli* K12. *Can J Microbiol.* 39:1155-1161.
39. **Shepherd M & Dailey HA.** 2005. A continuous fluorimetric assay for protoporphyrinogen oxidase by monitoring porphyrin accumulation. *Anal Biochem.* 344:115-121.
40. **Steensma E, Heering HA, Hagen WR, & Van Mierlo CP.** 1996. Redox properties of wild-type, Cys69Ala, and Cys69Ser *Azotobacter vinelandii* flavodoxin II as measured by cyclic voltammetry and EPR spectroscopy. *Eur J Biochem.* 235:167-172.
41. **Sun J, et al.** 2002. Hemoprotein Bach1 regulates enhancer availability of heme oxygenase-1 gene. *EMBO J.* 21:5216-5224.

42. **Tahara T, Sun J, Igarashi K, & Taketani S.** 2004. Heme-dependent up-regulation of the alpha-globin gene expression by transcriptional repressor Bach1 in erythroid cells. *Biochem Biophys Res Commun.* 324:77-85.
43. **Torres VJ, et al.** 2007. A *Staphylococcus aureus* regulatory system that responds to host heme and modulates virulence. *Cell Host Microbe.* 1:109-119.
44. **Wissenbach U, Kroger A, & Uden G.** 1990. The specific functions of menaquinone and demethylmenaquinone in anaerobic respiration with fumarate, dimethylsulfoxide, trimethylamine N-oxide and nitrate by *Escherichia coli*. *Arch Microbiol.* 154:60-66.
45. **Yin L, et al.** 2007. Rev-erbalpha, a heme sensor that coordinates metabolic and circadian pathways. *Science.* 318:1786-1789.



Scheme 2.1. Reaction catalyzed by oxygen-dependent protoporphyrinogen oxidase.

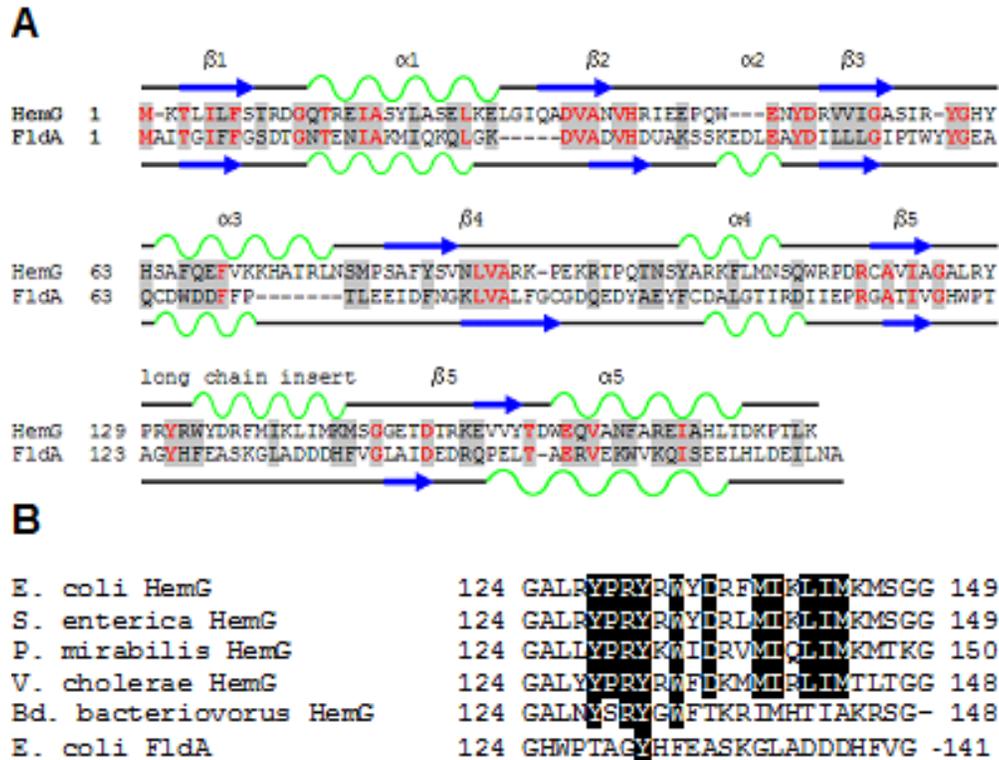


Figure 2.1. Sequence analysis of HemG. **A.** Sequence and secondary structure similarity of HemG with Flavodoxin A of *E. coli*. HemG is predicted by Jpred (5) to possess a typical flavodoxin-fold motif but with one notable difference, the presence of an alpha helical domain within the long chain insert loop. In the diagram similar residues are highlighted in grey, with conserved residues in red. Alpha helices and beta sheets are shown in green and blue respectively for each protein. **B.** Sequence alignment of the long chain insert loop of *E. coli* HemG with that of other annotated HemG proteins and FldA. Identical residues within the predicted alpha helix are highlighted. Significant change in this region (i. e. *Bd. bacteriovorus*) results in loss of function.

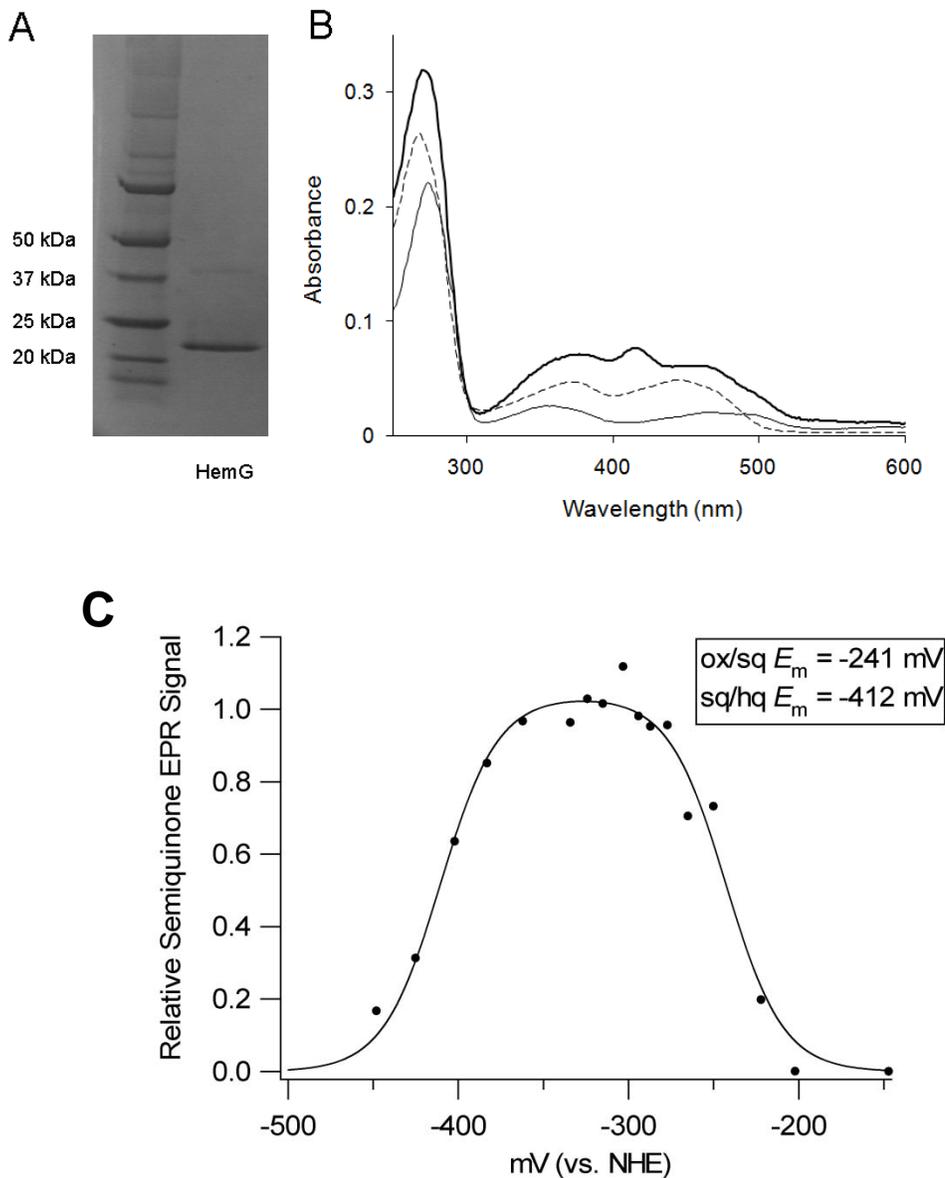


Figure 2.2. Initial characterization of HemG and its cofactor. **A.** SDS-page analysis of purified recombinant HemG. Left lane, molecular weight marker; Right lane, purified HemG. Recombinant HemG possesses a major band of approximately 22 kDa with a faint band seen at approximately 40 kDa. Both bands were identified as HemG by MS/MS analysis. **B.** UV-vis spectrum analysis of purified recombinant HemG. The holoprotein (top line) exhibits characteristic FMN features (dashed line), as well as the presence of a peak at 416 nm which is lacking in *E. coli* flavodoxin A (bottom line). **C.** EPR-monitored redox titration of HemG. Data were fitted to the Nernst equation for two consecutive one-electron reduction steps (Eq. 1). A relative semiquinone EPR signal of 1 indicates a sample where all FMN present is in the semiquinone state. Inset: Calculated redox potentials for the oxidized/semiquinone and semiquinone/hydroxyquinone couples of cofactor as determined by Eq. 1.

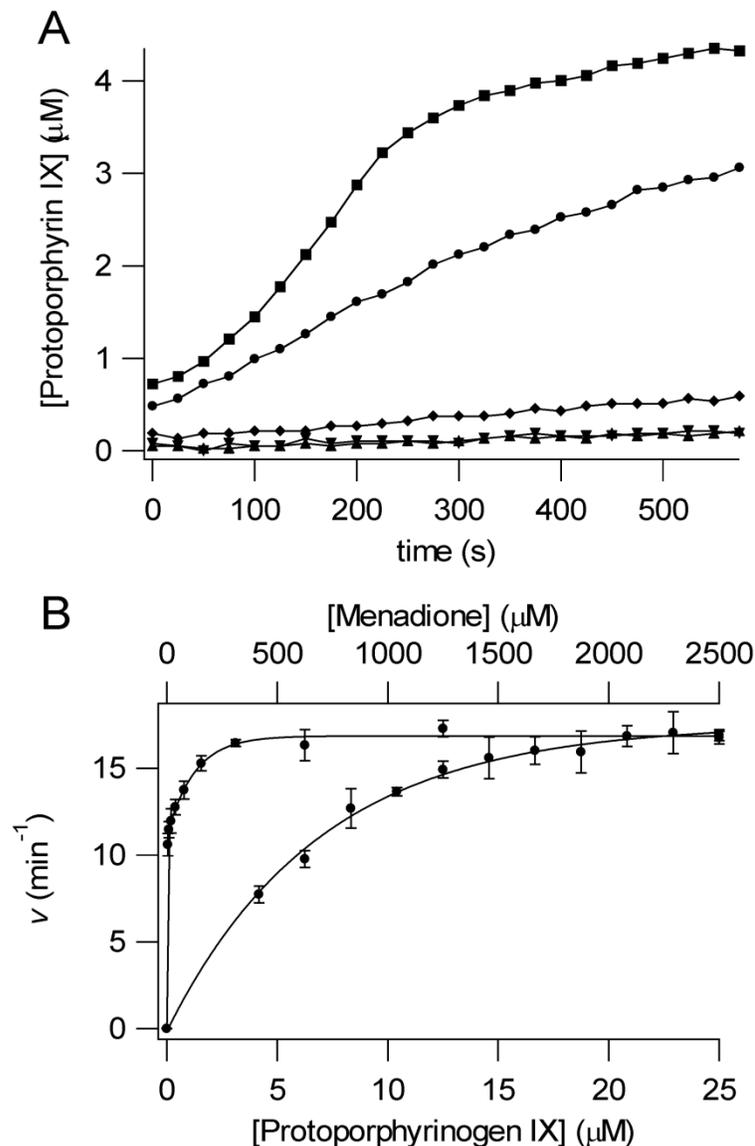


Figure 2.3. Enzymatic conversion of protoporphyrinogen IX to protoporphyrin IX by HemG. **A**. PPO assays are shown for human PPO (■), 100 nM HemG + 1 mM menadione (●), 100 nM HemG (◆), 1 mM menadione (▼), and protoporphyrinogen alone (▲). In the absence of menadione HemG is capable of generating only minimal amounts of product when compared to autooxidation. In the presence of saturating menadione concentrations, enzymatic activity is greatly increased to levels that are comparable to human PPO. **B**. Determination of apparent Michaelis constants for HemG. 100 nM HemG was assayed using varying concentrations of protoporphyrinogen IX (bottom curve) and menadione (top curve). Data were fitted to single rectangular hyperbolae (Eq. 2). K_m and K_{cat} for protoporphyrinogen IX = 7.0 μM and 17.52 min^{-1} ; K_m and K_{cat} for menadione = 3.76 μM and 16.87 min^{-1} .

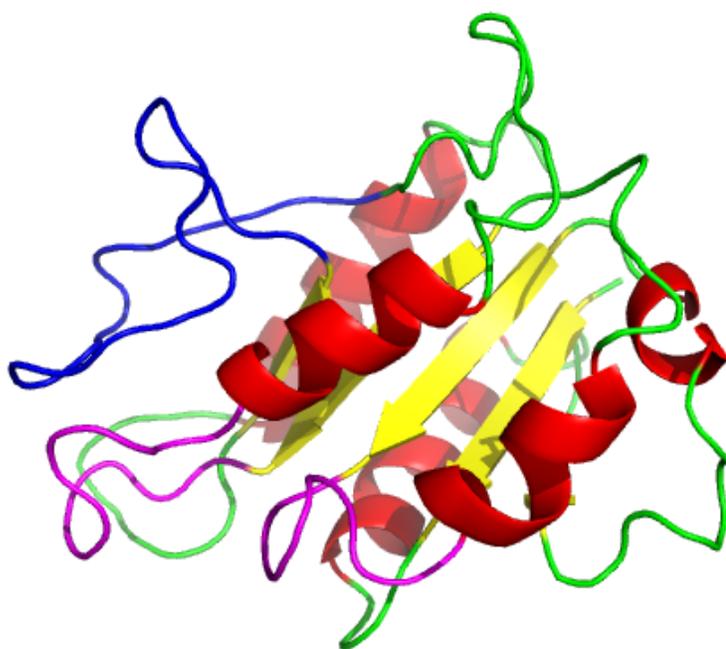


Figure 2.4. Predicted Structure of HemG. Structure was generated using 3D-JIGSAW (6). FMN-binding loops and the long chain insert loop are colored purple and blue respectively. The long chain insert loop sits above the binding pocket formed by the two loops allowing possible interaction with FMN.

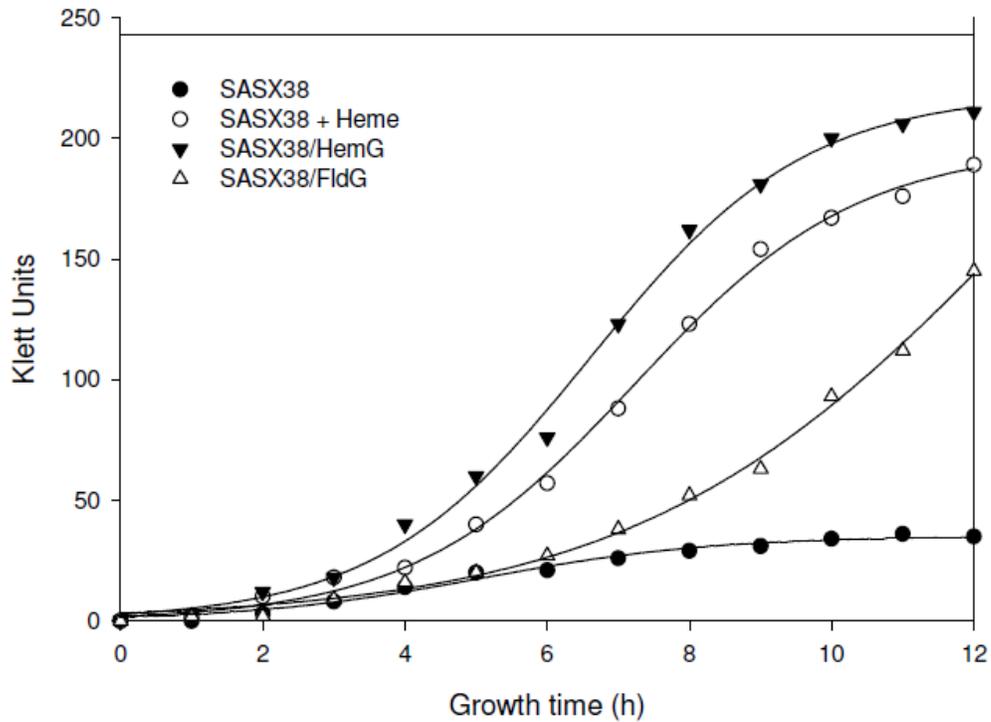


Figure S2.1. Complementation analysis of chimeric FldG. SAS38X cells are able to grow only when given heme or functional protoporphyrinogen oxidase. Cells containing FldG have restored, but slow, growth compared to control cells. Initial growth is due to residual heme in inoculums.

Table 2.1. Comparison of kinetic parameters for HemG and various PPOs

Protein	Km (μM)	Kcat (min^{-1})	Reference
<i>H. sapiens</i> PPO	1.7 / 3.8	10.5 / 5.7	(36)
<i>M. musculus</i> PPO	6.6	7.45	(37)
<i>B. subtilis</i> PPO	1.0	0.05	(35)
<i>D. gigas</i> PPO	21	17.7	(38)
<i>E. coli</i> HemG	7.0	17.52	

CHAPTER 3

REQUIREMENT OF HEMQ FOR PROPER FUNCTION OF HEMY AND HEMH IN THE FIRMICUTES AND ACTINOBACTERIA

Summary

The heme biosynthetic pathway was long thought to have been elucidated, with each step controlled by a characterized enzyme, though multiple forms exist in some cases. Herein we describe evidence for a novel enzyme outside of the classical pathway essential for heme synthesis to occur. This enzyme, HemQ, is present in the Firmicutes and Actinobacteria and is required for the terminal two enzymes in the pathway, protoporphyrinogen oxidase (HemY) and ferrochelatase HemH), to catalyze their reactions *in vivo*. *In vitro*, *Propionobacterium acnes* HemY was shown to contain little detectable activity when assayed alone, but was greatly stimulated by the addition of HemQ to reaction mixtures. Furthermore, synthesis of heme by HemH was not seen unless HemY and HemQ were present. Through the use of this three component system, it was observed that heme synthesis only occurred when protoporphyrinogen, the substrate of HemY, was available, suggesting that the three enzymes function in a complex. The existence of this complex was not supported by gel filtration techniques, but was not ruled out. Although not completely characterized, it is clear that HemQ represents a novel aspect of heme synthesis previously unconsidered. Finally, this enzyme represents an

intriguing antimicrobial target, present in such pathogenic organisms as *Mycobacterium tuberculosis* and *Staphylococcus aureus*. The data presented herein is part of a larger article by Dailey *et al.* previously published in the *Journal of Biological Chemistry*, 285:25978-25986.

Introduction

One of the hallmarks of prokaryotes is astounding diversity. In quantitative terms, a handful of soil has been estimated to contain millions of different species (6). Qualitatively, consider the capacity of these organisms then for genetic diversity. The wide range of metabolic pathways within microorganisms has enabled them to adapt and flourish in all known environments (16). One key example is the group of metabolic pathways that make up tetrapyrrole synthesis. As enzymatic cofactors, tetrapyrroles afford such integral metabolic processes as photosynthesis, aerobic and anaerobic respiration, oxidative stress response, sulfite assimilation, methanogenesis, and many others (11). These cofactors include heme (5), siroheme (17), chlorophyll (11), coenzyme F430 (7) and the highly complex cobalamin (vitamin B12) (13). The latter two are only synthesized in prokaryotes.

Perhaps the most well-characterized of these is heme biosynthesis (12). Like all tetrapyrroles, synthesis begins with the conversion of a precursor compound, d-aminolevulinic acid (ALA), to uroporphyrinogen III (uro-III). While ALA synthesis can occur two different ways in prokaryotes, the next three enzymes are shared by all pathways in all organisms. Next, terminal enzymes shared by both heme and chlorophyll pathways convert uro-III in to

protoporphyrin. Unique to heme synthesis, the final enzyme ferrochelatase (HemH) inserts iron into the macrocycle.

A recent discovery by Tamara A. Dailey in our group has proven that a protein resembling chlorite dismutase, now named HemQ, is present in the Firmicutes and Actinobacteria and is required for the soluble HemY and HemH they contain to function correctly. The initial question of HemY and HemH function in these organisms came from the observation that neither could complement an *E. coli* mutant lacking their respective steps in heme biosynthesis. HemQ was later discovered when the protein sequence of HemH within *Propionobacterium acnes* was seen to contain ~300 C-terminal amino acids unique to this group, and it was subsequently revealed that the coding sequence of *hemH* was fused to that of a hypothetical ORF, now named *hemQ*. Studies showed that a successful complementation of the *E. coli* mutants could only occur if plasmid-born *hemY*, *hemH*, and *hemQ* were all present (herein, *hemY* and *hemH* are used to refer to their respective genes only in organisms containing *hemQ*) (Table 1). This implies that HemY and HemH are not sufficient for their own respective activities. While the exact function of HemQ is not known, we have attempted to characterize the catalytic aspects of HemY and HemH with respect to HemQ essentiality. Presented herein, we have now shown that *in vitro* activity of HemY and HemH only occurs in the presence of HemQ. The existence of a complex between the three proteins was also examined. These observations indicate that enzymes superfluous to our current understand of catalytic steps can still be essential for heme biosynthesis.

Materials and Methods

Cloning, expression, and purification.

