1-1-2019

New Insights into Mycofactocin Biosynthesis, Structure and Function

Richard Selorm Ayikpoe
University of Denver

Recommended Citation

Ayikpoe, Richard Selorm, "New Insights into Mycofactocin Biosynthesis, Structure and Function" (2019). *Electronic Theses and Dissertations*. 1645.
https://digitalcommons.du.edu/etd/1645

New Insights into Mycofactocin Biosynthesis, Structure and Function

Abstract

Mycofactocin is a putative ribosomally synthesized and post-translationally modified peptide (RiPP)-derived redox cofactor. Its biosynthesis is accomplished through the dedicated actions of the products of six conserved genes, *mftABCDEF*. The mycofactocin pathway is one of the most widely distributed RiPP systems in bacteria however, this distribution is heavily skewed towards the *Mycobacteria* genus including human pathogenic variants such as *M. tuberculosis and M. ulcerans*. Gene expression studies have demonstrated the essentiality of the pathway in the ability of M. tuberculosis to utilize the host's cholesterol as sole carbon source during latency. However, the biosynthesis, structure and physiological function of mycofactocin remain enigmatic. Current efforts to elucidate the biosynthesis, structure and function of mycofactocin have focused on *in vitro* reconstitution of each enzyme in the pathway to gain insights into their role and function.

The biosynthesis of mycofactocin commences with the ribosomal production of the precursor peptide MftA containing conserved C-terminal residues - IDGMCGVY. In the presence of the RRE domain MftB, the RS-SPASM enzyme MftC, catalyzes the SAM-dependent oxidative decarboxylation and carbon-carbon bond formation on MftA to form MftA*. The roles of the auxiliary [Fe-S] clusters in MftC catalysis as well as the subsequent steps in the biosynthesis of mycofactocin are not known. Here, we have provided additional information regarding the roles of the auxiliary clusters in MftC. We showed that MftC contains three [4Fe-4S] clusters, all of which are required for catalysis. In addition, we measured the midpoint potentials of the clusters to provide insights into the redox flipping mechanism of MftC. Furthermore, we reconstituted the activity of MftE and showed that it selectively hydrolyzes MftA* to form MftA (1-28) and a 3-amino-5-(4-hydroxybenzyl)-4,4-dimethylpyrrolidin-2-one, herein referred to as AHDP. From this study, we have clarified the misunderstandings surrounding the accurate precursor for mycofactocin biosynthesis. Subsequently, we reconstituted the activity of MftD and showed that it catalyzes the oxidative deamination of AHDP to form an Î±-keto moiety herein referred to as premycofactocin. Lastly, we measured the midpoint potential of premycofactocin to be ~ -255 mV and demonstrated that it is used by mycofactocin-associated short chain dehydrogenases for multiple catalytic turnover.

Document Type
Dissertation

Degree Name
Ph.D.

Department
Chemistry and Biochemistry

First Advisor
John A. Latham, Ph.D.

Second Advisor
Martin Margiatti, Ph.D.

Third Advisor
Michelle Knowles, Ph.D.

Keywords
Mossbauer, Iron-sulfur clusters, MftC, Mycofactocin, Oxidative deamination, Redox cofactor, Ribosomally synthesized, Post translationally modified peptide

Subject Categories
Cell and Developmental Biology | Genetics and Genomics | Life Sciences

New Insights into Mycofactocin Biosynthesis, Structure and Function

A Dissertation

Presented to

the Faculty of Natural Sciences and Mathematics

University of Denver

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

by

Richard S. Ayikpoe

November 2019

Advisor: Dr. John A. Latham

Author: Richard S. Ayikpoe
Title: New Insights into Mycofactocin Biosynthesis, Structure and Function
Advisor: Dr. John A. Latham
Degree Date: November 2019

ABSTRACT

Mycofactocin is a putative ribosomally synthesized and post-translationally modified peptide (RiPP)-derived redox cofactor. Its biosynthesis is accomplished through the dedicated actions of the products of six conserved genes, *mftABCDEF*. The mycofactocin pathway is one of the most widely distributed RiPP systems in bacteria however, this distribution is heavily skewed towards the *Mycobacteria* genus including human pathogenic variants such as *M. tuberculosis and M. ulcerans*. Gene expression studies have demonstrated the essentiality of the pathway in the ability of *M. tuberculosis* to utilize the host's cholesterol as sole carbon source during latency. However, the biosynthesis, structure and physiological function of mycofactocin remain enigmatic. Current efforts to elucidate the biosynthesis, structure and function of mycofactocin have focused on *in vitro* reconstitution of each enzyme in the pathway to gain insights into their role and function.

The biosynthesis of mycofactocin commences with the ribosomal production of the precursor peptide MftA containing conserved C-terminal residues –IDGMCGVY. In the presence of the RRE domain MftB, the RS-SPASM enzyme MftC, catalyzes the SAM-dependent oxidative decarboxylation and carbon-carbon bond formation on MftA to form MftA*. The roles of the auxiliary [Fe-S] clusters in

MftC catalysis as well as the subsequent steps in the biosynthesis of mycofactocin are not known. Here, we have provided additional information regarding the roles of the auxiliary clusters in MftC. We showed that MftC contains three [4Fe-4S] clusters, all of which are required for catalysis. In addition, we measured the midpoint potentials of the clusters to provide insights into the redox flipping mechanism of MftC. Furthermore, we reconstituted the activity of MftE and showed that it selectively hydrolyzes MftA* to form MftA (1-28) and a 3-amino-5-(4-hydroxybenzyl)-4,4-dimethylpyrrolidin-2-one, herein referred to as AHDP. From this study, we have clarified the misunderstandings surrounding the accurate precursor for mycofactocin biosynthesis. Subsequently, we reconstituted the activity of MftD and showed that it catalyzes the oxidative deamination of AHDP to form an α-keto moiety herein referred to as premycofactocin. Lastly, we measured the midpoint potential of premycofactocin to be ~ -255 mV and demonstrated that it is used by mycofactocin-associated short chain dehydrogenases for multiple catalytic turnover.

ACKNOWLEDGEMENTS

I would like to thank Dr. John Latham for giving me the opportunity to join his lab and for his continued support over the years. The guidance and expertise he offered me over the years have been invaluable along with his unparalleled example of scientific rigor, mentorship and work ethic. I would also like to thank my committee members Dr. Martin Margittai, Dr. Michelle Knowles and Dr. Brian Michel for their continued advice and constructive criticism to improve the quality of my research projects. I am also grateful to Drs. Gareth and Sandra Eaton, Dr. Maria Pandelia (Brandeis University) and Dr. Sean Elliots (Boston University) for their collaborative efforts regarding EPR, Mossbauer and Electrochemical experiments respectively. Financial support for the work discussed in this dissertation was through the University of Denver and the NIH RO1 grant awarded to Dr. John Latham.

I would like to thank my labmates, both past and present, especially Bulat Khaliullin and Hallie Stolte for invaluable support when I first joined the Lab. I am also grateful for the amazing friends I have made in the Chemistry and Biochemistry program, especially Dr. Hilary Weismiller and Mary Shonk.

Finally, I would like to thank my wife who has been an incredible support system for me. Her endless encouragement and invaluable advice have made the completion of this program possible. Special thanks to my parents for their continuous encouragement and instilling in me a set of values that made the attainment of this goal possible.

TABLE OF CONTENTS

LIST OF FIGURES

ABBREVIATIONS

5’-dA•	5’-deoxyadensoyl radical
anSME	anaerobic sulfatase maturating enzyme
ATP	adenosine 5’-triphosphate
Aux	auxiliary cluster
Cα	alpha-carbon atoms
Cβ	beta-carbon atoms
Cγ	gamma-carbon atoms
CAPS	N-Cyclohexyl-3-aminopropanesulfonic acid
CHAPS	3-((3-cholamidopropyl) dimethylammonium)-1-propanesulfonate
CHES	N-Cycloheyl-2-aminoethanesulfonic acid
CV	cyclic voltammetry
DCM	dichloromethane
DMF	dimethylformamide
DTH	sodium dithionite
DTT	dithiothreitol
EPG	edge-plane graphite
GST	glutathione transferase
HCl	hydrochloric acid
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HPLC	high pressure liquid chromatography
IPTG	isopropyl β-D-1-thiogalactopyranoside

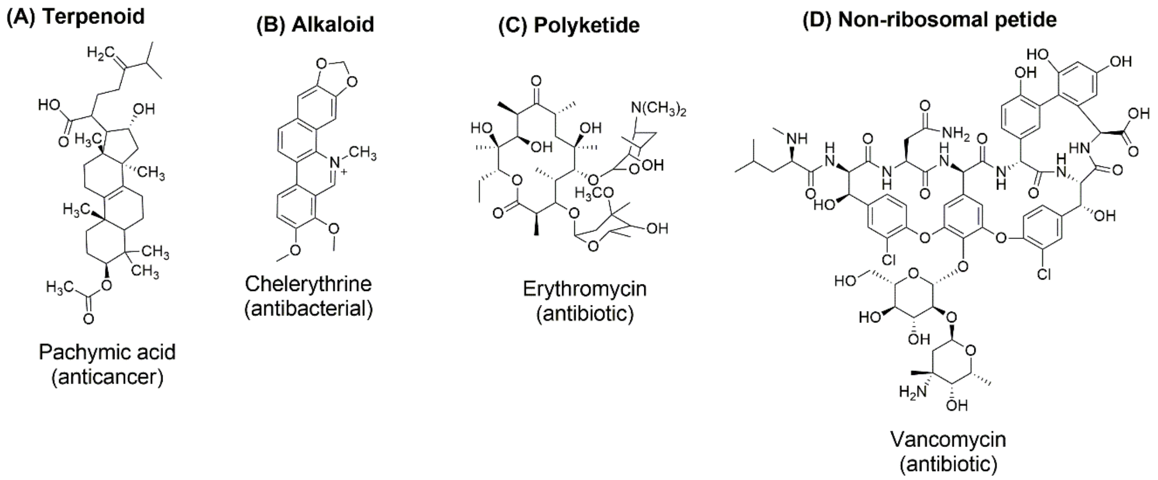
ICP-MS	Inductively coupled plasma mass spectrometry
LC-MS	liquid chromatography–mass spectrometry
MES	2-(N-morpholino) ethanesulfonic acid
NTA	Nitrilotriacetic acid
PMFT	Premycofactocin
$PMFTH_2$	Reduced premycofactocin
MOPS	3-(N-morpholino) propanesulfonic acid
m/z	mass-to-charge ratio (Mass spectrometry)
OD	Optical density
PBS	phosphate buffered saline
ppm	parts per million
PTM	post-translational modification
RiPPs	ribosomally synthesized and post-translationally modified peptide
RPM	revolutions per minute
RRE	RiPPs recognition element
RS	radical SAM
Sactionine	cysteine-sulfur to alpha-carbon thioether
sactipeptides	sulfur-to-alpha carbon thioether cross-linked peptides
SAM	*S*-adenosyl-L-methionine
SCIFF	six cysteines in forty-five residues
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SEC	size exclusion chromatography

Skf	sporulation killing factor
SPASM	subtilosin A/PQQ/anaerobic sulfatase and mycofactocin
SSN	sequence similarity network
SWCNTs	single-walled carbon nanotubes
MWCNTs	multi-walled carbon nanotubes
TAPS	(tris(hydroxymethyl)methylamino)propanesulfonic acid
TEV	tobacco etch virus
TFA	trifluoroacetic acid
UV	ultra-violet
wHTH	winged helix-turn-helix
WT	wild-t

CHAPTER ONE: INTRODUCTION

1.1 Natural Products and their Significance

Naturally occurring small molecules, herein referred to as natural products (NPs), make up a significant portion of the current pharmacopoeia especially regarding the treatment of infectious diseases, cancer, immunological disorders and hypercholesterolemia (1). They possess intricate molecular frameworks and structural diversities which offer them a range of potent biological activities. Owing to this, NPs have been recognized as the richest source of novel compounds with complex molecular scaffolds and potential therapeutic applications (2–4). In addition, NPs continue to play vital roles in augmenting our understanding of biology and chemistry (5–7). Unlike the synthetic compounds used in medicine, NPs are distinguished by their structural and functional diversity, thus expanding their biologically relevant chemical space (8–11). As a result, natural products remain an invaluable source of inspiration for drug design. Research in the 20th century identified several classes of natural products with four groups being the most prevalent. These groups include the terpenoids, alkaloids, polyketides, and non-ribosomal peptides.



Pachymic acid (anticancer)

Chelerythrine (antibacterial)

Erythromycin (antibiotic)

Vancomycin (antibiotic)

Figure 1.1 Representative members of the natural product classes highlighting their structural diversity (A) Pachymic acid, a terpenoid with anti-cancer properties (B) Chelerythrine, an alkaloid with antibacterial properties (C) Erythromycin, a polyketide used as antibiotics and (D) Vancomycin, a non-ribosomal peptide also used as antibiotics

Terpenoids, also referred to as isoprenoids constitute the largest class of natural products with diverse structures and potent biological activities (12–14). The majority of terpenoids are of plant origin and are composed of multicyclic structures with oxygen-containing functional groups (12, 15, 16). Terpenoids are involved in a plethora of biologically relevant activities such as cell membrane formation, intracellular electron transport (17), regulation of cell growth and defense (18), and anti-cancer drugs (19) (Figure 1.1A). They are biosynthesized from two universal five-carbon building blocks: isopentyl diphosphate (IPP) and its isomer dimethylallyl diphosphate (DMAPP) (14, 18, 20). Members of this class of natural products include monoterpenoids, diterpenoids, triterpenoids and tetraterpenoids (13, 15, 19, 21).

Alkaloids are another large class of natural products containing basic nitrogen atoms (22) (Figure 1.1B). They are produced by a large variety of organisms including bacteria, fungi, plants and animals. They are biosynthesized from complex pathways involving unique building blocks and a number of biosynthetic approaches (23). Members of this class are classified based on their biosynthetic precursor and heterocyclic ring systems and they include indoles, piperidines, tropanes, purines, pyrrolizidines, imidazole, quinolozidine, isoquinoline and pyrrolidine (24–26). Due to the extreme toxicity of alkaloids, they are mostly involved in the survival of the producing organism by exerting lethal effects on others. For instance, alkaloid-producing plants use alkaloids as a defense mechanism to ensure survival against microorganisms and herbivores (24). In spite of the high toxicity of alkaloids, this class of NPs has marked therapeutic effects in humans when used in small quantities. As a result, alkaloids are only used in medicine for their analgesic, antispasmodic and bactericidal effects as well as chemical transmitters in the nervous system (24, 25, 27, 28).

Polyketides (PK) belong to a structurally diverse class of natural products with broad range of pharmacological activities. Their structures are composed of highly oxygenated ring systems biosynthesized by multi-enzyme complexes referred to as polyketide synthases (PKS). Polyketides display a broad range of potent antibacterial activities. For example, erythromycin, a macrolide antibiotic, is used to treat a number of respiratory tract and skin infections as well as sexually transmitted infections such as chlamydia and syphilis (29–33) (Figure 1.1C). The standard biosynthetic model of polyketides involves the assembly of activated simple carboxylic acid precursor peptides such as malonyl-CoA on PKS through a stepwise condensation reaction (30, 31, 33, 34). Despite having three universal catalytic domains (ketosynthase, acyl transferase and acyl carrier protein) PKS are classified into two groups based on their similarity to fatty acid synthases : type I and the type II PKS, (35). Type I PKS are multifunctional enzymes that are organized into modules. Each module possess a discrete set of noniteratively acting activities and catalyze only a single cycle of chain elongation during the biosynthesis of polyketides (29–32). The type II on the other hand carry only a single set of iteratively acting activities and can catalyze several cycles of the chain elongation of polyketides (31, 32, 36, 37).

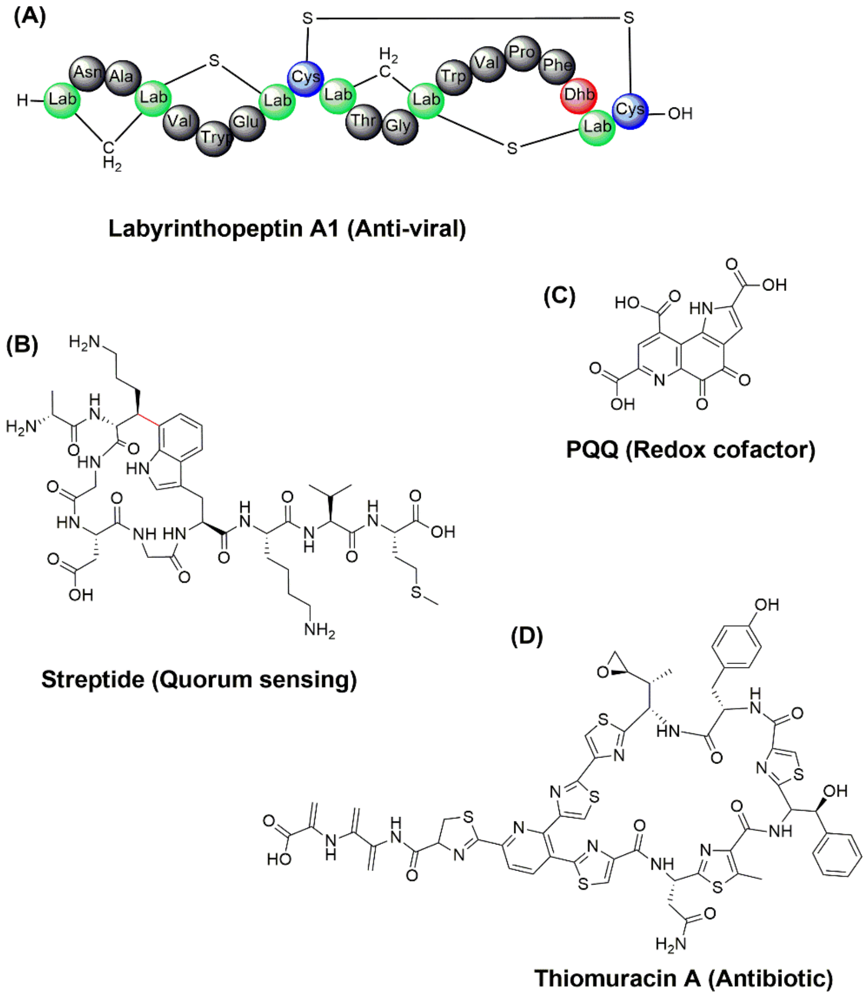
Non-ribosomal peptides are another diverse class of secondary metabolites with peculiar structures and pharmacological activities. This class of natural products is of profound interest because of its bioengineerability to generate new drugs with potent biological activities (38–40). The biological properties of non-ribosomal peptides can range from antibacterial e.g. vancomycin (41), immunosuppressive e.g. cyclosporin (42), cytostatic e.g. epothilones A and B (41), toxins e.g. thaxtomin (43) to iron chelators e.g. vibriobactin (44) (Figure 1.1D). The biosynthesis of non-ribosomal peptides is carried out by large modular multifunctional enzyme complexes with astonishing molecular weights called non-ribosomal peptide synthetases (42). Within each enzyme complex, there are several modules and each one is responsible for the incorporation of a single building block into the final structure of the polypeptide (38–40, 42, 45).

Although historically only four classes of NPs have been recognized, genome sequencing efforts by research scientists in the beginning of the 21st century have led to the discovery of a new class of natural products. This class of natural products is peptide based and is of ribosomal origin and called Ribosomally synthesized and Post translationally modified peptides (RiPPs).

1.2 Ribosomally Synthesized and Post-Translationally Modified Peptides (RiPPs)

RiPPs are a growing class of natural products that have garnered significant attention over the past decade due to their unique biosynthetic pathways and tailoring enzymes (46–50). Unlike polyketide synthases (PKS) and non-ribosomal peptide synthases (NRPS) which utilize large multi-modular enzyme complexes to incorporate unnatural amino acids into the backbone of their natural products (29, 34, 36, 38–40, 51), RiPPs are derived from genetically encoded precursor peptides. These precursor peptides undergo extensive chemical modifications by dedicated post translational modification (PTM) enzymes to achieve similar degrees of chemical diversities (47, 49, 50, 52) (Figure 1.2). RiPPs are of ribosomal origin and are produced in all three domains of life.

A single property of RiPPs that sets them apart from polyketides and non-ribosomal peptides is their remarkable biosynthetic malleability. This feature enables RiPPs pathways to be genetically engineered to produce new-to-nature compounds with higher potency and bioavailability (50), thus expanding the biologically relevant chemical space of RiPPs. Owing to this, RiPPs are fertile grounds for exploring molecular scaffolds and chemistries that are useful for chemotherapeutic applications (52, 53). The biological activities possessed by RiPPs are vast, ranging from anti-viral (54), anti-fungal (55), antibiotic (56), anti-cancer (57), quorum sensing molecules (58, 59), and redox cofactors (60–63) (Figure 1.2), to name a few.



Labyrinthopeptin A1 (Anti-viral)

Streptide (Quorum sensing)

PQQ (Redox cofactor)

Thiomuracin A (Antibiotic)

Figure 1.2 Classical examples of RiPP natural products highlighting their structural diversities and biological activities. (A) The Class III lanthipeptide Labryrinthopeptin A1 with anti-viral properties (B) Streptide involved in bacterial cell communication (C) The redox cofactor Pyrroloquinoline quinone (PPQ) (D) The thiopeptide Thiomuracin A with potent antibacterial properties**.**

The biosynthesis of RiPPs typically commences with the ribosomal production of the genetically encoded precursor peptide (Figure 1.3 middle row). This precursor peptide can house up to three domains: the leader sequence, the core sequence and the follower sequence (47). The leader sequence primarily serves as a recognition motif for the cognate biosynthetic machinery while the core sequence undergoes extensive chemical transformations and is incorporated into the final RiPP product (64). The follower sequence is not strictly conserved and if the precursor peptide has any, it often plays the role of the leader peptide (Figure 1.3 middle row). Upon translation and substrate engagement, the core sequence is extensively modified by post-translational modification (PTM) enzymes and the leader peptide is subsequently removed by proteolysis to allow for the maturation of the modified core peptide into the final natural product (Figure 1.3 bottom row). The PTMs impacted on these peptides endow them with non-proteinogenic structures that are not directly accessible by ribosomal or modular synthesis. The functionalized RiPP product becomes more conformationally restrained allowing for better target recognition, increased metabolic and chemical stability as well as increased chemical functionality (47). RiPPs are grouped into different subfamilies based on their biosynthetic machinery and structural characteristics and these include sactipeptides, lanthipeptides, thiopeptides, cyanobactins, methanobactins, bottromycins, streptides among others (47, 49). Sactipeptides can boast of potent antibacterial properties and their biosynthesis involve unique chemistries. Streptides are quorum sensing molecules involved in bacterial cell communication and possess peculiar biosynthetic machineries. For the purpose of this dissertation, only the Sactipeptide and streptide subfamily of RiPPs will be discussed.

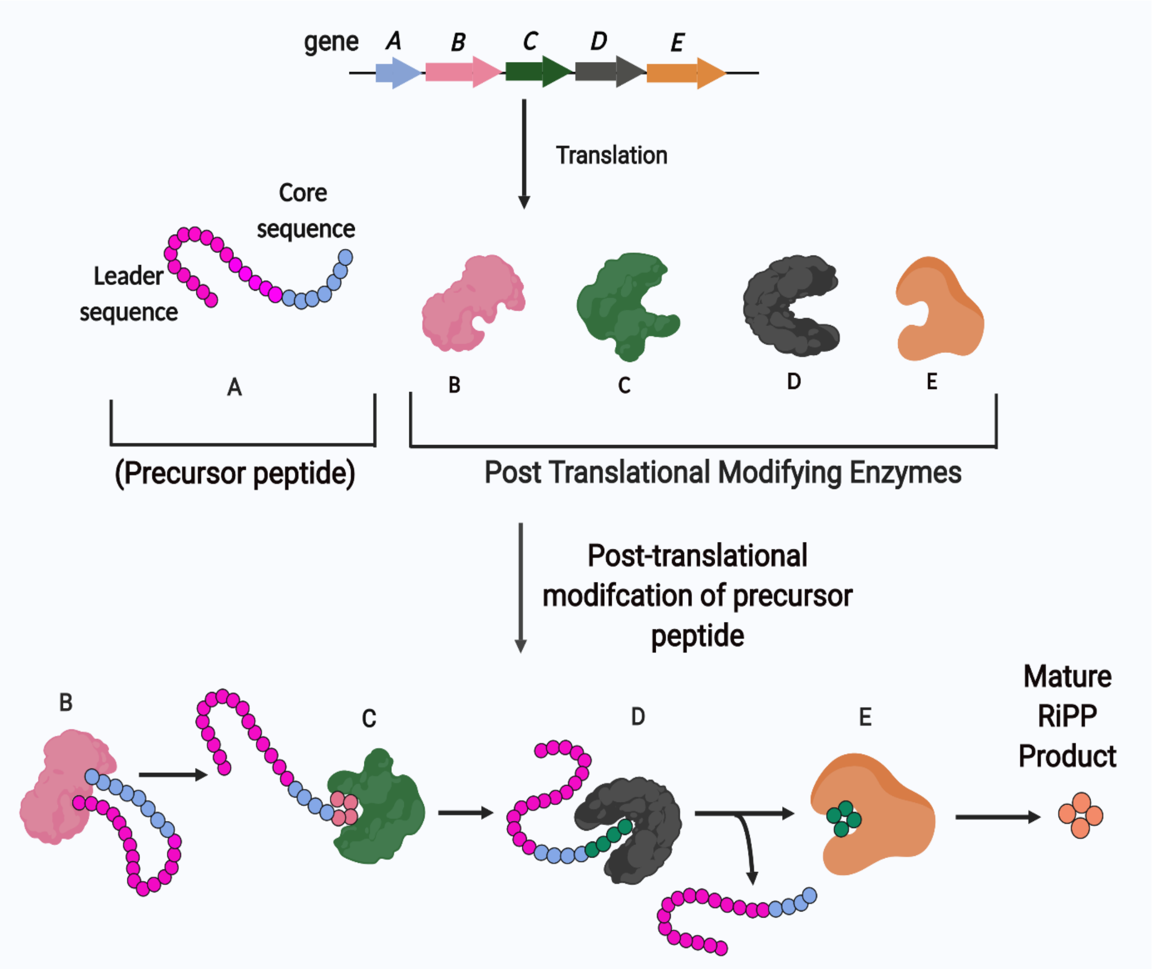


Figure 1.3 Graphical representation of a typical RiPP biosynthesis. (Top row) Typical gene organization of a RiPP biosynthetic pathway. (Middle row) Translation of the precursor peptide together with the RiPP biosynthetic machinery. (Bottom row) Extensive chemical modifications are made to the core sequence of the precursor peptide to form the mature RiPP product.

1.2.1 Sactipeptides

Sactipeptides are a diverse family of RiPPs natural products that possess potent antibacterial, spermicidal and hemolytic properties (65–71). They are characterized by a distinct intramolecular thioether bond that crosslinks the **s**ulfur atom of a cysteine residue to the **al**pha **c**arbon of an acceptor amino acid, hence the name sactipeptides. The thioether or sactionine linkages found in sactipeptides differ from those found in the well-characterized lanthipeptides, such as nisin. For instance, in lanthipeptide biosynthesis, the thioether crosslinks are formed between a sulfhydryl group of a cysteine residue and dehydrated side chains of threonine or serine residue on the precursor peptide through a Michael-like nucleophilic addition. The maturation of sactipeptides on the other hand involve a distinct group of enzymes called synthetases which install a thioether crosslink between a cysteine residue and an acceptor amino acid on their precursor peptide. Another difference between lanthipeptide and sactipeptide thioether linkages is that the thioether crosslinks in sactipeptides occur through a radical-mediated pathway involving a net two-electron oxidation of the precursor peptide whereas those found in lanthipeptides occur through a two-step redox neutral reaction mechanisms (72–74). Several subfamilies of sactipeptides have been characterized till date and they include Subtilosin A from *Bacillus subtilis* (Figure 1.4A), thurincin H from *Bacillus thuringiensis* (Figure 1.4B), sporulation killing factor from *Bacillus subtilis* (Figure 1.4C), and the two-component thuricin CD from *Bacillus thuringiensis* (Figure 1.4D). The sactisynthases AlbA, ThnB, SkfB and Trn-α/β are responsible for the installation of S-Cα thioether bonds during the maturation of subtilosin A, thurincin H and sporulation killing factor and thuricin CD respectively (66, 72, 75–79). Recently, another subfamily, Huazacin from *Bacillus thuringiensis* serova*r huazhongensis* was identified and shown to also contain a sulfur to alpha carbon thioether crosslink installed by the sactisynthases HuaB and

HuaC (80, 81) (**Error! Reference source not found.** E). Elegant bioinformatic work by Haft and Basu in 2011 identified another subfamily of the sactipeptides. This subfamily contains a variable N-terminal domain and a conserved C-terminal domain and it is characterized by the presence of a **s**ix **c**ysteines **i**n a **f**orty-**f**ive residue (SCIFF) motif (82). Notable members include freyrasin from *Paenibacillus polymyxa* (Figure 1.4F), thermocellum (Cte) from *Clostridium thermocellum* (Figure 1.4G) and the SCIFF peptide from Tte1186a from *Caldanaerobacter subterraneus* subsp. *tengcongensis* MB4 (Figure 1.4H) (74, 81, 83). Tte1186 was the first sactisynthase from this subfamily to be characterized. It installs a S-Cα thioether crosslink on the precursor peptide Tte1186a.

