Novel Enzymes and Catalytic Mechanisms in Tetrapyrrole Metabolism: a Nitrogenase Homolog and a Noncanonical Heme Oxygenase

by

Shuxin Li

A dissertation submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Auburn, Alabama
August 7, 2021

Keyword: heme oxygenase, *c*-type heme, coenzyme F430, nitrogenase, *in vivo* biosynthesis, CfbCD

Approved by

Steven O. Mansoorabadi, Chair, Associate Professor of Chemistry and Biochemistry
Eduardus C. Duin, Professor of Chemistry and Biochemistry
Douglas C. Goodwin, Professor of Chemistry and Biochemistry
Rashad Karimov, Assistant Professor of Chemistry and Biochemistry
Yi Wang, Associate Professor of Biosystems Engineering

Abstract

Methane is the major component of natural gas, which accounts for approximately one quarter of the world's energy demands. There are many advantages to using methane as a fuel. Methane is an abundance resource. A huge number of methane hydrates are on the seafloor and microorganism break up organic matter to produce methane. Methane is the simplest organic compound and produces more heat per unit mass (55.7 kJ/g) than any other hydrocarbon. Methane is also a clean fuel. It releases less sulfur, carbon, and nitrogen than coal or oil, and leaves little ash. However, methane formations are usually distributed in remote areas, which require an extensive and complex pipeline network to transport natural gas to end users. Natural gas is highly flammable, so the transportation from wellheads to homes and businesses is dangerous. Meanwhile, methane is a greenhouse gas that has thirty times the global warming potential of carbon dioxide. Thus, there is great current interest in strategies to convert methane to liquid fuel or other more easily transported commodity chemicals.

Methanogens are anaerobic archaea widely found in wetlands and the digestive tracts of animals, which produce around 1 billion tons of methane. Methanogenesis is balanced by aerobic or anaerobic methane oxidation. Anaerobic methanotrophic archaea (ANME) catalyze the anaerobic oxidation of methane (AOM), often with the aid of sulfate or nitrate reducing bacteria. Both methanogenesis and AOM use a key enzyme, methyl-coenzyme M reductase (MCR), which has a unique nickel-containing coenzyme F430 for activity. MCR catalyzes the interconversion of coenzyme B and methyl-coenzyme M to the mixed heterodisulfide CoB-S-S-CoM and methane. ANME have potential for use in natural gas-to-liquid fuel conversion strategies, but they have extremely slow growth rates, and they cannot be currently obtained in

pure culture. A possible solution is to engineer the methane-conversion pathway and its key enzyme, MCR, into a more robust microorganism. Currently, MCR cannot be produced in an active form in a heterologous host, in part because the *in vivo* biosynthesis of coenzyme F430 is still a challenge.

Our lab has identified the coenzyme F430 biosynthesis (*cfb*) genes and characterized the encoded enzymes *in vitro*. They include a sirohydrochlorin cobaltochelatase homolog, CfbA, which catalyzes the specific Ni-chelation of sirohydrochlorin; a cobyrinic acid *a,c*-diamide synthase homolog, CfbB, that catalyzes the adenosine triphosphate (ATP)-dependent amidation of the *a*- and *c*-acetate side chains of Ni-sirohydrochlorin using glutamine as an ammonia source; a nitrogenase homolog, CfbCD, which converts Ni-sirohydrochlorin *a,c*-diamide to 15,17^3-seco-F430-17^3-acid; and a Mur-ligase homolog, CfbE, that carries out the ATP-dependent cyclization of the *g*-proprionate side chain to generate the carbocyclic F ring.

Chapter Two details the construction of compatible expression vectors containing the *cfb* genes and putative accessory factors from *Methanosarcina acetivorans* and experiments designed to produce coenzyme F430 in *Escherichia coli*. Key proteins in the coenzyme F430 biosynthesis pathway were extracted from the engineered cell line and tested for activity *in vitro*. Different physiological reducing systems, including ferredoxin (Fd)/ferredoxin:NADP$^+$ reductase (FNR) and pyruvate:ferredoxin oxidoreductase (PFOR), were tested for their ability to support CfbCD catalysis *in vivo.* The effects of supplementation with an additional copy of the *cfbCD* genes from *Methanosarcina thermophila* (which encode a more stable/active variant of the CfbCD complex) and coexpression with 8 putative MCR maturation genes on the heterologous production of coenzyme F430 were also examined, as were various concentrations of supplementary chemicals and growth times/conditions.

Coenzyme F430 is the most highly reduced tetrapyrrole in nature, and this is achieved by a single multi-electron redox reaction catalyzed by the nitrogenase homolog, CfbCD. The CfbCD-catalyzed conversion of Ni-sirohydrochlorin *a,c*-diamide to 15,17^3-seco-F430-17^3-acid involves a 6-electron reduction of the isobacteriochlorin ring system, cyclization of the *c*-acetamide side chain to form the γ-lactam E ring, and the formation of 7 stereocenters. The chemical mechanism of this unprecedented transformation is unknown, and there is debate as to whether γ-lactam ring formation is catalyzed by CfbCD or occurs non-enzymatically. In Chapter Three, mechanistic studies of the CfbCD reaction is presented. The CfbCD reaction was performed in deuterium oxide to label 15,17^3-seco-F430-17^3-acid and determine the stereochemistry of the reaction using nuclear magnetic resonance (NMR) spectroscopy. A ^{15}N-labelled version of the immediate product of the CfbCD reaction was prepared and purified for structural characterization using NMR. Meanwhile, a novel tetrapyrrole, Ni-didecarboxysirohydrochlorin *a,c*-diamide, was synthesized for use as a mechanistic probe by combining the CfbA and CfbB reactions with an enzyme from the heme d_1 biosynthetic pathway.

Chapter Four describes a noncanonical heme oxygenase specific for the degradation of *c*-type heme. Heme oxygenases (HOs) play a critical role in recouping iron from the labile heme pool. The acquisition and liberation of heme iron is especially important for the survival of pathogenic bacteria. All characterized HOs, including those belonging to the HugZ superfamily, preferentially cleave free *b*-type heme. Another common form of heme found in nature is *c*-type heme, which is covalently linked to proteinaceous cysteine residues. However, mechanisms for direct iron acquisition from the *c*-type heme pool are unknown. We identified a HugZ homolog from the oligopeptide permease (*opp*) gene cluster of *Paracoccus denitrificans* that lacks any observable reactivity with heme *b* and show that it instead rapidly degrades *c*-type hemopeptides.

This *c*-type heme oxygenase catalyzes the oxidative cleavage of the model substrate microperoxidase-11 at the β- and/or δ-meso position(s), yielding the corresponding peptide-linked biliverdin, CO, and free iron. X-ray crystallographic analysis suggests that the switch in substrate specificity from *b*- to *c*-type heme involves loss of the N-terminal α/β domain and C-terminal loop containing the coordinating histidine residue characteristic of HugZ homologs, thereby accommodating a larger substrate that provides its own iron ligand. These structural features are also absent in certain heme utilization/storage proteins from human pathogens that exhibit low or no HO activity with free heme. This study thus expands the scope of known iron acquisition strategies to include direct oxidative cleavage of heme-containing proteolytic fragments of *c*-type cytochromes and helps to explain why certain oligopeptide permeases show specificity for the import of heme in addition to peptides.

Acknowledgments

Five years ago, when I first came to the Department of Chemistry and Biochemistry, I was full of uncertainties and worries. I was worried that professors would not accept me because of my poor English; I was worried that I would not be able to pass all my required courses; I was worried that my weak background in biochemistry would hinder me on the road to scientific research. Today, I am full of confidence and gratitude working on my dissertation. I would like to dedicate this big change to my mentor and professor, Dr. Steven Mansoorabadi. He is a genius who constantly conceives and explores new and novel ideas. So, I have never been at a loss in the lab. He is always patient and can explain the basic things to me over and over again. My English is not good, and he always encourages me to communicate with him and other lab members. I have been in the United States for eight years, but my English has greatly improved only in his lab. He has a very good personality and has never shown any bad emotion to us. He is humorous, approachable, and handsome. I do not know if I will be lucky enough to meet such a good mentor in the future.

Besides my advisor, I would like to thank Dr. Eduardus Duin, who shared his instruments with me and provided solutions to my problems. I would like to thank Dr. Douglas Goodwin, who shared his knowledge of enzyme kinetics with me. Also, he accessed me at my point of entry to the program and processed my transfer within a short time. I would like to thank Dr. Rashad Karimov, who taught me organic chemistry, and I really enjoyed the class.

I would like to extend my gratitude to my previous lab colleagues Victoria, Kaiyuan, Trey, Xingchen, Selam, Eta. I learned many techniques from them, and they were all patient to answer my questions when I was new in the lab. I also thank my current colleagues Kenny,

Chidinma, Patrick, Prosenjit, Chelsea, and Emily. They all provided an enjoyable lab environment and brought in diversity of cultures to the lab.

I would like to express the deepest appreciation to my mother, my father, and my husband. Without their support, I would not have had the opportunity to look at the world and finish my PhD studies.

Lastly, I would like to thank the Department of Chemistry and Biochemistry at Auburn University. My research career started here and will continue.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

5-ALA	5-aminolevulinic acid
PLP	Pyridoxal 5′-phosphate
PBG	Porphobilinogen
HMB	Hydroxymethylbilane
MFR	Methanofuran
H_4MPT	Tetrahydromethanopterin
CoM	Coenzyme M
CoA	Coenzyme A
CoB	Coenzyme B
CODH	Carbon monoxide dehydrogenase
FBeB	Flavin-based electron bifurcation
CDeT	Cytochrome-dependent electron transfer
MP	Methanophenazine
RuMP	Ribulose monophosphate
ANME	Anaerobic methanotrophic archaea
AOM	Anaerobic oxidation of methane
SRB	Sulfate-reducing bacteria
MCR	Methyl-coenzyme M reductase
EPR	Electronic paramagnetic resonance
BPS	3-Bromopropanesulfonate
RFQ	Rapid freeze-quench

MCD	Magnetic circular dichroism
PTM	Post-translational modification
SAM	*S*-adenosylmethionine
Mm10	Methanogenesis marker protein 10
ATP	Adenosine triphosphate
NMR	Nuclear magnetic resonance
cfb	Coenzyme F430 biosynthesis
UMA	UDP-*N*-acetylmuramoyl-L-alanine
ROS	Reactive oxygen species
OM	Outer membrane
CM	Cytosolic membrane
ABC	ATP-binding cassette
PBP	Periplasmic binding proteins
HO	Heme oxygenase
ABM	Antibiotic biosynthesis monooxygenase
NCBI	National Center for Biotechnology Information
PMP	Pyridoxamine 5′-phosphate
FMN	Flavin mononucleotide
FAD	Flavin adenine dinucleotide
F_o	8-hydroxy-5-deazaflavin
F_{420}	deazaflavin
CHO	*c*-type heme oxygenase
MP-11	Microperoxidase-11

LB	Luria-Bertani
HPLC	High-performance liquid chromatography
DAD	Diode array detector
MS	Mass spectrometry
ESI	Electrospray ionization
SDS-PAGE	Sodium dodecyl-sulfate polyacrylamide gel electrophoresis
PNP	Pyridoxine 5′-phosphate
Fd	Ferredoxin
FNR	Ferredoxin:$NADP^+$ reductase
MCS	Multiple cloning site
Isc	Iron-sulfur cluster
PFOR	Pyruvate:ferredoxin oxidoreductase
IPTG	Isopropyl β-D-1-thiogalactopyranoside
NAD^+	Nicotinamide adenine dinucleotide
NADPH	Nicotinamide adenine dinucleotide phosphate
PCR	Polymerase chain reaction
MALDI-TOF	Matrix-assisted laser desorption/ionization time-off-light
COR	Chlorophyllide oxidoreductases
NOESY	Nuclear Overhauser effect spectroscopy
HSQC	Heteronuclear single quantum coherence spectroscopy
FPLC	Fast protein liquid chromatography
DPOR	Dark-operative protochlorophyllide oxidoreductase

Chapter One: Literature Review

1.1 Tetrapyrroles

Tetrapyrroles are the most abundant pigments in nature, which provide widely diverse functions in living organisms. Many well-known molecules are tetrapyrroles.[1] Hemes are the oxygen carrier in blood and function in electron transport.[2] Chlorophyll converts light energy to chemical energy in plants during photosynthesis.[3] Cobalamin is an essential nutrient involved in fatty and amino acid metabolism in humans.[4] Siroheme is a cofactor in the assimilatory six electron reduction of both nitrite to ammonia and sulfite to sulfide.[5] Other tetrapyrroles include coenzyme F430, and heme d_1. Tetrapyrroles contain four pyrrole-type rings that are typically linked in linear or cyclic arrangement by single carbon bridges.[1] The only exceptions are corrinoids, which lack a carbon bridge between the first and the fourth pyrrole ring.[1] In general, the pyrrole rings of all cyclic tetrapyrroles are named A to D in a clockwise direction. Most cyclic tetrapyrroles coordinate a metal ion such as Fe, Mg, Co, or Ni, which can be cleaved to release the metal and generate a linear tetrapyrrole.[1]

The tetrapyrrole biosynthesis pathway is conserved in both prokaryotes and eukaryotes (Fig 1.1).[6] Uroporphyrinogen III is the last common intermediate in the biosynthesis of all tetrapyrroles (Fig 1.1). To synthesize uroporphyrinogen III, 5-aminolevulinic acid (5-ALA) is the first common precursor that can be formed via two different pathways (Fig 1.1).[6, 7] Glycine is converted to 5-ALA through a condensation reaction with succinyl-coenzyme A in the members of the α-proteobacteria and achloroplastic eukaryotes. This reaction is catalyzed by the pyridoxal 5′-phosphate (PLP)-containing enzyme 5-ALA synthase (HemT).[8] Glutamate forms 5-ALA by ligation to $tRNA^{Glu}$, reduction, and transamination in plants, algae, and bacteria that are not in the α-proteobacterial class.[8]

After 5-ALA formation, a series of catalytic reactions occur to form the first macrocyclic intermediate in the biosynthesis of tetrapyrroles, uroporphyrinogen III (Fig 1.1).[7] Two molecules of 5-ALA are combined through an asymmetric condensation reaction by porphobilinogen synthase (HemB) to form porphobilinogen (PBG).[9] *E. coli* HemB is a magnesium activated zinc metalloenzyme. The zinc ion is coordinated by three cysteines within the active site of the enzyme and is thought to help bind and activate the second ALA molecule.[10] Then, porphobilinogen deaminase (HemC) catalyzes the condensation of four porphobilinogen molecules into the linear hydroxymethylbilane and releases ammonia. HemC contains a unique cofactor, dipyrromethane, which consist of two PBG molecules linked together and covalently bound to the enzyme by Cys-242.[11] Finally, the A and D rings of hydroxymethylbilane (HMB) are linked together by uroporphyrinogen III synthase (HemD), which results in the formation of uroporphyrinogen III.[12] Free HMB can form the nonphysiological product uroporphyrinogen I because of spontaneously cyclization. This implies that HemC and HemD must act in a highly coordinated fashion to achieve uroporphyrinogen III formation.[13] In the biosynthesis of hemes and chlorophylls, uroporphyrinogen III is converted into coproporphyrinogen III (Fig 1.1). This reaction is catalyzed by uroporphyrinogen III decarboxylase (HemE), which eliminates carboxyl groups from the four acetate side chains of uroporphyrinogen III.[6] In the biosynthesis of heme d_1, cobalamin and coenzyme F430, uroporphyrinogen III is converted into precorrin 2 (Fig 1.1). This reaction is catalyzed by uroporphyrinogen III *C*-methyltransferase (SirA), which catalyzes two methylation reactions.[6] The most important tetrapyrroles involved in this study are heme and coenzyme F430. F430 is the key coenzyme of methanogenesis, the biological pathway of methane formation.

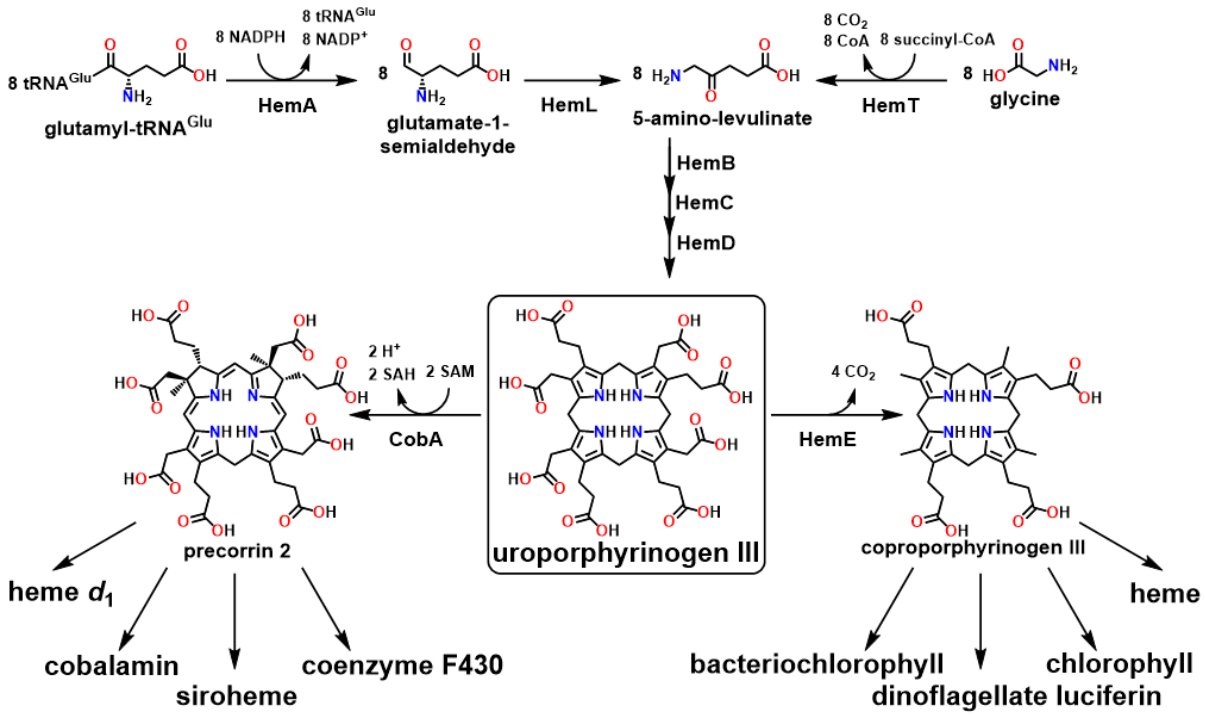


Figure 1.1 Tetrapyrrole biosynthesis pathway from reference 6.

1.2 Methane and its significance

Methane is the major component of natural gas, which accounts for approximately one quarter of the world's energy demands.[14] According to the *International Energy Outlook 2019* (IEO2019) reference case, consumption of natural gas is predicted to increase more than 40% between 2018 and 2050.[15] Most natural gas is used for heating and electricity generation. It also can be used as a fuel for cars, organic chemical production, and fertilizer. Methane is the simplest organic compound and produces more heat per mass unit than any other hydrocarbons.[15]

Compared with other energy sources, natural gas is an abundant resource.[16] Natural gas can be obtained from associated gas, which is a by-product of oil production, or non-associated gas, which are shale gas, sour gas, tight gas, and coalbed methane.[17] A large amount of natural gas is in the form of hydrates under the ocean. Methane hydrates are methane molecules trapped in water cages. They require low temperature and high pressure to form.[18] Almost all biological

sources of methane derive from methanogenic archaea. Some of the methanogenic archaea are symbiotic, which live in the digestive tracts of termites and ruminants and produce methane during the breakdown of cellulose.[19] Methanogenic archaea also live in crop fields, marshes, and landfills. They can degrade organic matter to produce methane.[20] Burning natural gas produces carbon dioxide and water, but burning coal and oil mainly produces carbon dioxide. Natural gas releases less carbon dioxide every unit of energy, which is about half the carbon dioxide produced from burning coal and 30% less carbon dioxide than burning petroleum.[21] Also, burning natural gas leaves less nitrogen, sulfur, and ash than coal and oil. Thus, natural gas is a more efficient, abundant, and cleaner fuel.[22]

Natural gas formations are usually distributed in remote areas, which require an extensive and complex pipeline network to transport to end users.[23] Raw natural gas usually has contaminants, which should be removed before delivered to pipelines.[24] Natural gas pipelines are impossible to cover all the locations and across oceans. In that case, natural gas is usually transported in its liquefied form, which requires low temperature and is returned to gas form at the terminal.[25] Natural gas is highly flammable, so transportation from wellheads to homes and businesses is dangerous.[23] In the past 20 years, thousands of serious incidents occurred, most of which were related to gas pipelines. There are three major reasons for pipeline leaks: excavation damage, corrosion, and material flaws.[26]

Methane is the second largest greenhouse gas after carbon dioxide. Although the lifetime of methane is relatively shorter than carbon dioxide, it can trap more heat.[27] Thus, methane has approximately 30 times the global warming potential of carbon dioxide in 100 years, and approximately 80 times in 20 years.[27] This means methane is the strongest carbon-based greenhouse gas in the short term.[27] Fossil fuel mining is responsible for about 40% of all

methane emissions (Fig 1.2). The decomposition of wastewater and solid waste accounts for 18% of methane emission (Fig 1.2).[28]

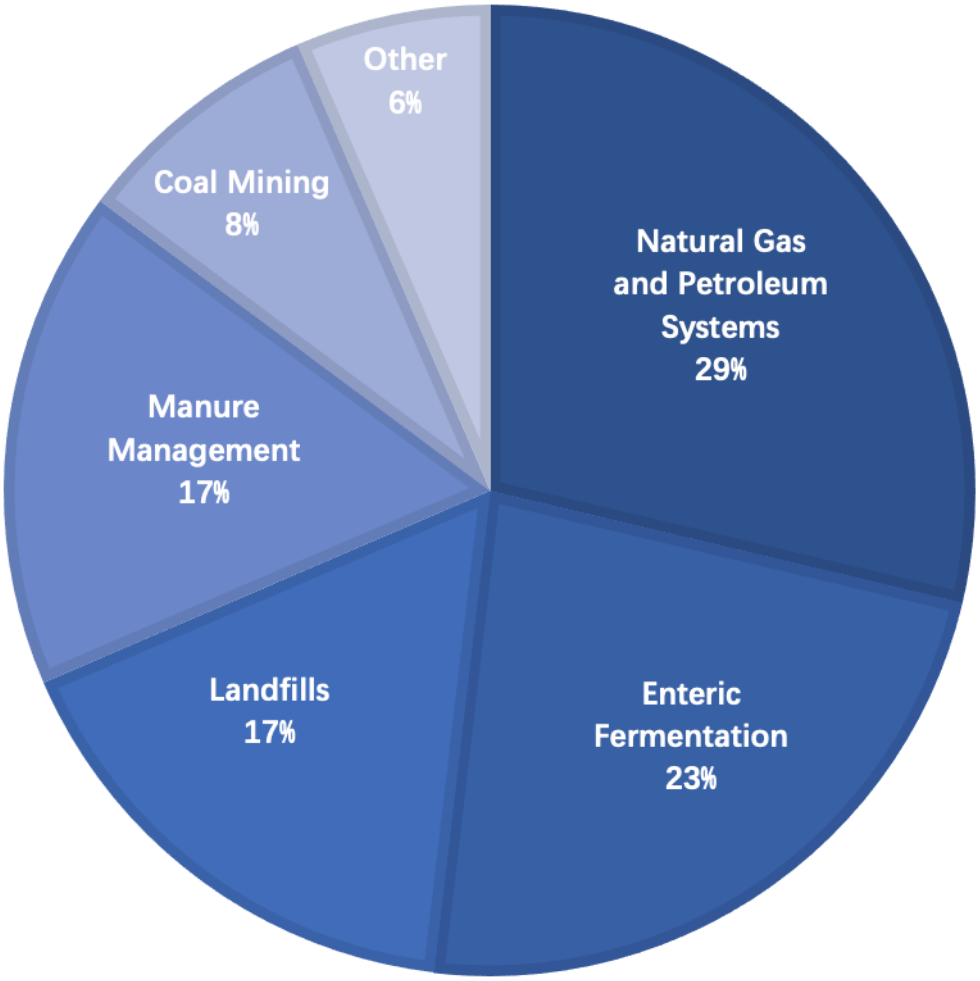


Figure 2.2 Global methane emission from reference 28.

1.3 Methanogenesis and methanogenic Archaea

Methanogens are the microorganisms that produce methane as a byproduct of their anaerobic metabolism. They belong to Archaea, one of the three domains of life along with Bacteria and Eukarya.[29] Methanogens are classified into seven orders, *Methanococcales*, *Methanobacteriales*, *Methanosarcinales*, *Methanopyrales*, *Methanomicrobiales*, *Methanocellales*, and *Methanomassiliicoccales*.[30-35] Hydrogenotrophic methanogens that grow

on H_2, and CO_2 are found in almost all methanogenic orders except *Methanomassiliicoccales*. This type of methanogen is speculated as the ancestor of methanogens due to its broad distribution.[36] Aceticlastic methanogens that grow on acetate are found only in the order Methanosarcinales.[37] Methylotrophic methanogens use different methylated compounds as substrates for methanogenesis. They are found in the orders *Methanomassiliicoccales*, *Methanobacteriales*, and *Methanosarcinales*.[38]

1.3.1 Diversity of methanogens

Methanogens are widely found in wetlands, landfills, marshlands, rice fields, and the digestive tracts of animals.[29] Most methanogens cannot survive under aerobic conditions. Some species can live longer than other methanogens in the presence of oxygen due to a superoxide dismutase.[39] Methanogens can be found in some extreme environments such as hydrothermal vents, saline lakes, and deep oceans, which means they can survive across a wide range of thermochemical gradients.[20] Many methanogens are mesophilic. However, hyperthermophilic methanogens, like *Methanopyrus kandleri* can grow up to 110 °C.[32] In contrast, cold-loving methanogenic strains still have metabolic activity at 0 °C.[40] Some methanogens are halophilic. They can synthesize osmoprotectants, like glutamate to protect themselves from losing water.[20] Another way to balance the cytoplasm osmotically is to accumulate potassium and chloride or take up glycine betaine from its environment.[41, 42] Besides the wide range of temperature and salinity, methanogens also can be isolated from acidophilic or alkaliphilic environments.[20]

1.3.2 Metabolism of methanogens

Methanogenesis is an anaerobic respiration that uses carbon as an electron acceptor to produce methane. Based on the different substrates, methanogenesis can be separated into three types (Fig 1.3).[20] Hydrogenotrophic methanogenesis uses CO_2 as the substrate and H_2 as an

electron donor. In the first step, CO_2 is reduced and combined with methanofuran (MFR) to form formylmethanofuran (CHO-MFR).[43] In the second step, the formyl group is transferred from MFR to tetrahydromethanopterin (H_4MPT), which forms methylene-H_4MPT (HC≡H_4MPT). Two subsequent reduction processes yield methyltetrahydromethanopterin (H_3C-H_4MPT). Then, the methyl group is transferred from H_4MPT onto coenzyme M (HS-CoM). Finally, methyl-coenzyme M (CH_3-S-CoM) is reduced to methane with coenzyme B (HS-CoB) as electron donor.[44, 45] Aceticlastic methanogenesis uses acetate as substrate. Acetate is first converted with ATP and coenzyme A (CoA) to acetyl-CoA, which further converts to CO_2 and a corrinoid protein-bound methyl group that is transferred to H_4MPT by the carbon monoxide dehydrogenase (CODH) and acetyl-CoA synthase (ACS) complex.[46] Methylotrophic methanogenesis uses methylated compounds such as methanol, methanethiol, and methylamines as substrate. The methyl groups are directly transferred to a corrinoid protein and subsequently to HS-CoM to form methyl-CoM by specific methyltransferase complexes.[47]

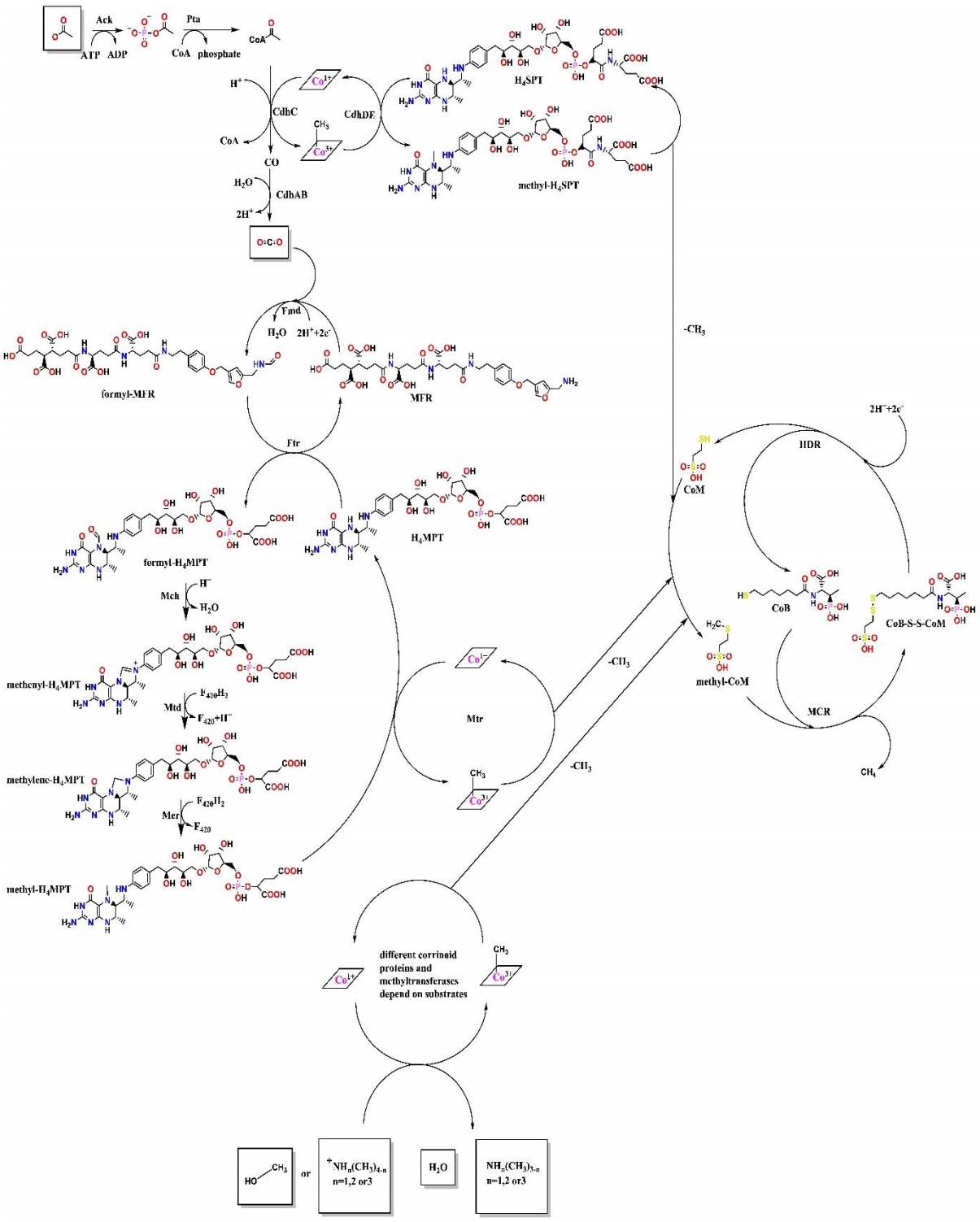


Figure 3.3 Methanogenesis pathway from reference 20.

1.3.3 Energy conservation in methanogens

The major energy conserving step in methanogenesis is the conversion of heterodisulfide back to coenzyme B and coenzyme M.[20] Its mechanism can be divided into two groups depending on whether or not the methanogen contains cytochromes (Fig 1.4).[45, 48] In most hydrogenotrophic methanogens, a cytochrome-independent or flavin-based electron bifurcation (FBeB) system forms a very low-potential ferredoxin during the reduction of the heterodisulfide.[49] Although ATP is not directly synthesized, the consumption of electrochemical sodium ion potential ($\Delta\mu Na^+$) for the ferredoxin reduction is avoided and energy is saved for the reduction of CO_2.[50] In methylotrophic, aceticlastic, and some hydrogenotrophic methanogens, a cytochrome-dependent electron transfer (CDeT) system reduces a lipophilic electron carrier methanophenazine (MP) and forms an electrochemical proton potential ($\Delta\mu H^+$), which can be used for the generation of a low-potential ferredoxin for CO_2 reduction or for ATP synthesis.[45, 46, 48]

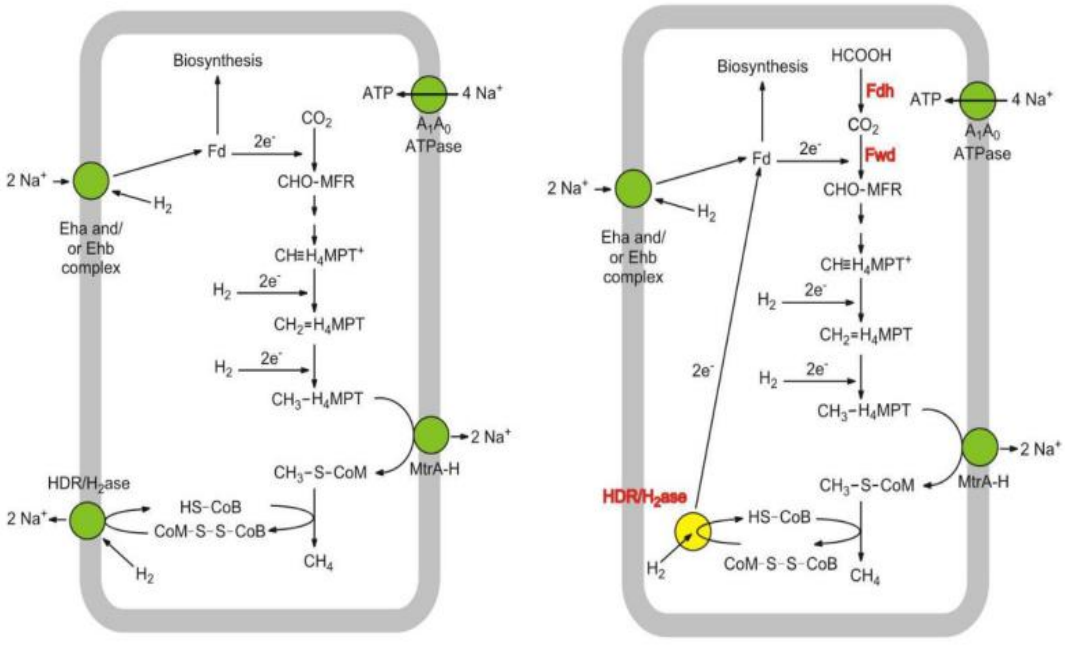


Figure 4.4 Energy conservation in methanogens from reference 45.

1.4 Anaerobic methane oxidation and methanotrophic Archaea

Methane can be oxidized by aerobic and anaerobic processes. The aerobic oxidation of methane involves a diverse group of gram-negative bacteria, which combine oxygen and methane to form formaldehyde. The formaldehyde can be used in the serine pathway or the ribulose monophosphate (RuMP) pathway, which releases CO_2.[51] Anaerobic methanotrophic archaea (ANME) catalyze the anaerobic oxidation of methane (AOM). According to the analysis of 16S rRNA sequences, ANME are divided into three groups (ANME-1, -2, -3) involving the methanogenic orders *Methanomicrobiales* and *Methanosarcinales.*[52-54] ANME are often found as part of syntrophic consortia with sulfate-reducing bacteria (SRB).[55] However, the mechanism of the ANME–SRB association remains unclear. AOM is also coupled to the reduction of other electron acceptors (i.e., Mn(IV), Fe(III), nitrate, and nitrite).[56, 57]

AOM is proposed to operate as a reverse methanogenesis pathway.[58] All the genes involved in the methanogenesis process have been found in ANME, except the N_5, N_{10}-methylenetetrahydromethanopterin (methylene-H_4MPT) reductase (Mer), which catalyzes the reduction of methylene-H_4MPT to methyl-H_4MPT.[54]

1.5 Methyl-coenzyme M reductase

Methyl-coenzyme M reductase (MCR) is the key enzyme in both methanogenesis and the anaerobic oxidation of methane. In methanogenesis, MCR catalyzes the final step in the formation of methane, which converts methyl-coenzyme M (methyl-mercaptoethanesulfonate, CH_3-S-CoM) and coenzyme B (N-7-mercaptoheptanoylthreonine phosphate, HS-CoB) into methane and the mixed heterodisulfide (CoM-S-S-CoB) (Fig 1.5).[59] The biochemical standard Gibbs free energy change of this reaction is about -30 ± 10 kJ/mol, which indicates that it is a

highly exergonic reaction.[58] Most methanogens contain more than one MCR isozymes, which are encoded by different operons and expressed under different conditions.[60]

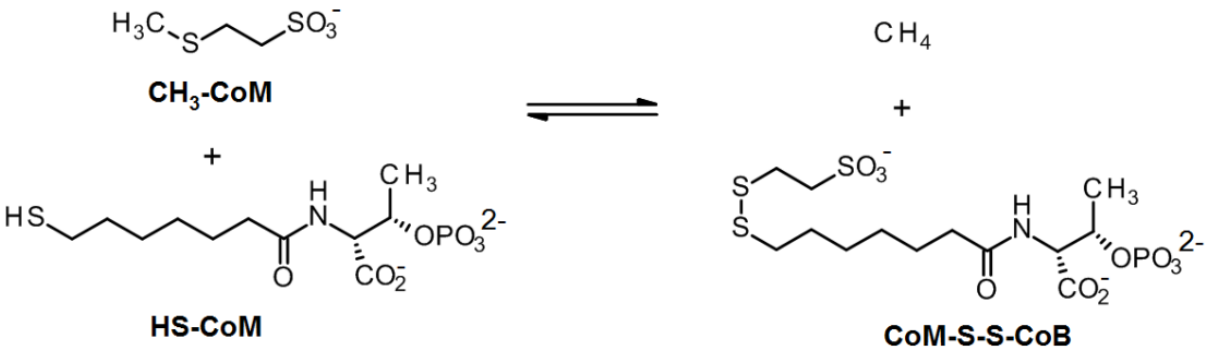


Figure 5.5 Reaction catalyzed by methyl-coenzyme M reductase.

1.5.1 Structure of MCR

The first solved crystal structure of MCR was from *Methanothermobacter marburgensis*.[61] It shows that MCR consists of a dimer of heterotrimers comprised of α, β, and γ subunits, which are encoded by the *mcrABG* genes (Fig 1.6). Each trimer binds one equivalent of coenzyme F430, a nickel-containing tetrapyrrole that is required for activity.[58, 62] Studies showed that the Ni(II)-MCR, which is silent in electronic paramagnetic resonance (EPR) spectroscopy, is inactive. Ni(I)-MCR is the active form of MCR.[58]

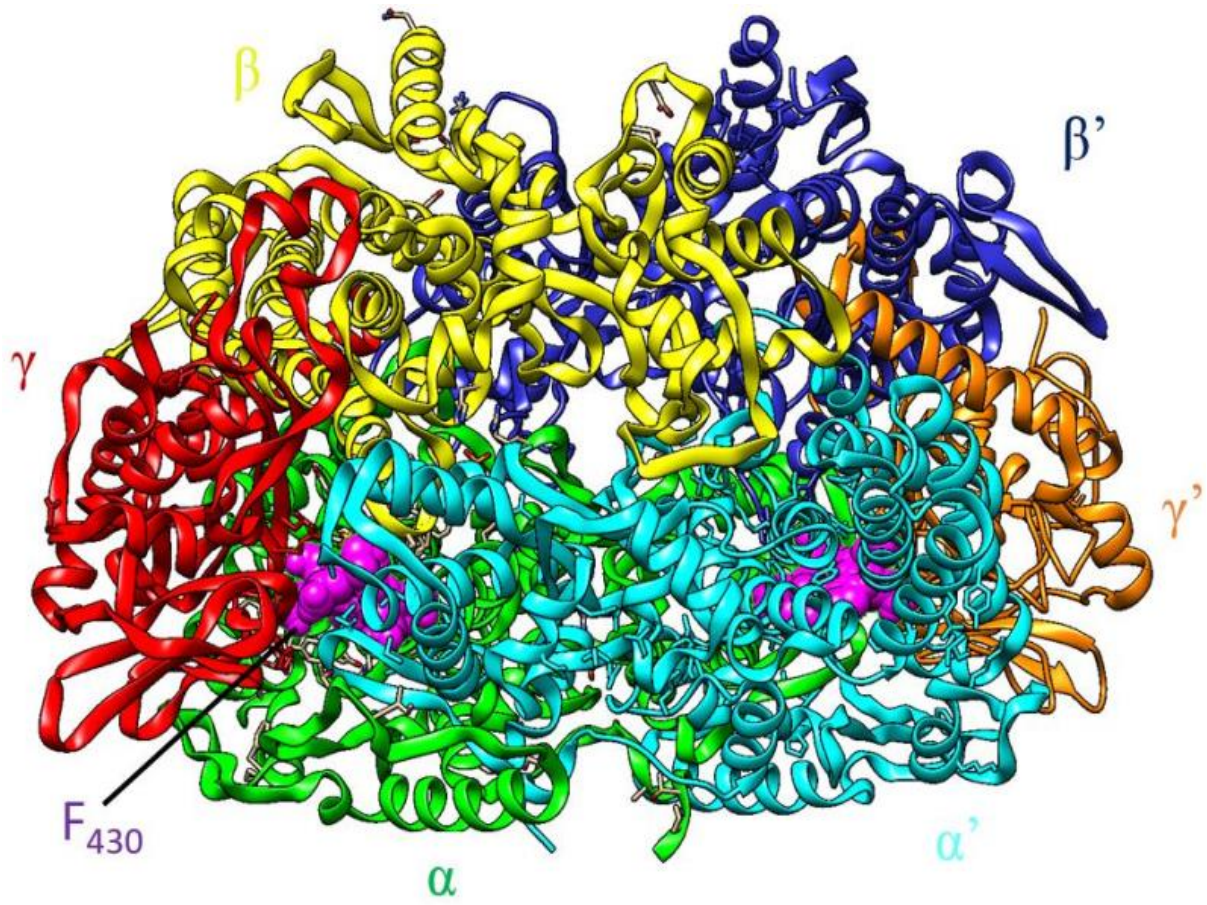


Figure 6.6 The structure of methyl-coenzyme M reductase from reference 61.

The crystal structure of MCR incorporated with its substrate has been solved. The active site containing coenzyme F430 is at the end of a narrow 50 Å long channel, in which the substrates enter and react (Fig 1.7).[58] In the absence of substrates, the channel is filled with water molecules.[63] High-resolution crystal structures revealed that MCR contains five unprecedented post-translationally modified amino acids surrounding the active site, including 2-(*S*)-methylglutamine, 5-(*S*)-methylarginine, 1-*N*-methylhistidine, *S*-methylcysteine, and thioglycine residues.[64]

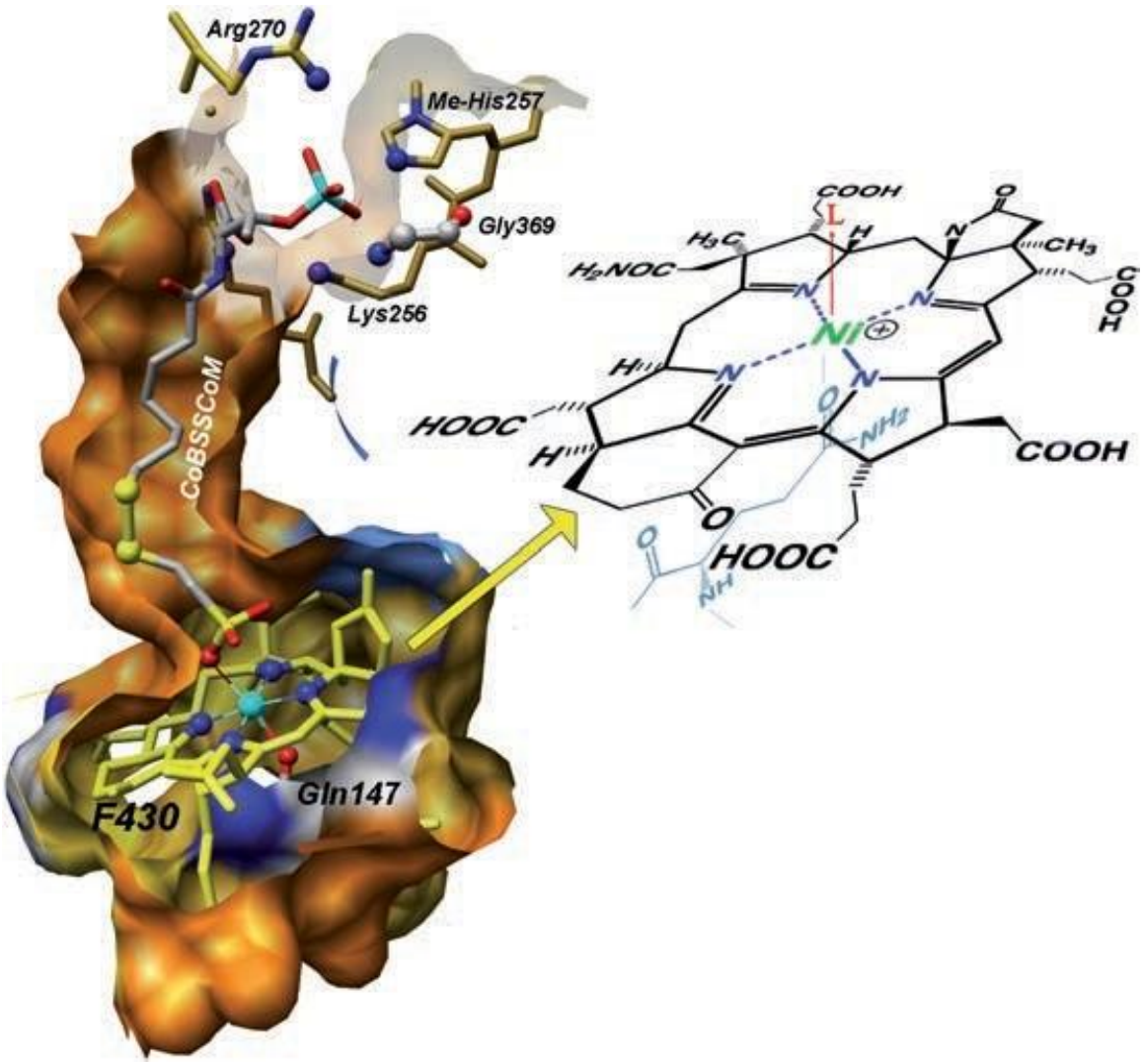


Figure 7.7 Structures of the channel in MCR and the coenzyme F430 from reference 58.

1.5.2 Mechanism of MCR catalysis

The mechanism of MCR catalysis has not been fully elucidated. Based on spectroscopic and computational analysis, there have been several proposed mechanisms including three involving a methyl-Ni(III) (mechanism I), a methyl radical (mechanism II), or a methyl anion (mechanism III) intermediate (Fig 1.8).[58]

In mechanism I, the reaction is initiated by a nucleophilic attack of the Ni(I) on the methyl group of methyl-CoM to produce a methyl-Ni(III) intermediate (Fig 1.8).[65] Then, an

electron is transferred from coenzyme M to methyl-Ni(III), which generates a methyl-Ni(II) and a coenzyme M thiyl radical. Finally, a proton is abstracted by the methyl group of methyl-Ni(II) from coenzyme B to form methane. Meanwhile, the coenzyme M thiyl radical reacts with CoB to form a disulfide anion radical (CoBS-SCoM$^{\bullet-}$) that transfers an electron to Ni(II) to regenerate Ni(I) for the next cycle.[58] This mechanism was supported by an inhibition study of MCR with methyl halides and 3-bromopropanesulfonate (BPS).[66-69]

In mechanism II, the reaction is initiated by attack of the Ni(I) on the sulfur atom of methyl-SCoM to cleave the methyl–sulfur bond, generating a Ni(II)-SCoM complex and a methyl radical (Fig 1.8). The methyl radical abstracts a hydrogen atom from HS-CoB to generate methane.[70-72] Then, the resulting thiyl radical reacts with the Ni(II)-SCoM complex to form a heterodisulfide anion radical (CoBS-SCoM$^{\bullet-}$). Finally, an electron is transferred to Ni(II) to regenerate Ni(I) for the next cycle.[58] This mechanism is based on density functional theory (DFT) calculations, transient kinetics, rapid freeze-quench (RFQ) EPR, and magnetic circular dichroism (MCD) methods.[69, 73] However, a mechanistic conundrum is that the distance between the thiol group of HS-CoB and the thiol of methyl-SCoM is too long for such a reaction. Recently, Ragsdale and coworkers proposed a new mechanism that involves methyl-SCoM binding to the Ni(I) through its sulfonate group and long-range electron transfer to generate the methyl radical.[74]

In mechanism III, the reaction is initiated by a nucleophilic attack of the Ni(I) on the sulfur atom of methyl-CoM to generate a Ni(III)-SCoM complex and a methyl anion, which undergoes protonation from HS-CoB to form methane and a CoBS^{-} anion (Fig 1.8).[58] This mechanism is based on the Ni(III)-SR intermediate that has been generated and characterized.[69, 75, 76]

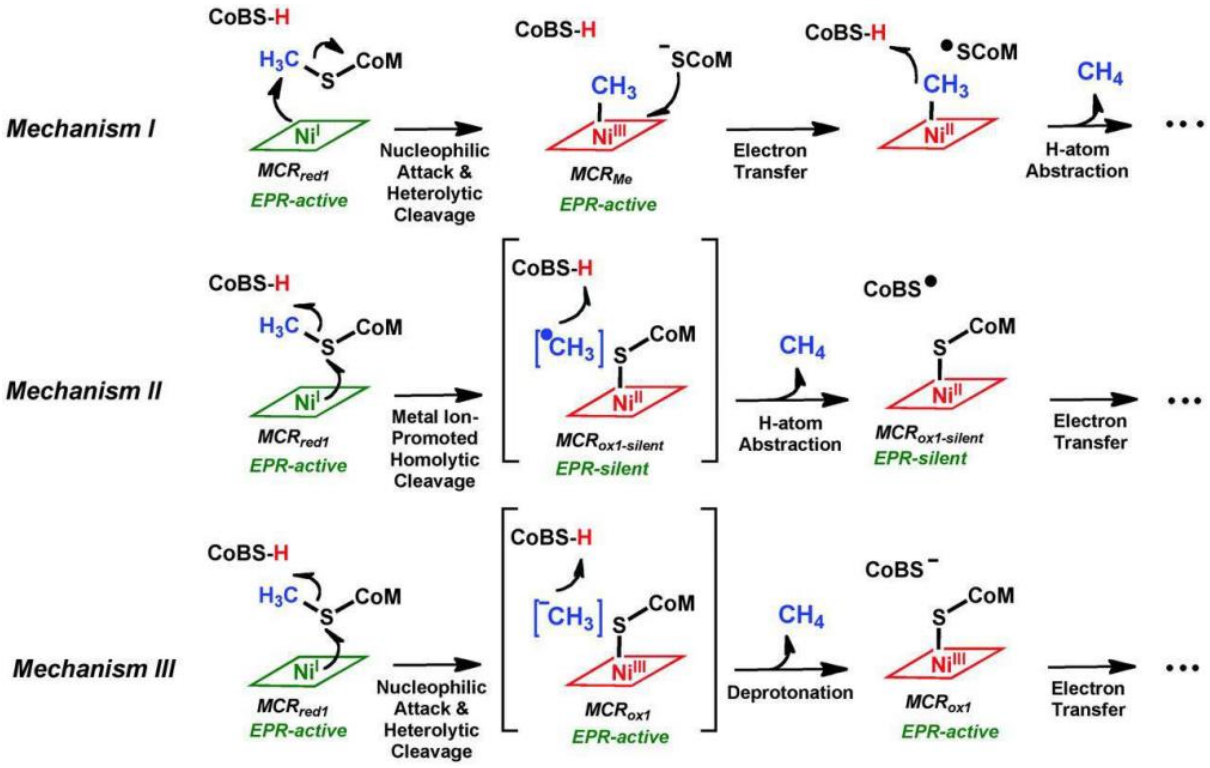


Figure 8.8 Initial steps in the three mechanisms of MCR catalysis from reference 69.

1.5.3 Post-translational modifications of MCR

As noted above, the active site of MCR has five post-translational modifications (PTMs) in the active site, including 2-(*S*)-methylglutamine, 5-(*S*)-methylarginine, 1-*N*-methylhistidine, *S*-methylcysteine, and thioglycine (Fig 1.9). Recently, two other PTMs, didehydroaspartate in *M. marburgensis* and 6-hydroxy-tryptophan in *Methanotorris formicicus* were also found, but their function and biosynthesis are still unclear (Fig 1.9).[73, 77] These PTMs are not conserved in all methanogens, except for the 1-*N*-methylhistidine and thioglycine PTMs.[78]

Based on a study from Selmer and coworkers, 1-*N*-methylhistidine and *S*-methylcysteine PTMs are likely to proceed via methyl group transfer from *S*-adenosylmethionine (SAM) to substitute the hydrogen of the cysteine thiol group or the ring nitrogen of histidine.[79] The 2-(*S*)-methylglutamine and 5-(*S*)-methylarginine PTMs involve the methylation of a carbon atom, and

the protons on these carbons are not sufficiently acidic to generate a carbon nucleophile. Thus, putative cobalamin-dependent radical SAM enzymes have been proposed to introduce these PTMs.[80] Layer group revealed methanogenesis marker protein 10 (Mm10) from *M. maripaludis* plays an important role in methanogenesis.[81] Whitman and Duin groups knocked out Mm10, which caused the McrA subunit to lose the 5-(*S*)-methylarginine PTM and the maximal rate of methane production was reduced by 40 to 60%.[82] Recently, Booker and coworkers showed that Mm10 from *M. acetivorans* catalyzes the methylation of Arg in a 13 amino acid peptide mimic of the McrA subunit in the presence of cobalamin.[83]

Thioglycine residues are very rare in nature and in proteins is only observed in MCR.[84] The biosynthesis of thioglycine is proposed to involve two proteins, YcaO and TfuA. Metcalf and coworkers knocked out both genes and *ycaO* only in *M. acetivorans,* resulted in the specific loss of the thioglycine PTM. This mutation also led to a conformational change near the active site.[85] Later, Mitchell and coworkers reported the *in vitro* thioamidation of a peptide substrate using heterologous YcaO and TfuA proteins from *M. acetivorans* and confirmed that the reaction proceeds via adenosine triphosphate (ATP)-dependent backbone *O*-phosphorylation.[84]

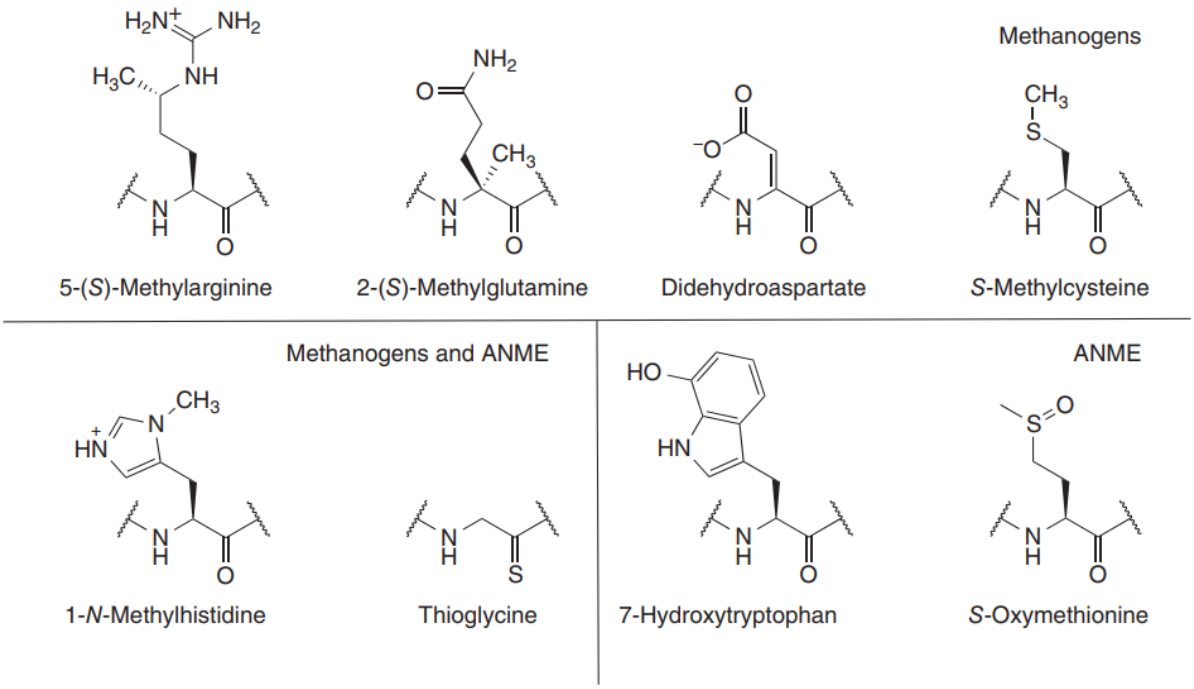


Figure 9.9 Posttranslational modifications found in the active site region of MCR from methanogens and ANME from reference 76.