In most cases, respective genes were obtained by PCR of genomic DNA (*M. tuberculosis* gDNA gifted by Fred Quinn, University of Georgia and *P. acnes* gDNA gifted by A. Strittmatter, Georg-August-Universitat, Germany) and cloned into the NheI/HindIII sites of pTrcHisA (Initrogen, Carlsbad, CA) and then expressed in *E. coli*. The exceptions to this are *P. acnes hemY*, which was engineered to contain a c-terminal 6x his-tag and cloned into the NcoI/HindIII site of pTrcHisA, and *B. subtilis hemQ*, which was cloned into the NheI/XhoI sites of pTrcHisA. All constructs were verified by direct sequencing. Protein expression and purification of recombinant proteins was done as previously described (1). HemQ referred to as “heme-loaded” indicates that 1 mg of hemin in DMSO was added per mL to ultracentrifuge supernatant prior to being loaded on the the HisPur cobalt chelate column (Pierce, Rockford, IL).

Enzymatic Assays.

Protoporphyrinogen oxidase assays were done as previously described (15) with the addition of 5 mM EDTA added to the reaction mixtures. Briefly, 100 nM of *P. acnes* HemY alone or with either 100 nM HemH/Q or 100 nM HemH/Q preloaded with heme was assayed and activity was monitored at 37 °C using A Synergy HTI plate reader (Biotek, Winooski, VT) fitted with a 550 nm highpass filter and two 528 and 545 nm bandpass filters on the excitation light and a 635 bandpass filter on the emission light. Data (v vs [S]) were fitted to eq 1:

$$v = V_{app}[S]/(K_{app} + [S])$$

where V_{app} is the apparent maximum rate and K_{app} is the apparent Michaelis-Menten constant.

The coupled assays of protoporphyrinogen oxidase and ferrochelatase were done by measuring the amount of heme produced by HemQ, HemH, and HemY in the presence of protoporphyrin IX (p'gen IX) and iron. Reaction mixtures consisted of 1 μ M enzymes, 66 mM Tris HCl pH 8.0, 5 mM glutathione, 3.3% Tween 20, 10 mM beta-mercaptoethanol, 100 μ M ferrous ammonium sulfate, and ~25 μ M p'gen IX. In the case of one control, 25 μ M protoporphyrin IX was added instead. 1 mL reactions were conducted at 25 °C and terminated after 30 minutes with 500 μ L of 0.1M iodoacetamide and vortexed. Final heme concentrations were determined by the pyridine-hemochromogen assay (3). Here, 500 μ L of a 200 mM NaOH and 40% pyridine solution was added to the terminated reactions, and the resulting 2 mL was split into two cuvettes. Final samples were analyzed using a Cary 1G UV spectrophotometer (Varian, Palo Alto, CA) and quantified using the extinction coefficient hemochrome (20.7 mM^{-1}).

Analysis of Complex Formation.

To determine if HemY, HemH, and HemQ formed a multiprotein complex, fast protein liquid chromatography (FPLC) was used. FPLC was carried out using an Aktaprime machine equipped with a Hi-prep Sephacryl S-300 column (GE healthcare, Piscataway, NJ, USA) under low ionic strength buffers of 10 mM NaMOPS pH 7.2. Combinations of HemQ, HemH, and HemY were run together

and separately using purified recombinant protein from *B. subtilis*. Prior to loading, proteins were incubated at 37 °C for 30 min. All peaks obtained were collected, concentrated, and examined by SDS-PAGE analysis.

Results

HemQ requirement for protoporphyrinogen oxidase activity.

As stated above, the presence of *hemY*, *hemH*, and *hemQ* from organisms containing *hemQ* is required for functional complementation of an *E. coli* strain previously characterized as lacking PPO activity. HemY alone does not restore growth. This is surprising given that a homologous gene, human PPO, will function in these cells. To determine how these results applied *in vitro*, PPO assays were performed on purified protein. *P. acnes* HemY displayed a V_{\max} of 3.0 min^{-1} and a K_m for protoporphyrinogen IX of $19 \mu\text{M}$. *B. subtilis* HemY is reported to display a low V_{\max} as well (0.05 min^{-1}), but these values are much less than those of other reported PPO enzymes. Addition of the heme-loaded *P. acnes* HemH/Q fusion protein resulted in an increased V_{\max} of 19.8 min^{-1} and lowered the K_m to $10.1 \mu\text{M}$ (Figure 3.1). These changes in the catalytic values clearly indicate that the presence of HemH/Q influence the activity levels of HemY *in vitro*.

Heme synthesis assays.

We next sought to analyze the ability of *M. tuberculosis* HemY and HemH to form heme in the presence of protoporphyrinogen IX and ferrous iron. No detectable activity was observed. Significant activity could only be seen when

heme-loaded HemQ was added to reaction mixtures, once again confirming its essentiality (figure 3.2). If the HemQ added was not pre-loaded with heme before purification, overall activity was lessened. Correct quantification of these levels cannot be determined however, since initial heme formed in the reaction would immediately be bound by apo-HemQ present. The three component HemY/H/Q system from *B. subtilis* and *P. acnes* enzymes functioned similarly. Interestingly, the type of HemQ present in reactions did not matter despite overall differences in protein homology, as *M. tuberculosis* HemQ allowed *B. subtilis* HemY and HemH to function properly.

Analysis of proposed multi-protein complex.

Fast protein liquid chromatography (FPLC) was used to investigate the possibility that HemY, HemH, and HemQ exist as a soluble complex *in vivo*. Proteins were run under low ionic conditions to promote complex formation. FPLC of HemQ alone yields peaks at 35 and 48 ml (figure 3.3A). Presence of HemQ within these fractions was confirmed by SDS-PAGE. The peak at 48 mL corresponds to roughly 200,000 kDa, in-line with HemQ existing as a pentamer or hexamer such as crystal structures of HemQ homologues adopt. HemY, HemH, and HemQ were incubated together and then subjected to FPLC. The peaks obtained from these can be seen in figure 3.3B. SDS-PAGE analysis revealed the presence of HemQ in the same fractions as before, 35 and 48 mL. HemH was seen to elute at 55 mL, separate from HemQ. The presence of HemY was not detected in any of the major fractions.

Discussion

Each individual step in heme biosynthesis, from the initial synthesis of d-aminolevulinic acid (ALA) to the insertion of iron into protoporphyrin, have been extensively studied and characterized. Structures and established mechanisms exist for the vast majority of the enzymes involved. As a whole, the pathway is well-conserved among all life, understandable since the toxic intermediates produced must be handled with care. Still, variation among enzymes occurs at some steps, as outlined earlier. Given the grand diversity that exists among prokaryotes, this is quite understandable. What was not expected, however, was the discovery of an enzyme outside of the catalytic cycle, yet still essential for heme biosynthesis. As evident by our findings, HemQ affords soluble HemY and HemH their function in the Firmicutes and Actinobacteria. The conversion of protoporphyrinogen to heme cannot occur unless HemQ is available. While the exact mechanism of this remains unknown, it is clear that a catalytic role in the heme biosynthetic pathway is not apparent. The remaining conclusion then is the protein must stabilize the other two enzymes in some way. Our understanding of the heme biosynthesis pathway must be reevaluated in light of the discovery of HemQ. Other enzymes may yet exist that, like HemQ, there was no previous reason to believe unaccounted.

In eukaryotes, PPOX and FECH are homologous to HemY and HemH respectively. These enzymes are localized to the inner mitochondrial membrane on opposite sides, with PPOX on the cytosolic side and FECH on the matrix side (4, 8). For HemY and HemH of the Gram positives, however, these enzymes are

not membrane associated. If these enzymes must form a similar complex, one possibility is that the ring like structure of HemQ may act as a bridge to stabilize this interaction. With the active fusion of HemH and HemQ in *P. acnes*, it is likely that each HemQ protein interacts with one monomer of HemH inside this complex. Our findings that heme can only be produced when protoporphyrinogen IX, ferrous iron, and all three proteins are present further strengthens the multi-protein complex idea, as protoporphyrin, the substrate of ferrochelatase cannot be used despite its functionality. Unfortunately, our FPLC data currently does not support the existence of the multiprotein complex. This may be due to a variety of factors, including improper stoichiometry and unfavorable *in vitro* conditions (such as pH or salt concentrations). Another problematic factor is the expression of HemY from these organisms. Through current purification methods, clean, high yields of this protein are not readily obtainable. The interactions between these three enzymes may also be transient in nature, as has been suggested previously for HemY and HemH of other organisms, complicating matters to observe this formation *in vitro*. Further investigations will be required for accurate determination.

From a practical standpoint, one of the more intriguing features of the newly-discovered HemQ is its potential as a drug target. HemQ is present in a number of human pathogens, most notably *Mycobacterium tuberculosis* and *Staphylococcus aureus*. Both of these organisms have a significant health and economic impact globally, with rates continuing to climb. It is estimated that 30% of the world's population is infected with *M. tuberculosis*, with 1.7 million deaths

occurring in 2009 (14) In the case of *S. aureus*, the methicillin resistant strain (MRSA) has become the most frequent cause of skin and soft tissue infections in the United States (9) and is associated with higher mortality rates and increased hospitals cost levels (2). The presence of HemQ in these organisms represents a unique opportunity for antibiotic development. The potent herbicide acifluorfen inhibits protoporphyrinogen oxidase causing cell death through oxidative stress and is widely used in soybean, peanut, pea, and rice production (10). Likewise, drugs inhibiting the function of HemQ would effectively block the same step in the pathway. These then would fulfill many of the ideal characteristics of an effective antibiotic agent: Narrow spectrum, bacteriocidal, and limited effect to the host (heme synthesis within humans and most normal flora would be unaffected).

References

1. **Boynton TO, Daugherty LE, Dailey TA, & Dailey HA.** 2009. Identification of *Escherichia coli* HemG as a novel, menadione-dependent flavodoxin with protoporphyrinogen oxidase activity. *Biochemistry*. 48:6705-6711.
2. **Cosgrove SE, et al.** 2005. The impact of methicillin resistance in *Staphylococcus aureus* bacteremia on patient outcomes: mortality, length of stay, and hospital charges. *Infect Control Hosp Epidemiol*. 26:166-174.
3. **Dailey HA & Fleming JE.** 1983. Bovine ferrochelatase. Kinetic analysis of inhibition by N-methylprotoporphyrin, manganese, and heme. *J Biol Chem*. 258:11453-11459.
4. **Ferreira GC, Andrew TL, Karr SW, & Dailey HA.** 1988. Organization of the terminal two enzymes of the heme biosynthetic pathway. Orientation of protoporphyrinogen oxidase and evidence for a membrane complex. *J Biol Chem*. 263:3835-3839.
5. **Furuyama K, Kaneko K, & Vargas PD.** 2007. Heme as a magnificent molecule with multiple missions: heme determines its own fate and governs cellular homeostasis. *Tohoku J Exp Med*. 213:1-16.
6. **Gans J, Wolinsky M, & Dunbar J.** 2005. Computational improvements reveal great bacterial diversity and high metal toxicity in soil. *Science*. 309:1387-1390.
7. **Ghosh A, Wondimagegn T, & Ryeng H.** 2001. Deconstructing F(430): quantum chemical perspectives of biological methanogenesis. *Curr Opin Chem Biol*. 5:744-750.

8. **Harbin BM & Dailey HA.** 1985. Orientation of ferrochelatase in bovine liver mitochondria. *Biochemistry*. 24:366-370.
9. **Klevens RM, et al.** 2007. Invasive methicillin-resistant *Staphylococcus aureus* infections in the United States. *JAMA*. 298:1763-1771.
10. **Matringe M, Camadro JM, Labbe P, & Scalla R.** 1989. Protoporphyrinogen oxidase as a molecular target for diphenyl ether herbicides. *Biochem J*. 260:231-235.
11. **Mochizuki N, et al.** 2010. The cell biology of tetrapyrroles: a life and death struggle. *Trends Plant Sci*. 15:488-498.
12. **O'Brian MR & Thony-Meyer L.** 2002. Biochemistry, regulation and genomics of haem biosynthesis in prokaryotes. *Adv Microb Physiol*. 46:257-318.
13. **Raux E, Schubert HL, & Warren MJ.** 2000. Biosynthesis of cobalamin (vitamin B12): a bacterial conundrum. *Cell Mol Life Sci*. 57:1880-1893.
14. **Shenoi S, Heysell S, Moll A, & Friedland G.** 2009. Multidrug-resistant and extensively drug-resistant tuberculosis: consequences for the global HIV community. *Curr Opin Infect Dis*. 22:11-17.
15. **Shepherd M & Dailey HA.** 2005. A continuous fluorimetric assay for protoporphyrinogen oxidase by monitoring porphyrin accumulation. *Anal Biochem*. 344:115-121.
16. **Stackebrandt E.** 2004. Will we ever understand? The undescrivable diversity of the prokaryotes. *Acta Microbiol Immunol Hung*. 51:449-462.

17. **Tripathy BC, Sherameti I, & Oelmuller R.** 2010. Siroheme: an essential component for life on earth. *Plant Signal Behav.* 5:14-20.

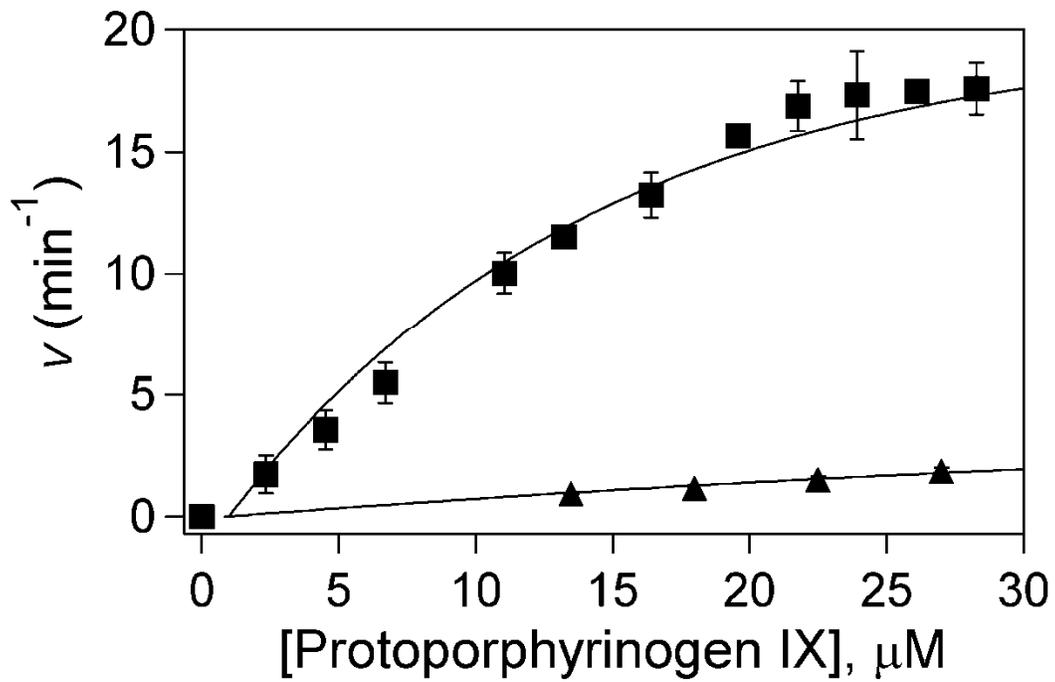


Figure 3.1. Stimulation of HemY activity by HemQ. HemY from *P. acnes* was assayed as described in the text with heme-loaded HemQ (upper curve). The activity exhibited by this system is greatly increased compared to that of HemY alone (lower curve).

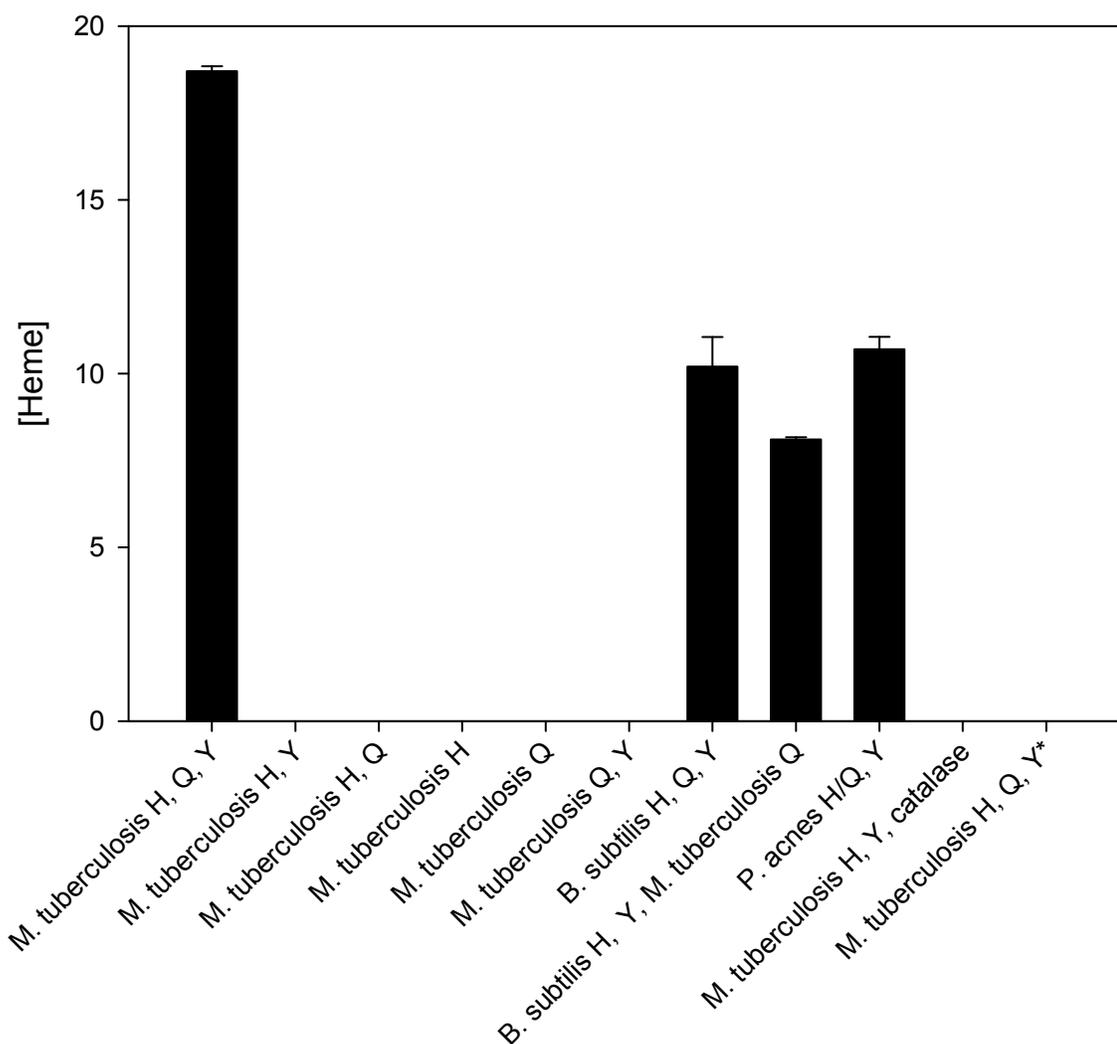


Fig. 3.2. Heme synthesis from protoporphyrinogen IX and ferrous iron. The bar graph shows heme formation in vitro by 1 nmol HemY and HemH. When heme-loaded HemQ was employed in the assay the amount of heme (present in holo-HemQ) added was quantitated and accounted for less than 5% of total heme in assays with heme-forming activity. In one experiment (far right) HemQ was replaced with catalase and in one experiment (denoted by asterisk) protoporphyrin, rather than the porphyrinogen, was supplied. In neither of these instances was heme formation observed.

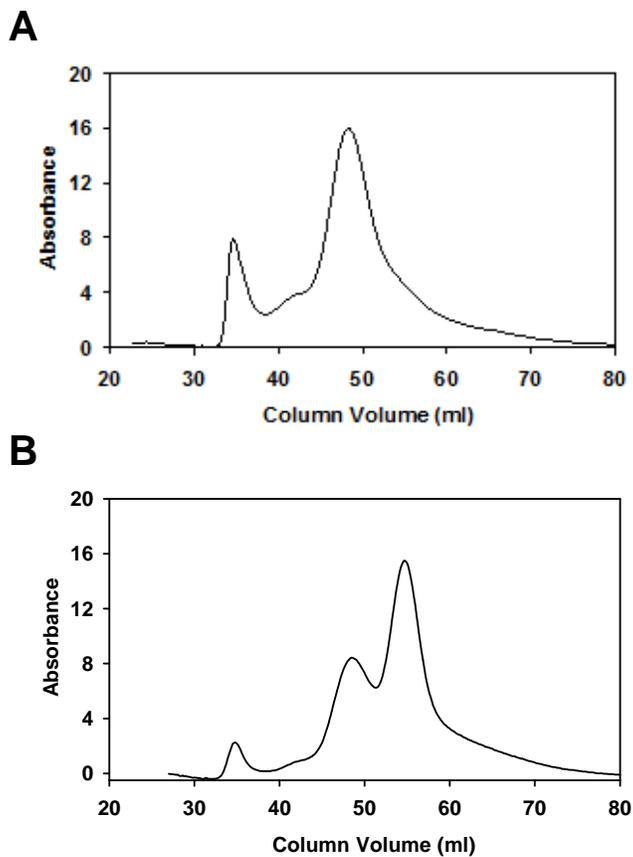


Figure 3.3. Sephacryl S-300 size exclusion chromatography of HemQ. **A.** FPLC of HemQ alone gives a major peak corresponding to 200,000 daltons. The peak at roughly 35 mL represents the void volume of the column. **B.** FPLC of HemQ in the presence of HemY and HemH yields an additional, separate eluted peak containing HemH.

Table 3.1. Complementation of *E. coli* Δ *hemG* with various expression plasmids. Plasmids for recombinant protein expression in *E. coli* were transformed into a *hemG*-deficient strain. Positive complementation resulted in the growth of wild-type sized colonies after overnight incubation at 37 °C. (From Dailey et al. 2010. *J. Biol. Chem.* 285:25978-25986.)