A recent hallmark publication by the Mitchell lab demonstrated that not all members of the SCIFF subfamily possess S-Cα thioether crosslinks as previously proposed. For example, the radical SAM enzyme, CteB from *Clostridium thermocellum* was found to install an S-Cγ thioether crosslink on its precursor peptide, CteA during the maturation of thermocellum. In addition, the radical SAM enzyme PapB from *Paenibacillus polymyxa* was also found to install S-Cβ thioether crosslinks on its precursor peptide papA during the maturation of freyrasin (84). This interesting finding led the Mitchell lab to the proposal of a more appropriate nomenclature, a **ra**dical **n**on-α **thi**oether peptides (ranthipeptides) for this subfamily to structurally and biosynthetically distinguish them from the canonical sactipeptides (81).

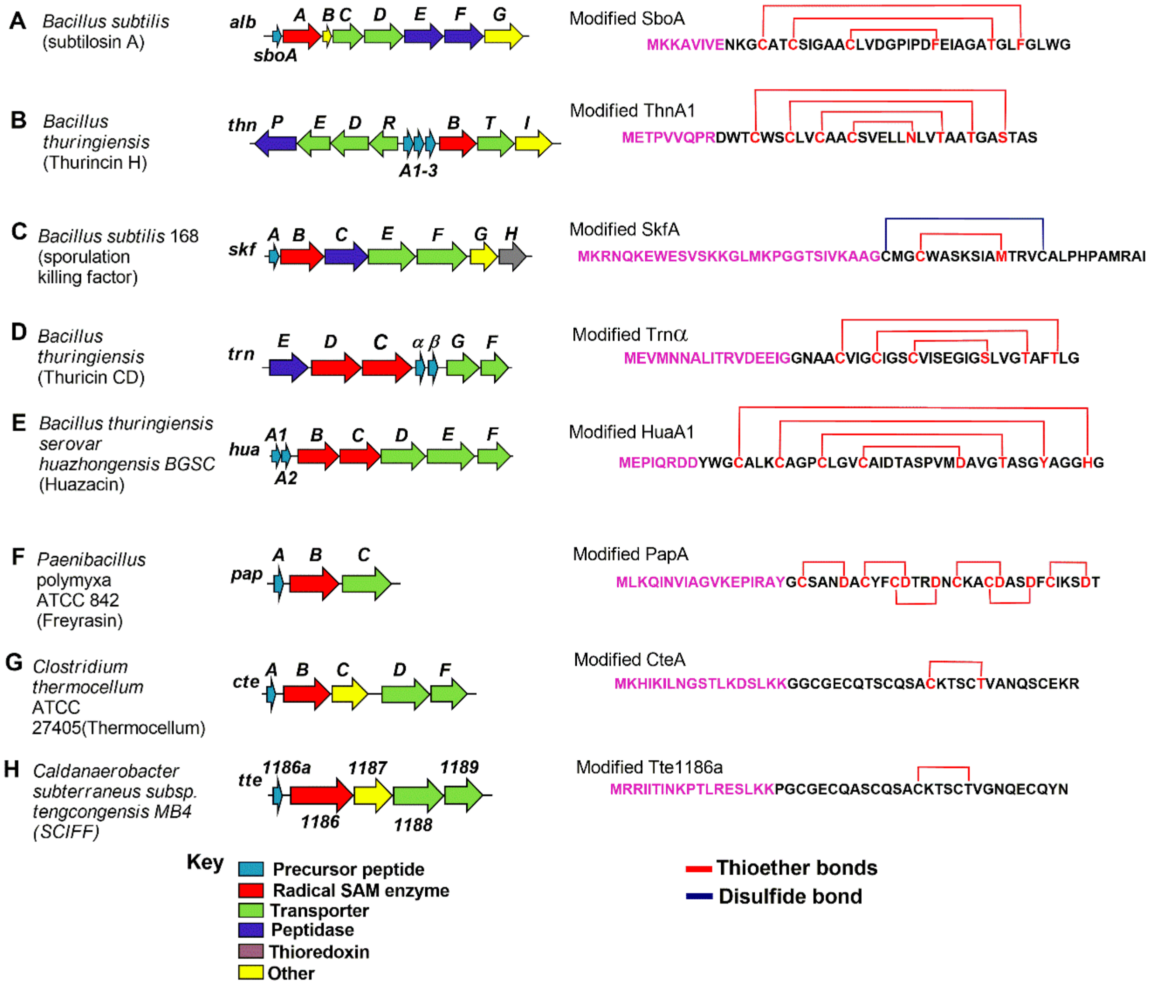


Figure 1.4 Gene scheme of several RiPP biosynthetic gene clusters involved in the biosynthesis of ranthipeptides and sactipeptides. (Left column) Different species containing biosynthetic genes encoding the precursor peptides (cyan), radical SAM enzymes (red) responsible for installing S-Cα, S-Cβ and S-Cγ thioether bonds, peptidases (blue) responsible for cleaving the leader peptide and transporters (green). (Right column) Sactisynthase and ranthisynthase modified precursor peptide from each biosynthetic cluster highlighting the leader peptide sequence (pink) and the mature RiPP product (blank). The thioether linkages formed between the sulfur atom of a cysteine residue and the Cα, Cβ, or Cγ of the acceptor residue are shown in red. Disulfide bonds are shown in blue.

The biosynthesis of sactipeptides and ranthipeptides begins with the leader peptide-dependent post translational modification of the precursor peptide core by sactisynthases. These sactisynthases house an N-terminal [4Fe-4S] radical SAM cluster and a C-terminal extension that can house up to two additional [4Fe-4S] clusters (72, 83, 85). During sactionine linkage formation, the radical SAM cluster is used to homolytically cleave SAM to generate a highly reactive 5’-deoxyadenosine radical. This radical abstracts a hydrogen atom from the substrate to form a substrate radical and 5’-deoxyadenosine. The substrate radical further undergoes an oxidation to form a more stable ketoimine intermediate (74) (Figure 1.5). Subsequent stereoselective intramolecular nucleophilic attack of a cysteine thiol on the unsaturated ketoimine via conjugate addition mechanism yields the corresponding crosslinked thioether bonds in the precursor peptide.

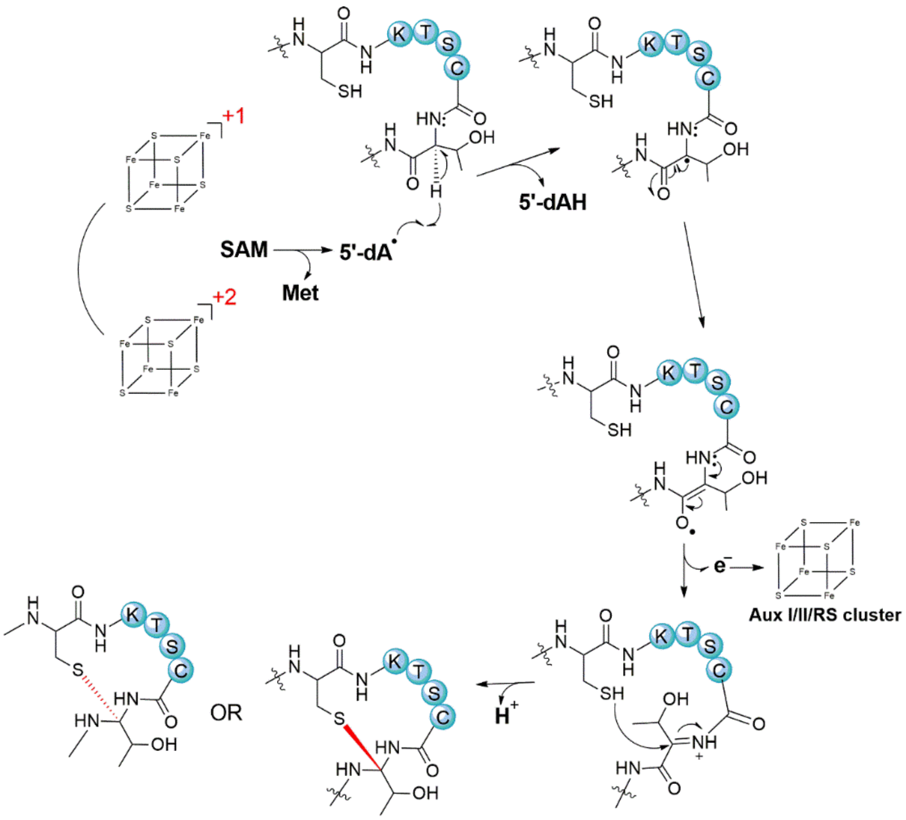


Figure 1.5 General reaction mechanism proposed for the biosynthesis of sactipeptides and ranthipeptides. This mechanism involves the abstraction of a hydrogen atom from the αβγ position of the precursor peptide. This is followed by the loss of an electron to form a ketoimine intermediate. Subsequent nucleophilic attack by a vicinal thiol on the ketoimine results in the formation of the matured RiPP product.

1.2.2 Streptides

Another family of RiPPs that has attracted significant attention due to their interesting biological activities is the Streptide family. Streptide is a 9-mer peptide that was first identified and characterized in 2007 by Ibrahim *et al.* (58). Streptides are pheromones produced by a number of pathogenic bacteria including *Streptococcus thermophilus,* and contain lysine to tryptophan crosslink in their peptide backbone to form an N-terminal macrocyclic structure (58, 59, 86–89).

They function as quorum sensing molecules in processes such as cell communication between a number of *Streptococci* (58, 59). The Streptide biosynthetic machinery is composed of the genes *strABC* which encode the precursor peptide StrA, the radical S-adenosylmethionine (SAM) enzyme StrB and the protease/transporter StrC respectively (88, 89) (Figure 1.6A).

The biosynthesis of Streptide commences with the ribosomal production of the genetically encoded precursor peptide StrA containing the conserved motif KGDGW (90). Following the reductive cleavage of SAM by StrB and the subsequent formation of the 5'-deoxyadenosine radical, a hydrogen atom is abstracted from the Cβ of Lys 16 on the precursor peptide StrA to form a lysyl radical and 5'-deoxyadenosine (Figure 1.6B). In the presence of a base, the lysyl radical then attacks the 3-position of the indole ring of Trp 20 to form an intramolecular carbon-carbon crosslink followed by re-aromatization of the ring and an electron transfer into one of auxiliary clusters. The modified core is then hydrolyzed from the rest of the precursor by the protease/transporter StrC to liberate the matured biological active product, Streptide (Figure 1.6B). The successful maturation of RiPPs natural products is contingent on the RiPP biosynthetic machinery, working independently in a sequential manner. One of the most critical components of this biosynthetic machinery is the class of radical SAM enzymes.

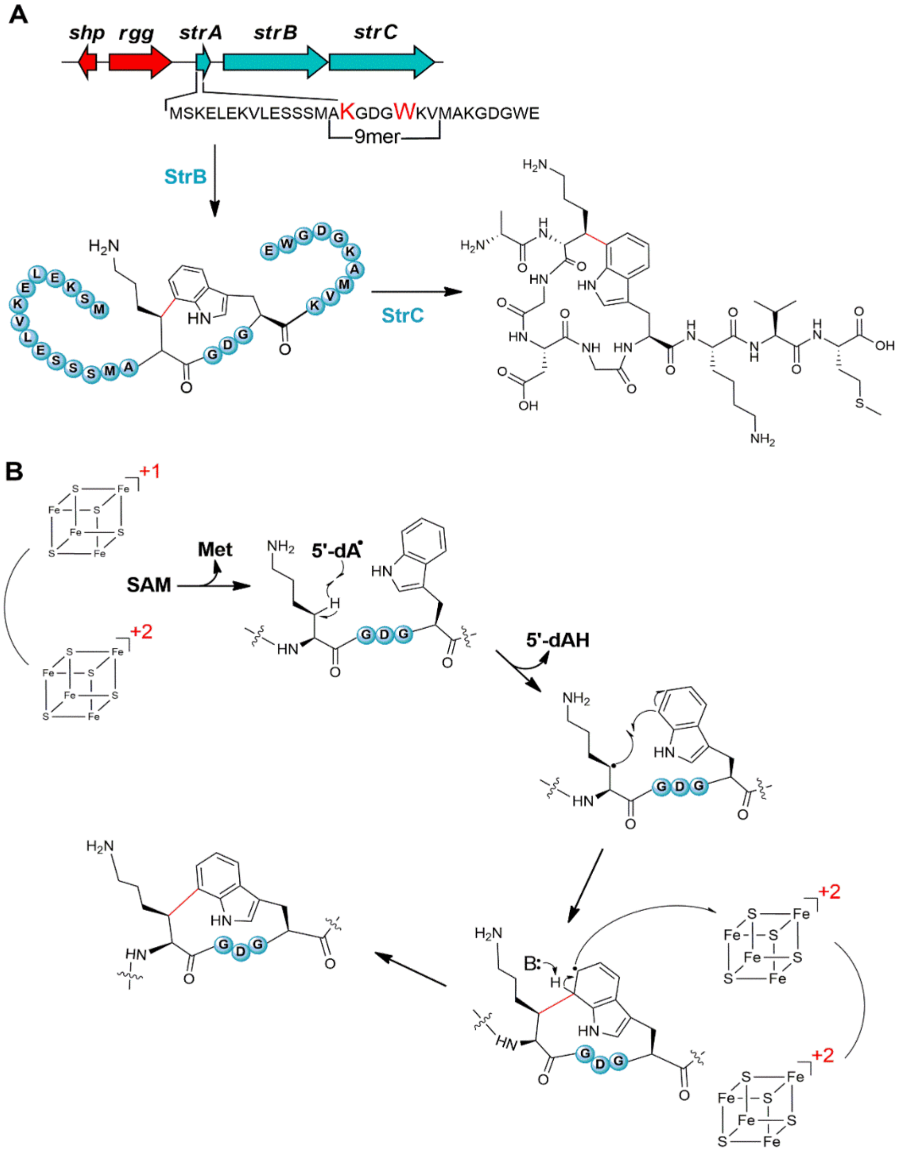


Figure 1.6 Streptide biosynthesis. (A) Streptide biosynthetic gene cluster activated by *shp/rgg* quorum sensing system and a condensed representation of the biosynthesis of Streptide. (B) Detailed reaction mechanism for Lys-to-Trp crosslink on the precursor peptide StrA, catalyzed by StrB during Streptide biosynthesis.

1.3 Structural Architecture of radical S-adenosylmethionine (RS) enzymes

Of the post-translational modification enzymes involved in the biosynthesis of RiPP natural products, one class has garnered enormous attention over the years. This is the so called class of radical SAM (RS) enzymes which are well known for their ability to anaerobically catalyze chemically difficult transformations (91, 92). These enzymes utilize a [4Fe−4S] cluster and S-adenosylmethionine to initiate an array of radical reactions via generation of a highly potent 5′-deoxyadenosyl radical intermediate. Structurally, RS enzymes can house up to three domains, all of which are required for catalysis. The first domain is an N-terminal extension that is used for binding substrate or cofactors. Following the N-terminal extension is the radical SAM domain, catalytic core that houses an active site [4Fe-4S] cluster. The third domain is a functionalized C-terminal extension for housing additional cofactors.

1.3.1 N-terminal Extensions of RS enzymes.

The N-terminal extensions of radical SAM enzymes adopt a typical winged-helix-turn (wHTH) motif (Figure 1.7). This was first observed in the crystal structures of the RiPP biosynthetic enzymes LynD and NisB where the leader sequences of their precursor peptides were found bound by a (wHTH) motif (93, 94). Subsequent bioinformatic analysis by the Mitchell group in 2015 also identified another conserved motif linking the homologous wHTH motifs in NisB and LynD

(95). This motif was found to be conserved among a number of prokaryotic RiPP classes and they are either fused to the RS enzyme or exist separately of the RS enzyme (93–96). Sequence homology alignment of wHTH motif in LynD and NisB revealed its structural homology to the protein PqqD (95) (Figure 1.7). PqqD is a small peptide chaperone involved in the biosynthesis of the peptide-derived redox cofactor pyrroloquinoline quinone (PQQ) (96). This PqqD-like domain was also previously identified in AlbA from subtilosin biosynthesis, however its exact role in RiPP biosynthesis remained largely enigmatic (47). In the biosynthesis of PQQ, the radical SAM enzyme PqqE was found to only bind to and modify the precursor peptide PqqA in the presence of the peptide chaperone PqqD (96). An alternate scenario was also observed in the biosynthesis of mycofactocin where the radical SAM enzyme MftC only bound to, and modified the precursor peptide MftA in the presence of the peptide chaperone MftB (97–99). The crystal structure of MftB is yet to be determined, however, sequence homology alignment revealed it possesses high similarity to PqqD. Additional bioinformatic analysis also demonstrated that these motifs were routinely found in RiPP biosynthetic proteins but absent in homologs not involved in processing of RiPPs. A more compelling evidence came from the finding that RiPP enzymes which act after the cleavage of the leader sequence from the precursor peptide lacked this motif (95). The aggregated bioinformatic data led the Mitchell group to propose that the PqqD-like domains associated with peptide modifying enzymes are used to recognize and bind the leader sequence of RiPP precursor peptides. This domain was therefore named **R**iPP precursor peptide **r**ecognition **e**lement (RRE) to highlight its role in the biosynthesis of RiPPs (95). This motif is composed of three antiparallel consecutive β-strands adjoining three consecutive α-helices (Figure 1.7)

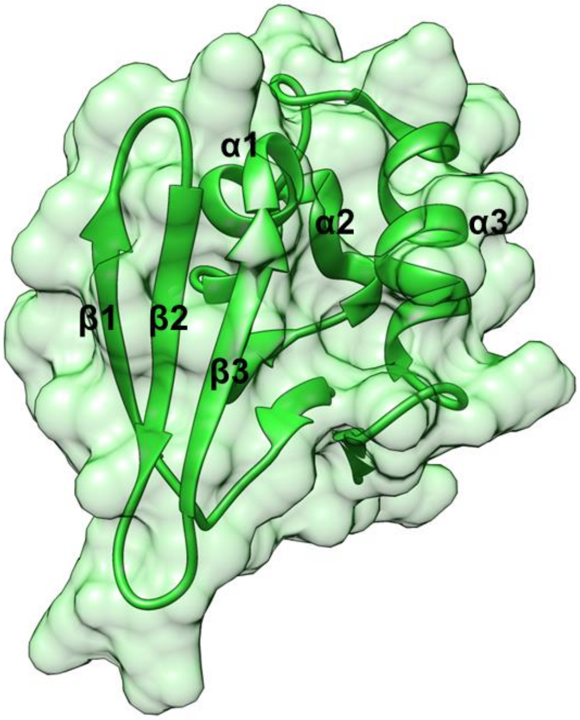


Figure 1.7 Crystal structure of the RRE domain of LynD (PDBID 4V1T). The three antiparallel β-strands and three α-helices adopt a winged-helix-turn-helix motif and constitute the conserved RiPP precursor peptide recognition element (RRE).

1.3.2 Radical SAM domain of RS enzymes

The radical SAM domain is responsible for binding SAM and the generation of the 5’-deoxyadenosine radical. This domain is comprised of a six (β/α) motifs arranged in a fashion that is reminiscent of triose-phosphate isomerase (TIM) barrel fold. The six α-helices make up the outside of the partial barrel and the six β-strands form the inner surface of the TIM barrel (100–102) (Figure 1.8). Although some radical SAM enzymes are composed of the complete $(\beta/\alpha)_8$ TIM barrel fold as found in HydE (103), BioB (104, 105) and ThiC (106), most of them have the partial TIM barrel fold. Whereas the partial TIM barrel fold is comprised of six repeating β/α motifs, the complete TIM barrel fold is composed of eight repeating β/α motifs. There are four distinct motifs within this domain that are responsible for coordinating SAM. The most conserved and important among them is the $CX_3CX\phi C$ (where ϕ is an aromatic residue) motif which binds the [4Fe-4S] cluster responsible for initiating radical chemistry (Figure 1.8). The three cysteines of this motif reside on the loop linking the β1 strand to the α1 strand, referred to as the cluster-binding loop and are responsible for ligating three of the irons from the [4Fe-4S] cluster. The fourth iron, termed the “unique iron” is ligated to the carboxylate and amino group of SAM in a bidentate fashion (Figure 1.8). Another motif, which is not strictly conserved is the Gly-Gly-Glu motif named after the residues in the radical SAM enzymes MoaA and pyruvate formate lyase activating enzyme. In these enzymes, the Gly-Gly-Glu motif resides at C-terminal end of the β2 strand in the radical SAM core and forms hydrogen bonds with the amino group of SAM molecule. This interaction allows the bidentate ligation of the carboxylate and amino groups of SAM to the unique iron of the [4Fe-4S] cluster (101), thus orienting the methionyl moiety of SAM. On the β4 strand of the radical SAM core is another motif termed “ribose motif (Figure 1.8). Residues on this motif form hydrogen bonds with both the 2’ and 3’ hydroxyl groups of the ribose moiety on SAM and orient the 5’-deoxyadeonsyl radical towards the substrate for hydrogen atom abstraction during catalysis (107). The last motif which is only partially conserved within this domain is the “GXIXGXXE motif” named after the corresponding sequence in BioB (104, 105), located on β5 strand. This motif provides hydrophobic interactions with the adenine ring of SAM (Figure 1.8). Taken together, all four motifs work in tandem to ensure adequate positioning of SAM for homolytic cleavage, radical generation and subsequent hydrogen abstraction from the substrate during radical SAM catalysis.

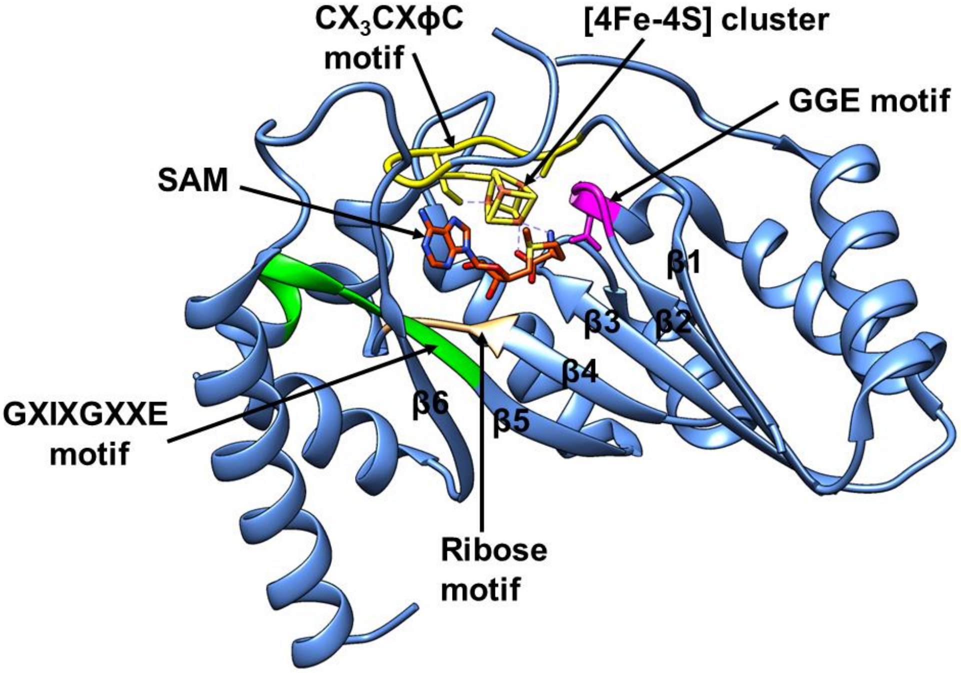


Figure 1.8 Radical SAM domain of anSME from Clostridium perfringens (PDB ID 4K37). The typical partial $(\beta/\alpha)_6$ TIM barrel fold of the radical SAM core. Shown are the CX$_3$CXϕC motif (yellow) for binding [4Fe-4S] cluster (yellow sticks), and the bidentate coordination of the unique iron to SAM (orange/blue sticks), the GGE motif (magenta), the ribose motif and the GXIIGXXE motif, all of which are necessary for orientation and cleavage of SAM during catalysis.

The canonical function of the radical SAM domain is the generation of a radical using a [4Fe-4S] cluster and SAM. This is a two-step process and requires SAM to be placed in the immediate vicinity of the [4Fe-4S] cluster to directly coordinate the cluster and allow for electron transfer. The first step in the radical generation is the reduction of the [4Fe-4S] cluster from the resting +2 oxidation state to the active +1 oxidation state. This is facilitated by the presence of an electron donor such as flavodoxin/flavodoxin reductase, a physiological reductant or sodium dithionite (DTH), a non-physiological reductant (108). The final step is the cleavage of SAM. This step involves the transfer of an electron from the reduced [4Fe-4S] cluster to the sulfonium antibonding orbital of SAM through an inner sphere electron transfer mechanism (109), resulting in the homolytic cleavage of SAM to form methionine and a highly potent 5'-deoxyadenosine radical (Figure 1.9). In a substrate coupled SAM cleavage, the generated 5'-deoxyadeonsine radical abstracts a hydrogen atom from the precursor peptide, thus priming the substrate for the formation of desired chemical transformations.

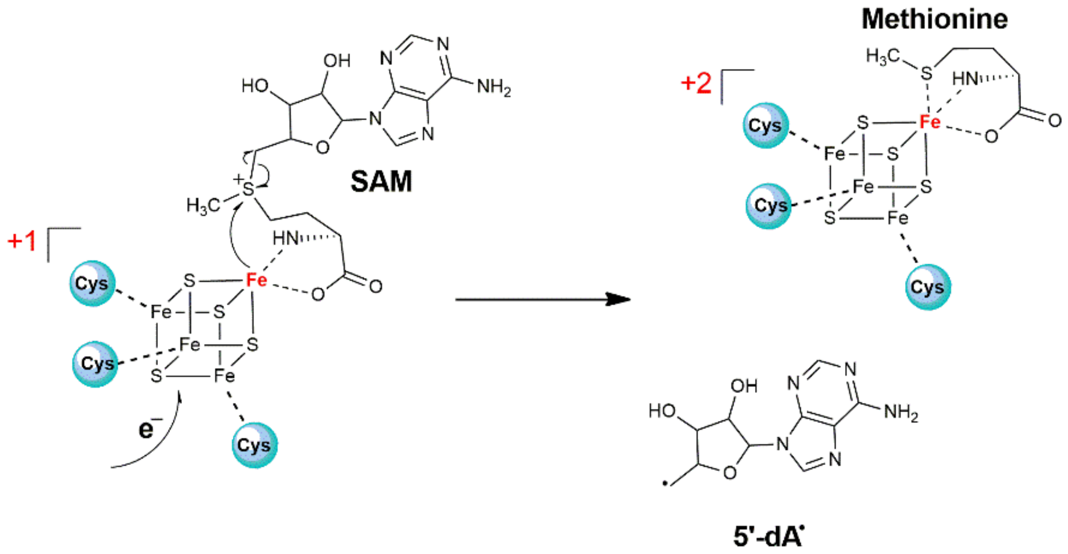


Figure 1.9 Proposed reaction mechanism of SAM cleavage. The reductive cleavage of SAM in RS enzymes showing an electron transfer from the reduced [4Fe-4S] cluster into the antibonding orbital of SAM to yield the 5'-dA radical and methionine.

1.3.3 C-terminal extensions of RS enzymes

On top of the RRE domain and the radical SAM domain, some RS enzymes have C-terminal extensions that are functionalized for binding substrates or other cofactors (83, 85, 110, 111). Radical SAM enzymes with functionalized C-terminal extensions are classified as either RS-TWITCH enzymes or RS-SPASM enzymes. The RS-TWITCH enzymes harbor one additional [Fe-S] cluster in their C-terminal extensions while the RS-SPASM enzymes house two additional [Fe-S] clusters in their C-terminal extensions (83, 101, 110, 112–115) (Figure 1.10). The C-terminal TWITCH domain is an abridged C-terminal SPASM domain which typically contains a conserved three/four cysteine motif for coordinating one additional [Fe-S] cluster. Examples of such protein fold can be found in BtrN (112) (Figure 1.10A) and MoaA (116) (Figure 1.10B) and recently characterized enzyme SkfB (85).

Unlike BtrN and MoaA, both of which contain a [4Fe-4S] cluster as the Aux I, SkfB contains a [2Fe-2S] cluster as the Aux I (101). The Aux I cluster in MoaA has an open coordination site for substrate binding (Figure 1.10B). The protein fold and functions of SPASM RS enzymes are described in detail below.