1.6 Coenzyme F430 biosynthesis

Coenzyme F430 is a nickel-containing tetrapyrrole that is required by methyl coenzyme M reductase (MCR). It was first discovered in *M. thermoautotrophicum* and was named after its absorption maximum at 430 nm.[86] The structure of coenzyme F430 was solved by X-ray crystallography and nuclear magnetic resonance (NMR) spectroscopy, which showed that it has a highly reduced isobacteriochlorin ring system with two additional rings, a γ-lactam E ring and a keto-containing carbocyclic F ring.[86]

In our lab, the coenzyme F430 biosynthesis (*cfb*) genes were identified using comparative genomics and the encoded enzymes were characterized *in vitro*.[87] They include homologs of *cbiX*^{S} (small sirohydrochlorin cobaltochelatase) and *cbiA* (cobyrinic acid a,c-diamide synthase), two genes involved in the cobalamin biosynthetic pathway; *nifH* (nitrogenase Fe protein) and

nifD (nitrogenase MoFe protein, α-subunit), two components of the key enzyme of nitrogen fixation, and *murD*, an ATP-dependent ligase involved in peptidoglycan biosynthesis.[87]

1.6.1 Sirohydrochlorin cobaltochelatase

Most cyclic tetrapyrroles coordinate a metal ion such as Fe, Mg, Co, or Ni.[1] They need a specific metal ion chelatase that is responsible for inserting the correct metal ion. Based on the size and energy requirements, chelatases are divided into three broad classes (Fig 1.10).[88] The class I chelatases require three subunits and couple the hydrolysis of ATP to metal ion insertion. They include the magnesium chelatase (ChlHID) in chlorophyll biosynthesis and the cobaltochelatase (CobNST) in the cobalamin biosynthetic pathway.[88] The class II chelatases are either monomers or homodimers and ATP-independent. They include chelatases from heme and cobalamin biosynthesis. The class III chelatases are ATP-independent multifunctional proteins, which are involved in siroheme biosynthesis.[88, 89]

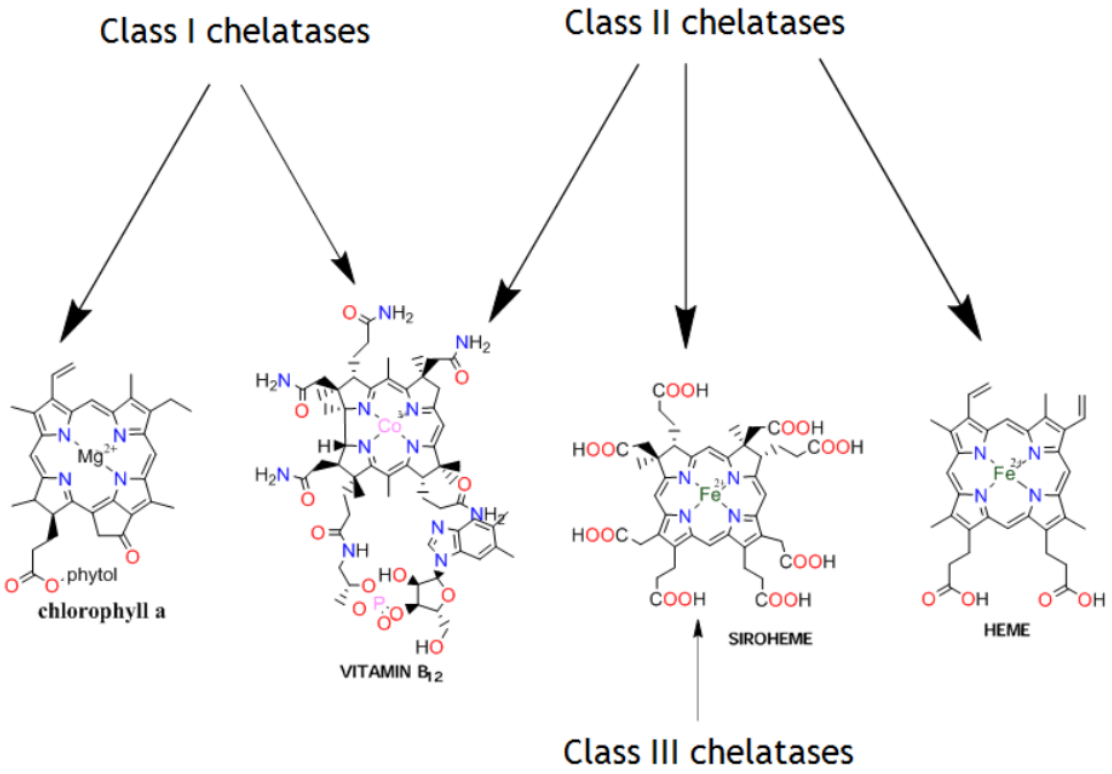


Figure 10.10 The three classes of chelatase from reference 88.

CbiX belongs to the class II chelatases and is involved in the cobalamin biosynthesis pathway. It chelates Co^{2+} into sirohydrochlorin, forming cobalt-sirohydrochlorin and releases two protons (Fig 1.11).[90] CfbA catalyzes the specific Ni-chelation of sirohydrochlorin (Fig 1.11).[87]

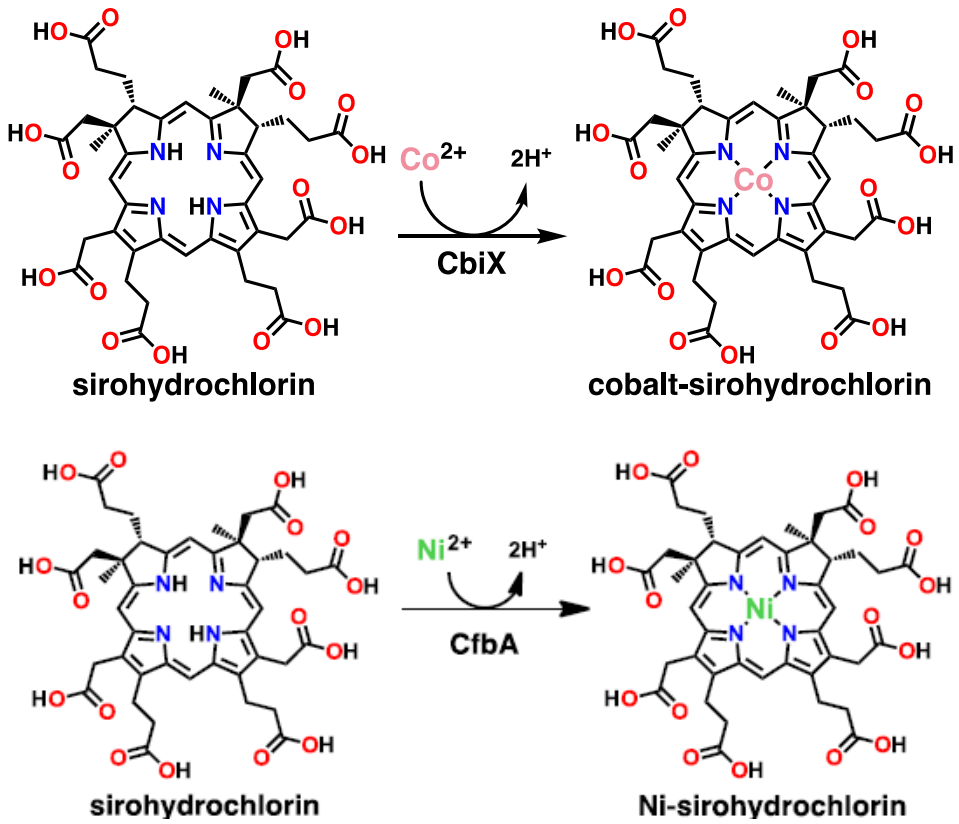


Figure 11.11 Reaction catalyzed by sirohydrochlorin cobaltochelatase and CfbA from reference 87.

1.6.2 Cobyrinic acid a,c-diamide synthase

Cobyrinic acid *a,c*-diamide synthetase catalyzes the adenosine triphosphate (ATP)-dependent amidation of the *a*- and *c*-acetate side chains of cobyrinic acid using glutamine as the nitrogen source (Fig 1.12). It is the first glutamine amidotransferase in the anaerobic cobalamin biosynthesis pathway.[91, 92] Based on a study by Raushel and coworkers, the reaction is initiated by the activation of the *a*- and *c*-carboxylate groups through phosphorylation with ATP, followed

by nucleophilic substitution of the phosphate groups with ammonia released through the hydrolysis of glutamine.[91] The amidation at the *a*- and *c*-carboxylate groups of cobyrinic acid is sequential, which is accommodated by a simple rotation of cobyrinic acid in the active site. The cobyrinic acid *c*-monoamide is released into solution prior to the formation of the *a,c*-diamide product.[91] CfbB catalyzes the adenosine triphosphate (ATP)-dependent amidation of the *a*- and *c*-acetate side chains of Ni-sirohydrochlorin using glutamine as an ammonia source (Fig 1.12).[87]

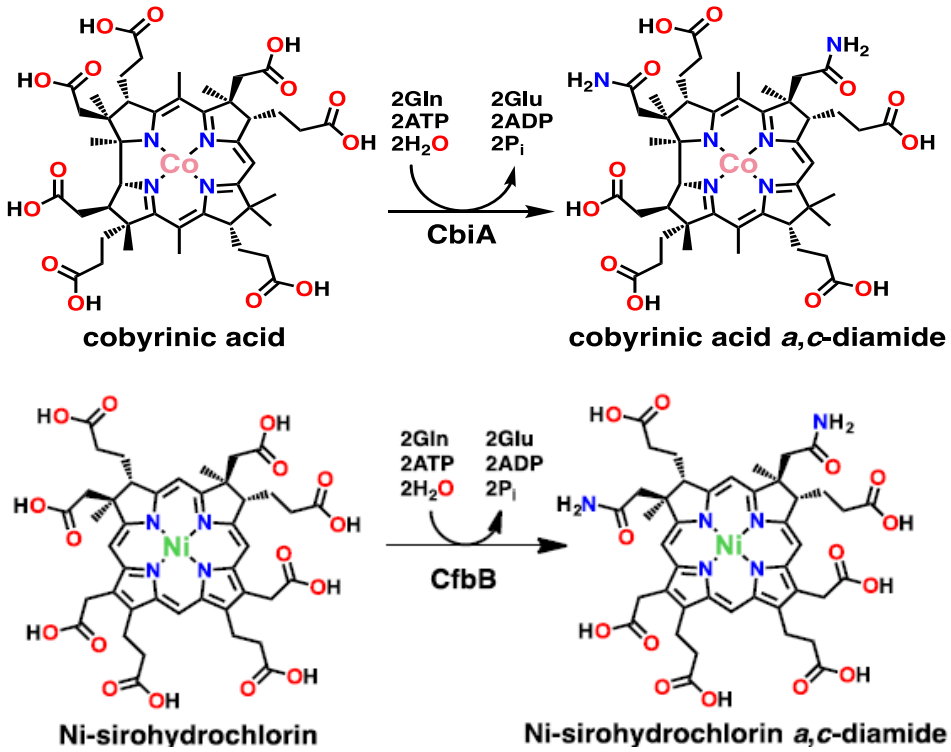


Figure 12.12 Reaction catalyzed by cobyrinic acid *a,c*-diamide synthetase and CfbB from reference 87.

1.6.3 Nitrogenase

Nitrogenase is responsible for biological N_2 fixation, which converts inorganic nitrogen gas to ammonia that can be assimilated into nitrogen-containing compounds such as proteins and nucleic acids.[93, 94] Nitrogenase is produced by free-living bacteria like cyanobacteria, as well as

symbiotic bacteria like *Rhizobium*.[95] Nitrogenase couples the hydrolysis of 16 ATPs to the 8-electron reduction of dinitrogen and 2 protons to ammonia and hydrogen gas, respectively (Fig 1.13).[96] Other than its physiological substrates, nitrogenase can reduce a wide range of small doubly or triply bonded substrates, such as acetylene (C_2H_2), ethylene (C_2H_4), cyanide (CN^-), nitrous oxide (N_2O), nitrite (NO_2^-), azide (N_3^-) and hydrazine (N_2H_4).[97] There are three types of nitrogenase found in various nitrogen-fixing bacteria: molybdenum (Mo) nitrogenase, vanadium (V) nitrogenase, and iron-only (Fe) nitrogenase.[97]

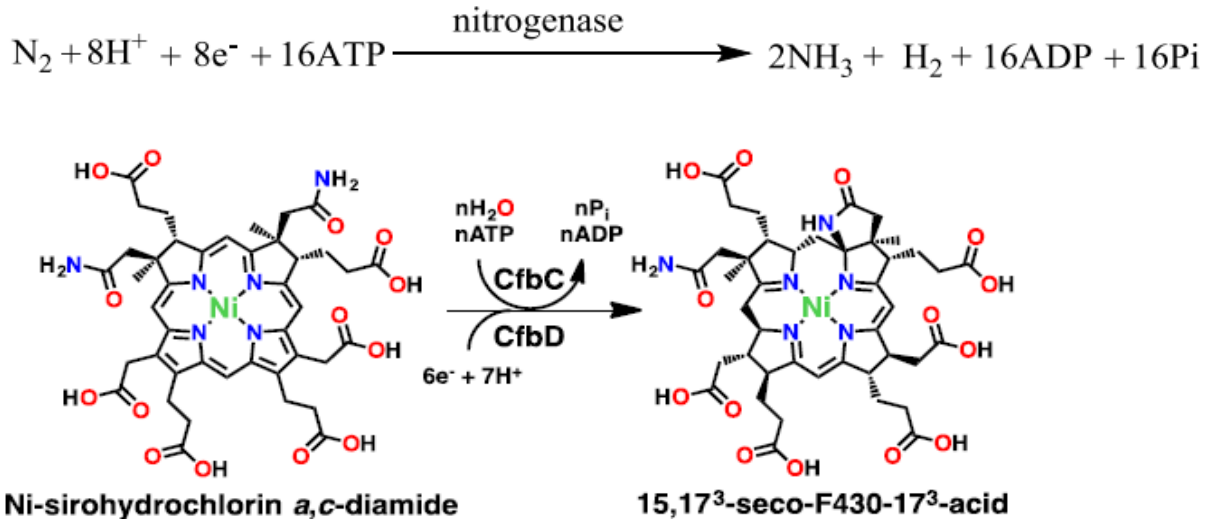


Figure 13.13 Reaction catalyzed by nitrogenase and CfbCD from reference 87.

The molybdenum nitrogenase is the most well characterized member of this enzyme family. It consists of a homodimeric Fe protein and a heterotetrameric MoFe-protein (Fig 1.14). There are two MgATP binding sites in each subunit of the Fe protein (Fig 1.14).[97, 98] The binding and hydrolysis of ATP influences the association/dissociation of the Fe and MoFe proteins. The two subunits of the Fe protein are bridged by a [Fe_4S_4] cluster, which can transfer electrons from a reducing agent (e.g., ferredoxin or flavodoxin) to the MoFe protein.[98] The MoFe protein contains two P-clusters at the interface between the α and β subunits and two M-clusters (FeMo cofactors)

in each α subunit (Fig 1.14).[97] The electrons are transferred sequentially from the [Fe_4S_4] cluster of the Fe protein, via the P-cluster, to the M-cluster of the MoFe protein.[97, 98]

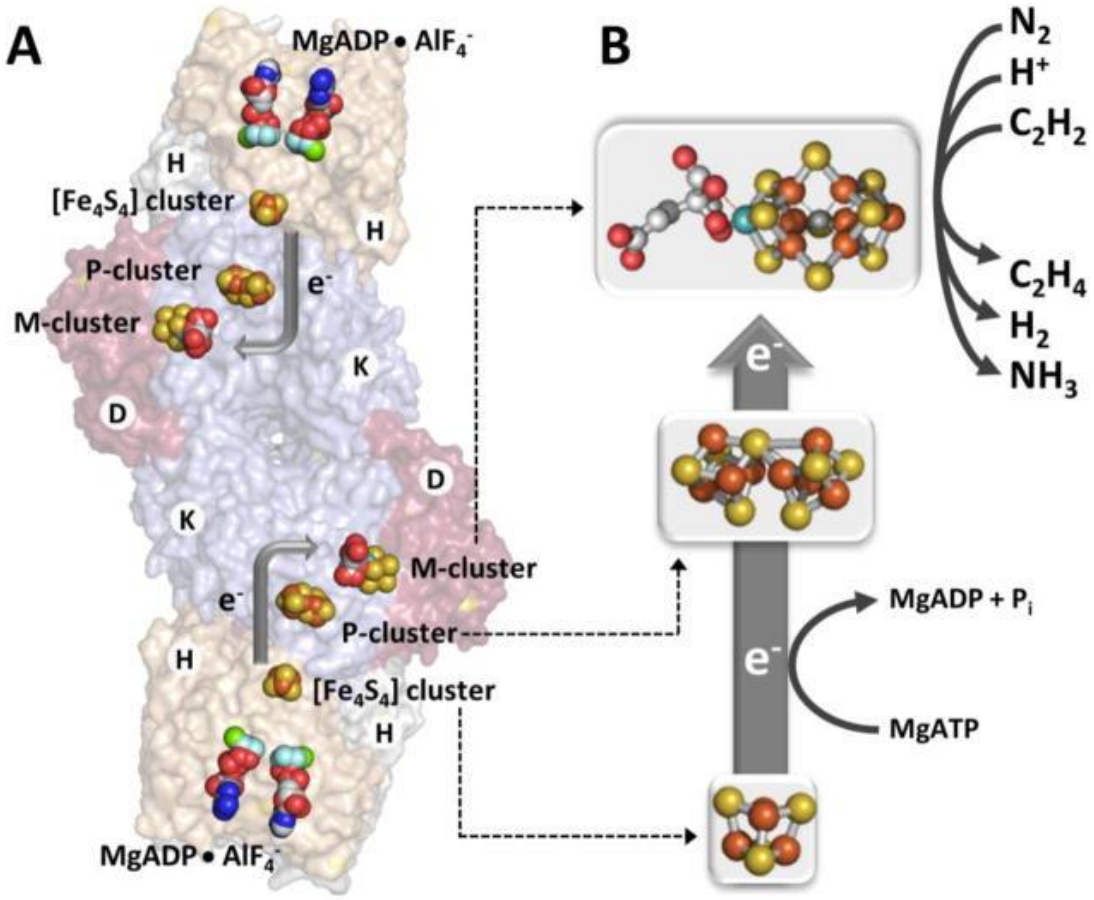


Figure 14.14 The structure of nitrogenase and its active site. Adapted from reference 96.

A Lowe-Thorneley kinetic model was developed to describe the delivery of eight protons and eight electrons through the reaction based on steady state, freeze quench, and stopped-flow kinetic measurements.[93, 99] According to the model, the protons and electrons bind to the metal cluster and are stored as hydrides until the E_4 state, which binds N_2 release of H_2. Addition of four additional protons and electrons produces ammonia, which is released to go back to the resting state (Fig 1.15).[93, 99] The mechanism of nitrogenase prior to the E_4 state in the Lowe-Thorneley kinetic model is generally accepted. However, there are currently two hypotheses

about the mechanism after the E_4 state: the "distal" and the "alternating" pathways (Fig 1.16).[100] In the distal pathway, the three hydrogens are sequentially added to the terminal nitrogen, which is released as ammonia, then another three hydrogens are transferred to the nitrogen directly bound to the metal to generate the second ammonia.[101, 102] In the alternating pathway, one hydrogen first binds to the terminal nitrogen, then another hydrogen bind to the nitrogen directly bound to the metal. This alternating pattern continues until two equivalents of ammonia are released.[103]

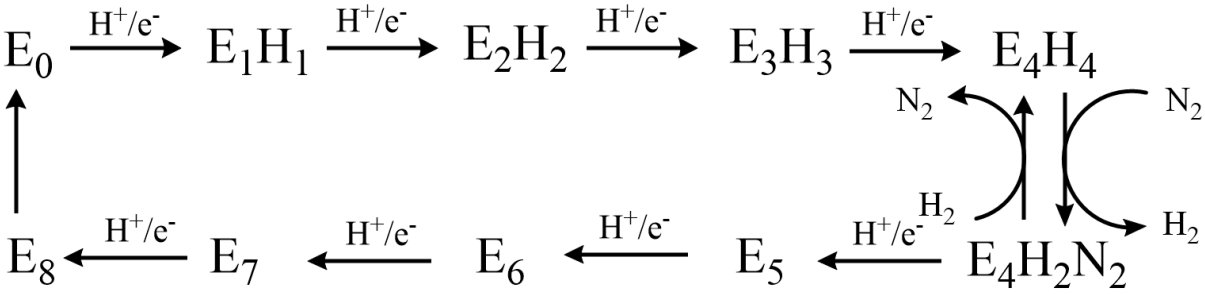


Figure 15.15 Lowe-Thorneley kinetic model for reduction of nitrogen to ammonia by nitrogenase from reference 93, 99.

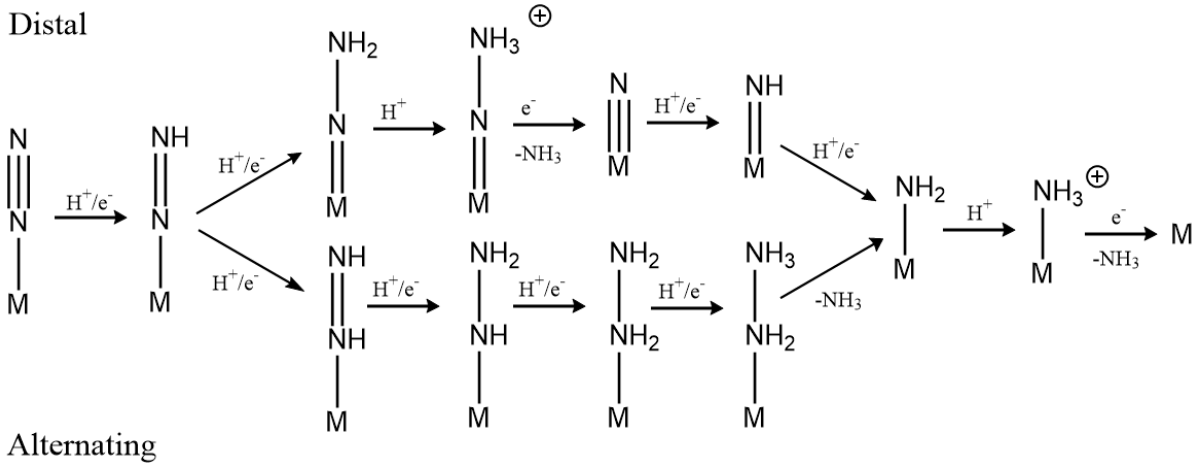


Figure 16.16 Distal vs. alternating mechanism pathway in nitrogenase from reference 100.

The vanadium nitrogenase and iron-only nitrogenase can both be found in select species of *Azotobacter* as an alternative nitrogenase since they are less efficient in N_2 reduction and need more electrons/protons in the reaction than molybdenum nitrogenase.[104] They also consist of two components as molybdenum nitrogenase, but the MoFe protein is replaced with VFe or FeFe proteins in vanadium and iron-only nitrogenase, respectively.[94] The vanadium nitrogenase has the same mode of action as molybdenum nitrogenase because they have similar primary sequences and metal centers.[97] The difference between vanadium nitrogenase and molybdenum is shown on the abilities to reduce CO and CO_2. The vanadium nitrogenase can reduce CO and CO_2 more efficiently and reduce to more various substrates.[105, 106]

CfbCD converts Ni-sirohydrochlorin *a,c*-diamide to 15,17^3-seco-F430-17^3-acid. This remarkable transformation involves a 6-electron reduction of the isobacteriochlorin ring system, cyclization of the *c*-acetamide side chain to form a γ-lactam ring, and the formation of 7 stereocenters.[87] In Chapter 3, a mechanism of CfbCD catalysis was proposed and a series of experiments were performed to study the mechanism. Phylogenetic analysis of the CfbCD complex suggests that it is representative of a more ancient lineage of the nitrogenase superfamily, and a thorough investigation of its structure and function is likely to shed light on the mechanisms and evolution of these important metalloenzymes.

1.6.4 Mur Ligase

UDP-*N*-acetylmuramoyl-pentapeptide is a precursor of the peptidoglycan cell wall. UDP-*N*-acetylmuramoyl-L-alanine:D-glutamate ligase (MurD) catalyzes the ATP-dependent formation of a peptide bond between UDP-*N*-acetylmuramoyl-L-alanine (UMA) and D-glutamate during peptidoglycan biosynthesis (Fig 1.17).[107, 108] MurD consists of three domains. Domain 1 provides the binding site for the UDP moiety of UMA. Domain 2 was characterized as a GTPase domain

involved in ATP binding. The active site is located between the GTPase domain and Domain 3.[107] CfbE carries out the ATP-dependent cyclization of the *g*-proprionate side chain to generate the carbocyclic F ring (Fig 1.17).[87]

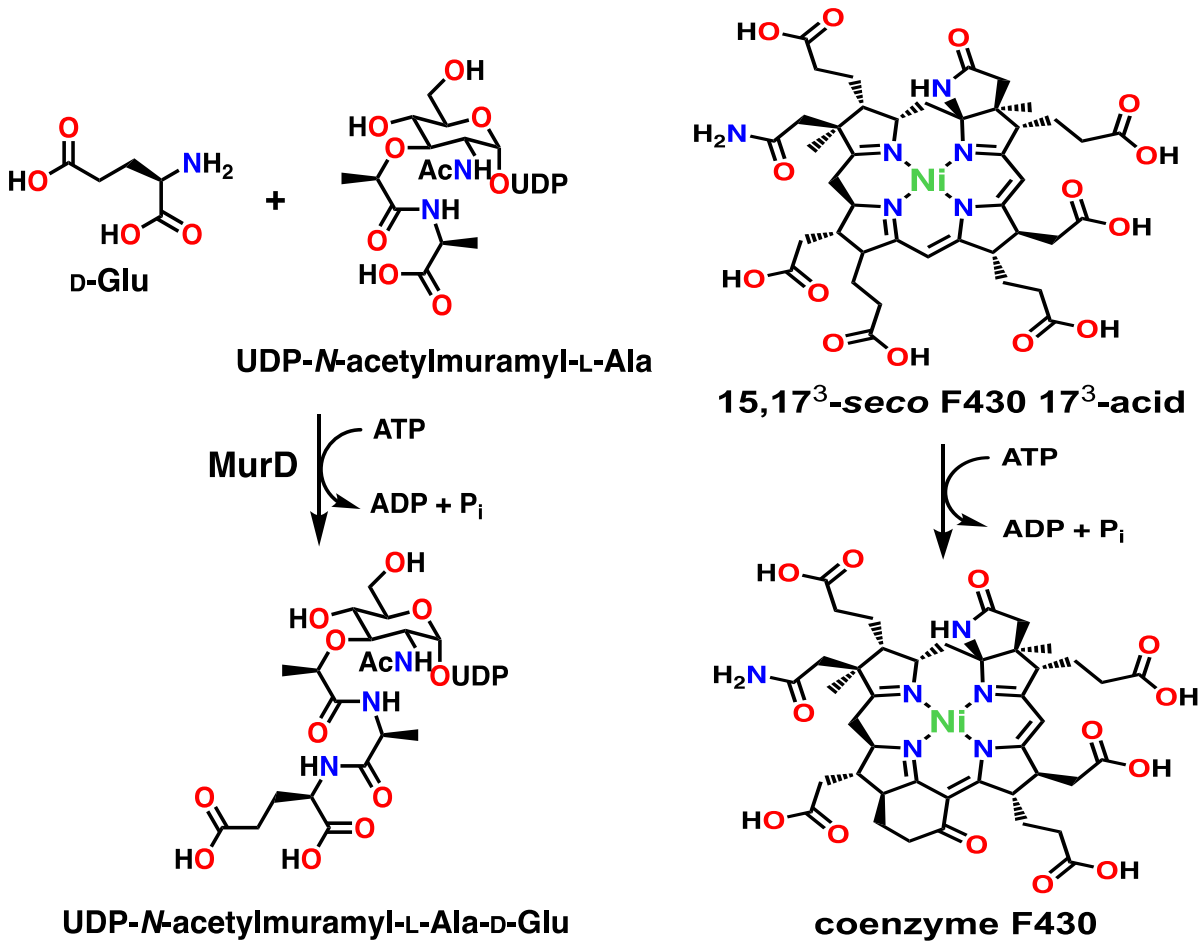


Figure 17.17 Reaction catalyzed by UDP-*N*-acetylmuramoyl-L-alanine:D-glutamate ligase and CfbE from reference 87.

In Chapter 2, experiments designed to introduce the coenzyme F430 biosynthesis pathway into a heterologous host are described, which provide a critical step towards metabolic engineering efforts using MCR for methane conversion. Previously, we constructed four plasmids that carry all the *cfb* genes and some required accessory genes. All the plasmids were transformed into *E. coli* BL21 (DE3) cells, and the enzymes were coexpressed under various

culture conditions. The synthesized tetrapyrroles were extracted with methanol and analyzed using high-performance liquid chromatography (HPLC). Unfortunately, Ni-sirohydrochlorin *a*,*c*-diamide was the last biosynthetic intermediate detected, indicating that the pathway stopped at the CfbCD reaction. We therefore constructed a plasmid, pRSFDuet:*Mt_cfbCD-porABDG*, which contains an additional copy of the coenzyme F430 biosynthesis genes *cfbCD* from *Methanosarcina thermophila* and genes encoding the native physiological reducing system pyruvate:ferredoxin oxidoreductase (PFOR). Several cell lines were constructed combining this plasmid with the *cfb*, *mcr*, *sirAC* (for sirohydrochlorin biosynthesis), *isc* (for iron sulfur cluster biogenesis), and/or methanogenesis marker (*mm*) genes (which may play a role in MCR assembly/activation) and tested for coenzyme F430 biosynthesis activity under a variety of growth conditions.

1.7 Iron

Iron is an essential element for life. There are around 4 to 5 grams of iron in the human body. More than 70% of iron is in the form of hemoglobin, which is a key protein involved in oxygen transport.[109] Iron also functions in energy production, DNA synthesis, and cell growth and replication. Most of the remaining iron is stored. Either too much or too little iron can be detrimental. Thus, organisms have developed a complicated and precise iron regulation system.[109]

Hemoglobin and myoglobin are heme-containing proteins involved in the transport and storage of oxygen. Hemoglobin is the main protein found in red blood cells. It is heterotetrametric and each monomer contains a heme.[110] The iron in heme can reversibly bind oxygen by a coordinate bond. When one oxygen molecule binds to the iron, it causes a slight conformational change and helps oxygen bind to the three remaining hemes in the other subunits.

Thus, the binding of oxygen is cooperative, which can obtain oxygen rapidly in the lungs.[110] Myoglobin is monomeric and consists of eight α-helices with a heme in the center. This heme binding pocket helps myoglobin have a high affinity for oxygen and allows it to function in oxygen storage in muscle cells.[111]

Iron-sulfur clusters are protein cofactors involved in many physiological processes. They accept or donate electrons by changing the oxidation state of iron atoms between the +2 and +3 oxidation states.[112] Ferredoxin contains iron-sulfur clusters that can transfer electrons in photosynthesis or the nitrogen fixation pathway. In the electron transport chain, NADH dehydrogenase (Complex I) contains iron-sulfur clusters that are used to couple the oxidation of NADH to the generation of an electrochemical proton gradient that can be used for ATP synthesis.[113] Cytochromes are heme-containing enzymes. They catalyze many redox reactions and serve as electron carriers. Cytochrome P450 is a heme-containing oxygenase, which catalyze the oxygenation of fatty acids, steroids, and xenobiotics.[114]

Catalase is a heme-containing enzyme that catalyzes the disproportionation of hydrogen peroxide to water and oxygen. Hydrogen peroxide is a reactive oxygen species (ROS), which can damage cells and tissues. Thus, catalase plays an important protective role. The mechanism of catalase involves the formation of high oxidation states of iron (i.e., the Fe(IV)-oxo of compound I).[115]

Although iron is essential for life, free iron inside the cell can be toxic and result in the formation of free radicals, which cause oxidative stress and cellular damage. Thus, organisms have developed the means of iron absorption, storage, utilization, and recycling.[116] Iron is typically found in two oxidation states, the insoluble ferric form (Fe^{3+}) and soluble ferrous form (Fe^{2+}). Bacteria secrete high-affinity ferric iron chelators (siderophores) outside the cell to trap

iron (Fig 1.18).[116] Gram-negative bacteria absorb ferrisiderophore complexes via specific outer membrane (OM) receptors. The TonB-ExbB-ExbD system provides energy for ferrisiderophore import. Then, ferrisiderophores are transferred from the OM receptors to cytosolic membrane (CM) ATP-binding cassette (ABC) transporters by periplasmic binding proteins (PBP). Ferrisiderophores can be reduced to ferrous form in the cytoplasm (Fig 1.18).[117] Gram-positive bacteria lack OM receptors and TonB-ExbB-ExbD systems, and instead have a special CM binding protein that is a lipoprotein tethered to the outer surface of the CM (Fig 1.18). This allows ferrisiderophores to pass through the ABC transporter.[118]

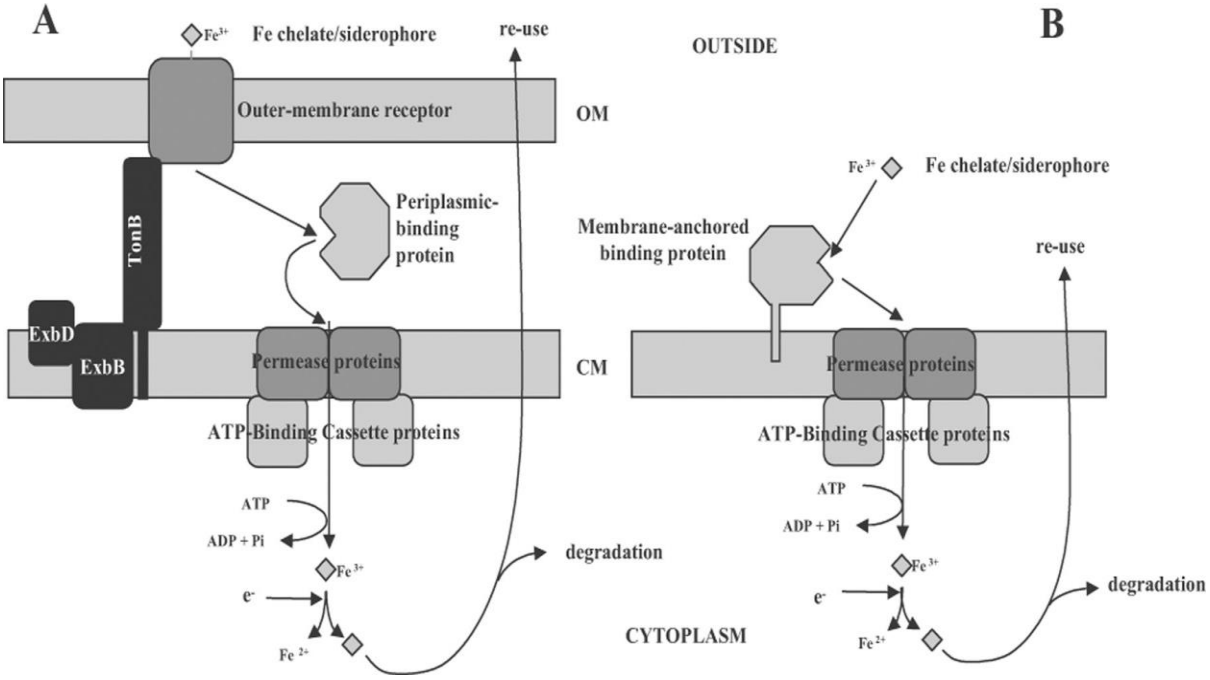


Figure 18.18 Siderophore mediated iron uptake in Gram-negative and Gram-positive bacteria from reference 115.

Bacteria store iron within iron storage proteins to prevent external iron deficiency.[119] There are two types of iron storage proteins, ferritins and bacterioferritins. They are composed of 24 subunits that form a spherical protein shell with a central cavity where the irons are stored (Fig 1.19). Iron is absorbed in the soluble ferrous form but stored in the ferric form within iron

storage proteins. This redox process is catalyzed by specific sites in the protein called ferroxidase centers, which are located in the middle of each subunit.[120, 121] When cellular iron is limiting to metabolism, stored iron can be released from ferritins or bacterioferritins. Based on a study by Rivera and coworkers, Fe^{3+} is reduced to Fe^{2+} by electrons shuttled from NADPH by way of a flavin mononucleotide (FMN) in ferritins.[121] In the bacterioferritins, ferredoxin promotes the release of iron and heme mediates electron transfer from the surface to the core.[122]

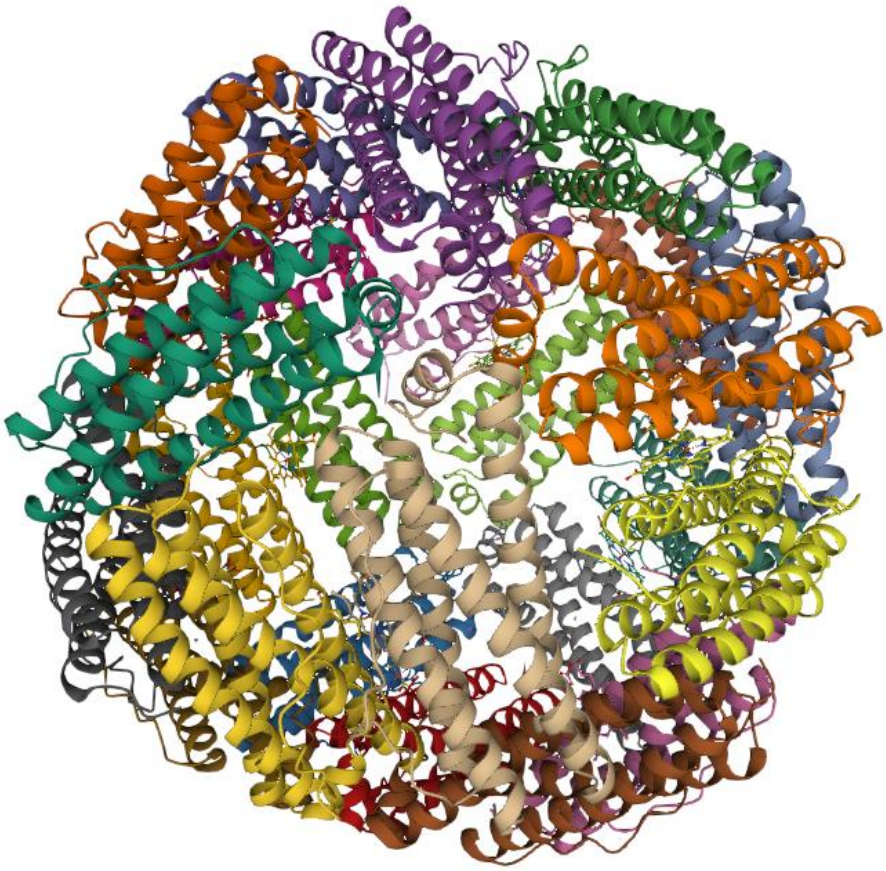


Figure 19.19 The structure of bacterioferritin from reference 121.

Although it is important to get enough iron for growth, it is just as important to ensure the iron is in a non-toxic form in the body. This requires that free Fe^{2+} is not allowed to interact with oxygen to form ROS and Fe^{3+}.[123] Reduced flavin is thought to directly transfer electrons to free iron, but it is affected by concentration of iron. Decreasing the production of siderophores or

increasing iron storage capacity can help limit the concentration of cellular iron.[124] Most bacteria regulate their iron metabolism by the ferric-uptake regulator protein (Fur). It acts as a positive repressor, which controls the iron-dependent expression of more than 90 genes. Fur represses transcription upon interaction with Fe^{2+} and derepresses transcription in the absence of Fe^{2+}.[125, 126]

1.8 Heme

Heme is a tetrapyrrole containing a central iron. The iron is held in place by coordinate bonds to the four nitrogen atoms of the porphyrin ring and has two open coordination sites for axial ligands.[127] Hemes are the most widely used metalloporphyrins and serve as prosthetic groups that define hemoproteins. Hemes are most well known as components of hemoglobin but are also found in many other important hemoproteins such as myoglobin, cytochromes, catalases, and heme peroxidase.[128]

There are several types of hemes such as heme *a*, heme *b*, heme *c*, heme d_1, and heme *o* (Fig 1.20). They all have porphyrin rings but differ in their side chains. The simplest and most common type of heme is heme *b*, which is present in hemoglobin and myoglobin. Each pyrrole ring of heme *b* contains one methyl group. Two of the pyrrole rings (A and B) have a vinyl group, and the other two (C and D) have a propionate group. The central iron can coordinate two axial ligands. Typically, one axial ligand is provided by the protein, while the other ligand can be from the protein or an exogenous substrate/inhibitor such as oxygen or carbon monoxide.[127] The difference between heme *c* and heme *b* is that the two vinyl side chains of heme *b* are modified through the formation of covalent thioether bonds with cysteine residues of the protein. Heme *c* is present in certain cytochromes (e.g., cytochrome *c*) and functions as an electron transporter in the respiratory chain. It also releases from mitochondrial and regulates cell apoptosis when the

cell is activated by an apoptotic stimulus.[129] Heme d_1 has two oxo groups on pyrrole rings A and B and an acrylate side chain on ring D. It is found in cytochrome cd_1 nitrite reductase, which catalyzes the one electron reduction of nitrite to nitric oxide and water during the denitrification.[130] Both heme *a* and heme *o* contain a 17-carbon hydroxyethyl-farnesyl side chain. They are cofactors of cytochromes and related to aerobic respiration.[131]

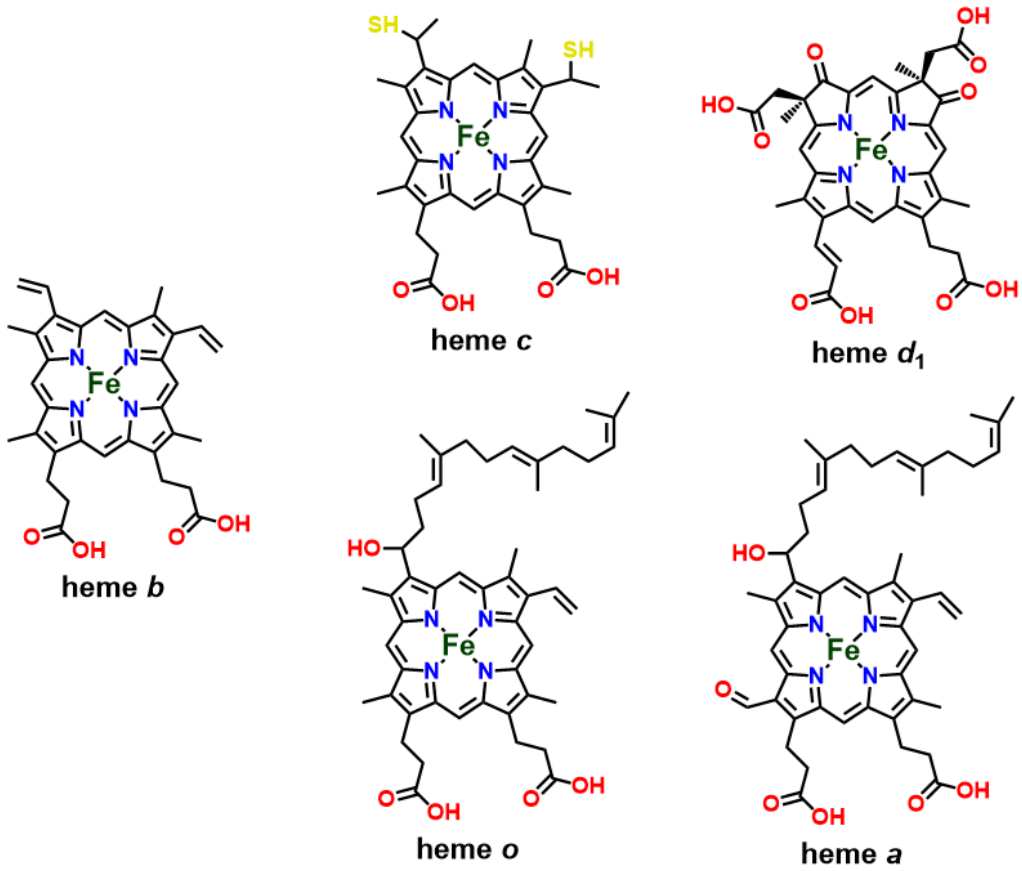


Figure 20.20 The structures of the various heme types.

As discussed above, heme is a versatile coenzyme that has multiple functions. It can be used for oxygen transport and storage in hemoglobin and myoglobin. In cytochromes, heme catalyzes redox reactions and serves as an electron carrier. In catalases and peroxidases, heme protects cells from damage by ROS.[128] Furthermore, heme can affect the expression of many genes and

regulate the differentiation and proliferation of various cell types. Thus, nearly all organisms must synthesize or acquire heme to meet these requirements.

Pathogenic bacteria acquire hemes from the host environment. They secrete proteins, known as hemophores, that bind heme and bring it to a receptor on the cell surface or they directly uptake heme with the receptor.[132] There are over 30 well-characterized outer membrane heme receptors in gram-negative bacteria. Hemes are transported into the periplasm by a TonB-ExbB-ExbD complex on the inner membrane.[133] A periplasmic binding protein (PBP) binds heme and acts as a soluble receptor for the ATP binding cassette (ABC) transporter that then transports heme into the cytoplasm.[134] In gram-positive bacteria, surface-exposed lipid-anchored proteins, cell wall–anchored proteins, and ABC transporters are required to transport heme through the peptidoglycan cell wall and across the membrane.[135]

As mentioned above, heme is a main source of iron. Both human and bacteria require heme oxygenase (HO) to liberate the tightly bound iron from heme. The canonical bacterial HO-1 catalyzes the degradation of heme to produce biliverdin, ferrous ion, and carbon monoxide.[136] Recently, Ikeda-Saito group reported three steps occur to degrade heme and each step requires O_2 activation. The first step is the hydroxylation of heme and results in an α-*meso*-hydroxyheme. In the second step, the α-*meso*-hydroxyheme reacts with O_2 to form verdoheme and CO. In the third step, the verdoheme is cleaved to release biliverdin and ferrous iron by third O_2 activation.[136] Some other heme oxygenases were mentioned in Chapter 4. Heme also can be degraded without oxygen. A *S*-adenosylmethionine methyltransferase uses flavodoxin as an electron donor to degrade heme and forms anaerobilin in *E. coli* O157:H7 (EHEC).[137] It benefits enteric pathogens to acquire iron under anaerobic environment.

Currently, all the heme oxygenases degrade *b*-type heme, which is the most abundant heme in the human body. It is unclear whether other types of heme can be utilized to recycle iron. In Chapter 4, we identified a noncanonical heme oxygenase specific for the degradation of *c*-type heme. It catalyzes the oxidative cleavage of *c*-type hemopeptides to a peptide-linked β- and/or δ-biliverdin with formation of CO and free iron. It may be involved in acquiring iron from degraded extracellular cytochromes and benefit the large number of *c*-type cytochromes utilized bacteria.

Chapter Two: Biosynthesis of coenzyme F430 in the heterologous host *Escherichia coli*

2.1 Introduction

Methane is the major component of natural gas, which is the cleanest fossil fuel on earth.[22] Most natural gas are obtained associated with oil and coal in remote areas, which requires an extensive and complex pipeline network to transport to end users.[23] Meanwhile, natural gas is often vented or flared, which wastes this important commodity and releases lots of carbon dioxide and methane into the atmosphere.[23] Methane is a greenhouse gas that has approximately thirty times the global warming potential of carbon dioxide.[138] Thus, there is great current interest in strategies to convert methane to liquid fuels or other more easily transported commodity chemicals.

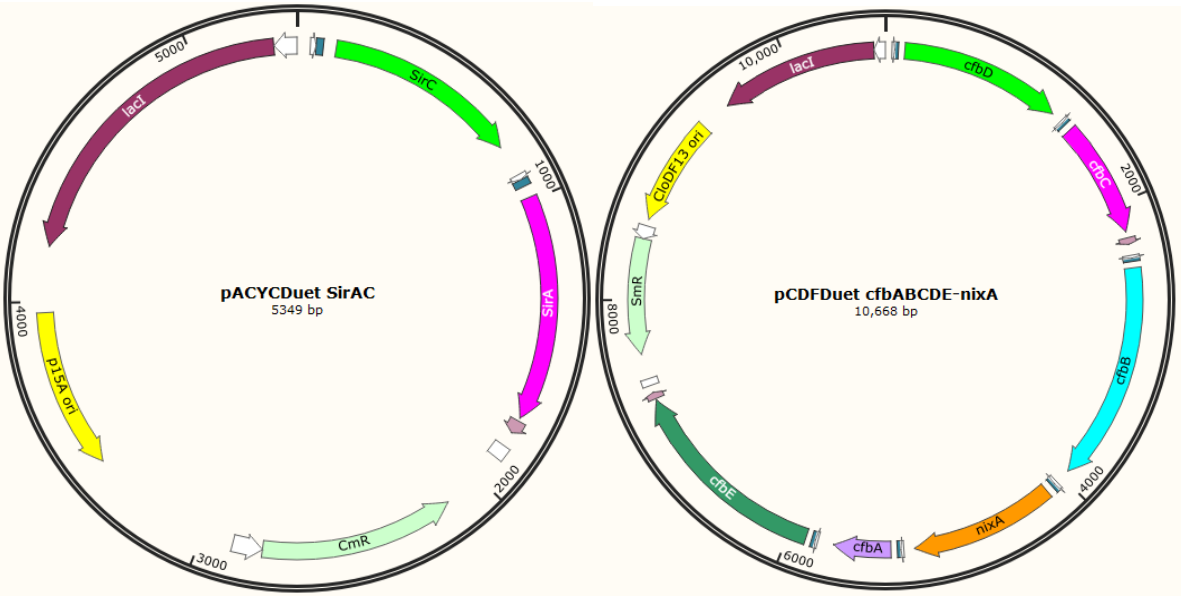
Nearly 1 billion tons of methane is produced by methanogenic archaea (methanogens) each year.[87] These microorganisms are widely found in wetlands and the digestive tracts of animals and produce methane as a metabolic byproduct.[20] The key methane-forming enzyme in methanogens is methyl-coenzyme M reductase (MCR), which uses a unique nickel-containing coenzyme F430 for activity.[61] MCR catalyzes the conversion of coenzyme B and methyl-coenzyme M to the mixed heterodisulfide CoB-S-S-CoM and methane.[61] Recently, MCR has been found in anaerobic methanotrophic archaea (ANME), which utilize reverse methanogenesis for the anaerobic oxidation of methane (AOM).[52] There is great potential for the use of ANME in natural gas-to-liquid fuel conversion strategies. Unfortunately, ANME have slow growth rates and have not been obtained in pure culture, which limit their use in large-scale conversion processes.[139] A solution is to engineer the methanogenesis pathway and its key enzyme, MCR, into a more robust microorganism for use in the bioconversion of natural gas. Currently, MCR

cannot be produced in an active form in a heterologous host, in part because the *in vivo* biosynthesis of coenzyme F430 is still a challenge.[87]

Our lab has identified the coenzyme F430 biosynthesis (*cfb*) genes and characterized the encoded enzymes *in vitro*.[87] They include a sirohydrochlorin cobaltochelatase homolog, CfbA, which catalyzes the specific Ni-chelation of sirohydrochlorin; a cobyrinic acid *a,c*-diamide synthase homolog, CfbB, that catalyzes the adenosine triphosphate (ATP)-dependent amidation of the *a*- and *c*-acetate side chains of Ni-sirohydrochlorin using glutamine as an ammonia source; a nitrogenase homolog, CfbCD, which converts Ni-sirohydrochlorin *a,c*-diamide to 15,17^3-seco-F430-17^3-acid; and a Mur-ligase homolog, CfbE, that carries out the ATP-dependent cyclization of the *g*-proprionate side chain to generate the carbocyclic F ring.[87] The *cfb* genes were cloned from *Methanosarcina acetivorans* and heterologously expressed in *Escherichia coli*. These recombinant enzymes together were shown to synthesize coenzyme F430 from sirohydrochlorin *in vitro.*[87]

For the biosynthesis of coenzyme F430 to proceed *in vivo*, we have to consider the substrates for the enzyme reactions and any required accessory factors or transporters. Thus, in addition to the *cfb* genes, we cloned *sirA* and *sirC* genes from the genomic DNA of *Methanosarcina acetivorans*, which convert uroporphyrinogen III (the last common precursor of all tetrapyrroles) to sirohydrochlorin. We synthesized the *nixA* gene from *Helicobacter pylori*, which is a nickel transporter that was shown to increase the *in vivo* availability of nickel in *E. coli*.[140] We also synthesized codon optimized ferredoxin (*fd*) and ferredoxin:NADP$^+$ reductase (*fnr*) genes from *Spinacia oleracea*. Spinach Fd and FNR was shown to support CfbCD catalysis *in vitro* using NADPH as an electron donor. Plasmids containing these genes were constructed using four compatible Duet vectors, each of which is designed for carrying two exogenous genes.

In order to insert more genes, a *Swa*I restriction enzyme site was introduced upstream of the multiple cloning sites (MCSs) and additional MCSs containing different genes were introduced. Finally, four plasmids containing multiple genes were constructed (Fig 2.1). The first plasmid, pCDFDuet:*cfbABCDE-nixA*, has all the coenzyme F430 biosynthesis (*cfb*) genes and the nickel transporter (*nixA*) gene. The second plasmid, pRSFDuet:*fd-fnr*, contains the spinach ferredoxin and ferredoxin:$NADP^+$ reductase system. The third plasmid, pACYCDuet:*sirAC*, encodes genes for the synthesis of the substrate for the Cfb enzymes. The fourth plasmid, pDB1282, contains the iron-sulfur cluster (*isc*) genes that can facilitate the production of *holo* CfbCD.[141] All the plasmids have been transformed into *E. coli* BL21 (DE3) cells, and enzymes were coexpressed under various culture conditions. The synthesized tetrapyrroles were extracted with methanol and analyzed using high-performance liquid chromatography (HPLC). Unfortunately, Ni-sirohydrochlorin *a,c*-diamide was the last biosynthetic intermediate detected, indicating that the pathway stopped at the CfbCD reaction.