Protein expressed from plasmid	Complementation
pTrcHisA (empty vector)	-
Human protoporphyrinogen oxidase	+
<i>S. coelicolor</i> HemY	-
<i>S. coelicolor</i> HemY + HemH	-
<i>S. coelicolor</i> HemY + HemQ	-
<i>S. coelicolor</i> HemY + HemH + HemQ	+
<i>M. tuberculosis</i> HemY	-
<i>M. tuberculosis</i> HemY + HemH	-
<i>M. tuberculosis</i> HemY + HemQ	-
<i>M. tuberculosis</i> HemY + HemH + HemQ	+
<i>P. acnes</i> HemY	-
<i>P. acnes</i> HemY + HemH/Q fusion	+
<i>B. subtilis</i> HemY	-
<i>B. subtilis</i> HemY + HemH	-
<i>B. subtilis</i> HemY + HemQ	-
<i>B. subtilis</i> HemY + HemH + HemQ	+
<i>P. acnes</i> HemY + <i>M. tuberculosis</i> HemH	-
<i>P. acnes</i> HemY + <i>M. tuberculosis</i> HemQ	-
<i>P. acnes</i> HemY + <i>M. tuberculosis</i> HemH + HemQ	+
<i>P. acnes</i> HemH/Q + <i>M. tuberculosis</i> HemY	+
<i>M. tuberculosis</i> HemY + HemH + <i>B. subtilis</i> HemQ	+
<i>M. tuberculosis</i> HemH + HemQ + <i>B. subtilis</i> HemY	+
<i>M. tuberculosis</i> HemY + HemQ + <i>B. subtilis</i> HemH	+
<i>B. subtilis</i> HemY + HemH + <i>M. tuberculosis</i> HemQ	+

CHAPTER 4

DISCOVERY OF A GENE INVOLVED IN A THIRD BACTERIAL PROTOPORPHYRINOGEN OXIDASE ACTIVITY THROUGH COMPARATIVE GENOMIC ANALYSIS AND FUNCTIONAL COMPLEMENTATION ¹

¹ Boynton TO, Gerdes S, Craven SH, Neidle EL, Phillips JD, and HA Dailey
- Submitted to: *Applied and Environmental Microbiology*. 03/01/2011

Abstract

Tetrapyrroles are ubiquitous molecules in nearly all living organisms. Heme, an iron containing tetrapyrrole, is widely distributed in nature including most characterized aerobic and facultative bacteria. A large majority of bacteria that contain heme possess the ability to synthesize it. Despite this capability, and the fact that the biosynthetic pathway has been well studied, enzymes catalyzing at least three steps have remained “missing” in many bacteria. In the current work we have employed comparative genomics via the SEED genomic platform, coupled with experimental verification utilizing *Acinetobacter baylyi* ADP1, to identify one of the missing enzymes, a new protoporphyrinogen oxidase, the penultimate enzyme in heme biosynthesis. COG1981 was identified by genomic analysis as a candidate protein family for the missing enzyme in bacteria that lacked HemG or HemY, two known protoporphyrinogen oxidases. The predicted amino acid sequence of COG1981 is unlike the known enzymes HemG and HemY, but in some genomes the gene encoding it is found neighboring other heme biosynthetic genes. When the COG1981 gene was deleted from the genome of *A. baylyi*, a bacterium that lacks both *hemG* and *hemY*, the organism became auxotrophic for heme. Cultures accumulated porphyrin intermediates, and crude cell extracts lacked protoporphyrinogen oxidase activity. The heme auxotrophy was rescued by the presence of a plasmid-borne protoporphyrinogen oxidase gene from a number of different organisms such as *hemG* from *Escherichia coli*, *hemY* from *Myxococcus xanthus*, or human protoporphyrinogen oxidase.

Introduction

With the availability of an increasing number of completed microbial genome sequences, metabolic pathways that were long considered to have been conserved across species have now been found to have missing pieces (enzymes, transporters, etc.) in some organisms (16, 36). Such is the case with microbial tetrapyrrole biosynthesis. By the start of 2011 over 500 sequenced microbial genomes were complete for organisms believed to possess functional heme biosynthetic pathways. Yet most bacteria lacked one or more identifiable genes for enzymes that are necessary to catalyze heme biosynthetic steps (40). The situation for the *Archaea* is even less settled although a recent bioinformatics-based approach has added significant clarity to the issue (47). This is important given the fact that hemes are essential biological molecules in virtually all living organisms.

Heme, the iron-containing tetrapyrrole protoporphyrin IX, functions as an enzymatic cofactor with a diverse background including electron transfer in cytochromes (4), catalysis in catalases and peroxidases (51), diatomic gas transport in hemoglobin (41), activation of oxygen in P450 enzymes (42), and signal molecule recognition (15). In prokaryotes, heme can also serve as an iron source (49), a regulator of pathogenicity and dormancy (18, 48), and a transcriptional regulator for heme-responsive global iron networks (46). The biosynthesis of protoheme from the first committed precursor, δ -aminolevulinic acid (ALA) is relatively well conserved. However, two distinct mechanisms exist for ALA synthesis, the so-called 5-C route from glutamyl tRNA that is present in

plants, most bacteria and the Archaea, and the C-4 route from glycine and succinyl-CoA found in the metazoan and a few bacteria (21, 22). However, the remaining seven enzymatic steps are highly conserved in eukaryotes and the entire system is tightly regulated. The pathway is only absent in a few organisms such as helminthes and trypanosomes, where it would appear heme biosynthesis was lost in favor of heme uptake (25, 43). In some instances eukaryotes lacking the ability to synthesize heme maintain a bacterial endosymbiont to synthesize and provide the necessary heme (12). In prokaryotes this essential biological pathway is more diverse and less well characterized.

Examination of available bacterial genomes for heme biosynthetic enzymes reveals several noteworthy gaps where recognizable forms of characterized enzymes cannot be located (40). Interestingly, this phenomenon is not limited to one specific taxon, but occurs across a diverse set of microorganisms. The first gap occurs during the cyclization of the linear tetrapyrrole, where uroporphyrinogen synthase cannot be identified in some α -proteobacteria and chlamydiaceae. The next missing step occurs at the antepenultimate step, where two distinct enzymes are known to catalyze this reaction (7): an oxygen-dependent coproporphyrinogen oxidase named HemF (3) and an oxygen-independent, radical SAM-utilizing coproporphyrinogen dehydrogenase named HemN (29). Some heme-synthesizing bacteria lack both of these proteins, so it is clear that there is an as yet undiscovered form of coproporphyrinogen oxidase.

Until recently there was a third example, where an identified enzyme for the penultimate step, the six electron oxidation of protoporphyrinogen IX to protoporphyrin IX, was lacking. Some bacteria such as *Myxococcus* and *Aquifex* possess an oxygen-dependent, membrane associated enzyme HemY which is homologous to eukaryotic protoporphyrinogen oxidase (PPOX) (6, 7, 50). Firmicutes and Actinobacteria contain a soluble version of HemY that is different in that it lacks the membrane binding domain (8) and the enteric γ -proteobacteria employ the oxygen-independent enzyme HemG (20, 44) that is a quinone-dependent flavodoxin with protoporphyrinogen dehydrogenase (PpdH) activity that can be coupled to aerobic or anaerobic respiration (2, 34). The δ -proteobacteria, such as *Desulfovibrio*, lack typical PPO activity since they utilize a different pathway for the terminal steps (19, 31). Interestingly, a homolog of either HemG or HemY is absent in many cyanobacteria and non-enteric Gram-negative bacteria, including many pathogenic microbial genera: *Bordetella*, *Brucella*, *Campylobacter*, *Fusobacterium*, *Helicobacter*, *Pseudomonas*, and *Rickettsia* for example.

The purpose of the present study was to identify and characterize the protein(s) responsible for PPO activity in this broad class of prokaryotes using a marriage of bioinformatics and experimental techniques. By employing the SEED platform (38), we identified a protein family, uncharacterized at the time, that we predicted may fulfill the “missing” function of protoporphyrinogen oxidase in these taxa. Utilizing the naturally transformable bacterium *Acinetobacter*

baylyi ADP1 (23), which lacks *hemG* and *hemY*, we determined that the protein family COG1981 possesses PPO activity *in vivo* and *in vitro*.

Materials and Methods

Identification of genes for putative bacterial PPO

The Signature Tool allows homology-blind identification of candidate protein families based solely on their phyletic occurrence profile (as described in www.nmpdr.org/FIG/wiki/view.cgi/FIG/SigGenes). It requires as input an inclusion and an exclusion set of genomes. The latter is formed of diverse genomes, all of which harbor clear homolog(s) of known PPO gene(s), and hence are not expected to contain a novel non-homologous form of PPO. The inclusion set consists of genomes that encode the majority of heme biosynthetic genes, yet lack any known form of PPO, and hence are expected to encode a yet undiscovered non-homologous form of the enzyme (Supplementary Table 4.1). The Signature Tool identifies and returns protein families represented in the genomes of the inclusion set, but absent from genomes of the exclusion set. In order to shorten and optimize the output list of candidate protein families, only the smallest genomes were chosen for the inclusion set, while larger genomes were selected for the exclusion set. After several iterations an output file was obtained containing only 31 candidate protein families.

Cloning

The sequences of PCR primers used in this study can be found in Supplementary Table 4.2. *Acinetobacter baylyi* ADP1 *hemJ* was obtained by PCR of genomic DNA from this organism. The DNA for the open reading frame (ORF) was ligated into the NheI and HindIII site of pTrcHisA (Invitrogen). *E. coli hemG*, *M. xanthus hemY*, and human PPOX were cloned as previously described (2, 7, 9). For expression in *A. baylyi* ADP1, these genes were inserted into the NheI and HindIII site of pBBR1sp. pBBR1sp was made by inserting the following short sequence into the KpnI and XhoI site of the broad-host-range cloning vector pBBR1MCS (26):

5'-CA**AGGAGCAAGTAATGGCTAGCC**-3'

3'-CATGGT**TCCTCGTTCATTACCGATCGGAGCT**-5'

This sequence contains an *Acinetobacter* ribosomal binding site (AGGA), an ATG start codon, and a unique NheI site (shown in bold). The unique ApaI site of pBBR1MCS between KpnI and XhoI was abolished in the process and was used for verification of the insertion. The plasmid was confirmed by sequencing by the Georgia Genomics Facility at the University of Georgia.

Construction of HemJ Knockout

The *A. baylyi hemJ* knockout strain, termed ACN1054, was constructed by a three-step gene targeting method. First, a DNA sequence containing flanking regions ~400 bp upstream and ~800 bp downstream of *hemJ* was amplified by

PCR and cloned into the EcoRI and HindIII sites of pUC19. This plasmid was named pBAC883. Next, a kanamycin resistance cassette was excised from pUI1637 with XbaI and inserted into the unique SpeI site within the *hemJ* sequence of pBAC883 and the resulting plasmid was named pBAC884. Finally, pBAC884 was introduced into the bacterial chromosome by allelic exchange as previously described (5). The wildtype strain was transformed with linear pBAC884 and recombination events were selected on LB agar containing 10 mg/ml hemin and 30 mg/ml kanamycin. Positive recombination was further screened by plating on media without hemin. A lack of growth was used to indicate a phenotype auxotrophic for heme. ACN1054 was verified by sequencing of genomic DNA.

Protoporphyrinogen Oxidase Assays

Enzymatic assays were carried out using a continuous fluorimetric assay as previously described (45). Briefly, crude extracts were obtained by sonicating 15g of cells three times for 30 seconds on ice in 60 ml solubilization buffer containing 100 mM Tris-MOPS pH 8.1, 100 mM KCl, and 0.1% Na cholate. The resulting extracts were diluted 1:100 in the same buffer and assayed for PPO activity in reaction mixtures consisting of 10 μ L extract, 50 mM NaH_2PO_4 pH 8.0, 0.2% (w/v) Tween 20, 2.5 mM glutathione, and 25 μ M protoporphyrinogen IX. Extract protein concentrations were in the order of 0.1 mg/mL. For soluble and membrane fractions, crude extracts were centrifuged at 100,000 x g and the soluble supernatant fraction was collected. The pelleted membrane fraction was

then resuspended in an equivalent volume of solubilization buffer containing 1% Triton X-100 and centrifuged again at 100,000 x g. The resulting supernatant was then collected as the membrane fraction. These extracts were diluted 1:10 before use in reaction mixtures. Protoporphyrinogen IX was generated by reducing protoporphyrin IX (Frontier Scientific, Logan, UT) with an amalgam of 3.5g sodium and 75g mercury and then adjusted to neutral pH with 1.0 M Na-MOPS. Product formation was monitored by accumulation of fluorescence at 545 nm using a Synergy HTI plate reader (BioTek, Winooski, VT).

Porphyrin Accumulation

To induce porphyrin accumulation *in vivo*, ACN1054 cells were grown on LB agar containing 25 µg/ml δ-aminolevulinic acid and 5 µg/ml hemin. After 72 hours at 30° in the dark, plates were examined under UV light and a visible red fluorescence could be seen in cells positive for accumulation. For identification and quantification, 1.0 ml of a 5.0 ml 48h culture of ACN1054 was grown under the same conditions as above and centrifuged at 10,000 rpm for 3 minutes. The resulting pellet was then suspended in a solution of 4:1 (v/v) ethyl acetate/acetic acid solution for extraction of porphyrins. This suspension was centrifuged at 10,000 × g for 3 min and the supernatant was collected. Porphyrins were then separated and quantified using HPLC as previously described (13). The lowest level of detection for any porphyrin in this system is 1 pmol.

Complementation and Growth Analysis

The plasmid pBBR1sp containing one of the listed PPOs was transformed into ACN1054 cells from a 5 ml overnight culture in LB broth to which 2.0 mM succinate had been added thirty minutes prior to the transformation procedure. For transformation, 2.0 μ l of each respective vector and 2.0 μ l of culture were mixed and spotted onto an LB plate. After 24 hours, the resulting growth was transferred to a fresh LB plate containing 10 μ g/ml hemin, 20 μ g/ml chloramphenicol, and 30 μ g/ml kanamycin. Select colonies arising from this procedure were transferred to LB with both antibiotics but no hemin to verify complementation. Transformation of the isolated cultures was verified by extraction of original plasmid and identification by restriction digest.

For growth analysis, 1.0 ml of a 1.0 OD₆₀₀ culture was used to inoculate 100 mL LB medium containing appropriate antibiotics. Flasks were shaken at 37°C and the culture density measured hourly for ten hours using a Klett-Summerson colorimeter bearing a number 66 filter.

Results

Identification of HemJ

Since protein-based searches identified no homologs of HemG or HemY in many genomes of heme-synthesizing prokaryotes, a non-homology based technique was employed using the SEED genomic platform (www.theseed.org (38)). A subsystem encoding the heme and siroheme biosynthetic pathways was constructed as described by Osterman and Begley (37) for the nearly 690

eubacteria with complete and nearly complete genomes available in the SEED database at the time. An updated analysis can be viewed at: http://theseed.uchicago.edu/FIG/seedviewer.cgi?page=Subsystems&subsystem=Heme_and_Siroheme_Biosynthesis. From this evaluation, 167 organisms were inferred to be incapable of heme biosynthesis based on their genomic data (with 123 of these incapable of any tetrapyrrole biosynthesis). The heme biosynthetic pathway could be asserted in the remaining 521 eubacteria with sequenced genomes. Four representative species are shown in Supplementary Figure 4.1, illustrating several common operational variants of this subsystem. Of the 521 potentially heme synthesizing bacteria, approximately half (254) lacked an identifiable homolog of *hemG* or *hemY*. The NMPDR Signature Tool (www.nmpdr.org/FIG/wiki/view.cgi/FIG/SigGenes (33)) with the similarity score cutoff of $1e^{-10}$ and default parameters was then utilized to identify hypothetical protein families with the following desired occurrence profile:

- (1) Present only in genomes encoding other enzymes of heme biosynthesis, specifically the terminal enzymes surrounding PPO
- (2) Not found in genomes with known forms of PPO
- (3) Present in nearly all candidate genomes missing *hemG* and *hemY*

The inclusion and exclusion sets of genomes required as input for the Signature tool were composed as described in Materials and Methods and shown in Supplementary Table 4.1. Thirty-one hypothetical protein families with the desired occurrence profile were identified. The list of possible candidates was further shortened by examining genomic context, as genes associated with

the same metabolic pathway or functional complex tend to co-localize in prokaryotic genomes (14, 39). Using the Compare Regions tool of the SEED database, the presence of other known heme biosynthetic genes in the immediate vicinity of each of the 30 candidate gene families was assessed. Only one gene family, encoding COG1981, was found to cluster with either uroporphyrinogen decarboxylase and ferrochelatase genes or the ferrochelatase gene alone. This clustering was seen in 26% of the genomes examined.

The ORF for COG1981 occurs in nearly all eubacterial genomes where a known PPO is absent and is not present in organisms incapable of heme biosynthesis. The COG1981 homologs are found in cyanobacteria and all subdivisions of proteobacteria except for the δ -proteobacteria. As stated previously, *Myxococcus xanthus* and its relatives contain the eukaryotic-like HemY, and others such as *D. gigas* employ a different route to heme that is based upon the siroheme pathway (19, 31). The enterics also lack COG1981, containing HemG instead (2, 20, 30, 44). In addition, COG1981 is present in a few subclasses of bacteroides, though the majority of these contain HemY. These results led us to select COG1981 for characterization and the gene was named *hemJ* (Supplementary Figure 4.2). A previous publication had predicted that the COG1981 protein family was associated with heme synthesis, but no function was assigned by these researchers (17).

HemJ knockout in Acinetobacter baylyi ADP1

To elucidate the function of HemJ *in vivo* we chose to knock out the gene in an organism that contains only *hemJ* and not *hemG* or *hemY*. *Acinetobacter baylyi* ADP1 was selected since it possesses only *hemJ* (ACIAD0878), is naturally competent, and affords an ease of genetic manipulation via homology-driven recombination (23). The *hemJ* gene was disrupted by the introduction of a linear plasmid containing a gene conferring kanamycin resistance. The resulting strain, named ACN1054, was selected for its ability to grow in the presence of kanamycin and hemin. ACN1054 cells do not grow in the absence of hemin but wild-type growth characteristics are restored by hemin supplementation of the medium (Figure 4.1A). This is consistent with the designation of this gene as essential in the complete *A. baylyi* knockout collection generated by de Berardinis et al (11).

Complementation of ACN1054

Functional complementation of heme auxotrophs has proven a valuable tool in identifying heme biosynthetic proteins (9, 10, 30, 35). To determine the role of HemJ, we cloned genes encoding several validated PPO enzymes into a newly created *A. baylyi*-active expression vector. HemG from *E. coli*, HemY from *Myxococcus xanthus*, and human PPOX were selected as candidates for complementation of ACN1054 on the basis that each is a characterized, fully functioning protoporphyrinogen oxidase. In addition, HemJ from *A. baylyi* ADP1 was used as a control. Soluble HemY of the Firmicutes or Actinobacteria was

not chosen since it has been shown that it is not functional in *E. coli* unless supplied concomitantly with HemH and HemQ (8). Each individual PPO-encoding plasmid was transformed into ACN1054 and resulted in the rescue of heme auxotrophy. Growth analysis further revealed that each plasmid restored growth to wild type status (Figure 4.1B).

To further confirm if the gene under examination possessed PPO activity two separate approaches were used: examination of intracellular porphyrin accumulation that is anticipated when PPO activity is not present, and in vitro assays of crude cell extracts. In other systems, such as *E. coli*, lack of PPO results in an increase of the protoporphyrin precursors uro-, copro-, and protoporphyrinogen (44). Accumulation of these intermediates is enhanced by the addition of δ -aminolevulinic acid, the initial precursor compound in heme biosynthesis.

When supplemented with ALA, ACN1054 cells exhibited an orange-red fluorescence under ultraviolet illumination after 24 hours of growth. This fluorescence increased up to approximately 72 hours of incubation (Figure 4.2). HPLC was used to separate and quantify oxidized samples obtained from cellular acid extracts. Data presented in Table 4.1 demonstrates an accumulation of protoporphyrin, coproporphyrin, uroporphyrin, and a heptacarboxyl porphyrin. The latter is an intermediate formed during the enzymatic decarboxylation of uroporphyrin. In comparison, quantifiable amounts of porphyrins were not found in extracts from wild-type cells supplemented with ALA in the same fashion.

In vitro PPO activity was determined in assays performed with crude cell lysates of wild type and ACN1054 cells (Figure 4.3A). No PPO activity was present in the mutant. *In vitro* assays were also conducted using soluble and membrane fractions of wild type *A. baylyi* extracts. Analysis of the HemJ amino acid sequence by the program TMHMM (27) yields a firm prediction for a membrane bound protein. However, no measurable activity was found for isolated membrane or soluble fractions. Activity was seen only when the soluble and membrane fractions were present together (Figure 4.3B).

Discussion

Over the past decades researchers have employed heme synthesis mutants of *E. coli* to identify enzymes of heme biosynthesis via complementation. This approach, while straightforward and robust, can only work if the sought-after enzyme is functional in *E. coli*. Another approach is to rely upon specific biochemical characteristics of the sought after protein. Kato et al. recently identified HemJ in *Synechocystis* in this fashion by introducing plant-derived acifluorfen-sensitive PPO into cells and then screening mutants for acifluorfen sensitivity. (24). This approach is suited for identification of novel enzymes where a second, well-characterized enzyme exists that catalyzes the reaction of interest and for which a novel inhibitor also exists, but it does not easily lend itself to all enzyme systems. A third approach is to employ bioinformatics to initially identify genes of potential interest and then experimentally test these. This technique has a broad range of application in biological systems for which

significant genomic data exists. It requires little knowledge of actual function to identify novel components, relying instead on gene context and functional complementation. That was the approach we employed herein to identify *hemJ*.

We utilized the SEED (<http://www.theseed.org>) platform in which a characterized biological system from one species can be accurately projected to other species within the database based on sequence similarity, nonhomologous contextual clues, and experimental evidence of enzymatic function (38). From this, “missing” steps become readily apparent as “gaps” for a given genome within the displayed subsystem. Once gaps are identified, the tools for comparative genome analysis within the SEED facilitate identification of candidate gene families that can then be experimentally verified (see 11, 15, 44, 45). As detailed above, this led to our identification of COG1981 as a strong candidate for the missing PPO that was then experimentally verified as HemJ.