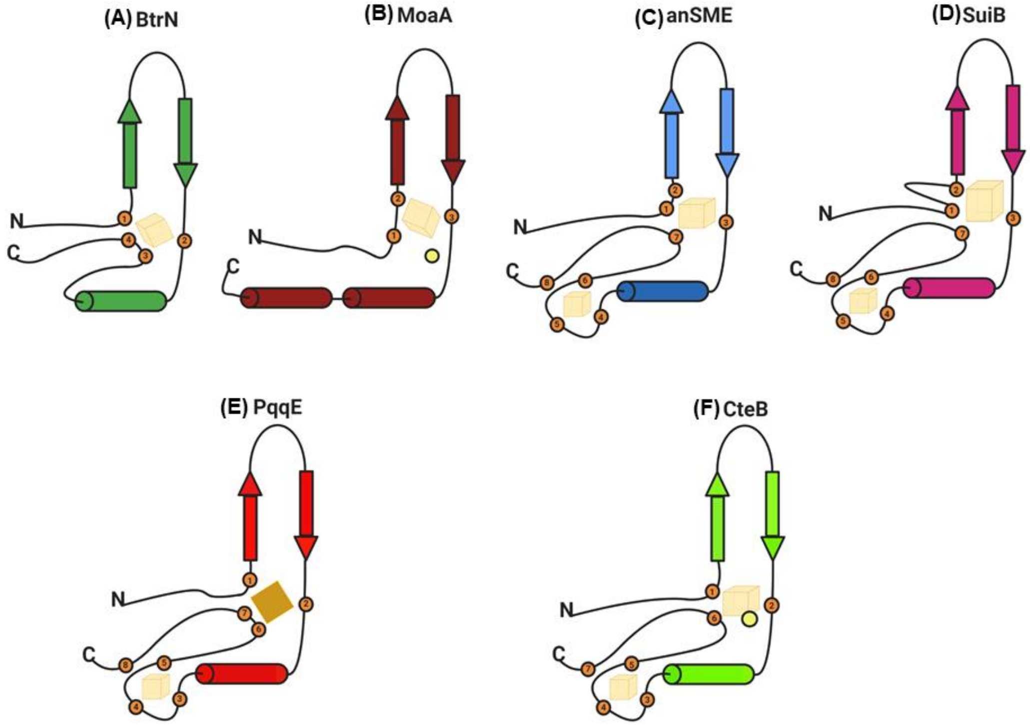


Figure 1.10 Comparison of TWITCH and SPASM domains of RS enzymes (A) The TWITCH domain of BtrN showing four cysteines for coordinating the Aux I cluster. (B) The TWITCH domain of MoaA showing three cysteines for the Aux I cluster with an open coordination site for substrate binding. (C) SPASM domain of anSME showing eight cysteines for coordinating Aux I and Aux II clusters. (D) SPASM domain of SuiB harboring two additional [4Fe-4S] clusters. (E) SPASM domain of PqqE showing the eight cysteine residues that are involved in coordinating the two auxiliary [Fe-S] clusters. (F) The functionalized C-terminal extension of CteB containing two auxiliary clusters with an open coordination site on the Aux I cluster for substrate binding.

1.4 RS-SPASM Enzymes

Bioinformatic studies by Haft and Basu in 2011 identified a domain in the C-terminal region of a subset of radical SAM enzyme class designated as TIGR08045 (82). This domain houses two additional cofactors suggesting its probable involvement in the post-translational modification of the corresponding precursor peptides. The cluster proximal to the RS-cluster is annotated as the Aux I cluster, and the distal cluster is annotated as the Aux II cluster (Figure 1.10 and Figure 1.11). The functionalized C-terminal extensions of these enzymes have a conserved seven cysteine $CX_{9-15}GXC$-gap-$CX_2CX_5CX_3C$-gap-C motif (83, 110, 114, 117, 118) (Figure 1.11). This motif is typically used for binding the two auxiliary [4Fe-4S] clusters, leaving a unique iron with an open coordination site for substrate binding (82, 83, 113) (Figure 1.10F). However, structural characterization of anSME, PqqE and SuiB showed full ligation of both Aux I and Aux II clusters by cysteine residues (110, 117) (Figure 1.10C, D and E). Unlike anSME and SuiB, both of which contains a [4Fe-4S] cluster in their Aux I site, PqqE contains a [2Fe-2S] cluster in its Aux I site (119, 120). The functions of these auxiliary clusters remain to be determined. However, they are proposed to play important catalytic roles such as binding and positioning of the substrate to the radical SAM cluster as found in CteB (83) and AlbA (76), or act as an electron shuttle for the controlled transfer of unpaired electrons out of the enzyme's active site during the oxidation of the substrate radical intermediate (113, 121). The functionalized C-terminal extensions of RS-SPASM proteins highlight both the diversity of reactions catalyzed by these enzymes as well as the range of functions of the iron-sulfur clusters.

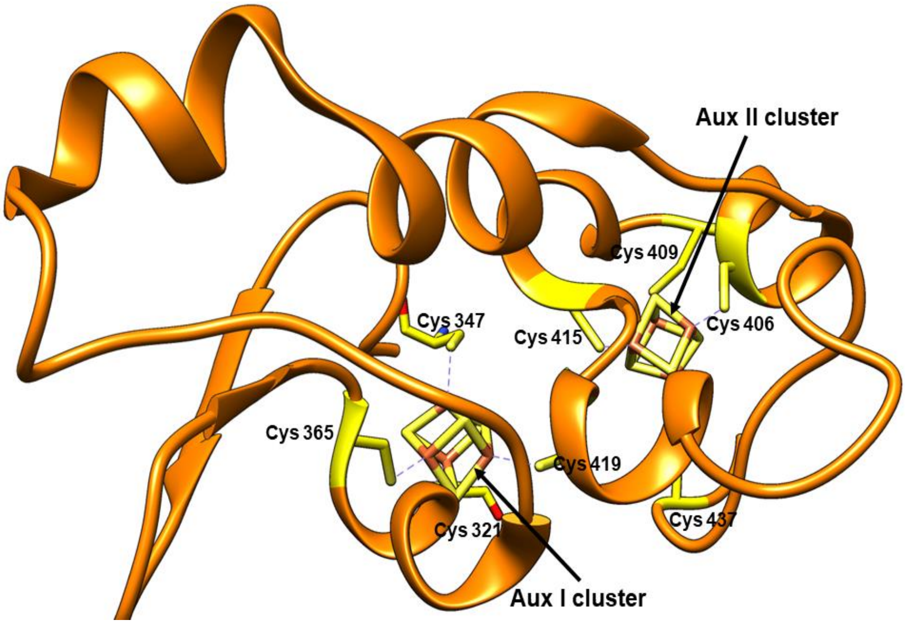


Figure 1.11 Architecture of the SPASM domain of SuiB from Streptococcus suis (PDB:5Q1V). Highlighted in yellow are the cysteine residues in the CX$_{9-15}$GXC-gap-CX$_2$CX$_5$CX$_3$C-gap-C motif used to coordinate two Auxiliary [4Fe-4S] clusters. The iron sulfur clusters are represented in sticks.

The name “SPASM” was coined for this subfamily of radical SAM enzymes as an abbreviation from the biochemically characterized founding members AlbA, PqqE, anSME and MftC which are radical SAM enzymes involved in the maturation of **S**ubtilosin A, **P**yrroloquinoline quinine, **A**nerobic sulfatase and **M**ycofactocin respectively. Anaerobic sulfatase maturation enzyme, AtsB is the first structurally characterized RS-SPASM protein and it has been shown to catalyze the two-step formal electron oxidation of a critical serine or cysteine residue into a C$_\alpha$-formylglycine, an *in situ* arylsulfatase cofactor required for the activity of AstA (117, 122, 123) (Figure 1.12A). The second member of this subfamily to be chemically characterized is AlbA, even though its crystal structure is yet to be solved. AlbA catalyzes the formation of sulfur to alpha carbon thioether bond (sactionine linkage) on its precursor peptide SboA during the maturation of the antibacterial peptide subtilosin A (72, 76) (Figure 1.12B). PqqE and MftC are both SPASM proteins involved in the biosynthesis of the peptide derived redox cofactors PQQ and Mycofactocin respectively. PqqE catalyzes the formation of a carbon-carbon bond between a glutamate and a tyrosine residue on its precursor peptide PqqA during the biosynthesis of PQQ (Figure 1.12C) while MftC catalyzes the oxidative decarboxylation of the C-terminal tyrosine residue followed by a carbon-carbon bond formation between the penultimate valine and the decarboxylated tyrosine during the biosynthesis of mycofactocin(Figure 1.12D) (97–99, 124).

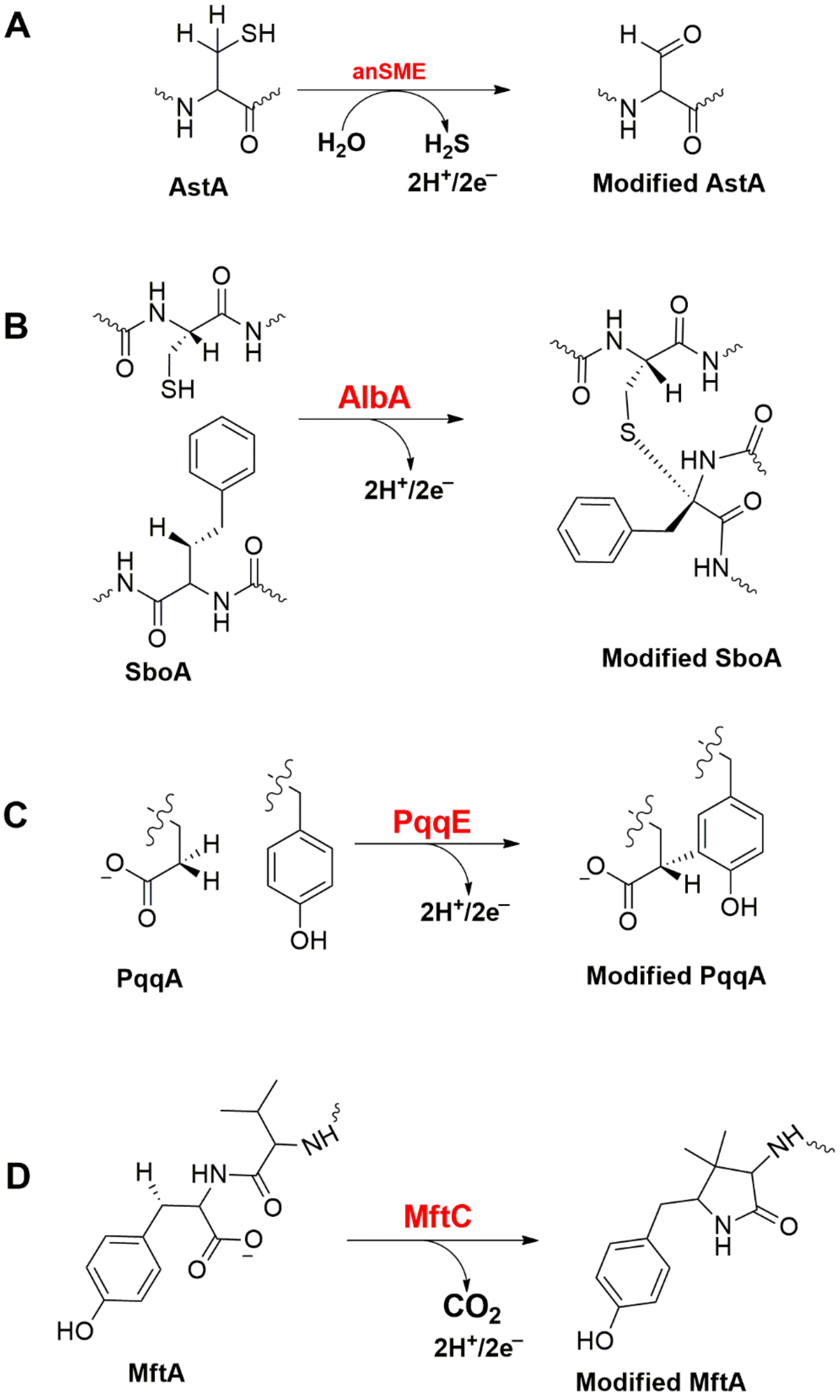


Figure 1.12 Reactions catalyzed by SPASM-domain radical SAM enzymes. (A) anSME (anaerobic Sulfatase-Maturating Enzyme: Cysteine or serine oxidation into Cα-formylglycine), (B) AlbA (Subtilosin A biosynthesis: S-Cα thioether bond formation), (C) PqqE (PQQ biosynthesis: C-C bond formation) (D) MftC (Mycofactocin biosynthesis: C-C bond formation).

More recently, the crystal structure of the first thioether bond forming SPASM radical SAM enzyme, CteB was determined (83). It was shown that its Aux I cluster is ligated by three cysteines leaving an open coordination site for substrate binding (Figure 1.10F and Figure 1.13A). CteB installs an sulfur to gamma carbon thioether bond on its precursor peptide CteA during the maturation of thermocellum (81, 83). Another RS-SPASM protein whose structure was recently solved is SuiB. The crystal structure of SuiB revealed the presence of two [4Fe-4S] clusters within its SPASM domain, both of which are fully ligated by four cysteine residues (110) (Figure 1.10D and Figure 1.13B). SuiB catalyzes the carbon-carbon crosslink formation between lysine and tryptophan residues on its precursor peptide SuiB during the maturation of the quorum sensing molecule Streptide (Figure 1.6).

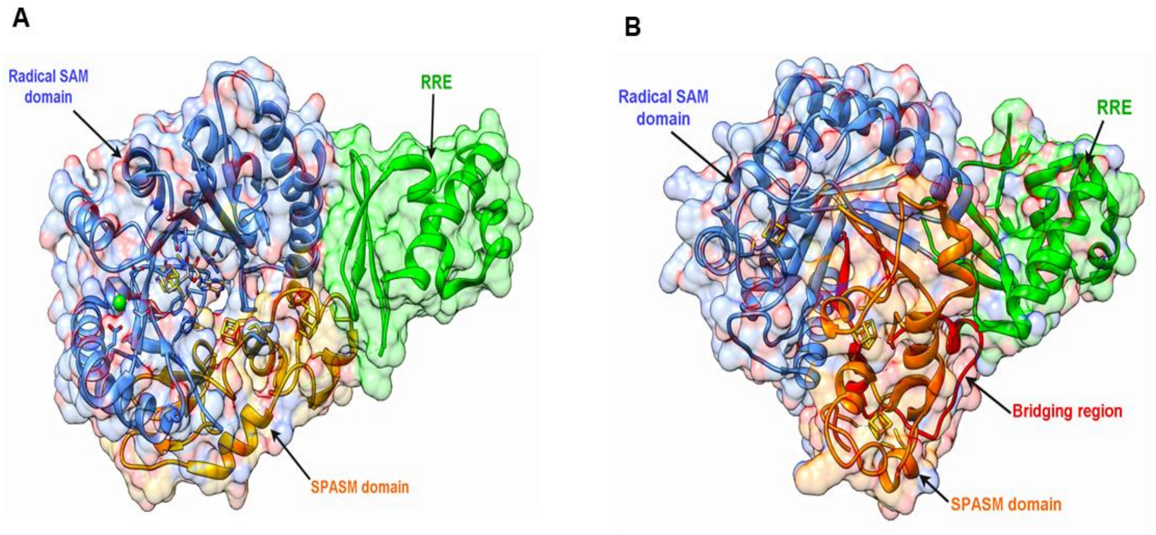


Figure 1.13 Complete crystal structures of the RS SPASM proteins. (A) CteB from *Clostridium thermocellum* (PDB: 5WHY) and (B) SuiB from *Streptococcus suis* (PDB:5Q1V). The N-terminal RRE region used for substrate recognition is colored green. The Radical SAM domain (blue) forms a partial TIM-barrel fold and houses the RS-[4Fe-4S] cluster. The C-terminal SPASM domain (orange) harbors two additional [4Fe-4S] clusters. The bridging region (red) which is unique to SuiB connects the radical SAM domain to the SPASM domain.

1.5 Peptide-derived Redox Cofactors

Peptide-derived redox cofactors are genetically encoded redox cofactors that are derived from post translational modification of the amino acid side chains within either a precursor peptide or a protein (125, 126). The best characterized peptide-derived redox cofactors are those that originate from tyrosine (127) and tryptophan (128) residues. Currently, there are five known peptide-derived redox cofactors and they include 2,4,5-trihydroxyphenylalanine quinone (TPQ), lysyl-tyrosine quinone (LTQ), tryptophan-tryptophylquinone (TTQ), cysteine-tryptophylquinone (CTQ), and pyrroloquinoline quinone (PQQ) (Figure 1.14). These quinocofactors participate in key biochemical processes in either prokaryotic and or eukaryotic organisms. For example, TPQ, which is also known as topaquinone spans both domains of life and functions as the prosthetic group of copper-containing amine oxidases (61, 127, 129). LTQ is present as a prosthetic group in lysyl oxidase, an enzyme involved in collagen and elastin production in mammals (61, 130). TTQ is used as the prosthetic group of aromatic amine dehydrogenases and methylamine dehydrogenases (126). CTQ is a protein-derived redox cofactor which functions as a prosthetic group in bacterial quinohemoprotein amine dehydrogenases(131, 132). PQQ is the prosthetic group of methanol and glucose dehydrogenases that catalyze the oxidation of alcohols and aldose sugars respectively (133–137). While the overall structures of these cofactors vary significantly, they do share a common feature – they are quinone based (Figure 1.14).

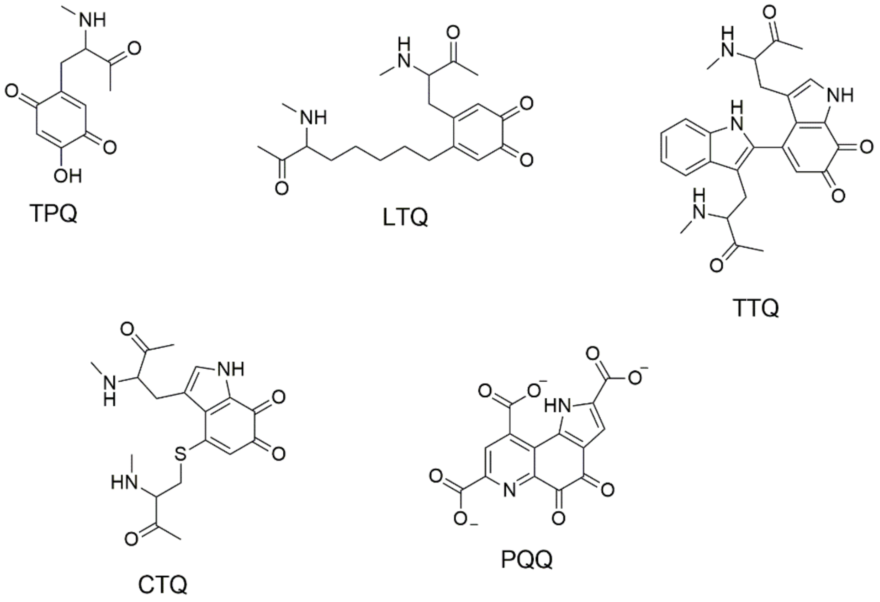


Figure 1.14 Structures of known peptide-derived quinocofactors. 2,4,5-trihydroxyphenylalanine quinone (TPQ), lysyl-tyrosine quinone (LTQ), tryptophan-tryptophylquinone (TTQ), cysteine-tryptophylquinone (CTQ), and pyrroloquinoline quinone (PQQ)

Of the five described peptide-derived quinocofactors, all except for PQQ are synthesized *in situ* of their functional enzymes. For instance, TPQ, LTQ, TTQ and CTQ are derived from tryptophan or tyrosine side chains of the enzymes that require them and found covalently bound to these enzymes as prosthetic groups (128, 131, 132, 138–140). PQQ on the other hand is derived from a RiPP precursor peptide after extensive chemical modification by dedicated posttranslational modification enzymes that are colocalized on the same operon as the precursor peptide. Because PQQ is the first known RiPP-derived redox cofactor, the following section focuses on its functions and biosynthesis.

1.6 Pyrroloquinoline quinone (PQQ)

1.6.1 Functions of PQQ

PQQ is a low molecular weight redox cofactor utilized by a variety of gram-negative bacteria. The majority of these bacteria catalyze oxidation of alcohols and aldose sugars in the periplasmic space via non-glycolytic pathways to produce cellular adenosine triphosphate (ATP) (137, 141, 142). These enzymes reduce the orthoquinone form of PQQ to its quinol form with the concomitant oxidation of their corresponding substrates (134, 141, 142) (Figure 1.15). The reduced PQQ is then re-oxidized by the release of two electrons (Figure 1.15) which are in turn used to reduce ubiquinone to produce ATP (142). The midpoint potential of PQQ was found to be -240 mV at pH 13 (143), making it more oxidative than nicotinamide adenine dinucleotide (NAD^+) and flavin mononucleotide (FMN) (144, 145). However, this midpoint potential is comparable to that of ubiquinone which is +66 mV at pH 7 (146, 147). Unlike most bacteria, PQQ is scarcely found in the tissues of animals due to their inability to biosynthesize it, and thus acquire PQQ from their diet. In mammals, the exact role of PQQ is yet to be fully understood however, some studies suggest its involvement in fertility and mitochondrial biogenesis (134, 148, 149).

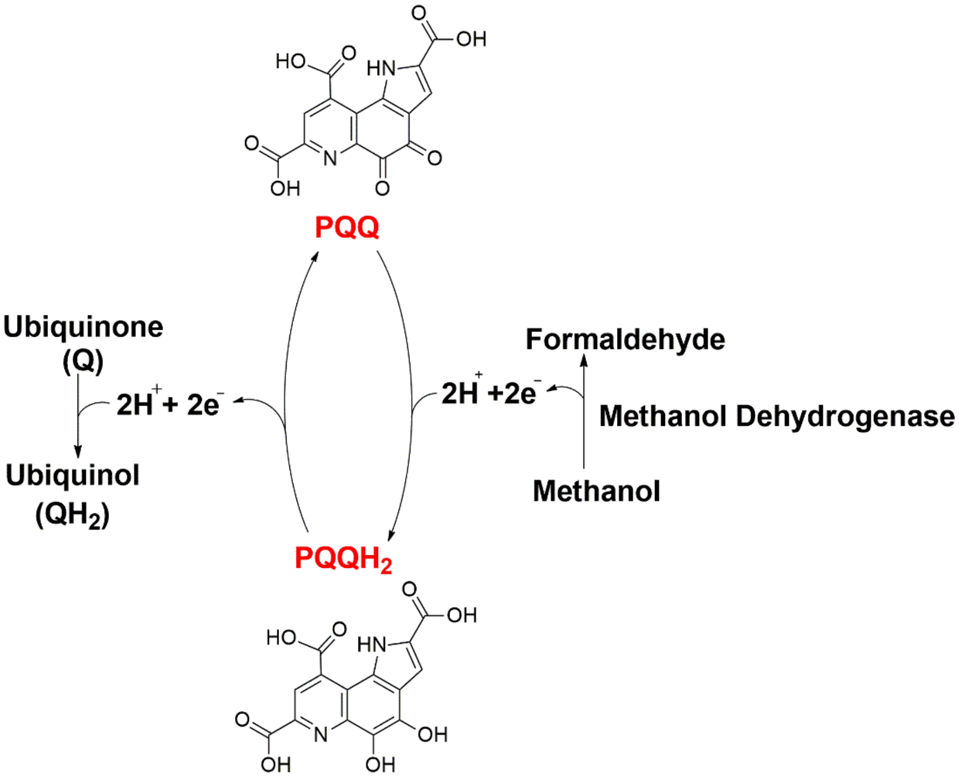


Figure 1.15 Function of PQQ in methanol dehydrogenases. Methanol dehydrogenases use PQQ to oxidize methanol to formaldehyde and in the process reduces PQQ to $PQQH_2$. The electrons in $PQQH_2$ are subsequently shuttled into the electron transport chain to generate ATP.

1.6.2 Biosynthesis of PQQ

PQQ is biosynthesized from the evolutionarily conserved glutamate and tyrosine side chains within the ribosomally produced precursor peptide PqqA. This is achieved through the dedicated actions of the products of five conserved genes in the PQQ biosynthetic gene cluster (150–153). These conserved genes annotated as *pqqABCDF/G* have been found in over 100 species of bacteria (47, 141). In all cases, *pqqA* encodes for the short precursor peptide, typically 20-30 amino acids in length containing conserved glutamate and tyrosine from which

PQQ is derived (141, 152) (Figure 1.16A). The genes *pqqB* and *pqqC* encode the hydrolase PqqB and cofactorless oxidase PqqC respectively(141, 154–156). The remaining genes *pqqDEF/G* encode the REE domain PqqD, the radical SAM enzyme PqqE and the peptidase PqqF/G respectively (96, 141, 154, 157, 158) (Figure 1.16A). PQQ biosynthesis commences with the ribosomal production of PqqA containing an N-terminal leader sequence and C-terminal core with conserved glutamate and tyrosine residues. The first step of the posttranslational modification occurs when the peptide chaperone PqqD engages both the leader sequence of the precursor peptide and the N-terminal RRE domain of the radical SAM enzyme PqqE, bringing both PqqA and PqqE into proximity and orienting the core region of PqqA into the active pocket of PqqE. In the presence of the physiological reductant flavodoxin/flavodoxin reductase, an electron is transferred from the radical SAM [4Fe-4S] cluster into the antibonding orbital of SAM resulting in the homolytic cleavage of SAM to form 5’-deoxyadenosine radical and methionine. This highly potent radical subsequently abstracts a hydrogen atom from the γ-position of the conserved glutamate to form a glutamyl radical and 5’-deoxyadenosine. The glutamyl radical attacks the phenyl ring of tyrosine followed by a subsequent aromatization of the ring in the presence of a base resulting in the formation carbon-carbon crosslink between glutamate and tyrosine on the PqqA (96, 158) The second step in the biosynthesis of PQQ is the removal of the leader peptide to liberate the small molecule for further processing. This step is carried out by the two component protease PqqF/G using three molecules of water

(157, 159). Other proteases are believed to be involved in the final hydrolysis step of biosynthesis after the removal of the leader peptide by PQQF/G (159). The next step is the oxidation of the phenyl ring of tyrosine by the iron-dependent hydroxylase PqqB. PqqB catalyzes the insertion of two atoms of oxygen into the phenyl ring using iron (II) thus facilitating the spontaneous ring closure of the PqqB reaction product to form 3a-(2-amino-2-carboxyethyl)-4,5-dioxo-4,5,6,7,8,9-hexahydroquinoline-7,9-dicarboxylic acid (AHQQ) (156). Following this is the final step in the biosynthesis of PQQ catalyzed by PqqC. PqqC is a cofactorless oxidase that uses molecular oxygen and hydrogen peroxide to catalyze the ring closure and an eight-electron oxidation of AHQQ to form PQQ (155, 160) (Figure 1.16B).

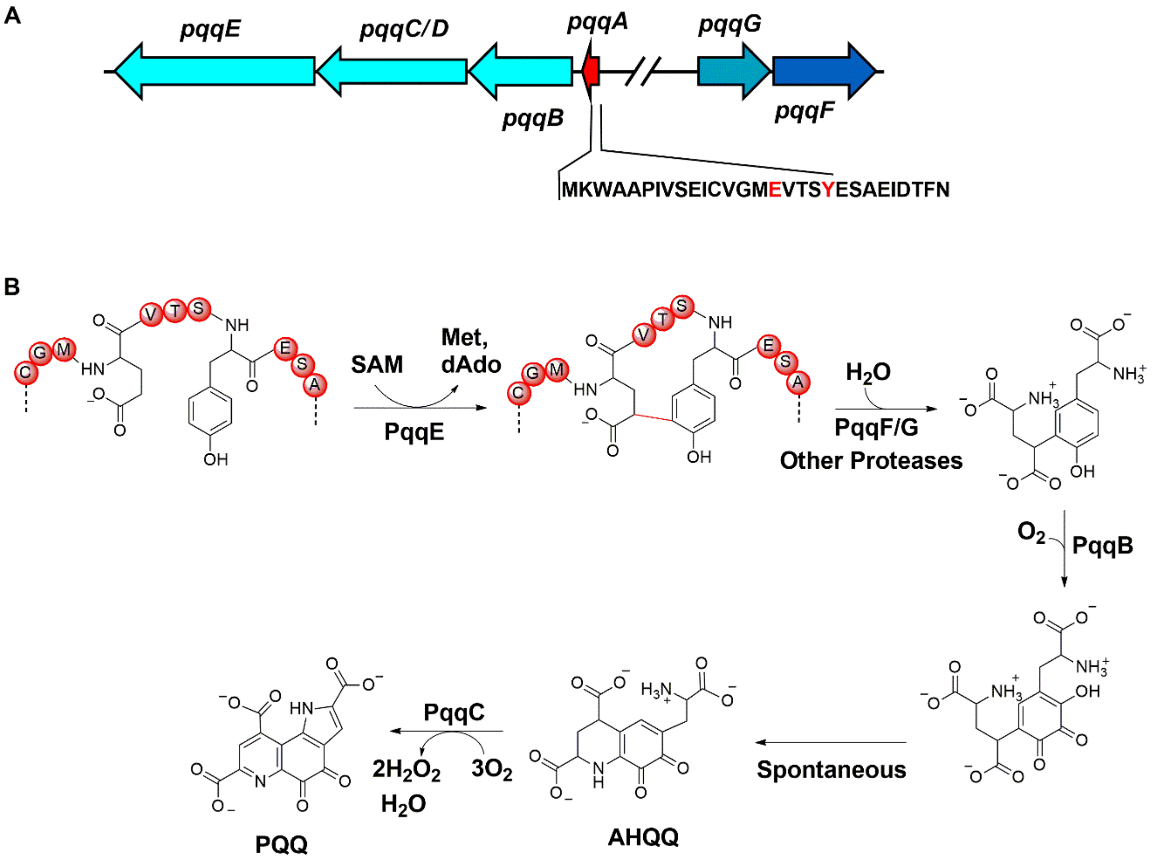


Figure 1.16 Condensed representation of PQQ the biosynthetic pathway. (A) PQQ biosynthetic gene cluster from *Methylobacterium extorquens* AM1. (B) Condensed representation of the formation of PQQ from conserved glutamate and tyrosine residues in the PqqA peptide. AHQQ is an intermediate that is further modified by the cofactorless oxidase, PqqC, to yield the biologically active cofactor PQQ.

Recently, a new and uncharacterized biosynthetic pathway which could also encode for a RiPP-derived redox cofactor has been identified. This biosynthetic pathway is widely distributed among the *Mycobacterium* genera and shares a striking resemblance to the PQQ cofactor and bacteriocin biosynthesis pathways, and thus the putative product of the pathway has been named mycofactocin. The focus of this dissertation is on the elucidation of the biosynthesis, structure and function of mycofactocin.