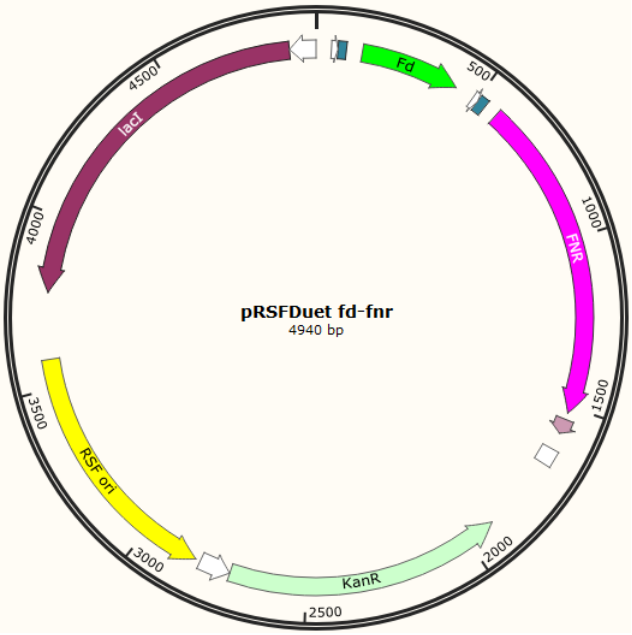


Figure 2.1 Plasmids constructed for the coexpression of *cfb* and accessory genes.

Two possibilities can be envisioned for why the *in vivo* biosynthesis of coenzyme F430 stops at the CfbCD reaction. First, a chemical reductant (i.e., sodium dithionite) is required for the CfbCD-catalyzed reaction *in vitro*, but *E. coli* cannot produce this compound and an appropriate physiological reductant is therefore needed for *in vivo* coenzyme F430 synthesis. The ferredoxin (Fd)/ferredoxin:$NADP^+$ reductase (FNR) system from spinach was shown to support CfbCD catalysis *in vitro*, but might not be functional in *E. coli* thus necessitating the need for an alternative reducing system. Second, CfbCD is an iron-sulfur enzyme, and its heterologous production requires the construction and insertion of iron-sulfur clusters. The pDB1282 plasmid, which includes all of the iron-sulfur cluster (*isc*) biosynthesis genes and has been shown to enhance iron-sulfur cluster incorporation into many recombinant proteins, was coexpressed with the *cfb* genes.[141] Although the activity of CfbCD from *M. acetivorans* can be detected *in vitro*, the enzyme complex and its iron-sulfur centers are unstable and readily degrade. Thus, the

stability and activity of CfbCD *in vivo* may also be low, so that 15,17^3-seco-F430-17^3-acid production could not be detected.

In this Chapter, we examined these possibilities. CfbCD and ferredoxin/ferredoxin:NADP$^+$ reductase from the coenzyme F430 cell line (which contains pCDFDuet:*cfbABCDE-nixA,* pACYCDuet:*sirAC,* pRSFDuet:*fd-fnr*, and pDB1282) were extracted and tested the activity *in vitro*. CfbCD and ferredoxin/ferredoxin:NADP$^+$ reductase were inactive in the conversion of Ni-sirohydrochlorin *a,c*-diamide to 15,17^3-seco-F430-17^3-acid. Thus, we replaced ferredoxin/ferredoxin:NADP$^+$ reductase with another physiological reducing system, pyruvate:ferredoxin oxidoreductase (PFOR). Meanwhile, the CfbCD from *M. acetivorans* were augmented with CfbCD from *Methanosarcina thermophila*.

Ferredoxin and flavodoxin are physiological electron donors for nitrogenase. CfbC and CfbD are homologous to the Fe protein (NifH) and the NifD subunit of the MoFe protein (NifDK) of nitrogenase, respectively.[142] PFOR catalyzes the thiamine pyrophosphate (TPP)-dependent oxidative decarboxylation of pyruvate to form acetyl-CoA (Fig 2.2).[143] Catalysis starts with nucleophilic attack on the 2-oxo carbon of pyruvate by C2 of TPP, which forms a lactyl-TPP intermediate.[143] The lactyl-TPP releases CO_2 and forms a hydroxyethylidene-TPP (HE-TPP) intermediate.[143] Then, this intermediate transfers an electron to a [4Fe-4S] cluster, which leads to a stable free radical intermediate.[143] The HE-TPP radical reacts with CoA and transfers another electron to a [4Fe-4S] cluster, forming the acetyl-CoA product and regenerating free TPP.[143] The reduced [4Fe-4S] clusters of PFOR then serve as an electron donor to ferredoxin.[143] Thus, we cloned the PFOR and ferredoxin genes from the genome of *M. acetivorans*.

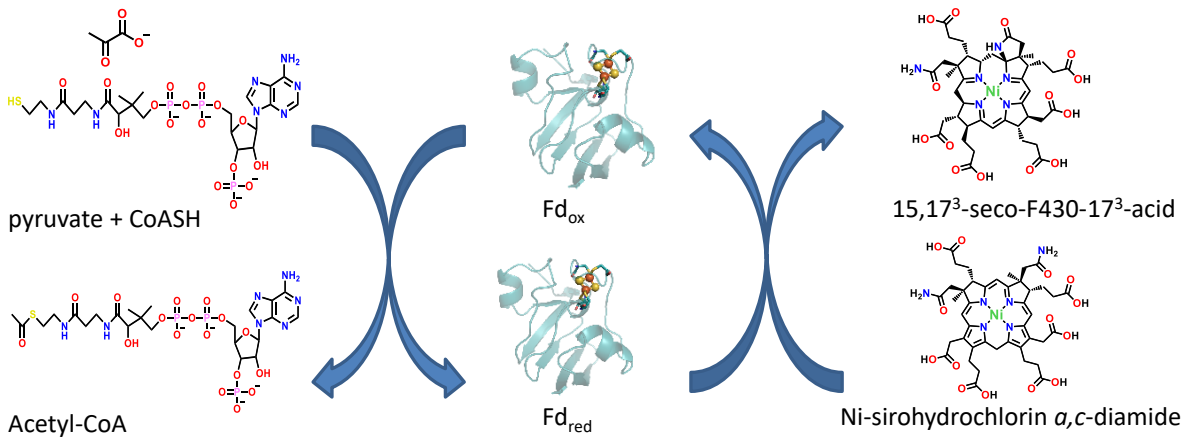


Figure 2.2 PFOR reducing system for supporting the CfbCD reaction.

M. thermophila is a thermophilic, methane-producing archaea.[144] It has been reported that *M. thermophila* possesses regulatory systems that enable it to adapt to various harsh environments.[144] Thermophilic versions of enzymes often display better stability and are more robust than their mesophilic counterparts. For CfbCD in particular, CfbC tends to precipitate readily. CfbC from *M. acetivorans* has an isoelectric point (pI) of 8.32, whereas CfbC from *M. thermophila* has a pI of 6.48, which is closer to that of CfbD, and thus might be less prone to precipitation. Thus, we cloned the *cfbCD* genes from *M. thermophila*. Here, the effects of switching to the PFOR reducing system and the thermophilic version of CfbCD on the synthesis of coenzyme F430 in *E. coli* is explored.

In this Chapter, we describe the construction of a plasmid, pRSFDuet:*Mt_cfbCD-porABDG*, containing a native physiological reducing system (PFOR) and CfbCD from *M. thermophila*. Two cell lines were prepared combining this plasmid with the following: 1) pCDFDuet:*cfbABCDE-nixA*, pACYCDuet:*sirAC*, and pDB1282, 2) pCDFDuet:*cfbABCDE-nixA*, pACYCDuet:*sirAC*, and pETDuet:*mm5-mm15-mm7-mm17-mm6-mm3-mm2-A2*. The plasmid containing the *mm* genes was constructed by Shadi Yavari. It encodes many methanogenesis marker proteins of unknown function, but which have been implicated in MCR

assembly/activation and could play a role in the biosynthesis/insertion of coenzyme F430. We tested for coenzyme F430 biosynthesis activity of these cell lines under a variety of growth conditions, including different growth times/temperatures after induction, addition of varying amounts of supplementary chemicals, different strains of *E. coli*, and anaerobic growth conditions.

2.2 Materials and Methods

2.2.1 Materials

Calcium chloride, arabinose, sodium sulfide, ferric chloride, cysteine, dithiothreitol, ferrous sulfate, nickel chloride, sodium dithionite, glutamine, *S*-adenosyl-L-methionine (SAM), and buffer components were from VWR Chemicals BDH® (Radnor, PA). 5-Aminolevulinic acid (5-ALA) and porphobilinogen (PBG) were from Frontier Scientific Inc. (Logan, UT). Isopropyl β-D-1-thiogalactopyranoside (IPTG) was from Teknova Inc. (Hollister, CA). Thrombin from bovine plasma and adenosine triphosphate (ATP) were purchased from Sigma-Aldrich (St. Louis, MO). Nicotinamide adenine dinucleotide (NAD^+) was from Amresco (Albany, NY). Reduced nicotinamide adenine dinucleotide phosphate (NADPH) was from Enzo Life Sciences Inc. Luria-Bertani (LB) media was from Becton, Dickinson and Company (BD) (Franklin Lakes, NJ), while buffer components were from VWR Chemicals BDH® (Radnor, PA). Primers were ordered from Sigma-Aldrich and *Methanosarcina acetivorans* C2A genomic DNA was purchased from DSMZ (Braunschweig, Germany). Phusion High-Fidelity PCR Kit, Gibson Assembly Master Mix, and restriction endonucleases were from New England Biolabs (Ipswich, MA). Gel extraction, PCR cleanup, and mini prep kits were purchased from Omega Bio-Tek (Norcross, GA). Profinity IMAC resin was obtained from Bio-Rad (Hercules, CA).

2.2.2 Construction of the plasmid containing the PFOR genes

The *porA* (MA0032), *porB* (MA0031), *porG* (MA0034), and *porD* (MA0033) genes were amplified by polymerase chain reaction (PCR) using designed primers for the genomic DNA of *Methanosarcina acetivorans* C2A. Primers were synthesized by Sigma-Aldrich and their sequences are provided in Table 2.1. We used four compatible Duet vectors, pRSFDuet-1, pETDuet-1, pCDFDuet-1, pACYCDuet-1, each of which are designed for carrying two exogenous genes, to coexpress more enzymes in *E. coli*. To construct the pRSFDuet:*porAB* plasmid, an empty pRSFDuet vector was amplified using pRSF *porA* primers. To construct the pETDuet:*porGD* plasmid, an empty pETDuet vector was amplified using pETD *porG* primers. All PCR products were purified via a Lonza FlashGel™ DNA Cassette (Basel, Switzerland) and recovered using an Omega Bio-tek E.Z.N.A.® Gel Extraction Kit (Norcross, GA). The *porA* gene was inserted into the first multiple cloning site (MCS-I) of pRSFDuet-1 using the Gibson assembly method. The *porG* gene was inserted into the MCS-I of pETDuet-1. The ligated mixture was then transformed into *E. coli* NEB 5α cells and plated on LB agar plates containing 50 μg/mL kanamycin for pRSFDuet and 100 μg/mL ampicillin for pETDuet. Colonies were picked, grown in liquid LB medium, and plasmid preparations were made using the Omega Bio-tek E.Z.N.A.® Plasmid Mini Kit. The plasmids pRSFDuet:*porA* and pETDuet:*porG* were sequence verified by Eurofins Genomics. Then, the plasmid pRSFDuet:*porA* was amplified using pRSF *porB* primers, and the plasmid pETDuet:*porG* was amplified using pETD *porD* primers. After PCR products were purified and recovered, the *porB* gene was inserted into the MCS-II of pRSFDuet:*porA* using the Gibson assembly method. The *porG* gene was inserted into the MCS-II of pETDuet:*porG*. The ligated mixtures were then transformed into *E. coli* NEB 5α cells and plated on LB agar plates containing 50 μg/mL kanamycin for pRSFDuet and 100

μg/mL ampicillin for pETDuet. Colonies were picked, grown in liquid LB medium, and plasmid preparations were made. The plasmids pRSFDuet:*porAB* and pETDuet:*porGD* were sequence verified by Eurofins Genomics.

To construct the pRSFDuet:*porABDG* plasmid, a *Swa*I restriction enzyme site was introduced approximately 100 bp upstream of the MCS-I of each Duet vector followed site-directed mutagenesis protocol. The pRSFDuet:*porAB* plasmid was treated with the SwaI restriction enzyme to linearize the plasmid.[145] A pair of specific SwaI Overlap primers were used to amplify the *porGD* fragments from pETDuet:*porGD.* After the PCR product and linearized vector were purified and recovered, the *porGD* fragment and linearized pRSFDuet:*porAB* plasmid were combined using the Gibson Assembly method. The constructed four gene plasmid pRSFDuet:*porABDG* was sent to Eurofins Genomics for sequence verification.

Table 2.1 Primers used for plasmid construction for the PFOR system and CfbCD from *M. thermophila.*

Gene	Primers
porA	Forward: 5'-GGAGATATACCatgccgctcaattcagc-3' Reverse: 5'-CGCAAGCTtcatagcaactcctccttgac-3'
pRSF *porA*	Forward: 5'-ttgctatgaAGCTTGCGGCCG-3' Reverse: 5'-cggcatGGTATATCTCCTTATTAAAGTTAAAC-3'
porB	Forward: 5'-AAGAAGGAGATATACATatgagtaaacccgcacc-3' Reverse: 5'-AGGGTACCttactgaagtccgtagtgttttatg-3'
pRSF *porB*	Forward: 5'-gacttcagtaaGGTACCCTCGAGTCTG-3' Reverse: 5'-gtttactcatATGTATATCTCCTTCTTATACTTAAC-3'
porD	Forward: 5'-GTATAAGAAGGAGATATACATatgaaggtcccaattggag-3' Reverse: 5'-CCAGACTCGAGttatttctcctccagaatcatttc-3'
pETD *porD*	Forward: 5'-gagaaataaCTCGAGTCTGGTAAAGAAACC-3' Reverse: 5'-cttcatATGTATATCTCCTTCTTATACTTAACTAATATAC-3'
porG	Forward: 5'-GGAGATATACCatgaaggaaatcagaatacacggtc-3' Reverse: 5'-CGGATCCtcatgcttcttccccccctg-3'
pETD *porG*	Forward: 5'-gaagcatgaGGATCCGAATTCGAGC-3' Reverse: 5'-atttccttcatGGTATATCTCCTTCTTAAAGTTAAAC-3'
Mt_cfbC	Forward: 5'-ACTTTAATAAGGAGATATACCatgaaaaaacagaaaatcatcgcg-3' Reverse: 5'-GCAAGCTtcatttggtcagttcacgc-3'

pRSF *Mt_cfbC*	Forward: 5'-gaccaaatgaAGCTTGCGGCCGC-3' Reverse: 5'-catGGTATATCTCCTTATTAAAGTTAAACAAAATTATTTC-3'
Mt_cfbD	Forward: 5'-AGGAGATATACATatgaccggtaaagaaatctctatc-3' Reverse: 5'-GGGTACCtcacgcttctttcgcaac-3'
pRSF *Mt_cfbD*	Forward: 5'-agcgtgaGGTACCCTCGAGTCTGG-3' Reverse: 5'-cggtcatATGTATATCTCCTTCTTATACTTAACTAATATAC-3'

2.2.3 Construction of the plasmid containing the CfbCD and PFOR genes

The *cfbC* (MSTHT2451), and *cfbD* (MSTHT2450) genes were amplified by PCR from the genomic DNA of *Methanosarcina thermophila* TM-1, a gift from Dr. Cathy Drennan (Massachusetts Institute of Technology, Cambridge, MA). To construct the pRSFDuet:*Mt_cfbCD* plasmid, an empty pRSFDuet vector was amplified using pRSF *Mt_cfbC* primers. After PCR products were purified and recovered, the *Mt_cfbC* gene was inserted into the MCS-I of pRSFDuet-1 using the Gibson Assembly method. The ligated mixture was then transformed into *E. coli* NEB 5α cells and plated on LB agar plates. Colonies were picked, grown in liquid LB medium, and plasmid preparations were made using the Omega Bio-tek E.Z.N.A.® Plasmid Mini Kit. The plasmid pRSFDuet:*Mt_cfbC* was sequence verified by Eurofins Genomics. Then, the plasmid pRSFDuet:*Mt_cfbC* was amplified using pRSF *Mt_cfbD* primers. After the PCR product was purified and recovered, the *Mt_cfbD* gene was inserted into MCS-II of pRSFDuet:*Mt_cfbC* using the Gibson assembly method. The ligated mixture was then transformed into *E. coli* NEB 5α cells and plated on LB agar plates containing 50 µg/mL kanamycin. Colonies were picked, grown in liquid LB medium, and plasmid preparations were made. The plasmid pRSFDuet:*Mt_cfbCD* was sequence verified by Eurofins Genomics.

To construct the pRSFDuet:*Mt_cfbCD-porABDG* plasmid, the pRSFDuet:*porABDG* plasmid was treated with the SwaI restriction enzyme to linearize it.[145] A pair of specific SwaI Overlap primers were used to amplify the *Mt_cfbCD* fragment from pRSFDuet:*Mt_cfbCD.* After

the PCR product and linearized vector were purified and recovered, the *Mt_cfbCD* fragment and linearized pRSFDuet:*porABDG* plasmid were combined using the Gibson Assembly method. The constructed six gene plasmid pRSFDuet:*Mt_cfbCD-porABDG* was sent to Eurofins Genomics for sequence verification (Fig 2.3).

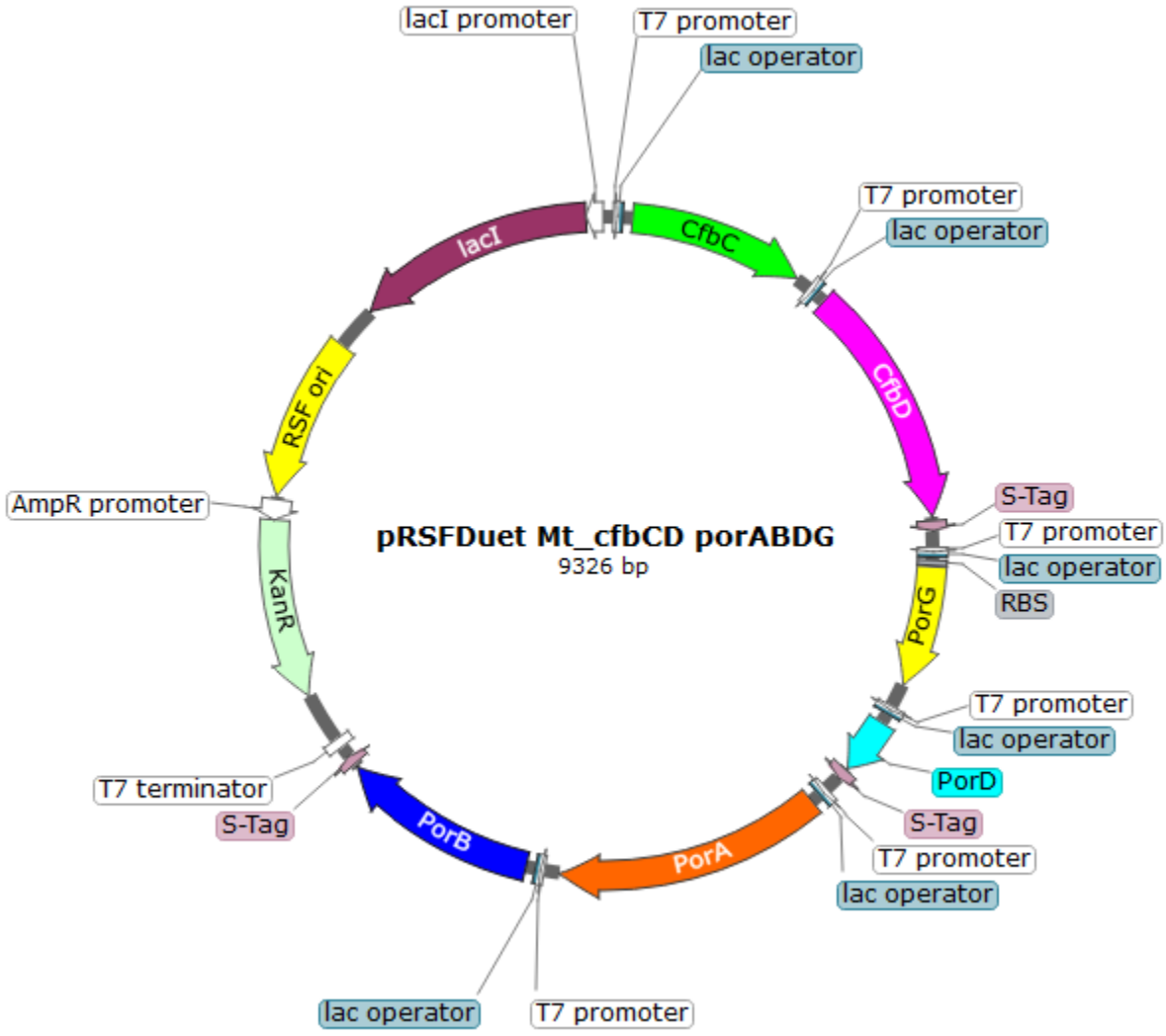


Figure 2.3 Plasmid construct for the PFOR system and CfbCD from *M. thermophila.*

2.2.4 Cell line construction

We constructed two *E. coli* cell lines containing either the pRSFDuet:*Mt_cfbCD-porABDG,* pCDFDuet:*cfbABCDE-nixA,* pACYCDuet:*sirAC,* and pDB1282 plasmids or the pRSFDuet:*Mt_cfbCD-porABDG,* pCDFDuet:*cfbABCDE-nixA,* pACYCDuet:*sirAC,* and pETDuet:*mm5-mm15-mm7-mm17-mm6-mm3-mm2-A2* plasmids. The first plasmid was transformed into *E. coli* BL21(DE3) cells or C41(DE3) (New England Labs) that were used as the expression host for *in vivo* coenzyme F430 biosynthesis. Successful transformation was confirmed using colony PCR. The cells were made competent to take up the other three recombinant plasmids using the following chemical competency protocol. First, the cells were cultured in 5 mL LB with the proper antibiotic(s) overnight. The cultures were transferred to 20 mL LB in a 1:100 ratio. When the OD_{600} was between 0.2 to 0.5, the cultures were chilled on ice for 20 min and then centrifuged for 3 min in a microcentrifuge at 5500 rcf to collect the cell pellets. The 5 mL sterilized 0.1 M $CaCl_2$ solution was prepared and used to resuspend the cells. The resuspension mixture was chilled on ice for 20 min and then centrifuged again for 3 min at 5500 rcf. The pellets were resuspended again and chilled on ice for another 20 min. Finally, the cells were pelleted and resuspended in a 2.5 mL sterilized 0.1 M $CaCl_2$ solution containing 15% glycerol. The resulting chemically competent *E. coli* BL21(DE3) or C41(DE3) cells with the first plasmid were separated into 100 μL aliquots for transforming the second plasmid. We then repeated this procedure to sequentially transform the four plasmids into one *E. coli* BL21(DE3) or C41(DE3) strains. Each step was confirmed with antibiotic resistance and PCR verification.

2.2.5 Expression and purification of Cfb enzymes

E. coli BL21(DE3) cells containing either pET-28b(+):*hemC,* pET-28b(+):*hemD,* pET-28b(+):*sirC,* pET-28b(+):*sirA,* pET-28b(+):*cfbA,* pET-28b(+):*cfbB,* pETSUMO:*cfbC,* and

pETSUMO:*cfbD* were grown in overnight cultures (2 mL of LB with 50 μg/mL kanamycin) that were used to inoculate 6 × 1 L of LB (50 mg/L kanamycin), separately. The cultures were incubated with shaking at 37 °C until the OD_{600} reached 0.6. The cultures were then induced with isopropyl β-D-1-thiogalactopyranoside (IPTG) at a final concentration of 0.1 mM and incubated overnight. The cells were centrifuged at 15,900 × g for 45 min. The supernatants were decanted, and the cells were stored in -80 °C prior to purification.

For His_6-tagged HemD, SirA, SirC, CfbA, and CfbB purification, the cells were resuspended in 50 mM sodium phosphate buffer (pH 8.0) containing 300 mM NaCl and 5 mM imidazole. The cells were then sonicated and centrifuged at 15,900 × g for 1 hour. The supernatants were loaded onto a Bio-Rad Econo-Pac column loaded with nickel-charged Profinity IMAC resin. The columns were washed with 50 mM sodium phosphate buffer (pH 8.0) containing 300 mM NaCl and 5 mM imidazole. The proteins were then eluted from the column by adding 50 mM sodium phosphate buffer (pH 8.0) containing 300 mM NaCl and 500 mM imidazole in 1 mL fractions. The protein containing fractions were combined and spin-concentrated/buffer exchanged into 100 mM Tris buffer (pH 8.0).

For His_6-tagged HemC, SUMO-tagged CfbC, and SUMO-tagged CfbD purification, the cells were transferred into a glovebox and resuspended in degassed 50 mM sodium phosphate buffer (pH 8.0) containing 300 mM NaCl and 5 mM imidazole. The cells were then sonicated and centrifuged at 30,000 × g for 1 hour. The supernatants were loaded onto a Bio-Rad Econo-Pac column loaded with nickel-charged Profinity IMAC resin. The columns were washed with 50 mM sodium phosphate buffer (pH 8.0) containing 300 mM NaCl and 5 mM imidazole. The proteins were then eluted from the column by adding 50 mM sodium phosphate buffer (pH 8.0) containing 300 mM NaCl and 500 mM imidazole in 1 mL fractions. The protein containing

fractions were combined and spin-concentrated/buffer exchanged into 100 mM Tris buffer (pH 8.0) containing 200 mM NaCl.

The purified His_6-tagged CfbA was incubation with thrombin (1:1000 [w:w]) overnight at room temperature to remove the N-terminal His-tag. After incubation, the cleaved protein was passed through a benzamidine/Ni-NTA column. The eluted protein was then spin-concentrated/buffer exchanged into 100 mM Tris buffer (pH 8.0) containing 15 % glycerol and separated into 100 µL aliquots.

The purified SUMO-tagged CfbC and CfbD were incubated with SUMO protease (1:1 [w:w]) overnight at room temperature to remove the N-terminal SUMO-tag in the glovebox. After incubation, the cleaved proteins were passed through a Ni-NTA column to remove the SUMO-tag and SUMO protease (which contained a N-terminal His_6-tag). The eluted proteins were then spin-concentrated/buffer exchanged into 100 mM Tris buffer (pH 8.0) containing 200 mM NaCl and 15 % glycerol.

The iron-sulfur clusters of cleaved CfbC and CfbD were reconstituted in the glovebox. Protein solutions and reagents were incubated under anaerobic conditions overnight before using. To reconstitute 0.2 mM CfbC and CfbD, the proteins were incubated with 3 mM dithiothreitol (DTT) and 3 mM Na_2S at room temperature for 20 min. Next, the mixture was incubated with 3 mM $FeCl_3$ for 20 min. This step was repeated three times and left overnight. The resulting reconstitution mixture was centrifuged at 7,000 × g for 20 min and then loaded onto a GE PD-10 Sephadex G-25 M column that was preequilibrated with 100 mM Tris buffer (pH 8.0). The reconstituted proteins were eluted with 100 mM Tris buffer (pH 8.0) and colored fractions were collected. The proteins were then spin-concentrated/buffer exchanged into 100 mM Tris buffer (pH 8.0) containing 200 mM NaCl and 15 % glycerol and separated into 100 µL aliquots.

2.2.6 Aerobically coexpression of F430 genes

The method for the coexpression of the F430 genes contained three stages. The cell line including pRSFDuet:*Mt_cfbCD porABDG,* pCDFDuet:*cfbABCDE-nixA,* pACYCDuet:*sirAC,* and pDB1282 plasmids were grown in overnight cultures (2 mL of LB with 50 μg/mL kanamycin) that were used to inoculate 6 × 1 L of Terrific Broth (TB) (50 mg/L kanamycin). The cultures were incubated with shaking at 37 °C until the OD_{600} reached 0.3 and then 3 g of arabinose, 2 mM ferrous sulfate, and 2 mM cysteine were added. The cultures were allowed to continue shaking for two hours and then 200 μM nickel chloride, 10 μM 5-aminolevulinic acid (5-ALA), 1 mM SAM, and 0.4 mM IPTG were added. The temperature was decreased to 18 °C. After 16 hours, the cultures were incubated with shaking at 37 °C for an additional 4 hours.

The cell line including pRSFDuet:*Mt_cfbCD porABDG,* pCDFDuet:*cfbABCDE-nixA,* pACYCDuet:*sirAC,* and pETDuet:*mm5-mm15-mm7-mm17-mm6-mm3-mm2-A2* plasmids were grown in overnight cultures (2 mL of LB with 50 μg/mL kanamycin) that were used to inoculate 6 × 1 L of Terrific Broth (TB) (50 mg/L kanamycin). The cultures were incubated with shaking at 37 °C until the OD_{600} reached 0.6 and then 200 μM nickel chloride, 10 μM 5-aminolevulinic acid (5-ALA), 1 mM SAM and 0.4 mM IPTG were added. The temperature was decreased to 18 °C. After 16 hours, the cultures were incubated with shaking at 37 °C for additional 4 hours.

The cells were centrifuged at 15,900 × g for 45 min. The supernatants were decanted, and the cells were transferred into a glovebox. The cells were resuspended in 50 mM sodium phosphate buffer (pH 8.0) containing 300 mM NaCl and 5 mM imidazole. The cells were then sonicated and centrifuged at 30,000 × g for 1 hour. The supernatants were then spin-concentrated/buffer exchanged into 100 mM Tris buffer (pH 8.0) containing 15 % glycerol.

2.2.7 Anaerobically coexpression of F430 genes

The cell line including pRSFDuet:*Mt_cfbCD porABDG,* pCDFDuet:*cfbABCDE-nixA,* pACYCDuet:*sirAC,* and pETDuet:*mm5-mm15-mm7-mm17-mm6-mm3-mm2-A2* plasmids were grown in overnight cultures (10 mL of LB with 50 μg/mL kanamycin) that were used to inoculate 1 L of degassed Terrific Broth (TB) (50 mg/L kanamycin). The cultures were incubated at 37 °C until the OD_{600} reached 0.3 and then 200 μM nickel chloride, 10 μM 5-aminolevulinic acid (5-ALA), 1 mM SAM and 0.4 mM IPTG were added. The cultures were incubated at 37 °C for additional 24 hours.

The cells were centrifuged at 15,900 × g for 45 min. The supernatants were decanted, and the cells were transferred into a glovebox. The products were extracted by adding 5 mL methanol to the cell and centrifuged at 7,000 × g for 30 min. The supernatants were analyzed using HPLC.

2.2.8 CfbA activity assay

The activity of CfbA in the engineered cell lines was examined by incubating 2.5 μM HemC, 2.5 μM HemD, 2.5 μM SirA, 2.5 μM SirC, 200 μL cell free extract, 0.88 mM porphobilinogen (PBG), 2 mM SAM, 1 mM $NAD(P)^+$, 4 mM $MgCl_2$, and 200 μM $NiCl_2$ in 1 mL of 100 mM Tris-HCl, pH 8.0 buffer. The reaction was performed overnight at 37 °C in the glovebox. The products were extracted by adding 100 μL methanol to 100 μL of assay mixture and centrifuged at 7,000 × g for 30 min. The supernatants were analyzed using HPLC (*vide infra*).

2.2.9 CfbB activity assay

The activity of CfbB in the engineered cell lines was examined by adding 4 mM glutamine and 5 mM ATP to the CfbA activity assay. The reaction was performed overnight at 37 °C in the glovebox. The products were extracted by adding 100 μL methanol to 100 μL of assay mixture

and centrifuged at 7,000 × g for 30 min. The supernatants were analyzed using HPLC (*vide infra*).

2.2.10 CfbCD activity assay

The activity of CfbCD in the engineered cell lineswas examined by adding 3 mM sodium dithionite and 5 mM ATP to the CfbB activity assay. The reaction was performed overnight at 37 °C in the glovebox. The products were extracted by adding 100 μL methanol to 100 μL of assay mixture and centrifuged at 7,000 × g for 30 min. The supernatants were analyzed using HPLC (*vide infra*).

2.2.11 Fd/FNR activity assay

The activity of Fd/FNR in the engineered cell lines was examined by adding 10 μM CfbC, 10 μM CfbD, 5 mM NADPH and 5 mM ATP to the CfbB activity assay. The reaction was performed overnight at 37 °C in the glovebox. The products were extracted by adding 100 μL methanol to 100 μL of assay mixture and centrifuged at 7,000 × g for 30 min. The supernatants were analyzed using HPLC (*vide infra*).

2.2.12 MALDI-TOF MS

The proteins Mt_CfbC, Mt_CfbD, PorA, PorB, PorG, and PorD were coexpressed and extracted from *E. coli* BL21(DE3) cells harboring the pRSFDuet:*Mt_cfbCD-porABDG*. To confirm their presence in the cell line, the proteins were purified using an acetone precipitation protocol and analyzed by matrix-assisted laser desorption/ionization time-off-light mass spectrometry (MALDI-TOF MS) as follows. The proteins were incubated with 4 volumes of pre-cooled acetone overnight at -20 °C. The protein pellet was centrifuged at 7,000 × g for 20 min. The supernatant was decanted, and the protein pellet was transferred in a fume hood to evaporate residual acetone. The resulting protein preparation was sent to the Proteomics and Mass

Spectrometry Facility at the University of Gerogia (UGA) for MS analysis. A trypsin or pepsin digestion was performed on the proteins, and the resultant peptide fragments were analyzed by MALDI-TOF MS for protein identification. MALDI-TOF MS analysis was performed on a Bruker Autoflex MALDI spectrometer. The resultant peptide fragments were also analyzed by liquid chromatography mass spectrometry (LC-MS) for further peptide verification. LC-MS analysis was performed on an Orbitrap Elite system with an electrospray ion source and an interchangeable nano spray source. The data was submitted to the publicly available Matrix Science software package. The MASCOT Peptide Mass Fingerprint server, MASCOT Sequence Query server and the University of Georgia Proteomics and Mass Spectrometry server were used for data analysis to identify proteins.

2.2.13 HPLC and LC-MS analysis

HPLC analysis was performed on an Agilent 1260 Infinity Quaternary LC System with a Diode Array Detector (DAD) VL+. The column was an Agilent Poroshell 120 EC-C18 (4.6 × 150 mm, 2.7 μm) reverse phase column. The Agilent OpenLAB ChemStation Edition software was used for data analysis. The mobile phase consisted of water with 0.5 % formic acid (solvent A) and acetonitrile with 0.5 % formic acid (solvent B). The chromatographic method for characterization of the coenzyme F430 biosynthetic reactions consisted of the following gradient: 0 % B for 2 min, linear gradient to 20 % B over 3 min, 20 % B for 5 min, linear gradient to 25 % B over 5 min, 25 % B for 5 min, linear gradient to 30 % B over 5 min, linear gradient to 100 % B over 2 min, and 100 % B for 3 min. A flow rate of 1.0 mL/min was used and the trace was acquired using detection at 400 nm.

MS analysis of the assays was performed at the Auburn University Department of Chemistry and Biochemistry Mass Spectrometry Center on an ultra-performance LC system

(ACQUITY, Waters Corp., USA) coupled with a quadrupole time-of-flight mass spectrometer (Q-TOF Premier, Waters Corp., USA) with electrospray ionization (ESI) in positive mode using the MassLynx software (v4.1). Injections of the samples were made directly into the mass spectrometer or onto a C4 column (Aeris™ 3.6 μm C4 200 Å, 50 × 2.1 mm, Phenomenex) with a 300 μL/min flow rate of mobile phase solution A (0.1 % formic acid in 95 % water and 5 % acetonitrile) and solution B (0.1 % formic acid in 95 % acetonitrile and 5 % water) using the following gradient: 0 % B for 2 min, linear gradient to 100 % B over 11 min, 100 % B for 1 min, linear gradient to 0 % B over 1 min, and 0 % B for 3 min. The column temperature was held at 40 °C and the UV detector monitored absorption at 420 nm. The capillary voltage was set at 3.1 kV, the sample cone voltage was 15 V, and the extraction cone was 4.0 V. The source and desolvation temperature were maintained at 95 and 300 °C, respectively, with the desolvation gas flow set to 600 L/h. The mass spectrometry scan was 0.5 s long from 50 to 2,000 *m*/*z* with a 0.02 s interscan delay using the centroid data format.

2.3 Results

2.3.1 Activity assay

The proteins SirA, SirC, CfbA, CfbB, CfbC, CfbD, CfbE, NixA, Fd, FNR, and Isc were coexpressed in *E. coli* using the three-stage expression method described in section 2.2.6. To test the activity of the proteins in the cell free extract, a CfbA assay was performed and analyzed using HPLC (Fig 2.4). No sirohydrochlorin was detected in the assay and all of the porphobilinogen (PBG) was converted to Ni-sirohydrochlorin, which has a retention time of 23 min, indicating that the CfbA protein in the extract retains activity (Fig 2.4). Ni-sirohydrochlorin *a,c*-diamide, which has a retention time of 19 min, was also detected in the assay. However, Ni-

sirohydrochlorin *a,c*-diamide was also found in the control without sirohydrochlorin, indicating that it is present in the cell free extract itself (Fig 2.5).

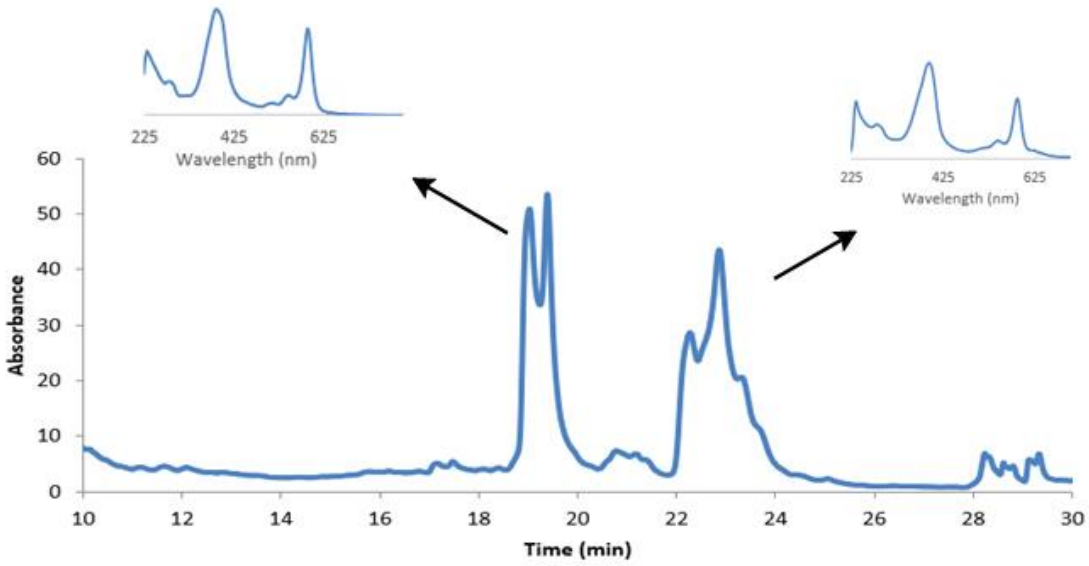


Figure 2.4 HPLC and UV-visible spectroscopic analysis of the CfbA reaction using cell free extracts of a strain of *E. coli* harboring the pCDFDuet:*cfbABCDE-nixA*, pACYCDuet:*sirAC*, pRSFDuet:*fd-fnr*, and pDB1282 plasmids. The reaction contained 2.5 μM HemC, 2.5 μM HemD, 2.5 μM SirA, 2.5 μM SirC, 200 μL protein extract, 0.88 mM porphobilinogen (PBG), 2 mM SAM, 1 mM $NAD(P)^+$, 4 mM $MgCl_2$, and 200 μM $NiCl_2$ in 1 mL of 100 mM Tris-HCl, pH 8.0 buffer.

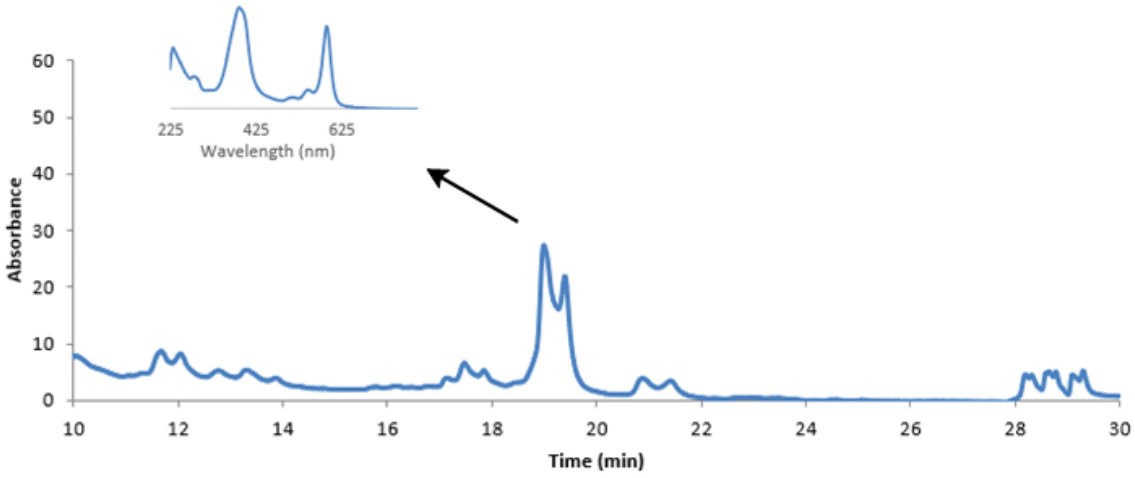


Figure 2.5 HPLC and UV-visible spectroscopic analysis of a control reaction using cell free extracts of a strain of *E. coli* harboring the pCDFDuet:*cfbABCDE-nixA*, pACYCDuet:*sirAC*, pRSFDuet:*fd-fnr*, and pDB1282 plasmids. The reaction contained 200 μL protein extracts in 1 mL of 100 mM Tris-HCl, pH 8.0 buffer.

A CfbB assay was performed to comfirm the activity of the next enzyme in the coenzyme F430 biosynthetic pathway. All the Ni-sirohydrochlorin was converted to Ni-sirohydrochlorin *a,c*-diamide, confirming the activity of CfbB in the cell free extracts (Fig 2.6).

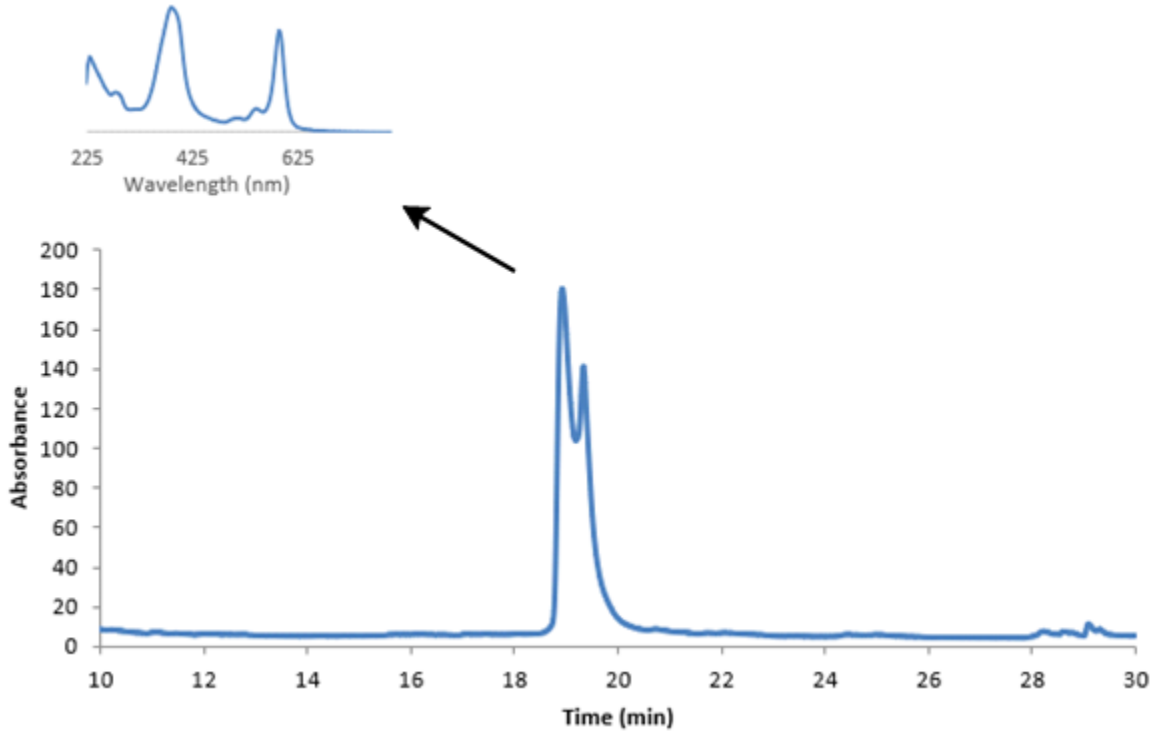


Figure 2.6 HPLC and UV-visible spectroscopic analysis of the CfbB reaction using cell free extracts of a strain of *E. coli* harboring the pCDFDuet:*cfbABCDE-nixA*, pACYCDuet:*sirAC*, pRSFDuet:*fd-fnr*, and pDB1282 plasmids. The reaction contained 4 mM glutamine and 5 mM ATP in addition to the components of the CfbA activity assay.

The CfbCD assay was then performed using sodium dithionite to provide electrons. However, no coenzyme F430 or 15,17^3-seco-F430-17^3-acid was detected, which has a retention time of 8 min, suggesting that CfbCD was inactive (Fig 2.7). There was a broad peak around 7 min, which was flavin based on the retention time and UV-visible absorption spectrum (Fig 2.7). The activity of Fd/FNR in the cell free extract was assayed using additional CfbCD and NADPH. However, no coenzyme F430 or 15,17^3-seco-F430-17^3-acid was detected (the HPLC and UV-visible spectra were identical to the CfbCD assay). From these results, both CfbCD and Fd/FNR

appear to be inactive in the *E. coli* cell line containing the pCDFDuet:*cfbABCDE-nixA*, pACYCDuet:*sirAC*, pRSFDuet:*fd-fnr*, and pDB1282 plasmids.

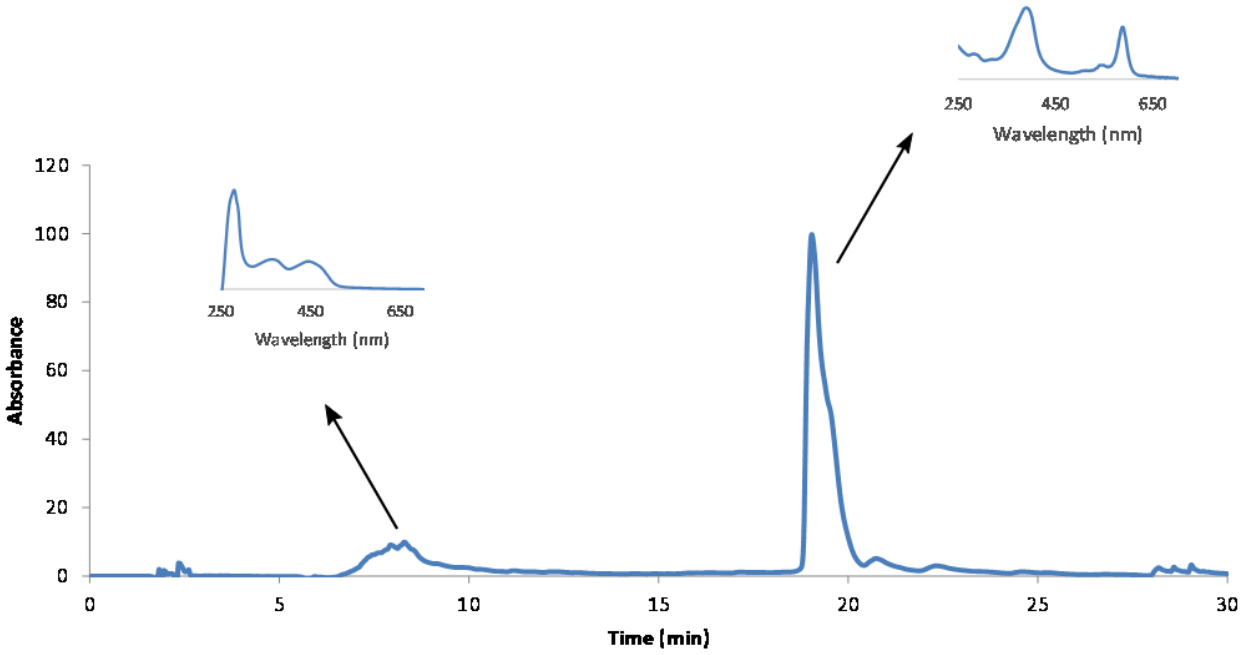


Figure 2.7 HPLC and UV-visisble spectroscopic analysis of the CfbCD reaction using cell free extracts of a strain of *E. coli* harboring the pCDFDuet:*cfbABCDE-nixA*, pACYCDuet:*sirAC*, pRSFDuet:*fd-fnr*, and pDB1282 plasmids. The reaction contained 3 mM sodium dithionite, and 5 mM ATP in the CfbB activity assay.

2.3.2 MALDI-TOF MS of Cfb and DPOR genes

To confirm whether the Cfb proteins were expressed successfully in the cell line, the cell free extract was analyzed using MALDI-TOF MS. From the results (Table 2.2), all the proteins were detected with high coverage and score using the MASCOT server. Thus, the lack of detectable CfbCD activity was not due to a lack of protein expression. Since thermophilic versions of enzymes often display better stability and are more robust than their mesophilic counterparts, augmentation of CfbCD from *M. acetivorans* with an additional copy from *M. thermophila* was pursued to ameliorate the issue with inactive CfbCD.

Table 2.2: MASCOT search results based on the MALDI-TOF MS spectrum of cell free extracts of a strain of *E. coli* harboring the pCDFDuet:*cfbABCDE-nixA* plasmid.

Accession	Description	Score	Coverage	# Unique Peptides	# Peptides	# PSMs	# AAs
AAM06986.1	cobalamin biosynthesis protein	1307.56	80.00	9	12	45	130
AAM06982.1	nitrogenase (iron protein)	1058.62	60.00	18	25	50	265
AAM06983.1	nitrogenase-related protein	653.34	36.49	9	23	29	370
AAM06981.1	cobyrinic acid a,c-diamide synthase	312.51	26.56	9	12	13	497
AAM06985.1	UDP-N-acetylmuramoylalanine-D-glutamate ligase	307.69	24.15	4	11	13	472

The ferredoxin (Fd)/ferredoxin:$NADP^+$ reductase (FNR) system was from spinach, which might not function in *E. coli*. To find the native physiological reductant for CfbCD, we searched for pyruvate:ferredoxin oxidoreductase (PFOR) and ferredoxin homologs in the genome of *M. acetivorans*, which support catalysis by the CfbCD homolog nitrogenase using pyruvate as a reductant. The *porA*, *porB*, and *porG* genes encode the three PFOR subunits, and *porD* encodes a ferredoxin. The *porA*, *porB*, *porG*, and *porD* genes were cloned into a pRSFDuet vector and coexpressed in *E. coli*. After induction, the cell free extract was analyzed using MALDI-TOF MS to ensure the proteins were expressed successfully. From the results (Table 2.3), all four proteins were detected with a high coverage and score using the MASCOT server.

Table 2.3: MASCOT search results based on the MALDI-TOF MS spectrum of cell free extracts of a strain of *E. coli* harboring the pRSFDuet:*porABDG* plasmid.

Accession	Description	Score	Coverage	# Unique Peptides	# Peptides	# PSMs	# AAs
AAM03488.1	pyruvate synthase, subunit gamma	2669.85	65.93	12	12	45	182
AAM03486.1	pyruvate synthase, subunit alpha	927.27	47.39	15	15	25	403
AAM03485.1	pyruvate synthase, subunit beta	231.34	27.03	5	5	6	296
AAM03487.1	pyruvate synthase, subunit delta	214.19	36.05	3	3	4	86

2.3.3 In vivo biosynthesis of coenzyme F430

The CfbC and CfbD from *M. thermophila* were overexpressed and purified, separately. They showed more stability than the CfbC and CfbD from *M. acetivorans*. The activity of CfbCD was also tested *in vitro* and showed high activity (not shown). Then, the genes encoding CfbCD from *M. thermophila* were incorporated into the pRSFDuet:*porABDG* plasmid and the resulting plasmid was used to replace pRSFDuet:*fd-fnr*. Thus, the new coenzyme F430 biosynthesis cell line included the pRSFDuet:*Mt_cfbCD-porABDG*, pCDFDuet:*cfbABCDE-nixA*, pDB1282, and pACYCDuet:*sirAC* plasmids. The proteins were coexpressed using the three-stage method. The harvested cell pellets had a dark purple color. The tetrapyrrole products were extracted and analyzed using HPLC. The retention time of the highest peak was 21 min, which corresponds to Ni-sirohydrochlorin *a,c*-diamide (Fig 2.8). The peaks around 7 min corresponding to flavin and an unknown were also detected (Fig 2.8). Unfortunately, there was still no detectable amount of

15,17^{3}-seco-F430-17^{3}-acid or coenzyme F430 produced. Another plasmid, pETDuet:*mm5-mm15-mm7-mm17-mm6-mm3-mm2-A2*, containing several uncharacterized genes proposed to be involved in the assembly/activation of MCR was constructed and used to replace the pDB1282 plasmid containing the *isc* genes in the *E. coli* cell line. Both cell lines were tested in *E. coli* BL21 and C41 cell strains, and under anaerobic growth conditions. However, HPLC analysis of the tetrapyrroles prepared by these cell lines still showed no 15,17^{3}-seco-F430-17^{3}-acid or coenzyme F430 formation (data not shown).

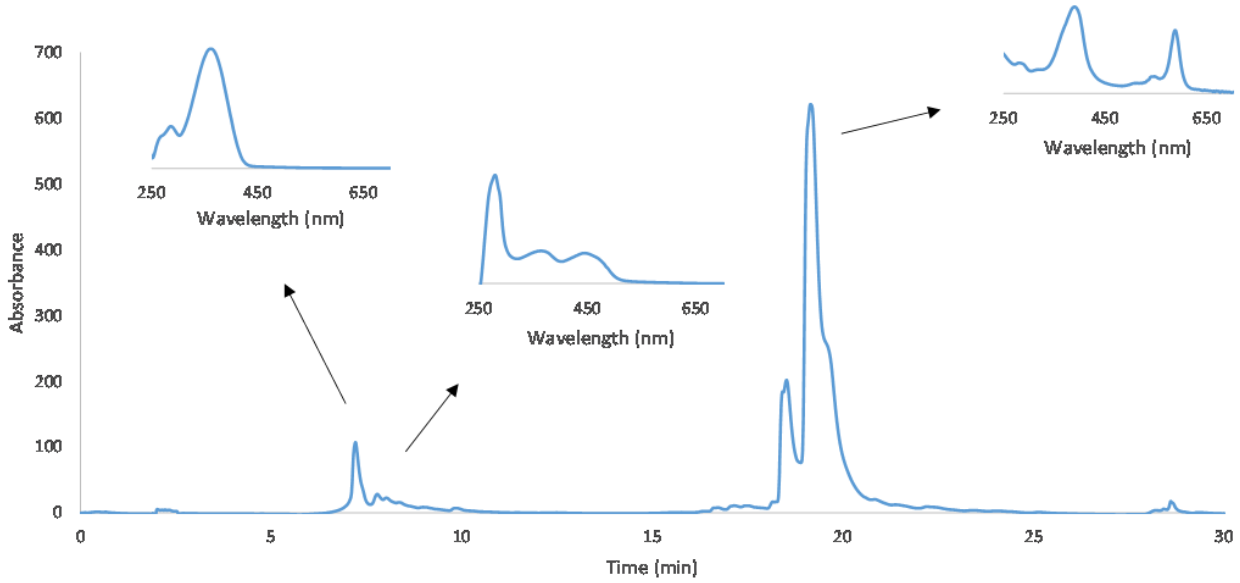


Figure 2.8 HPLC and UV-visible spectroscopic analysis of tetrapyrroles extracted from a strain of *E. coli* containing the pRSFDuet:*Mt_cfbCD-porABDG*, pCDFDuet:*cfbABCDE-nixA*, pDB1282, and pACYCDuet:*sirAC* plasmids. The cells were incubated with 2 mM ferrous sulfate and 2 mM cysteine during the initial induction of the *isc* operon, and then supplemented with 200 μM nickel chloride, 10 μM 5-ALA, and 1 mM SAM during the expression of the remaining genes.