With HemG, HemY and HemJ now identified, nearly all eubacterial genomes in the SEED database inferred to encode functional heme biosynthetic pathways have identifiable enzymes to catalyze the conversion of protoporphyrinogen IX to protoporphyrin IX (663 out of 671). The majority (46%) encode HemJ, with HemY or HemG present in 32% and 23%, respectively. The only two clear exceptions appear to be *Tropheryma* and *Desulfovibrionaceae*. For the *Desulfovibrionaceae* the explanation is simple. An alternative, ancient pathway for the biosynthesis of heme has been shown operational in the genus *Desulfovibrio* (19). Lobo et al. (31) presented evidence that protoheme biosynthesis in *D. vulgaris* occurs by branching off the common tetrapyrrole route

not at the level of precorrin-2 as was hypothesized by Ishida et al. (19) but at the level of sirohydrochlorin. In this scenario the activities of classic CPO, PPO, and possibly ferrochelatase are not involved in the terminal steps of protoheme synthesis. Instead the *D. vulgaris* homologs of the NirJ and NirD enzymes participate in this novel transformation (31, 47).

Tropheryma whipplei, the causative agent of Whipple's disease, is an obligate human pathogen that has not been cultured in the absence of living eukaryotic cells (28). Compared to free-living species of the same order (Actinomycetales), this organism has undergone drastic genome reduction (to merely ~ 930 kbp), resulting in the lack of key biosynthetic pathways and a reduced capacity for energy metabolism (1). The absence of a gene encoding any of the 3 known forms of PPO from the genome of this highly host-adapted organism might indicate a recent onset of pathway decay and the reliance on the host as a source of heme.

Kato et al. (24) noted that they could not identify *hemJ*, *hemG* or *hemY* in the Archaea, Actinobacteria, Acidibacteria, Deinococcus, Fusobacteria, Spirochaetes, and Thermotogae, and posited that these organisms may contain other yet to be identified novel PPO enzymes. This is clearly true in the Archaea where, with limited exceptions, these organisms do not possess HemF/N, HemG/Y/J, or HemH as the terminal pathway enzymes essential for protoheme biosynthesis, but instead utilize a different pathway from uroporphyrinogen to heme that involves *SUMT*, *PC2-DH* and the *ahb-nir* genes (47). While some Spirochaetes contain the siroheme biosynthetic pathway, but not the terminal

enzymes for protoheme synthesis, the remaining Spirochaetes and Thermotogae contain no tetrapyrrole biosynthetic enzymes. However, it should be noted that *hemY* is present in the genomes of Acidobacteria and Actinobacteria and has been expressed and characterized from the Actinobacteria (8). The examination of the remaining groups using the SEED Subsystem readily explains the absence of *hemG*, *hemY* or *hemJ* in these organisms.

The present work does not identify whether HemJ is a stand alone PPO enzyme or if it is a subunit of a larger complex. Although it is clear that HemJ is an essential protein for the cellular oxidation of protoporphyrinogen to protoporphyrin, our observation that both the membrane fraction containing HemJ and the soluble cellular fractions are required for measurable activity *in vitro* would argue for an additional component. Whether this additional component is an as yet unidentified protein that did not manifest itself in our genome screen or simply a redox cofactor is something that will require additional study. Interestingly, Kato et al. (24) reported that a partially purified HemJ-maltose binding protein fusion from *Rhodobacter sphaeroides* could catalyze oxidation of protoporphyrinogen to protoporphyrin without any added cofactors. This led them to speculate that HemJ could utilize molecular oxygen for the reaction. However, no description or evidence for the level of protein purity was presented and no ultraviolet/visible spectrum was presented that would have identified chromophoric cofactors such as FMN or FAD. During our studies we cloned and attempted to express and purify *hemJ* from *A. baylyi*, *Pseudomonas aeruginosa*, and *R. sphaeroides* using a variety of different

methods including maltose binding protein fusions. However, in all our efforts the membrane-associated enzyme expressed poorly and did not purify to a level that would allow confident kinetic studies. Since neither Kato et al. nor we managed to obtain pure enzyme a comparison of our results is not possible. However, the specific activity of our combined membrane and soluble fractions is greater than a factor of 10 higher than their published data for purified enzyme.

In addition to the use of SEED to identify target genes, the current work also has additional impact in that we have utilized *A. baylyi* ADP1 rather than *E. coli* in our gene target verification protocol. While our choice was in part driven by the fact that *A. baylyi* ADP1 lacks *hemG* and *hemY*, we were also attracted to the high natural transformability and ease with which one can produce gene mutations, deletions and homologous recombinations (23). Interestingly had we chosen *E. coli* Δ *hemG* for complementation with the sought after HemJ, we would have not been successful since we discovered that HemJ does not function in *E. coli* (data not shown).

HemJ represents a novel drug target for many human pathogens. Acifluoren, a potent herbicide, is already a widely-used plant PPO inhibitor (32) and it is imaginable that HemJ could present a high-specificity target in those organisms that possess it. Many clinically significant pathogens and opportunists that have HemJ are among those organisms that have evolved multi-antibiotic resistance and for these HemJ presents a new antimicrobial target. The prokaryotic heme biosynthetic pathway has previously been proposed as a drug

target for symbiont-containing parasites (12), and HemJ's absence in most human flora suggest it could be a prime candidate.

Acknowledgements: This work was supported by grant DK32303 (H.A.D.), NSF grant MCB-0920619 (E.L.N.), and NIAID Contract Number HHSN266200400042C (S.G.). We wish to acknowledge D. Vavilin for generous sharing of unpublished data related to HemJ of *Synechocystis* and TA Dailey for editing of the manuscript.

References

1. **Bentley SD, et al.** 2003. Sequencing and analysis of the genome of the Whipple's disease bacterium *Tropheryma whipplei*. *Lancet*. 361:637-644.
2. **Boynton TO, Daugherty LE, Dailey TA, & Dailey HA.** 2009. Identification of *Escherichia coli* HemG as a novel, menadione-dependent flavodoxin with protoporphyrinogen oxidase activity. *Biochemistry*. 48:6705-6711.
3. **Breckau D, Mahlitz E, Sauerwald A, Layer G, & Jahn D.** 2003. Oxygen-dependent coproporphyrinogen III oxidase (HemF) from *Escherichia coli* is stimulated by manganese. *J. Biol. Chem.* 278:46625-46631.
4. **Chance B.** 1972. The nature of electron transfer and energy coupling reactions. *FEBS Lett.* 23:3-20.
5. **Collier LS, Gaines GL, 3rd, & Neidle EL.** 1998. Regulation of benzoate degradation in *Acinetobacter* sp. strain ADP1 by BenM, a LysR-type transcriptional activator. *J. Bacteriol.* 180:2493-2501.
6. **Corradi HR, et al.** 2006. Crystal structure of protoporphyrinogen oxidase from *Myxococcus xanthus* and its complex with the inhibitor acifluorfen. *J. Biol. Chem.* 281:38625-38633.
7. **Dailey HA & Dailey TA.** 1996. Protoporphyrinogen oxidase of *Myxococcus xanthus*. Expression, purification, and characterization of the cloned enzyme. *J. Biol. Chem.* 271:8714-8718.
8. **Dailey TA, et al.** 2010. Discovery and characterization of HemQ: an essential heme biosynthetic pathway component. *J. Biol. Chem.* 285:25978-25986.

9. **Dailey TA & Dailey HA.** 1996. Human protoporphyrinogen oxidase: expression, purification, and characterization of the cloned enzyme. *Protein Sci.* 5:98-105.
10. **Dailey TA, Dailey HA, Meissner P, & Prasad AR.** 1995. Cloning, sequence, and expression of mouse protoporphyrinogen oxidase. *Arch. Biochem. Biophys.* 324:379-384.
11. **de Berardinis V, et al.** 2008. A complete collection of single-gene deletion mutants of *Acinetobacter baylyi* ADP1. *Mol. Syst. Biol.* 4:174.
12. **Foster J, et al.** 2005. The *Wolbachia* genome of *Brugia malayi*: endosymbiont evolution within a human pathogenic nematode. *PLoS Biol.* 3:e121.
13. **Franklin MR, Phillips JD, & Kushner JP.** 1997. Cytochrome P450 induction, uroporphyrinogen decarboxylase depression, porphyrin accumulation and excretion, and gender influence in a 3-week rat model of porphyria cutanea tarda. *Toxicol. Appl. Pharmacol.* 147:289-299.
14. **Galperin MY & Koonin EV.** 2000. Who's your neighbor? New computational approaches for functional genomics. *Nat. Biotechnol.* 18:609-613.
15. **Gilles-Gonzalez MA.** 2001. Oxygen signal transduction. *IUBMB Life.* 51:165-173.
16. **Hanson AD, Pribat A, Waller JC, & de Crecy-Lagard V.** 2009. 'Unknown' proteins and 'orphan' enzymes: the missing half of the engineering parts list--and how to find it. *Biochem. J.* 425:1-11.

17. **Harrington ED, et al.** 2007. Quantitative assessment of protein function prediction from metagenomics shotgun sequences. *Proc. Natl. Acad. Sci. USA.* 104:13913-13918.
18. **Ioanoviciu A, Meharena YT, Poulos TL, & Ortiz de Montellano PR.** 2009. DevS oxy complex stability identifies this heme protein as a gas sensor in *Mycobacterium tuberculosis* dormancy. *Biochemistry.* 48:5839-5848.
19. **Ishida T, et al.** 1998. A primitive pathway of porphyrin biosynthesis and enzymology in *Desulfovibrio vulgaris*. *Proc. Natl. Acad. Sci. USA.* 95:4853-4858.
20. **Jacobs NJ & Jacobs JM.** 1977. Evidence for involvement of the electron transport system at a late step of anaerobic microbial heme synthesis. *Biochim. Biophys. Acta.* 459:141-144.
21. **Jahn D & Heinz DW.** 2009. Biosynthesis of 5-aminolevulinic acid. p. 29-42. *In* M. J. Warren and A. G. Smith (ed.), *Tetrapyrroles: Birth, Life, and Death*, Landes Bioscience, Austin, Texas.
22. **Jordan PM.** 1990. The biosynthesis of 5-aminolevulinic acid and its transformation in coproporphyrinogen in animals and bacteria. p. 55-121. *In* H. A. Dailey (ed.), *Biosynthesis of Heme and Chlorophylls*, McGraw-Hill, New York, New York.
23. **Juni E & Janik A.** 1969. Transformation of *Acinetobacter calco-aceticus* (*Bacterium anitratum*). *J. Bacteriol.* 98:281-288.

24. **Kato K, Tanaka R, Sano S, Tanaka A, & Hosaka H.** 2010. Identification of a gene essential for protoporphyrinogen IX oxidase activity in the cyanobacterium *Synechocystis* sp. PCC6803. *Proc Natl. Acad. Sci. USA.* 107:16649-16654.
25. **Koreny L, Lukes J, & Obornik M.** 2010. Evolution of the haem synthetic pathway in kinetoplastid flagellates: an essential pathway that is not essential after all? *Int. J. Parasitol.* 40:149-156.
26. **Kovach ME, Phillips RW, Elzer PH, Roop RM, 2nd, & Peterson KM.** 1994. pBBR1MCS: a broad-host-range cloning vector. *Biotechniques.* 16:800-802.
27. **Krogh A, Larsson B, von Heijne G, & Sonnhammer EL.** 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* 305:567-580.
28. **La Scola B, et al.** 2001. Description of *Tropheryma whipplei* gen. nov., sp. nov., the Whipple's disease bacillus. *Int. J. Syst. Evol. Microbiol.* 51:1471-1479.
29. **Layer G, et al.** 2006. The substrate radical of *Escherichia coli* oxygen-independent coproporphyrinogen III oxidase HemN. *J. Biol Chem.* 281:15727-15734.
30. **Lermontova I, Kruse E, Mock HP, & Grimm B.** 1997. Cloning and characterization of a plastidal and a mitochondrial isoform of tobacco protoporphyrinogen IX oxidase. *Proc. Natl. Acad. Sci. USA.* 94:8895-8900.

31. **Lobo SA, Brindley A, Warren MJ, & Saraiva LM.** 2009. Functional characterization of the early steps of tetrapyrrole biosynthesis and modification in *Desulfovibrio vulgaris* Hildenborough. *Biochem. J.* 420:317-325.
32. **Matringe M, Camadro JM, Labbe P, & Scalla R.** 1989. Protoporphyrinogen oxidase as a molecular target for diphenyl ether herbicides. *Biochem. J.* 260:231-235.
33. **McNeil LK, et al.** 2007. The National Microbial Pathogen Database Resource (NMPDR): a genomics platform based on subsystem annotation. *Nucleic Acids Res.* 35:D347-353.
34. **Mobius K, et al.** 2010. Heme biosynthesis is coupled to electron transport chains for energy generation. *Proc. Natl. Acad. Sci. USA.* 107:10436-10441.
35. **Narita S, et al.** 1996. Molecular cloning and characterization of a cDNA that encodes protoporphyrinogen oxidase of *Arabidopsis thaliana*. *Gene.* 182:169-175.
36. **Osterman A & Overbeek R.** 2003. Missing genes in metabolic pathways: a comparative genomics approach. *Curr. Opin. Chem. Biol.* 7:238-251.
37. **Osterman AL & Begley TP.** 2007. A subsystems-based approach to the identification of drug targets in bacterial pathogens. *Prog. Drug. Res.* 64:131, 133-170.

38. **Overbeek R, et al.** 2005. The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. *Nucleic Acids Res.* 33:5691-5702.
39. **Overbeek R, Fonstein M, D'Souza M, Pusch GD, & Maltsev N.** 1999. The use of gene clusters to infer functional coupling. *Proc. Natl. Acad. Sci. USA.* 96:2896-2901.
40. **Panek H & O'Brian MR.** 2002. A whole genome view of prokaryotic haem biosynthesis. *Microbiology.* 148:2273-2282.
41. **Perutz M.** 1942. X-ray analysis of haemoglobin. *Nature.* 149:491-494.
42. **Raikhman LM, Annaev B, Mamedniazov ON, & Rozantsev EG.** 1973. [Molecular structure and mechanism of functioning of cytochrome P-450]. *Biofizika.* 18:228-234.
43. **Rao AU, Carta LK, Lesuisse E, & Hamza I.** 2005. Lack of heme synthesis in a free-living eukaryote. *Proc. Natl. Acad. Sci. USA.* 102:4270-4275.
44. **Sasarman A, et al.** 1979. Mapping of a new hem gene in *Escherichia coli* K12. *J. Gen. Microbiol.* 113:297-303.
45. **Shepherd M & Dailey HA.** 2005. A continuous fluorimetric assay for protoporphyrinogen oxidase by monitoring porphyrin accumulation. *Anal. Biochem.* 344:115-121.
46. **Singleton C, et al.** 2010. Heme-responsive DNA binding by the global iron regulator Irr from *Rhizobium leguminosarum*. *J. Biol. Chem.* 285:16023-16031.

47. **Storbeck S, et al.** 2010. A novel pathway for the biosynthesis of heme in Archaea: genome-based bioinformatic predictions and experimental evidence. *Archaea*. 2010:175050.
48. **Torres VJ, et al.** 2007. A *Staphylococcus aureus* regulatory system that responds to host heme and modulates virulence. *Cell Host Microbe*. 1:109-119.
49. **Wandersman C & Delepelaire P.** 2004. Bacterial iron sources: from siderophores to hemophores. *Annu. Rev. Microbiol.* 58:611-647.
50. **Wang KF, Dailey TA, & Dailey HA.** 2001. Expression and characterization of the terminal heme synthetic enzymes from the hyperthermophile *Aquifex aeolicus*. *FEMS Microbiol. Lett.* 202:115-119.
51. **Zamocky M, Furtmuller PG, & Obinger C.** 2010. Evolution of structure and function of Class I peroxidases. *Arch. Biochem. Biophys.* 500:45-57.

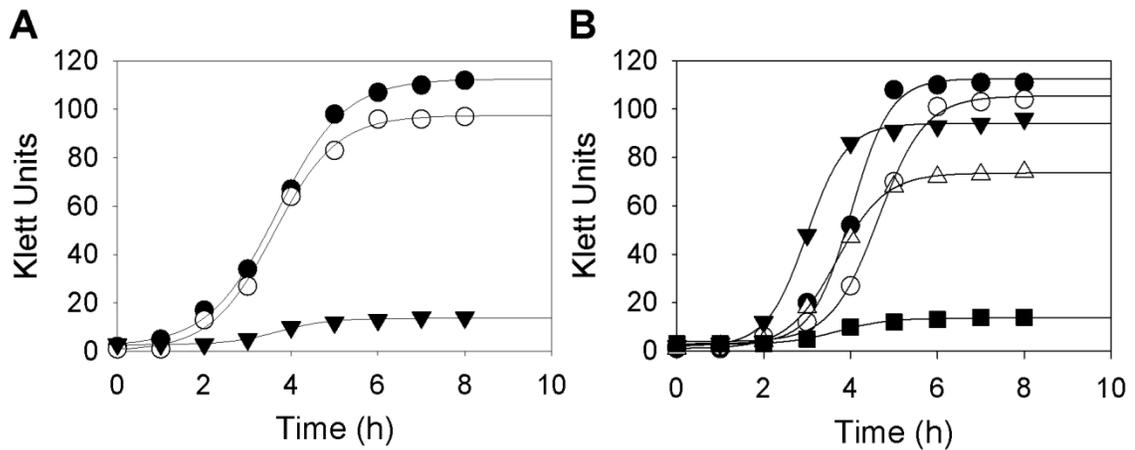


Figure 4.1. Comparison of *A. baylyi* ADP1 wild-type and ACN1054 cultures. **A.** Growth characteristics of wild type and ACN1054 mutant of *Acinetobacter baylyi* ADP1: ACN1054 cultures without added heme (▼) and supplemented with exogenous heme (○). Wild type cells without heme supplementation (●). **B.** Growth analysis of ACN1054 with expression of known protoporphyrinogen oxidase enzymes. Growth of ACN1054 cells alone (■) and with plasmid-encoded expression of human PPO (●), *M. xanthus* HemY (○), and *E. coli* HemG (△). Wildtype *A. baylyi* ADP1 is shown for comparison (▼).

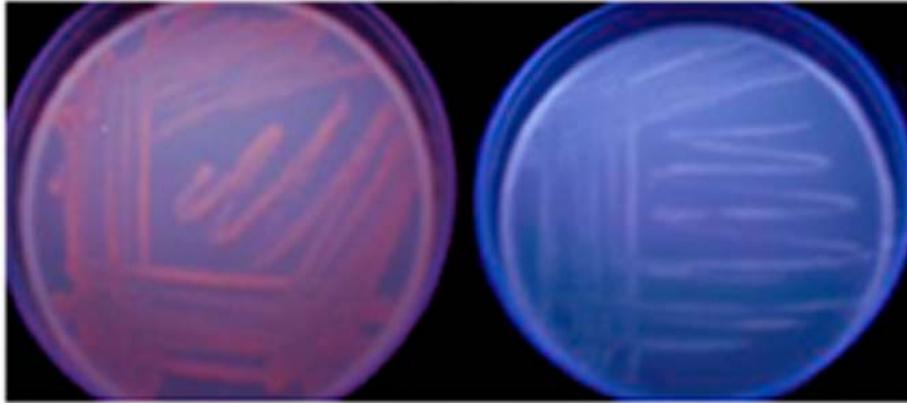


Figure 4.2. Porphyrin accumulation in mutant and wildtype *A. baylyi* cells. After stimulation of the heme biosynthetic pathway, porphyrin accumulation is clearly seen in ACN1054 cells (left) but not in wildtype cells (right). Fluorescence is visible with exposure to long wavelength UV light.

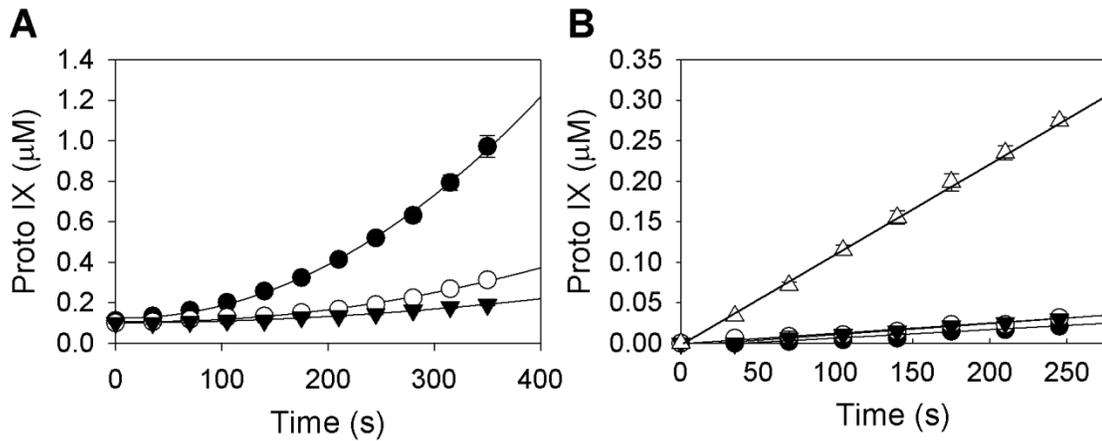
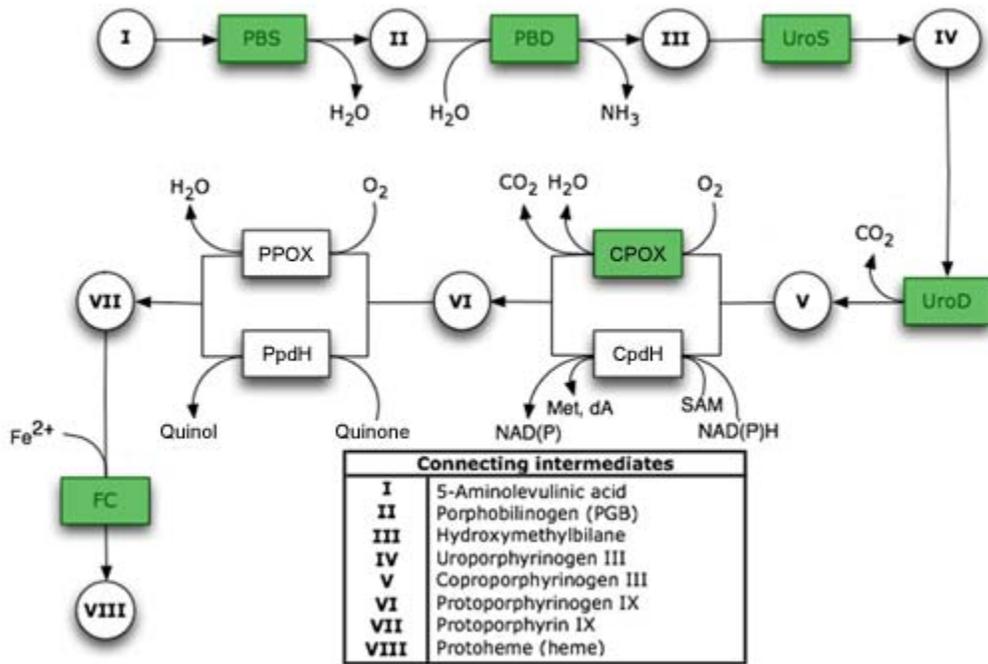


Figure 4.3. Protoporphyrinogen oxidase activity in *Acinetobacter* crude cell extracts as measured by continuous fluorescence spectroscopy. **A.** *In vitro* PPO assays are shown for wildtype (●) and ACN1054 (○) extracts, as well as an extract-free control (▼). **B.** Protoporphyrinogen oxidase activity in membrane and soluble fractions of *Acinetobacter* extracts. Activity levels for the soluble fraction (○), membrane fraction (▼), extract free control (●), and membrane plus soluble fractions together (△). Approximately 0.1 mg/ml protein was present in each extract.