1.7 Mycofactocin

1.7.1 Implications of the Mycofactocin Biosynthetic Pathway in Mycobacterium tuberculosis.

Many species within the *Mycobacterium* genus are prominent and opportunistic pathogens that cause severe skin and lung diseases in humans and animals. Notable human pathogenic species include *M. ulcerans,* the causative agent of the debilitating skin disease buruli ulcer, and *M. tuberculosis,* the causative agent of the deadly pulmonary infection tuberculosis (161, 162). Tuberculosis is the world's leading killer of infectious diseases, exceeding the human immunodeficiency virus (HIV) and malaria. More than a third of the world's population has been predicted to be infected with tuberculosis (163), with more than 1.6 million global deaths annually (164). This scenario has been complicated by the fact that many *Mycobacteria* have become increasingly drug-tolerant, resisting the current first-line antibiotics isoniazid and rifampin (165–167). Taken together, there is an urgent need to discover new therapeutic targets for the treatment of tuberculosis. This relies on complete understanding of the metabolic pathways in *Mycobacteria,* especially *M. tuberculosis*. Important physiological functions such as those essential for growth and/or survival of *M. tuberculosis* are potential targets for new antimicrobial drugs. A bioinformatic study identified a biosynthetic gene cluster that is widely distributed in the *Mycobacterial* genus and present in all strains of *M. tuberculosis* (168, 169). This biosynthetic gene cluster is called the mycofactocin biosynthetic pathway. In addition to the bioinformatic study, a recent pan-genomic analysis of 36 strains of *M. tuberculosis* revealed that the genes of the mycofactocin gene cluster are among the top 5% core genes of *Mycobacteria* (168). This suggests that the genes of the pathway and thereby the putative product of the pathway, participate in important physiological processes in these organisms. Therefore, a detailed understanding of the biosynthesis and function of the product of this pathway will provide meaningful insights into chemistry and physiology of these pathogens. This can serve as avenues for the development of tuberculosis intervention strategies such as designing of drug targets or silencing of important immunomodulatory or virulence genes.

Mycofactocin has been proposed through bioinformatic studies to function as a redox cofactor. Although there are some *in vivo* data regarding the essentiality of the pathway in M. tuberculosis, the exact physiological function of mycofactocin is yet to be determined. In an initial study using transposon site hybridization (TraSH), where the genes required for *M. tuberculosis* survival were idenitified and mutated and the resulting *M. tuberculosis* mutant strains were recovered on mice spleen, mycofactocin genes were found to be nonessential for *M. tuberculosis* survival during infection (170). This finding was further substantiated by a transposon sequencing (TnSeq) experiment where the mycofactocin genes were found to be nonessential when *M. tuberculosis* is grown on enriched media (171). However, when *M. tuberculosis* was grown on cholesterol, the mycofactocin genes were found to be critical for the growth of the pathogen (172, 173). This finding was very significant because the persistence of the pathogen in the host during the latent phase of tuberculosis infection is dependent on its ability to mobilize and utilize the host’s cholesterol as carbon source (172–175). This unusual ability of *M. tuberculosis* to catabolize the host’s cholesterol allows it to evade the host’s defense mechanism and thereby sustain a persistent infection by decreasing the permeability of the cell wall to antituberculosis drugs such as rifampin (172, 176) (Figure 1.17). *In vitro* gene expression studies have indicated that a number of genes are upregulated and necessary for *M. tuberculosis* H37Rv growth on cholesterol and among these genes are the mycofactocin genes *Rv0693, Rv0694, Rv0695, and Rv0696* (*mftC, mftD, mftE, and mftF,* respectively) (177). In addition, the mycofactocin-associated Zn/NAD$^+$ -dependent dehydrogenase encoding the gene *Rv0761c* was found to be also essential for the catabolism of cholesterol by *M. tuberculosis* (170, 178).

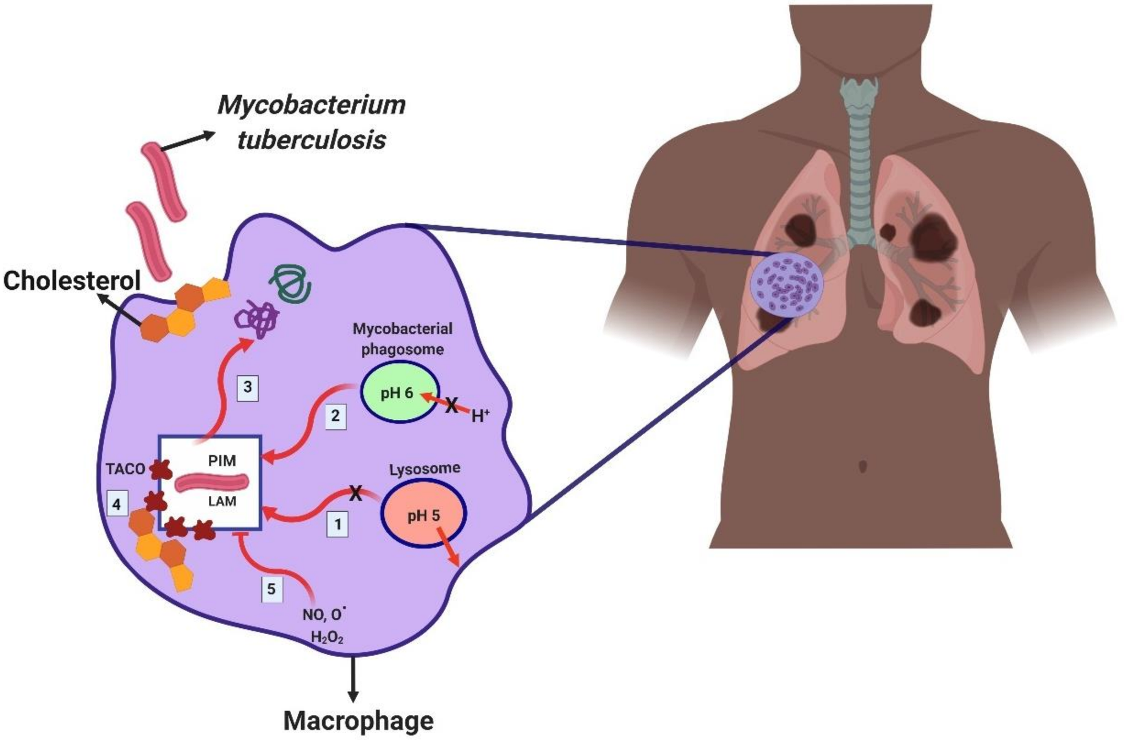


Figure 1.17 Ability of M. tuberculosis to utilize cholesterol allows it to circumvent the host's defense mechanism and persist within the phagosome. After internalization, the pathogen uses an array of strategies to modulate phagosome maturation. [1] Inhibition of macrophage fusion with the lysosome. [2] Inhibition of macrophage acidification. [3] Expression of the virulent Pro-Glu repetitive glycine-rich proteins. [4] Retention of cholesterol associated host protein (TACO) to prevent bacterial delivery to the lysosome. [5] While in the phagosome, the bacteria are protected from reactive oxidative radicals

While the role of Mycofactocin in the cholesterol catabolism is not very clear (175), these findings suggest that mycofactocin and mycofactocin-associated genes might be playing a supporting role in the virulence of *M. tuberculosis*. In support of this idea, recent gene knockout studies by the Kaufmann group have indicated that the genes *mftA-D and mftF*, along with the iron-dependent alcohol dehydrogenase *msmeg_6242* (TIGR04266), are essential for ethanol assimilation in the model organism *Mycobacterium smegmatis mc*^{2} *155*. In addition, the same study demonstrated that disruption of the mycofactocin genes led to imbalance of cellular concentrations of NAD$^+$/NADH (179). The results from these studies provided for the first time, *in vivo* evidence that mycofactocin may play a role in redox metabolism. A subsequent study by Jain group also demonstrated that the genes and the product of the pathway are indispensable for the growth of *M. smegmatis* on methanol as sole carbon source (180). Collectively, these findings suggest that mycofactocin genes and thereby the putative product of the pathway play an important role in the survival of *M. tuberculosis* during tuberculosis infection.

1.7.2 Occurrence and Architecture of the Mycofactocin Biosynthetic Pathway

Mycofactocin is a putative peptide derived redox cofactor derived from the actions of the products of six conserved genes, *mftABCDEF*. The pathway was first identified by Haft in 2011 through a bioinformatic study in which the sporadically distributed pathway was described (169). The study revealed that the gene product MftC, belongs to the RS-SPASM enzyme subfamily (82). Proteins of this subfamily are known to chemically modify peptides by installing intramolecular carbon-sulfur and carbon-carbon bonds (123, 181). Using the MftC clade as a starting point, Haft identified the remaining five mycofactocin genes in immediate neighborhood localized around *mftC* (Figure 1.18). From this bioinformatic study, it was shown that the association of *mftABCDEF* is such that they are either found all together or none at all, implying that they are part of the same biosynthetic pathway (169). In addition, Haft also identified other genes within the immediate neighborhood of the mycofactocin biosynthetic gene cluster, primarily a variety of dehydrogenases. These dehydrogenases, mainly members of TIGR03971, TIGR03989, and TIGR04266 families were found to be dependent upon the presence of the mycofactocin biosynthetic gene cluster. However, unlike mycofactocin genes, these ancillary genes do not share the same strict association with the pathway. Instead, they are found in various combinations only when the mycofactocin gene cluster is present, suggesting that they are dependent upon the product of this pathway, (169, 182). Significantly, members of the mycofactocin-associated dehydrogenase families were recently demonstrated to contain nonexchangeable NADH that were active *in vitro* only in the presence of nonphysiologically relevant redox mediators (182). In addition to these ancillary genes, the bioinformatic study identified also idenitified the gene mftR which is annotated as a member of the TetR family of regulators (183). Proteins within this family are single component regulators responsible for the regulation of a wide variety of biosynthetic pathways. Taken together, the bioinformatic study suggests that the expression of mycofactocin biosynthetic pathway is controlled by the gene product of this TetR like regulator denoted as MftR and the product of this pathway may be used by the pathway-associated short chain dehydrogenases (Figure 1.18).

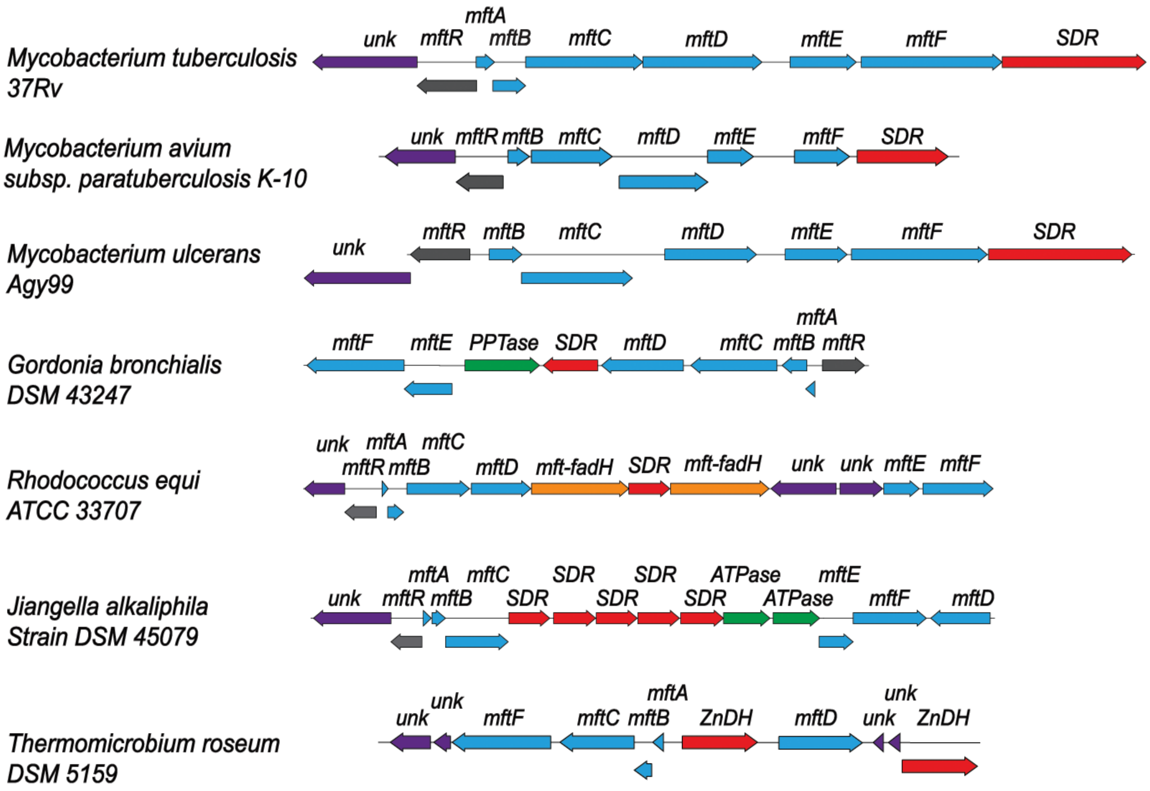


Figure 1.18 Gene organization of the mycofactocin biosynthetic pathway. The mycofactocin gene organization (blue) from various species indicates that the genes *mftABCDEF* are generally found together and that *mftR* (gray) is the regulator of the pathway. Ancillary genes (red), such as short-chain dehydrogenases (SDR) and zinc-dependent dehydrogenases (ZnDH) can be found in various combinations but are dependent upon the presence of the mycofactocin biosynthetic cluster. Additional genes, such as the mycofactocin fadH, are also associated with the pathway but for unclear reasons

The mycofactocin biosynthetic pathway is one of the most widely distributed RiPP systems in bacteria, however, this distribution is heavily skewed towards the *Mycobacterial* genus. A search of MftC in the Uniprot database, followed by careful examination of the gene context of putative homologs, led to the discovery of the mycofactocin biosynthetic pathway in 625 unique species and subspecies. Among these 625 species are representative members from the Proteobacteria, Chloroflexi, Euryarchaeota and Firmicutes phyla (Figure 15). However, the majority of these mycofactocin producing species belong to the Actinobacteria phylum. Within this phylum are a number of genera containing significant populations of species and subspecies that harbor the requisite mycofactocin genes. Notably, the *Streptomyces, Geodermatophilus, Nocardia, Nocardiodes, Frankia, Pseudonocardia, Rhodococcus* and *Gordonia* genera were found to contain 20, 17, 15, 17, 17, 15, 66 and 36 represented species and subspecies respectively. Interestingly, about half of all the mycofactocin-producing species identified were found to be concentrated in the *Mycobacterium* genus, with 300 represented species and subspecies (Figure 1.19). The mycofactocin gene cluster was also found to be evenly distributed in the genomes of both slow and rapid growing Mycobacteria (184), including human pathogenic variants such as *M. tuberculosis, M. avium, M. ulcerans* and *M. marinum.*

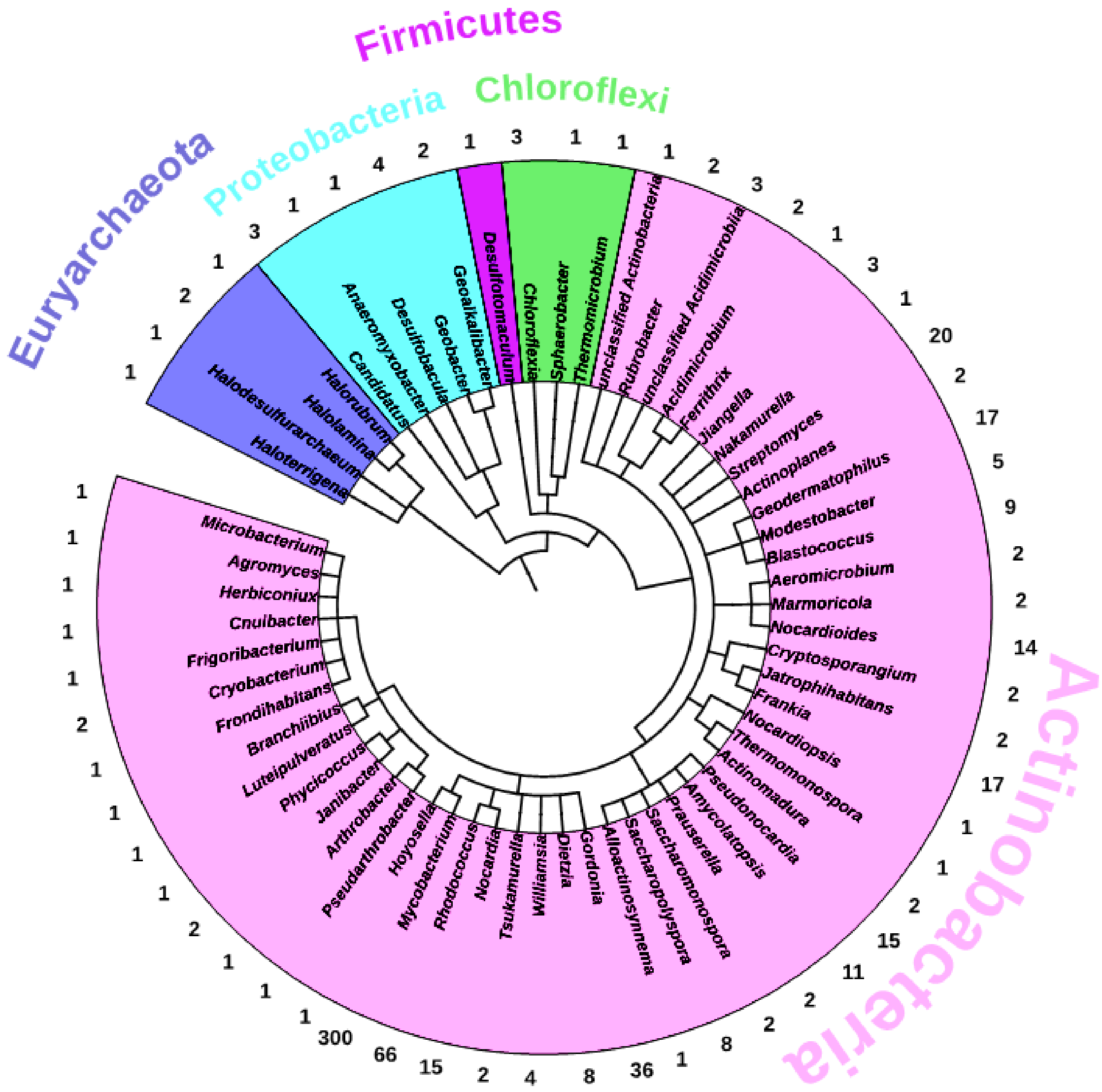


Figure 1.19 An iTOL-generated taxonomy based phylogenetic tree. The tree shows the number of mycofactocin encoding species identified in each genus. The indicated genera are color coordinated by phylum where pink represents Actinobacteria, green represents Chloroflexi, magenta represents Firmicutes, cyan represents Proteobacteria and purple represents Euryarchaeota.

1.7.3 Biosynthesis of Mycofactocin

Six strictly conserved genes, *mftABCDEF* were identified to be responsible for the biosynthesis of mycofactocin (169). The genes *mftA, mftB* and *mftC* encode the precursor peptide MftA, the peptide chaperone MftB and the radical SAM enzyme MftC respectively while the remaining genes *mftD*, *mftE* and *mftF* encode the flavin dehydrogenase MftD, the creatinase homolog MftE, and the glycosyltransferase MftF. Sequence homology alignment of MftA sequence from 250 different species revealed strict conservation of the C-terminal residues. The C terminal -SIDGMCGVY sequence was found to be strictly conserved among all 250 different species (Figure 1.20) (98, 169). The strict conservation of the C-terminal sequence led Haft to suggest that mycofactocin biosynthesis parallels that of PQQ. As described above, PQQ is a redox cofactor derived from extensive chemical modification of the evolutionarily conserved tyrosine and glutamate residues on the precursor peptide PqqA (see Figure 1.16 for details).

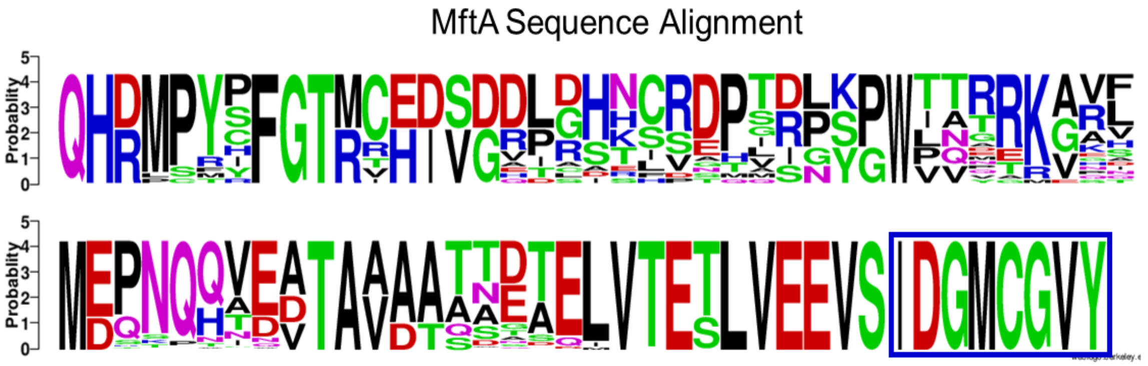


Figure 1.20 The mycofactocin precursor peptide and gene context. The amino acid sequence alignment of the MftA from 250 different species created by WebLogo. Highlighted in the blue box are the most conserved C-terminal resides of MftA.

In addition, the gene product, MftB was found to share high structural homology with the RRE domain PqqD, involved in the biosynthesis of PQQ (see Figure 1.7). Furthermore, sequence homology alignment of the radical SAM enzyme MftC revealed the similarity of its C-terminal extension to that of RS-SPASM enzymes known to catalyze carbon-carbon bonds and carbon-sulfur bonds (82). Taken together, the bioinformatic data indeed suggests that mycofactocin a RiPP derived product and its biosynthesis and function parallels to that of PQQ (169). To determine the structure of mycofactocin and thereby gain insight into its physiological role, recent efforts have aimed at *in vitro* reconstitution of each enzyme in the pathway, to understand their role and function in the biosynthesis of mycofactocin.

The biosynthesis of mycofactocin commences with the ribosomal production of the precursor peptide, MftA containing the strictly conserved C-terminal SIDGMCGVY motif. The first chemical transformation in the biosynthesis involves the RS enzyme MftC, and the RRE domain MftB in the presence of SAM. For this to occur, the RRE domain protein MftB, must engage both the radical SAM enzyme MftC and MftA to form a ternary complex, thereby bringing both MftA and MftC into catalytically relevant proximity and orientation for chemical transformations to occur. In the presence of SAM and a one-electron reductant such sodium dithionite (DTH), an electron is transferred from the RS-cluster into SAM resulting in the homolytic cleavage of SAM to form 5'-deoxyadenosine radical. This radical subsequently abstracts a hydrogen atom from the Cβ position of the C-terminal tyrosine to form a substrate radical. Delocalization of the radical into the phenyl ring followed by an oxidative decarboxylation of the tyrosine results in the formation of an α/β unsaturated product, MftA** (Figure 1.21A). Intriguingly, two independent studies, one from our lab demonstrated that MftC indeed, catalyzes the oxidative decarboxylation of the C-terminal tyrosine to form MftA** (74, 98). A subsequent study demonstrated that MftE, a member of the creatinine amidohydrolase family, was the likely candidate for being the next processing enzyme in the pathway. This was based on a biochemical study that concluded that MftE hydrolyzes the last two residues from MftA**, resulting in the formation of Val-p-(2-aminoethenyl) phenol moiety (185). However, in a follow-up study from our lab on the catalytic mechanism of MftC, it was shown that MftA** is an intermediate of a two-step reaction. In the second step, another molecule of SAM is used to generate a 5’deoxyadenosine radical which abstracts a hydrogen atom from the side chain of Val29 to form an alkyl radical. This newly formed radical subsequently attacks the adjacent α/β unsaturated bond of tyrosine to form a carbon-carbon crosslink and a radical at Cα position of tyrosine. This radical is then quenched by injecting an electron and a proton into the reaction to form a biologically relevant crosslinked product MftA* (99) (Figure 1.21B).

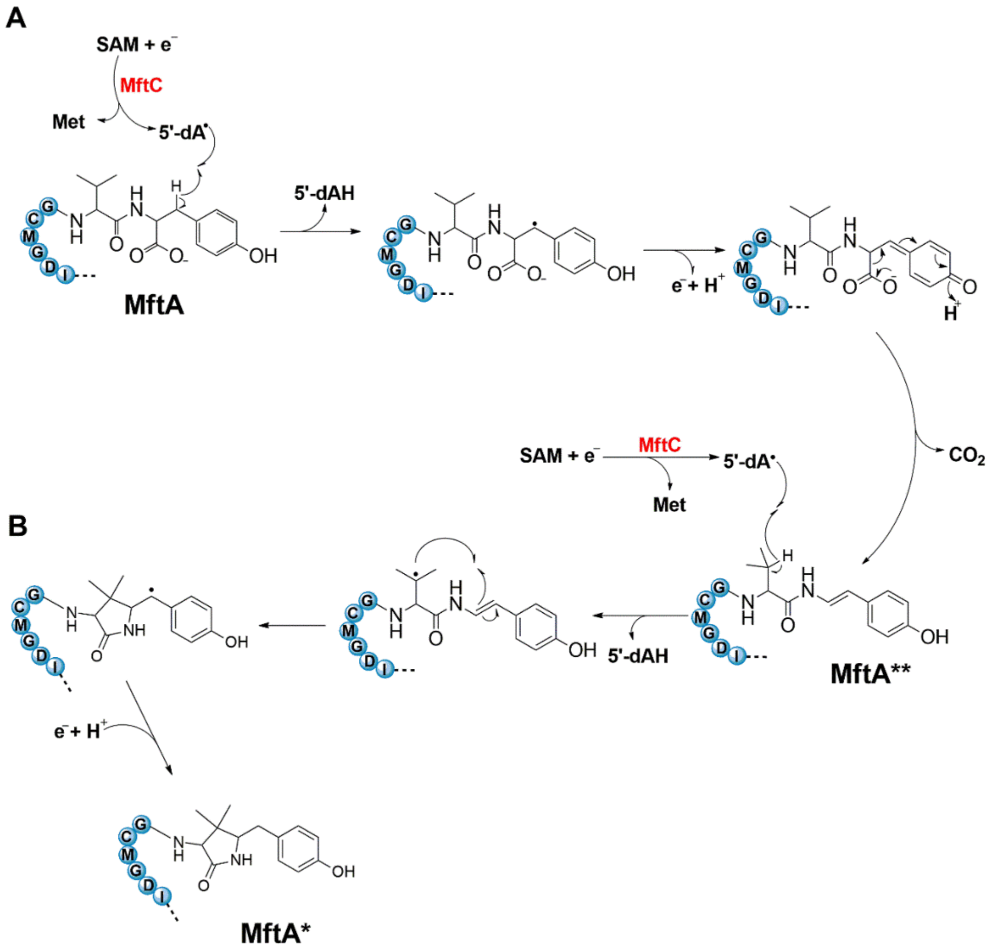


Figure 1.21 Proposed mechanism for MftC catalysis. The reaction mechanism of MftC catalysis showing the consumption of two equivalence of SAM to form the final product MftA*. (A) Oxidative decarboxylation reaction to form MftA** and (B) Redox neutral C-C crosslink reaction to form MftA*.

1.7.4 Objectives and scope of this dissertation

Because multiple products are produced by MftC and guiding experimental data regarding the structure of mycofactocin are lacking, a new question about the identity of the precursor for mycofactocin, MftA* or MftA**, has arisen. That is, which of the two products is recognized by MftE and carried forward in the biosynthesis of mycofactocin (Figure 1.22). In addition, the role of the remaining gene products MftD and MftF, in the biosynthesis of mycofactocin is unknown. Lastly, the notion that MftC catalyzes two distinct chemistries in the same active site, an oxidative decarboxylation of the C- terminus (Figure 1.21A) and a subsequent redox neutral C−C crosslink (Figure 1.21B) is interesting. This is because the reactions catalyzed by members of radical SAM protein family can be described as either oxidative or redox neutral, but MftC seems to be an outlier in that it can “redox-flip”. The ability of MftC to redox-flip, or accommodate both oxidative and redox neutral reactions in the same active site, is certainly unique to the RS-SPASM subfamily, in which all proteins characterized to date have been shown to catalyze net oxidative reactions on their peptide substrate (186).

The work presented in this dissertation focuses on addressing the concerns described above. This includes (i) investigation of the structural effect of MftC on its catalysis and how it’s able to redox-flip (ii) investigation of which product of MftC catalysis is carried forward in the biosynthesis of mycofactocin, (iii) investigation of the role of MftD in the biosynthesis of mycofactocin and the physiological role of the product of the pathway (Figure 1.22).