2.3.4 Optimization of different supplementary chemicals

The *E. coli* cell cultures used for the *in vivo* coenzyme F430 biosynthesis experiments were supplemented with 200 μM nickel chloride, 10 μM 5-aminolevulinic acid (5-ALA), and 1 mM SAM. Various concentrations of these supplementary chemicals were tested in their ability

to produce tetrapyrroles. The HPLC result with no supplementary chemicals showed no peak for Ni-sirohydrochlorin (Fig 2.9). Instead, there were a large unknown peak, a flavin peak and a small peak corresponding to sirohydrochlorin at 19 min (Fig 2.9). The HPLC result with half the concentration of supplementary chemicals showed a large Ni-sirohydrochlorin *a*,*c*-diamide peak, but still no 15,17^3-seco-F430-17^3-acid or coenzyme F430 (Fig 2.10). A similar chromatogram was observed when the concentration of supplementary chemicals was doubled (data not shown).

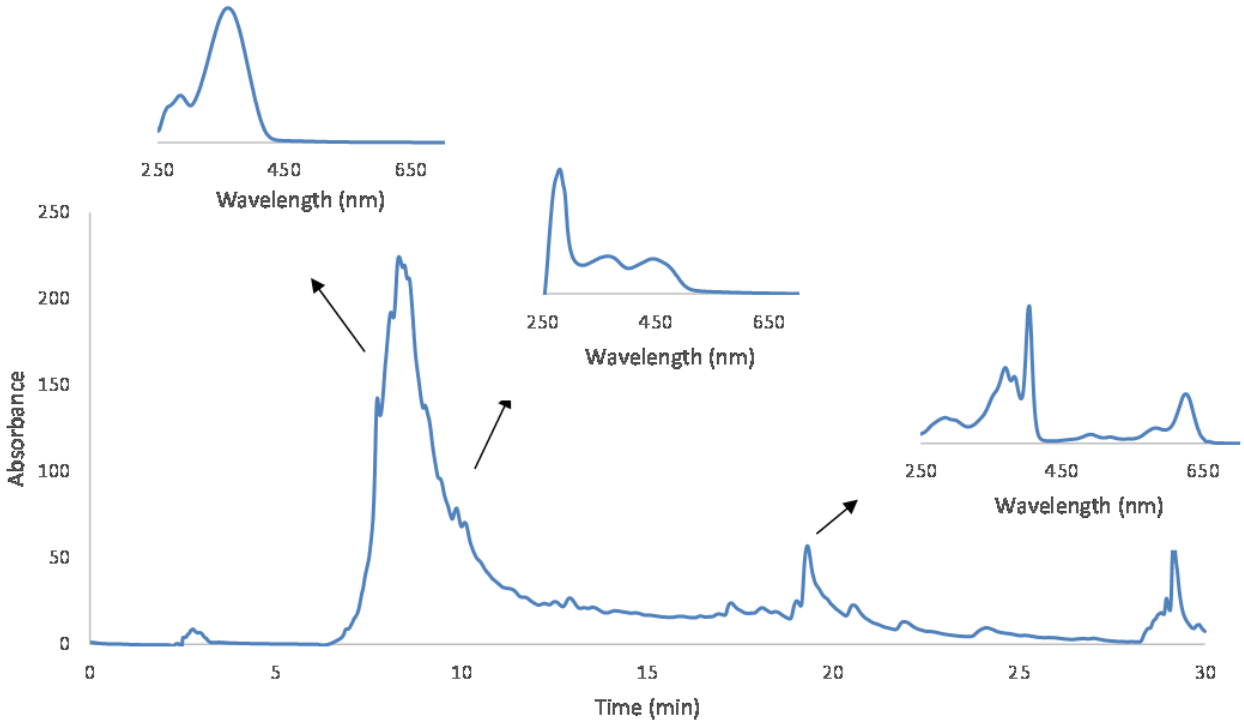


Figure 2.9 HPLC and UV-visible spectroscopic analysis of tetrapyrroles extracted from a strain of *E. coli* containing the pRSFDuet:*Mt_cfbCD-porABDG*, pCDFDuet:*cfbABCDE-nixA*, pDB1282, and pACYCDuet:*sirAC* plasmids incubated without supplementary chemicals. The cells were incubated with 2 mM ferrous sulfate and 2 mM cysteine during the initial induction of the *isc* operon, but were not supplemented with nickel chloride, 5-ALA, or SAM during the expression of the remaining genes.

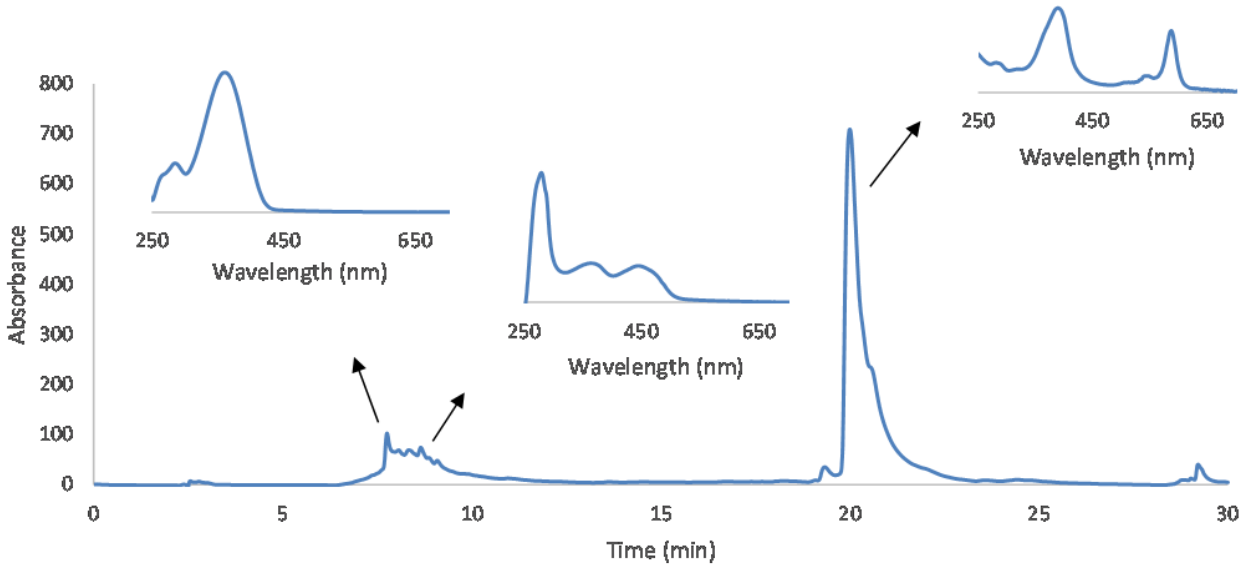


Figure 2.10 HPLC and UV-visible spectroscopic analysis of tetrapyrroles extracted from a strain of *E. coli* containing the pRSFDuet:*Mt_cfbCD-porABDG*, pCDFDuet:*cfbABCDE-nixA*, pDB1282, and pACYCDuet:*sirAC* plasmids incubated with half the amount of supplementary chemicals. The cells were incubated with 2 mM ferrous sulfate and 2 mM cysteine during the initial induction of the *isc* operon, and then supplemented with 100 μM nickel chloride, 5 μM 5-ALA, and 0.5 mM SAM during the expression of the remaining genes.

2.3.5 Optimization of different growth times/temperatures

There are many tetrapyrroles in the nature, but they are all synthesized from a common precursor, uroporphyrinogen III, and share many structural features. Thus, specific growth times and supplementary chemicals may be critical for controlling flux through the coenzyme F430 biosynthetic pathway and preventing the formation of shunt products and inhibiting the desired pathway. Therefore, several growth conditions with different incubation times and temperatures were tested. Cultures of the cell line containing the pRSFDuet:*Mt_cfbCD-porABDG*, pCDFDuet:*cfbABCDE-nixA*, pDB1282, and pACYCDuet:*sirAC* plasmids were incubated at 37 °C until the OD_{600} reached 0.5 and then 2 mM ferrous sulfate, 2 mM cysteine, and 3 g of arabinose were added for the expression of the *isc* operon. After a 2 hr incubation at 37 °C, 200 μM nickel chloride, 10 μM 5-aminolevulinic acid (5-ALA), 1 mM SAM, and 0.4 mM IPTG were

added, and the cultures were incubated for varying times/temperatures. In the first experiment the cultures were incubated with shaking for 2 hours at 37 °C. The cell pellets were collected and the tetrapyrroles were extracted and analyzed using HPLC (Fig 2.11). The result showed no 15,17^3-seco-F430-17^3-acid or coenzyme F430 and the amount of flavin produced was similar to that of Ni-sirohydrochlorin *a*, *c*-diamide.

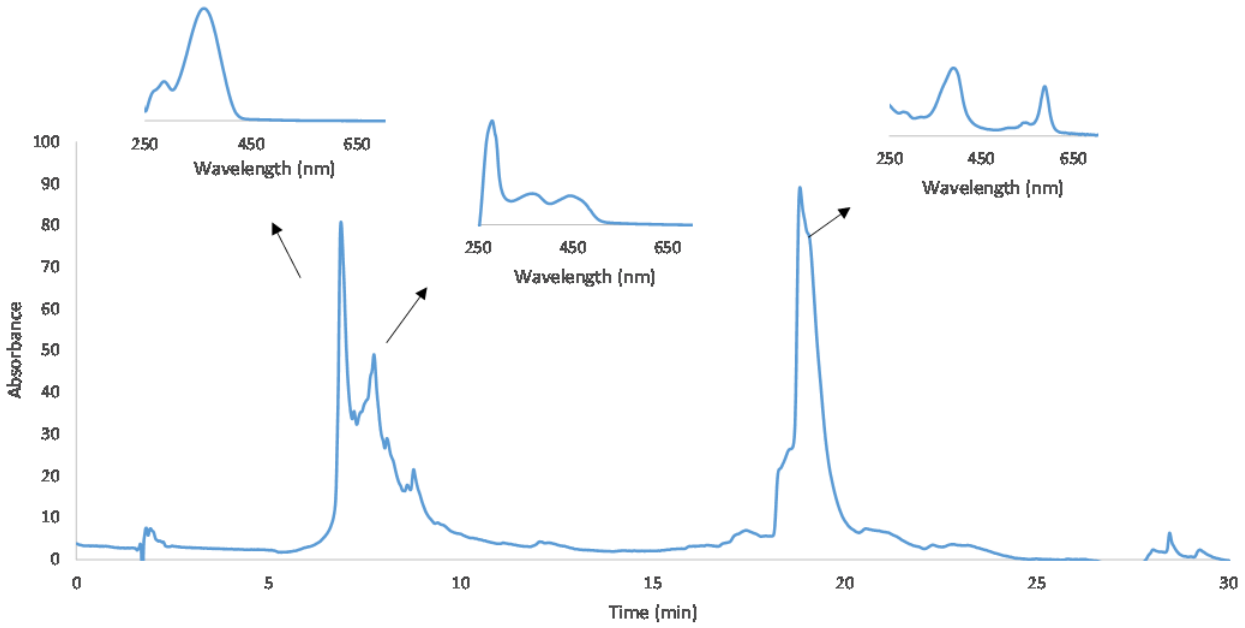


Figure 2.11 HPLC and UV-visible spectroscopic analysis of tetrapyrroles extracted from a strain of *E. coli* containing the pRSFDuet:*Mt_cfbCD-porABDG*, pCDFDuet:*cfbABCDE-nixA*, pDB1282, and pACYCDuet:*sirAC* plasmids grown at 37 °C for 2 hr after inducing with IPTG.

Subsequent experiments involved incubating with shaking for 4 hr at 37 °C, 6 hr at 37 °C, and 8 hr at 37 °C after inducing with IPTG. The cell pellets were all collected and the extracted tetrapyrroles were analyzed using HPLC (Fig 2.12). Each of these experiments yielded similar results wherein no 15,17^3-seco-F430-17^3-acid or coenzyme F430 was detected. However, compared with the first test, there was significantly more Ni-sirohydrochlorin *a*, *c*-diamide produced than flavin.

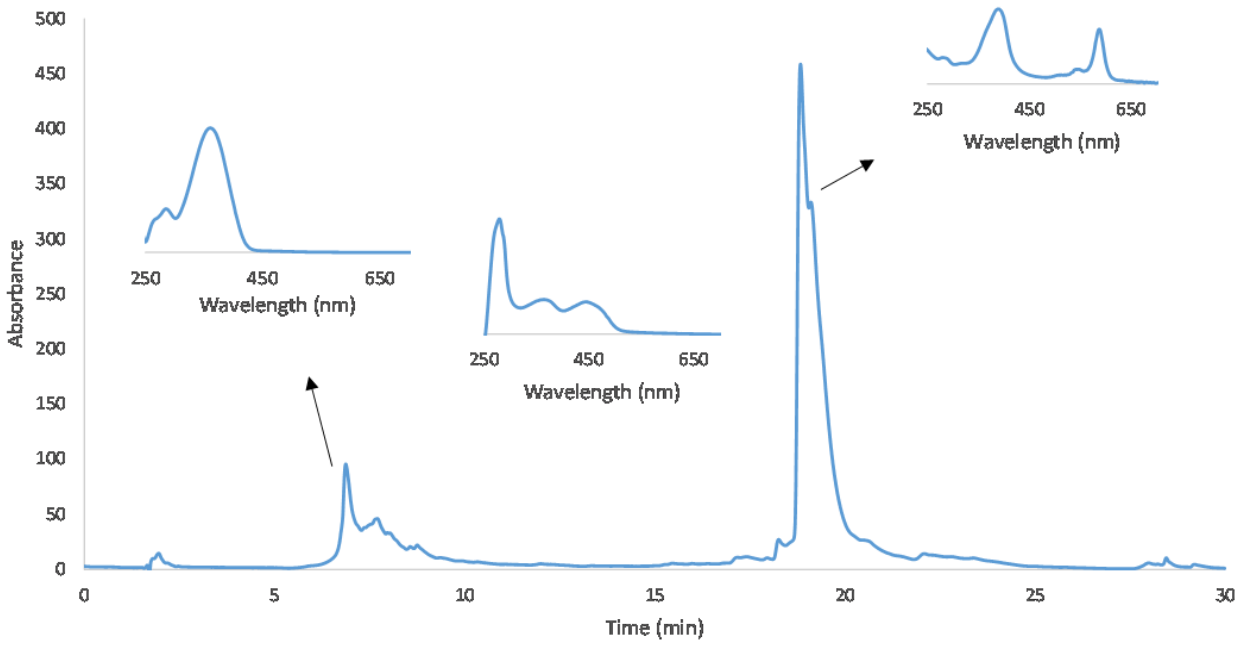


Figure 2.12 HPLC and UV-visible spectroscopic analysis of tetrapyrroles extracted from a strain of *E. coli* containing the pRSFDuet:*Mt_cfbCD-porABDG*, pCDFDuet:*cfbABCDE-nixA*, pDB1282, and pACYCDuet:*sirAC* plasmids grown at 37 °C for 4 hr after inducing with IPTG.

We also tested incubation with shaking overnight at 18 °C after inducing with IPTG, followed by an additional 2 or 4 hr incubation at 37°C. The tetrapyrrole products were extracted and analyzed using HPLC (Fig 2.13). These data were again similar to one another and no 15,17^3-seco-F430-17^3-acid or coenzyme F430 was detected.

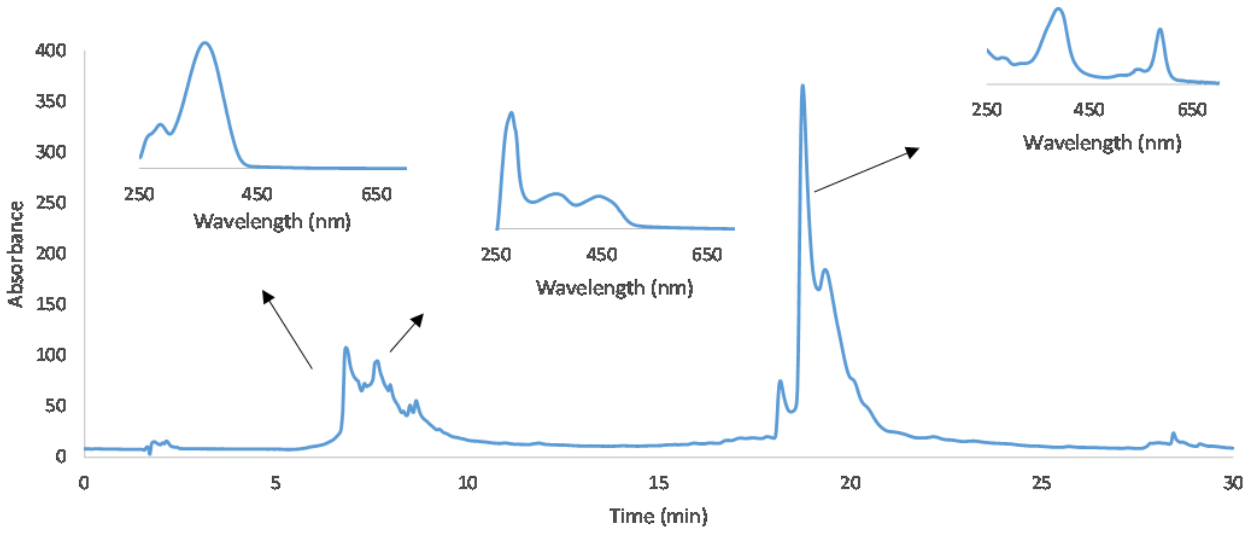


Figure 2.13 HPLC and UV-visible spectroscopic analysis of tetrapyrroles extracted from a strain of *E. coli* containing the pRSFDuet:*Mt_cfbCD-porABDG*, pCDFDuet:*cfbABCDE-nixA*, pDB1282, and pACYCDuet:*sirAC* plasmids grown at 18 °C overnight after inducing with IPTG, followed by incubation at 37 °C for 2 hr.

These preliminary experiments suggest that more Ni-sirohydrochlorin *a,c*-diamide was produced when the cells were incubated with shaking for 4 hr at 37 °C. Also, adding half of the amount of supplementary chemicals (100 μM nickel chloride, 5 μM 5-ALA, and 0.5 mM SAM) was enough to provide a high yield of Ni-sirohydrochlorin *a,c*-diamide. However, no 15,17^3-seco-F430-17^3-acid or coenzyme F430 was detected, and additional experiments are needed to explore the *in vivo* biosynthesis of coenzyme F430 in a heterologous host.

2.4 Discussion

Anaerobic methanotrophic archaea have potential utility in natural gas-to-liquid fuel conversion strategies, but they have not been obtained in pure culture and have exceptionally slow growth rates. A solution is to engineer the methane-forming pathway and its key enzyme, MCR, into a more robust microorganism. Thus, the *in vivo* biosynthesis of coenzyme F430 is a critical step. Our group has identified the coenzyme F430 biosynthetic enzymes: CfbA, CfbB, CfbC, CfbD, and CfbE. Here, we investigated the *in vivo* biosynthesis of coenzyme F430.

Previously, an *E. coli* cell line was constructed to examine the *in vivo* biosynthesis of coenzyme F430 that included the *cfb* genes (pCDFDuet:*cfbABCDE-nixA*), genes to produce the immediate substrate for the pathway (pACYCDuet:*sirAC*), genes to provide electrons for the CfbCD reduction reaction (pRSFDuet:*fd-fnr*), and genes to enhance iron-sulfur cluster incorporation into CfbCD (pDB1282). All the genes were coexpressed, and the tetrapyrroles that were produced were analyzed using HPLC. The result showed the pathway stopped at the third step, the CfbCD reaction. To help identify the source of the problem, we prepared cell free extracts containing all the proteins and tested the activity of CfbCD and Fd/FNR. Unfortunately, both of them were not active. Later, we searched the genome of *M. acetivorans* for a native physiological reducing system (PFOR) and cloned the genes. Meanwhile, a thermophilic version of CfbCD was cloned and combined with the PFOR plasmid (pRSFDuet:*Mt_cfbCD-porABDG*). These genes were coexpressed with the *cfb*, *sirAC*, and *isc* genes. However, the coenzyme F430 biosynthesis pathway still stopped at the CfbCD reaction. A plasmid encoding several methanogenesis marker proteins that are proposed to be involved in MCR assembly/activation was used to substitute the pDB1282 plasmid containing the *isc* operon and the cell line was tested for the production of coenzyme F430 and/or 15,17^3-seco-F430-17^3-acid. Both cell lines were tested in *E. coli* BL21 and C41 cell strains, and anaerobic growth condition. However, the obstacle was still not overcome.

To optimize the growth conditions, we incubated the cell line containing the pRSFDuet:*Mt_cfbCD-porABDG*, pCDFDuet:*cfbABCDE-nixA*, pDB1282, and pACYCDuet:*sirAC* plasmids with four different concentrations of supplementary chemicals (i.e., 0-200 μM nickel chloride, 0-10 μM 5-ALA, and 0-1 mM SAM) and six different growth times/temperatures. Without the addition of supplementary chemicals, the coenzyme F430

biosynthesis pathway was blocked at the first step, the CfbA reaction, which suggests that nickel is a limiting substrate for the pathway. Halving the amount of supplementary chemicals from the maximum tested showed similar results, indicating that the addition of extra chemicals does not greatly influence the yield of Ni-sirohydrochlorin *a,c*-diamide. When we grew the cell line at 37 °C for 2 hr after inducing with IPTG, the yield of Ni-sirohydrochlorin *a,c*-diamide was very low when compared with an incubation at 37 °C for 4 hr. However, when we continued to increase the growth time to 6 hr and 8 hr, the yield of Ni-sirohydrochlorin *a,c*-diamide was not enhanced. Also, we grew the cell line overnight at 18 °C after inducing with IPTG, followed by an incubation at 37 °C for 2 or 4 hours. Overall, the yield of Ni-sirohydrochlorin *a,c*-diamide was not enhanced in these cell lines. Thus, the best growth conditions tested for the *in vivo* synthesis of coenzyme F430 biosynthetic intermediates was at 37 °C for 4 hours after inducing with IPTG. Unfortunately, no 15,17^3-seco-F430-17^3-acid or coenzyme F430 was detected under any of the conditions tested.

In general, the *in vivo* synthesis of biomolecules in a heterogeneous host is much more complicated than *in vitro* biosynthesis. Although we cloned and expressed all of the *cfb* genes in *E. coli* and they were individually found to be active *in vitro*, there are still many obstacles to be solved to demonstrate their activity *in vivo*. First, we need a proper physiological reductant to support the CfbCD reaction. Previously, a ferredoxin/ferredoxin:NADP$^+$ reductase system from spinach was used, which was tested and found to support CfbCD activity *in vitro*. However, it was difficult to ascertain whether these eukaryotic proteins were active in *E. coli*. A new native physiological reducing system, PFOR, was tested in the cell line, but 15,17^3-seco-F430-17^3-acid or coenzyme F430 was still not detected. The recombinant PFOR needs to be tested *in vitro* to ensure it is produced in an active form. Second, CfbCD is an iron-sulfur enzyme, and its

heterologous production requires the construction and insertion of iron-sulfur clusters. Although CfbCD activity can be detected *in vitro*, the enzyme complex and its iron-sulfur centers are unstable and readily degrade. A thermophilic version of CfbCD was found to have better stability/activity *in vitro*, but it still did not exhibit activity *in vivo*. Third, physiological reactions need proper concentrations of substrates and enzymes to ensure proper flux towards the desired product. It is easy to adjust the concentration of chemicals and enzymes to drive reactions *in vitro*, but this task is more difficult to control *in vivo*. Here, we changed the concentration of supplementary chemicals in the cultures, but it is difficult to know how much is absorbed and utilized by *E. coli* for its intended purpose. Fourth, some biosynthesis pathways require many auxiliary genes to regulate and activate the pathway. Thus, some of the key enzymes related to the pathway may not be found in the heterologous host *E. coli*. Here, we introduced a plasmid, pETDuet:*mm5-mm15-mm7-mm17-mm6-mm3-mm2-A2*, that encoded many methanogenesis marker proteins of unknown functions, but which have been implicated in MCR assembly/activation and could play a role in the biosynthesis/insertion of coenzyme F430. Additional experiments are still needed to confirm the roles of these proteins and to identify others that may be necessary for this process. Fifth, there are many differrent tetrapyrroles in the nature, but they all have a common precursor, uroporphyrinogen III. Thus, a specific growth conditions are critical for controlling flux along the desired pathway and preventing the formation of side products and potential inhibitors of the main pathway. For example, an excess of iron could convert sirohydrochlorin to siroheme, which is a intermediate for the biosynthesis pathway of heme d_1, or zinc contamination could convert sirohydrochlorin to Zn-sirohydrochlorin, which could inhibit enzymes in the coenzyme F430 biosynthetic pathway. Several growth conditions with different times and temperatures were tested. However, none of

them resulted in the formation of 15,17^3-seco-F430-17^3-acid or coenzyme F430. The internal environment of bacteria is complex. It is possible that one of the many small molecules generated in the cytoplasm of *E. coli* could function as an inhibitor to block the coenzyme F430 biosynthesis pathway. More experiments are needed to test this hypothesis.

Chapter Three: Characterization of transient intermediates of CfbCD catalysis

3.1 Introduction

MCR catalyzes the conversion of coenzyme B and methyl-coenzyme M to the mixed heterodisulfide CoB-S-S-CoM and methane, which is the major component of natural gas and a potent greenhouse gas.[14, 61] MCR cannot currently be produced in an active form in a heterologous host because our knowledge of the biosynthesis and insertion of the unique nickel-containing prosthetic group of MCR, coenzyme F430, is incomplete.[87] Coenzyme F430 is the most highly reduced tetrapyrrole in nature, and this is achieved by a single multi-electron redox reaction catalyzed by the CfbCD complex in the key step of the pathway.[87] The CfbCD reaction involves a 6-electron reduction of the isobacteriochlorin ring system of Ni-sirohydrochlorin *a*,*c*-diamide, cyclization of the *c*-acetamide side chain to form the γ-lactam E ring, and the formation of 7 stereocenters.[87] Understanding the mechanism of CfbCD catalysis may aid in the development of specific inhibitors to reduce natural greenhouse gas emission and provide a critical step towards metabolic engineering efforts using MCR for methane conversion.

Moreover, CfbC and CfbD are homologous to the Fe protein (NifH) and the NifD subunit of the MoFe protein (NifDK) of nitrogenase.[142] Nitrogenase couples the hydrolysis of 16 ATPs to the 8-electron reduction of dinitrogen and 2 protons in the pathway of biological nitrogen fixation (Figure 3.1).[96] Several other homologs of nitrogenase are known, including nitrogenase assembly enzymes and (bacterio)chlorophyll biosynthesis enzymes.[97] The latter group includes the dark-operative protochlorophyllide oxidoreductase (DPOR), which catalyzes the ATP-dependent 2-electron reduction of the C17=C18 double bond of protochlorophyllide, and chlorophyllide oxidoreductases (COR), which catalyzes the reduction of the C7=C8 double bond of chlorophyllide (Figure 3.1).[146, 147] Each of these homologs contains heterotetrameric MoFe

protein-like components, while CfbD is homomeric.[97] Thus, we proposed that the CfbCD complex is representative of a more primitive lineage of the nitrogenase superfamily. Overall, understanding the mechanism of CfbCD will help to further understand the mechanism and evolution of the nitrogenase superfamily.

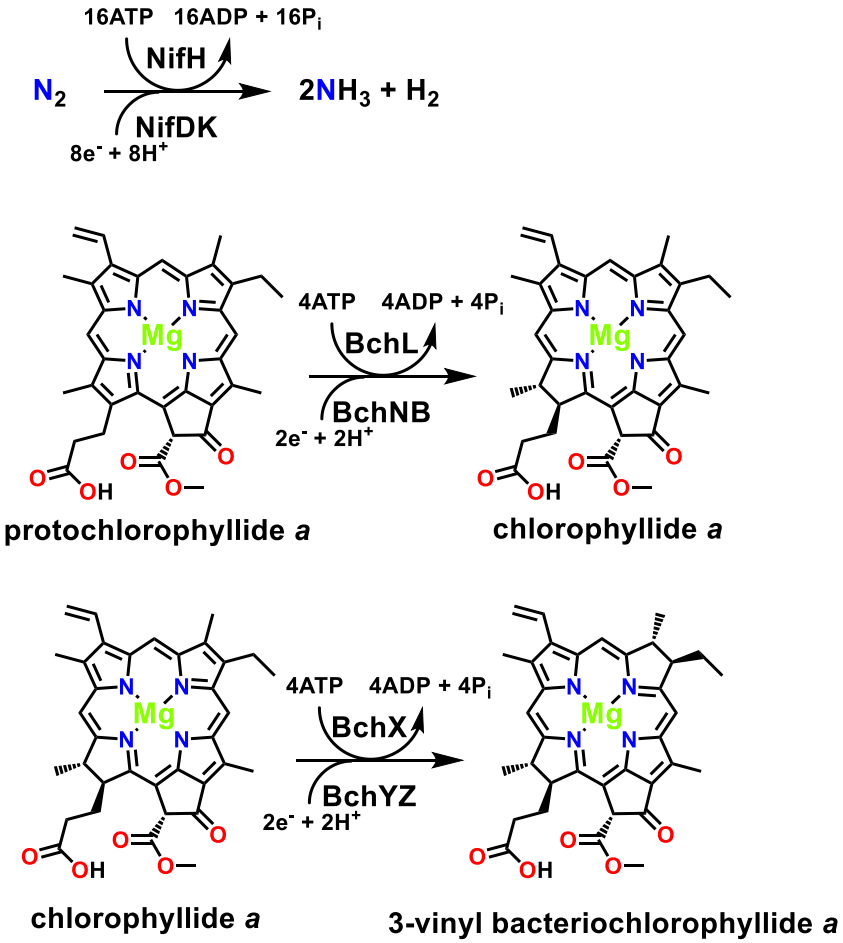


Figure 3.1 Reactions catalyzed by nitrogenase (NifHDK), DPOR (BchLNB), and COR (BchXYZ).

According to the high-resolution structures of nitrogenase, the Fe protein is homodimeric and contains two ATP-binding sites and a single [4Fe-4S] cluster (Figure 3.2).[148] ATP binding and hydrolysis is coupled to multiple cycles of complex formation/dissociation and

intercomponent electron transfer.[148] Electrons are transferred sequentially from the [4Fe-4S] cluster of the Fe protein to an [8Fe-7S] cluster, the P-cluster, and then to the iron-molybdenum cofactor (FeMo-co), which is the site of dinitrogen reduction in the MoFe protein heterotetramer.[148] DPOR also consists of two separable components, the L-protein homodimer and the NB-protein heterotetramer (Figure 3.2).[146] In the NB-protein, the P-clusters are replaced with [4Fe-4S] clusters (NB-clusters) with an unusual 3 cysteine/1 aspartate coordination environment and the tetrapyrrole substrate binds in place of FeMo-co.[146]

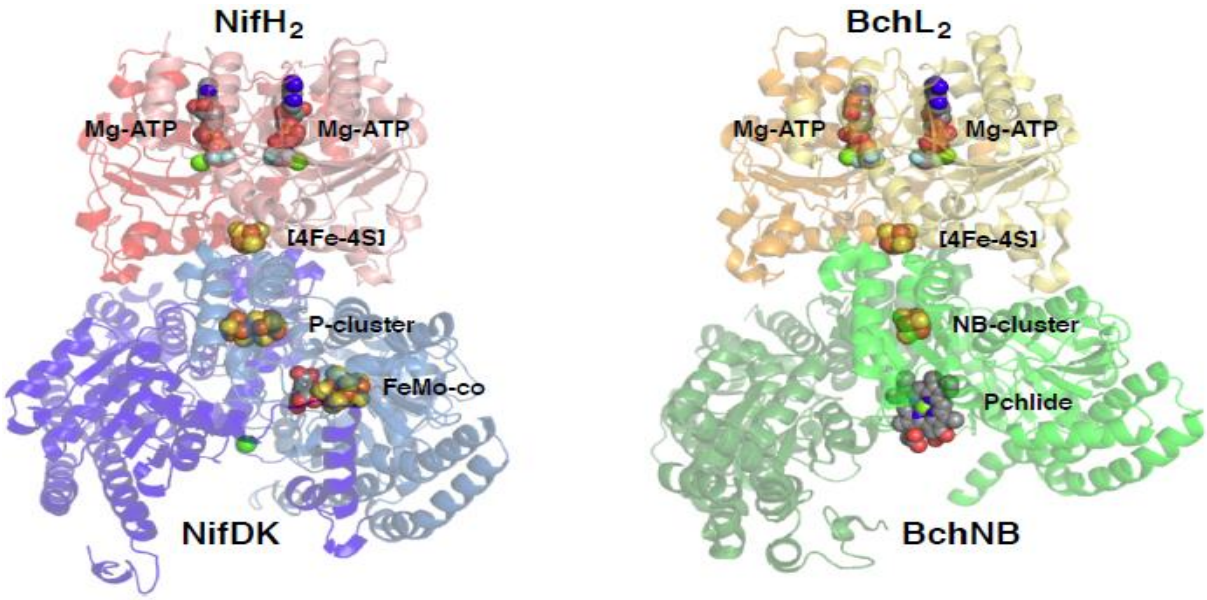


Figure 3.2 Structures of nitrogenase (left) and DPOR (right).

The CfbCD complex does not contain FeMo-co and the Ni-sirohydrochlorin *a,c*-diamide substrate is likely to bind in its stead, similar to what is observed in DPOR. Moreover, unlike all other nitrogenase homologs, the CfbD core is homomeric and it is unclear if there are also two active sites, or only one at the tetrameric interface. Based on the symmetry of the proton additions whose stereochemistry is known in 15,17^3-seco-F430-17^3-acid product, it appears as though one proton is added above and one below the ring for each of the four pyrroles. Thus, the

CfbCD reaction can be consistent with one active site at the tetrameric interface, where each monomer has two active site acid-base residues, one above the ring and one below. A proposed mechanism of CfbCD catalysis consistent with this active site geometry is shown in Figure 3.3. The four rounded squares represent the four monomers of CfbD. Each monomer is proposed to contain two acid-base residues (AH and BH) involved in substrate protonation, with the lighter colored residue positioned below and the darker colored residue above the plane of the isobacteriochlorin ring. Based on the structure of the isolated substrate and product, there are 6 electrons and 7 protons added during the CfbCD reaction. However, it is also possibile that CfbCD catalyzes the addition of 7 or 8 electrons/protons and an electron or hydrogen is lost during product isolation, respectively.

In the CfbCD reaction, the Ni-sirohydrochlorin *a,c*-diamide substrate and the 15,17^3-seco-F430-17^3-acid product have distinct UV-visible spectra.[87] As the conjugation of the π system changes during the reaction, it is likely that each intermediate will also have unique UV-visible absorption properties. The Layer and Warren groups observed that an initial product of the CfbCD reaction is formed that slowly converts to 15,17^3-seco-F430-17^3-acid.[149] It was proposed that this intermediate is Ni-hexahydrosirohydrochlorin *a,c*-diamide and that γ-lactam ring formation is the slow, non-enzymatic reaction that produces 15,17^3-seco-F430-17^3-acid.[149] However, another possibility is that this initial product is the Ni(II)-hydride form of 15,17^3-seco-F430-17^3-acid, which slowly evolves hydrogen. If this is the case, in addition to facilitating cyclization, CfbCD catalyzes the addition of 8 electrons and 8 protons to Ni-sirohydrochlorin *a,c*-diamide and produces H_2 gas, which is like nitrogenase, rather than the addition of 6 electrons and 7 protons, which is unusual stoichiometry. Yet another possibility is that CfbCD

catalyzes the addition of 7 electrons and 7 protons along with cyclization and the transient intermediate is the Ni(I) form of 15,17^3-seco-F430-17^3-acid.

To support these hypotheses and gain insight into the mechanism of CfbCD, four experiments were designed. The first involved a crystallization study of CfbCD. The complex of CfbCD with Ni-sirohydrochlorin *a,c*-diamide and a non-hydrolyzable analog of ATP was purified using fast protein liquid chromatography (FPLC) and crystallization conditions were determined using a sparse matrix screen. The second experiment involved characterization of the transient intermediate using ^{15}N-labeled Ni-sirohydrochlorin *a,c*-diamide. Whether the *c*-amide side chain of the transient intermediate is cyclized can be determined using ^{1}H-^{15}N heteronuclear single quantum coherence (HSQC) nuclear magnetic resonance (NMR) and comparison with both Ni-sirohydrochlorin *a,c*-diamide and 15,17^3-seco-F430-17^3-acid. The third experiment involved synthesizing a substrate analog, Ni-didecarboxysirohydrochlorin *a,c*-diamide, for mechanistic studies. A proton is provided by the the *h*-acetic acid arm of Ni-sirohydrochlorin *a,c*-diamide in the proposed mechanism (Figure 3.3). Thus, investigating the reaction products of CfbCD with Ni-didecarboxysirohydrochlorin *a,c*-diamide can provide insight into the order and mechanism of proton addition in CfbCD catalysis. Finally, a stereochemical study of proton addition to the two (α and δ) methylene bridges of 15,17^3-seco-F430-17^3-acid was initiated. The proposed mechanism (Figure 3.3) predicts that both protons are added to the *si* face during the reaction. Which proton is added during the CfbCD reaction can be determined by performing the reactions in D_2O and H_2O and seeing which becomes deuterated using ^{1}H-NMR spectroscopy.

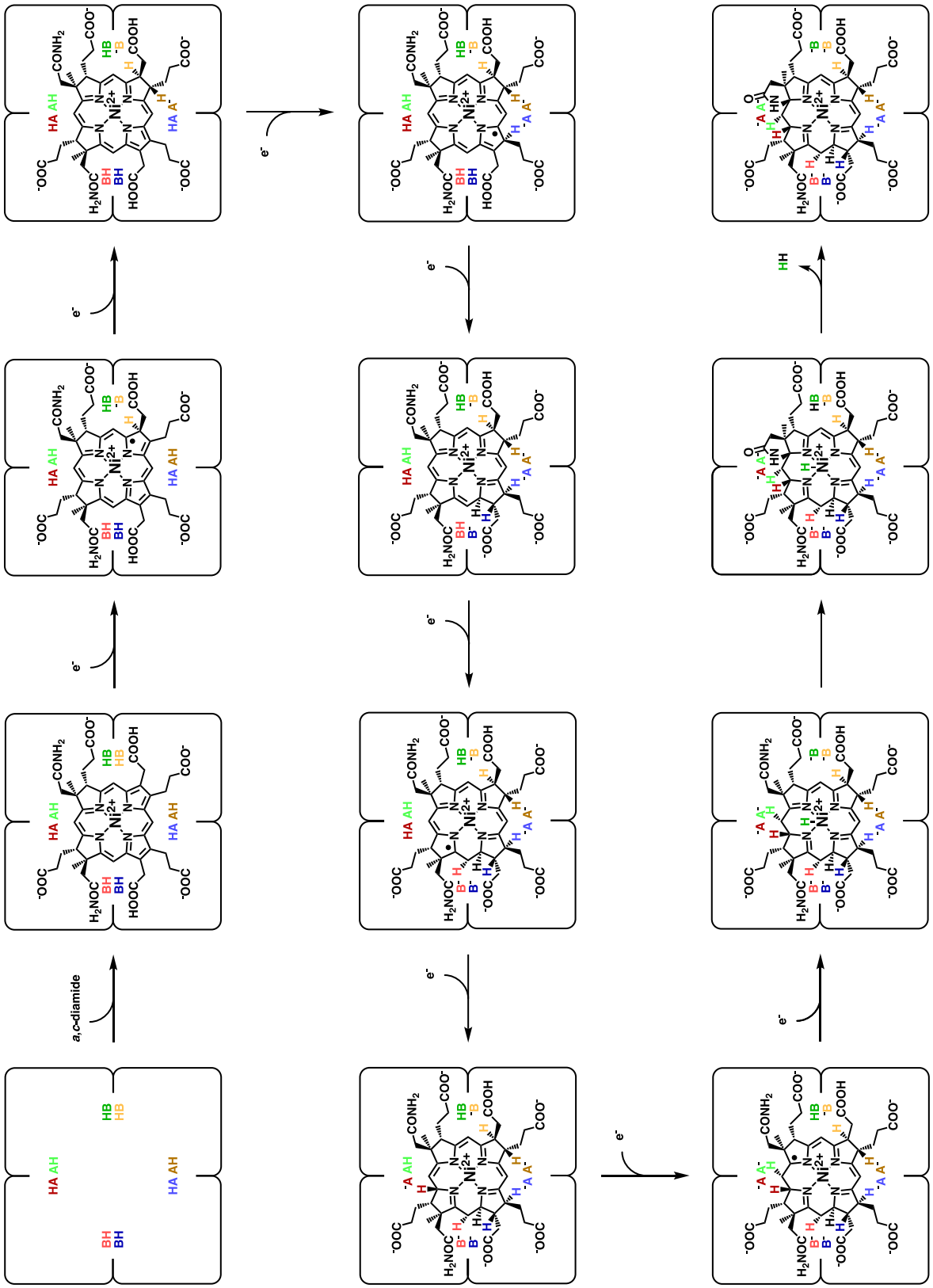


Figure 3.3 Proposed mechanism of CfbCD catalysis.

3.2 Materials and Methods

3.2.1 Materials

Most chemicals and buffer components were from VWR Chemicals BDH® (Radnor, PA). Nitrogen-15 ammonia and deuterium oxide were from Acros Organics (Fair Lawn, NJ). Porphobilinogen (PBG) was from Frontier Scientific Inc. (Logan, UT). Adenosine 5′-(β,γ-imido)triphosphate lithium salt hydrate and adenosine triphosphate (ATP) were purchased from Sigma-Aldrich (St. Louis, MO). Nicotinamide adenine dinucleotide ($NADP^+$) was from Amresco (Albany, NY). LB media was from Becton, Dickinson and Company (BD) (Franklin Lakes, NJ), while buffer components were from VWR Chemicals BDH® (Radnor, PA). Primers were ordered from Sigma-Aldrich and the synthesized *nirDL* gene from *Hydrogenobacter thermophilus* was purchased from GeneArt (Waltham, MA). Phusion High-Fidelity PCR Kit, Gibson Assembly Master Mix, and restriction endonucleases were from New England Biolabs (Ipswich, MA). Gel extraction, PCR cleanup, and mini prep kits were purchased from Omega Bio-Tek (Norcross, GA). Profinity IMAC resin was obtained from Bio-Rad (Hercules, CA).

3.2.2 Plasmid construction

The *nirDL* gene from *Hydrogenobacter thermophilus* was codon optimized for expression in *Escherichia coli*. Polymerase chain reaction (PCR) was then used with the Gibson Assembly method to insert the *nirDL* gene into the pETSUMO vector. The forward primer designed for *nirDL* was 5′-gagaacagattggtggtATGGACGACCTGGACCTGC-3′. The reverse primer was 5′- taaatacctaagcttgtcTCATGCGGCTACCTCGGC-3′. The forward primer designed for vector pETSUMO was 5′- GAGGCTCACAGAGAACAGATTGGTGGT-3′. The reverse primer was 5′-CTCGAGGACAAGCTTAGGTATTTATTCGGCGC-3′. The *nirDL* and pETSUMO PCR products were purified via a Lonza FlashGel™ DNA Cassette (Basel,

Switzerland) and recovered using an Omega Bio-tek E.Z.N.A.® Gel Extraction Kit (Norcross, GA). The resulting *nirDL* and pETSUMO were combined using Gibson Assembly method. The ligated mixture was then transformed into *E. coli* TOP10 cells and plated on Luria-Bertani (LB) agar plates containing 50 µg/mL kanamycin. Colonies were picked, grown in liquid LB medium, and plasmid preparations were made using the Omega Bio-tek E.Z.N.A.® Plasmid Mini Kit.

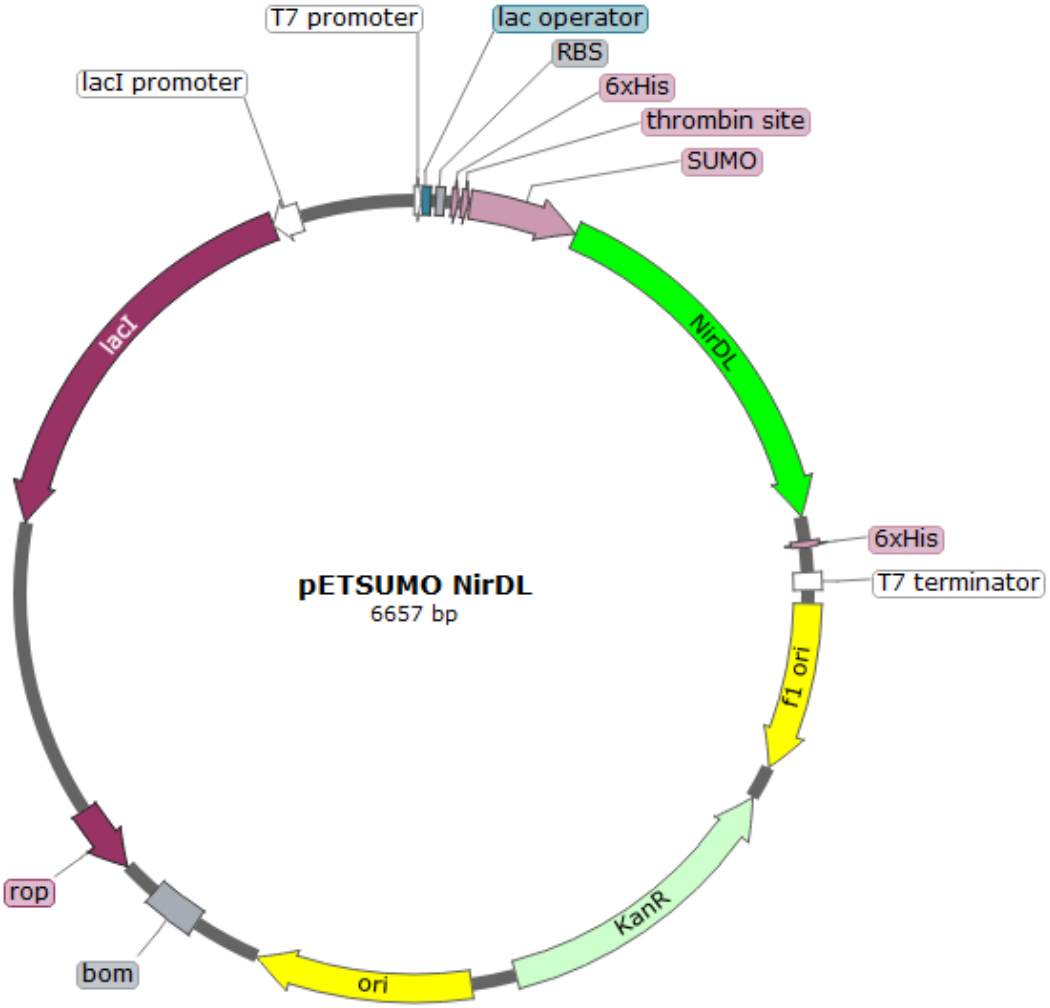


Figure 3.4 Plasmid construct for the *nirDL* gene in pETSUMO.

3.2.3 Expression and purification

The expression and purification of HemC, HemD, SirA, SirC, CfbA, CfbB, and CfbCD using the pET-28b(+):*hemC*, pET-28b(+):*hemD*, pET-28b(+):*sirC*, pET-28b(+):*sirA*, pET-28b(+):*cfbA*, pET-28b(+):*cfbB*, pETSUMO:*cfbC*, and pETSUMO:*cfbD* plasmids have been reported in the Chapter Two. The pETSUMO:*nirDL* expression vector was transformed into *E. coli* BL21(DE3) cells and grown in overnight cultures (2 mL of LB with 50 μg/mL kanamycin) that were used to inoculate 6 × 1 L of LB (50 mg/L kanamycin). The cultures were incubated with shaking at 37 °C until the OD_{600} reached 0.6. The cultures were then induced with isopropyl β-D-1-thiogalactopyranoside (IPTG) at a final concentration of 0.125 mM and incubated at 17 °C overnight. The cells were centrifuged at 15,900 × g for 45 min. The supernatants were decanted, and the cells were resuspended in 50 mM Tris buffer (pH 7.5) containing 300 mM NaCl, 5 % glycerol,, and 5 mM imidazole. The cells were then sonicated and centrifuged at 15,900 × g for 1 h. The supernatants were loaded onto a Bio-Rad Econo-Pac column loaded with nickel-charged Profinity IMAC resin. The columns were washed with 50 mM Tris buffer (pH 7.5) containing 300 mM NaCl, 5 % glycerol, and 5 mM imidazole. The protein was then eluted from the column by adding 50 mM Tris buffer (pH 7.5) containing 300 mM NaCl, 5 % glycerol, and 500 mM imidazole in 1 mL fractions. The protein containing fractions were combined and spin-concentrated/buffer exchanged. The purified SUMO-tagged NirDL was incubated with SUMO protease (1:1 [w:w]) overnight at room temperature to remove the N-terminal SUMO-tag. After incubation, the cleaved protein was passed through a Ni-NTA column to remove the SUMO-tag and SUMO protease. The eluted protein was then spin-concentrated/buffer exchanged into 100 mM Tris buffer (pH 8.0) containing 200 mM NaCl and 15 % glycerol.

3.2.4 Crystallization of CfbCD complex

Crystallization trials of CfbCD were performed in the glovebox. 200 μM iron-sulfur cluster reconstituted CfbC was incubated with 5 mM adenosine 5′-(β,γ-imido)triphosphate lithium salt and 3 mM sodium dithionite in 50 mM Tris buffer (pH 7.5) for 1 h. 200 μM iron-sulfur cluster reconstituted CfbD and 1 mM Ni-sirohydrochlorin *a,c*-diamide were incubated with the CfbC mixture for an additional 1 h. FPLC purification of the CfbCD complex was performed on an Cytiva AKTA go system with a Diode Array Detector (DAD) in the glovebox. The column was an Cytiva Superdex 200 size exclusion column. The Cytiva UNICORN software was used for data analysis. The mobile phase consisted of 100 mM Tris buffer (pH 8.0) with 100 mM NaCl. A flow rate of 0.5 mL/min was used and the trace was acquired using detection at 280 nm.

Three sparse matrix screens (The JCSG Core IV, The JCSG +, and The Protein Complex Suites, Qiagen) were used to determine cocrystallization conditions for CfbCD. Crystals were screened via the sitting-drop vapor diffusion method. Each drop consisted of 1 μL of the CfbCD complex and 1 μL of crystallization buffer.

3.2.5 Synthesis of 15N-labeled intermediate of CfbCD catalysis

To synthesize ^{15}N-labeled Ni-sirohydrochlorin *a,c*-diamide, we incubated 2.5 μM HemC, 2.5 μM HemD, 2.5 μM SirA, 2.5 μM SirC, 2.5 μM CfbA, 2.5 μM CfbB, 0.88 mM porphobilinogen (PBG), 2 mM SAM, 1 mM $NADP^+$, 4 mM $MgCl_2$, 200 μM $NiCl_2$, 4 mM nitrogen-15 ammonia, and 5 mM ATP in 30 mL of 100 mM Tris-HCl, pH 8.0 buffer. The reaction was performed overnight at 37 °C in the glovebox. The products were extracted by adding 100 μL methanol to 100 μL of the assay mixture and centrifuged at 7,000 × g for 30 min. The supernatant was analyzed using HPLC.

To obtain a ^{15}N-labeled version of the transient intermediate for structural characterization, we incubated 10 μM CfbC, 10 μM CfbD, 5 mM ATP, and 3 mM sodium dithionite with 0.2 mM of ^{15}N-labeled Ni-sirohydrochlorin *a,c*-diamide. The reaction was performed at 37 °C in the glovebox. The reaction was quenched by adding 60 μL methanol to 60 μL of the assay mixture at one hour time intervals from 1.5 h to 9.5 h, and after an 18 h incubation. The extraction was centrifuged at 7,000 × g for 30 min and the supernatants were analyzed using HPLC. Then, we incubated 10 μM CfbC, 10 μM CfbD, 5 mM ATP, and 3 mM sodium dithionite with 6 mM of ^{15}N-labeled Ni-sirohydrochlorin *a,c*-diamide. The reaction was spin concentrated and buffer exchanged with 50 mM phosphate buffer (pH 8.0) after a 5 h incubation. The products were extracted by adding 30 mL methanol and centrifuged at 7,000 × g for 30 min. The supernatants were analyzed using NMR.

3.2.6 Synthesis of Ni-didecarboxysirohydrochlorin a,c-diamide

The CfbA reaction was performed by incubating 2.5 μM HemC, 2.5 μM HemD, 2.5 μM SirA, 2.5 μM SirC, 2.5 μM CfbA, 0.88 mM porphobilinogen (PBG), 2 mM SAM, 1 mM $NADP^+$, 4 mM $MgCl_2$, and 200 μM $NiCl_2$ in 1 mL of 100 mM Tris-HCl, pH 8.0 buffer. The reaction was performed overnight at 37 °C in the glovebox. To synthesize Ni-didecarboxysirohydrochlorin, we incubated 10 μM NirDL with the CfbA reaction overnight at 37 °C. The products were extracted by adding 100 μL methanol to 100 μL of the assay mixture and centrifuged at 7,000 × g for 30 min. The supernatants were analyzed using HPLC. To synthesize Ni-didecarboxysirohydrochlorin *a,c*-diamide, we incubated 2.5 μM CfbB, 5 mM ATP, and 3 mM glutamine with the NirDL reaction overnight at 37 °C. The products were extracted by adding 100 μL methanol to 100 μL of the assay mixture and centrifuged at 7,000 × g for 30 min. The supernatants were analyzed using HPLC.

3.2.7 Synthesis of deuterated 15,173-seco-F430-173-acid

To synthesize Ni-sirohydrochlorin *a,c*-diamide, the substrate for the CfbCD reaction, we incubated 2.5 μM HemC, 2.5 μM HemD, 2.5 μM SirA, 2.5 μM SirC, 2.5 μM CfbA, 2.5 μM CfbB, 0.88 mM porphobilinogen (PBG), 2 mM SAM, 1 mM $NADP^+$, 4 mM $MgCl_2$, 200 μM $NiCl_2$, 4 mM glutamine, and 5 mM ATP in 30 mL of 100 mM Tris-HCl, pH 8.0 buffer. The reaction was performed overnight at 37 °C in the glovebox.