Supplementary Figure 4.1. Abbreviated version of the populated subsystem encoding heme biosynthesis (complete version can be found on the SEED server). **A:** General pathway diagram is shaded to show specific implementation of the heme biosynthetic pathway in *A. baylyi* ADP1. **B:** Subsystem spreadsheet fragment: only 4 representative organisms are shown (out of ~900 available in SEED at the time of publication). Plus and minus symbols denote the presence or absence respectively of genes within each organism. *E. coli* contains all genes within the pathway. *C. botulinum* synthesizes siroheme but not heme, and thus lacks the terminal enzymes. *H. influenza* does not synthesize heme but contains the final enzyme ferrochelatase. *A. baylyi* ADP1 possesses all but one gene within the pathway, indicating a novel gene must be present. Abbreviations employed are: PBS=porphobilinogen synthase, PBD=porphobilinogen deaminase, UroS=uroporphyrinogen synthase, UroD=uroporphyrinogen decarboxylase, CPOX=coproporphyrinogen oxidase, CpdH = coproporphyrinogen dehydrogenase, PPO=protoporphyrinogen oxidase, PpdH = protoporphyrinogen dehydrogenase, FC=ferrochelatase.

A



B

Organism	PBS	PBD	UroS	UroD	CPO	PPO	FC
<i>Escherichia coli</i>	+	+	+	+	+	+	+
<i>Acinetobacter baylyi</i> ADP1	+	+	+	+	+	-	+
<i>Clostridium botulinum</i>	+	+	+	-	-	-	-
<i>Haemophilus influenzae</i>	-	-	-	-	-	-	+

<i>Synechocystis</i>	MPKREYFSLPCPLSTFTMAYYWFKAFHLLIGIVVWFAGLFYLVRLFVYHAE	50
<i>Helicobacter</i>	---MEFLSG-----YFLVWKAFHVIAVISWMAALFYLPRLFVYHAE	38
<i>Nitrococcus</i>	-----MFYLWIKAFHIIIFVITWFAGLFYLPRLFVYHAM	33
<i>Pseudomonas</i>	-----MYMWLKAFHIIAVVCWFAGLFYLPRLFVYHAM	32
<i>Acinetobacter</i>	-----MDAPSDAFLVWKALHIIAVVCWFAALFYLPRLFVYHAM	38
<i>Rhodobacter</i>	MRMADHFEEETTMGTFLADYYLWTKSLHVISVLAWMAGLFYLPRLFVYHAE	50
	* * * * * * * * * * * * * *	
<i>Synechocystis</i>	ADQEPEPAKTILKKQYELMEKRLYNIITTPGMVVTVAMAIGLIFTEPEI-	99
<i>Helicobacter</i>	NAHKKEFVG----VVQIQEKKLYSFIASPAMGFTLITGILMLLIAPEM-	82
<i>Nitrococcus</i>	ATEAAGRER-----FKVMERKLYRGIMTPSGTLTVLFGGWLlyLSPG--	75
<i>Pseudomonas</i>	SEDQTSRER-----FCVMERKLYRGIMMPSMLATLVLGLWMLYLTPG--	74
<i>Acinetobacter</i>	SEDDTSHRR-----FEVMERKLYRGIMWPSMIATLITAHFLVDWGDATR	82
<i>Rhodobacter</i>	VVKAGTETD----ALFQTMERRLLRAIMNPAMIATWIFGLLLVFTPGIVD	96
	* * * * *	
<i>Synechocystis</i>	-LKSG-WLHIKLTfVALLLLYHFYCGRVMKKLAQGESQWSGQQFRALNEA	147
<i>Helicobacter</i>	-FKSGGWLHAKLALVLLLIYHFYCKKCMRELEKDPTGKNARFYRVFNEI	131
<i>Nitrococcus</i>	-WLSedwLHLKLLLVLLIIFYHLYCGILLRRFRDDRNTSHVWYRWFNEL	124
<i>Pseudomonas</i>	-WLSQGWLHAKLTLVLLIGYHHACGAMLKRFARGEPRSHVFYRWFNEV	123
<i>Acinetobacter</i>	HYHEALWFYIKVGLVGLLVIYHFVCGYYRKKLIGNAHYKSHKFWRYFNEM	132
<i>Rhodobacter</i>	--WSMLWPWTKAACVLAmtGFHMWLAARRRDFAAGANRHKGRTYRMMNEL	144
	* * * * * * * *	
<i>Synechocystis</i>	PTILLVIVILLAVFKN-NLPLDATTWLIIVALVIAMAASIQLYAKRRRDQ	196
<i>Helicobacter</i>	PTILMILIVILVVVK----PF-----	148
<i>Nitrococcus</i>	PVMGLIGIVILVVVK----PL-----	141
<i>Pseudomonas</i>	PVLFLLLIVLLVVLK----PF-----	140
<i>Acinetobacter</i>	PTLILFAVVILVVVK----PQF-----	150
<i>Rhodobacter</i>	PTLLMLVIVFSAVAKWNYWGF-----	165
	* * * *	
<i>Synechocystis</i>	ALLTEQQKAASAQN	210
<i>Helicobacter</i>	-----	
<i>Nitrococcus</i>	-----	
<i>Pseudomonas</i>	-----	
<i>Acinetobacter</i>	-----	
<i>Rhodobacter</i>	-----	

Supplementary Figure 4.2. Amino acid sequence comparisons of various HemJ enzymes. Overall sequence identity among HemJs is low. A carboxyl-terminal extension is present only in cyanobacterial HemJ.

Table 4.1. Porphyrin analysis of cell extracts from *A. baylyi* ADP1 wild-type and ACN1054

	Porphyrins (nM)				
	Total	Uro-	7-COOH	Copro-	Proto-
ACN1054	4802 ± 318	559 ± 81	81 ± 43	2429 ± 149	1734 ± 135
Wildtype	b.d.	b.d.	b.d.	b.d.	b.d.

*Quantitation of porphyrins was done as described in Materials and Methods. Samples for analysis were obtained from 1 mL of 48 hour *A. baylyi* ADP1 cultures grown in the presence of δ -aminolevulinic acid. b.d., below detection level.

Supplementary Table 4.1. Identification of candidate protein families for the missing role of PPO via NMPDR Signature Tool

Inclusion genome set (a novel PPO is expected to be present)	Exclusion genome set (presence of a novel PPO is unlikely)
<p><i>Bordetella bronchiseptica</i> RB50 <i>Bordetella pertussis</i> Tohama I <i>Brucella melitensis</i> 16M <i>Brucella suis</i> 1330 <i>Helicobacter pylori</i> J99 <i>Rickettsia conorii</i> str. Malish 7 <i>Rickettsia prowazekii</i> str. Madrid E <i>Helicobacter pylori</i> 26695 <i>Tropheryma whipplei</i> str. Twist <i>Wolbachia pipientis</i> wMel <i>Wolbachia</i> sp. endosymbiont of <i>Drosophila melanogaster</i></p>	<p><i>Bacillus halodurans</i> C-125 <i>Bacillus subtilis</i> subsp. <i>subtilis</i> str. 168 <i>Chlamydia trachomatis</i> D/UW-3/CX <i>Chlamydophila pneumoniae</i> CWL029 <i>Corynebacterium diphtheriae</i> NCTC 13129 <i>Deinococcus radiodurans</i> R1 <i>Mycobacterium leprae</i> TN <i>Staphylococcus aureus</i> subsp. <i>aureus</i> Mu50 <i>Thermus thermophilus</i> HB27 <i>Escherichia coli</i> K12 <i>Salmonella enterica</i> subsp. <i>enterica</i> serovar <i>Typhi</i> str. CT18 <i>Salmonella enterica</i> subsp. <i>enterica</i> serovar <i>Typhi</i> Ty2</p>

Supplementary Table 4.2. List of primers used in this study.

Description	Sequence
Amplification of <i>A. baylyi</i> ADP1 gDNA sequence pBAC883:	
0878F_EcoR	5' – CTTGAATTCGGAGCGCAGTTAT – 3'
0878R_Hind	5' – GTGACTGAACTCGATGAAGGC – 3'
Amplification of ACIAD0878 for insertion into pTrcHis:	
pTH0878F	5' – GTTCTTGCTAGCGATGCACCTTCTGA – 3'
pTH0878R	5' – GTTCTTAAGCTTTTAAAATTGTGGTTT – 3'

CHAPTER 5
DISCOVERY AND CHARACTERIZATION OF A COBALT-DEPENDENT,
OXYGEN UTILIZING BACTERIAL COPROPORPHYRINOGEN
DECARBOXYLASE ¹

¹ Boynton TO, Gerdes S, Dailey TA, Hobson D, Neidle EL, Phillips JD, and HA Dailey
- Submitted to: *PNAS*, 04/28/2011

Abstract

The antepenultimate step in protoheme biosynthesis is the oxidative decarboxylation of the propionates on the A and B ring to vinyl moieties. In prokaryotes this may occur via one of two distinct enzymes, an oxygen-dependent HemF that is homologous to eukaryotic coproporphyrinogen oxidase (CPOX), and an oxygen-independent member of the radical SAM superfamily, HemN. HemF is found in a minority of Gram negative bacteria and although hemN has been annotated in most bacterial genomes, it has been experimentally verified as a coproporphyrinogen decarboxylase (CpdH) only in *Escherichia coli*. Given the lack of evidence and the occurrence of postulated hemNs in bacteria that do not synthesize heme, we examined the validity of hemN annotations by utilizing the SEED bioinformatic tools and functional complementation of a Δ *hemF* *Acinetobacter baylyi* ADP1. Our data support the existence of four distinct “hemN” families with only one that encodes a bona fide HemN (CpdH). CpdH is present only in the γ -proteobacteria leaving all Gram positive bacteria without a verified enzyme for this step. The SEED tools were employed to identify new putative enzymes for this step and one of these was identified as a bona fide Firmacutes-specific coproporphyrinogen oxidative decarboxylase (CpoD). Expression and characterization of the recombinant CpoD demonstrated that this unique protein displays activity *in vivo* and *in vitro*, requires the presence of Co^{2+} , and is oxygen-dependent.

Introduction

Tetrapyrroles are one of the more ancient and broadly employed groups of cofactors (32). The most prevalent types of these are iron-bound heme and magnesium-bound chlorophyll. While the latter is essential for light absorption during photosynthesis, the former exhibits a wide range of chemical properties such as electron transport in cytochromes (10), anabolic catalysis in peroxidases and catalases (50), and binding of diatomic gases in a variety of sensor proteins (29, 37). Heme-containing proteins are also known to play a crucial role in pathogenesis in several bacteria, including truncated hemoglobins in *Mycobacterium* and *Campylobacter* that scavenge reactive nitrogen species produced by host immune systems (3, 4), and heme-sensing and –transporting proteins in *Staphylococcus* that modulate virulence (43, 46). The biosynthesis of these crucial molecules must be tightly regulated as the photo-reactive intermediates are highly toxic to cells (2).

The formation of all tetrapyrroles begins with the same four steps, resulting in the conversion of δ -aminolevulinic acid to uroporphyrinogen III (1). In protoheme IX and chlorophyll biosynthesis, uroporphyrinogen III is further decarboxylated and oxidized to protoporphyrin IX by the actions of uroporphyrinogen decarboxylase (UROD), coproporphyrinogen oxidase (CPOX), and protoporphyrinogen oxidase (PPOX). These last two steps are shown in Figure 5.1. The initially identified forms of both CPOX and PPOX were shown to utilize molecular oxygen during catalysis and hence were termed oxidases.

However, it is clear that heme-synthesizing organisms growing under anaerobic conditions must have alternate mechanisms to catalyze these two steps.

To date, an oxygen-dependent enzyme, HemF, and an oxygen-independent enzyme, HemN, have been described for the CPOX step (49). HemF is a cofactorless enzyme homologous to eukaryotic CPOX that requires molecular oxygen to convert the propionate side chains to vinyl groups (22). HemN utilizes a radical-S-adenosylmethionine (SAM) requiring system (26). Since HemN does not require oxygen, it is more appropriately named coproporphyrinogen dehydrogenase (CpdH) (23). For the final step in protoporphyrin synthesis an oxygen-dependent PPOX, HemY, has been identified in Actinobacteria, Firmicutes, and a few Gram-negative organisms. An oxygen-independent "PPOX," HemG, has been identified and characterized, but is restricted to the γ -proteobacteria (6). Like CpdH, this enzyme is oxygen-independent and is best referred to as protoporphyrinogen dehydrogenase (PpdH) rather than oxidase. Recently a third enzyme, HemJ, has been identified in the remaining Gram negative bacteria that lack HemY and HemG and while it is suggested to utilize molecular oxygen, its catalytic mechanism remains unknown (7, 21).

With HemY, HemG and HemJ, the vast majority of bacteria for which complete sequenced genomes exist have identifiable enzymes to catalyze the penultimate step in protoheme synthesis. The same cannot be stated for the antepenultimate step. HemF is present in only a minority of Gram negative bacteria and while most bacteria have annotated genes for "probable oxygen-

independent coproporphyrinogen III oxidases” there is experimental validation for a bona fide HemN only in *Escherichia coli*. The annotation within this gene family has never been critically examined and the probability is that many “*hemN*” genes do not code for CpdH leaving the possibility that an as yet unidentified and perhaps unique enzyme remains to be discovered and characterized.

Herein, we have employed the SEED platform (www.theseed.org) (34, 35) to analyze bacterial genes currently annotated as coproporphyrinogen oxidase and to identify putative new enzymes for this step in heme biosynthesis. This bioinformatics approach coupled with experimental verification has allowed us to identify a new enzyme with CPOX/CpdH-like activity in the Firmicutes. This soluble protein has no sequence resemblance to either HemF or HemN, is oxygen-dependent, and requires Co^{2+} for activity.

Results

Analysis of current hemN annotations

The legitimacy of many genes annotated as *hemN* is questionable given that they exist in bacteria that do not synthesize heme, or they co-exist with hemF in strict aerobes. Additionally it has been demonstrated experimentally that some genes annotated as *hemN* or “probable oxygen-independent coproporphyrinogen III oxidases” do not function as bona fide HemN (48). Our initial goal was to determine which currently annotated hemNs were true coproporphyrinogen dehydrogenases and which were incorrectly annotated

proteins that serve other functions in the organisms that possess them. To do this comparative genomics were employed via the SEED genomic platform.

All genes annotated as oxygen-independent coproporphyrinogen oxidases were projected across the previously populated SEED heme biosynthetic subsystem and examined using genomic context and amino acid sequence comparisons. The examined genes were found to group into four distinct families within the radical SAM enzyme superfamily (Table 5.1). The first of these, group A, represents the bona fide HemN and contains the experimentally validated *E. coli* HemN. Group B contains the proteins that are best exemplified by the *Vibrio cholerae* HutW protein that has been shown to not have HemN activity (48). The group C enzyme family is found in a wide variety of organisms, including many incapable of heme biosynthesis such as *Haemophilus* species, and these genes are often found clustered with genes related to nucleotide metabolism. Finally, group D contains genes labeled *hemN* in the Firmicutes and Actinobacteria, including *Streptococcus* species that do not synthesize heme.

The differences among these four groups are clearly shown when their amino acid sequences are viewed together. Multiple sequence alignment using ClustalW reveals little homology among these groups (Figure S5.1) and the conserved regions that do appear correspond to those found in all radical SAM enzymes, e.g. the Fe-S cluster motif (CXXXCXXC) and S-adenosyl methionine binding sites. Furthermore, the N-terminal $_{20}\text{GPRYTSYPTA}_{29}$ and $_{308}\text{KNFQGYTT}_{315}$ motifs of *E. coli hemN*, reported as integral to CpdH activity,

are present only in the group A family (25). Thus, these data strongly support the proposition that only group A contains a bona fide CpdH.

Creation of hemF knockout mutant

While our *in silico* predictions suggested that only the group A *hemN* family encodes functional oxygen-independent CpdHs, the possibility remained that genes within the other groups could also encode functional enzymes. To decipher the *in vivo* functionality of the various radical SAM protein families as potential CpdHs, we developed a complementation system in which to test members of the four families. To do this we deleted the gene ACIAD3250 (annotated as *hemF*) of *Acinetobacter baylyi* ADP1. The entire coding sequence of *hemF* was replaced via homologous recombination with a 3.6 kb cassette bearing both a kanamycin antibiotic resistance marker and a *sacB* sucrose lethality marker (Figure 5.2). While HemF of *Acinetobacter* has never been characterized, the protein bears 60% identity with the extensively studied *E. coli* and *Salmonella* enzymes. The resulting mutant strain, named ACN1156, is auxotrophic for heme with wildtype growth levels restored by the supplementation of exogenous hemin in growth media (Figure 5.3A). This is consistent with the encoded protein's function in heme biosynthesis and the designation of the gene as essential in the *A. baylyi* ADP1 knockout collection generated by de Berardinis *et al* (13). The heme auxotrophy exhibited by ACN1156 reinforces the designation of ACIAD3250 as a bona fide oxygen-dependent coproporphyrinogen oxidase (HemF).

Interestingly, *A. baylyi* ADP1 contains ACIAD0432, a member of the group B radical SAM “HemN” protein family. The heme auxotrophy of the ACIAD3250 knockout, despite the presence of ACIAD0432 in the genome, supports the contention that the putative group B hemNs are not functional CpdHs.

Experimental evaluation of putative hemNs via complementation

Functional complementation of heme biosynthetic mutants has provided valuable data in the identification of novel enzymes. Here, we have used the ACN1156 mutant as a platform for both rapid validation of *hemN*-annotated gene families and screening of putative coproporphyrinogen oxidases. Homologous recombination where the resistance marker of ACN1156 was replaced by DNA encoding the enzyme under examination resulted in single gene replacement driven by the natural promoter of *hemF*. Thus, unlike plasmid-driven complementation techniques, this system ensures the same physiological expression profile as the wild type gene.

Although *A. baylyi* ADP1 contains a number of annotated radical SAM proteins (including ACIAD0432 mentioned above), there was concern that given the reported oxygen sensitivity of radical SAM enzymes, even a functional HemN would not complement the aerobic *A. baylyi* ADP1. However, expression of the *E. coli hemN* (the chosen representative of Group A) eliminated the heme auxotrophy of ACN1156, thereby providing validation of this complementation system to evaluate putative *hemNs*.

The following ORFs were tested for complementation: *E. coli hemN* (Group A), *Chloroflexus aurantiacus* Caur_0209 (group C), *Mycobacterium tuberculosis hemN* (group D), *Bacillus subtilis hemN* (group D), and *B. subtilis hemZ* (group D). Genes from group B encode members of the HutW family and have previously been ruled out as functional CpdH enzymes, and thus were not retested here (48). Although a group C gene is present in the ACN1156 genome (ACIAD0432), the possibility exists that this gene is not functional under our selected growth conditions, so we tested Caur_0209 placed behind the *hemF* promoter. Of all genes tested, only *E. coli hemN* restored growth to the *hemF*-deficient ACN1156 (Figure 5.3B). These data provide strong support for the bioinformatics-based predictions that only Group A HemNs are functional CpdHs.

Identification of putative bacterial coproporphyrinogen oxidases

The above data demonstrated that only a subset of genes annotated as *hemN* encode functional CpdH enzymes. These data further reveal that heme-synthesizing organisms in the following taxa contain neither *hemF* nor group A *hemN*: Acidobacteria, Actinobacteria and Firmicutes. These organisms contain enzymes for all other steps of the heme biosynthetic pathway. Since genomic data on the Acidobacteria is limited, we focused on the Actinobacteria and Firmicutes.

The first step was to identify genes that might encode the missing enzyme(s). To do this the Signature Tool within the SEED database (www.nmpdr.org/FIG/wiki/view.cgi/FIG/SigGenes) was used to generate a list of

genes containing the desired occurrence profile: i.) candidate enzymes should be present only in genomes containing other genes of heme biosynthesis and should never be present in organisms incapable of heme biosynthesis and ii.) candidates may be present along with *hemF* and *hemN* since these are known to co-occur in some organisms (17, 38). Co-localization of candidates with other genes of heme biosynthesis was not used as a limiting factor since previous analysis of small-colony variants of *Staphylococcus aureus* established the existence of heme biosynthetic genes outside of the two known gene clusters (45). Additionally available gene expression patterns performed in *S. aureus* by the Dunman group (5, 30, 47) were used to limit the list of potential candidates to only hypothetical genes displaying expression patterns similar to that of other genes within the heme branch (i.e. *hemE*, *hemY*, *hemH*).

Analysis of the available 690 bacterial genomes yielded three prime candidates (Table 5.2). The first candidate gene family, HemQ, has previously been shown to be essential in heme synthesis, although a function has not been assigned to it. HemQ is present in both phyla examined and its relationship with HemY and HemH (11) makes it an excellent candidate. The second gene family consists of annotated hypothetical ORFs with no known function. Examples of this family are *ytpQ* in *Bacillus subtilis* and SAV1743 in *S. aureus*. Homologues of these genes are found in the Bacillales order of the Firmicutes and cover all heme-synthesizing organisms within the phyla. However, they are not present in the Actinobacteria. The final candidate family is classified as encoding a putative

lipoprotein and is exemplified by *ygaE* in *B. subtilis* and SAV1865 in *S. aureus*. It is also absent outside of the Bacillales order.