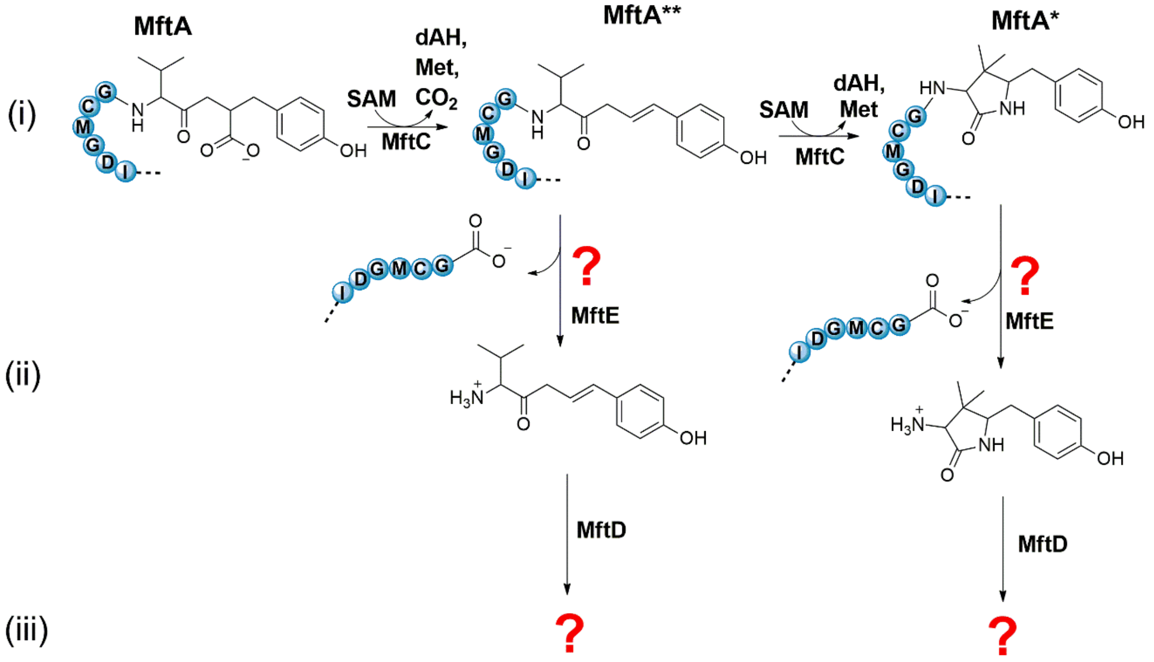


Figure 1.22 Steps in the biosynthesis of mycofactocin highlighting the focus of this dissertation. (i) MftC catalyzes the oxidative and redox neutral reactions in the same active site to form two products MftA* and MftA**. (ii) Which of the products of MftC is carried forward in the biosynthesis of mycofactocin? That is, does MftE recognize MftA* or MftA** and (iii) what kind of chemical modification does MftD impart on the product of MftE hydrolysis and what is the physiological role of mycofactocin.

I began by employing biochemical, spectroscopic and electrochemical techniques to investigate the roles of the auxiliary clusters in MftC catalysis. This was done using mutagenesis and activity assays to investigate the role of the auxiliary clusters in MftC. The number and type of auxiliary clusters in MftC were subsequently analyzed by Mossbauer spectroscopic techniques. This was a collaborative work with the Pandelia lab at Brandeis University. Furthermore, the redox potentials of all three clusters in MftC were measured using electrochemical techniques to gain insights into the redox flipping mechanism of MftC. This work was done in conjunction with the Elliot lab at Boston University.

Having investigated the role of the auxiliary clusters in MftC, the activity of MftE was revisited to clarify which product of MftC is carried forward in the biosynthesis of mycofactocin. To do this, activity assays were combined with high pressure liquid chromatography HPLC techniques to analyze MftE reactions with MftA, MftA* and MftA**. The structure of the product was solved by combining high resolution mass spectroscopy (HRMS) with NMR spectral workup. The effect of metals and other parameters on the rate of hydrolysis of MftE were also investigated. The results obtained from this study provided insights into the biologically relevant precursor for mycofactocin biosynthesis.

Having determined which product of MftC is the precursor for mycofactocin biosynthesis, the activity of the next processing enzyme, MftD was investigated. This was done by combining activity assays and biochemical techniques to demonstrate that MftD is active towards the product of MftE hydrolysis reaction. The type of modification that was imparted on the product of MftE hydrolysis by MftD was also investigated. This was done by combining HPLC and LCMS techniques with NMR spectral work up to solve structure of the reaction product. The midpoint potential for the new molecule was subsequently measured using cyclic voltammetry. Lastly, the physiological function of mycofactocin was probed by coupling it to the enzymatic activities of mycofactocin-associated short chain dehydrogenases. The experimental data presented in this dissertation have coalesced to provide the first direct evidence for the physiological role of mycofactocin.

CHAPTER TWO: METHODS

2.1 General Methods

2.1.1 *High Resolution Mass Spectrometry (HRMS)*

All samples for HRMS were analyzed by nano-flow LC-ESI-MS (Thermo Scientific LTQ Orbitrap Velos Pro) at the Biological Mass Spectrometry facility, University of Colorado Anschutz Medical Campus. Prior to analysis, samples were desalted using a C18 ZipTip (EMD Millipore). Data were analyzed using XCalibur Qual Browser v. 3.0.63 (Thermo Scientific).

2.1.2 Nuclear Magnetic Resonance (NMR) Assay

All NMR Spectra were recorded on Bruker UltraShield 500/54 Plus Spectrometer. Spectra were processed in Topspin v. 2.1 program (Bruker) using deuterated water as solvent. Water suppression signal was applied at a frequency of 2353.37 Hz. All NMR spectra were analyzed using MestReNova v. 12.0.0-20080 program (Mestrelab Research).

2.1.3 HPLC program

All reactions were filtered through 0.2 μm spin columns to remove all precipitated proteins. The supernatants were pipetted into 300 μL autosampler

HPLC vials and injected directly into a reverse-phase Jupiter C5, 5 µm, 4.6 × 250 mm column (Phenomenex) using Shimadzu Prominence-i LC-2030C HPLC and 5 mM sodium phosphate, pH 7.5 (buffer A) and 5 mM sodium phosphate in 70% acetonitrile pH 7.5 (buffer B). A linear gradient of solvent B from 0% to 35% was applied from 2 to 8 min, which was followed by a linear gradient to 100% solvent B from 8 to 17 min. Solvent B was then held constant at 100% from 17 to 19 min followed by linear decrease to 100% solvent A from 19 to 22 min. Solvent A was held constant for an additional 2 min to re-equilibrate the column before subsequent sample injections. Chromatograms were reported at 280 nm while monitoring wavelengths between 200 and 400 nm. New peaks were isolated, lyophilized, and analyzed by NMR and HRMS

2.1.4 Analytical size exclusion chromatography

Analytical size exclusion chromatography was run on purified wildtype and variants of MftC, MftE and MftD proteins by injecting 100 µl of the protein (200 µM) onto a Superdex 200 10/300 GL column (GE Life Sciences) at a flow rate of 0.8 mL/min and using the protein storage buffer as the mobile phase. Biorad Gel Filtration Standards containing protein standards for 670 kDa (thyroglobulin), 158 kDa (γ-globulin), 44 kDa (ovalbumin), 17 kDa (myoglobin) and 1.35 kDa (vitamin B-12) were used to prepare a standard curve.

2.2 Expression and purification of MftA from Mycobacterium ulcerans Agy99

The unannotated gene sequence of MftA from *Mycobacterium ulcerans* Agy99 was cloned into *pET28HST* vector using *NdeI* and *XhoI* restriction sites. The plasmid was designed to add a His_6-SUMO tag to the N-terminus of the protein. The sequence verified *mftA/pET28HST* plasmid was transformed into BL21 Star (DE3) (Invitrogen) competent cell line and grown overnight. The overnight culture was used to inoculate 4 L of LB medium. The cells were grown at 37°C with shaking at 200 rpm until OD_{600} ~ 0.6 was reached, at which point protein production was induced by the addition of 1 mM IPTG (isopropyl-β-thiogalactopyranoside) and grown at 21°C overnight with shaking at 200 rpm. The cell culture was centrifuged at 7000 rpm for 10 min, and the resulting cell pellet was suspended in 5-fold cell weight lysis buffer (50 mM Tris, 200 mM NaCl, 25 mM Imidazole, pH 7.6). To the suspension was added 0.1 mg/g of lysozyme, 0.1 mg/g of DNase and 1% w/v CHAPS and stirred for an additional 30 mins. Lysis was completed by sonicating the suspension for 5 mins and centrifuging at 13000 rpm for 15 min. The supernatant was loaded onto a 5 mL HisTrap column (GE Life Sciences) using an ÄKTA FPLC system. The column was washed with 5 column volumes of lysis buffer (50 mM Tris, 200 mM NaCl, 25 mM Imidazole, pH 7.6). Bound proteins were eluted using elution buffer containing 50 mM Tris, 200 mM NaCl, 300 mM Imidazole, pH 7.6. The purified SUMO-MftA protein fraction was collected and buffer exchanged into the sodium phosphate buffer (5 mM

Na_2HPO4, 1 mM TCEP, pH 7.5) using HiPrep 26/10 Desalting column (GE Life Sciences) and concentrated using 10 kDa spin columns (Amicon ultra). Tobacco etch virus (TEV) protease was added and incubated at 4°C overnight. The resulting protein/protease mixture was centrifuged at 14000 rpm for 5 min and the supernatant was analyzed by reverse phase HPLC using the program described above.

2.3 Expression and purification of MftB from Mycobacterium ulcerans Agy99

The *Mul_0772* gene sequence encoding for MftB from *Mycobacterium ulcerans* Agy99 (Uniprot: A0PM48) was cloned into *pGEX6p-1* vector using *BamHI* and *XhoI* restriction sites to obtain the GST-MftB fused protein. The sequence *verified MumftB/pGEX6p1* plasmid was transformed into *E. coli* T7Iq Express (NEB) and grown overnight at 37 ^{o}C and 200 rpm. The overnight culture was used to inoculate 4 L of LB medium containing 100 µg/mL of Ampicillin and 34 µg/ml of chloramphenicol. The culture was grown at 37 ^{o}C and 200 rpm. When an OD_{600} ~ 0.6 was attained, protein production was induced by addition of isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 1 mM and allowed to grow at 21 ^{o}C overnight with shaking at 200 rpm. Cells were harvested and suspended in five times the mass of pellets of lysis buffer containing 50 mM HEPES (pH 7.5), 300 mM NaCl and 1 mM tris(2-carboxyethyl) phosphine (TCEP). To this suspension was added 0.1 mg/g of lysozyme, 0.1 mg/g of DNase and 1 % w/v

CHAPS and stirred for 30 min. Lysis was completed by sonication for 5 min and the lysate was clarified by centrifugation at 13000 rpm for 10 min. The supernatant was introduced into Coy Lab anaerobic chamber maintained under an atmosphere of 90% nitrogen and 10% hydrogen gases (oxygen concentration maintained below 20 ppm by palladium catalysts). The supernatant was loaded onto a 5ml GSTrap FF column (GE Healthcare) pre-equilibrated with lysis buffer using an AKTA Start FPLC (GE Healthcare). The loaded column was washed with five column volumes of lysis buffer and bound protein was eluted using 50 mM HEPES buffer (pH 7.5) containing 300 mM NaCl, 30 mM Glutathione HCl and 1 mM TCEP. Fractions containing proteins were pooled together and buffer exchanged into 50 Mm HEPES storage buffer (pH 7.5) containing 300 mM NaCl, 10% glycerol and 1 mM TCEP using PD-10 columns (GE Healthcare). The protein was concentrated using 10 kDa spin concentrators (Millipore Amicon Ultra), aliquoted into 1.5 mL HPLC vials, flash-frozen and stored at -80 °C until use.

2.4 Generation of RS, Aux I, and Aux II KO Variants of MftC.

To produce MftC knockout variants, cysteine residues identified in the bioinformatic study described above were replaced with alanine to generate the radical SAM cluster knockout (RS KO), the auxiliary I cluster knockout (Aux I KO), and the auxiliary II cluster knockout (Aux II KO) variants of MftC. Site-directed mutagenesis and sequence verification were performed by Genscript. For the RS

KO variant, the C30A/C37A double mutant was used. For the Aux I KO variant, the single C323A mutant was used. For the Aux II KO variant, the C310A/C341A double mutant was used.

2.5 Expression and purification of MftC from Mycobacterium ulcerans Agy99

The *Mul_0773* gene sequence encoding for MftC from *Mycobacterium ulcerans* Agy99 (Uniprot: A0PM49) was cloned into *pET28a* vector using *NdeI* and *XhoI* restriction sites. The sequence verified *MumftC/pET28a* plasmid was transformed into *E. coli* BL21 (DE3) competent cells containing the iron-sulfur cluster builder plasmid pPH149 for protein expression. An overnight culture of E. coli BL21 (DE3)- *MumftC/pET28a/pPH149* was grown in 10 mL of LB medium containing 50 µg/mL of kanamycin and 34 µg/mL chloramphenicol at 37 ^oC overnight at 200 rpm. The overnight culture was used to inoculate 4 L of terrific broth (TB) growth medium and incubated at 37 ^oC with shaking at 200 rpm until an OD600 of ~1.0 was attained. Protein production was induced by the addition of IPTG to a final concentration of 1 mM, 0.75g/L of sodium fumarate and 1 mL/L of 1000 x trace metals. Following a 20 h induction period at 21 ^oC and 200 rpm, cells were harvested by centrifugation at 7000 rpm for 10 min and stored at -80 ^oC until lysis. All purification steps were carried out in the anaerobic chamber. The thawed cells were introduced into the anerobic chamber and resuspended in five times the mass of cell pellet of lysis buffer (pH 8.0) containing 50 mM HEPES, 300 mM KCl and 50 mM imidazole. To this suspension was added 0.1 mg/g of lysozyme, 0.1 mg/g of DNase and 1 % w/v CHAPS and stirred for 30 min at room temperature to complete the lysis process. The resulting lysate was clarified by centrifugation at 13000 rpm for 10 min at 4 ^{o}C. The supernatant was loaded onto a 5 ml HisTrap FF column (GE Healthcare) pre-equilibrated with lysis buffer using an AKTA Start FPLC. The loaded column was washed with five column volumes of lysis buffer and the bound protein was eluted using 50 mM HEPES buffer (pH 8.0) containing 300 mM KCl, 300 mM imidazole. Fractions containing MftC proteins were pooled together and buffer exchange in storage buffer (pH 8.0) containing 50 mM HEPES, 300 mM KCl, 10 mM DTT and 10% glycerol over PD-10 columns (GE Healthcare). *Mu*MftC cluster knockout strains were expressed and purified in the same manner as described above. The resulting volume of MftC protein was reconstituted as described below.

2.6 Reconstitution of MftC

Purified MftC protein was reconstituted anaerobically as previously described (98, 99). In summary, to a 15 mL solution of 150 µM MftC in storage buffer (pH 8.0) containing 50 mM HEPES, 300 mM KCl, 10% glycerol and 10 mM DTT, was added 1.8 mM $FeCl_3$ and 1.8 mM Na_2S and stirred for 30 min on ice. The reconstituted protein was centrifuged at 5000 rpm for 5 min to remove precipitates. The supernatant containing soluble MftC protein was buffer exchanged into fresh storage buffer (pH 8.0) containing 50 mM HEPES, 300 mM

KCl, 10% glycerol and 10 mM DTT using a PD-10 column (GE Healthcare). The resulting volume of reconstituted protein was concentrated using 30 kDa spin columns (Millipore Amicon Ultra), aliquoted into 1.5 mL airtight HPLC vials, flash-frozen and stored at -80 ^oC until use.

2. 7 Bioinformatic Analysis of MftC

To bioinformatically analyze the gene context and distribution of MftC, the amino acid sequence for *M. ulcerans* Agy99 MftC (Uniprot: A0PM49) was used in a BLAST search of the Uniprot database. The BLAST search used in default settings except that the E-threshold was set to 0.001 and the number of hits was set to 1000. Approximately 760 hits with a sequence identity of >35% were carried forward for the remaining analysis. Proteins with ≥85% sequence identity (~310) were classified as MftC by default. The genomes of the organisms that contained homologous proteins with sequence identities between 85 and 35% were examined to verify that the putative homologues were indeed MftC. The curated genomes of the organisms found on the EnsemblBacteria database or the NCBI Gene database were searched for the precursor peptide MftA. The homologues of the genomes containing an annotated MftA with the conserved sequence −SIDGMCGVY were confirmed as MftC (169). Those genomes that did not have an annotated MftA were further examined to identify at least two other genes associated with mycofactocin biosynthetic cluster, such as *mftB, mftD, mftE, or mftF* (182). If the genome containing the MftC homologue contained two of the aforementioned genes within a 10 kb window, the homologue was validated as MftC. If a curated genome in the two databases was unavailable for the organism that contained the putative homologue, the homologue was classified as annotated and carried forward in the analysis as if it were MftC. Of the remaining ~720 homologues, species duplicates were removed, leaving 625 unique species containing a putative MftC. To identify which cysteines on MftC are strictly conserved, a sequence alignment of all 625 unique MftC amino acid sequences was carried out using the Constraint-based Multiple Alignment Tool (COBALT) provided on the NCBI Web site (ncbi.nlm.nih.gov/tools/cobalt/cobalt.cgi?CMD=Web). The sequence alignment file was downloaded and used as the input in the server program, Weblogo (187) for visualization. MftC protein structure models were generated by the server program, Phyre2.0.1. Using the expert mode program, one-to-one threading, the amino acid sequence for M. ulcerans MftC was input along with the Protein Databank identifier for anSME (PDB ID: 4K36) (117), CteB (PDB ID: 5WHY) (83) or SuiB (PDB ID: 5V1T) (110). The threaded MftC structural models were overlaid with their respective parent structure using Chimera's Match Maker tool to visualize the conserved cysteines of MftC and their vicinity to the parent structure's [Fe−S] clusters.

2.8 Iron and sulfur quantification

Protein concentrations were estimated using a Bradford assay. To determine the iron concentration of reconstituted MftC, 100 μL of 20 μM protein was mixed with 10 μL of 3 M trichloroacetic acid. The precipitated protein was centrifuged at 15,000g for 10 min. The supernatant was carefully removed and diluted with 330 μL of diH2O. To the solution, 20 μL of sodium ascorbate, 20 μL of 10 mM ferrozine, and 20 μL of saturated sodium acetate were added. The absorbance change due to the ferrozine−iron complex was measured at 562 nm (**ε** = 27.9 mM^{-1} cm^{-1}). To determine the sulfur concentration, 200 μL of 10 μM protein was mixed with 600 μL of 1% (w/v) zinc acetate and 50 μL of 7% (w/v) sodium hydroxide. The solution was incubated at room temperature for 15 min after which 150 μL of 0.1% (w/v) N, N-dimethyl-p-phenylenediamine (in 5 M HCl) and 150 μL of 10 mM $FeCl_3$ (in 1 M HCl) were added. The solution was vortexed and incubated at room temperature for 20 min. The absorbance from methylene blue was measured at 670 nm (ε = 34.5 mM^{-1} cm^{-1}). For both assays, blanks were generated by substituting protein with deionized water.

2.9 *MftC activity assays*

All reactions were conducted in a Coy anaerobic chamber maintained under an atmosphere of 97% nitrogen and 3% hydrogen gases. MftC variants were assayed for uncoupled cleavage of SAM and complete modification of MftA following previously described procedures (99). For SAM cleavage assays, a reaction containing 2 mM SAM, 10 mM DTT, 2 mM dithionite, and 100 μM of wildtype MftC or variant in 50 mM HEPES, 200 mM NaCl was incubated in an anaerobic chamber overnight. The reactions were analyzed by HPLC over a Phenomenex Jupiter C4 5 μm 300 Å reverse phase column using 0.1% TFA and acetonitrile as solvents. For peptide modification assays, the same reaction was set up with the addition of 100 μM MftA and 100 μM MftB. The reactions were analyzed on reverse phase HPLC as described above.

2.10 *Preparation of MftC for Mossbauer Spectroscopy*

$^{57}Fe_2O_3$ (Cambridge Isotope Laboratories) was converted to $^{57}FeCl_3$ by incubating the oxide in 1 M HCl at 80 °C overnight. An overnight starter culture containing *E. coli* Bl21 (DE3) transformed with *pPH151* and *mftC/pET28a* was used to inoculate 8 L of M9 minimal media (Sigma-Aldrich) supplemented with 20 amino acids (0.2 g/L) and 1X auto induction metals, excluding iron. The large-scale culture was incubated at 37 °C with shaking until OD600 ~ 1.2 was reached. At this point, 1 mM IPTG, 0.75 g/L sodium fumarate, and ^{57}Fe (75 μM) were added to the cultures and the flasks were sealed. The induced cultures were incubated at 20 °C overnight, and the cells were harvested by centrifugation. The cell pellet was introduced into the anaerobic chamber and suspended in anaerobic lysis buffer, and MftC was purified as described above. Reconstitution of MftC was carried out as described above except that $^{57}FeCl_3$ was substituted for $FeCl_3$ and the reconstituted protein was buffer exchanged over a PD-10 column twice. Iron and sulfur quantification were performed for all protein samples prior to being characterized by Mossbauer spectroscopy. All Mossbauer spectroscopic experiments were performed by the Pandelia lab at Brandeis University.

2.11 *Electrochemical Measurements of MftC*

Electrochemical experiments on MftC were performed by the Elliott Lab at Boston University using the following procedures. They performed all experiments anaerobically in a MBraun Labmaster glovebox using a PGSTAT12 potentiostat. A three-electrode configuration was used in a water-jacketed glass cell. A platinum wire was used as the counter electrode, and a standard calomel electrode was used as the reference electrode. Reported potentials are relative to the standard hydrogen electrode. Baseline measurements were collected using an edge-plane graphite (EPG) electrode pretreated by mechanical polishing with sandpaper followed by incubating overnight with 14 μL of 3 mg of multiwalled carbon nanotubes (MWCNT) (Sigma, 10 ± 1 nm × 4.5 ± 0.5 nm × 3.6 μm) dissolved in 1 mL of dimethylformamide (DMF) by sonication for 15 min. The electrode was then rinsed and placed into a glass cell containing a mixed buffer solution (10 mM HEPES/CHES/ MES/TAPS/CAPS, 200 mM NaCl) and pH 7.5. A 4 μL aliquot of either wild-type MftC, the RS KO, or the Aux II KO variant was applied directly to the electrode surface and allowed to sit for approximately 2 min at room temperature before being placed into the buffer cell solution. Cyclic voltammograms were collected at 4 °C with a scan rate of 50 mV/s and electrochemical signals were analyzed using QSoas (188), where the data are presumed to represent all three iron−sulfur clusters of MftC. Voltammograms were fit to three Nernstian components as described where the surface coverage for all components is the same, and napp and the redox potentials are fit as floatable parameters (189). Square wave measurements were collected at 4 °C with a frequency of 10 Hz and an amplitude of 20 mV.

2.12 ***Production of MftA* and MftA*****

All reactions were conducted in a Coy anaerobic chamber maintained under an atmosphere of 97% nitrogen and 3% hydrogen gases. The oxygen concentration was maintained below 20 ppm by palladium catalysts. MftC reactions were conducted in a total volume of 500 µL and contained 50 mM HEPES reaction buffer (pH 8.0), 300 mM KCl, 10 mM dithiothreitol (DTT), 10% glycerol. The order of addition of reagents is as follows: 300 µM of substrate MftA and 300 µM of GST-MftB were added sequentially followed by 5 min of incubation period. To the reaction was added 300 µM MftC, 2 mM S-adenosylmethionine (SAM) and 2 mM sodium dithionite (DTH) consecutively and the resulting mixture was incubated at room temperature overnight. The reaction was then filtered with 10 kDa cut off 500 µL spin columns to remove all precipitated proteins. The filtrate was pipetted into 1.5 mL HPLC autosampler vials and MftA* was separated by reverse-phase chromatography on a Shimadzu Prominence-*i*-LC-2030C HPLC as described above. MftA* and MftA** peaks were collected, lyophilized and analyzed by NMR or used for subsequent reactions.

2.13 Generation of a sequence similarity network for the creatinine amidohydrolase family.

A sequence similarity network (SSN) of the creatinine amidohydrolase family was generated using the Enzyme Function Initiative Enzyme Similarity Tool (ESI-EST) (190). The Interpro identifier IPR003785 was used in Option B method and sequences were restricted to those of UniRef 90 seed sequences. The network was constrained using an E-value of 10-35, sequence lengths between 220-280 and sequence identities of 50%. The resulting network, containing 2,344 nodes, was colored using the Color SSN tool on the EFI-EST website. The network was visualized in Cytoscape3.7.2 and the mycofactocin cluster was re-colored to green.

2.14 Expression and purification of MftE from Mycobacterium ulcerans Agy99

The amino acid sequence of MftE (Uniprot: A0PM51) from Mycobacterium ulcerans Agy99 was used for *E. coli* optimized gene synthesis (Genscript). The resulting *mftE* gene was cloned into *pGEX6p1* using the *BamHI* and *XhoI* restriction sites. The sequence verified *MumftE/pGEX6p1* plasmid was transformed into chemically competent E. coli BL21 star (DE3). An overnight culture of E. coli BL21 star (DE3)-MumftE/pGEX6p1 was grown in 10 mL of LB medium containing 100 µg/mL of Ampicillin at 37 ^{o}C and 200 rpm was used to inoculate 4 L of the same medium. The culture was incubated at 37 ^{o}C with shaking at 200 rpm until an OD_{600} ~ 0.8 was reached. Protein production was induced by the addition of ITPG to a final concentration of 1 mM, and the cells were incubated for an additional 24 h at 20 ^{o}C with shaking at 200 rpm. The cells were harvested and resuspended in lysis buffer (pH 7.7) containing 50 mM HEPES and 200 mM NaCl. To the suspension was added 0.1 mg/g of lysozyme, 0.1 mg/g of DNase and 1 % w/v CHAPS and stirred on ice for an additional 30 min. The cells were disrupted by sonication for 4 min and the resulting lysate was centrifuged at 13000 rpm for 10 min. The supernatant was loaded onto a 5 mL GSTrap FF column equilibrated with lysis buffer using an AKTA Pure FPLC. The column was washed with five column volumes of lysis buffer and the bound protein was eluted with 50 mM HEPES buffer (pH 7.7) containing 200 mM NaCl and 30 mM reduced glutathione. Fractions containing MftE proteins were pooled together and buffer exchanged into storage buffer (pH 7.7) containing 50 mM HEPES, 200 mM NaCl, 1 mM TCEP and 1% glycerol, using a HiPrep 26/10 desalting column. The protein was concentrated using a 30 kDa spin concentrators, aliquoted into 1.5 mL HPLC vials, flash-frozen and stored at -80 ^{o}C until use.

2.15 Expression and purification of MftE from Mycobacterium smegmatis mc^2 155

The *Ms_1425* gene sequence for *Ms*MftE (Uniprot: A0QSC0) was amplified from the genomic DNA isolated from *Mycobacterium smegmatis* mc^2 155 and cloned into *pGEX6P-1* (GE Life Sciences) using the *BamHI* and *XhoI* restriction sites. The *Ms*MftE protein was expressed and purified in a method similar to that described for *Mu*MftE above.

2.16 Analysis of metal ion content in MftE

The types and amounts of metal ions present in MftE was determined by ICP-MS in the Majestic Lab at the Department of Chemistry and Biochemistry, University of Denver. Prior to the analysis, a fraction of the as-purified MftE was reconstituted on ice with two eq of Fe $(NH_4)_2(SO_4)_2$ and/or $ZnCl_2$. The reconstituted protein was buffer exchanged into storage buffer (50 mM HEPES, 200 mM NaCl, 1 mM TCEP and 1% glycerol, pH 7.7) using PD-10 columns (GE Life Sciences) and concentrated using 10 kDa spin (Millipore Amicon Ultra). The Zn^{2+} and/or Fe^{2+} reconstituted MftE was used for protein concentration estimation by Bradford assay and for metal quantification by ICP-MS.

2.17 Substrate Screening for MftE and production of 3-amino-5-[(p-hydroxyphenyl) methyl]-4,4-dimethyl-2-pyrroli-dinone (AHDP).

MftE modification reactions were set up as follows: 50 µM of MftE was added to a reaction buffer containing 50 mM HEPES, 300 mM NaCl, 1% glycerol, 1 mM TCEP, pH 7.7. For substrate specificity screening, 50 µM (1 eq) of ferrous ammonium sulfate and zinc chloride were added and incubated-on ice for 5 mins. Modified and unmodified MftA (50 µM MftA, 50 µM of MftA* and or 50 µM MftA**) were added to the reaction mixture where appropriate. Control reactions were set up in similar fashion except buffer was substituted for individual reagents or proteins. Reactions were incubated at room temperature for 3 hr after which they were filtered through a 0.2 µm spin columns to remove all precipitated proteins. The supernatant was either flash frozen and stored at -80 ^{o}C until ready to be analyzed or injected directly onto a reverse-phase HPLC system and separated according to the above HPLC described program. New peaks were isolated and analyzed by NMR and LC/MS. Activity assays for MftE from *M. smegmatis* were set up in the manner described for *Mu*MftE.