The synthesized Ni-sirohydrochlorin *a,c*-diamide bound to CfbB was spin-concentrated/buffer exchanged into 20 mM phosphate buffer (pH 8.0). The Ni-sirohydrochlorin *a,c*-diamide was released from CfbB by adding 30 mL methanol and centrifuged for 30 min. The methanol was removed from the supernatant by rotary evaporation and the Ni-sirohydrochlorin *a,c*-diamide was redissolved in D_2O. The Ni-sirohydrochlorin *a,c*-diamide was lyophilized and redissolved in D_2O. Purified CfbCD was buffer exchanged into 100 mM Tris buffer (pH 8.0) containing 200 mM NaCl that was made with deuterium chloride and D_2O. Sodium dithionite and ATP were dissolved in D_2O. The 15,17^3-seco-F430-17^3-acid synthesis was performed by incubating 10 μM CfbC, 10 μM CfbD, 5 mM ATP, 3 mM sodium dithionite, and 30 mL of Ni-sirohydrochlorin *a,c*-diamide in 90 mL of 100 mM Tris-HCl, pH 8.0 buffer. The reaction was performed overnight at 37 °C in the glovebox. The products were extracted by adding 100 μL methanol to 100 μL of the assay mixture and centrifuged at 7,000 × g for 30 min. The supernatants were analyzed using HPLC. The synthesized deuterated 15,17^3-seco-F430-17^3-acid was extracted by adding 30 mL methanol and centrifuged for 30 min. The methanol was removed from the supernatant by rotary evaporation and the 15,17^3-seco-F430-17^3-acid redissolved in water. The 15,17^3-seco-F430-17^3-acid was then loaded onto a self-packaged C18 column to separate it from unreacted Ni-sirohydrochlorin *a,c*-diamide. Finally, the 15,17^3-seco-

F430-17[3]-acid was loaded onto a GE PD-10 Sephadex G-10 M column and eluted with water to remove small molecules.

3.2.8 HPLC and LC-MS analysis

HPLC analysis was performed on an Agilent 1260 Infinity Quaternary LC System with a Diode Array Detector (DAD) VL+. The column was an Agilent Poroshell 120 EC-C18 (4.6 × 150 mm, 2.7 μm) reverse phase column. The Agilent OpenLAB ChemStation Edition software was used for data analysis. The mobile phase consisted of water with 0.5 % formic acid (solvent A) and acetonitrile with 0.5 % formic acid (solvent B). The chromatographic method for characterization of the coenzyme F430 biosynthetic reactions consisted of the following gradient: 0 % B for 2 min, linear gradient to 20 % B over 3 min, 20 % B for 5 min, linear gradient to 25 % B over 5 min, 25 % B for 5 min, linear gradient to 30 % B over 5 min, linear gradient to 100 % B over 2 min, and 100 % B for 3 min. A flow rate of 1.0 mL/min was used and the trace was acquired using detection at 400 nm.

MS analysis of the assays was performed at the Auburn University Department of Chemistry and Biochemistry Mass Spectrometry Center on an ultra-performance LC system (ACQUITY, Waters Corp., USA) coupled with a quadrupole time-of-flight mass spectrometer (Q-TOF Premier, Waters Corp., USA) with electrospray ionization (ESI) in positive mode using MassLynx software (v4.1). Injections of the samples were made directly into the mass spectrometer or onto a C4 column (Aeris™ 3.6 μm C4 200 Å, 50 × 2.1 mm, Phenomenex) with a 300 μL/min flow rate of mobile phase solution A (0.1 % formic acid in 95 % water and 5 % acetonitrile) and solution B (0.1 % formic acid in 95 % acetonitrile and 5 % water) using the following gradient: 0 % B for 2 min, linear gradient to 100 % B over 11 min, 100 % B for 1 min, linear gradient to 0 % B over 1 min, and 0 % B for 3 min. The column temperature was held at

40 °C and the UV detector monitored at absorption at 420 nm. The capillary voltage was set at 3.1 kV, the sample cone voltage was 15 V, and the extraction cone was 4.0 V. The source and desolvation temperature were maintained at 95 and 300 °C, respectively, with the desolvation gas flow set to 600 L/h. The mass spectrometry scan was 0.5 s long from 50 to 2,000 *m*/*z* with a 0.02 s interscan delay using the centroid data format.

3.2.9 NMR analysis

All the synthesized products were lyophilized and redissolved in D_2O before NMR analysis. Proton nuclear magnetic resonance (^{1}H-NMR), ^{1}H-^{13}C heteronuclear single quantum coherence (HSQC), and ^{1}H-^{15}N HSQC spectra were acquired on either a 400 MHz or 600 MHz AV Bruker NMR spectrometer at the Auburn University NMR facility.

3.3 Results

3.3.1 Expression and purification

The HemC, HemD, SirA, SirC, CfbA, CfbB, CfbC, CfbD, and NirDL enzymes were expressed in *Escherichia coli* as His_6-tagged fusion proteins. HemD, SirA, SirC, CfbA, CfbB, and NirDL were purified under aerobic conditions from the cell extracts using a Ni-affinity column, while HemC, CfbC, and CfbD were purified anaerobically. The His-tag/SUMO-tag of CfbA, NirDL, CfbC, and CfbD were cleaved and the CfbC and CfbD enzymes were reconstituted with iron and sulfide. The purity of the proteins was determined by sodium dodecyl-sulfate polyacrylamide gel electrophoresis (SDS-PAGE). HemC is 33.9 kDa with His-tag, HemD is 27.8 kDa with His-tag, SirA is 28 kDa with His-tag, SirC is 24 kDa with His-tag, CfbA is 13.8 kDa, CfbB is 53 kDa with His-tag, CfbC is 28.5 kDa, CfbD is 41 kDa, and NirDL is 39 kDa (Fig 3.5).

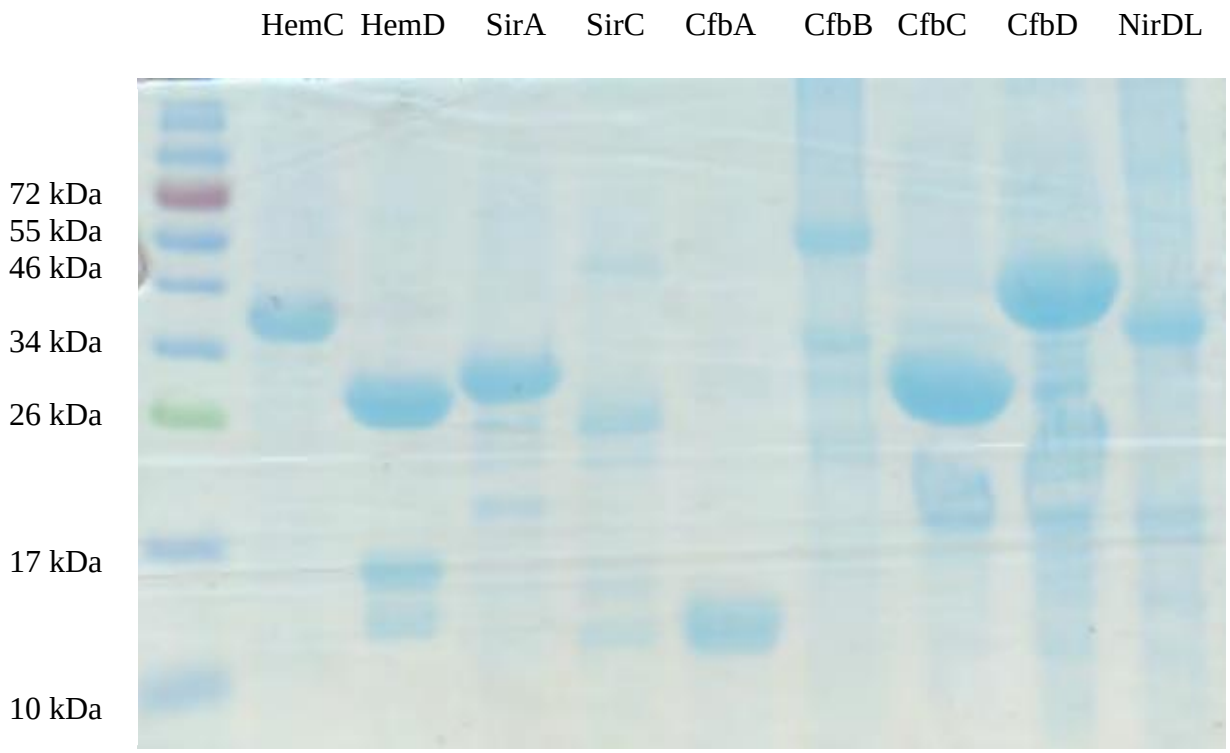


Figure 3.5 SDS-PAGE gel of HemC, HemD, SirA, SirC, CfbA, CfbB, CfbC, CfbD, and NirDL.

3.3.2 Crystallographic studies of the CfbCD complex

CfbC is a homolog of the Fe protein of nitrogenase and thus has two proposed ATP binding sites and a [4Fe-4S] cluster. Dithionite was incubated with CfbC to reduce the iron-sulfur cluster and a stable/nonhydrolyzable ATP analogue was added to CfbC in an attempt to trap CfbC in complex with CfbD. CfbC was then incubated with CfbD with or without the substrate Ni-sirohydrochlorin *a,c*-diamide. FPLC analysis showed that CfbCD with or without substrate had four major peaks (Fig 3.6). The proteins in the collected fraction 6 through 11 were concentrated and analyzed using SDS-PAGE (Fig 3.7) (Fig 3.8). For CfbCD without substrate, fractions 6 to 8 contained only CfbD; fraction 9 contained the CfbCD complex, and fraction 10 and 11 contained CfbC. For CfbCD with substrate, only fraction 6 contained CfbD. Fractions 7 through 9 contained the CfbCD complex, while fraction 10 contained CfbC and fraction 11 had a purple

color, suggesting it contained most of the free substrate. From the results, most of CfbD formed multimer aggregates, resulting in the earlier retention time. More CfbCD complex formed after adding Ni-sirohydrochlorin *a,c*-diamide, suggesting that the substrate helps CfbC and CfbD binding. In the absence of substrate, fraction 11 had high absorbance at 280 nm, suggesting that excess ATP and dithionite also eluted in this fraction. The CfbCD complex was collected for the crystallization experiments.

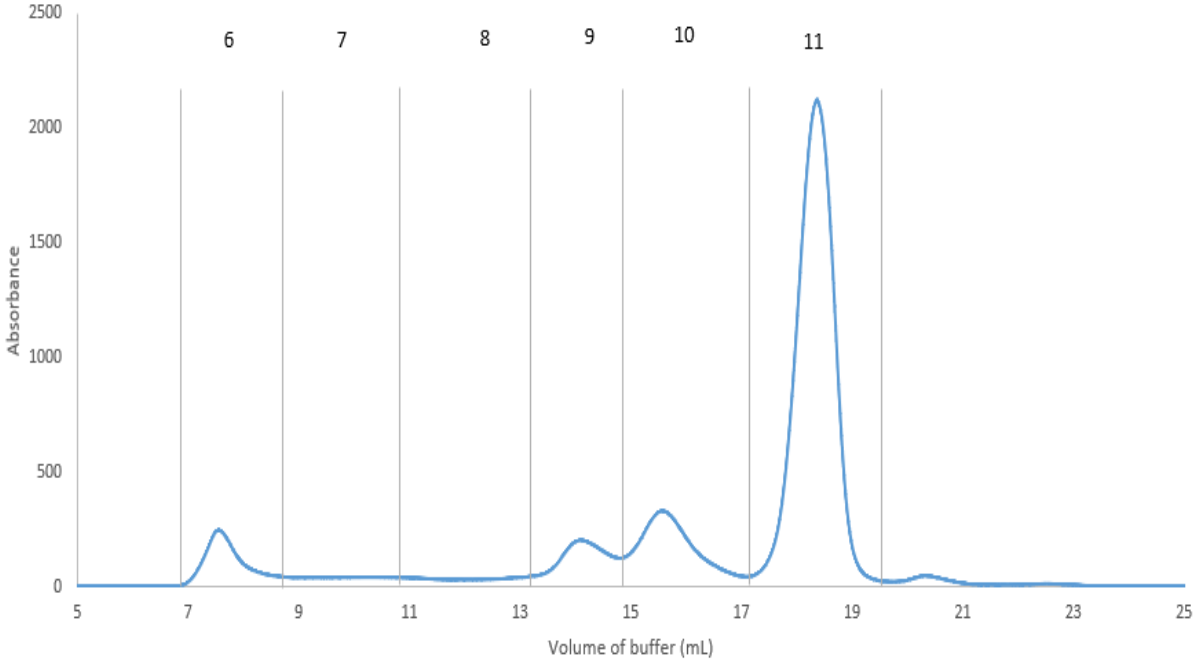


Figure 3.6 FPLC purification of the CfbCD complex. The CfbCD mixture contained CfbC, CfbD, adenosine 5′-(β,γ-imido)triphosphate lithium salt hydrate, sodium dithionite, and Ni-sirohydrochlorin *a,c*-diamide.

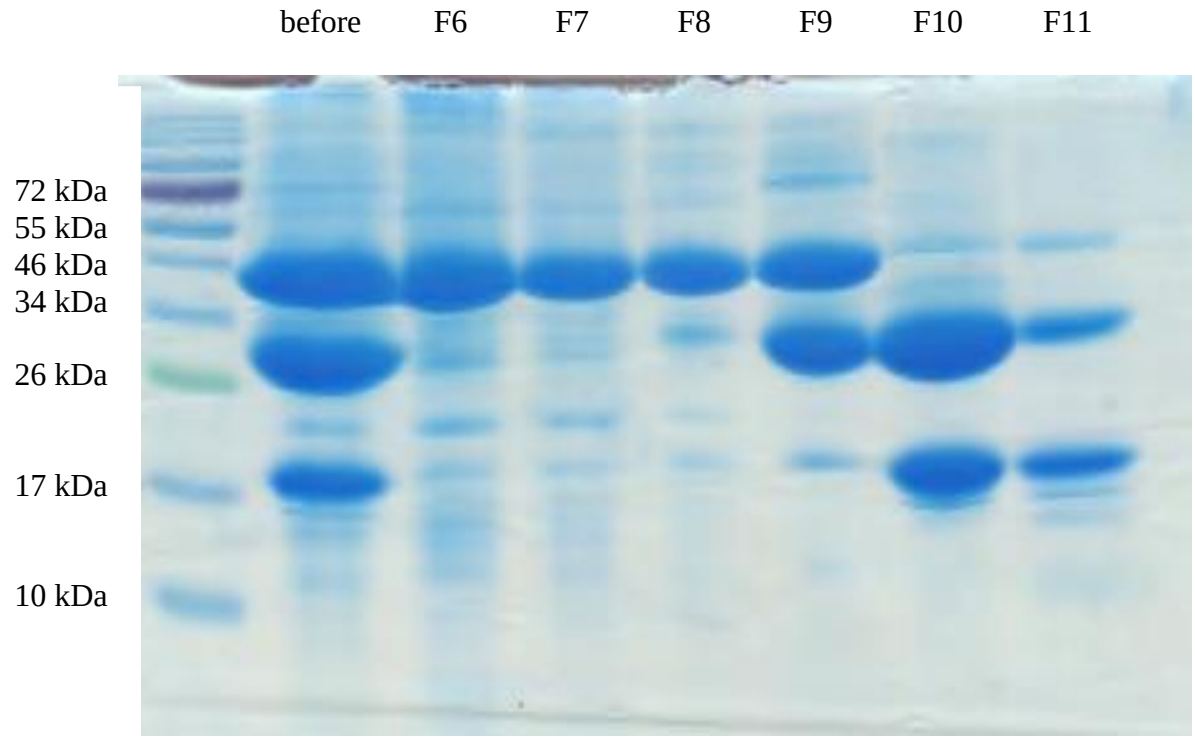


Figure 3.7 SDS-PAGE analysis of the CfbCD mixture without Ni-sirohydrochlorin *a,c*-diamide after FPLC purification. The CfbCD mixture contained CfbC, CfbD, adenosine 5′-(β,γ-imido)triphosphate lithium salt hydrate, and sodium dithionite.

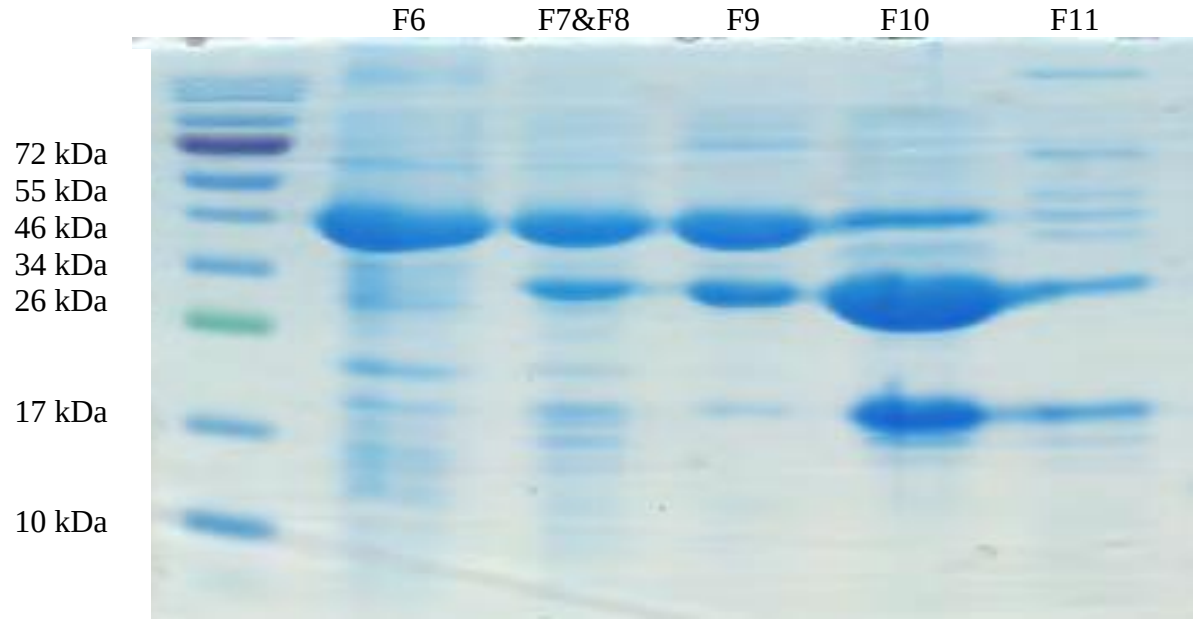


Figure 3.8 SDS-PAGE analysis of the CfbCD mixture with Ni-sirohydrochlorin *a,c*-diamide after FPLC purification. The CfbCD mixture contained CfbC, CfbD, adenosine 5′-(β,γ-imido)triphosphate lithium salt hydrate, sodium dithionite, and Ni-sirohydrochlorin *a,c*-diamide.

3.3.3 Synthesis of 15N-labeled intermediate of CfbCD catalysis

Glutamine is the physiological ammonia source for the CfbB reaction. To label the nitrogens on the *a*- and *c*-amide side chains of the coenzyme F430 biosynthetic intermediates, ^{15}N-labeled ammonium chloride was used as an alternative ammonia source for the CfbB reaction. HPLC analysis of the reaction showed that all the Ni-sirohydrochlorin was converted to Ni-sirohydrochlorin *a,c*-diamide (Fig 3.9).

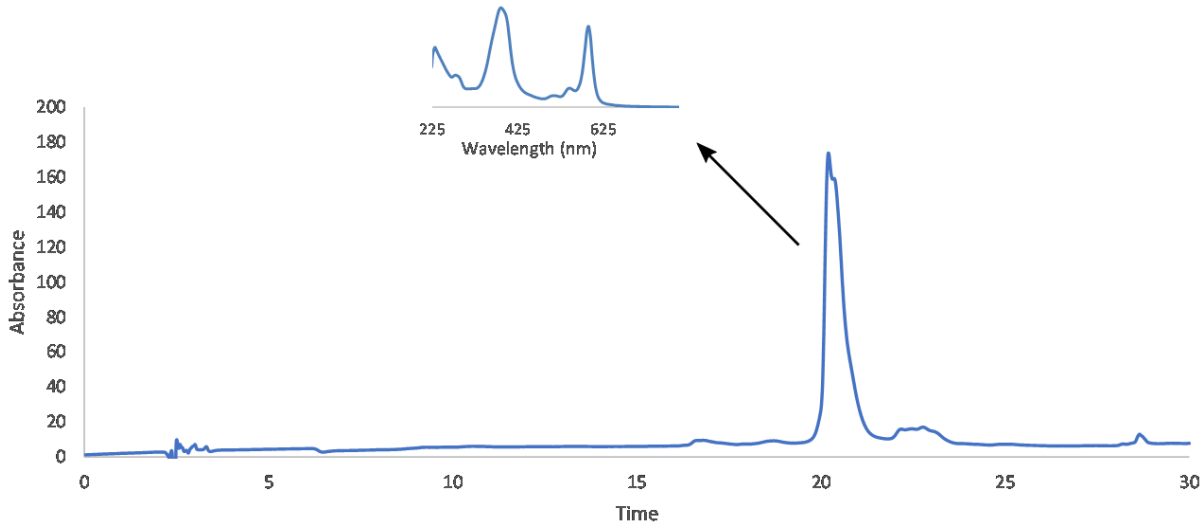


Figure 3.9 HPLC and UV-visibile spectroscopic analysis of the CfbB reaction using ^{15}N-labeled ammonium chloride as an ammonia source. The reaction contained 2.5 µM HemC, 2.5 µM HemD, 2.5 µM SirA, 2.5 µM SirC, 2.5 µM CfbA, 2.5 µM CfbB, 0.88 mM PBG, 2 mM SAM, 1 mM $NADP^+$, 4 mM $MgCl_2$, 200 µM $NiCl_2$, 4 mM nitrogen-15 ammonia, and 5 mM ATP in 30 mL of 100 mM Tris-HCl, pH 8.0 buffer.

NMR analysis requires concentrated and pure product. To remove SAM, $NADP^+$, $MgCl_2$, $NiCl_2$, NH_4Cl, ATP, glycerol, and Tris in the reaction, the Ni-sirohydrochlorin *a,c*-diamide bound with proteins were concentrated and buffer exchanged with sodium phosphate buffer (pH 8). Then, methanol was added to precipitate the proteins (HemC, HemD, SirA, SirC, CfbA, and CfbB). Purified Ni-sirohydrochlorin *a,c*-diamide was lyophilized and redissolved in D_2O. The ^{1}H-^{15}N HSQC NMR spectra showed a broad nitrogen-15 peak at a chemical shift of 116 ppm

that was coupled to four protons (Fig 3.10). Based on the NMR spectrum from Warren's group, the protons with a chemical shift of 7.5 ppm and 6.7 ppm are coupled to the nitrogen on the *a*-amide side chain of Ni-sirohydrochlorin *a,c*-diamide, and the protons with a chemical shift of 7.0 ppm and 6.6 ppm should be on the *c*-amide side chain.[149] To further confirm the structure of Ni-sirohydrochlorin *a,c*-diamide, a ^{1}H-^{13}C HSQC NMR experiment was performed, and the result showed a pure Ni-sirohydrochlorin *a,c*-diamide spectrum consistent with what was observed by the Warren group (Fig 3.11).[149]

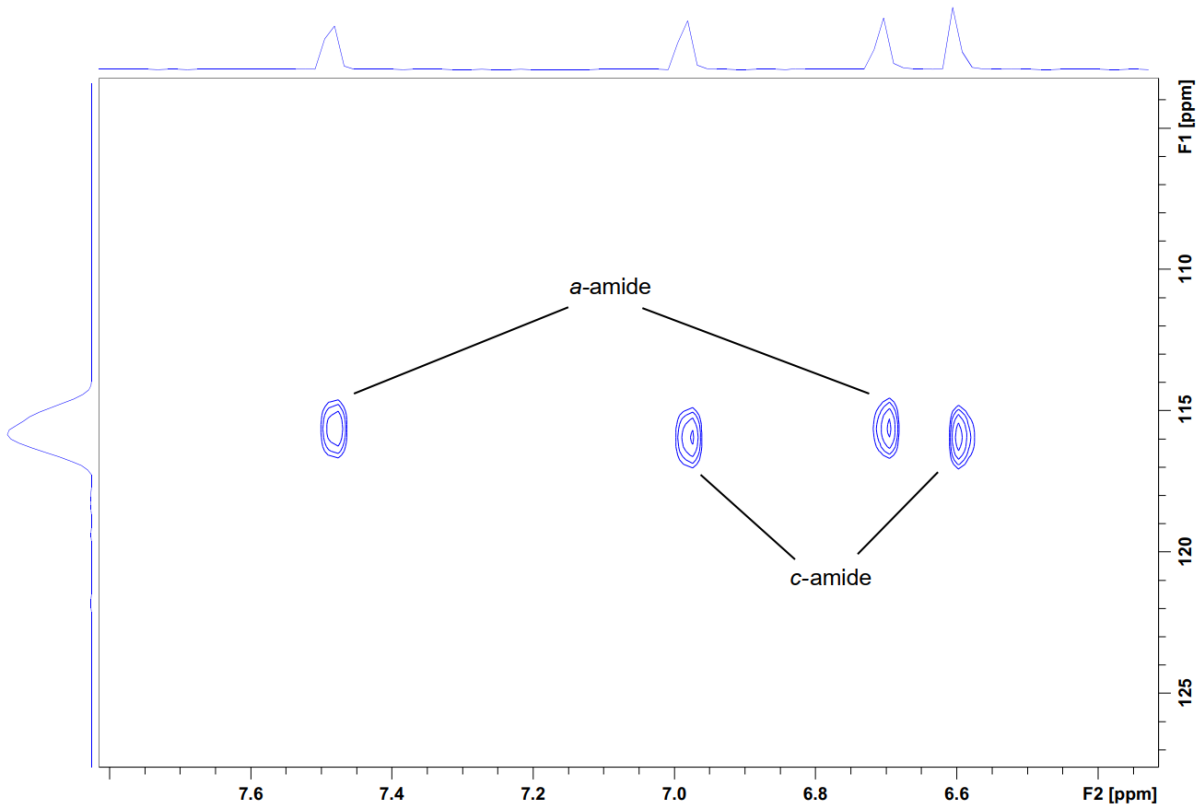


Figure 3.10 The ^{1}H-^{15}N HSQC NMR spectrum of ^{15}N-labeled Ni-sirohydrochlorin *a,c*-diamide.

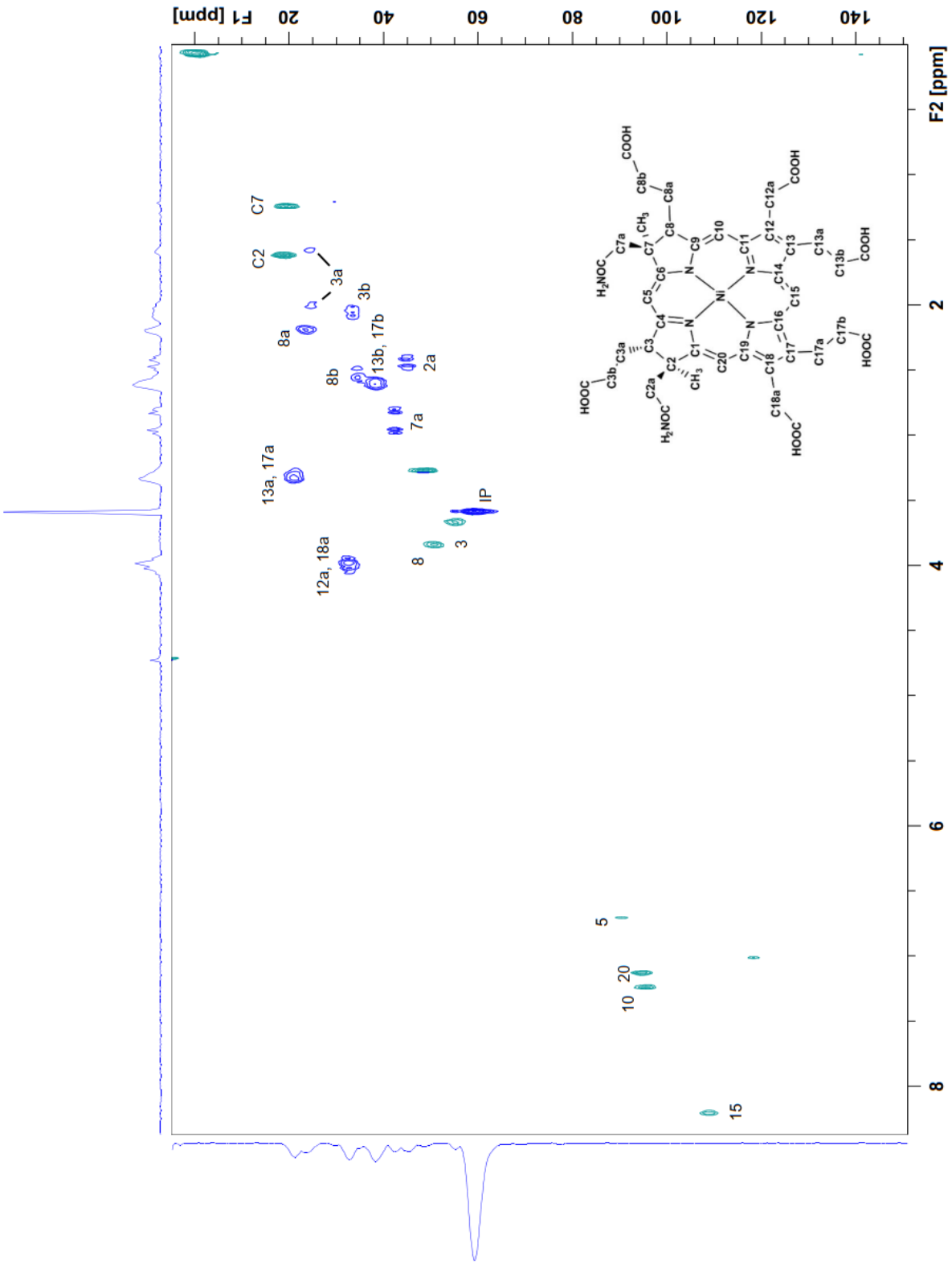


Figure 3.11 The ^{1}H-^{13}C HSQC NMR spectrum of ^{15}N-labeled Ni-sirohydrochlorin *a,c*-diamide.

A small amount of ^{15}N-labeled Ni-sirohydrochlorin *a,c*-diamide was used for CfbCD reactions to determine the time course of intermediate formation. The HPLC results showed the Ni-sirohydrochlorin *a,c*-diamide peak, which has a retention time of 18.9 min, decreased and a product peak with a retention time of 9.1 min increased with time (Fig 3.12). The UV-vis spectrum of the product showed a peak at 446 nm that increased in intensity for the first 5.5 h before giving way to a new peak at 428 nm after 6.5 h (Fig 3.13). Thus, the intermediate was formed in the first 5.5 h and then 15,17^3-seco-F430-17^3-acid started formating. The rest of the ^{15}N-labeled Ni-sirohydrochlorin *a,c*-diamide was used for a CfbCD reaction and quenched at 5 h. Both intermediate and unreacted substrate were purified following the protocol above. The intermediate was separated from substrate via a self-packaged C18 column. The yellow fraction was collected and analyzed using NMR. The ^{1}H-^{15}N HSQC NMR spectrum showed a nitrogen-15 peak at a chemical shift of 116 ppm that was coupled to two protons. Based on the NMR spectrum of Ni-sirohydrochlorin *a,c*-diamide and the spectrum of 15,17^3-seco-F430-17^3-acid from the Warren group, the protons with a chemical shift of 7.5 ppm and 6.7 ppm should be coupled to the nitrogen on the *a*-amide side chain.[149] Surprisingly, both the protons on the *c*-amide side chain were not observed, which is consistent with Ni(I) form seco-F430 and the lactam nitrogen is subjected to paramagnetic relaxation due to the closer proximity to the Ni ion (Fig 3.14).

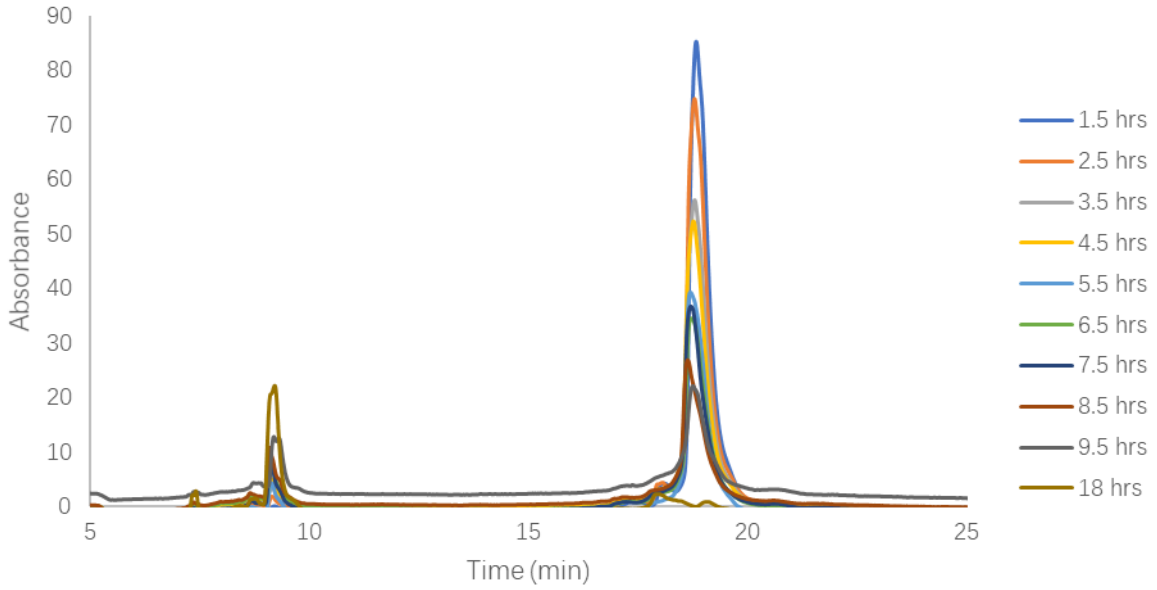


Figure 3.12 HPLC analysis of the time course of the CfbCD reaction. The reaction was quenched every hour from 1.5 h to 9.5 h, and 18 h after incubation, and analyzed by HPLC with detection at 400 nm.

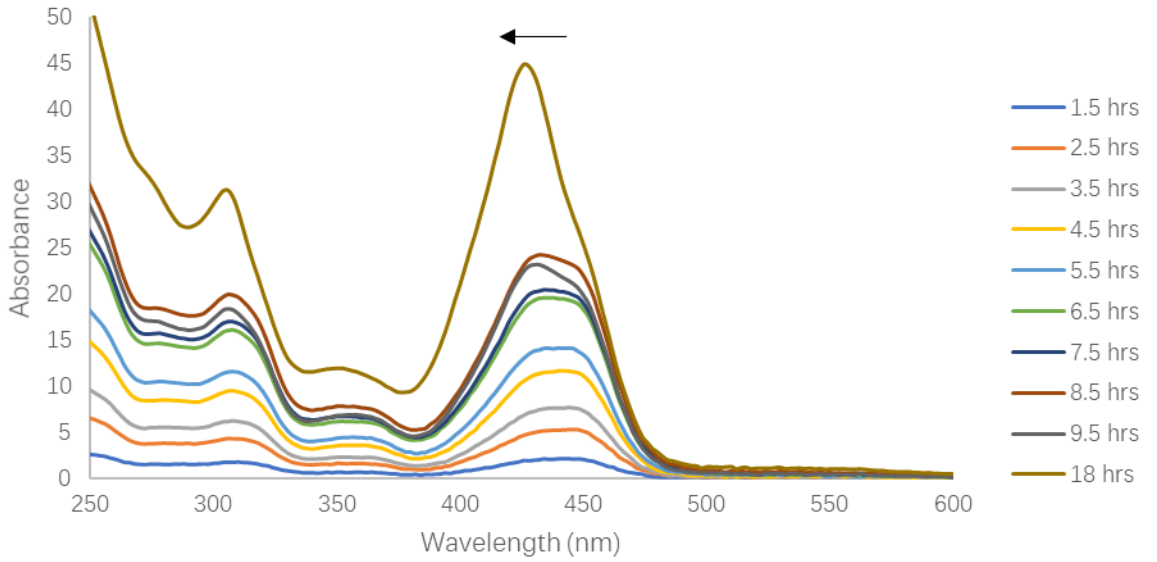


Figure 3.13 UV-visible spectra of the product peak of the CfbCD reaction having a HPLC retention time of 9.1 min.

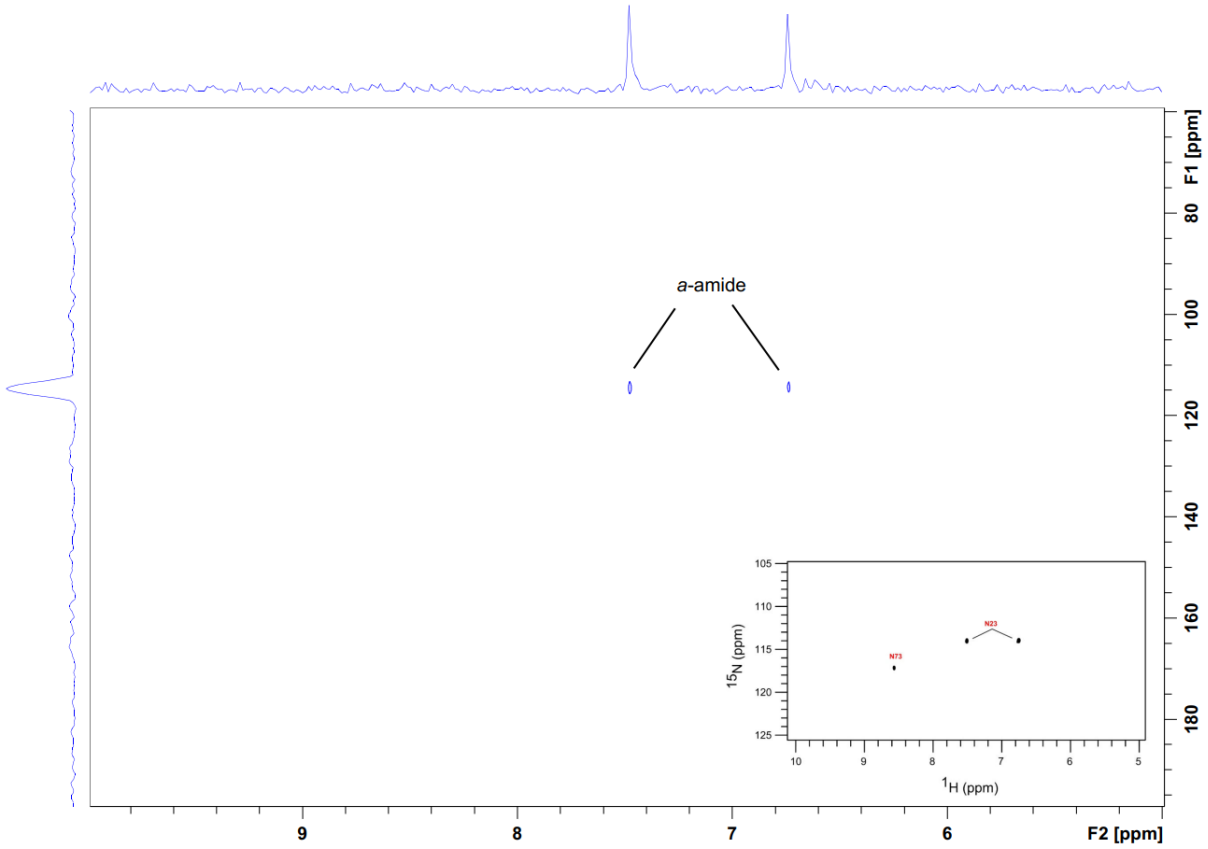


Figure 3. 6 The ^{1}H-^{15}N HSQC NMR spectrum of the intermediate of the CfbCD reaction. Insert is the reference spectrum of 15,17^{3}-seco-F430-17^{3}-acid from reference 148.

3.3.4 Synthesis of Ni-didecarboxysirohydrochlorin a,c-diamide

Ni-didecarboxysirohydrochlorin *a,c*-diamide was proposed as an alternative substrate that would stop the CfbCD reaction at the C19 protonation step and provide insight into the mechanism and order of proton addition. There are two steps to convert Ni-sirohydrochlorin to Ni-didecarboxysirohydrochlorin *a,c*-diamide, didecarboxylation of the C12 and C18 side chains, and amidation on the C2 and C7 side chains. NirDL from *Hydrogenobacter thermophilus* was expressed and purified in *E. coli* and catalyzes the didecarboxylation of siroheme during the biosynthesis of heme d_1. We tested to see whether NirDL exhibits substrate flexibility and could also catalyze the didecarboxylation of Ni-sirohydrochlorin or Ni-sirohydrochlorin *a,c*-diamide.

When Ni-sirohydrochlorin *a*,*c*-diamide was used as substrate for NirDL, HPLC and MS analysis did not show any new products were formed (data not shown). However, when Ni-sirohydrochlorin was used as substrate for NirDL, HPLC showed a new peak with a retention time of 28.4 min was formed. MS analysis of the new peak showed it had a mass of 845.22 *m/z*, which is consistent with an oxidized form of Ni-didecarboxysirohydrochlorin (Fig 3.16). These peaks were not observed in a control reaction lacking NirDL (Fig 3.15).

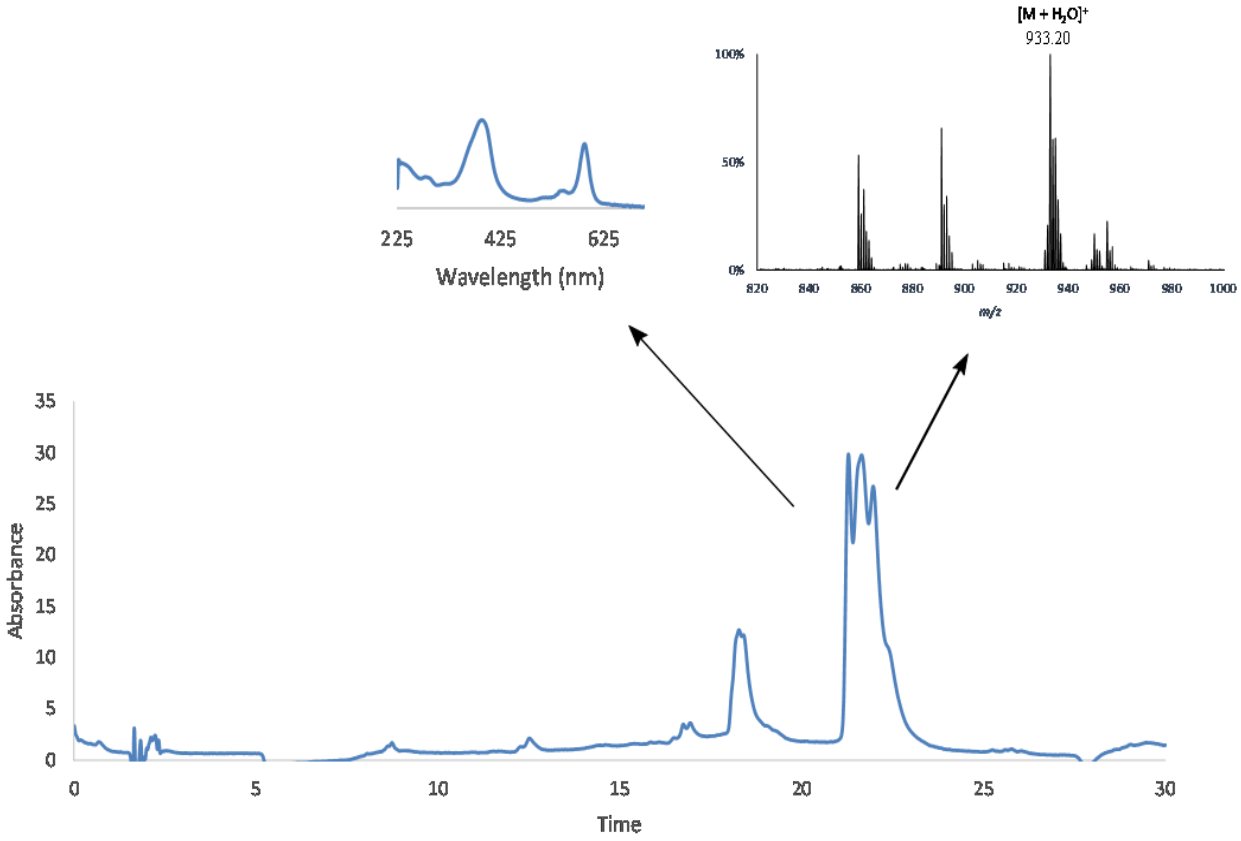


Figure 3.7 HPLC and MS analyses of the no enzyme control reaction of NirDL with Ni-sirohydrochlorin. The reaction contained 2.5 μM HemC, 2.5 μM HemD, 2.5 μM SirA, 2.5 μM SirC, 2.5 μM CfbA, 0.88 mM porphobilinogen (PBG), 2 mM SAM, 1 mM $NADP^+$, 4 mM $MgCl_2$, and 200 μM $NiCl_2$ in 1 mL of 100 mM Tris-HCl, pH 8.0 buffer.

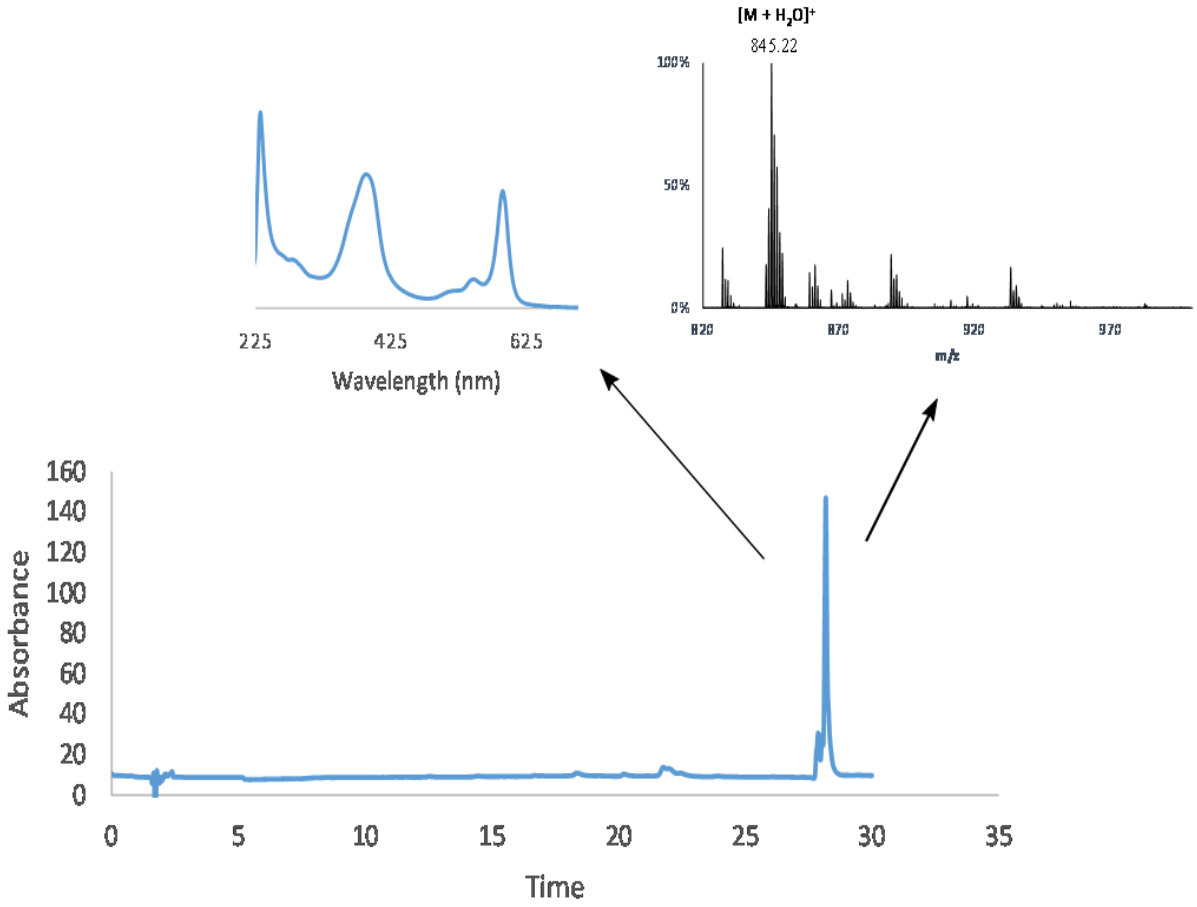


Figure 3.8 HPLC and MS analyses of the NirDL reaction with Ni-sirohydrochlorin. The reaction contained 10 μM NirDL, 2.5 μM HemC, 2.5 μM HemD, 2.5 μM SirA, 2.5 μM SirC, 2.5 μM CfbA, 0.88 mM porphobilinogen (PBG), 2 mM SAM, 1 mM $NADP^+$, 4 mM $MgCl_2$, and 200 μM $NiCl_2$ in 1 mL of 100 mM Tris-HCl, pH 8.0 buffer**.**

To synthesize Ni-didecarboxysirohydrochlorin *a,c*-diamide, CfbB, ATP, and glutamine were added to the Ni-didecarboxysirohydrochlorin assay. HPLC showed a major peak with a retention time of 18.7 min (Fig 3.17). MS analysis showed two major peaks with a mass difference of 88.01 *m/z*, which is consistent with a mixture of oxidized forms of Ni-didecarboxysirohydrochlorin *a,c*-diamide and Ni-sirohydrochlorin *a,c*-diamide (Fig 3.18). The resulting Ni-didecarboxysirohydrochlorin *a,c*-diamide can be used for future mechanistic studies of CfbCD catalysis.

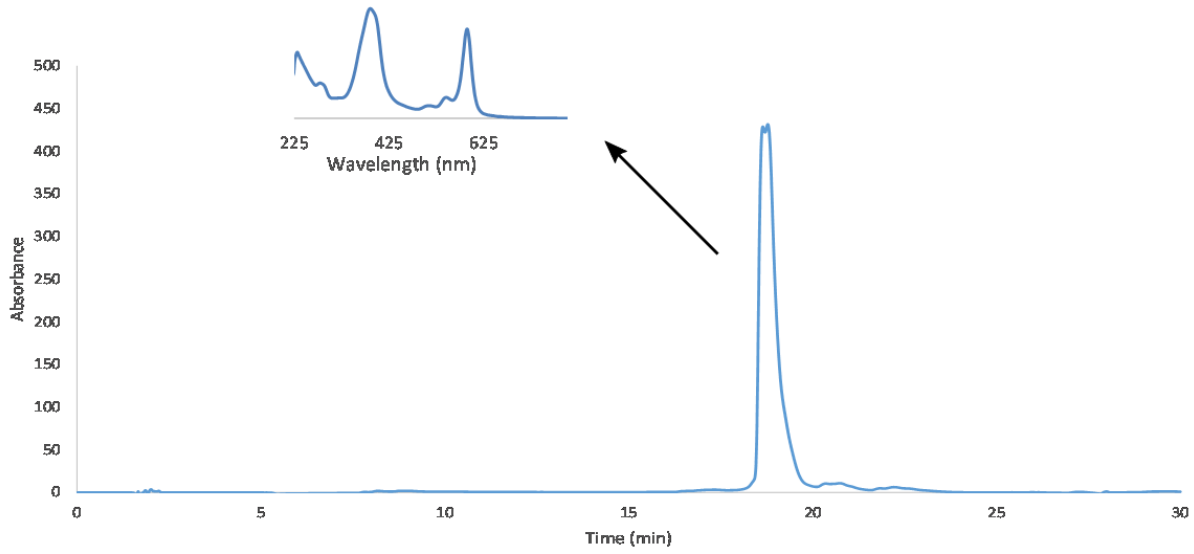


Figure 3.9 HPLC and UV-visible spectroscopic analyses of the CfbB reaction with Ni-didecarboxysirohydrochlorin. The reaction contained 2.5 μM CfbB, 5 mM ATP, 3 mM glutamine, 10 μM NirDL, 2.5 μM HemC, 2.5 μM HemD, 2.5 μM SirA, 2.5 μM SirC, 2.5 μM CfbA, 0.88 mM porphobilinogen (PBG), 2 mM SAM, 1 mM $NADP^+$, 4 mM $MgCl_2$, and 200 μM $NiCl_2$ in 1 mL of 100 mM Tris-HCl, pH 8.0 buffer.

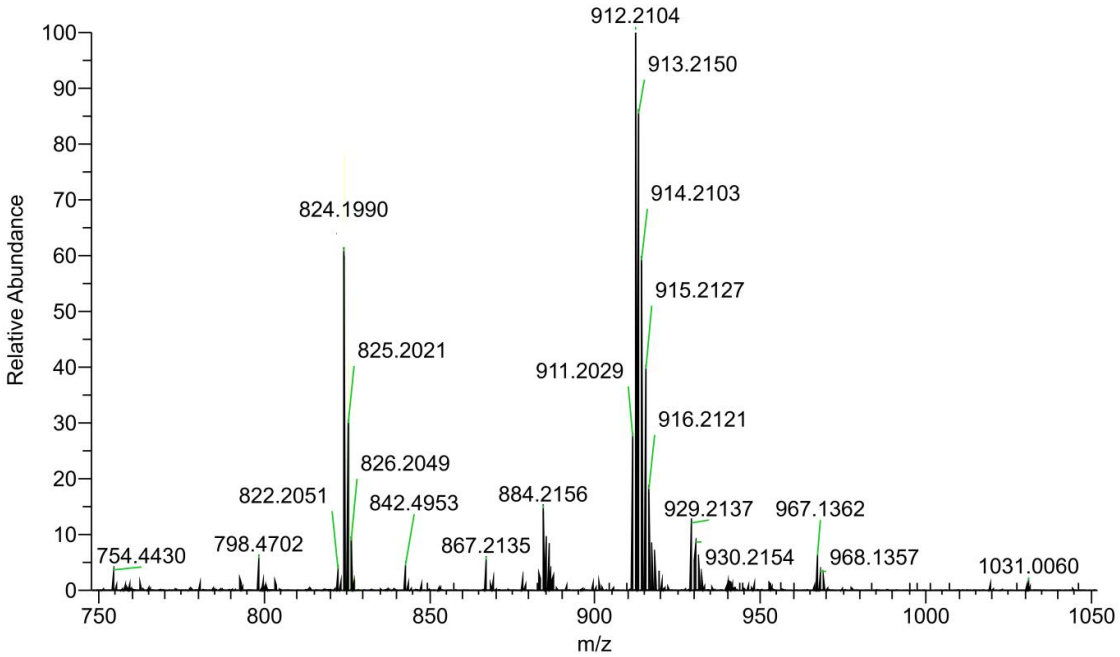


Figure 3.10 MS analysis of the CfbB reaction with Ni-didecarboxysirohydrochlorin.

3.3.5 Stereochemical study of 15,173-seco-F430-173-acid

A high concentration CfbCD reaction was performed in D_2O to produce enough 15,17^3-seco-F430-17^3-acid for NMR analysis. HPLC analysis showed pure 15,17^3-seco-F430-17^3-acid was formed. The proteins were removed by adding methanol and the 15,17^3-seco-F430-17^3-acid was purified using a self-packaged C18 column in the glovebox. The yellow color fraction, which is 15,17^3-seco-F430-17^3-acid, was separated from the purple color fraction, which is the unreactive substrate, Ni-sirohydrochlorin *a,c*-diamide. Unfortunately, 15,17^3-seco-F430-17^3-acid was not retained by the column, so it was not purified away from all of the small molecules (Fig 3.19). The ^{1}H-NMR spectrum showed a high concentration contaminant. Comparison with a reference from the database suggests that the large contaminant was glycerol, which was used to stabilize the proteins (Fig 3.20).