One or more members of each of these families was selected to be tested in the complementation system outlined above. The genes selected were: *Mycobacterium tuberculosis hemQ*, *B. subtilis hemQ*, *B. subtilis ytpQ*, and *B. subtilis ygaE*. The only successful complementation was from *ytpQ* (Figure 5.3B), leading us to propose that this gene encodes an enzyme capable of converting coproporphyrinogen III to protoporphyrinogen IX.

Expression and biochemical characterization of B. subtilis YtpQ

Sequence analysis of *B. subtilis* YtpQ, which is a member of the DUF1444 gene superfamily, reveals a protein sequence that is conserved in the Firmicutes (Figure S5.2) but with only 30% sequence identity. Phylogenetic clustering of this gene family is shown in Figure S5.3. Homologues are not present outside this group nor in Firmicutes that do not synthesize heme.

YtpQ was cloned as an in-frame fusion behind a amino-terminal six histidine-tagged *E. coli* maltose binding protein (MBP). Cleavage of the purified fusion protein with TEV protease followed by another round of purification yielded an untagged protein purified to homogeneity at yields of approximately 30 mg protein/L culture. SDS-PAGE analysis revealed a roughly 30 kDa band, which is in line with the calculated theoretical molecular weight of 31 kDa (Figure 5.4A). FPLC of YtpQ under low ionic conditions gave a single peak corresponding to its observed molecular weight, suggesting that YtpQ exists in solution as a soluble

monomer. The UV-visible spectrum of active recombinant YtpQ (see below) displayed a peak associated with the protein and a small shoulder feature in the 325-350 nm region (Figure 5.4B).

To initially test YtpQ for enzyme activity we utilized a two-enzyme linked assay described previously (33). This assay employs the use of both a functional CPOX enzyme and the following enzyme in the heme biosynthetic pathway, protoporphyrinogen oxidase (PPOX), to convert coproporphyrinogen to protoporphyrinogen and then protoporphyrinogen into protoporphyrin which can be quantitated by fluorescence (39). Purified YtpQ protein was assayed and displayed no measurable activity. A wide range of enzymatic cofactors (Table S5.1) were tested for their ability to stimulate enzyme activity and of these only Co^{2+} in the form of CoCl_2 resulted in support of measurable enzyme activity. In the presence of 10 μM CoCl_2 , 100 nM YtpQ displayed enzyme activity of 3.5 min^{-1} (Figure 5.4C). This activity compares with published rates of 2 min^{-1} and 0.11 min^{-1} for purified yeast and *E. coli* HemF, respectively (8, 9). Addition of 100 μM EDTA reversed the stimulation seen, resulting in no activity above background (data not shown). Although cobalt has been reported to substitute for physiological cofactors such as zinc or magnesium in some enzymes (16, 42), we were unable to obtain significant activity with other metals tested (Figure 5.4D). The fact that the as isolated YtpQ needs cobalt supplementation is not surprising given that the purification scheme employs a cobalt metal affinity column where the protein is washed off the column with high concentrations of free imidazole that might be expected to also strip Co from the protein. To

determine if Co is required as part of a holo-enzyme, purified YtpQ was incubated with excess CoCl_2 and then dialyzed overnight against cobalt-free buffer. The isolated enzyme exhibits spectral characteristics of a cobalt-containing protein (Figure 5.4B) and metal analysis of the protein revealed 0.80 equivalents of cobalt (Table 5.3).

The kinetic parameters with respect to the substrate coproporphyrinogen were determined for YtpQ in the presence of cobalt and yielded apparent K_m and K_{cat} values of $37 \mu\text{M}$ and 14 min^{-1} respectively (Figure 5.4E). The K_m value is in line with those reported for various CPO enzymes, which range from $<1\text{-}50 \mu\text{M}$ (19). HemF, but not HemN, uses oxygen to carry out catalysis. Since the coupled YtpQ enzyme assays were carried out aerobically, the possibilities existed that YtpQ uses oxygen as part of its reaction mechanism, or that it is oxygen-independent and not affected by oxygen. To examine this, additional assays were performed with YtpQ in which buffers were held under vacuum and sparged with nitrogen gas, and reactions carried out under nitrogen. While not anaerobic, the limited oxygen remaining in these assays may be insufficient to support full activity if molecular oxygen is required in catalysis by YtpQ. Alternatively, lowered oxygen concentrations would have no impact on an oxygen-independent enzyme. The results (Table 5.4) show that when oxygen is limited, the specific activity of the enzyme is decreased ten-fold, which is consistent with YtpQ utilizing molecular oxygen as a substrate.

Discussion

Bacteria are known to possess two enzymes, HemF and HemN, capable of converting coproporphyrinogen III into protoporphyrinogen IX. Bacterial HemF is an oxygen-dependent CPO similar in structure and function to eukaryotic CPOX and requires no cofactor to catalyze the decarboxylation and oxidation of two propionate side chains to yield vinyl groups at the 2, 4 positions on the A and B pyrrole rings (27, 36). However, the gene for HemF is absent in all Firmicutes and Actinobacteria and some Gram negative bacteria. HemN has been identified as an oxygen independent CpdH that is a member of the “radical SAM” class of enzymes and possesses a [4Fe-4S] cluster (24). Many organisms have an annotated hemN and/or hemZ as an “oxygen independent CPO-like” gene in their genome and because of this, there has been a widespread assumption that the conversion of coproporphyrinogen to protoporphyrinogen in bacteria is catalyzed by either HemF or HemN/Z. As detailed above, we demonstrate that HemN is present in only a subclass of the γ -proteobacteria including *Aeromonas*, *Enterobacteria*, *Shewanella*, and *Vibrio* and not in the Firmicutes or Actinobacteria.

Our data demonstrating that neither *B. subtilis* *hemN* nor *hemZ* complement a *hemF* knockout of *A. baylyi* ADP1 are not consistent with an earlier publication that proposed both *hemN* and *hemZ* code for functional CpdHs (18). These workers reported that a *Salmonella typhimurium* *hemF/hemN* double mutant was complemented by either *B. subtilis* *hemN* or *hemZ*, but the doubling time with either of these (i.e. 2-2.5 hrs under aerobic conditions and 5-11.5 hrs under anaerobic conditions) was considerably slower than that of wild-type *S.*

typhimurium. In the previous study neither HemN nor HemZ of *B. subtilis* was demonstrated *in vitro* to have CpdH activity and a knockout of both *hemN* and *hemZ* was not lethal as one would expect. Perhaps most significant, however, is that the *Bacillus* proteins lack over forty amino-terminal residues as well as a significant number of conserved residues that have been identified from the *E. coli* HemN crystal structure as being involved in substrate binding and domain interactions (25). These investigators proposed that a third coproporphyrinogen decarboxylase was present in *B. subtilis*. From our results it is clear that *ytpQ* encodes the bona fide coproporphyrinogen decarboxylase of *Bacillus* with HemN and HemZ being proteins whose primary functions are ones other than heme synthesis.

YtpQ is a member of the DUF1444 superfamily, a group of bacterial proteins with no previously known function that are found only in the Firmicutes. We propose to name proteins in this group coproporphyrinogen oxidative decarboxylase (CpoD). YtpQ contains 270 amino acids and its sequence has no similarity to either HemF or HemN. Within the DUF1444 superfamily there is less than 30% amino acid sequence identity and among identical residues there are no cysteine or tryptophan residues and only two glutamate, five aspartate, two histidine, and two arginine residues (Figure S5.2). It is distinct from HemN in that it does not contain a [4Fe-4S] cluster and requires oxygen for activity. It differs from HemF in that it requires the presence of a metal for activity. Thus, the reaction it catalyzes may be unique from that of HemN or HemF.

Both HemF and HemN are reasonably well characterized enzymes that catalyze the formation of protoporphyrinogen IX from coproporphyrinogen III. In doing this, both release CO₂ and function in a stepwise fashion starting at the A pyrrole ring with the production of harderoporphyrinogen followed by decarboxylation of the B ring propionate (44). For human CPOX, experiments employing mesoporphyrinogen VI, which has ethyl rather than propionate side chains on the C and D rings, and site directed mutagenesis of select Arg residues, showed that the 6, 7 propionates on the C and D rings are not required for the first decarboxylation to occur, but were essential to decarboxylate the B ring propionate (44). These data clearly indicate that reorientation of the harderoporphyrinogen must occur prior to the second decarboxylation and that one or both of the non-reacting propionates is essential for proper spatial positioning of the intermediate substrate. HemF uses molecular oxygen in the absence of any metal or organic cofactor. Based upon the crystal structure of yeast and human CPOX (without substrate bound), decades of kinetic studies, and recent density function theory analysis (41), a model for catalysis proposed by Lash (22) has current acceptance. In this model the A ring pyrrole nitrogen is first deprotonated, thereby allowing the formation of a pyrrole peroxide anion by reaction with molecular oxygen. The pyrrole nitrogen is proposed to interact with an active site aspartate side chain in a fashion similar to what has been shown with uroporphyrinogen decarboxylase, the preceding pathway enzyme that carries out a non-oxidative decarboxylation without a cofactor (40). In contrast, HemN, a radical SAM enzyme, is oxygen independent. Jahn and colleagues

have presented evidence that formation of a substrate radical (derived directly from the 5'-deoxyadenosyl radical) by hydrogen abstraction at the β carbon of the propionate side chain is the first step in the reaction sequence (26). Thus, HemN and HemF differ in their mechanisms for conversion of coproporphyrinogen to protoporphyrinogen.

Since the activity catalyzed by CpoD is the oxidative decarboxylation of two propionate side chains to two vinyl groups and presumably carbon dioxide, the requirement for cobalt is interesting. One other decarboxylase, α -amino- β -carboxymuconate- ϵ -semialdehyde decarboxylase (ACMSD), was initially reported to contain an essential cobalt, although the crystal structure of the recombinant protein expressed in *E. coli* shows a zinc (15, 28, 31) (a cobalt-containing structure was also determined). Extended soaking of the apo-ACMSD with zinc yields an active enzyme similar in properties to the Co-containing enzyme. The role of the active site Zn is unknown but it is coordinated by three conserved histidyl side chains, which are lacking in the CpoD family of proteins. One catalytic model for ACMSD has the metal interacting with the leaving carboxylate group while a second model has the leaving carboxylate interacting with an Arg side chain and a metal-bound hydroxyl group attacking the C2 carbon. However, unlike ACMSD, CpoD carries out an oxidative decarboxylation and requires molecular oxygen along with cobalt.

If the cobalt in CpoD is redox inactive and serves only a non-catalytic role, such as interactions to immobilize a propionic side chain or maintain protein structure, then the Lash catalytic model may be appropriate. However, the

stringent requirement for cobalt seems more in line with a role for the metal in catalysis, possibly involving a radical mechanism and Co^{2+} - Co^{3+} cycling being driven by O_2 . If cobalt functions in radical formation, then a mechanism involving a substrate radical similar to what has been proposed for HemN may be appropriate. Future studies will resolve this issue.

CpoD represents yet another example of the diversity exhibited by prokaryotic organisms in the terminal steps of heme synthesis. Separate, distinct enzymes are now known to catalyze the antepenultimate (HemF, HemN, CpoD) and penultimate (HemG, HemJ, HemY) steps in this otherwise well-conserved pathway. Since CpoD is only present in the Firmicutes, an as yet to be identified enzyme must exist in the Actinobacteria.

Materials and Methods

Bacterial strains, plasmids, and growth conditions

Bacterial strains and plasmids used or produced during the course of this study are listed in Table S5.2. *A. baylyi* ADP1 and *E. coli* strains were grown in LB medium at 37 °C unless otherwise noted. Where appropriate, media was supplemented with 30 µg/ml kanamycin, 100 µg/ml ampicillin, and 10 µg/ml hemin. For initial pUGA1 construction, the 500 bp segment immediately upstream of the *A. baylyi* ADP1 *hemF* gene was PCR amplified to contain a 5' HindIII site and a 3' BamHI site. The 500 bp segment immediately downstream of the gene was PCR amplified to contain a 5' BamHI site and a 3' EcoRI site. The two segments were ligated in tandem using the BamHI sites and inserted

into pUC19 using the HindIII and EcoRI sites. The resulting plasmid was termed pUGA1. Plasmids derived from pUGA1 were created by PCR amplification of target genes with 5' and 3' BamHI restriction sites and insertion into the BamHI site of pUGA1, except for pUGA2, where the *sacB*-Kan^R cassette was excised directly from pRMJ1 (20), and *M. tuberculosis hemQ*, where BglII gene flanking sites were used. Orientation of each insertion was screened via natural restriction sites within each gene. For protein expression, pTHMBPYtpQ was created by inserting *Bacillus subtilis ytpQ* into pTrcHisA as an in-frame fusion with the *E. coli malE* gene, with a TEV protease cleavage site between the two ORFs.

Generation of a hemF knockout in A. baylyi ADP1

The *A. baylyi* ADP1 *hemF* knockout was generated by homologous recombination, replacing ORF ACIAD3250 (*hemF*) with a *sacB*/Kan^R cassette allowing for selection by antibiotic resistance and sucrose lethality. Wildtype *A. baylyi* ADP1 was naturally transformed with a PCR amplicon containing the 500 bp upstream and 500 bp downstream of ACIAD3250 with the cassette inserted in between. Transformations were conducted as previously described (7). Briefly, 1 µl of DNA was mixed with 1 µl of an overnight *A. baylyi* culture and spotted onto LB agar and incubated for 24 h. Resulting growth was then plated onto LB agar containing hemin and kanamycin. Successful replacement of *hemF* was verified by acquired antibiotic resistance and auxotrophy for heme, as well as PCR of genomic DNA.

Complementation of hemF knockout

For complementation and analysis of gene function, the newly created ACN1156 ($\Delta hemF$) strain was transformed in the same fashion as described above, but with various genes (Table S5.2) cloned into pUGA1 in place of the *sacB/Kan^R* cassette. The DNA and overnight culture mixture was spotted onto LB agar containing kanamycin and hemin, and after 24 hours replated on LB agar. Growth comparable to wildtype levels without exogenous hemin and lack of Kan^R was used to assess positive complementation. Direct sequencing of each replacement was used for final verification of the genotype.

Protein expression, purification, and initial characterization

Expression and purification of *B. subtilis* YtpQ was done as previously described for HemG (6), except that the protein was fused to maltose binding protein with a TEV cleavage site between them. For expression, *E. coli* cells harboring pTHMBPYtpQ were grown for 8 hours at 30° C in 100 mL Circlegrow media (MP Biomedical, Solan, OH) with ampicillin and then transferred to 1L media and grown for an additional 22 hours at the same temperature. Cells were harvested by centrifugation at 5000 x *g* for 10 minutes and frozen at -80° C. For purification, cell pellets were suspended in previously described solubilization buffer, sonicated, and centrifuged at 100,000 x *g* for 30 minutes. The resulting supernatant was applied to a HisPur Cobalt resin (Thermo Sci., Rockford, IL) column, washed, and eluted. The resulting fusion protein was incubated with AcTEV protease (Invitrogen, Carlsbad, CA) per company directions and applied

to a HisPur Cobalt resin column to remove MBP and the his-tagged protease. Protein purity was assessed by SDS-PAGE. Spectrophometric analysis was done using a Cary 1G spectrophotometer (Varian, Palo Alto, CA, USA). FPLC was done using an Aktaprime machine equipped with a Hi-Prep Sephacryl S-300 column (GE Healthcare, Piscataway, NJ) with a running buffer consisting of 100 mM Tris-HCl pH 8.0 and 20 mM KCl.

Coproporphyrinogen oxidase assays

Coproporphyrinogen oxidase activity was measured using a combined method of two published procedures (33, 39), employing a coupled reaction with YtpQ and human protoporphyrinogen oxidase (PPOX). PPOX was added in ten-fold excess to ensure any rates observed were attributable to the YtpQ enzyme. Reaction mixtures consisted of 50 mM NaH₂PO₄ pH 8.0, 0.2% (w/v) Tween 20, 2.5 mM glutathione, 100 nM YtpQ, 1 μM PPOX, and varying amounts of the substrate coproporphyrinogen III. Coproporphyrinogen III was produced from coproporphyrin III using sodium amalgam as described previously (39). For cofactor determination, 10 μM of the cofactors listed in Table S5.1 were separately tested in reaction mixtures initially. Subsequent assays contained 1 μM CoCl₂ once the requirement for cobalt was determined. Product formation was monitored by fluorescence with a Synergy HTI plate reader (BioTek, Winooski, VT).

In assays to determine oxygen requirements, a coupled reaction was not possible since PPOX requires molecular oxygen for activity (12). For these

reactions assays were conducted as above except that buffers, reagents, and protein were all sparged with nitrogen gas before use and no PPOX was present. Reactions were carried out in sealed tubes under nitrogen and stopped after 5 minutes with an equal volume of 2.7 N HCl. HPLC was then used to quantify protoporphyrin levels as previously described (14).

Metal analysis

For metal analysis, 1 mL 300 μ M purified YtpQ with 1 mM CoCl_2 was incubated for 1 hour and then dialyzed against 4 L of solubilization buffer overnight at 4 °C. After dialysis, samples were analyzed using a Thermo Jarrell-Ash 965 (Franklin, MA) inductively coupled argon plasma (ICP) spectrometer at the University of Georgia Chemical Analysis Laboratory.

References

1. **Ajioka RS, Phillips JD, & Kushner JP.** 2006. Biosynthesis of heme in mammals. *Biochim Biophys Acta.* 1763:723-736.
2. **Anzaldi LL & Skaar EP.** 2010. Overcoming the heme paradox: heme toxicity and tolerance in bacterial pathogens. *Infect Immun.* 78:4977-4989.
3. **Ascenzi P, Bolognesi M, Milani M, Guertin M, & Visca P.** 2007. Mycobacterial truncated hemoglobins: from genes to functions. *Gene.* 398:42-51.
4. **Ascenzi P & Visca P.** 2008. Scavenging of reactive nitrogen species by mycobacterial truncated hemoglobins. *Methods Enzymol.* 436:317-337.
5. **Bischoff M, et al.** 2004. Microarray-based analysis of the *Staphylococcus aureus* sigmaB regulon. *J Bacteriol.* 186:4085-4099.
6. **Boynton TO, Daugherty LE, Dailey TA, & Dailey HA.** 2009. Identification of *Escherichia coli* HemG as a novel, menadione-dependent flavodoxin with protoporphyrinogen oxidase activity. *Biochemistry.* 48:6705-6711.
7. **Boynton TO, et al.** (in press). Discovery of a unique bacterial protoporphyrinogen oxidase enzyme through comparative genomic analysis and functional complementation. *Appl Environ Microbiol.*
8. **Breckau D, Mahlitz E, Sauerwald A, Layer G, & Jahn D.** 2003. Oxygen-dependent coproporphyrinogen III oxidase (HemF) from *Escherichia coli* is stimulated by manganese. *J Biol Chem.* 278:46625-46631.

9. **Camadro JM, Chambon H, Jolles J, & Labbe P.** 1986. Purification and properties of coproporphyrinogen oxidase from the yeast *Saccharomyces cerevisiae*. Eur J Biochem. 156:579-587.
10. **Chance B.** 1972. The nature of electron transfer and energy coupling reactions. FEBS Lett. 23:3-20.
11. **Dailey TA, et al.** 2010. Discovery and Characterization of HemQ: an essential heme biosynthetic pathway component. J Biol Chem. 285:25978-25986.
12. **Dailey TA & Dailey HA.** 1996. Human protoporphyrinogen oxidase: expression, purification, and characterization of the cloned enzyme. Protein Sci. 5:98-105.
13. **de Berardinis V, et al.** 2008. A complete collection of single-gene deletion mutants of *Acinetobacter baylyi* ADP1. Mol Syst Biol. 4:174.
14. **Franklin MR, Phillips JD, & Kushner JP.** 1997. Cytochrome P450 induction, uroporphyrinogen decarboxylase depression, porphyrin accumulation and excretion, and gender influence in a 3-week rat model of porphyria cutanea tarda. Toxicol Appl Pharmacol. 147:289-299.
15. **Garavaglia S, Perozzi S, Galeazzi L, Raffaelli N, & Rizzi M.** 2009. The crystal structure of human alpha-amino-beta-carboxymuconate-epsilon-semialdehyde decarboxylase in complex with 1,3-dihydroxyacetonephosphate suggests a regulatory link between NAD synthesis and glycolysis. FEBS J. 276:6615-6623.

16. **Gonzalez JM, Buschiazzo A, & Vila AJ.** 2010. Evidence of adaptability in metal coordination geometry and active-site loop conformation among B1 metallo-beta-lactamases. *Biochemistry*. 49:7930-7938.
17. **Goto T, Aoki R, Minamizaki K, & Fujita Y.** 2010. Functional differentiation of two analogous coproporphyrinogen III oxidases for heme and chlorophyll biosynthesis pathways in the cyanobacterium *Synechocystis* sp. PCC 6803. *Plant Cell Physiol*. 51:650-663.
18. **Homuth G, Rompf A, Schumann W, & Jahn D.** 1999. Transcriptional control of *Bacillus subtilis* hemN and hemZ. *J Bacteriol*. 181:5922-5929.
19. **Jones MA, He J, & Lash TD.** 2002. Kinetic studies of novel di- and tri-propionate substrates for the chicken red blood cell enzyme coproporphyrinogen oxidase. *J Biochem*. 131:201-205.
20. **Jones RM & Williams PA.** 2003. Mutational analysis of the critical bases involved in activation of the AreR-regulated sigma54-dependent promoter in *Acinetobacter* sp. strain ADP1. *Appl Environ Microbiol*. 69:5627-5635.
21. **Kato K, Tanaka R, Sano S, Tanaka A, & Hosaka H.** 2010. Identification of a gene essential for protoporphyrinogen IX oxidase activity in the cyanobacterium *Synechocystis* sp. PCC6803. *Proc Natl Acad Sci USA*. 107:16649-16654.
22. **Lash TD.** 2005. The enigma of coproporphyrinogen oxidase: how does this unusual enzyme carry out oxidative decarboxylations to afford vinyl groups? *Bioorg Med Chem Lett*. 15:4506-4509.