2.18 Labeling of MftA

The generation of the ^{13}C, ^{15}N labelled MftA followed previously reported procedures, with some modifications (99). In short, M1W *mftA/pET28HST* plasmid was transformed into *E. coli* BL21 Star (DE3) (Invitrogen). An overnight culture was used to inoculate 4 L of M9 minimal media containing 4 mL/L glycerol, 25 mM

Na_2HPO_4, 25 mM KH_2PO_4, 50 mM NH_4Cl, 5 mM Na_2SO_4, 1 mM $MgSO_4$, 100 µM $CaCl_2$, 50 µM trace metals, 50 mg/L each of the 18 unlabeled amino acids and 50 mg/L of ^{13}C and ^{15}N-labeled tyrosine and valine (Cambridge isotope laboratory) pH 6.8. The cells were grown at 37 ^oC A_{600} ~ 0.6. His$_6$-SUMO-MftA production was induced with 1 mM IPTG and the temperature was reduced to 21 ^oC. After ~12 hr, cells were harvested by centrifugation at 7000 rpm for 10 mins at 4 ^oC and the resulting pellet was suspended in lysis buffer (50 mM HEPES, 200 mM NaCl, 25 mM imidazole pH 7.6). CHAPS (1%), lysozyme (1mg/g of cell paste) and DNase (0.5mg/g of cell paste) were added and stirred on ice for 20 mins. Cell lysis was completed by sonication and the lysate was centrifuged at 13000 rpm for 10 mins at 4 ^oC. The supernatant was loaded onto a 5 mL HisTrap column using an AKTA Pure FPLC (GE Healthcare) system. The column was washed with lysis buffer followed by elution buffer (50 mM HEPES, 200 mM NaCl, 300 mM Imidazole, pH 7.6). The fractions containing the His$_6$-SUMO-MftA were pooled together and buffer exchanged into storage buffer (50 mM Hepes, 200 mM NaCl, 1 mM TCEP, pH 7.6) using a HiPrep 26/10 Desalting column (GE Life Sciences). The protein was concentrated using a 10 kDa spin columns (Millipore Amicon Ultra). Tobacco etch virus (TEV) protease was added to the protein and incubated at 4 ^oC overnight. The MftA/SUMO/TEV protease mixture was separated on reverse-phase HPLC using the program described above. Peaks corresponding to the heavy labeled peptide were collected, lyophilized and used for further analysis.

2.19 Kinetic Analysis of MftE

Kinetic analysis of MftE was carried out under single turnover conditions (Figure SI4). A 500 µL reaction was setup to contain 60 µM MftE reconstituted with Zn^{2+} and/or Fe^{2+}, in the reaction buffer described above, and 50 µM of MftA*. At interval times, 50 µL aliquots were extracted, filtered through a 0.2 µm spin columns and either flash frozen or immediately used directly in HPLC analysis. Frozen samples were rapidly thawed and immediately used in HPLC analysis. HPLC analysis of the reactions were carried out as described above. The MftA* peaks were integrated and plotted versus time. Rates of the single turnover reaction were extracted from single exponential decay fits (GraphPad Prism 7.0).

2.20 Generation of sequence similarity network for alpha-hydroxy acid oxidase family.

A sequence similarity network (SSN) of the alpha-hydroxy acid oxidase family was generated using the Enzyme Function Initiative Enzyme Similarity Tool (ESI-EST) (190). The Interpro identifier IPR012133 was used in Option B method and sequences were restricted to those of UniRef 90 seed sequences. The network was constrained using an E-value of 40, sequence lengths between 320-450 and sequence identities of 50%. The resulting network, containing 2,215 nodes, was colored using the Color SSN tool on the EFI-EST website. The network was visualized in Cytoscape4 and the mycofactocin system oxidoreductase MftD was colored to yellow.

2.21 Cloning, Expression and Purification of MftD from M. ulcerans Agy99

The *mul_0774* gene sequence encoding for MftD from *Mycobacterium ulcerans* Agy99 (Uniprot: 0PM50) was cloned into *pET28a-TEV* using *BamHI* and *XhoI* restriction sites. The sequence verified *MumftD/pET28a-TEV* plasmid was transformed into *E. coli* ArcticExpress (DE3) (Stratagene) for protein production. A 50 mL overnight culture of the *E. coli* ArcticExpress (DE3)-MumftD/pET28a-TEV was used to inoculate 4 L of the same medium containing 50 μg/mL of kanamycin and 10 μg/mL of tetracycline and supplemented with 100 μM riboflavin. The culture was incubated at 37 °C with shaking at 200 rpm until an $OD_{600} \approx 0.6$ was attained, at which point MftD production was induced by addition of IPTG to a final concentration of 1 mM. Following a 24 h induction at 13 °C with shaking at 200 rpm, the cells were harvested by centrifugation at 7000 rpm for 10 min and stored at −80 °C until lysis. All purification steps were carried out in a Coy Lab anaerobic chamber maintained under an atmosphere of 97% N_2 and 3% H_2 gases. The harvested cells were thawed and resuspended in 5 times pellet mass of lysis buffer containing 2X phosphate-buffered saline (PBS) buffer (pH 8.0), 30 mM imidazole and 2 mM dithionite (DTH). To the suspension was added 0.1 mg/g of lysozyme, 0.1 mg/g of DNase and 1% w/v CHAPS and stirred on ice for an additional 30 min.

Cells were disrupted by sonication for 5 min to complete the lysis. The resulting lysate was clarified by centrifugation at 13,000 rpm for 10 min and the supernatant was loaded onto a 5 mL HisTrap FF Ni-NTA column (GE Healthcare) pre-equilibrated with lysis buffer using an AKTA Start FPLC (GE Healthcare). The column was washed with 2X PBS buffer (pH 8.0) containing 50 mM imidazole, and the bound protein was eluted using 2X PBS buffer (pH 8.0) containing 300 mM imidazole and 2 mM DTH. Fractions containing MftD proteins were immediately buffer exchanged into 2× PBS storage buffer (pH 8.0) containing 10% glycerol and 2 mM DTH using PD-10 columns (GE Healthcare). The resulting volume containing MftD protein was concentrated using 30 kDa spin concentrators (Millipore). Tobacco etch virus (TEV) protease was added to the protein and incubated a room temperature for 3 h to hydrolyze the N-terminal His-tag. The protein/protease mixture was loaded onto a 5 mL HisTrap FF Ni-NTA column pre-equilibrated with storage buffer. The flow through was collected, and the bound protein was eluted with 2X PBS (pH 8.0) containing 50 mM imidazole and 2 mM DTH. Fractions containing MftD proteins were pooled together and buffer exchanged into a fresh 2X PBS storage buffer (pH 8.0) containing 10% glycerol using PD-10 columns (GE Healthcare). The resulting protein fraction was concentrated using 30 kDa spin concentrators and subjected to a final size exclusion chromatography purification step as described above.

2.22 Cloning, Expression, and Purification of MftD from Mycobacterium smegmatis mc^{2}155.

The *msmeg_1424* gene encoding for MftD from *Mycobacterium smegmatis* mc^{2}155 (Uniprot: A0QSB) was cloned into pET28a vector using *NdeI* and *XhoI* restriction sites. The sequence verified *msmeg_1424/pET28a* plasmid was transformed into *E. coli* ArcticExpress (DE3) for protein expression. *Ms*MftD proteins were expressed and purified as described above for *Mu*MftD with thrombin being substituted for TEV.

2.23 Flavin Determination in MftD.

To determine the type of flavin contained in MftD, 100 μL of 200 μM of the purified protein was first denatured by heating at 80 °C for 10 min. The denatured protein was centrifuged at 14000 rpm for 10 min to remove all precipitated proteins. The supernatant was pipetted into a 1.5 mL HPLC autosampler vial and analyzed by reverse-phase chromatography on a Shimadzu Prominence-*i* LC-2030C HPLC using a Jupiter C 18, 5 μm, 4.6 X 250 mm column (Phenomenex) and 10 mM sodium phosphate monobasic, pH 5.5 (buffer A), and 90% methanol (buffer B) as the mobile phase. A linear gradient of solvent B from 0% to 35% was applied from 2 to 7 min, which was followed by another linear gradient to 95% solvent B from 7 to 15 min. A third gradient of solvent B from 95% to 100% was applied from 15 to 17 min, after which solvent B was held constant at 100% from 17 to 19 min. This was followed by linear decrease to 100% solvent A from 19 to 22 min. Solvent A was held constant for an additional 2 min to re-equilibrate the column before subsequent sample injections. Chromatograms were reported at 450 nm while monitoring wavelengths between 200 and 600 nm. Authentic samples of riboflavin, flavin mononucleotide (FMN), and flavin adenine dinucleotide were run as controls.

2.24 Accessing the lability of MftD bound flavin mononucleotide (FMN).

HPLC analytical size exclusion chromatography was run on MftD protein under both aerobic and anaerobic conditions to assess the stability of MftD bound FMN. For the anaerobic assay, 100 µL of 200 µM purified protein containing 3 mM DTH was centrifuged at 14000 rpm for 5 min in an anaerobic chamber. The supernatant was transferred into a 1.5 mL autosampler HPLC vial anaerobically and analyzed by reverse-phase chromatography on a Shimadzu Prominence-*i* LC-2030C HPLC using a BioSep-SEC-S 2000, 5 µm, 7.8 x 300 mm column (Phenomenex) and anaerobic PBS buffer pH 8.0 (Buffer A) as the mobile phase. Isocratic elution was applied, and chromatograms were reported at 280 nm for MftD and 450 nm for FMN while monitoring absorbance between 200 and 600 nm. The same procedure was repeated for the aerobic assay except MftD was oxidized by exposing it air at room temperature for 1 h and this was confirmed by a color change from colorless to deep yellow of the protein and an increase in 450 nm absorbance of UV spectrum. The mobile phase (PBS pH 8.0) was made aerobic by purging with air for 30 min. Biorad Gel Filtration Standards containing protein standards for 670 kDa (thyroglobulin), 158 kDa (γ-globulin), 44 kDa (ovalbumin), 17 kDa (myoglobin) and 1.35 kDa (vitamin B-12) were used to prepare a standard curve.

2.25 AHDP Modification Reactions

AHDP modification reactions were conducted in volumes of 200 μL under aerobic and anaerobic conditions and contained 50 mM sodium phosphate (pH 8.0) as the reaction buffer. Consecutively, 100 μM MftD, 200 μM FMN, and 100 μM AHDP were added, and the reactions were incubated under aerobic and anaerobic conditions for 12 h. Control reactions were set up in a similar manner except buffer was substituted for individual reagents, protein, or substrate. Reactions were subsequently filtered through 0.2 μm spin columns to remove all precipitated proteins. The supernatant was pipetted into 300 μL autosampler HPLC vials and injected directly into a reverse-phase HPLC and analyzed as described above. New peaks were isolated, lyophilized, and analyzed by NMR and HRMS.

2.26 ***Oxygen Requirement for MftD Reaction with AHDP***

To investigate the dependence of MftD catalysis on oxygen, a total of 200 μL of MftD reaction was set up under anaerobic conditions as follows. To the reaction buffer containing 2 × PBS (pH 8.0) were added consecutively 10 mM glucose and 10 μM glucose oxidase followed by incubation at room temperature for 10 min to deplete oxygen from the buffer. Sequentially, 100 μM MftD, 200 μM FMN, and 200 μM were added to the reaction followed by an additional 12 h incubation period at room temperature. Reactions were filtered through 0.2 μm spin columns to remove all precipitated proteins. The supernatant was pipetted into 300 μL autosampler HPLC vials and injected directly into a reverse-phase HPLC system. HPLC analysis was carried out as described above

2.27 *Assessing for MftD activity with NAD$^+$*

To assess if NAD$^+$ could serve as the final electron acceptor in MftD reaction with AHDP, a total of 200 μL reaction was set up anaerobically as follows. To a reaction buffer containing 2 x PBS (pH 8.0) was added consecutively10 mM glucose and 10 μM of glucose oxidase (GOX) followed by incubation at room temperature for 10 min to deplete oxygen from the buffer. Subsequently, 100 μM of MftD, 200 μM of FMN, 200 μM of AHDP and 1 mM NAD+ were added stepwise to the reaction followed by an additional 12 h of incubation period at room temperature. Reactions were filtered through 0.2 μm spin columns to remove all precipitated proteins. The supernatant was pipetted into 300 μL autosampler HPLC vials and injected directly onto a reverse-phase HPLC system. HPLC analysis was carried out as described above.

2.28 ***^{18}O Isotope Labeling of PMFT***

To determine the source of oxygen atom incorporated into the product of MftD catalysis, reactions were carried out in either $^{18}OH_2$ or $^{18}O_2$. For reactions carried out in $^{18}OH_2$, a buffer solution of 25 mM ammonium acetate was first prepared by lyophilizing 500 µL of the buffer solution to powder and resuspending into an equal volume of $^{18}OH_2$ (PET grade, Sigma-Aldrich). Reactions consisted of 90 µL of $^{18}OH_2$ buffer, 100 µM MftD (5 µL), 200 µM FMN (1 µL), and 100 µM AHDP (4 µL). Following a 1 h aerobic incubation at room temperature, the reaction was analyzed by HRMS by direct infusion onto a Thermo Scientific Q Exactive. Reactions carried out in the presence of $^{18}O_2$ proceeded similarly but were performed in $^{16}OH_2$ buffer that was bubbled with $^{18}O_2$ for 5 min prior to the addition of MftD. A control reaction was performed to monitor the back-exchange of labeled oxygen. To do so, ^{16}O-PMFT was incubated in $^{18}OH_2$ buffer for 60 min prior to analysis by HRMS.

2.29 ***Electrochemical Characterization of PMFT***

To determine the redox potential of PMFT, a 5 mg/mL suspension of functionalized SWCNTs was made by dispersing 50 mg of the nanotubes in 10 mL of dimethylformamide and sonicating in an ultrasonic bath for 1 h to obtain a uniform suspension. Glassy carbon electrodes were polished with 0.05 µm alumina on nylon and microbroth polishing pads. The electrode was sonicated in 10 mL of water in an ultrasonic bath for 5 min followed by 10 mL of ethanol for an additional 5 min and allowed to thoroughly dry at room temperature. The dried electrode was coated with 20 μL of the SWCNTs suspension and allowed to dry at room temperature. This was repeated two more times allowing for thorough drying in between casts. The dried SWCNTs-modified glassy carbon electrode was incubated in phosphate buffer (pH 7.0) as control or in phosphate buffer (pH 7.0) containing 500 μM PMFT for 3 h prior to running cyclic voltammetry (CV) experiments. A conventional three-electrode system was used in a water-jacketed glass cell. The counter electrode was platinum wire, and Ag/AgCl was the reference electrode. CV measurements were carried out in a Coy Lab anaerobic chamber using a CHI 600E Series electrochemical analyzer. Measurements were recorded at 0.05 V/s with a potential sweep window of +0.2 to −0.8 V at 22 °C.

2.30 *Cloning, Expression, and Purification of Carveol Dehydrogenase (CDH) from Mycobacterium smegmatis mc^{2}155*

The *msmeg_1410* gene sequence encoding CDH from *M. smegmatis* mc^{2}155 (Uniport: A0QSA5) was cloned into pET28a using *NdeI* and *HindIII* restriction sites. The sequence verified *msmeg_1410/pET28a* plasmid was transformed into the *E. coli* BL21 star (DE3) and grown overnight at 37 °C in 50 mL of LB containing 50 μg/mL of kanamycin. The overnight culture was used to inoculate 4 L of LB growth medium containing 50 μg/mL of kanamycin. The culture was incubated at 37 °C with shaking at 200 rpm until an OD_{600} of ~0.6, at which point *Ms*CDH production was induced by addition of IPTG to a final concentration of 1 mM and allowed to grow overnight at 20 °C with shaking at 200 rpm. The cells were harvested by centrifugation at 7000 rpm for 10 min, and the pellet was resuspended in 5 times pellet mass of lysis 50 mM MOPS buffer (pH 7.0) containing, 250 mM NaCl and 30 mM imidazole. To the suspension were added 0.1 mg/g of lysozyme, 0.1 mg/g of DNase, and 1% w/v 3-[(3-cholamidopropyl) dimethylammonio]-1-propanesulfonate (CHAPS), and the mixture was stirred for 30 min. Cells were disrupted by sonication for 5 min on ice to complete the lysis. The resulting lysate was clarified by centrifugation at 13000 rpm for 10 min, and the supernatant was loaded onto a 5 mL HisTrap HP Ni-NTA column pre-equilibrated with lysis buffer using an AKTA Pure FPLC. The column was washed with 50 mM MOPS buffer (pH 7.0) containing 250 mM NaCl and 30 mM imidazole, and the bound protein was eluted using 50 mM MOPS buffer (pH 7.0) containing 250 mM NaCl and 300 mM imidazole. Fractions containing *Ms*CDH proteins were pooled together and immediately buffer exchanged into 50 mM MOPS storage buffer (pH 7.0) containing 250 mM NaCl and 10% glycerol using a HiPrep 26/10 desalting column. The resulting volume of protein was concentrated using 10 kDa spin concentrators. Thrombin protease was added to the concentrated protein and incubated at room temperature for 1 h to cleave off the N-terminal His-tag. The protein/protease mixture was loaded onto a Superdex 200 10/300 GL analytical size exclusion chromatographic column using an AKTA Pure FPLC. The mobile phase was storage buffer containing 50 mM MOPS (pH 7.0), 250 mM NaCl, and 10% glycerol at a flow rate of 1.0 mL/min. Protein peaks were collected and concentrated using 10 kDa spin concentrators. The concentrated protein was aliquoted into 1.5 mL vials, flash-frozen in liquid nitrogen, and stored at −80 °C until use.

2.31 Monitoring MsCDH UV Vis Spectral Changes in the Presence of PMFT

To determine if PMFT could be used as a cofactor by *Ms*CDH for carveol modification, a UV spectroscopic assay was performed using Shimadzu TCC-240A UV−visible spectroscopy as follows. An initial UV absorbance spectrum of 200 μL reaction buffer containing 50 mM MOPS (pH 7.0), 250 mM NaCl, 10% glycerol, and 150 μM *Ms*CDH was taken from 190 to 500 nm. Following the initial scan, 30 μM PMFT was added to *Ms*CDH and incubated for 5 min, and the UV−visible spectra were measured. This was repeated until a final concentration of 150 μM PMFT was achieved. A control experiment was carried out using the same procedures but with NADH free in solution.

2.32 MsCDH-Carveol Modification Reactions.

Carveol dehydrogenase modification reactions were conducted in a total volume of 200 μL containing 50 mM MOPS buffer (pH 7.0), 250 μM NaCl, 10% glycerol, 100 μM *Ms*CDH, 500 μM carveol, and 500 μM PMFT. The reaction was incubated at room temperature for one hour. Control reactions were set up in a similar fashion except buffer was substituted for individual reagents, protein, or substrate. Reactions were filtered through 3 kDa spin columns to remove all precipitated proteins. The supernatant was pipetted into 300 μL autosampler HPLC vials and injected directly onto a reverse-phase HPLC system. The HPLC analysis was as described above and UV absorbance readings were monitored between 200 and 400 nm.

2.33 Stopped Flow Spectrophotometry.

A DX.17MV sequential stopped-flow spectrometer (Applied Photophysics, Leatherhead, UK) with a 3 ms deadtime was used to measure the single turnover oxidation experiments of *Ms*CDH by PMFT and DCPIP. All reactions were carried out in 50 mM Hepes, pH 7.0, 250 mM NaCl, and 10% glycerol. For reactions with PMFT, the absorbance of the *Ms*CDH bound NADH was monitored at 340 nm (Δε NADH 340 nm = 6.22 mM^{-1} cm^{-1}). For reactions with DCPIP, the absorbance of the oxidant was monitored at 600 nm (Δε DCPIP 600 nm = 20.7 mM^{-1} cm^{-1}). Reactions were set up so that the final concentrations were 160 μM MsCDH and 25 or 50 μM PMFT, or 22 or 60 μM DCPIP. Reactions were run in triplicate at each concentration, averaged, and fitted to a single exponential decay or linear equations using the accompanying software.

CHAPTER THREE: RESULTS

3.1 Identification of Conserved Cysteines in MftC

To identify possible cysteine residues involved in the coordination of [Fe-S] clusters in MftC, a bioinformatic analysis was carried out on all probable MftC homologues obtained from a BLAST query of the Uniprot database. From the BLAST, a total of 625 putative MftC sequences were retrieved and aligned, and 11 cysteines were found to be strictly conserved. Based on their sequence-wise location and residue spacing, Cys30, 34, and 37 were assigned as the RS ligands (characteristic $CX_3CX\phi C$ sequence motif). In addition to the RS sequence motif, residues Cys251, 258, 269, 310, 313, 319, 323, and 341 were also found to be strictly conserved among all 625 species (characteristic $CX_{9\text{-}15}GXC$-gap-$CX_2CX_5CX_3C$-gap-C motif of the SPASM domain) (Figure 3.1). These last eight cysteine residues are expected to reside in the C-terminal SPASM domain and could bind one or two auxiliary [Fe−S] clusters. The conserved cysteine residues were then mapped onto a homology model of MftC that was generated on the basis of the structures for anSME (PDB ID: 4K36) (117), CteB (PDB ID: 5WHY) (83) or SuiB (PDB ID: 5V1T) (110). In the representative overlay between the modeled MftC and the parent anSME (sequence identity ~25%) shown in Figure 3.2A, the predicted RS binding residues Cys 30, 34, and 37 align accordingly with the corresponding RS binding motif of anSME. Similarly, the modeled MftC residues Cys 310, 313, 319, and 341 align well with those from anSME involved in ligating the Aux II cluster. Of the remaining four modeled MftC Cys residues, only residues 269 and 323 overlay well with the Cys residues of anSME involved in binding the Aux I cluster. Residue Cys 251 is modeled in the vicinity of the Aux I cluster of anSME, but the modeled residue Cys 258 is located in the second shell around the Aux I cluster of anSME (Figure 3.2A). All models generated for MftC provided similar findings (Figure 3.2 B & C, sequence identities ~20%). The poor overlay of two of the four Cys residues used to bind Aux I is likely due to the low sequence similarity, and neither exclude nor confirm binding of a cluster in the Aux I position in MftC.

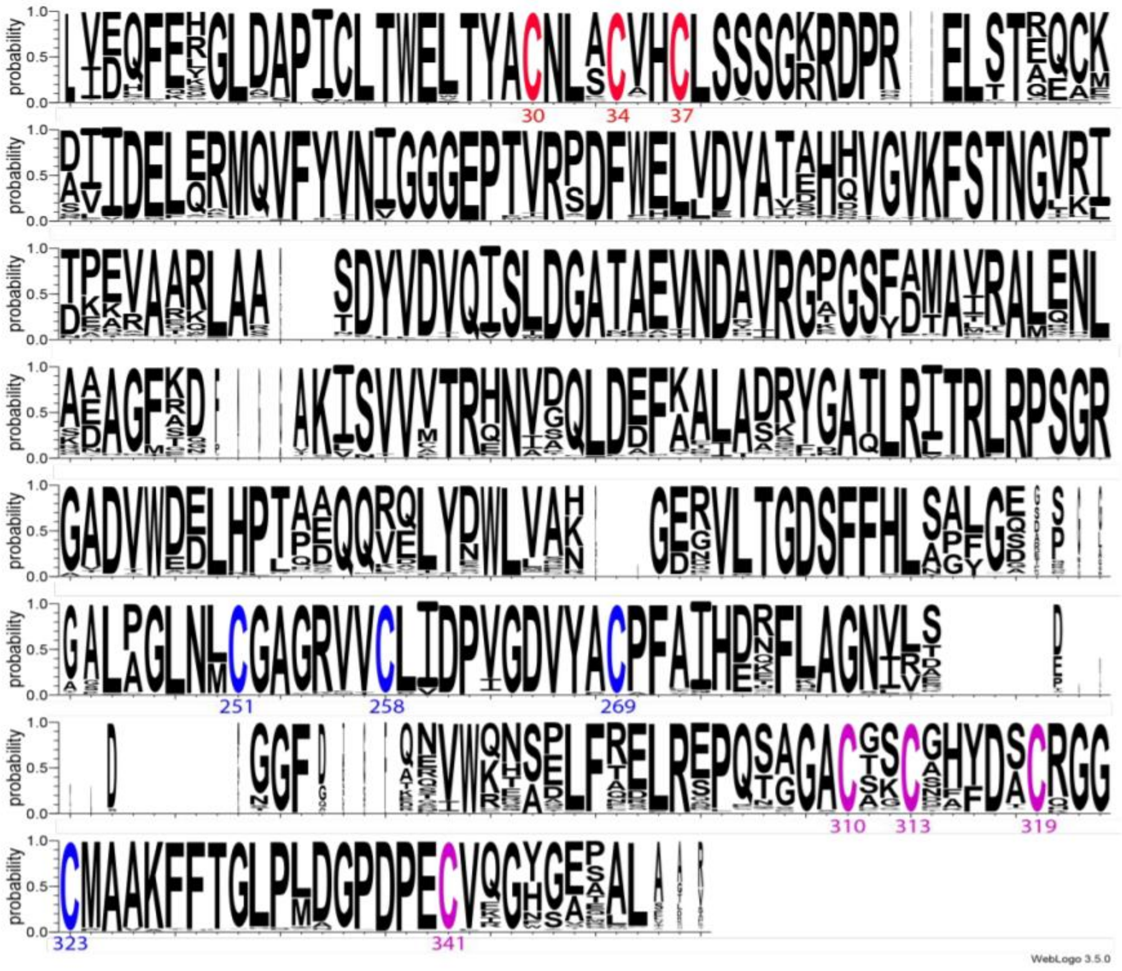
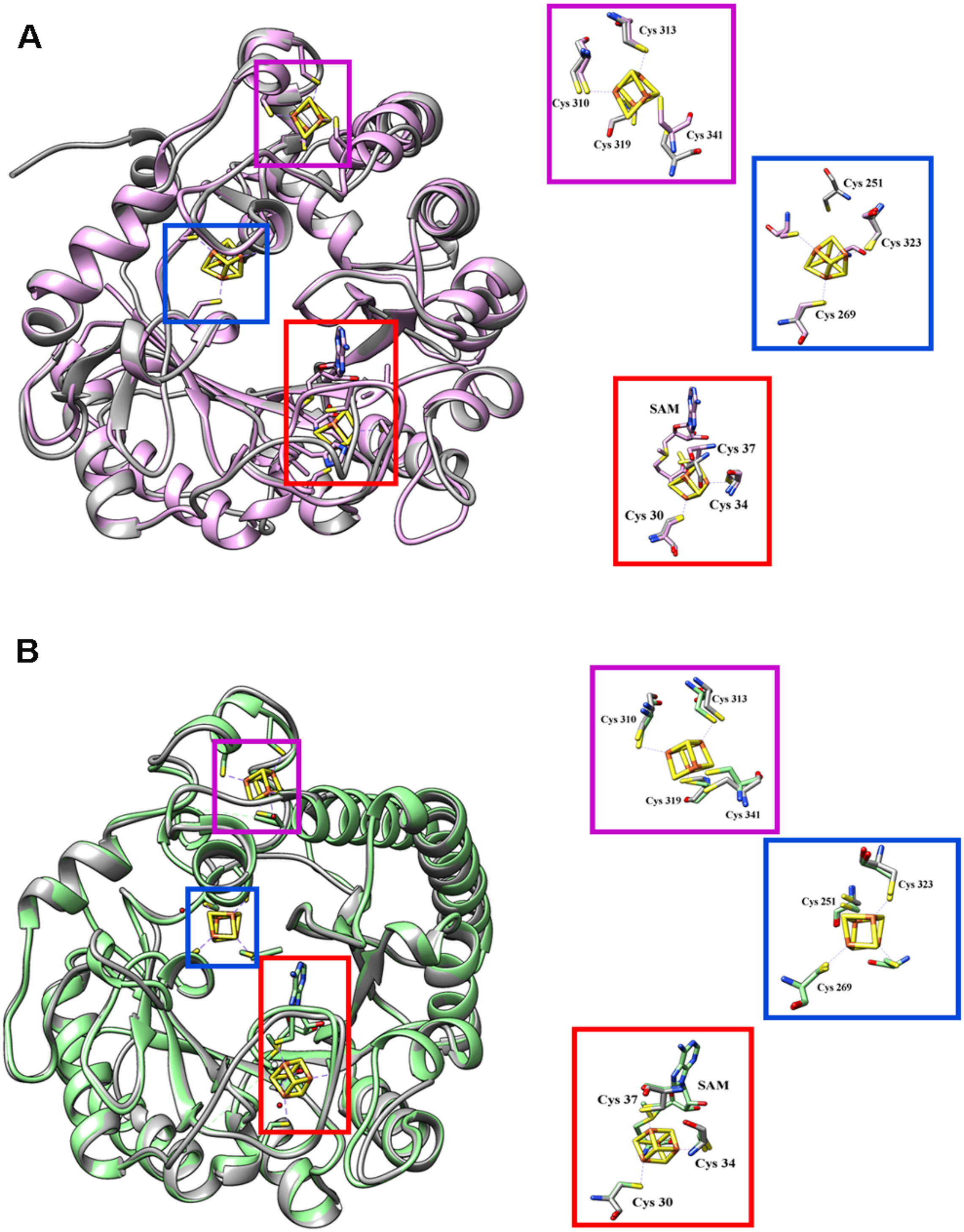


Figure 3.1 Multiple sequence alignment of MftC visualized by Weblogo. The probability of a residue in a position is indicated on the y-axis. Cysteine residues found to be strictly conserved were colored by predicted cluster binding (red – RS cluster, blue – Aux I cluster, and magenta – Aux II cluster) and numbered according to the *M. ulcerans* MftC amino acid sequence.