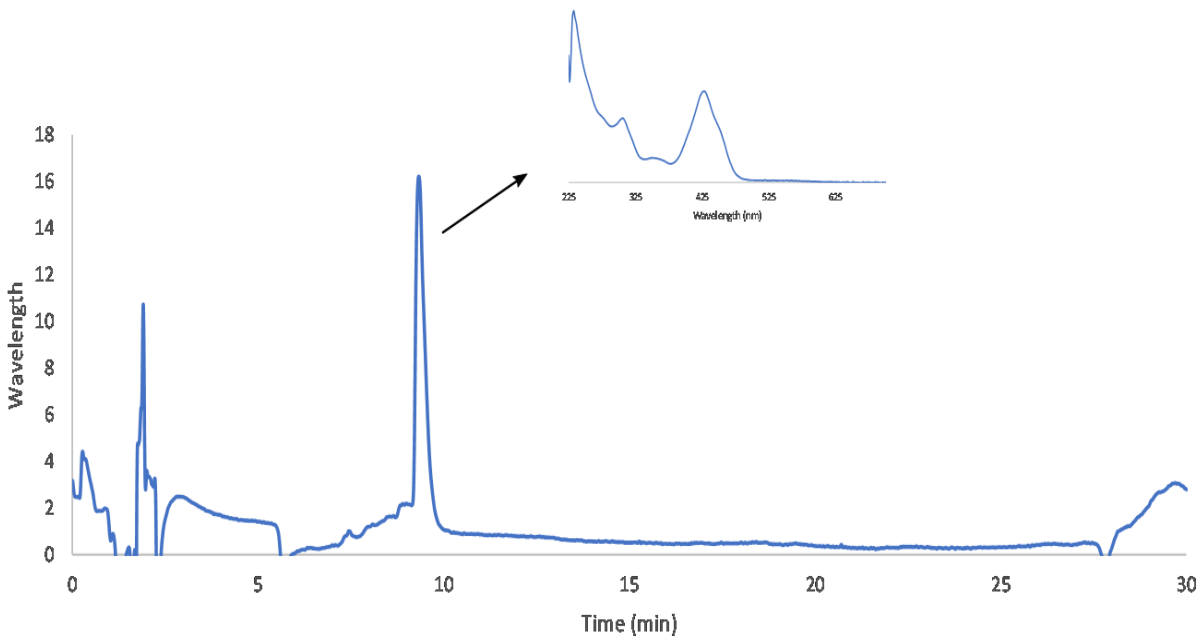


Figure 3.11 HPLC and UV-visible spectroscopic analysis of a large-scale CfbCD reaction carried out in D_2O. The reaction contained 10 µM CfbC, 10 µM CfbD, 5 mM ATP, 3 mM sodium dithionite, and 6 mM Ni-sirohydrochlorin *a,c*-diamide in 90 mL of 100 mM Tris-HCl, pH 8.0 buffer.

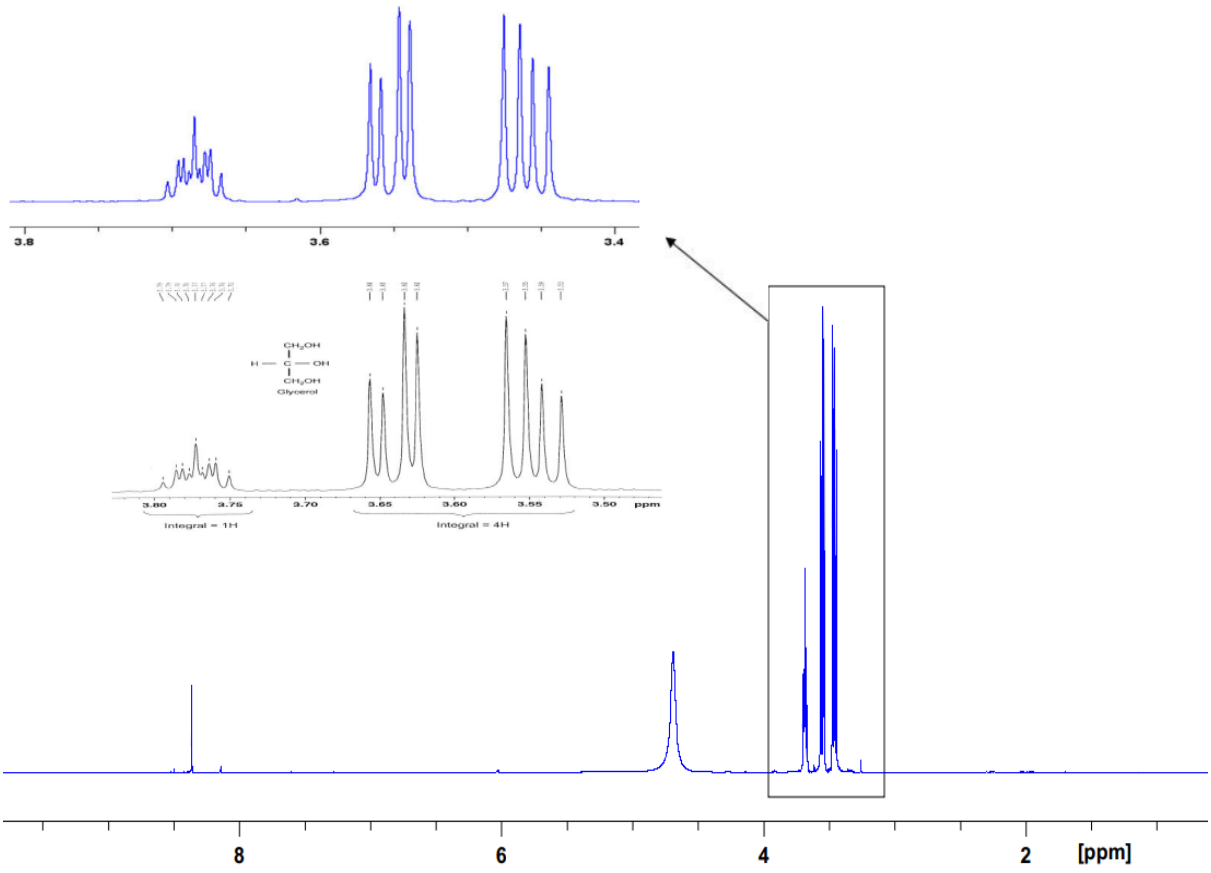


Figure 3.12 ^{1}H-NMR spectrum of a large-scale CfbCD reaction carried out in D_2O, showing a large amount of contaminating glycerol.

3.4 Discussion

The key step in coenzyme F430 biosynthesis was found to involve an unprecedented reductive cyclization reaction that converts Ni-sirohydrochlorin *a*,*c*-diamide to 15,17^{3}-seco-F430-17^{3}-acid. This reaction involves a 6-electron reduction of the isobacteriochlorin ring system, cyclization of the *c*-acetamide side chain to form a γ-lactam ring, and the formation of 7 stereocenters, which is catalyzed by a primitive homolog of nitrogenase (CfbCD). Nitrogenase is an important enzyme that catalyzes reduction of dinitrogen to ammonia. Homologs of nitrogenase are also involved in the biosynthesis of the photosynthetic pigments chlorophyll and bacteriochlorophyll. Thus, understanding the mechanism and structure of CfbCD will help to

further understand the mechanism and evolution of the nitrogenase superfamily. A mechanism and structure of CfbCD was proposed based on the known stereochemistry of 15,17^3-seco-F430-17^3-acid and homologs of CfbCD: 1) docking of the reduced ATP-bound CfbC onto the substrate-bound CfbD, 2) a conformationally gated electron transfer, 3) ATP hydrolysis and release of inorganic phosphate, 4) dissociation of the oxidized ADP-bound CfbC, and 5) reduction of CfbC by ferredoxin and nucleotide exchange. This cycle is repeated multiple times to affect the reduction of Ni-sirohydrochlorin *a*,*c*-diamide. However, the exact number of electrons transferred during the CfbCD reaction is unknown, as is whether CfbCD catalyzes the γ-lactamization reaction or if this process occurs non-enzymatically.

For the structural study of the CfbCD complex, we co-crystallized CfbC, CfbD, and the substrate Ni-sirohydrochlorin *a*,*c*-diamide. A nonhydrolyzable ATP analogue was used to stabilize CfbCD complex and prevent it from dissociating. Dithionite was used to reduce the iron-sulfur cluster of CfbC. The CfbCD complex was purified using FPLC and confirmed by SDS-PAGE. However, the CfbCD complex formation was not efficient. Most of the CfbC, CfbD, and substrate eluted separately. One possibility is that CfbD was aggregated since it eluted earlier than the CfbCD complex. Not enough active CfbD may have prevented CfbC and the substrate from binding properly and caused them to elute separately. Buffers and incubation conditions should be optimized in future experiments. Another possibility is that the concentration of dithionite, nonhydrolyzable ATP, and substrate were not optimized. More experiments will be needed to find the best conditions for CfbCD complex formation. The CfbCD complex was concentrated and crystallized. Unfortunately, we have not yet found conditions that yield crystals of the CfbCD complex. A more concentrated and purer sample of the CfbCD complex will be needed to identify and optimize crystallization conditions.

An initial product of the CfbCD reaction was reported by the Warren group. This initial product was assumed to be Ni-hexahydrosirohydrochlorin *a,c*-diamide, which lacks the γ-lactam ring and implies that the cyclization occurs spontaneously or requires an unidentified enzyme. However, the exact structure of this species has yet to be confirmed.[149] Other possibilities are that the initial product of the CfbCD reaction is the Ni(I) or Ni(II)-hydride form of 15,17^3-seco-F430-17^3-acid, which slowly oxidizes or evolves hydrogen, respectively. If this is the case, in addition to facilitating cyclization, CfbCD catalyzes the addition of either 7 electrons and 7 protons or 8 electrons and 8 protons to Ni-sirohydrochlorin *a,c*-diamide. The synthesis of concentrated and purified ^{15}N-labeled Ni-sirohydrochlorin *a,c*-diamide was confirmed using NMR. Meanwhile, a time course of the CfbCD reaction confirmed the initial product was observable at 5.5 h. The initial product from a large-scale CfbCD reaction was purified and concentrated from a 5-6 h quench. Surprisingly, the peak corresponding to the *c*-amide side chain of the initial product was not observed in the ^{1}H-^{15}N HSQC NMR spectra. A ^{1}H-^{15}N HSQC NMR experiment with 15,17^3-seco-F430-17^3-acid needs to be performed to confirm the chemical shift of the *c*-amide side chain. The purification of the initial product needs to be started from 4.5 h to ensure no 15,17^3-seco-F430-17^3-acid formation occurs during the purification step and the experiment needs to be repeated to confirm the result. If the ^{1}H-^{15}N HSQC NMR spectrum of 15,17^3-seco-F430-17^3-acid shows the peak corresponding to the *c*-amide side chain but the initial product still does not, the initial product might be a novel product. This may suggest that the Ni of initial product might be in the +1 (or +3) oxidation state and the lactam nitrogen is subjected to paramagnetic relaxation due to the closer proximity to the Ni ion. Electron paramagnetic resonance (EPR) spectroscopy could be used to confirm this hypothesis.

Analysis of the stereochemistry and pseudosymmetry of 15,17^3-seco-F430-17^3-acid suggests that the protons could be added *anti* (one above and one below the ring) to each of the four pyrroles within an active site containing 4-fold symmetry (e.g., at the tetrameric interface of a $(CfbD)_4$ core). To test these hypotheses, we carried out a CfbCD reaction in D_2O and used NMR to determine the positions of deuteron addition to Ni-sirohydrochlorin *a,c*-diamide. The two positions that are non-evident in the structure of 15,17^3-seco-F430-17^3-acid are the two (α and δ) methylene bridges, and *anti* addition predicts protons/deuterons will be added to the 'bottom' face of the ring on the side opposite to the γ-lactam ring. Unfortunately, purity and concentration are two difficulties for NMR analysis, and 15,17^3-seco-F430-17^3-acid was not purified sufficiently using C18 column chromatography. One possibility to aid in 15,17^3-seco-F430-17^3-acid purification is to concentrate and buffer exchange all the proteins into buffer lacking the glycerol contaminant or to incubate 15,17^3-seco-F430-17^3-acid with a binding protein (e.g., an inactive variant of CfbE) that will allow buffer exchange of the CfbCD reaction and removal of small molecules prior to purification. These experiments, together with a study of the reaction products of the CfbCD reaction with the substrate analog Ni-didecarboxysirohydrochlorin *a,c*-diamide, should provide great insight into the unprecedented reaction catalyzed by this primitive homolog of nitrogenase.

Chapter Four: A noncanonical heme oxygenase specific for the degradation of *c*-type heme

4.1 Introduction

Iron is an essential element for life. Organisms have therefore developed sophisticated machinery for obtaining, transporting, and storing iron.[116] Iron acquisition is particularly important for bacterial pathogens to establish and maintain infections.[150] Many pathogenic bacteria synthesize and secrete siderophores, high-affinity iron chelators that scavenge free ferric iron.[151] The Fe(III)-siderophore complex is then imported into the cell, where the iron is released (e.g., via reduction to the less tightly bound ferrous state).[152] However, in many hosts, including humans, iron is predominantly found complexed with protoporphyrin IX in the form of heme. Thus, bacteria also typically contain transport systems to acquire heme from the environment.[153] A dedicated enzyme, heme oxygenase (HO), is then required to liberate the tightly bound iron from heme under aerobic or microaerophilic conditions.[154]

In addition to serving as the substrate of HO, heme is utilized as a cofactor that facilitates its own oxidative cleavage.[155] There are several distinct groups of HO, which differ in their structural folds and in the regiospecificity of their ring cleavage reactions (Fig 4.1).[156] The first HO, HO-1, was isolated from rat microsomes in the late 1960s and is a monomeric α-only protein.[157-159] HO-1 and most other canonical HOs from this group (such as HmuO from *Corynebacterium diphtheriae* and HemO from *Neisseria meningitidis*) cleave heme at the α-meso position, releasing ferrous iron and producing biliverdin IXα (α-biliverdin) and carbon monoxide as coproducts (Fig 4.1).[160, 161] However, the HemO homolog PigA from *Pseudomonas aeruginosa* was shown to cleave heme at both the β- and δ-meso positions to yield a mixture of β- and δ-biliverdins (Fig 4.1).[162]

The second group of HOs are structurally and functionally related to the iron-regulated surface determinants protein, IsdG, from *Staphylococcus aureus*.[163] IsdG belongs to the antibiotic biosynthesis monooxygenase (ABM) superfamily and contains a ferredoxin-like fold.[164] In contrast to canonical HOs, IsdG cleaves heme at the β- or δ-meso positions to form a mixture of β- and δ-staphylobilins and formaldehyde (Fig 4.1).[165, 166] The structure of staphylobilin differs from that of the corresponding biliverdin by the presence of an oxo group at the unopened β- or δ-meso position. Another member of this group, MhuD from *Mycobacterium tuberculosis*, cleaves heme at the α-position; however, the meso carbon is retained in the product as a formyl group, yielding a mixture of mycobilins *a* and *b* (Fig 4.1).[167] Mycobilin *a* and *b* also contain an oxo group at the β- and δ-meso position, respectively.

The third group of HOs contain tandem repeats of a motif named after its founding member, HemS.[168] The HemS homolog HmuS from *Yersinia pseudotuberculosis* was found to convert heme to a mixture of β- and δ-biliverdins, analogous to the reaction catalyzed by PigA (Fig 4.1).[169] In contrast, ChuS, a HemS homolog from *Escherichia coli* O157:H7, was shown to facilitate the peroxide-dependent conversion of heme to hematinic acid and a tripyrrole via ring cleavage at adjacent meso positions (i.e., γ/β and/or γ/δ) (Fig 4.1).[170]

The final group of HOs belong to the HugZ superfamily, as typified by the enzyme from *Helicobacter pylori*, which cleaves heme specifically at the δ-meso position and produces δ-biliverdin (Fig 4.1).[171] Homologs of HugZ are abundant in Proteobacteria and belong to a larger superfamily of dimeric split β-barrel enzymes that include flavin (FMN/FAD)- and deazaflavin (F_{420})-dependent oxidoreductases.[172] The two heme binding sites of HugZ are located at the intermonomer interface (Fig 4.1).[173] Notably, the active site is framed by the N-terminal α/β-

type domain, which interacts with both the α-meso edge of heme and the C-terminal loop containing a histidine residue that serves as the axial ligand to the heme substrate (Fig 4.1).[173]

Linking all known HOs across these structural and functional differences is a preferential degradation of free, *b*-type heme as their substrate.[156] Yet in addition to *b*-type heme, one of the most prevalent forms of heme found in nature is *c*-type heme, wherein the vinyl groups of the porphyrin macrocycle are covalently linked via thioether bonds to proteinaceous cysteine residues from a conserved CXXCH motif.[174] The terminal histidine residue of this *c*-type heme binding motif serves as the axial ligand to the heme iron.[174] HO-1 was previously shown to have some activity with heme *c* derivatives, although the catalytic efficiency was more than an order of magnitude less than with free heme.[175, 176] Thus, specific mechanisms for direct iron acquisition from the *c*-type heme pool remain unknown.

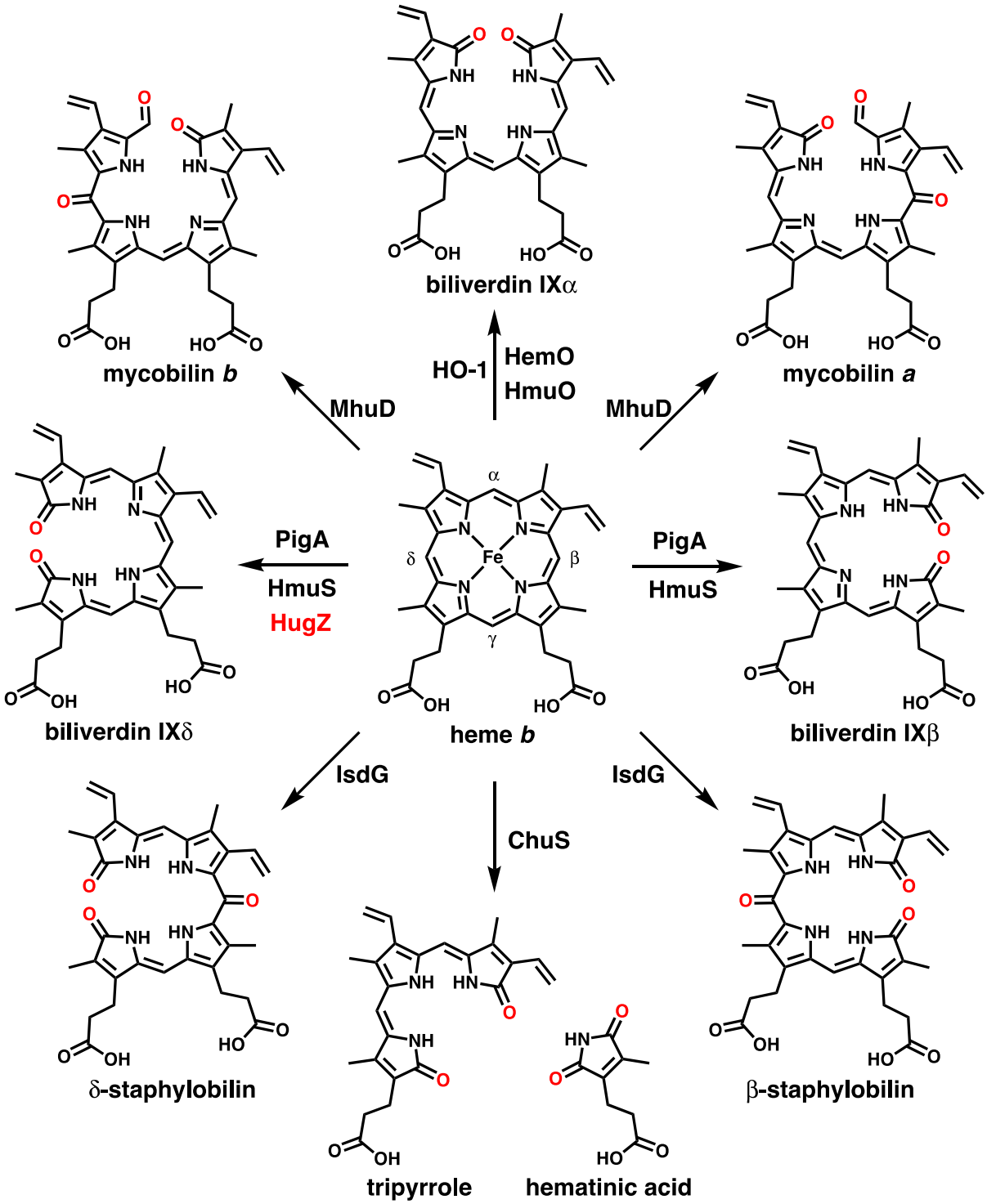


Figure 4.1 Products of free *b*-type heme degradation generated by the various groups of heme oxygenase (HO).

Paracoccus denitrificans is a metabolically versatile gram-negative alphaproteobacterium, capable of chemoautotrophic growth on hydrogen/thiosulfate and C_1 compounds (e.g., CO_2, methanol, methylamines) or heterotrophic growth using oxygen (aerobic respiration) or nitrate/nitrite (denitrification) as terminal electron acceptors.[177] In each of these metabolic pathways, *c*-type cytochromes (heme *c* binding proteins such as cytochrome c_{550}, bc_1, and cd_1) play critical roles as components of the relevant electron transport chains.[178]

The genome of *P. denitrificans* contains a cluster of genes that encode a homolog of the oligopeptide permease (Opp) system (Fig 4.2).[179] The Opp system is an ATP-binding cassette (ABC)-type transporter consisting of a membrane-associated substrate-binding lipoprotein (OppA), a hydrophobic transmembrane channel (OppBC), and two intracellular ATPase subunits (OppDF). OppA is a cluster C type substrate-binding protein, members of which have recently been shown to bind and facilitate the transport of heme in addition to short peptides.[180, 181] Within the *opp* operon is a gene (*pden_1323*) annotated by the National Center for Biotechnology Information (NCBI) as a pyridoxamine 5′-phosphate (PMP) oxidase-related FMN-binding protein (Fig 4.2). PMP oxidases belong to the same superfamily as HugZ-like HOs, which motivated investigation of Pden_1323 as a HO.[172]

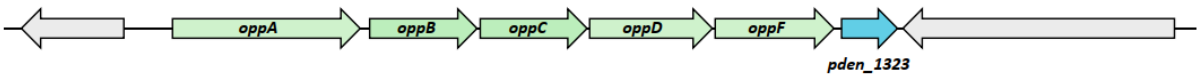


Figure 4.2 Gene cluster showing the location of *pden_1323* within the oligopeptide permease (*opp*) operon.

Here, we show that Pden_1323 is an unusual HO that shows specificity for the degradation of *c*-type heme instead of *b*-type heme, differentiating it from all previously

characterized HOs. The crystal structure of Pden_1323 contains only a partial active site relative to HugZ, suggesting that it can accommodate a *c*-type hemopeptide substrate that provides its own histidine ligand. Activity assays of Pden_1323 with the hemopeptide substrate microperoxidase-11 (MP-11) show *c*-type heme degradation to free iron, CO, and a peptide-linked β- and/or δ-biliverdin. The structural and kinetic results together demonstrate Pden_1323 as a novel *c*-type heme oxygenase (CHO), indicating the ability of *P. denitrificans* to acquire iron from direct oxidative cleavage of *c*-type cytochrome fragments containing heme.

4.2 Materials and Methods

4.2.1 Materials

Biliverdin hydrochloride, hemin chloride, imidazole, and sodium ascorbate were acquired from Alfa Aesar (Tewksbury, MA). MP-11, cytochrome *c*, catalase, myoglobin, and FMN were purchased from Sigma-Aldrich (St. Louis, MO). FAD was obtained from Tokyo Chemical Industry (TCI) (Tokyo, Japan). F_{420} and F_o were generous gifts from Dr. Kayunta Johnson-Winters (University of Texas, Arlington). The FerroZine iron reagent and neocuproine hemihydrate used for iron determination were from Acros Organics (Fair Lawn, NJ). LB media was from Becton, Dickinson and Company (BD) (Franklin Lakes, NJ), while buffer components were from VWR Chemicals BDH® (Radnor, PA). Primers were ordered from Sigma-Aldrich and the synthesized *pden_1323* gene was purchased from GeneArt (Waltham, MA). Phusion High-Fidelity PCR Kit, T4 DNA ligase, and restriction endonucleases were from New England Biolabs (Ipswich, MA). Gel extraction, PCR cleanup, and mini prep kits were purchased from Omega Bio-Tek (Norcross, GA). Profinity IMAC resin was obtained from Bio-Rad (Hercules, CA).

4.2.2 Vector construction

The *pden_1323* gene from *Paracoccus denitrificans* PD1222 was codon optimized for expression in *Escherichia coli* with NdeI and XhoI restriction sites included at the 5′ and 3′ ends, respectively. The vector containing *pden_1323*, as well as an empty pET-28b(+) vector, were digested with these restriction enzymes. The *pden_1323* product was purified via a Lonza FlashGel™ DNA Cassette (Basel, Switzerland) and recovered using an Omega Bio-tek E.Z.N.A.® Gel Extraction Kit (Norcross, GA), while the linearized pET-28b(+) vector was purified using an Omega Bio-tek E.Z.N.A.® Cycle Pure Kit. The *pden_1323* insert and the digested pET-28b(+) vector were then ligated together using T4 DNA Ligase (New England Biolabs, Ipswich, MA) using the manufacturer's protocol. The ligated mixture was then transformed into *E. coli* TOP10 cells and plated on Luria-Bertani (Luria-Bertani) agar plates containing 50 μg/mL kanamycin. Colonies were picked, grown in liquid LB medium, and plasmid preparations were made using the Omega Bio-tek E.Z.N.A.® Plasmid Mini Kit.

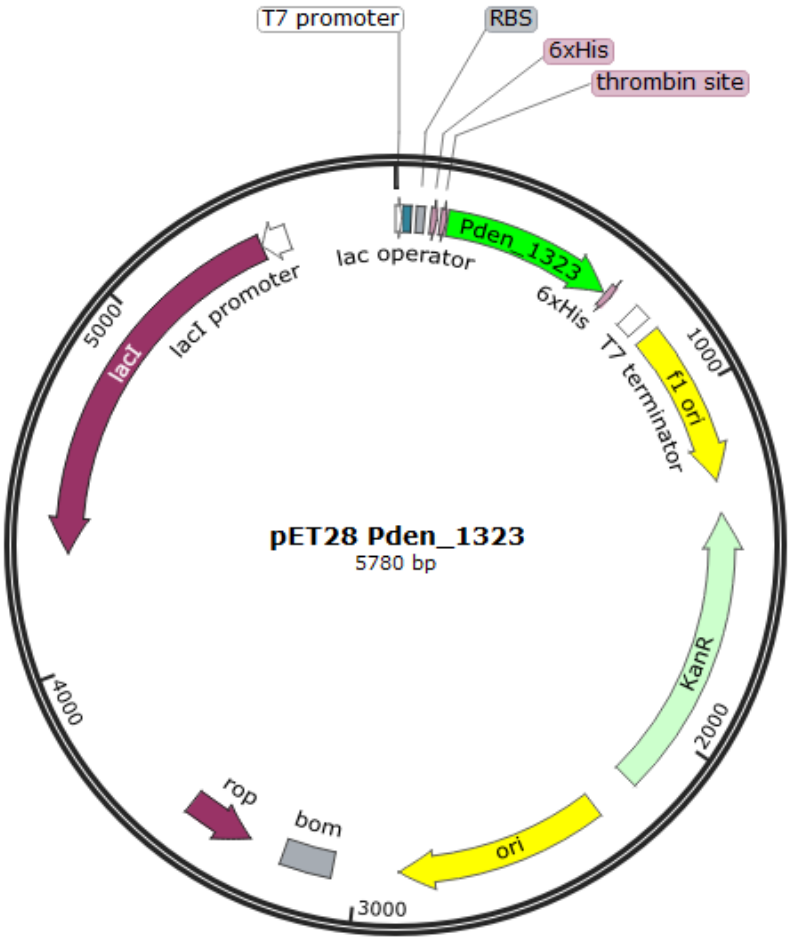


Figure 4.3 Vector map of the pET-28b(+):*pden_1323* vector.

4.2.3 Mutagenesis

The H64A, R92A, and D127A Pden_1323 mutants were generated following the QuikChange site-directed mutagenesis protocol (Fig 4.4). The forward primer designed for the H64A substitution was 5′- CTGGCAGGTCTGGCAGCA**GC**TGCCCGTGCGCTGGCAG-3′. The reverse primer was 5′-CTGCCAGCGCACGGGCA**GC**TGCTGCCAGACCTGCCAG-3′. The forward primer designed for the R92A substitution was 5′-GCAATGACCCATGCA**GC**TCTGAGCATTCTGGGTC-3′. The reverse primer was 5′-GACCCAGAATGCTCAGA**GC**TGCATGGGTCATTGC-3′. The forward primer designed for

the D127A substitution was 5′-GTTTATCTGGATCTGCCGG**C**TTTTCGTTTTTGGCG-3′. The reverse primer was 5′-CGCCAAAAACGAAAA**G**CCGGCAGATCCAGATAAAC-3′. The mutated nucleotides are shown in bold. After DNA polymerase extended and incorporated the mutagenic primers using PCR, the parental template pET-28b(+):*pden_1323* was digested with DpnI. The H64A, R92A, and D127A mutants were transformed into *E. coli* TOP10 cells and plated on Luria-Bertani (LB) agar plates containing 50 µg/mL kanamycin. Colonies were picked, grown in liquid LB medium, and plasmid preparations were made using the Omega Bio-tek E.Z.N.A.® Plasmid Mini Kit.

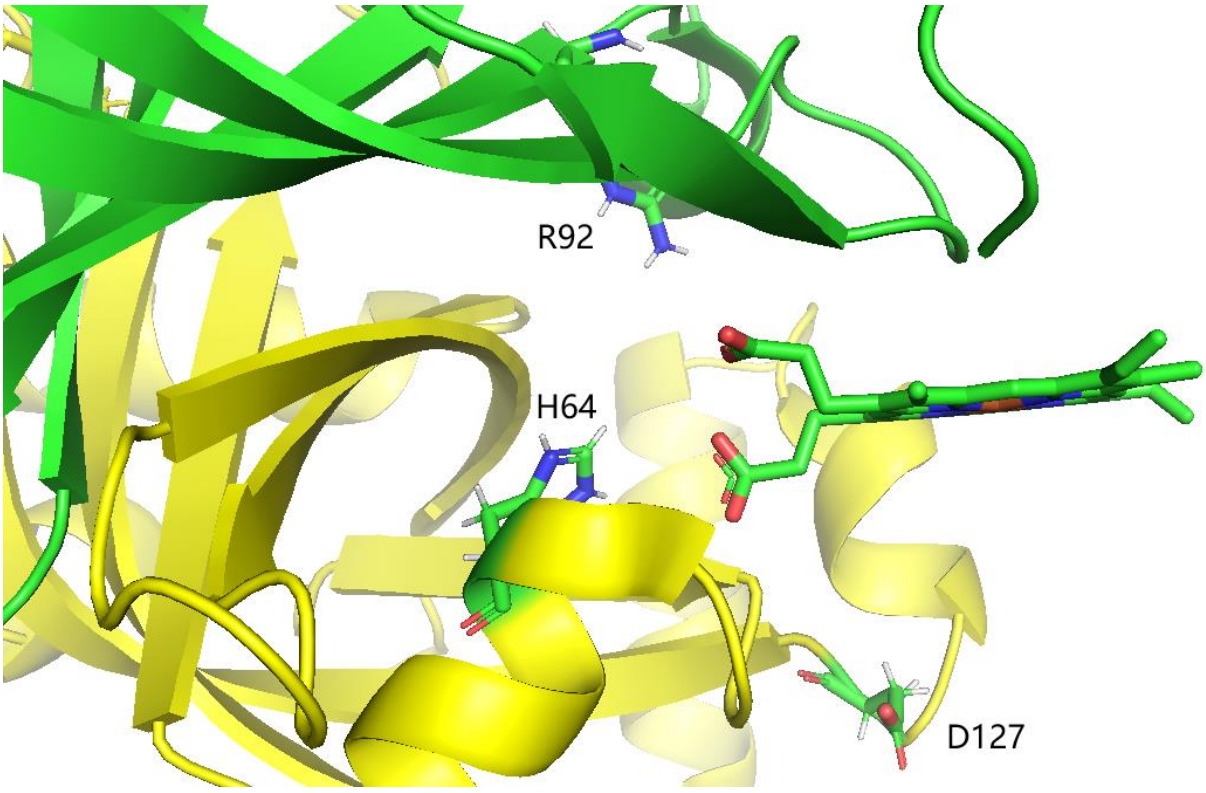


Figure 4.4 Crystal structure of Pden_1323 showing the mutated residues.

4.2.4 Protein expression and purification

The pET-28b(+):*pden_1323*, H64A, R92A, and D127A expression vectors were transformed into *E. coli* BL21(DE3) cells. The transformed expression hosts were grown in overnight cultures (2 mL of LB with 50 μg/mL kanamycin) that were used to inoculate 6 × 1 L of LB (50 mg/L kanamycin). The cultures were incubated with shaking at 37 °C until the OD_{600} reached 0.6. The cultures were then induced with isopropyl β-D-1-thiogalactopyranoside (IPTG) at a final concentration of 0.1 mM and incubated for an additional 4 hours. The cells were centrifuged at 15,900 × g for 45 min. The supernatants were decanted, and the cells were resuspended in 50 mM sodium phosphate buffer (pH 8.0) containing 300 mM NaCl and 5 mM imidazole. The cells were then sonicated and centrifuged at 15,900 × g for 1 hour. The supernatants were loaded onto a Bio-Rad Econo-Pac column loaded with nickel-charged Profinity IMAC resin. The columns were washed with 50 mM sodium phosphate buffer (pH 8.0) containing 300 mM NaCl and 5 mM imidazole. The proteins were then eluted from the column by adding 50 mM sodium phosphate buffer (pH 8.0) containing 300 mM NaCl and 500 mM imidazole in 1 mL fractions. The protein containing fractions were combined and spin-concentrated/buffer exchanged into 100 mM Tris buffer (pH 8.0) followed by incubation with thrombin (1:1000 [w:w]) overnight at room temperature to remove the N-terminal His-tag. After incubation, the cleaved proteins were passed through a benzamidine/Ni-NTA column. The eluted proteins were then spin-concentrated/buffer exchanged into 100 mM Tris buffer (pH 8.0) containing 15 % glycerol and separated into 100 μL aliquots.

For crystallographic studies, the transformed cultures were induced with 0.4 mM IPTG and grown for an additional 3 h. Cells were harvested via centrifugation (15,900 × g for 10 min) and the pelleted cells were stored frozen at -80 °C. The frozen cells were thawed and

resuspended in lysis buffer (50 mM Tris (pH 7.5) buffer with 10 % (v/v) glycerol and 500 mM NaCl) for sonication. The lysed cells were centrifuged (15,900 × g for 20 min) and the supernatant was incubated with Ni-NTA agarose resin (Qiagen) at 4 °C for 30 min and loaded into a column. The column was washed with 15 column volumes of 20 mM imidazole in lysis buffer before Pden_1323 was eluted with 500 mM imidazole in lysis buffer. Pden_1323 was desalted and buffer exchanged into a thrombin cleavage buffer (20 mM Tris (pH 8.0) containing 25 mM $CaCl_2$ and 100 mM NaCl) followed by incubation with thrombin (1:1000 [w:w]) overnight at 25 °C to remove the N-terminal His-tag. After incubation, the cleaved protein was passed through a benzamidine/Ni-NTA column. Pden_1323 was further purified over a gel filtration column (Superdex 200, GE Healthcare Life Sciences) equilibrated with 50 mM Tris buffer (pH 7.5) containing 5 % (v/v) glycerol and 150 mM NaCl. An Amicon® Stirred Cell protein concentrator was used to exchange the protein into 10 mM Tris (pH 7.5) buffer containing 5 % (v/v) glycerol and 25 mM NaCl and achieve a final concentration of 12 mg/mL. Aliquots were flash-frozen in liquid nitrogen and stored at -80 °C until further use.

4.2.5 Heme oxygenase activity assays

The *c*-type heme oxygenase (CHO) activity of Pden_1323 was examined by incubating 2.8 µM enzyme and 5.4 µM MP-11 with 10 mM sodium ascorbate in a cuvette containing 20 mM potassium phosphate buffer (pH 7.6). UV-visible absorption spectra were recorded every two seconds for 5 min using a NanoDrop 2000c (Thermo Fisher). Control reactions were performed with no enzyme or reductant and with MP-11 replaced by 25 µM hemin (with or without 1.8 mM imidazole), 5.4 µM cytochrome *c*, or 1.7 µM MP-9. MP-9 was obtained by incubating 5 mg of cytochrome *c* with 5 mg of trypsin in 1 mL of 20 mM potassium phosphate buffer (pH 7.6) overnight at room temperature. The MP-9 hemopeptide was then purified by high-performance

liquid chromatography (HPLC) using an Agilent 1260 Infinity Quaternary LC System equipped with a Diode Array Detector (DAD) VL+, an analytical scale 1260 Faction Collector, and an Agilent Poroshell 120 EC-C18 (4.6 × 150 mm, 2.7 μm) column. The mobile phase consisted of water with 0.5 % formic acid (solvent A), and acetonitrile with 0.5 % formic acid (solvent B). The chromatographic method used to isolate MP-9 consisted of a linear gradient of 0-100 % solvent B over 30 mins, with a flow rate was 1.0 mL/min and detection at 400 nm.

Another reaction containing 135 μM MP-11, 100 μM Pden_1323, 5 mM sodium ascorbate, and 8 μM catalase in 200 mL of 20 mM potassium phosphate buffer (pH 7.6) was quenched after 30 min with 200 mL methanol and subjected to HPLC analysis. The same LC system and mobile phases as above were utilized with the following gradient: 0 % B for 2 min, linear gradient to 20 % B over 3 min, 20 % B for 5 min, linear gradient to 25 % B over 5 min, 25 % B for 5 min, linear gradient to 30 % B over 5 min, linear gradient to 100 % B over 2 min, and 100 % B for 3 min. A flow rate of 1.0 mL/min was used and the trace was acquired using detection at 280 nm. The reactant and product peaks were analyzed with UV-visible absorption spectroscopy (via the in-line diode array detector) and subjected to mass spectrometry (MS) analysis as detailed below.

The dependence of the initial rate of the CHO reaction on the concentration of enzyme was determined by incubating 9 μM MP-11, 10 mM sodium ascorbate, 2 μM catalase and 1.2-4.8 μM Pden_1323 in a cuvette containing 20 mM potassium phosphate buffer (pH 7.6). The dependence of the initial rate of the CHO reaction on the concentration of MP-11 was determined by incubating 0.6 μM enzyme, 10 mM sodium ascorbate, 2 μM catalase, and 1-10 μM MP-11 in a cuvette containing 20 mM potassium phosphate buffer (pH 7.6) in a total volume

of 400 µL. The reactions were monitored at 406 nm and were performed in triplicate using an Agilent 8453 UV-visible Spectrophotometer. The latter data were then fit to the Hill equation,

$$v = \frac{k_{cat}[\text{Pden\_1323}][\text{MP-11}]^n}{K_m{}^n + [\text{MP-11}]^n}$$

to obtain estimates for the turnover number (k_{cat}) and apparent dissociation (Michaelis) constant (K_m) of Pden_1323/MP-11.

4.2.6 CO detection assays

To detect the production of CO in the Pden_1323 reaction, a mixture of 50 µM Pden_1323, 50 µM MP-11, 10 mM sodium ascorbate and 2 µM catalase was placed in a sealed vial containing 20 mM potassium phosphate buffer (pH 7.6). A control reaction was also performed without Pden_1323. After one hour, 100 µL of both the Pden_1323 and control reactions were mixed with 300 µL of an anaerobic myoglobin solution (10 µM final concentration) in an anaerobic chamber. UV-visible absorption spectra were then recorded using a NanoDrop One^C (Thermo Fisher).

4.2.7 Iron determination assay

Ferrozine-based colorimetry was used for determination of the released iron in the Pden_1323 reaction. All containers and pipette tips used were boiled in 1 M HCl and washed with ultra-pure water. Iron standards (0, 10, 20, 30, 40, 50 µM) were prepared by dissolving ferrous ethylenediammonium sulfate in 10 mM HCl. The iron chelating and reducing reagent contained 6.5 M ferrozine, 13.1 mM neocuprine, 2 M ascorbic acid, and 5 M ammonium acetate in a 25 ml solution. The reaction sample contained 1.2 µM Pden_1323, 10 mM sodium ascorbate, 2-10 µM MP-11, 2 µM catalase, and 40 µL iron chelating and reducing reagent in a cuvette containing a total of 400 µL of 20 mM potassium phosphate buffer (pH 7.6). The control

reaction was also performed without enzyme. The reactions were monitored at 562 nm using an Agilent 8453 UV-visible Spectrophotometer.

4.2.8 Mass spectrometry analysis

MS analysis of the Pden_1323 reaction mixtures was performed at the Auburn University Chemistry and Biochemistry Mass Spectrometry Center on an ultra-performance LC system (ACQUITY, Waters Corp., USA) coupled with a quadrupole time-of-flight mass spectrometer (Q-TOF Premier, Waters Corp., USA) with electrospray ionization (ESI) in positive mode using MassLynx software (v4.1). Injections of the samples were made directly into the mass spectrometer or onto a C4 column (Aeris™ 3.6 µm C4 200 Å, 50 × 2.1 mm, Phenomenex) with a 300 µL/min flow rate of mobile phase solution A (0.1 % formic acid in 95 % water and 5 % acetonitrile) and solution B (0.1 % formic acid in 95 % acetonitrile and 5 % water) using the following gradient: 0 % B for 2 min, linear gradient to 100 % B over 11 min, 100 % B for 1 min, linear gradient to 0 % B over 1 min, and 0 % B for 3 min. The column temperature was held at 40 °C and the UV detector monitored at absorption at 420 nm. The capillary voltage was set at 3.1 kV, the sample cone voltage was 15 V, and the extraction cone was 4.0 V. The source and desolvation temperature were maintained at 95 and 300 °C, respectively, with the desolvation gas flow set to 600 L/h. The mass spectrometry scan was 0.5 s long from 50 to 2,000 *m*/*z* with a 0.02 s interscan delay using the centroid data format. Tandem mass spectra were taken for selected masses with a 0.5 s long scan and 0.025 s interscan delay over the same range. The collision energy varied as described in the text. The lock mass was used to correct instrument accuracy with a 2 ng/µL solution of leucine enkephalin.

4.2.9 Crystallization, data processing, and structure determination

Three sparse matrix screens (The JCSG Core IV, The JCSG + and The Protein Complex Suites, Qiagen) were used to determine crystallization conditions for Pden_1323. Several conditions produced crystals, and optimization plates were set up for the following condition: 14 % (w/v) glycerol, 18-25 % (w/v) PEG 4000, 0.3 M sodium acetate, and 0.1 M Tris (pH 8.8). Crystals grew overnight via the sitting-drop vapor diffusion method. Each drop consisted of 1 μL of Pden_1323 (12 mg/mL) and 2 μL of crystallization buffer. Crystals were cryoprotected with 2 μL of glycerol in 8 μL mother liquor for 5 min then flash frozen in liquid nitrogen prior to data collection. For the cocrystallization with ligand, 12 mg/mL Pden_1323 was incubated with 10 μM MP-11 at 4 ^{0}C for 1 hour. The crystallization conditions were determined using three sparse matrix screens (The JCSG Core IV, The JCSG + and The Protein Complex Suites, Qiagen).

Data for Pden_1323 were collected on the Advanced Photon Source Beamline 23-ID-B and subsequently processed with HKL-3000. The structure of Pden_1323 was determined by molecular replacement with the Atu2129 monomer (PDB ID code: 3DNH) as a search model using Phaser. The model was initially refined with ARP/wARP and then built and refined through several cycles with Coot and Refmac. The final structure was optimized using the PDB_REDO server.

4.3 Results

4.3.1 Expression and purification

The codon optimized version of *pden_1323* and mutant variants were synthesized for heterologous expression in *Escherichia coli*. Pden_1323 and mutants were purified from the cell extract using a Ni-affinity column. The purity of the protein was determined by sodium dodecyl-

sulfate polyacrylamide gel electrophoresis (SDS-PAGE). Pden_1323 and the mutant variants have molecular weights of 19.1 kDa with the His-tag, and 17.3 kDa without the His-tag (Fig 4.5).

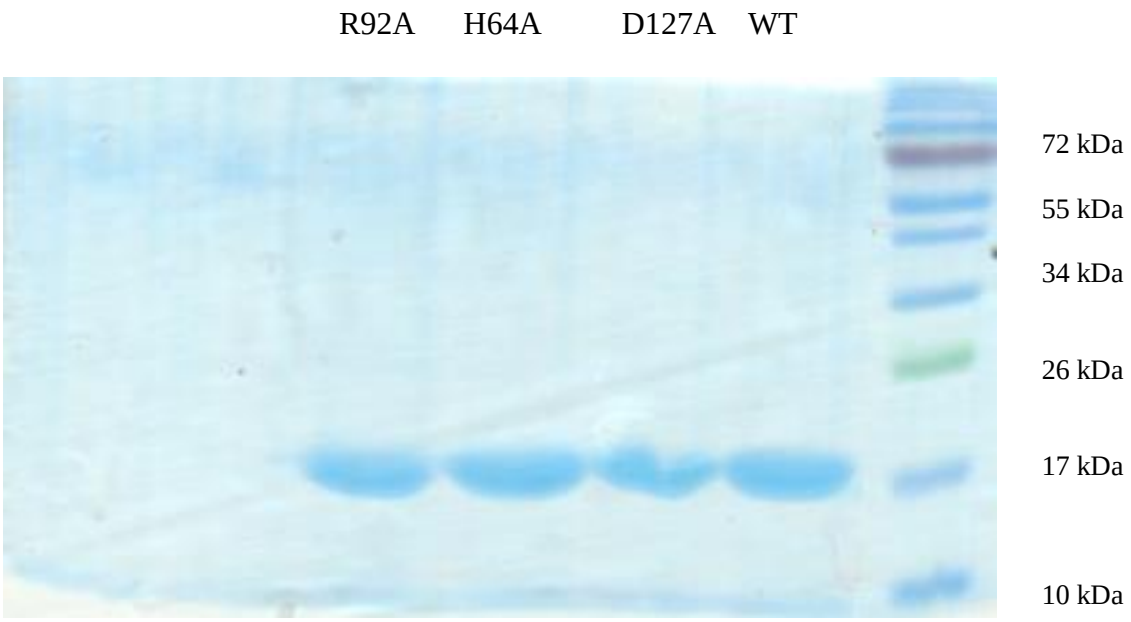


Figure 4.5 SDS-PAGE gel of Pden_1323 and mutant variants.

4.3.2 Pden_1323 Lacks PMP/PNP Oxidase and HO Activity

The purified enzyme was tested for its ability to oxidize PMP and pyridoxine 5′-phosphate (PNP), but no activity was observed with either substrate using FMN, FAD, F_{420}, or 8-hydroxy-5-deazaflavin (F_o). Pden_1323 was also tested for HO activity using hemin as the substrate. Again, no significant activity above the background level was detected using ascorbate as the reductant (Fig 4.6).

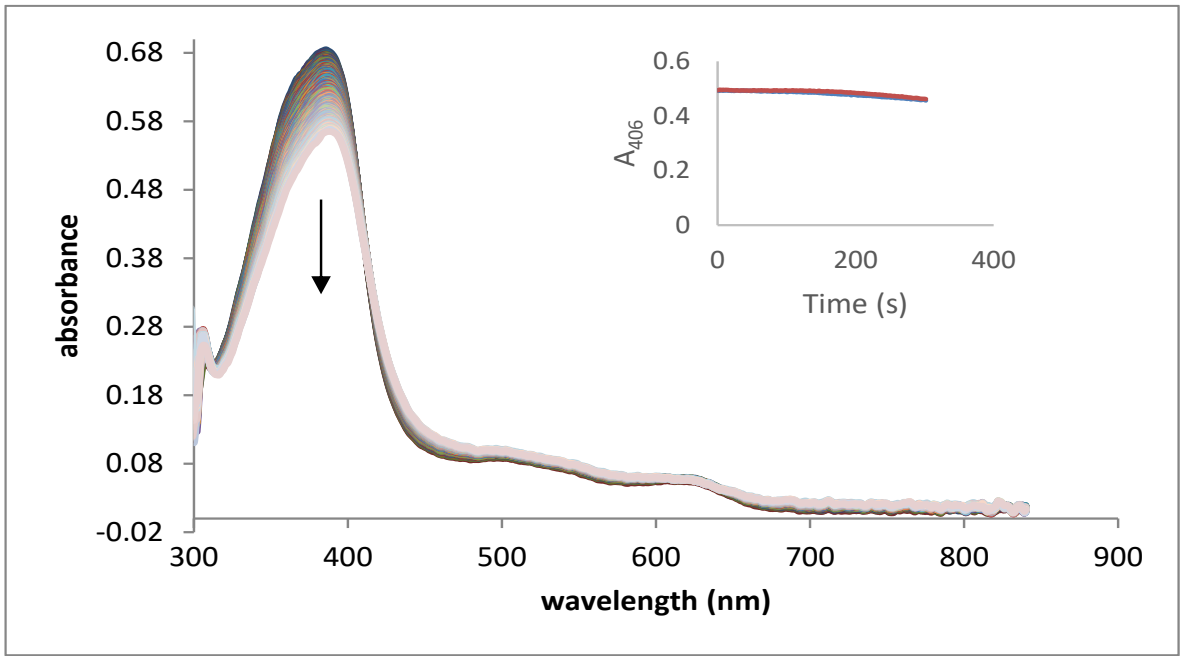


Figure 4.6 Spectrophotometric assays of Pden_1323 with hemin as substrate. UV-visible absorption spectra of the Pden_1323 reaction with 25 μM hemin and 10 mM ascorbate taken every 2 seconds for 5 minutes. Inset: Time course of hemin degradation measured by the decrease in intensity of the Soret band in the presence (blue trace) or absence (red trace) of Pden_1323.

4.3.3 Crystral Structure of Pden_1323

The X-ray crystal structure of Pden_1323 was solved and found to form a dimeric split β-barrel similar to HugZ (Fig 4.7).[173] However, comparison of the active sites of Pden_1323 and HugZ show stark differences. While the two residues of HugZ that form salt bridging interactions with the carboxylates of heme, His163 and Arg166, are conserved in the structure of Pden_1323 (His64 and Arg92), the entire N-terminal α/β domain of HugZ is not present in the structure of Pden_1323 (Fig 4.7).[173] Furthermore, the C-terminal loop that contains the coordinating heme ligand is not found in the Pden_1323 structure (Fig 4.7).[173] Taken together, the lack of the N-terminal domain and C-terminal loop indicate that the presumed active site of Pden_1323 is only partially formed, which provides a rationale for Pden_1323's lack of HO

activity with free heme. Given these structural characteristics, combined with the fact that *pden_1323* is found within a gene cluster that is likely involved in the transport of both heme and peptides, it was reasoned that the substrate of Pden_1323 may instead be a *c*-type hemopeptide.

Table 4.1. Data collection and refinement statistics (molecular replacement).

	Pden_1323
Data collection	
Space group	C2
Cell dimensions	
a, *b*, *c* (Å)	72.84, 117.04, 59.38
α, β, γ (°)	90, 99.38, 90
Resolution (Å)	58.66 – 2.20 (2.27 – 2.20)
R_{merge}	0.037(0.378)
Completeness (%)	99.6(99.3)
Redundancy	2.0(1.9)
Refinement	
Resolution (Å)	2.20
No. reflections	24905
R_{work} / R_{free}	0.212/0.232
No. atoms	
Protein	6562
Ligand/ion	
Water	47
B-factors	
Protein	69
Ligand/ion	
Water	47
R.m.s. deviations	
Bond lengths (Å)	0.0067
Bond angles (°)	1.181

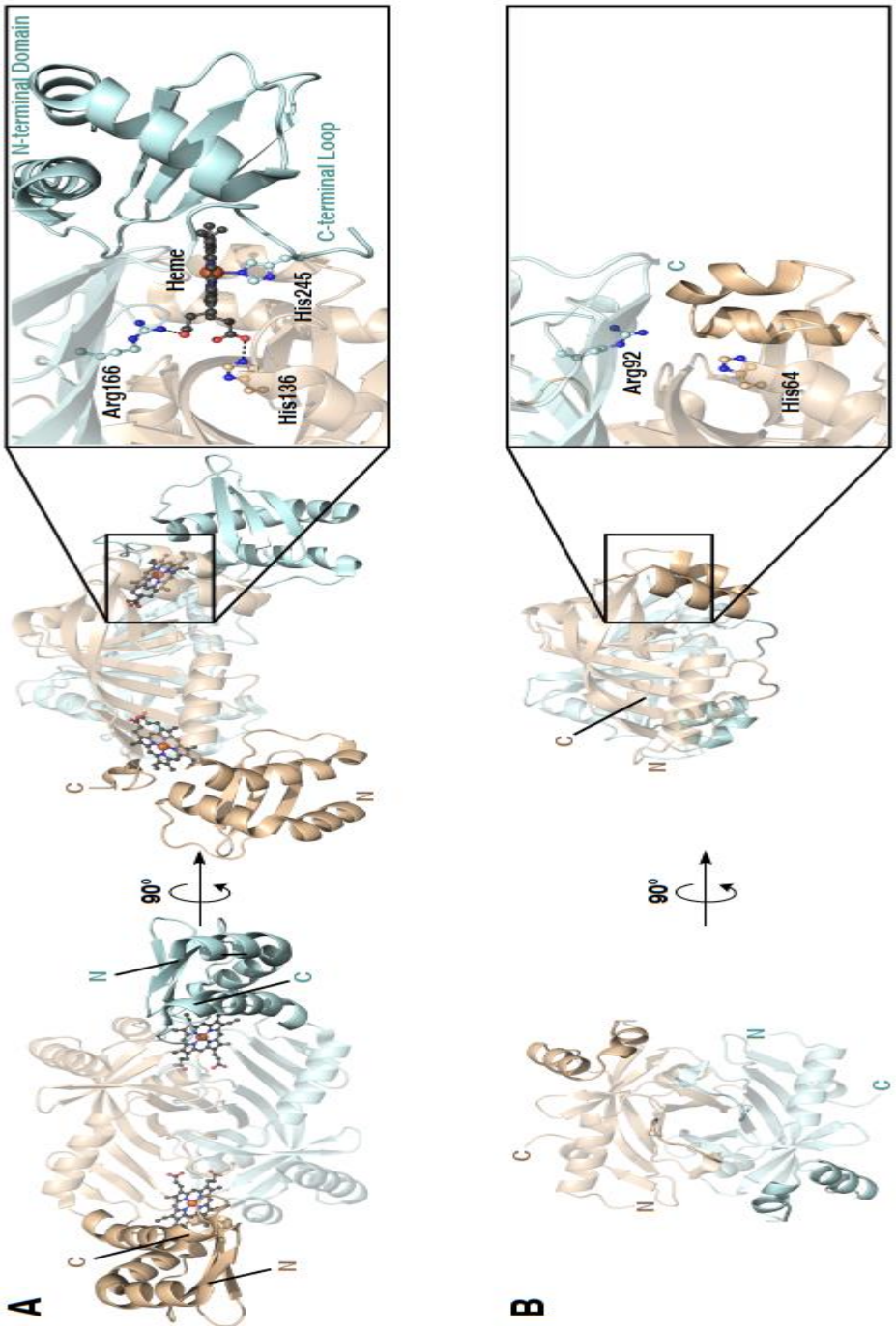


Figure 4.7 Structural comparison of HugZ and Pden_1323. *A*, X-ray crystal structure of HugZ (PDB ID code: 3GAS). The N-terminal α/β domain and C-terminal loop that are absent in the structure of Pden_1323 are highlighted. The heme substrate and several interacting active site residues (His136, Arg166, and His 245) are shown as ball-and-sticks. *B*, X-ray crystal structure of Pden_1323 (deposited PDB ID code: 6VNA). Dynamic regions showing varying degrees of disorder within the monomers of the asymmetric unit are highlighted. The His and Arg residues that interact with the carboxylates of heme are conserved in the structure of Pden_1323 and are shown as ball-and-sticks (His64 and Arg92).