23. **Layer G, Heinz DW, Jahn D, & Schubert WD.** 2004. Structure and function of radical SAM enzymes. *Curr Opin Chem Biol.* 8:468-476.
24. **Layer G, et al.** 2005. Structural and functional comparison of HemN to other radical SAM enzymes. *Biol Chem.* 386:971-980.
25. **Layer G, Moser J, Heinz DW, Jahn D, & Schubert WD.** 2003. Crystal structure of coproporphyrinogen III oxidase reveals cofactor geometry of Radical SAM enzymes. *EMBO J.* 22:6214-6224.
26. **Layer G, et al.** 2006. The substrate radical of *Escherichia coli* oxygen-independent coproporphyrinogen III oxidase HemN. *J Biol Chem.* 281:15727-15734.
27. **Layer G, Reichelt J, Jahn D, & Heinz DW.** Structure and function of enzymes in heme biosynthesis. *Protein Sci.* 19:1137-1161.
28. **Li T, Walker AL, Iwaki H, Hasegawa Y, & Liu A.** 2005. Kinetic and spectroscopic characterization of ACMSD from *Pseudomonas fluorescens* reveals a pentacoordinate mononuclear metallocofactor. *J Am Chem Soc.* 127:12282-12290.
29. **Lukin JA & Ho C.** 2004. The structure--function relationship of hemoglobin in solution at atomic resolution. *Chem Rev.* 104:1219-1230.
30. **Luong TT, Dunman PM, Murphy E, Projan SJ, & Lee CY.** 2006. Transcription Profiling of the mgrA Regulon in *Staphylococcus aureus*. *J Bacteriol.* 188:1899-1910.
31. **Martynowski D, et al.** 2006. Crystal structure of alpha-amino-beta-carboxymuconate-epsilon-semialdehyde decarboxylase: insight into the

- active site and catalytic mechanism of a novel decarboxylation reaction. *Biochemistry*. 45:10412-10421.
32. **Mauzerall DC**. 1998. Evolution of porphyrins. *Clin Dermatol*. 16:195-201.
 33. **Medlock AE & Dailey HA**. 1996. Human coproporphyrinogen oxidase is not a metalloprotein. *J Biol Chem*. 271:32507-32510.
 34. **Osterman A & Overbeek R**. 2003. Missing genes in metabolic pathways: a comparative genomics approach. *Curr Opin Chem Biol*. 7:238-251.
 35. **Overbeek R, et al**. 2005. The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. *Nucleic Acids Res*. 33:5691-5702.
 36. **Phillips JD, et al**. 2004. Crystal structure of the oxygen-dependant coproporphyrinogen oxidase (Hem13p) of *Saccharomyces cerevisiae*. *J Biol Chem*. 279:38960-38968.
 37. **Reinking J, et al**. 2005. The *Drosophila* nuclear receptor e75 contains heme and is gas responsive. *Cell*. 122:195-207.
 38. **Rompf A, et al**. 1998. Regulation of *Pseudomonas aeruginosa* hemF and hemN by the dual action of the redox response regulators Anr and Dnr. *Mol Microbiol*. 29:985-997.
 39. **Shepherd M & Dailey HA**. 2005. A continuous fluorimetric assay for protoporphyrinogen oxidase by monitoring porphyrin accumulation. *Anal Biochem*. 344:115-121.

40. **Silva PJ & Ramos MJ.** 2005. Density-functional study of mechanisms for the cofactor-free decarboxylation performed by uroporphyrinogen III decarboxylase. *J Phys Chem B.* 109:18195-18200.
41. **Silva PJ & Ramos MJ.** 2008. A comparative density-functional study of the reaction mechanism of the O₂-dependent coproporphyrinogen III oxidase. *Bioorg Med Chem.* 16:2726-2733.
42. **Smith RA, Latchney LR, & Senior AE.** 1985. Tight divalent metal binding to *Escherichia coli* F1-adenosinetriphosphatase. Complete substitution of intrinsic magnesium by manganese or cobalt and studies of metal binding sites. *Biochemistry.* 24:4490-4494.
43. **Stauff DL & Skaar EP.** 2009. The heme sensor system of *Staphylococcus aureus*. *Contrib Microbiol.* 16:120-135.
44. **Stephenson JR, et al.** 2007. Role of aspartate 400, arginine 262, and arginine 401 in the catalytic mechanism of human coproporphyrinogen oxidase. *Protein Sci.* 16:401-410.
45. **Tien W & White DC.** 1968. Linear sequential arrangement of genes for the biosynthetic pathway of protoheme in *Staphylococcus aureus*. *Proc Natl Acad Sci USA.* 61:1392-1398.
46. **Torres VJ, et al.** 2007. A *Staphylococcus aureus* regulatory system that responds to host heme and modulates virulence. *Cell Host Microbe.* 1:109-119.
47. **Weinrick B, et al.** 2004. Effect of mild acid on gene expression in *Staphylococcus aureus*. *J Bacteriol.* 186:8407-8423.

48. **Wyckoff EE, Schmitt M, Wilks A, & Payne SM.** 2004. HtzZ is required for efficient heme utilization in *Vibrio cholerae*. J Bacteriol. 186:4142-4151.
49. **Xu K, Delling J, & Elliott T.** 1992. The genes required for heme synthesis in *Salmonella typhimurium* include those encoding alternative functions for aerobic and anaerobic coproporphyrinogen oxidation. J Bacteriol. 174:3953-3963.
50. **Zamocky M, Furtmuller PG, & Obinger C.** 2010. Evolution of structure and function of Class I peroxidases. Arch Biochem Biophys. 500:74-81.

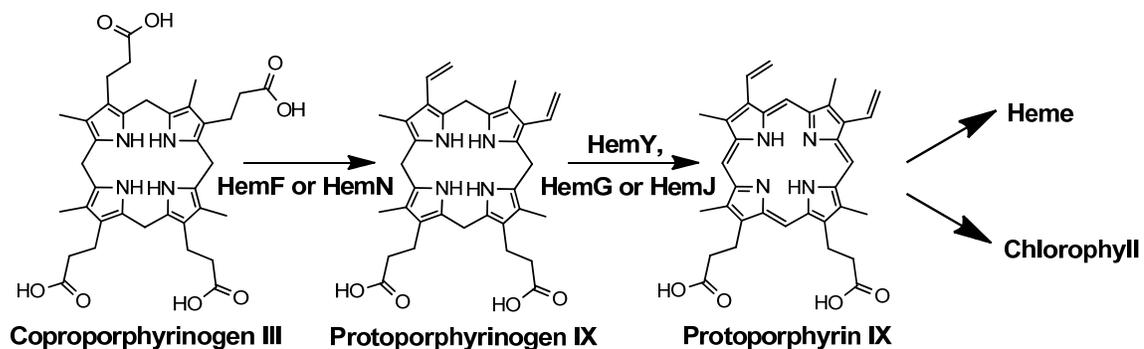


Figure 5.1. Terminal steps of protoporphyrin biosynthesis shared by the heme and chlorophyll pathways. Among prokaryotic organisms, multiple, distinct enzymes catalyze each of these steps. Oxidative decarboxylation of coproporphyrinogen III to protoporphyrinogen IX occurs through O_2 -dependent HemF or O_2 -independent HemN. Protoporphyrinogen is then oxidized to protoporphyrin by either O_2 -dependent HemY, O_2 -independent HemG, or the recently discovered HemJ.

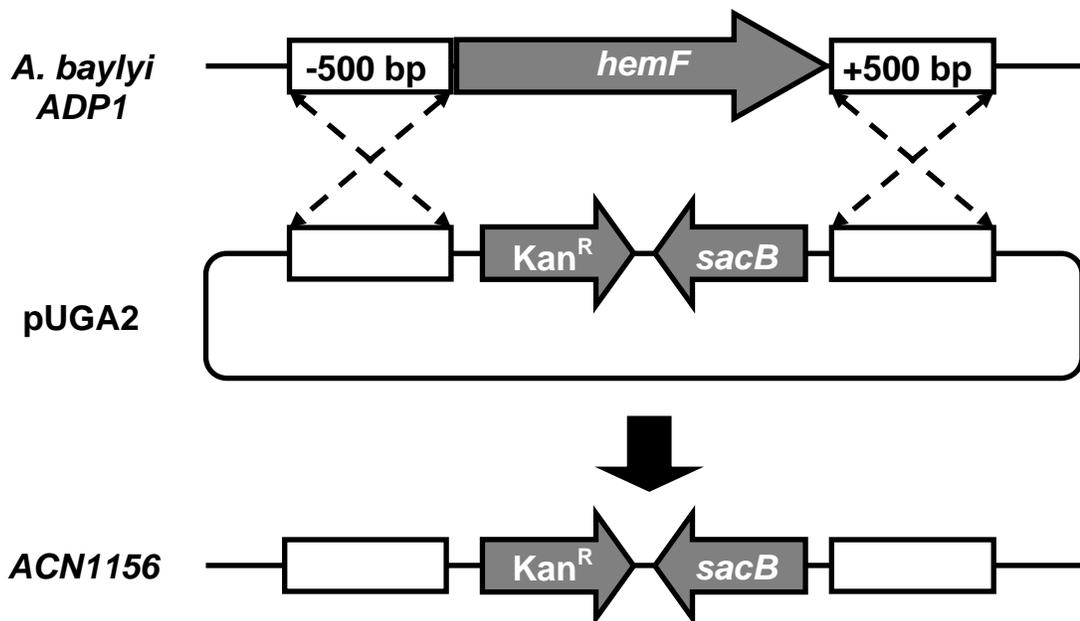


Figure 5.2. $\Delta hemF$ knockout generation in *A. baylyi* ADP1. The entire ACIAD3250 (*hemF*) ORF was replaced with a cassette containing kanamycin resistance and *sacB* sucrose lethality genes to create the strain ACN1156. Homologous recombination was driven by the presence of 500 bp ACIAD3250 flanking regions in pUGA2. Complementation of the knockout was achieved in a similar fashion, using test genes inserted into pUGA1 in place of the *sacB*/*Kan^R* cassette.

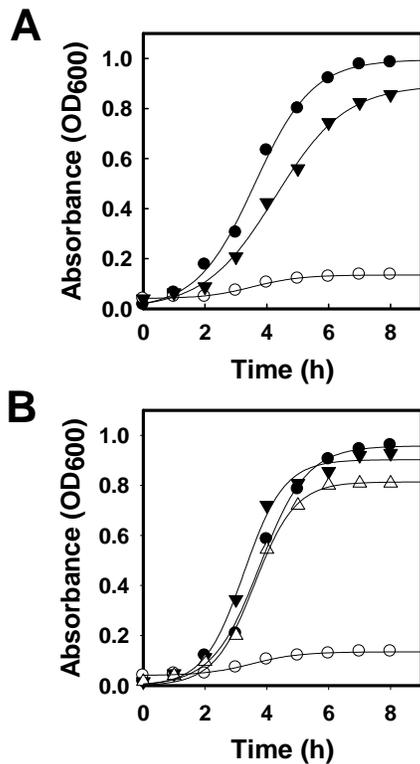


Figure 5.3. Characterization and functional complementation of the ACN1156 $\Delta hemF$ mutant, ACN1156. (A) Growth characteristics of wildtype *A. baylyi* ADP1 (●) and ACN1156 (○). Growth of ACN1156 required hemin supplementation (▼). (B) Functional complementation of ACN1156. Viability was restored to the mutant ACN1156 when the following genes were present (see Material and Methods): *E. coli hemN* (Δ), *Bacillus subtilis ytpQ* (▼), with *A. baylyi* ADP1 *hemF* (●) as a control. ACN1156 grown without exogenous heme is shown as a reference (○). Representative genes from the remaining radical SAM families failed to complement, as did the two remaining CPD candidates *hemQ* and *ygaE*.

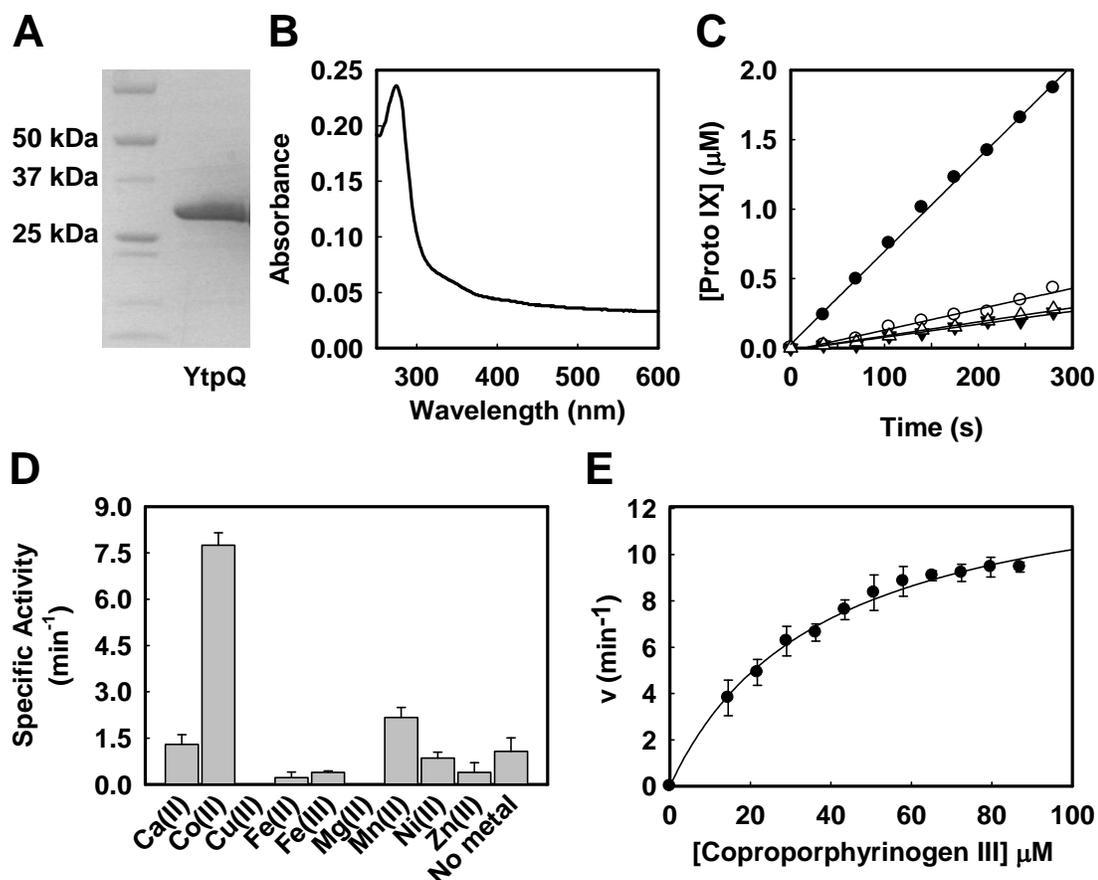


Fig. 4. Biochemical characterization of *Bacillus subtilis* ytpQ. (A) SDS-PAGE analysis of purified recombinant YtpQ. The band observed corresponds to an estimated molecular size of 30 kDa. (B) UV-Vis spectrum of YtpQ. The protein sample examined was incubated with cobalt and then dialyzed overnight as described in the text. The shoulder feature at ~ 350 nm is consistent with a cobalt-containing protein. (C) Typical enzymatic reaction of 100 nM YtpQ in the presence (\bullet) and absence (\circ) of 10 μM CoCl_2 . In the absence of YtpQ, Co^{2+} alone (\blacktriangledown) displays no activity as does a reaction containing only substrate (Δ). (D) Stimulation of YtpQ by various metals. (E) Determination of apparent Michaelis-Menton kinetics for YtpQ. 100 nM YtpQ was assayed with varying concentrations of substrate. Activity (v) is expressed as nmol protoporphyrin IX produced per minute per nmol YtpQ. The apparent K_m and K_{cat} values are 37 μM and 14 min^{-1} respectively.

```

Group A -----MSVQQIDWDLALIQKYNYSGPRYTSYPTALEFSEDFGEQAFLLQAVARYP 49
Group B MIMLILDNFDESILGANTPDPLRFQFNKHSAGGIAMPVPSHEQANVWQHITQQVSQR 60
Group C -----MVKLP 5
Group D -----

Group A ERPLSLYVHIPFCHKLCYFCGCNKIVTRQQHKADQYLDALQEIIVHRAPLFAGRH--VSQ 107
Group B QQVRCLYIHVPFCRVRCTFCNFFQNAAS-RQLVDAYFAALLEEIKQKAALPWTQTGVFHA 119
Group C --PLSLYIHIPWCVQKCPYCDFNSHALKGEVPHDDYVQHLLNDLNDVAYAQGRE--VKT 61
Group D --MKSAYIHIPFCEHICHYCDFNKYFIQ-SQPVDEYLNALQEMINTIAKTGQPD--LKT 55
          * * * * * * * * * *

Group A LHWGGGTPTYLNKAQISRLMKLLRENQFNADA-EISIEVDPREIELDVLHLRAEGFNR 166
Group B VYIGGGTPTELSPEQIRQLGTAIRESFPLTPDC-EITLEGRIHRFSDEMFFENALEGGFNR 178
Group C IFIGGGTPSLLSGPAMQTLLDGVRARLPLAADA-EITMEANPGTVEADRFVDYQRAGVNR 120
Group D IFIGGGTPTSLSEEQLKKLMDMINRVLKPPSSDLSEFAVEANPDDLSAEKLKILKEAGVNR 115
          ***** * * * * * * * * * *

Group A LSMGVQDFNKEVQRLVNREQDEEFIFALLNHAREIGFTSTNIDLIYGLPKQTPESFAFTL 226
Group B FSFGVQSFNTQVRRRAKRLDDREVVMERIASLAATQQAPIVIDLLYGLPYQTAQVFEQDL 238
Group C ISIGVQSFSEEKLRKLGRIHGPFQAKRAAKLASGLGLRSFNLDLMHGLPDQSLEEALGDL 180
Group D LSFGVQTFEDDLLEKIGRVHKQKDVFTSFERAREIGFENISLDLMFGLPGQTLKHLEHSI 175
          * * * * * * * * * *

Group A KRVAELNPDRLSVFNAYALP-TIFAAQRKIKDADLPSPQ-QKLDILQETIAFLTQSGYQF 284
Group B QDFMQTGAQGIDLYQLVVGGSAPMLNLVEKGLPPPATTDPKASLYQIGVEFMAKHHLRP 298
Group C RQAIELNPPHLSWYQLTIEPNTLFG----SRPPVLPDDD-ALWDFEQGHQLLTAAGYQQ 235
Group D NTALSLDAEHYSVYSLIVEPKTVFYNLQKGRHLHLPQE-QEAEMYEIVMSKMEAHGIHQ 234
          *

Group A IGMDFHARPDDELAVAQREGVLHKNFQGYTTQGDTDLLGMGVS AISMIGDCYAQNQKELK 344
Group B LSVNHWTRDNR-----ERSLYNSLAKTYAEVLPICGAGGNMGYSLSMQHRQLD 347
Group C YETSAYAKPG-----YQCQHNLNYWRFQDYIGIGCGAHGKVTFPDGRILRTTK 283
Group D YEISNFAKAG-----MESKHNLTYWSNEQYFGFGAGAHGYIGGTRTVNVGPVK 282
          * *

Group A QYYQQVDEQGNALWRGIALTRDDCIRRDVIKSLICNFRLDYAPIEKQWDLHFADYFAEDL 404
Group B TYLDAMK-NGQPLVA---MMARQHEYEPLFAALKAGFDSGVIQAKRQRLPKFYHHQTFDWM 403
Group C TRHPRGFMQGRYLES----QRDVEATDKPFEEFMNRFRLLLEAAPRVEFIAYTGLCEDVIR 339
Group D HYIDLIAEKGFPPYRD----THEVTTEEQIEEEMFLGLRKTAGVSKKRFAEKYGRSLDGLF 338
          *

Group A K-LLAPLAKDGLVDVDEKGIQVTAKGRLLIRNICMCFDYLRLQKARMQQFSRVI 457
Group B P-LFLRWQQIGLVEVEQDYLTLTAGRFWSVSLAQACIQVLIHSYKYQQORIA- 455
Group C P-QLDEAIAQGYLTECADYWITEHGKLFNLSLLELFLAE----- 378
Group D PSVLKDLAEKGLIHNSESAVYLTHQKLLGNEVFGAFLGEL----- 379
          * * *

```

Figure S5.1. Multiple sequence alignment of proteins annotated as “oxygen-independent CPOs” within databases. Each group represents the subfamilies found in Table 5.1. Among these groups, homologous regions are involved in Fe-S clusters (CXXXCXXC motif, highlighted black) and SAM binding (highlighted blue). Regions reported as integral for HemN function are found only in true HemN sequences (group A, highlighted red).

```

B. subtilis      -MKMTSRKLS DILKQRLQ-HENRSFLFDREKDTLRVEDQTTKKGITLDLP 48
G. stearotherm. ---MNSRQMYEKIKERLVAHPHWTFRFDAQDAMRVEDRRTKKGVTISLP 47
S. aureus       ---MNTFQMRDKLKERLS-HLDVDFKFNREEETLRIYRTDNNKGITIKLN 46
L. monocytogenes MAKMTTLKMKKEKLEKELQ-APNRQFSYNRDNDTLTVAQ--NGKKVTLTIP 47
                *                *                *                *      *

B. subtilis      PIIAKWELKKDEAIDEIVYYVSEAMTAMEGKAQEMTGKETRIYPVIRSTS 98
G. stearotherm.  GVIAKWHEQKDEAVREVVYYVEQTLKTMEEDA-ALSGNERNIYPVIRSAS 96
S. aureus       AIVAKYEDKKEKIVDEIVYYVDEAIAQMADKTLESISS-SQIMPVIRATS 95
L. monocytogenes QIIANFENDGNAAVEKIVYYVEEGFRAAAGNV-ELENNKASIYPVVRATS 96
                *                ****                * * * * *

B. subtilis      FPDKSSEDIPLIYDDHTAETRIYYALDLGKTYRLIDQRMLEKENWTKERI 148
G. stearotherm.  FPTETKEGVPLLFDDHTAETRIYYALDLGKTYRLIDERMLEKDKWSRERV 146
S. aureus       FDKKTKQGVPIYDEHTAETA VYYAVDLGKSYRLIDESMLEDLKLTEQQI 145
L. monocytogenes FPDETKAGEALLTDDHTAETKIFYAVDLGKSYRFIEESMLKKAQLTHKEI 146
                *                * * * * *      * * * * *      * *

B. subtilis      RETAAFNLRLSLPTVVKEDTVAGNYFYFFRANRNDGYDASRILNEAILNEYKQ 198
G. stearotherm.  KEIARFNVRSLPTPVKEARVADNVFYFVNPNDGYDASRILNESFLADMRA 196
S. aureus       REMSLFNVRKLSNSYTTDEVKGNIFYFINSNDGYDASRILNTAFLNEIEA 195
L. monocytogenes REVAFNNLANLEIPLKKDSVNGNDFYFVRTNDGYDASRLLNEAFLREMRE 196
                *      *      *                *      * * * *      * * * * * * * * *

B. subtilis      HAEGELAISVPHQDVLILADIRNESGYDILGQMSMSFFAGGTVPITALSF 248
G. stearotherm.  RVEGTMAVAVPHQDVLVIADVRNDIGYDVLAQMTMSFFAGGRVPITALSF 246
S. aureus       QCQGEMLVAVPHQDVLI IADIRNKTGYDVM AHLTMEFFTKGLVPITSLSF 245
L. monocytogenes KLTGEMVLAVPHQDVLIIGDIQDNTGYDVLAHMTMDFADGLVPITSLPF 246
                *                * * * * *      *      * * *      * * * * * * * *

B. subtilis      LYNEGKLEPVFILAKSRPKD----- 269
G. stearotherm.  LYENGKLEPIFILGKRRRT----- 265
S. aureus       GYKQGHLEPIFILGKNNKQRDPNVIQRLEANRRKFNKDK 285
L. monocytogenes VYNNGKLEPIFIMAKNRLKE----- 266
                *      * * * * *      *

```

Figure S5.2. Multiple sequence alignment of YtpQ. Shown are representative sequences from *Bacillus subtilis*, *Geobacillus stearothermophilus*, *Staphylococcus aureus*, and *Listeria monocytogenes*.