A
Cys 313
Cys 310
Cys 319
Cys 341
Cys 251
Cys 323
Cys 269
SAM
Cys 37
Cys 34
Cys 30
B
Cys 310
Cys 313
Cys 319
Cys 341
Cys 251
Cys 323
Cys 269
Cys 37
SAM
Cys 34
Cys 30

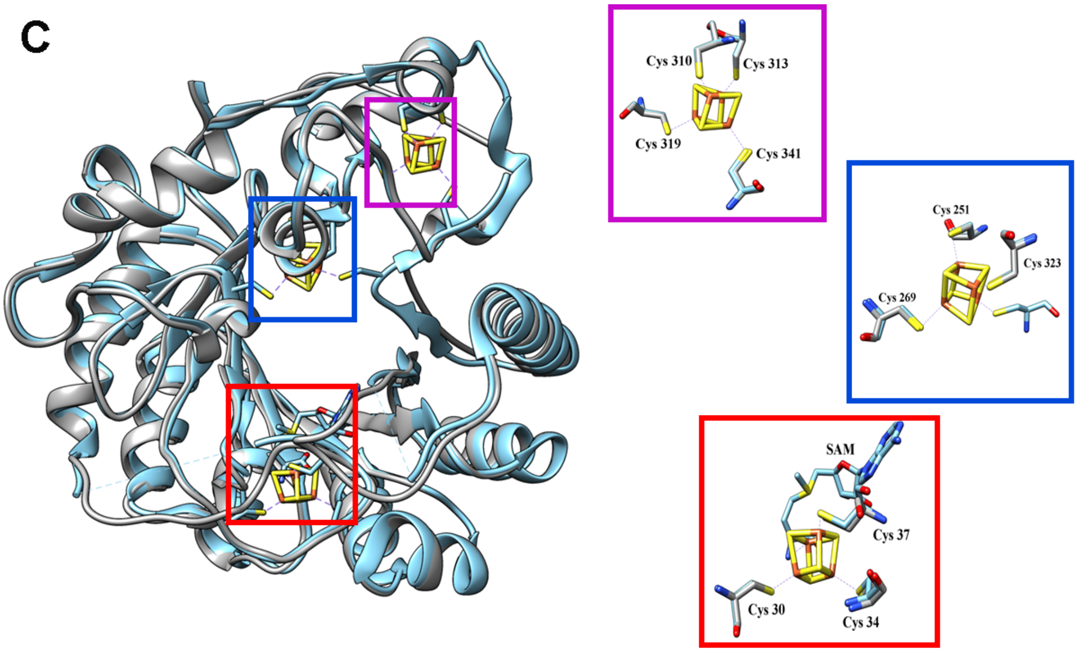


Figure 3.2 The model of MftC (grey), generated by Phyre 2.0. MftC was modeled on the structure of (A) AnSME (PDBID: 4K36), (B) CteB (PDBID: 5WHY) and (C) the SuiB (PDBID: 5V1T) using one-to-one threading on phyre2. when overlain with the same CteB (green) or SuiB (cyan) structure implies that three [4Fe-4S] could be bound. The blowout figures show the cysteine residues around the RS cluster (red box), the Aux I cluster (blue box), and the Aux II cluster (magenta box).

3.2 Generation of Soluble MftC Single- and Multiple Point Cysteine Variants.

To gain insights into the role of the auxiliary clusters in MftC catalysis and to identify the Cys residues involved in binding of the clusters, Ala variants of MftC were constructed. Using the bioinformatic information described above, Cys residues were systematically replaced with Ala to generate soluble RS KO, Aux I KO, and Aux II KO variants. For the RS KO variant, Cys30 and Cys37 were mutated to generate the C30A/C37A double mutant. The triple mutant C30A/C34A/C37A was insoluble. For the Aux I KO variant, only the single mutants C269A and C323A were soluble. Both the C251A single mutant and the double C269A/C323A mutant resulted in insoluble protein. The C258A variant did not impact protein solubility or MftC activity and was not considered an important mutation. Since the C269A and C323A variants were found to have identical enzymatic activity and since C323A could be produced in larger quantities, the C323A variant was used as in all experiments as the Aux I KO. To make the Aux II KO, Cys310 and Cys341 were mutated to generate the C310A/C341A double mutant. Additional mutations of the Aux II site led to insoluble protein. The purity of the MftC variants were checked by SDS-PAGE and size exclusion chromatography (SEC) (Figure 3.3). For all variants of MftC, the proteins were purified to near homogeneity as determined by SDS-PAGE (Figure 3.3A). In addition, the SEC profile of MftC variants showed a major peak at a retention time consistent with ~42 kDa (Figure 3.3B). This suggests that the RS KO, Aux I KO, and Aux II KO had a minimal effect on the oligomeric state of MftC.

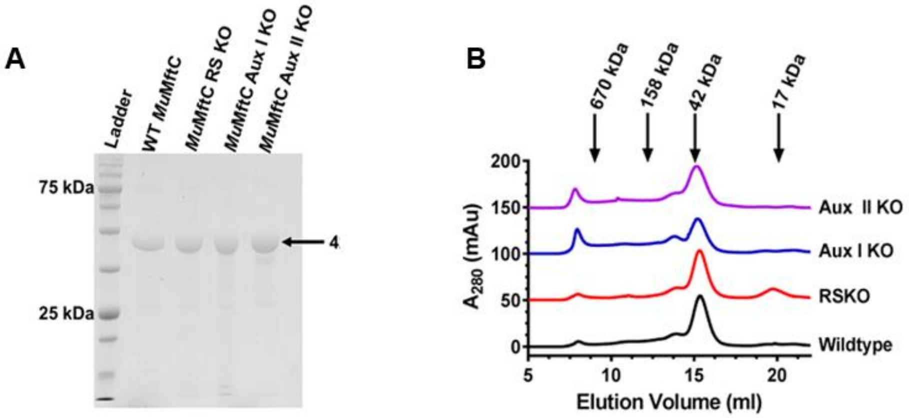


Figure 3.3 Characterization of purified MuMftC proteins. (A) An SDS-PAGE gel containing a protein ladder (Gold Bio Blue Stain Plus) shows that all MftC variants have been purified to near homogeneity and that they are all approximately 42 kDa. (B) The chromatogram of size exclusion chromatography experiments monitoring protein absorbance at 280 nm shows that MftC (black), RS KO (red), Aux I KO (blue), and the Aux II KO (violet) are predominantly in the monomeric state. Protein standard elution volumes for thyroglobulin (670 kDa), γ-globulin (158 kDa), oval albumin (42 kDa) and myoglobin (17 kDa) are indicated by black arrows.

3.3 MftC Contains Three [4Fe−4S] Clusters.

UV-vis spectroscopic and iron-sulfur quantification assays from previous experiments suggested that MftC contains [Fe−S] clusters; however, the exact number and configuration (e.g., [4Fe− 4S] vs [2Fe−2S]) of the clusters have not been established (97, 98). To address this, Mössbauer spectroscopic experiments were carried out of MftC and MftC variants to gain independent additional insights into the chemical nature and number of [Fe−S] clusters. The 4.2 K/78 mT Mössbauer spectrum of the as-isolated wild-type MftC is shown in Figure 3.4 and can be best analyzed considering two quadrupole doublets, which account for 88% of the total iron in the sample. The dark green shaded doublet has an isomer shift of δ = 0.26 mm/s and quadrupole splitting of ΔEQ = 0.55 mm/s, parameters that are characteristic of all-ferric [2Fe−2S] clusters in their 2^+ state. The light-green-shaded doublet has an average isomer shift of δ = 0.45 mm/s and an average quadrupole splitting of ΔEQ = 1.01 mm/s, parameters that are characteristic of valence delocalized $Fe^{2+}-Fe^{3+}$ cysteine-coordinated pairs in oxidized $[4Fe-4S]^{2+}$ clusters. The mismatches between the summations of the theoretical spectra with the experimental data reveal the presence of a minor and poorly defined sub spectrum (∼10% intensity), which we attribute to nonspecifically bound iron. Considering the number of Fe ions per cluster and those obtained from elemental analysis the as-isolated MftC sample contains 0.67 $[2Fe-2S]^{2+}$ clusters and 0.61 $[4Fe-4S]^{2+}$ clusters per polypeptide, respectively.

Since the as-purified protein already contains an appreciable number of [Fe−S] clusters, the protein employed for Mössbauer analysis to probe the number of clusters was chemically reconstituted with 2 equivalence of [Fe−S] centers (see Methods 2.7). The 4.2 K/78 mT Mössbauer spectrum of this sample is shown in Figure 3.4 (bottom) and can be best analyzed considering a single quadrupole doublet that accounts for 92% of the total iron in the sample. The light green shaded quadrupole doublet has an isomer shift of δ = 0.45 mm/s and a quadrupole splitting of ΔEQ = 1.11 mm/s, parameters consistent with the presence of oxidized $[4Fe- 4S]^{2+}$ clusters. The remaining small broad absorption in the spectrum accounts for ∼8% and corresponds to nonspecifically bound iron. There are no signals detectable for $[2Fe-2S]^{2+}$ clusters, suggestive that their presence in the as-purified MftC could be an oxidative byproduct. After correction for the number of Fe ions per cluster and elemental analysis, the sample was found to contain 2.7 ± 0.15 [4Fe−4S]$^{2+}$ clusters per polypeptide confirming that that MftC contains three [4Fe−4S]. This finding was further corroborated by data from the quantification of iron and sulfur in the chemically reconstituted wildtype MftC protein (Table 1). Taken together, these results provide direct evidence that MftC contains thee [4Fe-4S] clusters, however more information regarding the crystal structure of MftC is required to determine the coordination state (open versus full ligation) of the two Auxiliary clusters.

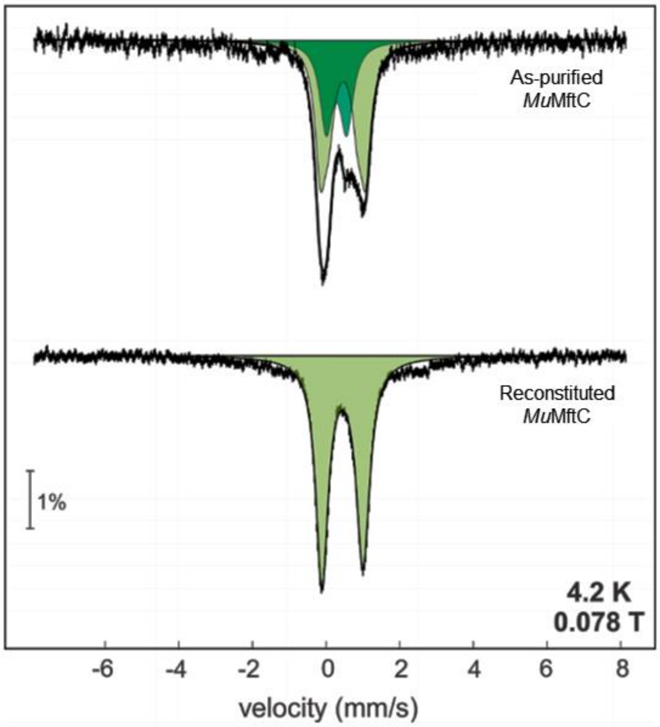


Figure 3.4 Mössbauer spectra of the as-purified and chemically reconstituted wildtype MuMftC at 4.2 K. (Top) The spectrum of the as-isolated MftC (black) consists of a quadrupole doublet corresponding to the presence of [2Fe−2S] clusters (dark green highlight) and [4Fe−4S] clusters (light green highlight). The overall simulation is overlaid on the experimental spectrum as a black solid line. (Bottom) The spectrum of the chemically reconstituted with 2 equivalence of [Fe−S] clusters MftC (black) consists of a single quadrupole doublet corresponding to the presence of [4Fe−4S] clusters. The spectra were recorded in the presence of a small external magnetic field (0.078 T), applied parallel to the orientation of the γ beam.

	WT MftC	MftC RS KO	MftC Aux I KO	MftC Aux II KO
Iron	10.0 ± 0.2	7.4 ± 1	8.8 ± 0.2	8.2 ± 0.1
Sulfide	8.1 ± 0.1	7.9 ± 0.1	10.0 ± 0.4	7.7 ± 0.1

Table 1 Iron and sulfur quantification for reconstituted MftC proteins.

3.4 All Three [Fe−S] Clusters Are Required for MftC Modification of MftA

Having demonstrated that MftC contains three [4Fe-4S] clusters, we next set out to investigate the role of the Aux I and Aux II clusters in MftC catalysis. This was done by carrying out activity assays for the RS KO, Aux I KO, and Aux II KO mutants. Reactions containing the appropriate components (see the Methods and Materials) were analyzed by reverse phase HPLC, monitoring the absorbance at 260 nm (SAM cleavage) or 280 nm (peptide modification). As shown in Figure 3.5, Wildtype MftC catalyzed the reductive cleavage of SAM to form dAdo and further converted the substrate MftA to MftA*. Conversely, the RS KO could neither cleave SAM nor modify MftA, consistent with the successful knockout of the RS cluster. Activity assays for Aux I and Aux II KOs also provided insightful results. Both Aux I and Aux II KOs were capable of catalyzing the reductive cleavage of SAM to form dAdo (Figure 3.5A), suggesting that the RS cluster remained intact and in an active conformation in the mutated proteins. However, when assayed against MftA, both Aux I and Aux II KOs were incapable of converting MftA to MftA* or MftA** (Figure

3.5B). This suggests that the Aux I and Aux II are required for catalysis, consistent with results shown for QhpD, AnSME, SCIFF, and PqqE (74, 110, 158, 191).

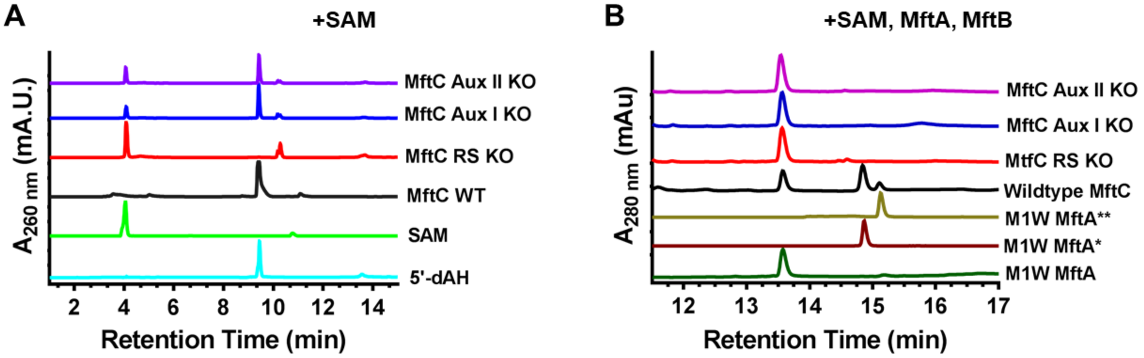


Figure 3.5 Activity assays of wildtype and mutant MuMftC. Reactions consisting of (A) 2 mM SAM, 10 mM DTT, 2 mM DTH, and 100 μM of the MftC variant or (B) 2 mM SAM, 10 mM DTT, 2 mM DTH, 100 μM MftA, 100 μM MftB, and 100 μM of the MftC variant suggest that Aux I and Aux II clusters are not required for SAM turnover and that all three clusters are required for MftA modification.

3.5 Protein Film Electrochemistry Measured the Midpoint Potentials for All Clusters in MftC.

MftC catalyzes both oxidative and redox neutral reactions in the same active site. This unique ability of MftC suggests that the Auxiliary clusters could partake in the reaction through an electron transfer mechanism. To determine if the redox potentials of the Auxiliary clusters could accommodate the reversible electron transfer between the clusters, protein-film electrochemistry was carried out on MftC and its variants. Wild-type MftC was noncovalently adsorbed onto a plan-edge graphite electrode using multiwalled carbon nanotubes as a co-adsorbent. A single and broad feature in the cyclic voltammogram for wild-type

MftC was observed between −425 and −600 mV (Figure 3.6A). This suggests that the three [4Fe−4S] clusters have similar reduction potentials that appear as a single envelope when summed up. The envelope signal was fitted using QSoas software, and three, one electron transfers with reduction potentials of ~ −550, −500, and −460 mV vs SHE were extracted (Table 2) (192). The three reduction potentials represent three redox active species and are expected to represent each of the three [4Fe−4S] clusters found on MftC. To resolve which reduction potential belonged to which cluster, the same analysis was conducted on the RS KO and Aux II KO. In experiments containing the RS KO, fitted data suggest that two redox active species, with reduction potentials of −507 and −502 mV vs SHE, are present (Figure 3.6B and Table 2) Notably, the highest potential of ~ −460 mV was no longer apparent, suggesting that this potential belongs to the RS cluster. It remained unclear which of the two remaining signals belong to the Aux I or Aux II cluster. To clarify this, the cyclic voltammogram for the Aux II KO variant was also acquired and fitted. The fitted data from this experiment indicated that two redox active species, with reduction potentials of −520 and −454 mV vs SHE, were present (Figure 3.6C and Table 2). The reappearance of the ~ −460 mV species further suggests that the midpoint potential for the RS cluster is indeed ~ −460 mV. Additionally, the remaining signal of −520 mV could be assigned to the Aux I cluster. It is likely that the −520 mV species found in the Aux II KO data set corresponds to the −550 mV species found in the wild-type MftC data set. Therefore, the resting state reduction potentials measured for MftC could be assigned as the following: ~ −460 mV, RS cluster; −550 mV, Aux I cluster; −500 mV, Aux II cluster. These potentials are similar to those found for the only other RS-SPASM protein measured to date, SCIFF maturase. SCIFF maturase catalyzes the formation of intramolecular thioether bonds on its precursor peptide and also contains three [4Fe− 4S] clusters (74). Similar to MftC, the reduction potentials measured for the three [4Fe−4S] clusters on SCIFF maturase were −490, −540, and −585 mV vs SHE (193).

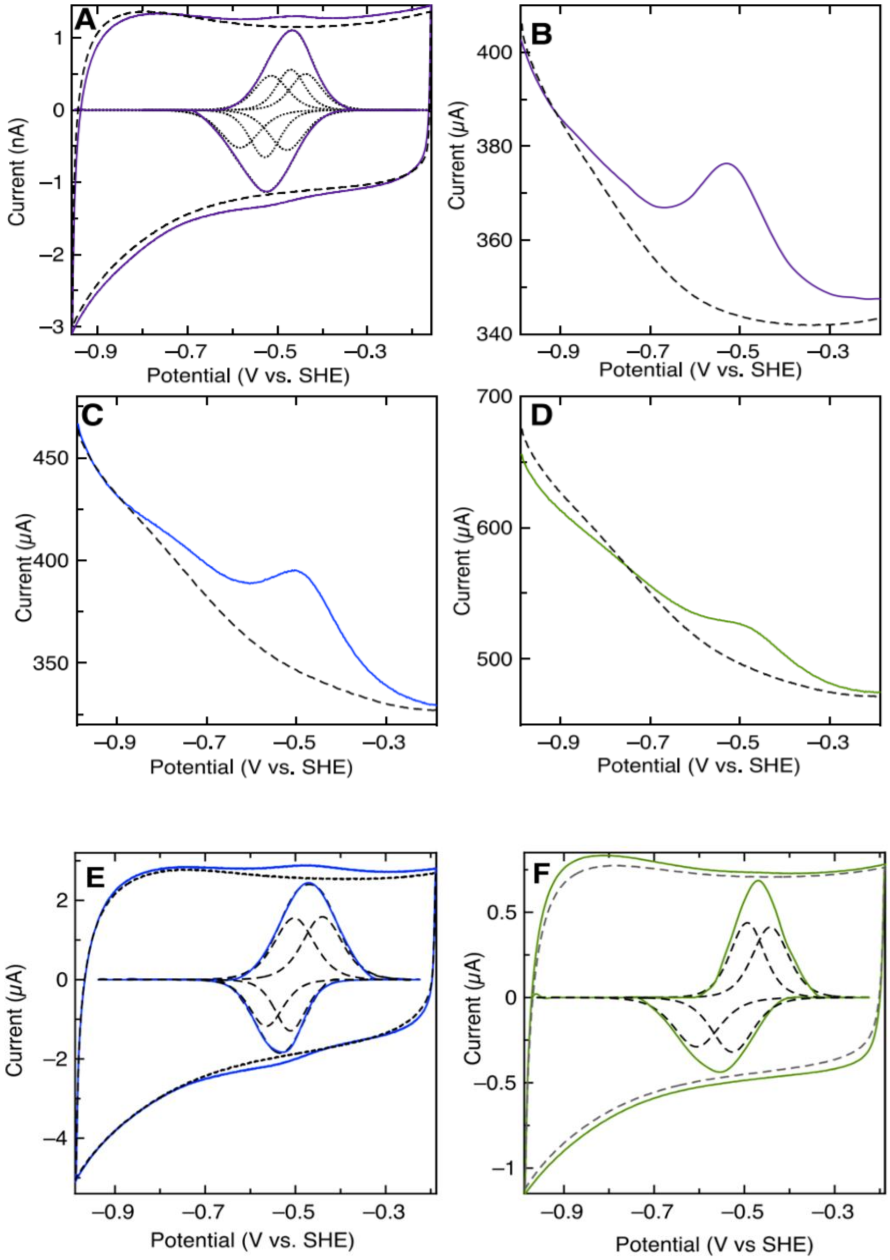


Figure 3.6 Voltammetry of MftC and variants measured at pH 7.5 and 4 °C. (A) Cyclic voltammogram measured with a scan rate of 50 mV/s for wild-type MftC (solid line), fitting for three, one electron transfers (dotted line), and EPG baseline (dashed line). Square wave voltammograms measured with a frequency of 10 Hz and an amplitude of 20 mV for (B) wild-type MftC (solid line), (C) RS KO, and (D) Aux II KO with EPG baselines (dashed line). Additional CV data of (E) the RS KO and (B) the Aux II KO constructs.

Variant	$E_{m,1}$ (mV vs SHE)	$E_{m,2}$ (mV vs SHE)	$E_{m,3}$ (mV vs SHE)	$E_{mid,}$ (mV vs SHE)
WT MftC	- 549	- 499	- 458	- 489
RS KO	- 507	- 502		- 484
Aux II KO	- 520		-454	- 475

Table 2 Reduction potentials and midpoint potentials for the RS, Aux I, and Aux II [4Fe−4S]$^{2+/1+}$ clusters in MftC

Taken together, these results suggest that MftC contains three [4Fe-4S] clusters, all of which are required for catalysis. In addition, the redox potentials of these clusters are within -100 mV apart, particularly the Auxiliary clusters which are within 50 mV apart. The similarity of the redox potentials of the auxiliary clusters of MftC could allow for the reversible electron transfer between the clusters, hence the ability of MftC to catalyze both oxidative and redox neutral reactions in the same active site. Based on this, a revised mechanism for MftC catalysis is proposed. This mechanism involves an electron being shuttled from the substrate into the Aux I cluster to form a benzenone intermediate. The Aux I cluster subsequently shuttles this electron into the Aux II cluster to complete the first step of MftC catalysis (Figure 3.7 black). In the second step, the electron is back into the reaction to quench the substrate radical to complete MftC catalysis resulting in the formation of MftA* (Figure 3.7 blue). This mechanism fits the ability of MftC to redox flip, however, there is only one caveat. That is, in the absence of the Aux II cluster, we should expect accumulation of the intermediate MftA** in MftC peptide modification assays, however, that is not the case (Figure 3.5B).

Instead, we observed aborted catalytic reaction of MftC, which could suggest that the Aux II cluster is playing some structural role in MftC and without it, the catalytically structural scaffold of MftC is distorted. It should be noted that without additional information regarding the crystal structure of MftC, this is nothing more than speculation.

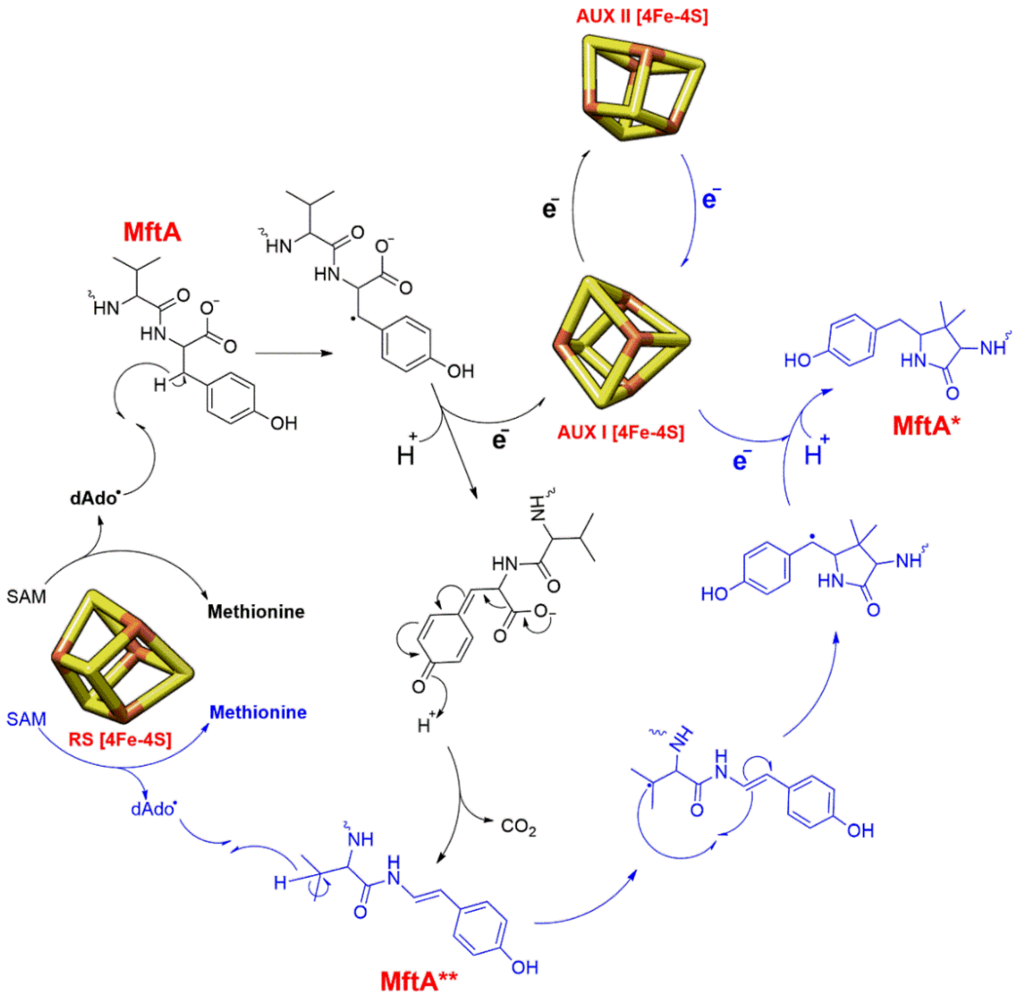


Figure 3.7 Revised mechanism for MftC catalysis. All three clusters in MftC are required for catalysis. An electron is shuttled from Aux I to Aux II during the decarboxylation step (black) and this electron is shuttled back to quench the substrate radical in the carbon-carbon crosslink step (blue).

3.6 *MftE is a member of the creatinine amidohydrolase family (IPR003785)*

MftE is a member of the creatinine amidohydrolase family of proteins (185). This is a large protein family consisting of more than 11000 protein sequences annotated as IPR003785. Creatinine amidohydrolases are widely distributed binuclear Zn^{2+} enzymes responsible for the hydrolysis of creatinine to form creatine (194, 195). Within this family are the FAPy deformylase and mycofactocin peptidase MftE subfamilies. Interestingly, both subfamilies are associated with cofactor biosynthesis. While MftE is associated with mycofactocin biosynthesis, FAPy deformylase on the other hand is involved in the Fe^{2+}-dependent deformylation of 2-amino-5- formylamino-6-ribosylamino-4(3H)-pyrimidinone 5′-monophosphate during the biosynthesis of the redox cofactors F_0 and F_{420} (196). To observe the relationship between the mycofactocin peptidase, MftE, and the remaining members of the amidohydrolase fold family, a sequence similarity network (SSN) was generated for the IPR003785 family using the Enzyme Function Initiative-Enzyme Similarity Tool (ESIEST) (190). The analysis, based on 6279 Uniref 90 sequences, generated a network consisting of 3432 nodes where each node represents a set of protein sequences that are 50% identical (Figure 3.8). Notably, the mycofactocin (green, 43 sequences) stand out from the central creatinine amidohydrolase clusters (1735 sequences). Indeed, the majority of sequences within the network belong to uncharacterized clusters that are independent of the creatinine amidohydrolase central clusters, suggesting that this family of proteins may have evolved to hydrolyze a variety of substrates.

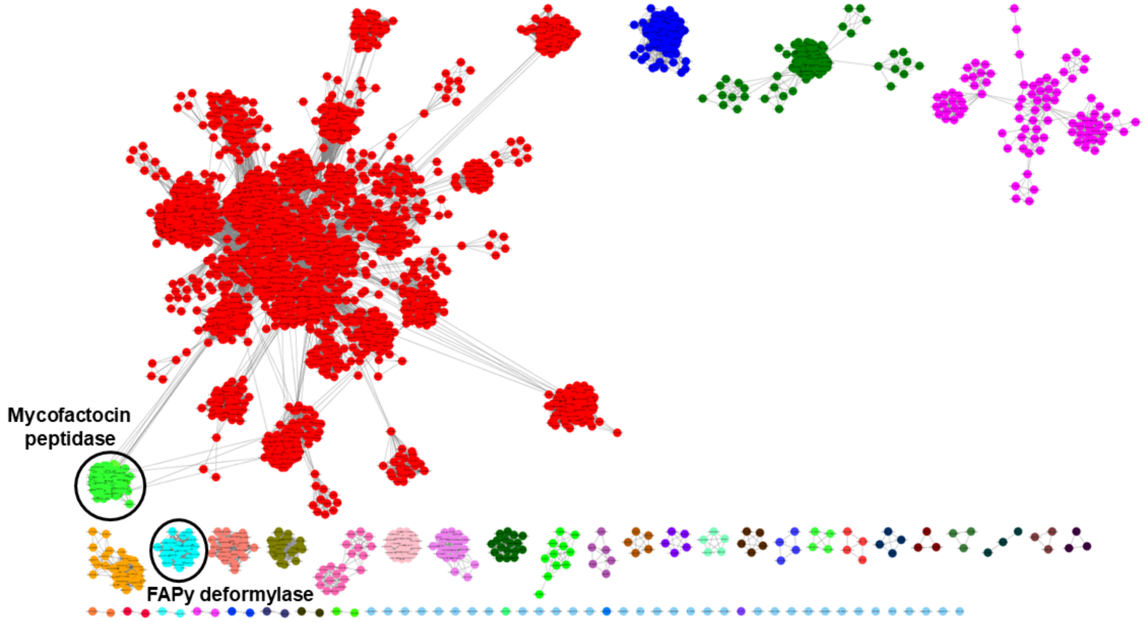


Figure 3.8 A sequence similarity network for MftE. A sequence similarity network for the creatinine amidohydrolase family (IPR003785) shows the divergence of the mycofactocin peptidase (green) and the FAPy deformylase (cyan) from the main body of creatinine amidohydrolases.