4.3.4 Activity assay of Pden_1323

When Pden_1323 was incubated with ascorbate and microperoxide-11 (MP-11), an 11 amino acid heme-containing fragment of equine heart cytochrome *c*, a rapid decrease in the Soret band of MP-11 was observed (Fig 4.8).[182] It means Pden_1323 degraded MP-11 and formed a product without 400 nm peak. The negative control without Pden_1323 or ascorbate showed no activity.

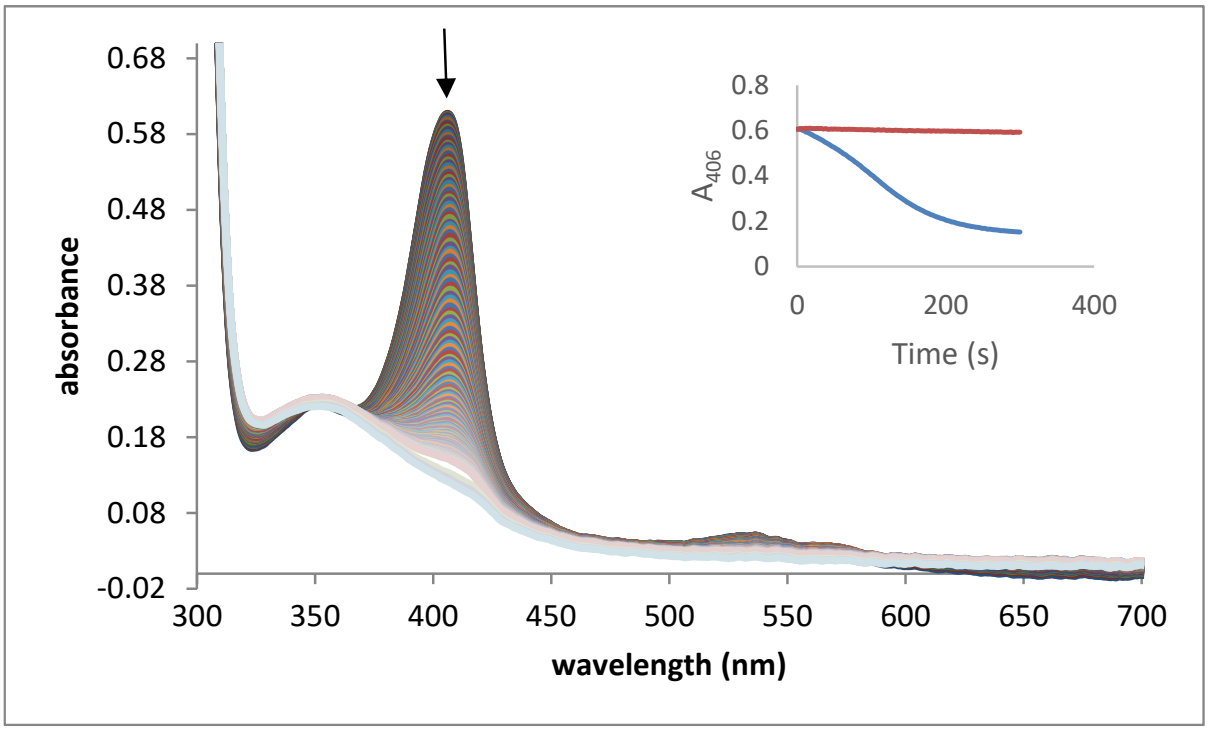


Figure 4.8 Spectrophotometric assays of Pden_1323 with MP-11 as substrate. UV-visible absorption spectra of the Pden_1323 reaction with 5.4 μM MP-11 and 10 mM ascorbate taken every 2 seconds for 5 minutes. Inset: Time course of MP-11 degradation measured by the decrease in intensity of the Soret band in the presence (blue trace) or absence (red trace) of Pden_1323.

We set up the reaction with four different concentrations of Pden_1323 using MP-11 as substrate. The velocity of the reaction was remarkably correlated linearly with the concentration of Pden_1323 (Fig 4.9). The loss of the *c*-type heme chromophore was both enzyme and reductant dependent.

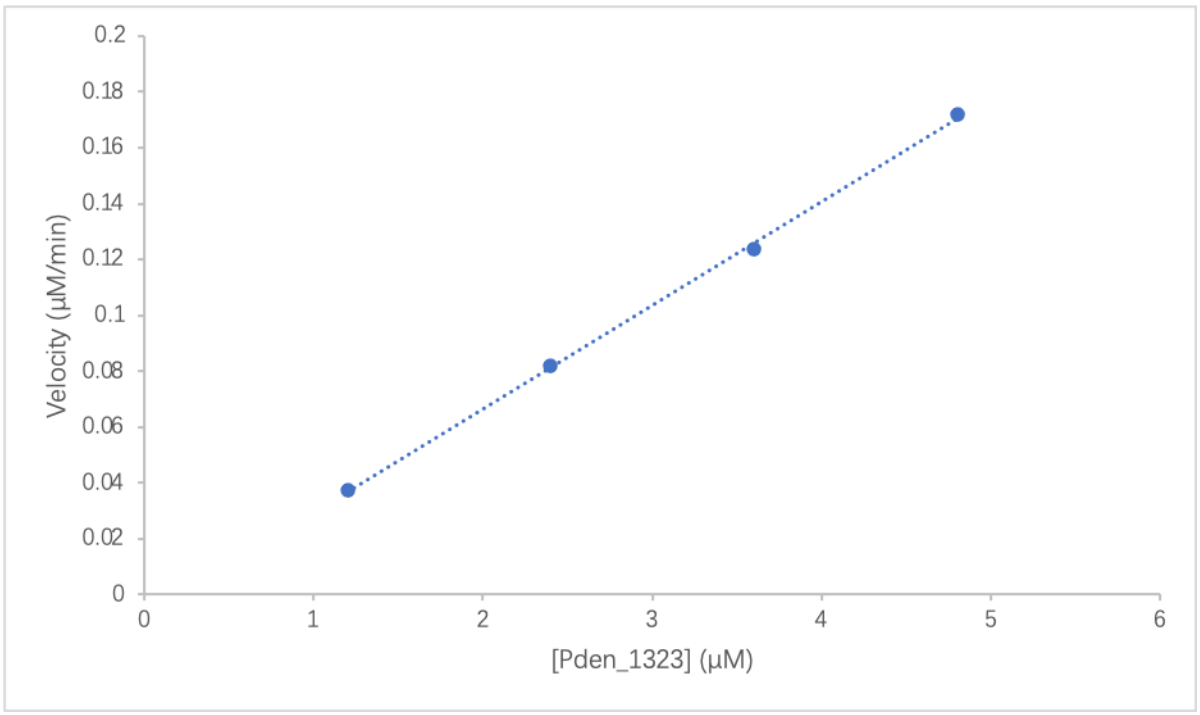


Figure 4.9 The velocity changes of the reaction based on the concentration of Pden_1323. Comparison of the rates of degradation of 9 μM MP-11 as a function of Pden_1323 concentration in the presence of 2 μM catalase.

4.3.5 Kinetics of Pden_1323

In the presence of catalase, the Pden_1323-catalyzed degradation of MP-11 occurred with a specific activity of 3.9 nmol/min/mg and displayed sigmoidal kinetics with a turnover number (k_{cat}) of 0.076 ± 0.0006 min^{-1}, an apparent dissociation (Michaelis) constant (K_m) of 2.0 ± 0.03 μM, and a Hill coefficient of 2.6 ± 0.08 (Fig 4.10).

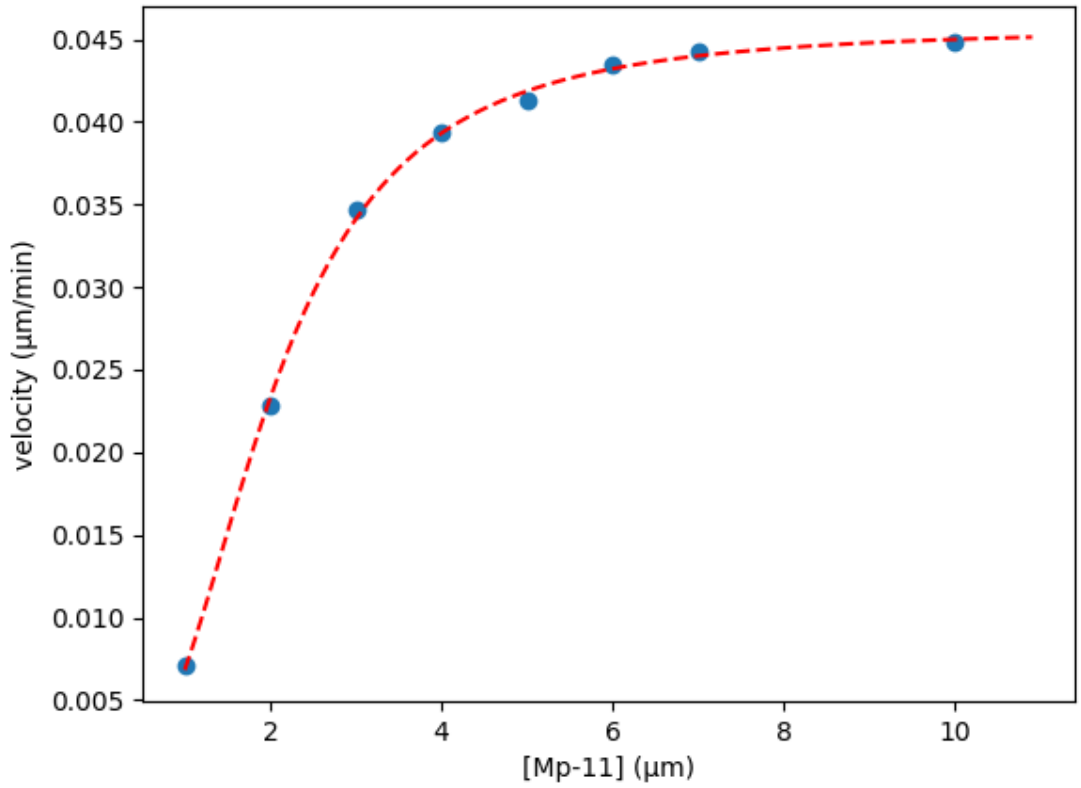


Figure 4.10 The velocity changes of the reaction based on the concentration of MP-11. Initial rate of the Pden_1323 reaction as a function of [MP-11] in the presence of 2 µM catalase. The dashed line shows best fit to the Hill equation using the kinetic parameters k_{cat} = 0.076 ± 0.0006 min^{-1}, K_m = 2.0 ± 0.03 µM, and n = 2.6 ± 0.08.

4.3.6 Products of Pden_1323

To detect products, we used high-performance liquid chromatography (HPLC) to separate the product from the reaction mixture. MP-11 is a peptide that has Soret band at 280 nm. We observed two peaks at 280 mm. The retention time of 15.9 min had a biliverdin like spectrum. And the retention time of 21.6 min had a hemin like spectrum (Fig 4.11). Both of them were analyzed by mass spectrometry (MS). The peak at 15.9 min had a mass of 1827.81 *m/z*, which is consistent with a peptide-linked biliverdin (theoretical $[M + H]^+$ *m/z* = 1827.82) (Fig 4.11). The

peak at 21.6 min had a mass of 1861.74 *m/z*, which is consistent with MP-11 (theoretical $[M + H]^+$ *m/z* = 1861.75) (Fig 4.11).

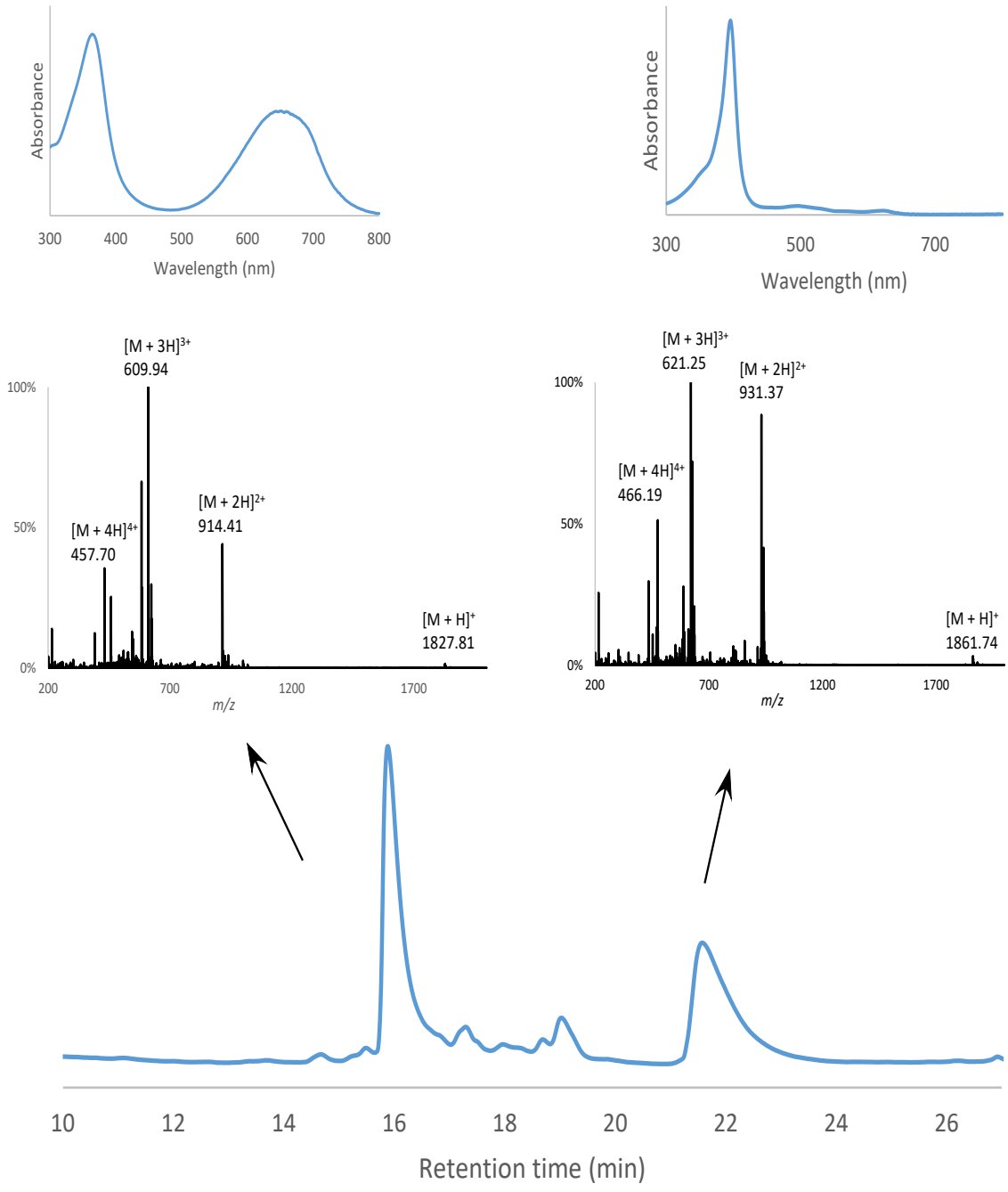


Figure 4.11 HPLC and MS analyses the Pden_1323 reaction with MP-11. HPLC analysis of a Pden_1323 reaction mixture, showing MS and UV-visible absorption spectra of the reactant (21.6 min) and product (15.9 min) peaks.

To further confirm the product, we analyzed the product using tandem mass spectrometry (MS/MS). The data showed that the 1827.82 ion could be cleaved into peptide (theoretical *m*/*z* = 1245.57) and bilin (theoretical *m*/*z* = 583.26) components by increasing the collision energy (Fig 4.12). LC-MS/MS analysis of the bilin peak yielded a fragmentation pattern that is distinct from that obtained with α-biliverdin and consistent with oxidized biliverdin fragments containing A-B-C (or B-A-D) rings (theoretical *m*/*z* = 421.20) and C-B-D (or A-D-C) rings (theoretical *m*/*z* = 495.20) (Fig 4.13). Thus, Pden_1323 converts MP-11 to a product with a mass and UV-visible absorption spectrum that are consistent with a peptide-linked biliverdin. And the MS data suggest that the product of the reaction is a peptide linked β- and/or δ-biliverdin.

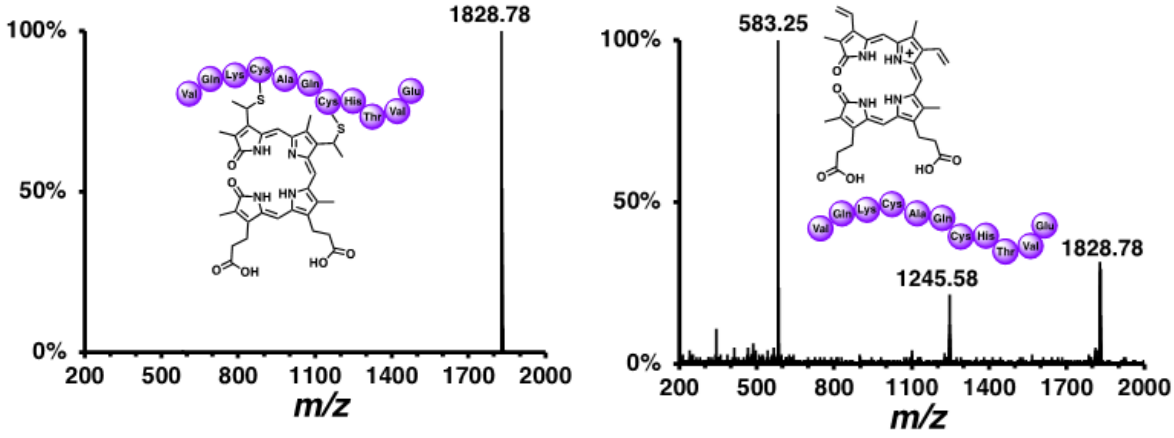


Figure 4.12 MS/MS analyses confirming the product of the Pden_1323 reaction with MP-11. Left: simplified MS/MS spectrum of the product of the Pden_1323 reaction with MP-11 using a linear range of collision energies from 10-30 eV. Right: simplified MS/MS spectrum of the product of the Pden_1323 reaction with MP-11 using a linear range of collision energies from 60-90 eV.

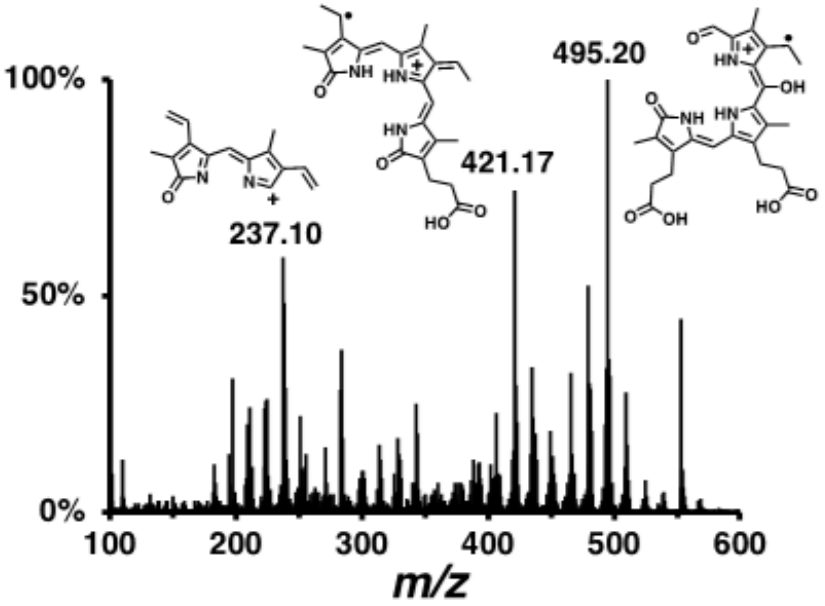


Figure 4.13 LC-MS/MS spectrum of the 583.25 *m/z* peak of the Pden_1323 reaction with MP-11.

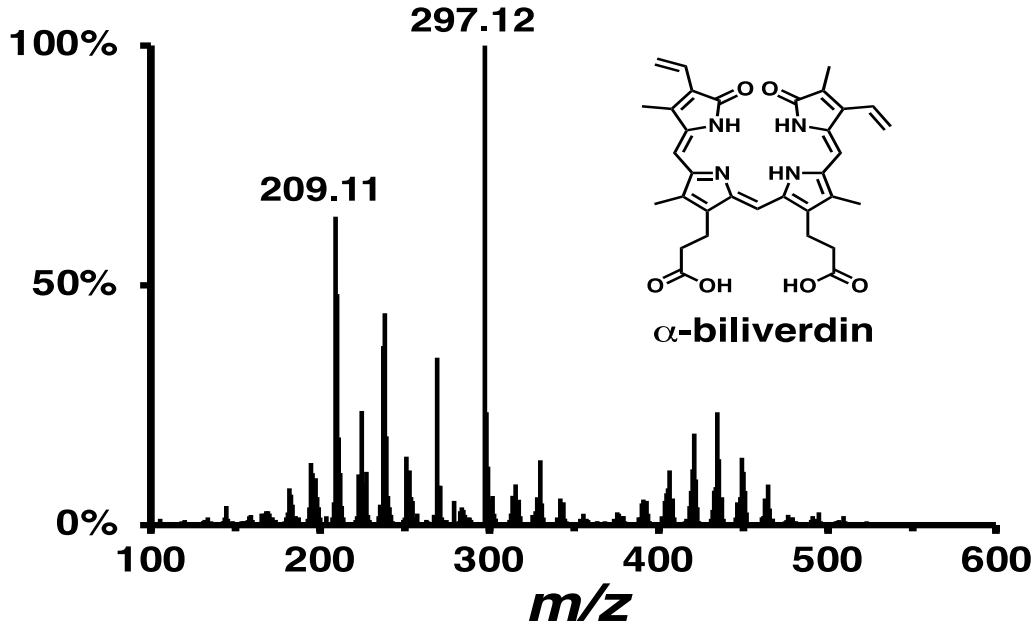


Figure 4.14 LC-MS/MS spectrum of a biliverdin IXα standard.

HO-1 cleaves heme at the α-meso position, releasing ferrous iron and producing biliverdin IXα (α-biliverdin) and carbon monoxide as coproducts. From MS data, the mass of the product is consistent with a peptide-linked biliverdin, which suggests that CO may also be formed in the Pden_1323 reaction. We used myoglobin, which has a high affinity for CO, to detect the coproduct CO. The spectrum of myoglobin-CO complex has a Soret band at 421 nm and two small bands at 540 nm and 580 nm (Fig 4.15). After incubating myoglobin in the Pden_1323 reaction, the Soret band of MP-11 shifted from 406 to 421 and appeared the bands at 540 nm and 580 nm (Fig 4.15). The conversion indicates that CO was generated as a product of the Pden_1323 reaction.

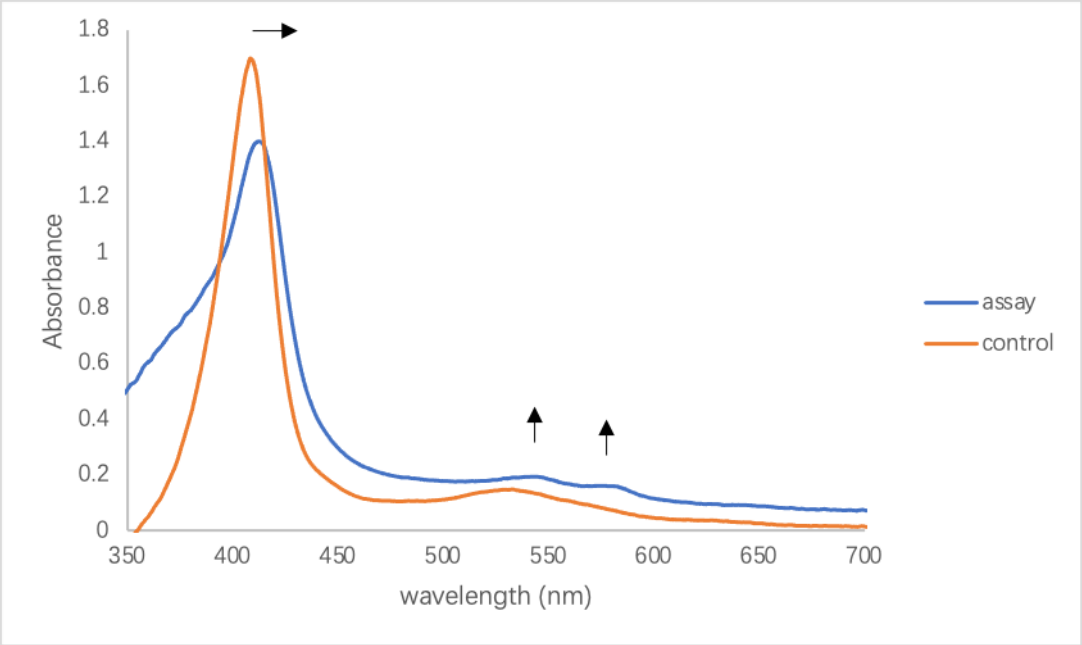


Figure 4.15 Changes in the absorption spectra of the Pden_1323 reaction in the presence of myoglobin. The assay contained 50 µM Pden_1323, 50 µM MP-11, 10 mM sodium ascorbate and 2 µM catalase. The control reaction was also performed without Pden_1323. After one hour, 100 µL of both the Pden_1323 and control reactions were mixed with 300 µL of a myoglobin solution (10 µM).

We used ferrozine assay to confirm the concentration of free iron increase. Ferrozine is a chelator of ferrous iron. It can form an iron-Ferrozine complex and absorb strongly at 562 nm.

We incubated the iron chelating and reducing reagent in Pden_1323 reaction for 10 mins. Both assay and control were measured at 562 nm using UV-visible spectrophotometry. The corrected concentration of the released iron with control was correlated linearly with the concentration of substrate MP-11 (Fig 4.16). The data indicates that iron was also generated as a product of Pden_1323 reaction.

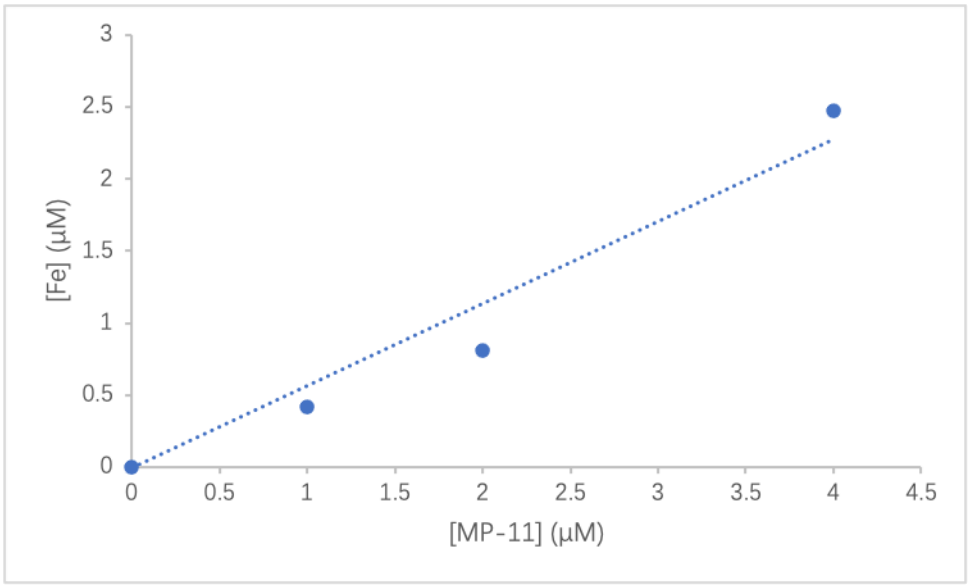


Figure 4.16 Ferrozine assay showing Pden_1323-catalyzed production of free iron as a function of MP-11 concentration. The sample contained 1.2 μM Pden_1323, 10 mM sodium ascorbate, 2-10 μM MP-11, 2 μM catalase, and 40 μL iron chelating and reducing reagent. The control reaction was also performed without enzyme. The reactions were monitored at 562 nm using a UV-visible Spectrophotometer.

4.3.7 Various substrates of Pden_1323

Various substrates were tested for Pden_1323 catalysis. Full-length cytochrome *c* was incubated with Pden_1323 and showed no spectral changed (Fig 4.17). This indicates the peptide linkage must be small enough so that the porphyrin macrocycle is not buried and is able to access the active site of Pden_1323. To test substrate with different lengths of *c*-type hemopeptide, cytochrome *c* was digested with trypsin and the peptides were purified by HPLC to obtain a short 9 amino acid hemopeptide (MP-9). When MP-9 was incubated with Pden_1323, the intensity of

the Soret band at 400 nm decreased as observed with MP-11 (Fig 4.18). The MP-9 sequence is missing the three N-terminal residues of MP-11 (VQK) and has a C-terminal lysine does not present in the MP-11 sequence. This suggests that the exact sequence of the hemopeptide is not critical for recognition and cleavage by Pden_1323, though the coordinating histidine residue is present in both MP-11 and MP-9.

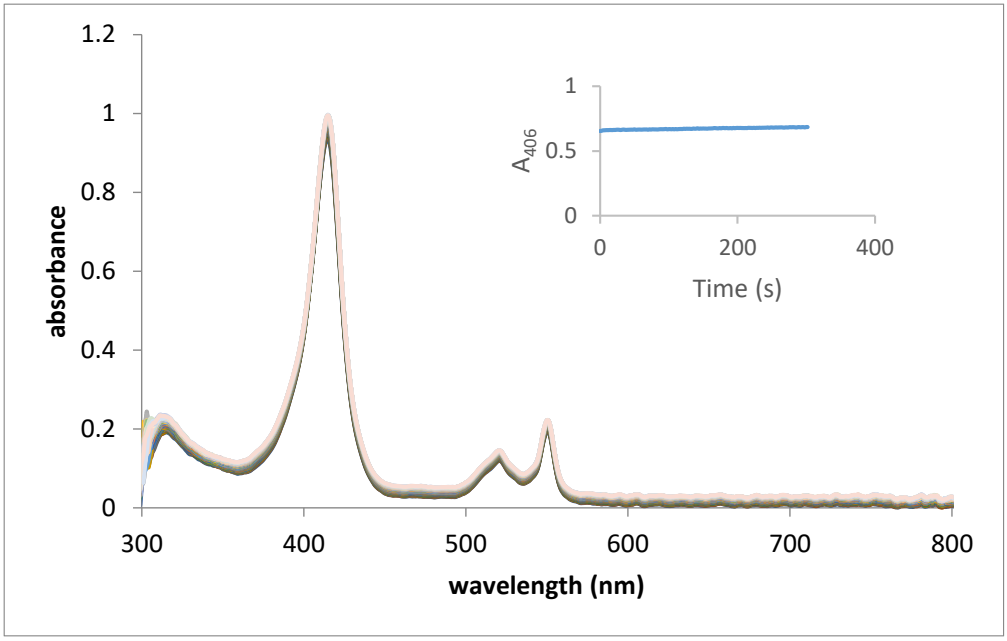


Figure 4.17 Spectrophotometric assays of Pden_1323 with full-length cytochrome *c* as substrate. UV-visible absorption spectra of the Pden_1323 reaction with 5.4 μM cytochrome *c* and 10 mM ascorbate taken every 50 seconds for 5 minutes. Inset: Time course of the reaction followed by monitoring the intensity of the Soret band.

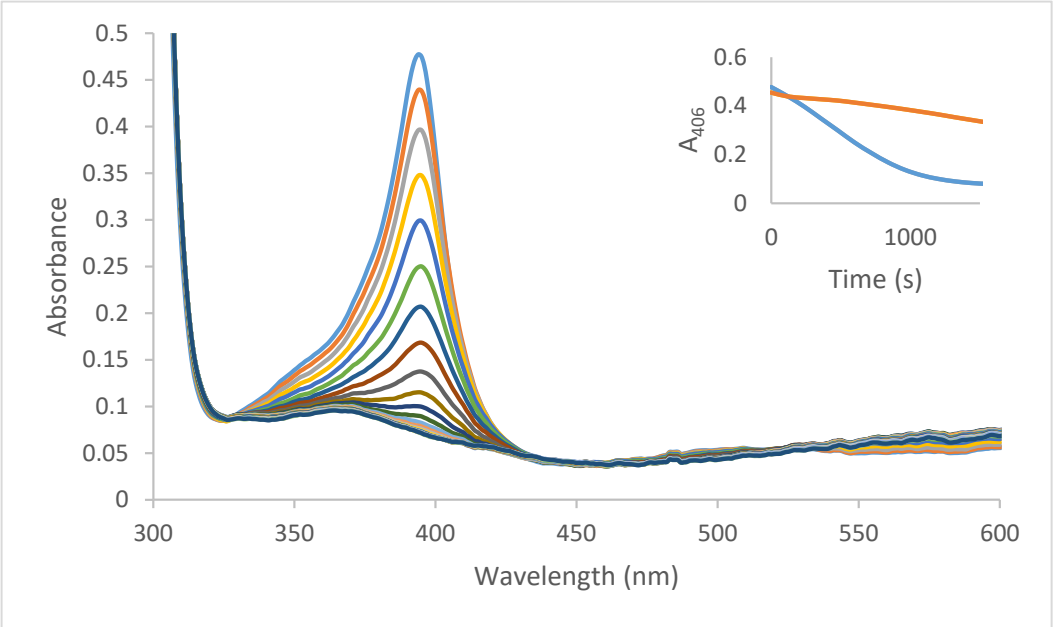


Figure 4.18 Spectrophotometric assays of Pden_1323 with MP-9 as substrate. UV-visible absorption spectra of the Pden_1323 reaction with 1.7 μM MP-9 and 10 mM ascorbate taken every 4 minutes. Inset: Time course of MP-9 degradation measured by the decrease in intensity of the Soret band in the presence (blue trace) or absence (black trace) of Pden_1323.

To confirm the importance of histidine lower axial ligand, hemin was incubated with a high concentration of imidazole (where ~90 % is expected to be coordinated given an estimated K_d ~ 200 μM) (Fig 4.19). Pden_1323 can use imidazole coordinated hemin as substrate.

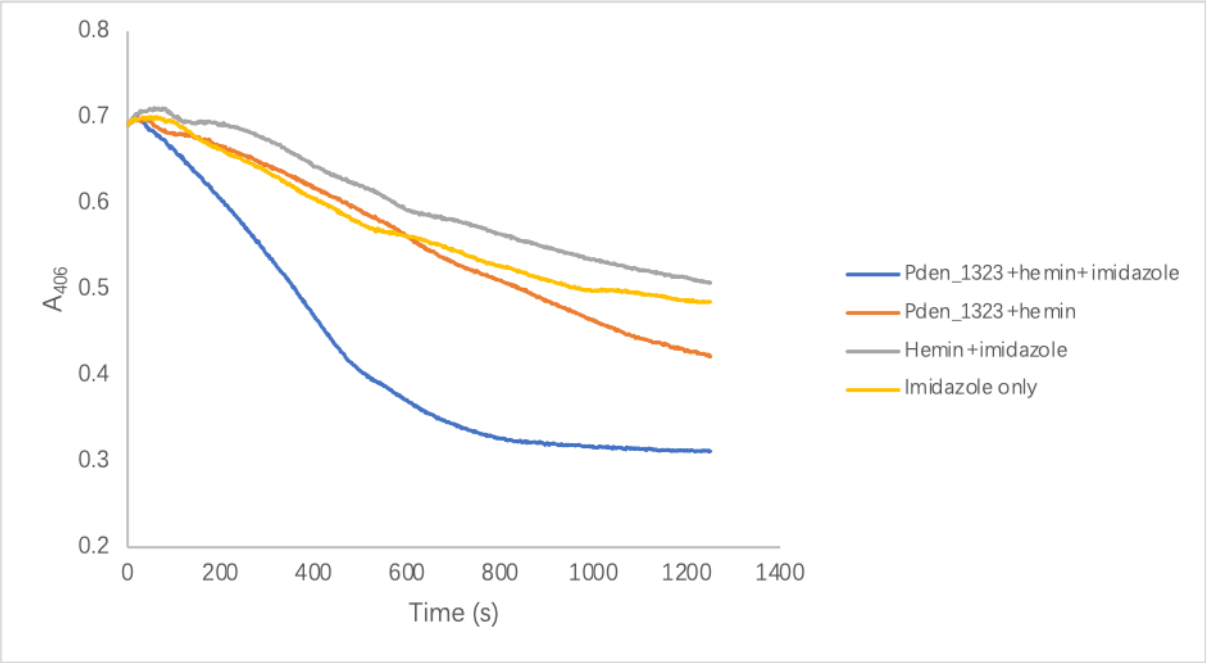


Figure 4.19 Spectrophotometric assays of Pden_1323 with coordinated hemin as substrate. Blue: UV-visible absorption spectra of the Pden_1323 reaction with 25 μM hemin, 1.8 mM imidazole and 10 mM ascorbate every 2 seconds for 5 mins. Red: control without imidazole. Grey: control without Pden_1323. Yellow: control with only imidazole.

4.3.8 Kinetics of Pden_1323 mutants

Residues His64 and Arg92 are proposed to form salt bridging interactions with the carboxylates of heme, so they could play critical roles in binding and orienting the heme substrate of Pden_1323. Asp127 is proposed to form a hydrogen bond with the proximal histidine ligand of the heme substrate, which is expected to be important in the heme oxygenase mechanism of Pden_1323. To investigate their importance, we prepared H64A, R92A, and D127A mutants and tested their activities.

The D127A mutant catalyzed degradation of MP-11 displayed sigmoidal kinetics with a turnover number (k_{cat}) of 0.02 min^{-1}, an apparent dissociation (Michaelis) constant (K_m) of 1.36 μM, and a MP-11 oxygenase activity (k_{cat}/k_M) of 0.015 $\mu M^{-1}min^{-1}$ (Fig 4.20). Compared with wild type, Pden_1323 lost more than half of catalytic activity.

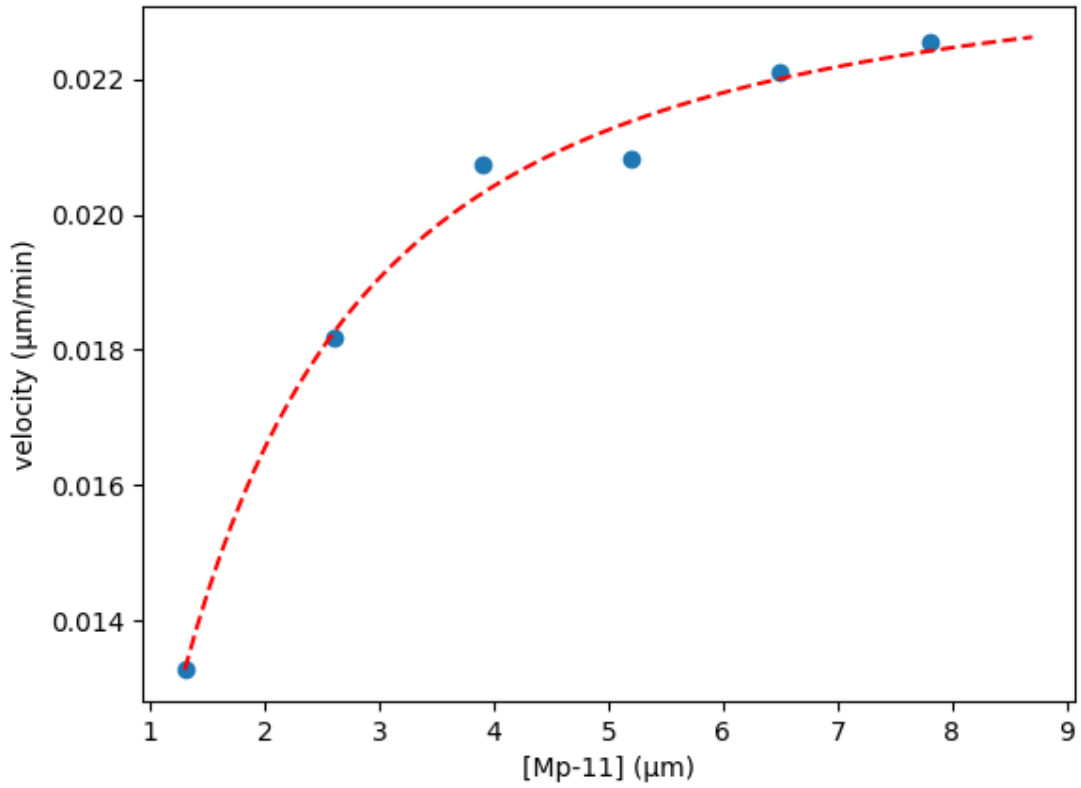


Figure 4.20 Initial rate of the Pden_1323 D127A reaction as a function of [MP-11] in the presence of 2 µM catalase. The dashed line shows best fit to the Hill equation using the kinetic parameters k_{cat} = 0.02 min^{-1}, K_m = 1.36 µM, and n = 1.11.

The H64A mutant catalyzed degradation of MP-11 displayed sigmoidal kinetics with a turnover number (k_{cat}) of 0.14 min^{-1}, an apparent dissociation (Michaelis) constant (K_m) of 1.83 µM, and a MP-11 oxygenase activity (k_{cat}/k_M) of 0.076 $\mu M^{-1}min^{-1}$ (Fig 4.21). Compared with wild type, Pden_1323 gained more than twice the catalytic activity.

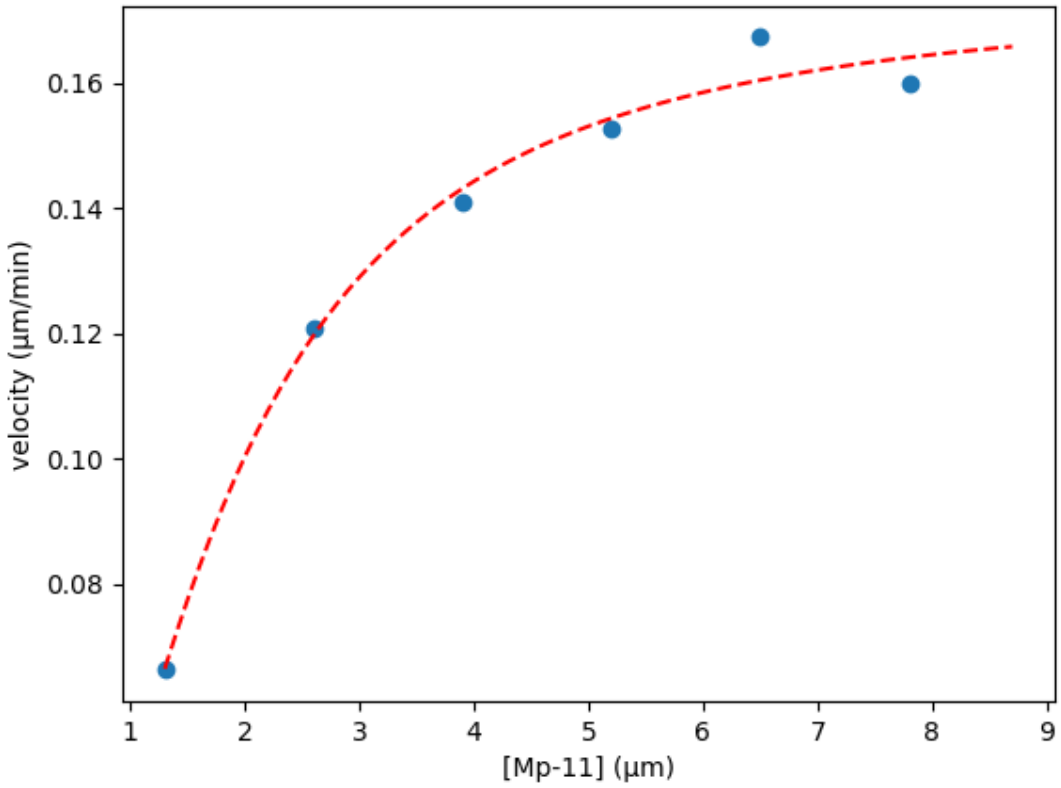


Figure 4.21 Initial rate of the Pden_1323 H64A reaction as a function of [MP-11] in the presence of 2 µM catalase. The dashed line shows best fit to the Hill equation using the kinetic parameters k_{cat} = 0.14 min^{-1}, K_m = 1.83 µM, and n = 1.69.

The R92A mutant catalyzed degradation of MP-11 displayed sigmoidal kinetics with a turnover number (k_{cat}) of 0.033 min^{-1}, an apparent dissociation (Michaelis) constant (K_m) of 1.9 µM, and a MP-11 oxygenase activity (k_{cat}/k_M) of 0.017 $\mu M^{-1}min^{-1}$ (Fig 4.22). Compared with wild type, Pden_1323 lost more than half of catalytic activity.

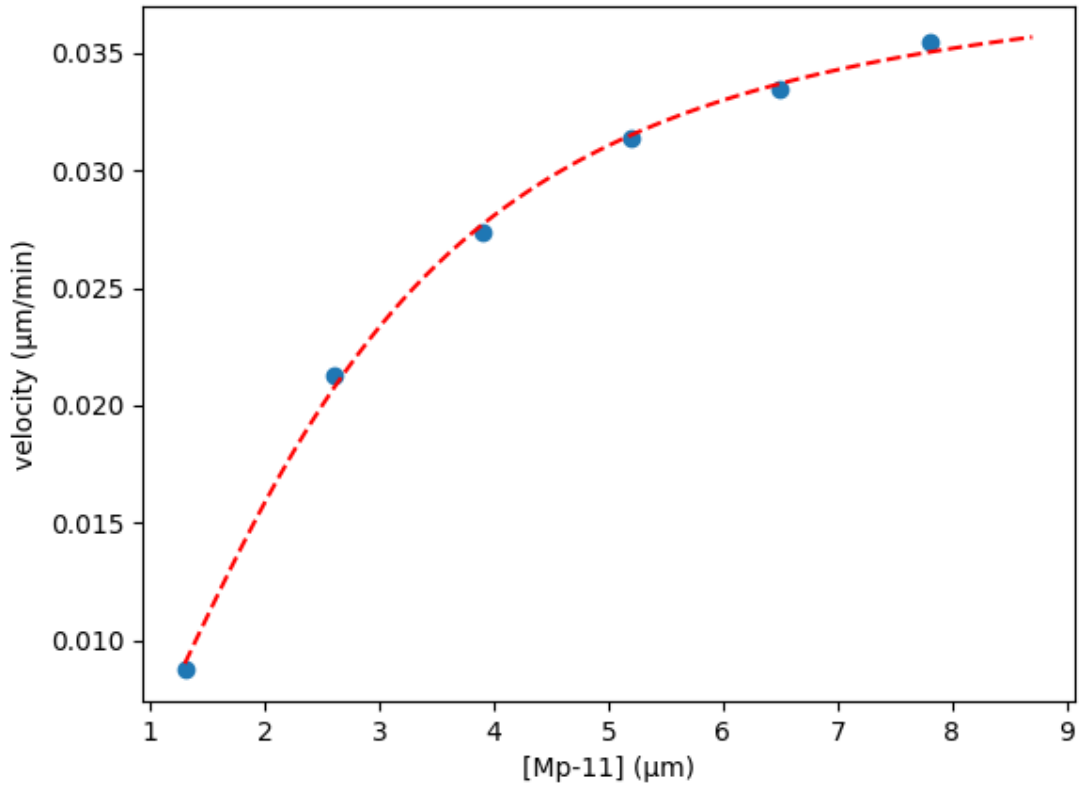


Figure 4.22 Initial rate of the Pden_1323 R92A reaction as a function of [MP-11] in the presence of 2 µM catalase. The dashed line shows best fit to the Hill equation using the kinetic parameters k_{cat} = 0.03 min^{-1}, K_m = 1.9 µM, and n = 2.41.

Table 4.2 Kinetic parameters for wild-type and mutant variants of Pden_1323 with MP-11.

Activity Parameter	WT	D127A	H64A	R92A
k_{cat} (min^{-1})	0.076	0.020	0.140	0.033
K_m (μM^{-1})	2.00	1.36	1.83	1.9
k_{cat}/K_m ($\mu M^{-1}min^{-1}$)	0.038	0.015	0.076	0.017

4.3.9 Crystral Structure of Pden_1323/MP-11

The cocrystallization conditions of Pden_1323 with MP-11 were determined using three sparse matrix screens (The JCSG Core IV, The JCSG + and The Protein Complex Suites, Qiagen). Crystals were produced under several conditions, and optimization plates were set up for the following condition: 12-18 % (w/v) PEG 6000, 0.1 M magnesium acetate, and 0.1 M sodium cacodylate (pH 6.5).

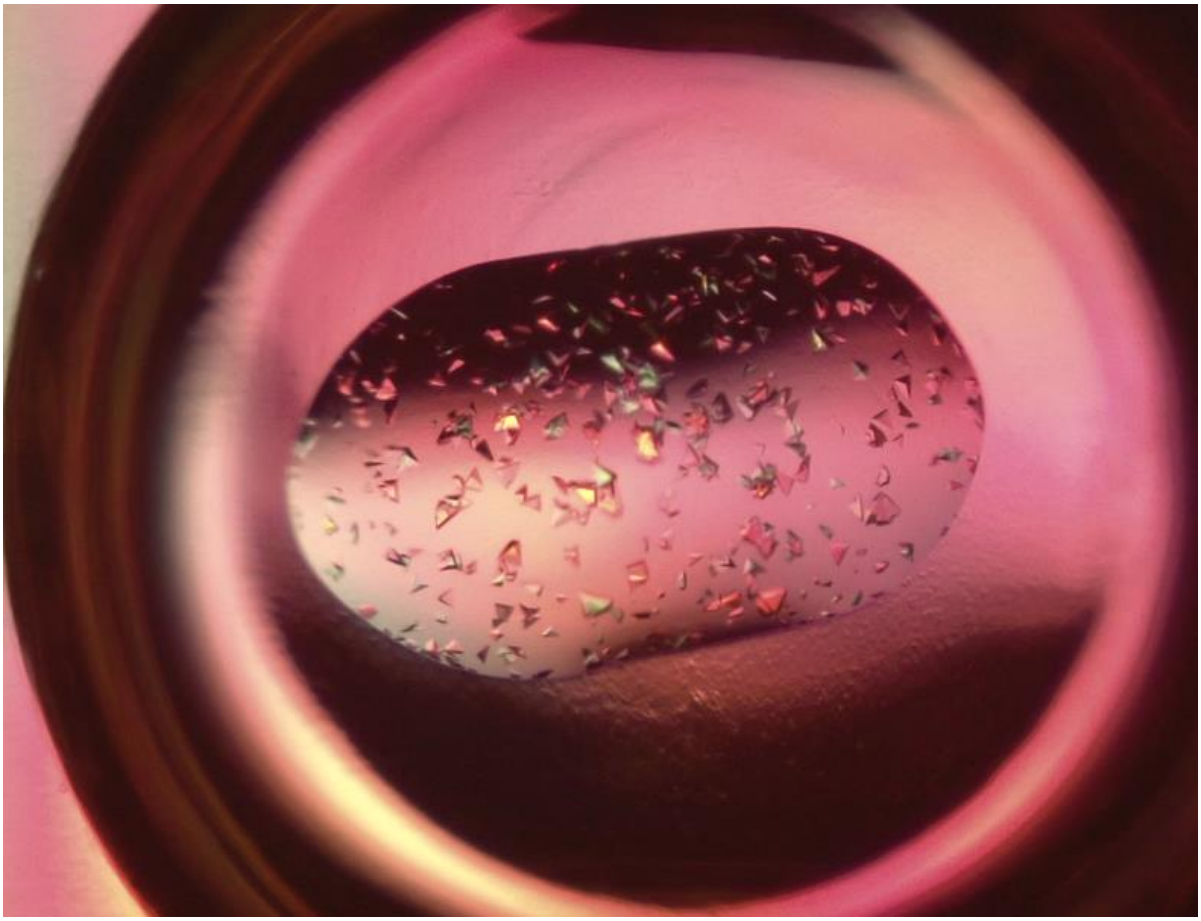


Figure 4.23 Crystals of Pden_1323 with MP-11. The crystal growth conditions consist of 15 % (w/v) PEG 6000, 0.1 M magnesium acetate, and 0.1 M sodium cacodylate (pH 6.5).

4.4 Discussion

Taken together, the activity assays demonstrate that Pden_1323 is a novel *c*-type heme oxygenase (CHO) that shows specificity for the degradation of *c*-type hemopeptides over free *b*-type heme. The results of the MS experiments and myoglobin/ferrozine assays further indicate that Pden_1323 catalyzes the oxidative cleavage of *c*-type hemopeptides to a peptide-linked β- and/or δ-biliverdin with concomitant formation of CO and free iron. The inability of Pden_1323 to cleave full-length cytochrome *c* shows that the peptide linkage must be small enough so that the porphyrin macrocycle is not buried and is able to access the active site, while the cleavage of hemin at high imidazole concentration indicates that the histidine lower axial ligand (rather than the *c*-type thioether linkages or the peptide sequence) is the crucial factor enabling catalysis. This suggests that Pden_1323 may bind a range of hemopeptide substrates in a way that creates a HugZ-like active site, with the histidine residue provided by the peptide substrate coordinating and activating the heme iron. After substrate binding, the mechanism of Pden_1323-catalyzed *c*-type heme degradation is expected to proceed analogously to that of HugZ.[173] The exact physiological role(s) of Pden_1323 in the soil microbe *P. denitrificans* has yet to be established. It may be involved in acquiring iron from environmental hemopeptides (e.g., from degraded extracellular cytochromes). However, given the large number *c*-type cytochromes utilized by this metabolically versatile bacterium, Pden_1323 may also play an important role in general cell maintenance by recycling iron from damaged cytochromes.[178]

The structural comparison of HugZ and Pden_1323 suggests that the change in substrate specificity from free heme *b* to a *c*-type hemopeptide involves the loss of the C-terminal loop containing the coordinating heme ligand and an opening of the active site to accommodate the larger substrate via loss of the N-terminal α/β domain.[173] Similar structural features are also

observed in the heme storage protein HutZ from *Vibrio cholerae* and the heme utilization protein HupZ from Group A streptococcus (GAS).[183, 184] HutZ is required for optimal heme utilization in *V. cholerae*, although no HO activity was initially observed with hemin.[183] However, subsequent work by Uchida and coworkers suggests that HutZ functions as a HO at pH values below 8.0, producing the typical oxyferrous heme, meso-hydroxyheme, and verdoheme intermediates with regiospecificity similar to HugZ.[185, 186] While HutZ lacks the N-terminal α/β domain, it retains the C-terminal loop containing the conserved histidine lower axial ligand, which is important for its heme degradation activity.[187] In contrast, HupZ lacks both the N-terminal α/β domain and C-terminal loop, and thus more closely resembles Pden_1323.[184] Interestingly, HupZ was shown to copurify with heme when it was overexpressed in *E. coli* grown in the presence of 5-aminolevulinic acid and iron and demonstrated modest hemin degrading activity.[184] However, this enzyme was expressed as a C-terminal His-tagged fusion protein, and it is possible that the affinity tag enhanced the activity of HupZ with free heme.[184] Indeed, during the revision of this manuscript, Traore et al. published a paper reinvestigating GAS HupZ and found that the weak heme-degradation activity (and an observed oligomerization state change) is associated with heme binding to the C-terminal His-tag[188]. Thus, it is possible that HupZ and other heme storage/utilization proteins may also show preference for the degradation of heme *c*.