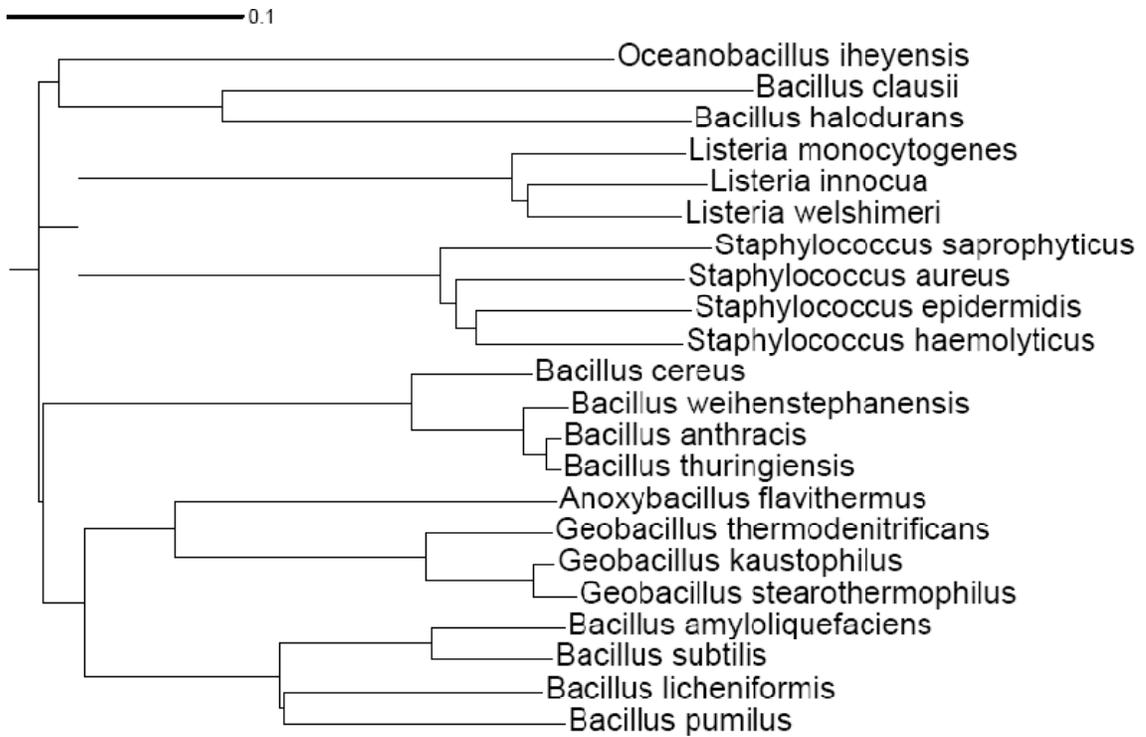


Figure S5.3. ClustalW phylogenetic tree showing the relationship of YtpQ gene family among heme-synthesizing Firmicutes.

Table 5.1. Protein families currently annotated as “coproporphyrinogen III oxidase, oxygen independent” in public databases. These families are grouped based on sequence comparison and genomic context. *Percent amino acid sequence identity to *E. coli* HemN. As a reference, distant bona fide HemN sequences exhibit 50% identity.

Group	Representative genes	Sequence identity*	Putative Function
A	<i>E. coli hemN</i> , <i>S. typhimurium hemN</i>		Bona fide oxygen-independent coproporphyrinogen dehydrogenase
B	<i>V. cholera hutW</i> , <i>E. coli hutW</i>	25%	Found exclusively with heme uptake gene clusters
C	<i>A. baylyi</i> ADP1 ACIAD0432, <i>C. aurantiacus</i> Caur_0209	23%	Associated with the genes of nucleoside metabolism RdgB and ribonuclease PH
D	<i>B. subtilis hemN</i> , <i>hemZ</i> <i>M. tuberculosis hemN</i>	23%	Clustered with stress response genes and present in organism lacking heme biosynthesis

Table 5.2. Candidate gene families for novel coproporphyrinogen oxidase enzymes based on phyletic and expression profiles.

	Representative genes	Description
Candidate family #1	<i>B. subtilis hemQ</i> , <i>M. tuberculosis hemQ</i>	Found in the Firmicutes and Actinobacteria, essential for heme biosynthesis
Candidate family #2	<i>B. subtilis ytpQ</i> , <i>S. aureus SAV1743</i>	Present only in the Firmicutes and a member of the DUF1444 superfamily.
Candidate family #3	<i>B. subtilis ygaE</i> , <i>S. aureus SAV1865</i>	Present only in the Firmicutes and annotated as “putative lipoprotein.”

Table 5.3. ICAP spectrometric metal analysis of YtpQ. YtpQ was preloaded with Co^{2+} and dialyzed to remove unbound metal.

Metal	Content (mol metal/mol enzyme)
Cobalt	0.8026
Copper	0.01
Iron	<0.01
Magnesium	0.11
Manganese	<0.01
Nickel	0.02
Zinc	0.01

Table 5.4. Oxygen requirements for YtpQ coproporphyrinogen oxidase activity. Oxygen was reduced by sparging reaction mixtures with nitrogen prior to substrate addition as described in the text.

Enzyme	Reaction condition	Specific Activity (min ⁻¹)
YtpQ	Aerobic	9.75 ± 0.44
	Limited oxygen	1.05 ± 0.16

Table S5.1. List of potential biological cofactors tested for stimulation of *B. subtilis* YtpQ activity. 100 nM *B. subtilis* YtpQ was assayed for activity in the presence of 10 μ M cofactor.

Cofactors	Metals
ATP	CaCl ₂
FAD	CoCl ₂
FMN	CuCl ₂
NAD	FeSO ₄
NADH	FeCl ₂
NADP	MgSO ₄
NADPH	MnCl ₂
PLP	NiCl ₂
PMP	ZnSO ₄
TPP	

Table S5.2. Bacterial Strains and plasmids

Strain or Plasmid	Description
Strains	
<i>E. coli</i>	
JM109	Used for cloning of plasmids and protein expression
<i>A. baylyi</i>	
ADP1	Wild-type strain
ACN1156	$\Delta hemF$, obtained by gene replacement, Kan ^R
Plasmids	
pUC19	Backbone vector, Amp ^R
pRMJ1	Plasmid harboring <i>sacB</i> -Kan ^R cassette (48)
pUGA1	500 bp ACIAD0432 flanking regions cloned into pUC19
pUGA2	<i>sacB</i> -Kan ^R cassette from pRMJ1 inserted into pUGA1
pUGA3	<i>E. coli hemN</i> inserted into pUGA1
pUGA4	<i>Chloroflexus aurantiacus hemN</i> inserted into pUGA1
pUGA5	<i>Bacillus subtilis hemN</i> inserted into pUGA1
pUGA6	<i>Bacillus subtilis hemZ</i> inserted into pUGA1
pUGA7	<i>Mycobacterium tuberculosis hemN</i> inserted into pUGA1
pUGA8	<i>Bacillus subtilis hemQ</i> inserted into pUGA1
pUGA9	<i>Mycobacterium tuberculosis hemQ</i> inserted into pUGA1
pUGA11	<i>Bacillus subtilis ytpQ</i> inserted into pUGA1
PUGA12	<i>Bacillus subtilis ygaE</i> inserted into pUGA1
pTrcHisA	<i>E. coli</i> expression vector
pTHMBP	Maltose binding protein inserted into the NheI site of pTrcHisA
pTHMBPYtpQ	<i>Bacillus subtilis ytpQ</i> inserted into pTHMBP

CHAPTER 6

CONCLUSIONS

Heme, a major biologically functional form of iron, is a molecule integral in catalytic and regulatory functions of all cells. While the most basic of the tetrapyrroles, it is also the most diverse, with its iron-chelated porphyrin structure utilized in a wide array of molecular functions, as has been described throughout this text. In eukaryotes, the synthesis of this molecule has long been established, beginning with the ground-breaking work of David Shemin and David Rittenberg in 1945 (14) that ignited the further research that would follow. Synthesis of the precursor compound δ -aminolevulinic acid (ALA) was eventually seen to occur differently in plants than in animals, but the remaining seven steps are catalyzed by conserved enzymes. An important fact to remember about heme biosynthesis is the toxicity of the intermediate products, as this highlights the principle that these intermediates should never exist freely within a cell. Thus, evolutionary conservation of the enzymes makes sense in this regard.

In prokaryotes, however, the situation is much more complex. With the advent of genomic sequencing, bioinformatic analysis has revealed that homologues of certain enzymes within the pathway could not be located in bacteria, and in some cases, such as *Desulfovibrio* species and all of the archaea, the entire terminal portion of the pathway appeared to be missing. Among those missing enzymes are the antepenultimate enzyme

coproporphyrinogen oxidase (CPO) and protoporphyrinogen oxidase (PPO). The organisms lacking these enzymes contain the remaining enzymes including the terminal ferrochelatase (FC), therefore it was clear that they must also contain novel, unidentified forms of CPO and PPO to carry out their respective chemical conversions. When this research began, bona fide protoporphyrinogen oxidase could only be asserted in a subset of bacteria, mainly in the Firmicutes and Actinobacteria (2). Only a handful of other species such as *Myxococcus xanthus* also harbor HemY. This is especially notable as the absence of this enzyme is so taxonomically widespread, despite the presence of all other biosynthetic enzymes within these same organisms. As noted by Panek et al. (12), “This (lack of identifiable PPO) suggests a major gap in our understanding of haem biosynthetic pathways, even in organisms that are well studied such as *Pseudomonas aeruginosa*, *Caulobacter crescentus*, and the rhizobia.” The specific aim of this dissertation, then, was to identify potential variant enzymes that must exist, thereby solving one of the major problems facing prokaryotic heme biosynthesis, as well as characterize the overall diversity.

In our first main goal, we set out to ascertain the function of HemG, a protein in *E. coli* and other enterics, previously implicated in PPO activity by way of an *E. coli* mutation in the respective gene’s nucleotide sequence discovered in 1979 (13). This mutant exhibited an auxotrophy for heme, and subsequent use of the strain in identifying eukaryotic PPO enzymes revealed the auxotrophy could be reversed by complementation with functional PPO (10). While this proved that removal of HemG also removed PPO activity, no other information on

HemG was ever discovered. The amino acid sequence of HemG puts it in a category of proteins known as flavodoxins, non-catalytic electron carriers that shuttle electrons between two systems. During our initial attempts at characterizing this protein as a flavodoxin, we determined that it did in fact contain many of the expected traits. HemG bound FMN as a cofactor and maintained a redox potential consistent with other known flavodoxins. We then discovered that, despite it resembling non-catalytic flavodoxins, HemG was in fact capable of converting protoporphyrinogen ix into protoporphyrin IX. This activity was dependent on menadione, a soluble form of the quinone. The dependence on the menaquinone pool *in vivo* explained all the early work by the Jacobs group (4-8) showing that PPO activity in *E. coli* crude cell extracts was linked to electron transport, both aerobic and anaerobic, since enzymes such nitrate reductase, fumarate reductase, and the terminal oxidases of *E. coli* turn this pool over for HemG to reuse. Astoundingly, this activity can be transferred to a non catalytic flavodoxin by replacing the long chain insert loop, roughly 20 amino acids, with that of HemG. This region clearly imparts activity, though our *in vivo* results show that menadione utilization is not transferred. With this discovery, PPO could be fully asserted in the enteric subdivision of the γ -proteobacteria.

Unfortunately, HemG only covered a small percentage of those missing HemY. Roughly half the bacteria with sequenced genomes were still left missing an identifiable enzyme, and filling this void became our second main objective. Here, though, there was no previous knowledge of PPO activity as with *E. coli*.

To address the problem, a different approach was needed. Standard homology-driven methods at finding these enzymes would be fruitless, as it was clear homologues of HemY or HemG were not present. Instead, we chose to utilize the SEED comparative genomics database, a system designed to work with no working knowledge of unidentified enzymes except function (11). The SEED utilizes a database of subsystems, defined in this regard as biosynthetic pathways, and populates a list of these across all sequence genomes available. With it, the genomes containing or lacking specific steps are easily highlighted. Using a carefully curated heme biosynthetic subsystem, we were able to identify a novel candidate for PPO, now named HemJ, in the remaining organisms based on its occurrence profile and genomic context. We subsequently showed the gene encoding HemJ to be essential for PPO activity within *Acinetobacter baylyi* ADP1. This enzyme satisfied the remaining heme-synthesizing bacteria for presence of a functional PPO, completing our objective and removing perhaps the biggest hurdle left in our knowledge of prokaryotic heme biosynthesis.

Next, we turned our attention to coproporphyrinogen oxidase, the enzyme catalyzing the step prior to that of PPO. Two forms exist, the oxygen dependent HemF and the oxygen independent HemN (1, 9). Annotation of HemN within prokaryotes had been in question prior to our use of the SEED, but examination within the Heme subsystem revealed that these misannotation levels were much more widespread than previously thought. Genomic analysis revealed that function CPO was missing mainly in the Actinobacteria and Firmicutes. Through methods similar to those used for HemJ, we were able to find several novel

candidates for this function in the Firmicutes, one of which was shown to contain oxygen-dependent CPO activity *in vivo* and *in vitro* and also required a mononuclear cobalt cofactor. Our findings on *B. subtilis* YtpQ are first descriptions of this enzyme in any regard. The requirement of both cobalt and oxygen presents an intriguing mechanism that will require further investigation. YtpQ homologues cover the entire heme-synthesizing Firmicutes, leaving only the Actinobacteria and Acidobacteria with a missing enzyme for this step.

While these novel enzymes were being discovered and characterized, we also learned valuable information on the function of HemY, the PPO resembling that of eukaryotes, and the final enzyme, ferrochelatase. In the Firmicutes and Actinobacteria, these enzymes are soluble rather than membrane-associated like their other homologues. The discovery that they could not catalyze their respective reactions *in vivo* alone sparked our interest, as they clearly function differently than their counterparts. We found that these enzymes could only properly catalyze their respective reactions in the presence of each other and a previously unidentified third enzyme, HemQ. HemQ binds heme, and it was shown that the presence of this heme group is necessary for full activity. Thus, it is plausible that in this system, heme is needed to make heme. While we were unable to correlate a multi-protein complex with our data, our findings that heme can only be made by providing iron and the substrate of HemY suggest that a complex is formed. Otherwise the substrate of HemH (ferrochelatase) would be sufficient.

In his work *Principles of Science: A Treatise on Logic and Scientific Method*, William Stanley Jevons wrote, "Science arises from the discovery of Identity amid Diversity." This goal of this dissertation can be drawn from this, as here we have discerned novel heme biosynthetic enzymes where we knew diversity among them existed. As stated, a large gap in our scientific understanding of heme biosynthesis was missing in a broad range of prokaryotes prior to this work. Now, these "holes" in the pathway have been replaced with newly characterized, unique enzymes and opened up a entirely new avenue of research to pursue. With the discovery of YtpQ, we have found even more diversity present among these terminal steps, with even more yet to be discovered. In addition to our findings, other groups have now shown that an entire novel terminal pathway exists in organisms such as *Desulfovibrio* species and the heme-synthesizing Archaea, who first convert uroporphyrinogen-III into precorrin and are suspected to use a pathway of distinctive *nir*-like genes to produce heme. It is exciting to be a part of the current field in heme biosynthesis, as much can yet be learned concerning the enzymes outlined above. With the discovery of HemQ, an essential enzyme outside the actual catalytic pathway, the possibility now exists that more cases such of this are out there, waiting to be found.

This research has yielded several avenues for future work. In regards to HemG, it will be interesting to obtain a crystal structure. We were able to obtain diffraction data from HemG crystals at 2 Å, but were not able to phase this data despite attempts at crystal soaking to allow isomorphous replacement. Once a

structure is obtained, it will be intriguing to investigate the function of the long chain insert loop we have shown is responsible for activity. For HemJ, we must overcome the hurdle of protein expression to characterize the recombinant protein *in vitro*. Interestingly, this enzyme resembles a four helix bundle, which are known to be capable of binding heme (3). HemJ may bind protoporphyrinogen in a similar manner and allow catalysis, but we must first obtain purified protein before this can be determined. Furthermore, cyanobacterial HemJs contain a C-terminal extension whose function is unknown. Finally, the reliance on cobalt of YtpQ will need to be examined through extensive electrochemical studies, in addition to crystallographic structure determination.

Perhaps the most valuable tool extending from this work is the SEED heme subsystem. Once created, this tool enabled us to rapidly and accurately evaluate heme synthesis in all sequenced genomes, an application that has been invaluable. With it, we are continuing to examine other proteins that will potentially provide even more insight into this metabolic system in the various areas that remain unfilled, such as coproporphyrinogen oxidase in the Actinobacteria. Another potential application is the discovery of a missing, novel uroporphyrinogen III synthase in the α -proteobacteria, which has remained unresearched. A further application of our work is the use of the described enzymes as drug targets, as has been briefly described in the proceeding chapters. When viewed together, the vast majority of microbial pathogens contain either HemG, HemJ, or YtpQ.

During the course of this work, it has also become clear that heme biosynthetic enzyme nomenclature must also soon be addressed. It has been stated earlier that HemG and HemJ cannot be described accurately as protoporphyrinogen oxidases, despite this term being applied to the catalytic step in general. The same can be said of YtpQ, whose mechanism cannot currently be confirmed as an actual oxidase. In addition to these problems, the standard nomenclature of designating these enzymes as Hem followed by a designated letter has reached its inevitable limit. Currently, this system cannot be used to designate YtpQ as the remaining letters have been attributed to various enzymes, in some cases erroneously. As a consequence, a new system must be agreed upon to encase current and future heme biosynthetic enzymes and avoid confusion.

References

1. **Breckau D, Mahlitz E, Sauerwald A, Layer G, & Jahn D.** 2003. Oxygen-dependent coproporphyrinogen III oxidase (HemF) from *Escherichia coli* is stimulated by manganese. *J Biol Chem.* 278:46625-46631.
2. **Dailey TA, et al.** 2010. Discovery and Characterization of HemQ: an essential heme biosynthetic pathway component. *J Biol Chem.* 285:25978-25986.
3. **Gibney BR, Rabanal F, Reddy KS, & Dutton PL.** 1998. Effect of four helix bundle topology on heme binding and redox properties. *Biochemistry.* 37:4635-4643.
4. **Jacobs JM & Jacobs NJ.** 1977. The late steps of anaerobic heme biosynthesis in *E. coli*: role for quinones in protoporphyrinogen oxidation. *Biochem Biophys Res Commun.* 78:429-433.
5. **Jacobs NJ & Jacobs JM.** 1975. Fumarate as alternate electron acceptor for the late steps of anaerobic heme synthesis in *Escherichia coli*. *Biochem Biophys Res Commun.* 65:435-441.
6. **Jacobs NJ & Jacobs JM.** 1976. Nitrate, fumarate, and oxygen as electron acceptors for a late step in microbial heme synthesis. *Biochim Biophys Acta.* 449:1-9.
7. **Jacobs NJ & Jacobs JM.** 1977. Evidence for involvement of the electron transport system at a late step of anaerobic microbial heme synthesis. *Biochim Biophys Acta.* 459:141-144.

8. **Jacobs NJ & Jacobs JM.** 1978. Quinones as hydrogen carriers for a late step in anaerobic heme biosynthesis in *Escherichia coli*. *Biochim Biophys Acta.* 544:540-546.
9. **Layer G, Verfurth K, Mahlitz E, & Jahn D.** 2002. Oxygen-independent coproporphyrinogen-III oxidase HemN from *Escherichia coli*. *J Biol Chem.* 277:34136-34142.
10. **Nishimura K, Taketani S, & Inokuchi H.** 1995. Cloning of a human cDNA for protoporphyrinogen oxidase by complementation in vivo of a hemG mutant of *Escherichia coli*. *J Biol Chem.* 270:8076-8080.
11. **Overbeek R, et al.** 2005. The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. *Nucleic Acids Res.* 33:5691-5702.
12. **Panek H & O'Brian MR.** 2002. A whole genome view of prokaryotic haem biosynthesis. *Microbiology.* 148:2273-2282.
13. **Sasarman A, et al.** 1979. Mapping of a new hem gene in *Escherichia coli* K12. *J Gen Microbiol.* 113:297-303.
14. **Shemin D & Rittenberg D.** 1945. The utilization of glycine for the synthesis of a porphyrin. *J Biol Chem.* 159:567-568.