The crystal structure of HugZ shows residues His64 and Arg92 form salt bridging interactions with the carboxylates of heme, and Asp127 forms a hydrogen bond with the proximal histidine.[189] We proposed these residues play analogous roles in the *c*-type heme oxygenase enzymatic mechanism of Pden_1323. The D127A mutation lost more than half of catalytic activity. The loss of the hydrogen bond between Asp127 and the proximal histidine may decrease electron donation to the iron, thereby compromising its ability to be reduced/react with

O_2 and resulting in loss of heme degradation. His64 and Arg92 are proposed to stabilize the heme propionates and account for the heme orientation. The R92A mutation lost more than half of catalytic activity, which is consistent with this hypothesis. However, the H64A mutation gained more than double the catalytic activity. Goulding and coworkers reported a single mutation on a residue that interacts with the heme propionate that caused an orientation change and produced a different product.[190] This suggests that the H64A mutation may change the orientation of Mp-11 and result in a more rapid and uncontrolled product oxidation. This could also be due to the production of more ROS that damages the substrate non-enzymatically or maybe greater enzyme stability. Additional experiments are required to confirm this hypothesis.

In developed countries, dietary heme accounts for approximately two-thirds of an average person's iron stores, though heme *b* present in consumed hemoglobin and myoglobin is considered to be the largest iron source.[191] However, acid or enzymatic hydrolysis of cytochromes in the gastrointestinal tract can result in soluble hemopeptides, as well as polypeptides known to increase the solubility of free heme.[192] Enterocytes utilize heme carrier protein 1 (HCP1) to move free heme into the cell from the intestinal lumen.[193] To the best of the authors' knowledge, HCP1 has not been shown to transport hemopeptides. Therefore, enteric bacteria with an appropriate oligopeptide permease (Opp) system and a homolog of Pden_1323 could gain a competitive advantage within the gut microbiome by exploiting this alternative iron source resulting from cytochrome degradation, with little competition for hemopeptide substrates from enterocytes. In the case of pathogenic bacteria (such as the aforementioned Group A streptococcus) that rely on this pathway for iron acquisition, inhibiting CHOs could lead to greatly reduced iron availability for the pathogen while leaving enterocytes relatively unaffected.

Further mechanistic studies of Pden_1323, and its possible inhibition mechanisms, is needed to investigate this hypothesis.

In summary, a novel HO from *P. denitrificans* with specificity for the degradation of *c*-type hemopeptides was identified and structurally characterized. The oxygenation reaction with MP-11 rapidly yields free iron, CO, and a peptide-linked β- and/or δ-biliverdin as products. The structural features conferring specificity for heme *c* are also observed in a number of heme storage/utilization proteins from pathogenic bacteria, suggesting that these organisms may also target the *c*-type heme pool as a source of iron to sustain infections.

Chapter Five: Discussion and future work

Natural gas is an important fuel that meets around a quarter of the word's energy demands.[14] Natural gas is an abundant resource across the United States. In recent years, the production of natural gas from shale gas has increased exponentially due to the discovery of new formations and extraction methods.[16] Given current production rates, this resource should provide enough natural gas to supply the United States for the next century. However, most natural gas is formed in remote areas and requires complicated transportation systems to deliver to end users.[23] As a result, natural gas is often flared or vented, which wastes this important commodity and causes large amounts of methane to be released into the atmosphere. Methane is more than 30 times as potent as carbon dioxide at trapping heat in the atmosphere, which is the second most abundant anthropogenic greenhouse gas.[27] Therefore, it is of great significant to convert methane into liquid fuels or other commodities that are easier to transport.

Methyl-coenzyme M reductase (MCR) catalyzes the final step in the formation of methane and the first step in the anaerobic oxidation of methane (AOM), which has much promise in natural gas to liquid fuel conversion.[59] Currently, MCR cannot be produced in an active form in a heterologous host, due to challenges associated with the *in vivo* biosynthesis of coenzyme F430 and the posttranslational modification (PTM) and maturation of MCR. In Chapter Two, we constructed a plasmid, pRSFDuet:*Mt_cfbCD-porABDG*, which contains an additional copy of the coenzyme F430 biosynthesis genes *cfbCD* from *Methanosarcina thermophila* and genes encoding the native physiological reducing system pyruvate:ferredoxin oxidoreductase (PFOR).[143] Several cell lines were constructed combining this plasmid with the *sirAC, cfb, isc, mcr*, and/or *mm* genes. The coenzyme F430 biosynthesis activity of these cell lines was tested using different growth times/temperatures after inducing with IPTG, different

amounts of supplementary chemicals, different strains of *E. coli*, and different growth conditions (aerobic or anaerobic). The cell line with the *mm* genes grown at 37 ˚C for 4 hours after inducing with IPTG gave the highest yield of the substrate Ni-sirohydrochlorin *a,c*-diamide. Unfortunately, coenzyme F430 or 15,17^3-seco-F430-17^3-acid were not detected in any of the cell lines, which indicates the pathway stops at the CfbCD reaction. There may be several potential reasons for this. We cloned the PFOR genes from *Methanosarcina acetivorans* and confirmed that they were successfully expressed in *E. coli*. However, the activity of recombinant PFOR has not been confirmed. We will extract PFOR from the coenzyme F430 biosynthesis cell lines and test its *in vitro* activity in future work. When the coenzyme F430 biosynthesis cell lines were grown under anaerobic conditions only a small amount of *E. coli* was obtained. The anaerobic growth conditions should therefore be optimized. Generating a cell line that combines all of the above genes (together with other potentially required accessory genes) may necessary. This can be facilitated by combining the pACYCDuet:*sirAC* and pRSFDuet:*Mt_cfbCD-porABDG* plasmids.

In Chapter Three, we explored the mechanism of CfbCD catalysis, which involves a multi-electron reduction of the isobacteriochlorin ring system of Ni-sirohydrochlorin *a,c*-diamide, cyclization of the *c*-acetamide side chain to form the γ-lactam E ring, and the formation of 7 stereocenters.[87] CfbC and CfbD are homologous to the Fe protein (NifH) and the NifD subunit of the MoFe protein (NifDK) of nitrogenase, respectively.[87] Nitrogenase and its homologs contain heterotetrameric MoFe protein-like components, while CfbD is homomeric. Meanwhile, Nitrogenase has a P-cluster that is thought to have evolved from two [4Fe-4S] clusters, while CfbD has a regular [4Fe-4S] cluster.[87] Thus, the CfbCD complex is representative of a more primitive lineage of the nitrogenase superfamily. Indeed, a molecular phylogenic analysis of the nitrogenase homologs placed the ancestor of CfbCD in the last common ancestor of modern

organisms and basal to the emergence of the lineages involved in nitrogen fixation and photosynthesis.[87] Overall, understanding the mechanism of CfbCD catalysis will provide significant insight into this unprecedented reductive cyclization reaction and will shed light on the mechanisms and evolution of the nitrogenase superfamily. NMR spectroscopy was utilized to gain insight into the structure of the immediate product of CfbCD catalysis and the stereochemistry of 15,17^3-seco-F430-17^3-acid. The missed nitrogen peak on the *c*-amide side chain of the initial product suggests that the Ni center might be in the +1 oxidation state, which is the active form of the coenzyme in MCR. Electron paramagnetic resonance (EPR) spectroscopy will be used to confirm this hypothesis. We found that the products of the CfbCD reaction remains bound to the enzyme complex and can be purified via buffer exchange. This method will be used to remove glycerol and other small molecules for the stereochemistry study of CfbCD catalysis. A novel substrate analog, Ni-didecarboxysirohydrochlorin *a*,*c*-diamide, was synthesized for future mechanistic studies of CfbCD catalysis. If this analog is accepted as a substrate of CfbCD, the structures of any intermediates/products formed will be determined using NMR, which may give insight into the order of proton addition and the involvement of substrate assisted catalysis.

In Chapter Four, we identified a heme oxygenase (HO) that is specific for the degradation of *c*-type hemopeptides. Iron is an essential element for life. Organisms have developed complicated iron regulation systems to obtain, transport, and store iron.[116] Iron acquisition is important for bacterial pathogens to establish and maintain infections.[150] Many pathogens synthesize and release siderophores to capture free ferric iron.[116] The Fe(III)-siderophore complex is reduced to the less tightly bound ferrous state and released in the cell.[152] However, in many hosts, iron is complexed with protoporphyrin IX in the form of heme.[153] Bacteria have

developed transport systems to intake heme from the environment and use HOs to break down heme and release the tightly bound iron.[116] All known HOs degrade free *b*-type heme.[156] Pden_1323 from *Paracoccus denitrificans* catalyzes the oxidative cleavage of *c*-type hemopeptides to a peptide-linked β- and/or δ-biliverdin with concomitant formation of CO and free iron. This study thus expands the scope of known iron acquisition strategies to include direct oxidative cleavage of heme-containing proteolytic fragments of *c*-type cytochromes and helps to explain why certain oligopeptide permeases show specificity for the import of heme in addition to peptides. To further study this novel enzyme and its catalysis mechanisms, we will cocrystallize wild-type and/or mutant variants of Pden_1323 with the substrate MP-11, which will help to understand the functions of key amino acids in the active site. Different *c*-type hemopeptides with varying lengths and amino acid sequences will be tested to explore the substrate specificity of this novel enzyme.

References

1. Battersby, A. R., Tetrapyrroles: the pigments of life. *Natural product reports* **2000,** *17* (6), 507-526.

2. Ponka, P., Cell biology of heme. *The American journal of the medical sciences* **1999,** *318* (4), 241-256.

3. Von Wettstein, D.; Gough, S.; Kannangara, C. G., Chlorophyll biosynthesis. *The Plant Cell* **1995,** *7* (7), 1039.

4. Roth, J. R.; Lawrence, J.; Bobik, T., Cobalamin (coenzyme B12): synthesis and biological significance. *Annual review of microbiology* **1996,** *50* (1), 137-181.

5. Stroupe, M. E.; Leech, H. K.; Daniels, D. S.; Warren, M. J.; Getzoff, E. D., CysG structure reveals tetrapyrrole-binding features and novel regulation of siroheme biosynthesis. *Nature Structural & Molecular Biology* **2003,** *10* (12), 1064-1073.

6. Jeon, K. W., *International review of cell and molecular biology*. Academic Press: 2013.

7. Bali, S.; Palmer, D. J.; Schroeder, S.; Ferguson, S. J.; Warren, M. J., Recent advances in the biosynthesis of modified tetrapyrroles: the discovery of an alternative pathway for the formation of heme and heme d 1. *Cellular and molecular life sciences* **2014,** *71* (15), 2837-2863.

8. Kwon, S. J.; De Boer, A. L.; Petri, R.; Schmidt-Dannert, C., High-level production of porphyrins in metabolically engineered Escherichia coli: systematic extension of a pathway assembled from overexpressed genes involved in heme biosynthesis. *Applied and environmental microbiology* **2003,** *69* (8), 4875-4883.

9. Smith, A.; Witty, M., *Heme, chlorophyll, and bilins: methods and protocols*. Springer Science & Business Media: 2002.

10. Jaffe, E. K.; Lawrence, S. H., Allostery and the dynamic oligomerization of porphobilinogen synthase. *Archives of biochemistry and biophysics* **2012,** *519* (2), 144-153.

11. Jordan, P. M.; Warren, M. J., Evidence for a dipyrromethane cofactor at the catalytic site of E. coli porphobilinogen deaminase. *FEBS letters* **1987,** *225* (1-2), 87-92.

12. Warren, M. J.; Stolowich, N. J.; Santander, P. J.; Roessner, C. A.; Sowa, B. A.; Scott, A. I., Enzymatic synthesis of dihydrosirohydrochlorin (precorrin‐2) and of a novel pyrrocorphin by uroporphyrinogen III methylase. *FEBS letters* **1990,** *261* (1), 76-80.

13. Sassa, S.; Kappas, A., Molecular aspects of the inherited porphyrias. *Journal of internal medicine* **2000,** *247* (2), 169-178.

14. Modern shale gas development in the United States: A primer. U.S. Department of Energy, Office of Fossil Energy, National Energy Technology Laboratory, 2009.

15. International energy outlook 2019 (IEO2019). U.S. Energy Information Administration, Office of Integrated Analysis and Forecasting, U.S. Department of Energy, 2019.

16. Annual Energy Outlook 2012: With Projections to 2035. U.S. Energy Information Administration, Office of Integrated and International Energy Analysis, U.S. Department of Energy, 2012.

17. Shale Gas and Other Unconventional Sources of Natural Gas. Union of Concerned Scientists, 2015.

18. Ruppel, C. D., Methane hydrates and contemporary climate change. *Nature Eduction Knowledge* **2011,** *2* (12), 12.

19. H. Karasov, W.; Douglas, A. E., Comparative digestive physiology. *Comprehensive Physiology* **2013,** *3* (2), 741-783.

20. Enzmann, F.; Mayer, F.; Rother, M.; Holtmann, D., Methanogens: biochemical background and biotechnological applications. *AMB Express* **2018,** *8* (1), 1-22.

21. World Carbon Dioxide Emissions from the Use of Fossil Fuels. U.S. Energy Information Administration, 2016.

22. Haynes, C. A.; Gonzalez, R., Rethinking biological activation of methane and conversion to liquid fuels. *Nature chemical biology* **2014,** *10* (5), 331.

23. Natural Gas Flaring, Processing, and Transportation. Union of Concerned Scientists, 2015.

24. Natural gas explained Delivery and storage of natural gas. U.S. Energy Information Administration, 2021.

25. Ulvestad, M.; Overland, I., Natural gas and CO2 price variation: impact on the relative cost-efficiency of LNG and pipelines. *International Journal of Environmental Studies* **2012,** *69* (3), 407-426.

26. Fessler, R. R., Pipeline corrosion. *Report, US Department of Transportation Pipeline and Hazardous Materials Safety Administration, Baker, Evanston, IL* **2008**.

27. Myhre, G.; Shindell, D.; Pongratz, J., Anthropogenic and natural radiative forcing. **2014**.

28. Allison, E.; Mandler, B., Methane Emissions in the Oil and Gas Industry. American Geosciences Institute, 2018.

29. Garcia, J.-L.; Patel, B. K.; Ollivier, B., Taxonomic, phylogenetic, and ecological diversity of methanogenic Archaea. *Anaerobe* **2000,** *6* (4), 205-226.

30. Stadtman, T. C.; Barker, H., STUDIES ON THE METHANE FERMENTATION X.: A New Formate-Decomposing Bacterium, Methanococcus vannielii1. *Journal of bacteriology* **1951,**

62 (3), 269.

31. Balch, W.; Fox, G. E.; Magrum, L. J.; Woese, C. R.; Wolfe, R. S., Methanogens: reevaluation of a unique biological group. *Microbiological reviews* **1979,** *43* (2), 260.

32. Kurr, M.; Huber, R.; König, H.; Jannasch, H. W.; Fricke, H.; Trincone, A.; Kristjansson, J. K.; Stetter, K. O., Methanopyrus kandleri, gen. and sp. nov. represents a novel group of hyperthermophilic methanogens, growing at 110 C. *Archives of Microbiology* **1991,** *156* (4), 239-247.

33. Sakai, S.; Imachi, H.; Hanada, S.; Ohashi, A.; Harada, H.; Kamagata, Y., Methanocella paludicola gen. nov., sp. nov., a methane-producing archaeon, the first isolate of the lineage 'Rice Cluster I', and proposal of the new archaeal order Methanocellales ord. nov. *International Journal of Systematic and Evolutionary Microbiology* **2008,** *58* (4), 929-936.

34. Dridi, B.; Fardeau, M.-L.; Ollivier, B.; Raoult, D.; Drancourt, M., Methanomassiliicoccus luminyensis gen. nov., sp. nov., a methanogenic archaeon isolated from human faeces. *International journal of systematic and evolutionary microbiology* **2012,** *62* (8), 1902-1907.

35. Iino, T.; Tamaki, H.; Tamazawa, S.; Ueno, Y.; Ohkuma, M.; Suzuki, K.-i.; Igarashi, Y.; Haruta, S., Candidatus Methanogranum caenicola: a novel methanogen from the anaerobic digested sludge, and proposal of Methanomassiliicoccaceae fam. nov. and Methanomassiliicoccales ord. nov., for a methanogenic lineage of the class Thermoplasmata. *Microbes and environments* **2013**, ME12189.

36. Bapteste, É.; Brochier, C.; Boucher, Y., Higher-level classification of the Archaea: evolution of methanogenesis and methanogens. *Archaea* **2005,** *1* (5), 353-363.

37. Fotidis, I. A.; Karakashev, D.; Kotsopoulos, T. A.; Martzopoulos, G. G.; Angelidaki, I., Effect of ammonium and acetate on methanogenic pathway and methanogenic community composition. *FEMS microbiology ecology* **2013,** *83* (1), 38-48.

38. Borrel, G.; O'Toole, P. W.; Harris, H. M.; Peyret, P.; Brugere, J.-F.; Gribaldo, S., Phylogenomic data support a seventh order of methylotrophic methanogens and provide insights into the evolution of methanogenesis. *Genome biology and evolution* **2013,** *5* (10), 1769-1780.

39. Brioukhanov, A.; Thauer, R.; Netrusov, A., Catalase and superoxide dismutase in the cells of strictly anaerobic microorganisms. *Microbiology* **2002,** *71* (3), 281-285.

40. Zhang, G.; Jiang, N.; Liu, X.; Dong, X., Methanogenesis from methanol at low temperatures by a novel psychrophilic methanogen,"Methanolobus psychrophilus" sp. nov., prevalent in Zoige wetland of the Tibetan plateau. *Applied and environmental microbiology* **2008,** *74* (19), 6114.

41. Roessler, M.; Pflüger, K.; Flach, H.; Lienard, T.; Gottschalk, G.; Müller, V., Identification of a salt-induced primary transporter for glycine betaine in the methanogen Methanosarcina mazei Gö1. *Applied and environmental microbiology* **2002,** *68* (5), 2133.

42. Oren, A., Microbial life at high salt concentrations: phylogenetic and metabolic diversity. *Saline systems* **2008,** *4* (1), 1-13.

43. Wagner, T.; Ermler, U.; Shima, S., The methanogenic CO2 reducing-and-fixing enzyme is bifunctional and contains 46 [4Fe-4S] clusters. *Science* **2016,** *354* (6308), 114-117.

44. Liu, Y.; Whitman, W. B., Metabolic, phylogenetic, and ecological diversity of the methanogenic archaea. *Annals of the New York Academy of Sciences* **2008,** *1125* (1), 171-189.

45. Thauer, R. K.; Kaster, A.-K.; Seedorf, H.; Buckel, W.; Hedderich, R., Methanogenic archaea: ecologically relevant differences in energy conservation. *Nature Reviews Microbiology* **2008,** *6* (8), 579-591.

46. Welte, C.; Deppenmeier, U., Bioenergetics and anaerobic respiratory chains of aceticlastic methanogens. *Biochimica et Biophysica Acta (BBA)-Bioenergetics* **2014,** *1837* (7), 1130-1147.

47. Burke, S. A.; Krzycki, J. A., Reconstitution of monomethylamine: Coenzyme M methyl transfer with a corrinoid protein and two methyltransferases purified fromMethanosarcina barkeri. *Journal of Biological Chemistry* **1997,** *272* (26), 16570-16577.

48. Mayer, F.; Müller, V., Adaptations of anaerobic archaea to life under extreme energy limitation. *FEMS microbiology reviews* **2014,** *38* (3), 449-472.

49. Buckel, W.; Thauer, R. K., Energy conservation via electron bifurcating ferredoxin reduction and proton/Na+ translocating ferredoxin oxidation. *Biochimica et Biophysica Acta (BBA)-Bioenergetics* **2013,** *1827* (2), 94-113.

50. Lyu, Z.; Shao, N.; Akinyemi, T.; Whitman, W. B., Methanogenesis. *Current Biology* **2018,** *28* (13), R727-R732.

51. Hanson, R. S.; Hanson, T. E., Methanotrophic bacteria. *Microbiological reviews* **1996,** *60* (2), 439.

52. Michaelis, W.; Seifert, R.; Nauhaus, K.; Treude, T.; Thiel, V.; Blumenberg, M.; Knittel, K.; Gieseke, A.; Peterknecht, K.; Pape, T., Microbial reefs in the Black Sea fueled by anaerobic oxidation of methane. *Science* **2002,** *297* (5583), 1013-1015.

53. Knittel, K.; Lösekann, T.; Boetius, A.; Kort, R.; Amann, R., Diversity and distribution of methanotrophic archaea at cold seeps. *Applied and environmental microbiology* **2005,** *71* (1), 467.

54. Cui, M.; Ma, A.; Qi, H.; Zhuang, X.; Zhuang, G., Anaerobic oxidation of methane: an “active” microbial process. *MicrobiologyOpen* **2015,** *4* (1), 1-11.

55. Schreiber, L.; Holler, T.; Knittel, K.; Meyerdierks, A.; Amann, R., Identification of the dominant sulfate‐reducing bacterial partner of anaerobic methanotrophs of the ANME‐2 clade. *Environmental microbiology* **2010,** *12* (8), 2327-2340.

56. Beal, E. J.; House, C. H.; Orphan, V. J., Manganese-and iron-dependent marine methane oxidation. *Science* **2009,** *325* (5937), 184-187.

57. Haroon, M. F.; Hu, S.; Shi, Y.; Imelfort, M.; Keller, J.; Hugenholtz, P.; Yuan, Z.; Tyson, G. W., Anaerobic oxidation of methane coupled to nitrate reduction in a novel archaeal lineage. *Nature* **2013,** *500* (7464), 567-570.

58. Ragsdale, S. W.; Raugei, S.; Ginovska, B.; Wongnate, T., Biochemistry of methyl-coenzyme M reductase. In *The Biological Chemistry of Nickel*, 2017; pp 149-169.

59. DiMarco, A. A.; Bobik, T. A.; Wolfe, R. S., Unusual coenzymes of methanogenesis. *Annual review of biochemistry* **1990,** *59* (1), 355-394.

60. ROSPERT, S.; LINDER, D.; ELLERMANN, J.; THAUER, R. K., Two genetically distinct methyl‐coenzyme M reductases in Methanobacterium thermoautotrophicum strain Marburg and ΔH. *European journal of Biochemistry* **1990,** *194* (3), 871-877.

61. Ermler, U.; Grabarse, W.; Shima, S.; Goubeaud, M.; Thauer, R. K., Crystal structure of methyl-coenzyme M reductase: the key enzyme of biological methane formation. *Science* **1997,** *278* (5342), 1457-1462.

62. Grabarse, W.; Mahlert, F.; Duin, E. C.; Goubeaud, M.; Shima, S.; Thauer, R. K.; Lamzin, V.; Ermler, U., On the mechanism of biological methane formation: structural evidence for conformational changes in methyl-coenzyme M reductase upon substrate binding. *Journal of molecular biology* **2001,** *309* (1), 315-330.

63. Cedervall, P. E.; Dey, M.; Pearson, A. R.; Ragsdale, S. W.; Wilmot, C. M., Structural insight into methyl-coenzyme M reductase chemistry using coenzyme B analogues. *Biochemistry* **2010,** *49* (35), 7683-7693.

64. Grabarse, W.; Mahlert, F.; Shima, S.; Thauer, R. K.; Ermler, U., Comparison of three methyl-coenzyme M reductases from phylogenetically distant organisms: unusual amino acid modification, conservation and adaptation. *Journal of molecular biology* **2000,** *303* (2), 329-344.

65. Horng, Y.-C.; Becker, D. F.; Ragsdale, S. W., Mechanistic studies of methane biogenesis by methyl-coenzyme M reductase: evidence that coenzyme B participates in cleaving the C− S bond of methyl-coenzyme M. *Biochemistry* **2001,** *40* (43), 12875-12885.

66. Kunz, R. C.; Dey, M.; Ragsdale, S. W., Characterization of the Thioether Product Formed from the Thiolytic Cleavage of the Alkyl− Nickel Bond in Methyl-Coenzyme M Reductase. *Biochemistry* **2008,** *47* (8), 2661-2667.

67. Dey, M.; Kunz, R. C.; Lyons, D. M.; Ragsdale, S. W., Characterization of alkyl-nickel adducts generated by reaction of methyl-coenzyme M reductase with brominated acids. *Biochemistry* **2007,** *46* (42), 11969-11978.

68. Dey, M.; Telser, J.; Kunz, R. C.; Lees, N. S.; Ragsdale, S. W.; Hoffman, B. M., Biochemical and Spectroscopic Studies of the Electronic Structure and Reactivity of a Methyl−

Ni Species Formed on Methyl-Coenzyme M Reductase. *Journal of the American Chemical Society* **2007,** *129* (36), 11030-11032.

69. Wongnate, T.; Sliwa, D.; Ginovska, B.; Smith, D.; Wolf, M. W.; Lehnert, N.; Raugei, S.; Ragsdale, S. W., The radical mechanism of biological methane synthesis by methyl-coenzyme M reductase. *Science* **2016,** *352* (6288), 953-958.

70. Pelmenschikov, V.; Blomberg, M. R.; Siegbahn, P. E.; Crabtree, R. H., A mechanism from quantum chemical studies for methane formation in methanogenesis. *Journal of the American Chemical Society* **2002,** *124* (15), 4039-4049.

71. Chen, S.-l.; Pelmenschikov, V.; Blomberg, M. R.; Siegbahn, P. E., Is there a Ni-methyl intermediate in the mechanism of methyl-coenzyme M reductase? *Journal of the American Chemical Society* **2009,** *131* (29), 9912-9913.

72. Pelmenschikov, V.; Siegbahn, P. E., Catalysis by methyl-coenzyme M reductase: a theoretical study for heterodisulfide product formation. *JBIC Journal of Biological Inorganic Chemistry* **2003,** *8* (6), 653-662.

73. Wagner, T.; Kahnt, J.; Ermler, U.; Shima, S., Didehydroaspartate Modification in Methyl‐Coenzyme M Reductase Catalyzing Methane Formation. *Angewandte Chemie* **2016,** *128* (36), 10788-10791.

74. Patwardhan, A.; Sarangi, R.; Ginovska, B.; Raugei, S.; Ragsdale, S. W., Nickel–Sulfonate Mode of Substrate Binding for Forward and Reverse Reactions of Methyl-SCoM Reductase Suggest a Radical Mechanism Involving Long-Range Electron Transfer. *Journal of the American Chemical Society* **2021,** *143* (14), 5481-5496.

75. Albracht, S.; Ankel-Fuchs, D.; Böcher, R.; Ellermann, J.; Moll, J.; Van der Zwaan, J.; Thauer, R., Five new EPR signals assigned to nickel in methyl-coenzyme M reductase from Methanobacterium thermoautotrophicum, strain Marburg. *Biochimica et Biophysica Acta (BBA)-Protein Structure and Molecular Enzymology* **1988,** *955* (1), 86-102.

76. Becker, D. F.; Ragsdale, S. W., Activation of methyl-SCoM reductase to high specific activity after treatment of whole cells with sodium sulfide. *Biochemistry* **1998,** *37* (8), 2639-2647.

77. Mansoorabadi, S. O.; Zheng, K.; Ngo, P. D., Biosynthesis of coenzyme F430 and the posttranslational modification of the active site region of methyl-coenzyme M reductase. *Metalloprotein Active Site Assembly* **2017**, 323.

78. Wagner, T.; Wegner, C.-E.; Kahnt, J.; Ermler, U.; Shima, S., Phylogenetic and structural comparisons of the three types of methyl coenzyme M reductase from Methanococcales and Methanobacteriales. *Journal of bacteriology* **2017,** *199* (16).

79. Selmer, T.; Kahnt, J.; Goubeaud, M.; Shima, S.; Grabarse, W.; Ermler, U.; Thauer, R. K., The biosynthesis of methylated amino acids in the active site region of methyl-coenzyme M reductase. *Journal of Biological Chemistry* **2000,** *275* (6), 3755-3760.

80. Bauerle, M. R.; Schwalm, E. L.; Booker, S. J., Mechanistic diversity of radical S-adenosylmethionine (SAM)-dependent methylation. *Journal of Biological Chemistry* **2015,** *290* (7), 3995-4002.

81. Deobald, D.; Adrian, L.; Schöne, C.; Rother, M.; Layer, G., Identification of a unique radical SAM methyltransferase required for the sp 3-C-methylation of an arginine residue of methyl-coenzyme M reductase. *Scientific reports* **2018,** *8* (1), 1-12.

82. Lyu, Z.; Shao, N.; Chou, C.-W.; Shi, H.; Patel, R.; Duin, E. C.; Whitman, W. B., Posttranslational methylation of arginine in methyl coenzyme M reductase has a profound impact on both methanogenesis and growth of Methanococcus maripaludis. *Journal of bacteriology* **2020,** *202* (3).

83. Radle, M. I.; Miller, D. V.; Laremore, T. N.; Booker, S. J., Methanogenesis marker protein 10 (Mmp10) from Methanosarcina acetivorans is a radical S-adenosylmethionine methylase that unexpectedly requires cobalamin. *Journal of Biological Chemistry* **2019,** *294* (31), 11712-11725.

84. Mahanta, N.; Liu, A.; Dong, S.; Nair, S. K.; Mitchell, D. A., Enzymatic reconstitution of ribosomal peptide backbone thioamidation. *Proceedings of the National Academy of Sciences* **2018,** *115* (12), 3030-3035.

85. Nayak, D. D.; Mahanta, N.; Mitchell, D. A.; Metcalf, W. W., Post-translational thioamidation of methyl-coenzyme M reductase, a key enzyme in methanogenic and methanotrophic Archaea. *Elife* **2017,** *6*, e29218.

86. Gunsalus, R.; Wolfe, R., Chromophoric factors F342 and F430 of Methanobacterium thermoautotrophicum. *FEMS Microbiology Letters* **1978,** *3* (4), 191-193.

87. Zheng, K.; Ngo, P. D.; Owens, V. L.; Yang, X.-p.; Mansoorabadi, S. O., The biosynthetic pathway of coenzyme F430 in methanogenic and methanotrophic archaea. *Science* **2016,** *354* (6310), 339-342.

88. Brindley, A. A.; Raux, E.; Leech, H. K.; Schubert, H. L.; Warren, M. J., A story of chelatase evolution: identification and characterization of a small 13–15-kDa “ancestral” cobaltochelatase (CbiXS) in the archaea. *Journal of Biological Chemistry* **2003,** *278* (25), 22388-22395.

89. Leech, H.; Raux-Deery, E.; Heathcote, P.; Warren, M., Production of cobalamin and sirohaem in Bacillus megaterium: an investigation into the role of the branchpoint chelatases sirohydrochlorin ferrochelatase (SirB) and sirohydrochlorin cobalt chelatase (CbiX). *Biochemical Society Transactions* **2002,** *30* (4), 610-613.

90. Schubert, H. L.; Raux, E.; Wilson, K. S.; Warren, M. J., Common chelatase design in the branched tetrapyrrole pathways of heme and anaerobic cobalamin synthesis. *Biochemistry* **1999,** *38* (33), 10660-10669.

91. Fresquet, V.; Williams, L.; Raushel, F. M., Mechanism of cobyrinic acid a, c-diamide

synthetase from Salmonella typhimurium LT2. *Biochemistry* **2004,** *43* (33), 10619-10627.

92. Debussche, L.; Thibaut, D.; Cameron, B.; Crouzet, J.; Blanche, F., Purification and characterization of cobyrinic acid a, c-diamide synthase from Pseudomonas denitrificans. *Journal of bacteriology* **1990,** *172* (11), 6239.

93. Burgess, B. K.; Lowe, D. J., Mechanism of molybdenum nitrogenase. *Chemical reviews* **1996,** *96* (7), 2983-3012.

94. Eady, R. R., Structure− function relationships of alternative nitrogenases. *Chemical reviews* **1996,** *96* (7), 3013-3030.

95. Christiansen, J.; Dean, D. R.; Seefeldt, L. C., Mechanistic features of the Mo-containing nitrogenase. *Annual review of plant biology* **2001,** *52* (1), 269-295.

96. Hoffman, B. M.; Lukoyanov, D.; Yang, Z.-Y.; Dean, D. R.; Seefeldt, L. C., Mechanism of nitrogen fixation by nitrogenase: the next stage. *Chemical reviews* **2014,** *114* (8), 4041-4062.

97. Hu, Y.; Ribbe, M. W., Nitrogenase and homologs. *JBIC Journal of Biological Inorganic Chemistry* **2015,** *20* (2), 435-445.

98. Howard, J. B.; Rees, D. C., Structural basis of biological nitrogen fixation. *Chemical reviews* **1996,** *96* (7), 2965-2982.

99. Thorneley, R. N.; Lowe, D. J., Kinetics and mechanism of the nitrogenase enzyme system. *Molybdenum enzymes* **1985,** *7*, 89-116.

100. Hoffman, B. M.; Lukoyanov, D.; Dean, D. R.; Seefeldt, L. C., Nitrogenase: a draft mechanism. *Accounts of chemical research* **2013,** *46* (2), 587-595.

101. Schrock, R. R., Catalytic reduction of dinitrogen to ammonia at a single molybdenum center. *Accounts of chemical research* **2005,** *38* (12), 955-962.

102. Kästner, J.; Blöchl, P. E., Ammonia production at the FeMo cofactor of nitrogenase: results from density functional theory. *Journal of the American Chemical Society* **2007,** *129* (10), 2998-3006.

103. Hinnemann, B.; Nørskov, J. K., Catalysis by enzymes: the biological ammonia synthesis. *Topics in Catalysis* **2006,** *37* (1), 55-70.

104. Robson, R. L.; Eady, R. R.; Richardson, T. H.; Miller, R. W.; Hawkins, M.; Postgate, J. R., The alternative nitrogenase of Azotobacter chroococcum is a vanadium enzyme. *Nature* **1986,** *322* (6077), 388-390.

105. Rebelein, J. G.; Hu, Y.; Ribbe, M. W., Differential reduction of CO2 by molybdenum and vanadium nitrogenases. *Angewandte Chemie* **2014,** *126* (43), 11727-11730.

106. Hu, Y.; Lee, C. C.; Ribbe, M. W., Extending the carbon chain: hydrocarbon formation

catalyzed by vanadium/molybdenum nitrogenases. *Science* **2011,** *333* (6043), 753-755.

107. Bertrand, J. A.; Auger, G.; Fanchon, E.; Martin, L.; Blanot, D.; Van Heijenoort, J.; Dideberg, O., Crystal structure of UDP‐N‐acetylmuramoyl‐L‐alanine: D‐glutamate ligase from Escherichia coli. *The EMBO journal* **1997,** *16* (12), 3416-3425.

108. PRATVIEL‐SOSA, F.; MENGIN‐LECREULX, D.; van HEIJENOORT, J., Over‐production, purification and properties of the uridine diphosphate N‐acetylmuramoyl‐l‐alanine: d‐glutamate ligase from Escherichia coli. *European journal of biochemistry* **1991,** *202* (3), 1169-1176.

109. Wallace, D. F., The regulation of iron absorption and homeostasis. *The Clinical Biochemist Reviews* **2016,** *37* (2), 51.

110. Mihailescu, M.-R.; Russu, I. M., A signature of the T→ R transition in human hemoglobin. *Proceedings of the National Academy of Sciences* **2001,** *98* (7), 3773-3777.

111. Ordway, G. A.; Garry, D. J., Myoglobin: an essential hemoprotein in striated muscle. *Journal of Experimental Biology* **2004,** *207* (20), 3441-3446.

112. Lu, Y., Assembly and transfer of iron–sulfur clusters in the plastid. *Frontiers in plant science* **2018,** *9*, 336.

113. Knaff, D. B.; Hirasawa, M., Ferredoxin-dependent chloroplast enzymes. *Biochimica et Biophysica Acta (BBA)-Bioenergetics* **1991,** *1056* (2), 93-125.

114. Denisov, I. G.; Makris, T. M.; Sligar, S. G.; Schlichting, I., Structure and chemistry of cytochrome P450. *Chemical reviews* **2005,** *105* (6), 2253-2278.

115. Tehrani, H. S.; Moosavi-Movahedi, A. A., Catalase and its mysteries. *Progress in Biophysics and Molecular Biology* **2018,** *140*, 5-12.

116. Andrews, S. C.; Robinson, A. K.; Rodríguez-Quiñones, F., Bacterial iron homeostasis. *FEMS microbiology reviews* **2003,** *27* (2-3), 215-237.

117. Richard, K. L.; Kelley, B. R.; Johnson, J. G., Heme uptake and utilization by gram-negative bacterial pathogens. *Frontiers in cellular and infection microbiology* **2019,** *9*, 81.

118. Fukushima, T.; Allred, B. E.; Raymond, K. N., Direct evidence of iron uptake by the Gram-positive siderophore-shuttle mechanism without iron reduction. *ACS chemical biology* **2014,** *9* (9), 2092-2100.

119. Andrews, S. C., Iron storage in bacteria. *Advances in microbial physiology* **1998,** *40*, 281-351.

120. Rivera, M., Bacterioferritin: structure, dynamics, and protein–protein interactions at play in iron storage and mobilization. *Accounts of chemical research* **2017,** *50* (2), 331-340.

121. Weeratunga, S. K.; Gee, C. E.; Lovell, S.; Zeng, Y.; Woodin, C. L.; Rivera, M., Binding of Pseudomonas aeruginosa apobacterioferritin-associated ferredoxin to bacterioferritin B promotes heme mediation of electron delivery and mobilization of core mineral iron. *Biochemistry* **2009,** *48* (31), 7420-7431.

122. Yao, H.; Wang, Y.; Lovell, S.; Kumar, R.; Ruvinsky, A. M.; Battaile, K. P.; Vakser, I. A.; Rivera, M., The structure of the BfrB–Bfd complex reveals protein–protein interactions enabling iron release from bacterioferritin. *Journal of the American Chemical Society* **2012,** *134* (32), 13470-13481.

123. Fridovich, I., Superoxide radical and superoxide dismutases. *Annual review of biochemistry* **1995,** *64* (1), 97-112.

124. Woodmansee, A. N.; Imlay, J. A., Reduced Flavins Promote Oxidative DNA Damage in Non-respiringEscherichia coli by Delivering Electrons to Intracellular Free Iron. *Journal of Biological Chemistry* **2002,** *277* (37), 34055-34066.

125. Hantke, K.; Braun, V., The art of keeping low and high iron concentrations in balance. *Bacterial stress responses. ASM Press, Washington, DC* **2000**, 275-288.

126. Hantke, K., Iron and metal regulation in bacteria. *Current opinion in microbiology* **2001,** *4* (2), 172-177.

127. Casiday, R.; Frey, R., Hemoglobin and the heme group: Metal complexes in the blood for oxygen transport. *Retrieved March* **1998,** *28*, 2008.

128. Paoli, M.; Marles-Wright, J.; Smith, A., Structure–function relationships in heme-proteins. *DNA and cell biology* **2002,** *21* (4), 271-280.

129. Ow, Y.-L. P.; Green, D. R.; Hao, Z.; Mak, T. W., Cytochrome c: functions beyond respiration. *Nature reviews Molecular cell biology* **2008,** *9* (7), 532-542.

130. Allen, J. W.; Ferguson, S. J.; Fülöp, V., Cytochrome cd 1 nitrite reductase. *Encyclopedia of Inorganic and Bioinorganic Chemistry* **2011**.

131. Yoshikawa, S.; Muramoto, K.; Shinzawa-Itoh, K.; Mochizuki, M., Structural studies on bovine heart cytochrome c oxidase. *Biochimica et Biophysica Acta (BBA)-Bioenergetics* **2012,** *1817* (4), 579-589.

132. Letoffe, S.; Ghigo, J.; Wandersman, C., Iron acquisition from heme and hemoglobin by a Serratia marcescens extracellular protein. *Proceedings of the National Academy of Sciences* **1994,** *91* (21), 9876-9880.

133. Occhino, D. A.; Wyckoff, E. E.; Henderson, D. P.; Wrona, T. J.; Payne, S. M., Vibrio cholerae iron transport: haem transport genes are linked to one of two sets of tonB, exbB, exbD genes. *Molecular microbiology* **1998,** *29* (6), 1493-1507.

134. Wandersman, C., Protein and peptide secretion by ABC exporters. *Research in*

microbiology **1998,** *149* (3), 163-170.

135. Brown, J. S.; Holden, D. W., Iron acquisition by Gram-positive bacterial pathogens. *Microbes and infection* **2002,** *4* (11), 1149-1156.

136. Matsui, T.; Unno, M.; Ikeda-Saito, M., Heme oxygenase reveals its strategy for catalyzing three successive oxygenation reactions. *Accounts of chemical research* **2010,** *43* (2), 240-247.

137. LaMattina, J. W.; Nix, D. B.; Lanzilotta, W. N., Radical new paradigm for heme degradation in Escherichia coli O157: H7. *Proceedings of the National Academy of Sciences* **2016,** *113* (43), 12138-12143.

138. Meyer, L.; Brinkman, S.; van Kesteren, L.; Leprince-Ringuet, N.; van Boxmeer, F., IPCC, 2014: climate change 2014: synthesis report. IPCC, Geneva, Switzerland, 2014.

139. Mueller, T. J.; Grisewood, M. J.; Nazem-Bokaee, H.; Gopalakrishnan, S.; Ferry, J. G.; Wood, T. K.; Maranas, C. D., Methane oxidation by anaerobic archaea for conversion to liquid fuels. *Journal of industrial microbiology & biotechnology* **2015,** *42* (3), 391-401.

140. Mobley, H. L.; Garner, R. M.; Bauerfeind, P., Helicobacter pylori nickel‐transport gene nixA: synthesis of catalytically active urease in Escherichia coli independent of growth conditions. *Molecular microbiology* **1995,** *16* (1), 97-109.

141. Frazzon, J.; Fick, J.; Dean, D., Biosynthesis of iron-sulphur clusters is a complex and highly conserved process. Portland Press Ltd.: 2002.

142. Staples, C. R.; Lahiri, S.; Raymond, J.; Von Herbulis, L.; Mukhophadhyay, B.; Blankenship, R. E., Expression and association of group IV nitrogenase NifD and NifH homologs in the non-nitrogen-fixing archaeon Methanocaldococcus jannaschii. *Journal of bacteriology* **2007,** *189* (20), 7392-7398.

143. Ragsdale, S. W., Pyruvate ferredoxin oxidoreductase and its radical intermediate. *Chemical Reviews* **2003,** *103* (6), 2333-2346.

144. Proctor, L. M.; Lai, R.; Gunsalus, R. P., The methanogenic archaeon Methanosarcina thermophila TM-1 possesses a high-affinity glycine betaine transporter involved in osmotic adaptation. *Applied and Environmental Microbiology* **1997,** *63* (6), 2252-2257.

145. Tolia, N. H.; Joshua-Tor, L., Strategies for protein coexpression in Escherichia coli. *Nature methods* **2006,** *3* (1), 55-64.

146. Muraki, N.; Nomata, J.; Ebata, K.; Mizoguchi, T.; Shiba, T.; Tamiaki, H.; Kurisu, G.; Fujita, Y., X-ray crystal structure of the light-independent protochlorophyllide reductase. *Nature* **2010,** *465* (7294), 110-114.

147. Nomata, J.; Mizoguchi, T.; Tamiaki, H.; Fujita, Y., A second nitrogenase-like enzyme for bacteriochlorophyll biosynthesis reconstitution of chlorophyllide a reductase with purified x-

protein (bchX) and yz-protein (bchY-bchZ) from Rhodobacter capsulatus. *Journal of Biological Chemistry* **2006,** *281* (21), 15021-15028.

148. Hu, Y.; Ribbe, M. W., Biosynthesis of the metalloclusters of nitrogenases. *Annual review of biochemistry* **2016,** *85*, 455-483.

149. Moore, S. J.; Sowa, S. T.; Schuchardt, C.; Deery, E.; Lawrence, A. D.; Ramos, J. V.; Billig, S.; Birkemeyer, C.; Chivers, P. T.; Howard, M. J., Elucidation of the biosynthesis of the methane catalyst coenzyme F 430. *Nature* **2017,** *543* (7643), 78-82.

150. Ratledge, C.; Dover, L. G., Iron metabolism in pathogenic bacteria. *Annual reviews in microbiology* **2000,** *54* (1), 881-941.

151. Wilson, B. R.; Bogdan, A. R.; Miyazawa, M.; Hashimoto, K.; Tsuji, Y., Siderophores in iron metabolism: from mechanism to therapy potential. *Trends in molecular medicine* **2016,** *22* (12), 1077-1090.

152. Schalk, I. J.; Guillon, L., Fate of ferrisiderophores after import across bacterial outer membranes: different iron release strategies are observed in the cytoplasm or periplasm depending on the siderophore pathways. *Amino acids* **2013,** *44* (5), 1267-1277.

153. Tong, Y.; Guo, M., Bacterial heme-transport proteins and their heme-coordination modes. *Archives of biochemistry and biophysics* **2009,** *481* (1), 1-15.

154. Frankenberg-Dinkel, N., Bacterial heme oxygenases. *Antioxidants & redox signaling* **2004,** *6* (5), 825-834.

155. Wilks, A.; Heinzl, G., Heme oxygenation and the widening paradigm of heme degradation. *Archives of biochemistry and biophysics* **2014,** *544*, 87-95.

156. Lyles, K. V.; Eichenbaum, Z., From host heme to iron: the expanding spectrum of heme degrading enzymes used by pathogenic bacteria. *Frontiers in cellular and infection microbiology* **2018,** *8*, 198.

157. Tenhunen, R.; Marver, H. S.; Schmid, R., The enzymatic conversion of heme to bilirubin by microsomal heme oxygenase. *Proceedings of the National Academy of Sciences of the United States of America* **1968,** *61* (2), 748.

158. Tenhunen, R.; Marver, H. S.; Schmid, R., Microsomal heme oxygenase: characterization of the enzyme. *Journal of Biological Chemistry* **1969,** *244* (23), 6388-6394.

159. Sugishima, M.; Omata, Y.; Kakuta, Y.; Sakamoto, H.; Noguchi, M.; Fukuyama, K., Crystal structure of rat heme oxygenase-1 in complex with heme. *FEBS letters* **2000,** *471* (1), 61-66.

160. Schmitt, M. P., Utilization of host iron sources by Corynebacterium diphtheriae: identification of a gene whose product is homologous to eukaryotic heme oxygenases and is required for acquisition of iron from heme and hemoglobin. *Journal of bacteriology* **1997,** *179*

(3), 838-845.

161. Zhu, W.; Wilks, A.; Stojiljkovic, I., Degradation of heme in gram-negative bacteria: the product of the hemO gene of Neisseriae is a heme oxygenase. *Journal of bacteriology* **2000,** *182* (23), 6783-6790.

162. Ratliff, M.; Zhu, W.; Deshmukh, R.; Wilks, A.; Stojiljkovic, I., Homologues of neisserial heme oxygenase in gram-negative bacteria: degradation of heme by the product of the pigA gene of Pseudomonas aeruginosa. *Journal of bacteriology* **2001,** *183* (21), 6394.

163. Wu, R.; Skaar, E. P.; Zhang, R.; Joachimiak, G.; Gornicki, P.; Schneewind, O.; Joachimiak, A., Staphylococcus aureus IsdG and IsdI, heme-degrading enzymes with structural similarity to monooxygenases. *Journal of Biological Chemistry* **2005,** *280* (4), 2840-2846.

164. Acharya, G.; Kaur, G.; Subramanian, S., Evolutionary relationships between heme-binding ferredoxin α+ β barrels. *BMC bioinformatics* **2016,** *17* (1), 1-11.

165. Reniere, M. L.; Ukpabi, G. N.; Harry, S. R.; Stec, D. F.; Krull, R.; Wright, D. W.; Bachmann, B. O.; Murphy, M. E.; Skaar, E. P., The IsdG‐family of haem oxygenases degrades haem to a novel chromophore. *Molecular microbiology* **2010,** *75* (6), 1529-1538.

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

167. Nambu, S.; Matsui, T.; Goulding, C. W.; Takahashi, S.; Ikeda-Saito, M., A new way to degrade heme: the Mycobacterium tuberculosis enzyme MhuD catalyzes heme degradation without generating CO. *Journal of Biological Chemistry* **2013,** *288* (14), 10101-10109.

168. Schneider, S.; Sharp, K. H.; Barker, P. D.; Paoli, M., An induced fit conformational change underlies the binding mechanism of the heme transport proteobacteria-protein HemS. *Journal of Biological Chemistry* **2006,** *281* (43), 32606-32610.

169. Onzuka, M.; Sekine, Y.; Uchida, T.; Ishimori, K.; Ozaki, S.-i., HmuS from Yersinia pseudotuberculosis is a non-canonical heme-degrading enzyme to acquire iron from heme. *Biochimica et Biophysica Acta (BBA)-General Subjects* **2017,** *1861* (7), 1870-1878.

170. Ouellet, Y. H.; Ndiaye, C. T.; Gagné, S. M.; Sebilo, A.; Suits, M. D.; Jubinville, É.; Jia, Z.; Ivancich, A.; Couture, M., An alternative reaction for heme degradation catalyzed by the Escherichia coli O157: H7 ChuS protein: Release of hematinic acid, tripyrrole and Fe (III). *Journal of inorganic biochemistry* **2016,** *154*, 103-113.

171. Guo, Y.; Guo, G.; Mao, X.; Zhang, W.; Xiao, J.; Tong, W.; Liu, T.; Xiao, B.; Liu, X.; Feng, Y., Functional identification of HugZ, a heme oxygenase from Helicobacter pylori. *BMC microbiology* **2008,** *8* (1), 1-11.

172. Ahmed, F. H.; Carr, P. D.; Lee, B. M.; Afriat-Jurnou, L.; Mohamed, A. E.; Hong, N.-S.; Flanagan, J.; Taylor, M. C.; Greening, C.; Jackson, C. J., Sequence–structure–function

classification of a catalytically diverse oxidoreductase superfamily in Mycobacteria. *Journal of molecular biology* **2015,** *427* (22), 3554-3571.

173. Hu, Y.; Jiang, F.; Guo, Y.; Shen, X.; Zhang, Y.; Zhang, R.; Guo, G.; Mao, X.; Zou, Q.; Wang, D.-C., Crystal structure of HugZ, a novel heme oxygenase from Helicobacter pylori. *Journal of Biological Chemistry* **2011,** *286* (2), 1537-1544.

174. Bowman, S. E.; Bren, K. L., The chemistry and biochemistry of heme c: functional bases for covalent attachment. *Natural product reports* **2008,** *25* (6), 1118-1130.

175. Yoshinaga, T.; Sassa, S.; Kappas, A., The oxidative degradation of heme c by the microsomal heme oxygenase system. *Journal of Biological Chemistry* **1982,** *257* (13), 7803-7807.

176. Kutty, R. K.; Maines, M., Oxidation of heme c derivatives by purified heme oxygenase. Evidence for the presence of one molecular species of heme oxygenase in the rat liver. *Journal of Biological Chemistry* **1982,** *257* (17), 9944-9952.

177. Baker, S. C.; Ferguson, S. J.; Ludwig, B.; Page, M. D.; Richter, O.-M. H.; van Spanning, R. J., Molecular genetics of the genus Paracoccus: metabolically versatile bacteria with bioenergetic flexibility. *Microbiology and Molecular Biology Reviews* **1998,** *62* (4), 1046-1078.

178. Ferguson, S. J., The functions and synthesis of bacterial c-type cytochromes with particular reference to Paracoccus denitrificans and Rhodobacter capsulatus. *Biochimica et Biophysica Acta (BBA)-Bioenergetics* **1991,** *1058* (1), 17-20.

179. Slamti, L.; Lereclus, D., The oligopeptide ABC-importers are essential communication channels in Gram-positive bacteria. *Research in microbiology* **2019,** *170* (8), 338-344.

180. Scheepers, G. H.; Lycklama a Nijeholt, J. A.; Poolman, B., An updated structural classification of substrate‐binding proteins. *FEBS letters* **2016,** *590* (23), 4393-4401.

181. Tanaka, K. J.; Pinkett, H. W., Oligopeptide-binding protein from nontypeable Haemophilus influenzae has ligand-specific sites to accommodate peptides and heme in the binding pocket. *Journal of Biological Chemistry* **2019,** *294* (3), 1070-1082.

182. Tu, A.; Reinosa, J.; Hsiao, Y., Peroxidative activity of hemepeptides from horse heart cytochrome c. *Experientia* **1968,** *24* (3), 219-221.

183. Liu, X.; Gong, J.; Wei, T.; Wang, Z.; Du, Q.; Zhu, D.; Huang, Y.; Xu, S.; Gu, L., Crystal structure of HutZ, a heme storage protein from Vibrio cholerae: A structural mismatch observed in the region of high sequence conservation. *BMC structural biology* **2012,** *12* (1), 1-10.

184. Sachla, A. J.; Ouattara, M.; Romero, E.; Agniswamy, J.; Weber, I. T.; Gadda, G.; Eichenbaum, Z., In vitro heme biotransformation by the HupZ enzyme from group A streptococcus. *Biometals* **2016,** *29* (4), 593-609.

185. Uchida, T.; Sekine, Y.; Matsui, T.; Ikeda-Saito, M.; Ishimori, K., A heme degradation enzyme, HutZ, from Vibrio cholerae. *Chemical communications* **2012,** *48* (53), 6741-6743.

186. Uchida, T.; Sekine, Y.; Dojun, N.; Lewis-Ballester, A.; Ishigami, I.; Matsui, T.; Yeh, S.-R.; Ishimori, K., Reaction intermediates in the heme degradation reaction by HutZ from Vibrio cholerae. *Dalton transactions* **2017,** *46* (25), 8104-8109.

187. Uchida, T.; Dojun, N.; Sekine, Y.; Ishimori, K., Heme proximal hydrogen bonding between His170 and Asp132 plays an essential role in the heme degradation reaction of HutZ from Vibrio cholerae. *Biochemistry* **2017,** *56* (21), 2723-2734.

188. Traore, E. S.; Li, J.; Chiura, T.; Geng, J.; Sachla, A. J.; Yoshimoto, F.; Eichenbaum, Z.; Davis, I.; Mak, P. J.; Liu, A., Heme binding to HupZ with a C-terminal tag from group A Streptococcus. *Molecules* **2021,** *26* (3), 549.

189. Uchida, T., Dojun, N., Sekine, Y., & Ishimori, K, Heme proximal hydrogen bonding between His170 and Asp132 plays an essential role in the heme degradation reaction of HutZ from Vibrio cholerae. *Biochemistry* **2017,** *56* (21), 2723-2734.

190. Chao, A.; Goulding, C. W., A single mutation in the Mycobacterium tuberculosis heme-degrading protein, MhuD, results in different products. *Biochemistry* **2019,** *58* (6), 489-492.

191. West, A. R.; Oates, P. S., Mechanisms of heme iron absorption: current questions and controversies. *World journal of gastroenterology: WJG* **2008,** *14* (26), 4101.

192. Hooda, J.; Shah, A.; Zhang, L., Heme, an essential nutrient from dietary proteins, critically impacts diverse physiological and pathological processes. *Nutrients* **2014,** *6* (3), 1080-1102.

193. Przybyszewska, J.; Żekanowska, E., The role of hepcidin, ferroportin, HCP1, and DMT1 protein in iron absorption in the human digestive tract. *Przeglad gastroenterologiczny* **2014,** *9* (4), 208.