3.7 Characterization of purified MftE.

To investigate the activity of MftE in the biosynthesis of mycofactocin, the MftE protein from *M. ulcerans* was cloned into, overexpressed in, and isolated from *E. coli*. Recombinant MftE was purified to >95% purity, as determined by SDS-PAGE (Figure 3.9A). The purified protein showed high homogeneity as determined by size exclusion chromatography, with >95% of the total protein appearing at retention a time consistent with that of an octamer (Figure 3.9 C). As described above, MftE is a member of the metal-binding creatine amidohydrolase protein fold family and was previously shown to bind zinc and iron (185). However, efforts to fully reconstitute the protein with iron and zinc and their influence on catalysis of MftE were not reported. To address this, metal quantification assays of as-purified and reconstituted MftE were carried out. Consequently, ICP-MS analysis of the as-purified MftE revealed that it binds 0.51 ± 0.01 mol equivalents of Zn^{2+} and 0.10 ± 0.01 mol equivalents of Fe^{2+}, consistent with the previously reported values (185). Following reconstitution with Zn^{2+}, MftE was found to bind 0.57 mol equivalents of Zn while maintaining ~0.10 mol equivalents of Fe. Likewise, when Fe^{2+} was added, MftE was found to bind 0.70 ± 0.02 mol equivalents of Fe and maintained 0.15 ± 0.01 mol equivalents of Zn. Upon reconstitution with both Fe^{2+} and Zn^{2+}, MftE was found to bind 0.71 ± 0.01 mol equivalents of Fe and 0.55 ± 0.01 mol equivalents of Zn per polypeptide chain (Table 3), that is one atom of Fe^{2+} and one atom of Zn^{2+} per polypeptide chain, similar to that of FAPy deformylase protein ArpB (196).

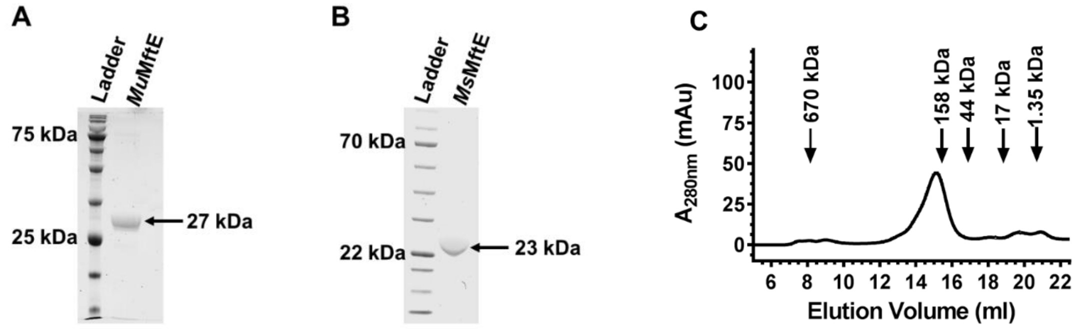


Figure 3.9 Characterization of MftE proteins. An SDS-PAGE gel of MftE from (A) *M. ulcerans* and (B) M. smegmatis shows that the protein is >95% pure and has a molecular weight of approximately 27 kDa. (C) The chromatogram from the size exclusion chromatography while monitoring absorbance at 280 nm shows that majority of the purified protein migrates as an octamer with a molecular weight of 230 kDa. Protein standard elution volumes for thyroglobulin (670 kDa), γ-globulin (158 kDa), ovalbumin (44 kDa), myoglobin (17 kDa) and vitamin B-12 (1.35 kDa) are indicated by the black arrows

Sample	Fe^{2+} eq	Zn^{2+} eq
As-purified MftE	0.10 ± 0.01	0.51 ± 0.01
MftE + Fe(NH$_3$)SO$_4$	0.70 ± 0.02	0.15 ± 0.01
MftE + ZnCl$_2$	0.08 ± 0.01	0.57 ± 0.01
MftE + Fe(NH$_3$)SO$_4$ and ZnCl$_2$	0.71 ± 0 .01	0.55 ± 0.01

Table 3 ICP-MS quantification of metal ions in MftE. Analysis of metal iron content of as-purified and reconstituted MftE shows that the protein copurifies with both iron and zinc.

*3.8 MftE is selectively active towards MftA**

Since multiple products are produced from MftC reaction, a new question regarding the correct precursor for mycofactocin biosynthesis has arisen. To address this question, we set out to probe the activity of MftE towards MftA* and MftA** or both. Previous studies have demonstrated that MftE from *Mycobacterium smegmatis* hydrolyzes MftA** to produce Val-p-(2-aminoethenyl)phenol and the remaining truncated MftA peptide (185). However, these results were largely dependent upon difficult HPLC−ESI-MS analysis of a complex mixture and with the inability to distinguish the substrate (e.g., MftA* vs MftA**) or the structure of the product. Additionally, the mass estimated for the small molecule is consistent with the hydrolyzed products from both MftA* and MftA**. To circumvent these obstacles, MftA* and MftA** were separately purified from reaction mixtures containing MftA, MftB, MftC, dithiothreitol (DTT), SAM, and dithionite (DTH), as previously described (Figure 3.10A) (99), and individually used in reactions with MftE. Purified and chemically reconstituted *Mu*MftE was incubated with MftA, MftA*, or MftA** for 3 h at room temperature and the reaction mixture was analyzed by HPLC while monitoring the absorbance of tryptophan and tyrosine at 280 nm. In the control experiment, MftA eluted at ∼13 min (Figure 3.10B, black) and the addition of *Mu*MftE to MftA did not result in a significant change in the chromatogram (Figure 3.10B, gray). Likewise, when *Mu*MftE was added to MftA**, no significant change was observed in the chromatogram (Figure 3.10B, blue and light blue). However, the addition of *Mu*MftE to MftA* resulted in the formation of two new product peaks (12.5 and 13.9 min) with the concomitant disappearance of the MftA* peak (Figure 3.10B, red and orange).

To confirm that this activity was not unique to *Mu*MftE, and to reconcile the different activities reported for *Ms*MftE versus those being discussed in this dissertation, MftE from *M. smegmatis* was cloned into and purified from *E. coli* (Figure 3.9 B), and tested its substrate specificity towards MftA, MftA* and MftA**. The reaction mixtures were analyzed on HPLC in the same manner described for *Mu*MftE. As expected, the HPLC chromatograms for the *Ms*MftE reaction showed results similar to that of *Mu*MftE (Figure 3.10C), indicating that AHDP is most likely the substrate for both *Ms* and MuMftE. To further probe the selectivity of MftE for its substrate, *Mu*MftE was incubated with MftA* in the presence of equal amounts of MftA** and reaction mixture was analyzed by HPLC. The chromatogram for the reaction showed that MftE retained its activity toward MftA* (Figure 3.10D). UV-vis spectral analysis of the products of MftE reaction showed that the new peak at 13.9 mins displayed absorption properties reminiscent of tyrosine (Tyr 30 of MftA*) while the peak at 12.5 mins was similar to that of the substrate MftA* (MftA (1-28) of MftA*). Taken together, these results demonstrate that MftE is highly specific for the substrate MftA* and does not recognize the unmodified peptide MftA or the intermediate MftA**. Therefore, we propose that MftA* is the correct precursor in mycofactocin biosynthesis.

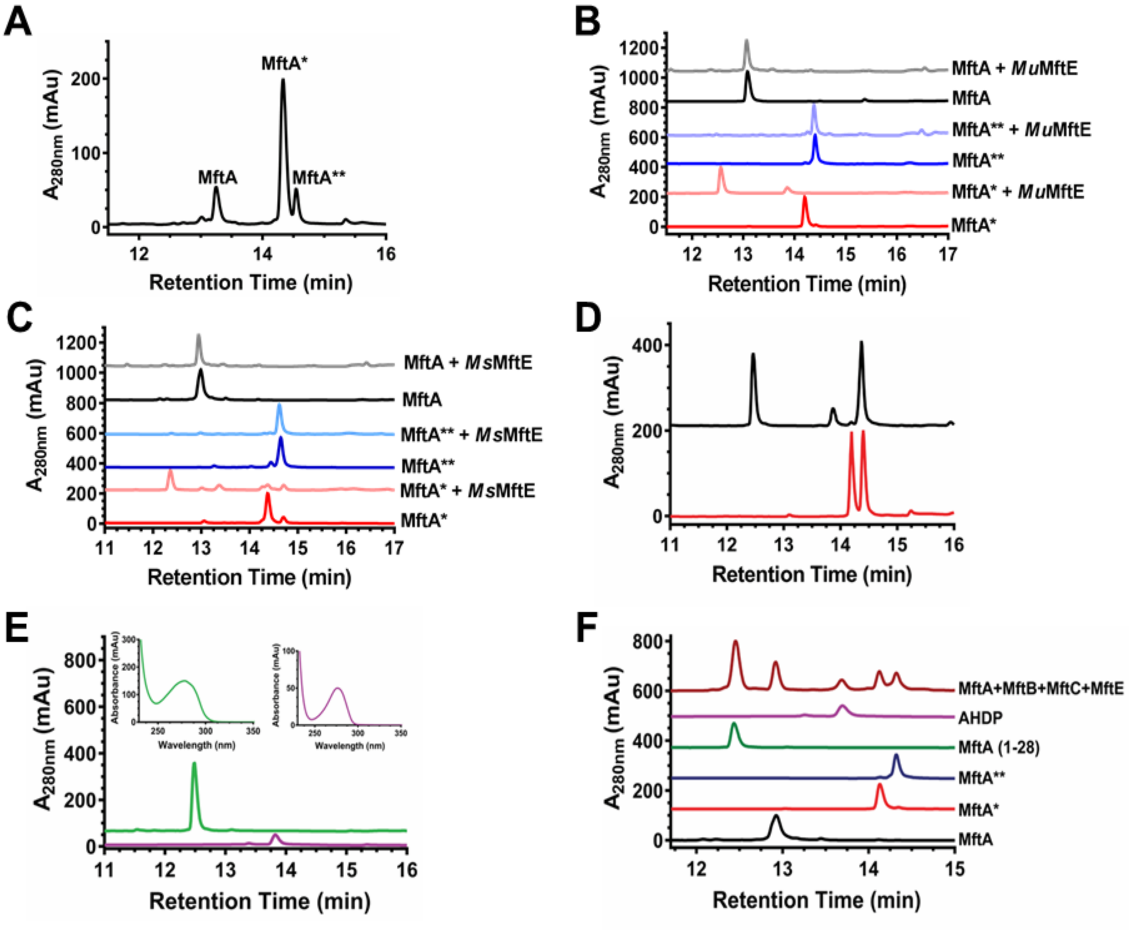


Figure 3.10 HPLC chromatograms of MuMftC and MuMftE reactions. (A) MftC reactions consisting of 2 mM SAM, 10 mM DTT, 2 mM DTH, and 300 µM of the MftC, 300 µM MftB and 300 µM MftA. (B) *Mu*MftE and (C) *Ms*MftE reactions containing 50 µM MftA* and MftA** with and without 50 µM MftE indicate that MftE preferentially hydrolyzes MftA* to form two new products with retention times of 12.5 and 13.9 min (orange). All MftE reactions were performed in the presence of 50 µM Zn^{2+} and 50 µM Fe^{2+}. (D) MftE selectively cleaves MftA* in the presence of MftA**. (E) UV-vis spectra of the products of MftE reaction. (F) A one-pot reaction mixture of MftC and MftE containing 2 mM SAM, 10 mM DTT, 2 mM DTH, and 300 µM of the MftC, 300 µM MftB, 300 µM MftA and 300 µM MftE.

3.9 MftE requires iron (II) for catalysis.

Having found from ICP-MS analysis that MftE binds both Zn^{2+} and Fe^{2+}, we next set out to determine the effects of these divalent cations on MftE catalysis. Initially, reactions were set up under steady state ([MftA*] ≫ [MftE]) conditions in the presence of 1 equivalent of Zn^{2+} and/or Fe^{2+} (Figure 3.11B); however, we observed that MftE catalyzes only single turnover reactions under these conditions. Attempts to modify reaction conditions to promote multiple turnovers (e.g., addition of ascorbate, DTT, Fe^{2+}, or Fe^{3+}) were unsuccessful (Figure 3.11A). Therefore, reactions were set up under single turnover conditions ([MftE] > [MftA*]) in the presence of Zn^{2+} and or Fe^{2+} and the reaction products were analyzed by HPLC. The relative integrated peak area of the substrate, MftA* was plotted versus time (Figure 3.11C), and the reaction rates were extracted from single-exponential decay fits and reported as k_{obs} (Table 4). Under these conditions, the rate of hydrolysis for as-purified MftE is a modest 0.001 min^{-1}. The addition of Zn^{2+} had limited effect on the reaction rate (k_{obs} = 0.018 min^{-1}), whereas the addition of Fe^{2+} increased the rate by >200-fold from 0.001 min^{-1} to 0.232 min^{-1}. The addition of both Zn^{2+} and Fe^{2+} had approximately the same effect on the reaction rate as Fe^{2+} alone. The aggregated data suggests that Fe^{2+} is catalytically active in MftE and that Zn^{2+} could be playing an auxiliary role. The observations made here are consistent with those reported for AprB where the addition of Fe^{2+} increased the catalytic rate 10-fold (196). In ArpB, it was proposed that the addition of Fe^{2+} replaced the catalytically inactive Fe^{3+}, resulting in an increased activity. However, unlike MftE, ArpB is capable of multiple turnovers, which suggests that we are missing an important cofactor in our reactions.

Protein	**k_{obs} (min^{-1})**
As-purified MftE	0.001 ± 0.001
MftE + $Fe(NH_3)SO_4$	0.018 ± 0.006
MftE + $ZnCl_2$	0.23 ± 0.04
MftE + $Fe(NH_3)SO_4$ + $ZnCl_2$	0.12 ± 0.01

Errors are reported as the sum of squared errors.

Table 4 Extracted initial reaction rates for MftE catalysis with the Addition of the Indicated Metals.

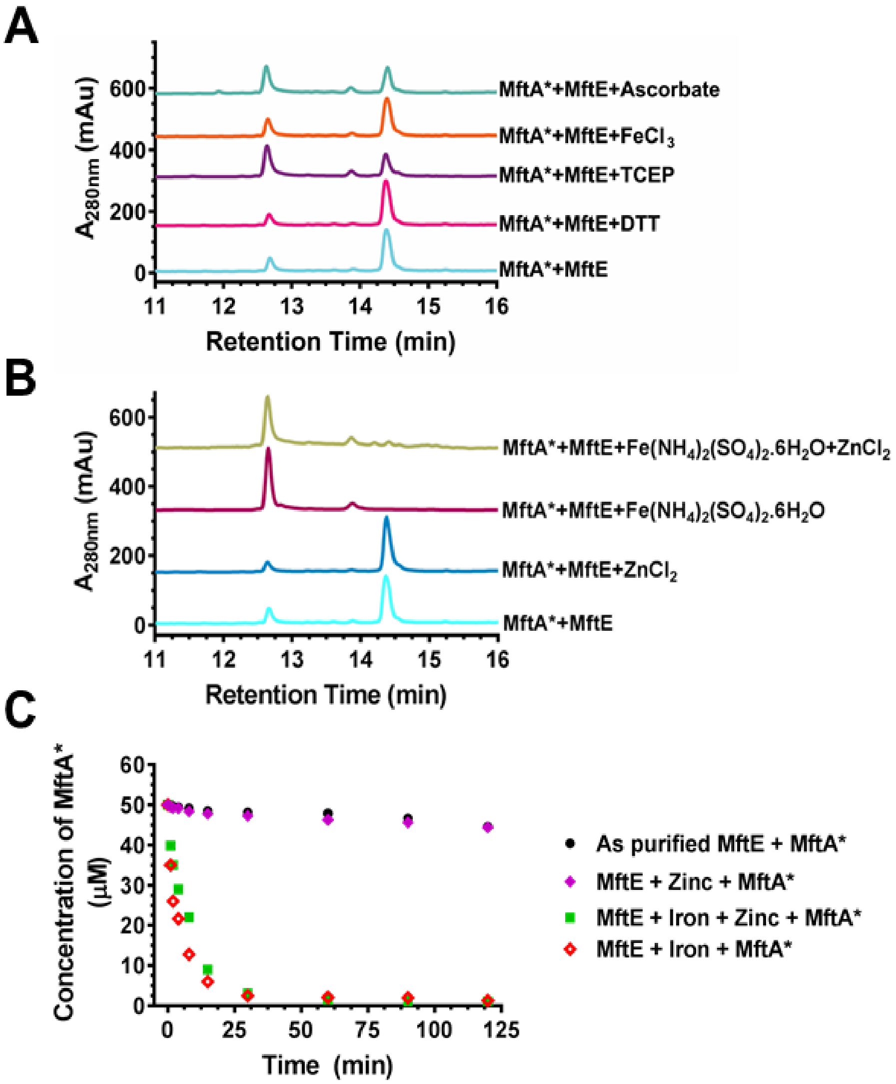


Figure 3.11 Screening of reaction conditions for MftE catalysis. (A) HPLC chromatograms of various reaction conditions to increase multiple turnovers. All reactions were carried out under steady-state conditions. (B) Addition of Fe^{2+} and or Zn^{2+} to MftE reaction indicates Fe^{2+} enhances the reaction rate of MftE. (C) The time course reaction of MftE with MftA* in the presence of various metals indicate that complete hydrolysis of MftA* is achieved in the presence of Fe^{2+}.

3.10 MftE hydrolyzes MftA to form 3-amino-5-[(p-hydroxyphenyl) methyl]-4,4-dimethyl-2-pyrroli-dinone (AHDP).*

Having demonstrated that MftE is specific for MftA*, we next set out to characterize the products of this reaction. To begin with, new peaks from HPLC analysis were isolated and analyzed by HRMS. First, the mass spectrum of MftA* shows an abundant species at *m/z* = 1111.85 ($[M + H]^{3+} + H_2O$, theoretical *m/z* = 1111.85) consistent with a 46 Da loss of mass from the unmodified peptide MftA (Figure 3.12A and B). The difference of 46 Da is indicative of the loss of 1C, 2O, and 2H, consistent with the decarboxylation of the C-terminal tyrosine and subsequent C−C bond formation between Val29 and Tyr30 as previously published (99). The mass spectrum of the 12.5 min peak from the MftE reaction showed a prominent species at *m/z* = 1550.21 ($[M + H]^{2+}$). This mass is 217 Da lower than the mass of MftA* (Figure 3.12C), and the difference in mass is consistent with the loss of the last two residues on MftA* plus the addition of an OH from the hydrolysis of the Gly28−Val29 amide bond (theoretical $[M + H]^{2+}$ *m/z* = 1550.21). The mass spectrum of the 13.9 min peak from the MftE reaction showed an abundant mass at *m/z* 235.1436 $[M + H]^{+}$ (Figure 3.12D). This new mass is consistent with the mass of AHDP (theoretical ($[M + H]^{+}$ *m/z* = 235.1441). These results demonstrate that MftE acts as a selective peptidase by hydrolyzing the amide bond between Gly28 and Val29 of MftA* to form two new products: MftA (residues 1−28) and, putatively, AHDP. To corroborate this finding, we turned to ^{1}H NMR to solve the structure of the two products from MftE reaction with MftA*.

The rationale for this experiment is that if MftE cleaves the last two residues of MftA*, we would expect to see the disappearance of the two aromatic doublet peaks from tyrosine in the ^{1}H NMR spectrum of the cleaved peptide MftA (1−28). Indeed, the ^{1}H NMR spectrum of MftA (1−28) shows the absence of the two-tyrosine aromatic doublets at δ = 7.12 and 6.79 ppm when compared to the ^{1}H NMR spectrum of MftA* (Figure 3.13A and B) (See Figure S 1A and B for full spectra). Currently, the only evidence of the existence of the AHDP precursor comes from deuterium labeling in reactions with MftC and limited ^{13}C NMR experiments with MftA*. Therefore, we set out to determine the structure of the putative AHDP molecule by using ^{1}H NMR, COSY, and ^{13}C NMR techniques. To do so, submilligram quantities of the molecule were purified from one-pot reaction mixtures containing MftA, MftB, MftC, MftE, SAM, DTT, and DTH (Figure 3.10F). To obtain the ^{13}C-enriched AHPD for ^{13}C NMR experiments, MftA labeled with both $^{13}C_9$ and ^{15}N-enriched tyrosine and $^{13}C_5$- and ^{15}N-enriched valine was used in the aforementioned one-pot reaction. Heavy MftA was produced via *in vivo* expression of the peptide in *E. coli* grown in M9 medium supplemented with the heavy amino acids. These procedures resulted in the heterogeneous incorporation of ^{13}C into AHDP. One-dimensional ^{1}H NMR analysis (Figure S 1C) of the putative AHDP is consistent with the structure containing a para-substituted aromatic (δ = 6.79 and 7.14 ppm), two nonequivalent methyl groups (δ = 0.82 and 1.03 ppm), an isolated hydrogen near an amino group (δ = 3.31 ppm), two nonequivalent benzylic hydrogens (δ = 2.498 and 2.79 ppm), and a hydrogen within a substituted lactam (δ = 3.56 ppm). COSY spectral analysis (Figure S 1E) indicates several crosspeak couplings between proton peaks. These include crosspeak couplings between aromatic hydrogens on C9−C10 (Figure 3.13C), and between aliphatic hydrogens on C7−C1 and C5−C1 (Figure 3.13D). Assignments of these couplings are highlighted in Figure 3.13E and are consistent with the proposed structure for AHDP. Additionally, ^{13}C NMR analysis (Figure S 1D) provided the chemical shifts of the carbon signals that are consistent with the proposed structure for AHDP. The aggregated MS data and NMR data presented here are consistent with the product of the MftE reaction with MftA* being AHDP.

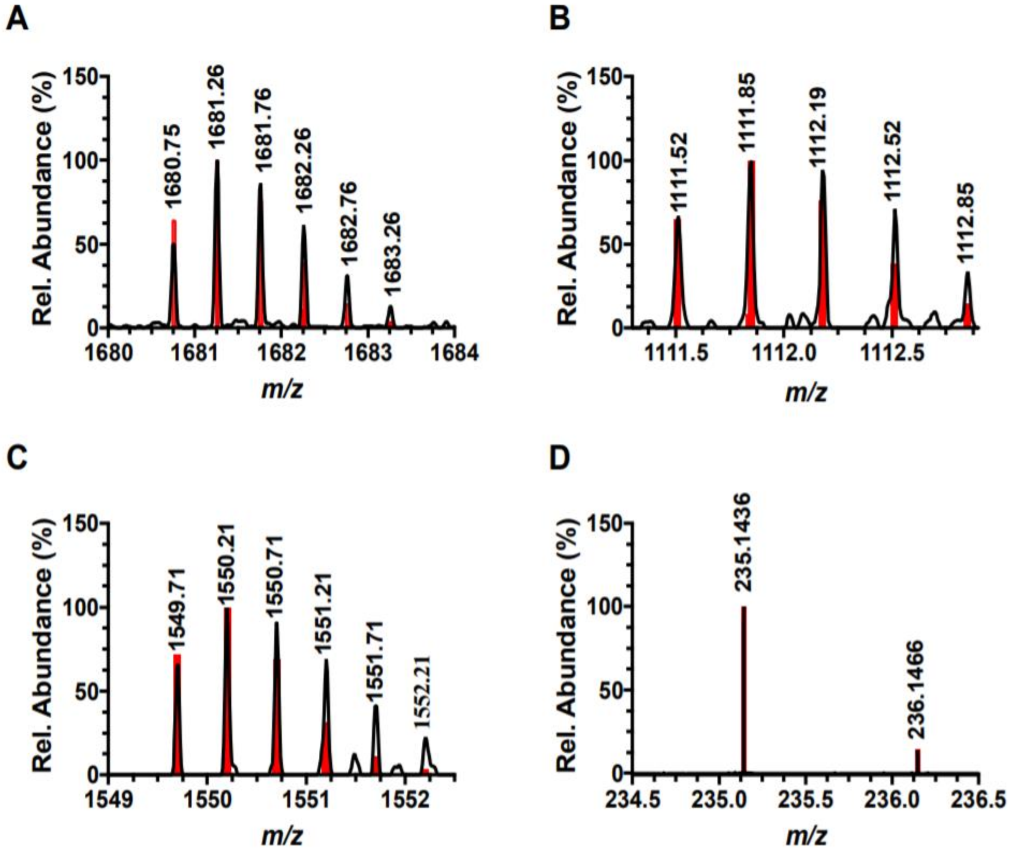


Figure 3.12 HRMS analysis of the products of MftA* reaction with MftE. (A) The extracted ion chromatogram of M1W MftA consistent with a $[M+2H]^{2+}$. (B) The extracted ion chromatogram of MftA* corresponding to a $[M+3H]^{3+}$ + H_2O. (C) The ion chromatogram of the remaining cleaved peptide MftA (1-28) corresponding to a $[M+2H]^{2+}$. (D) The ion chromatogram of the dipeptide AHDP corresponding to a $[M+H]^{1+}$. Simulated ion chromatograms are indicated in red and the experimental ion chromatograms are indicated in black.

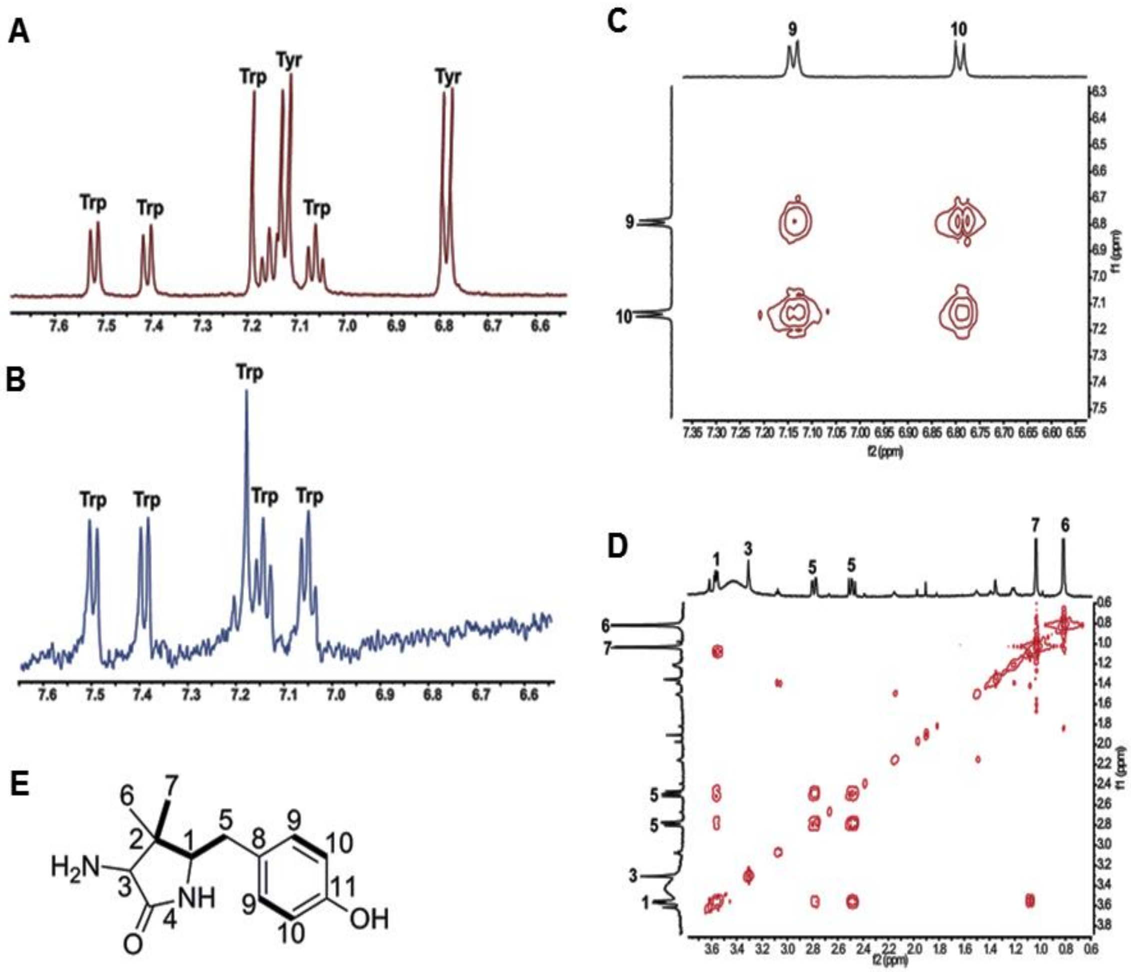
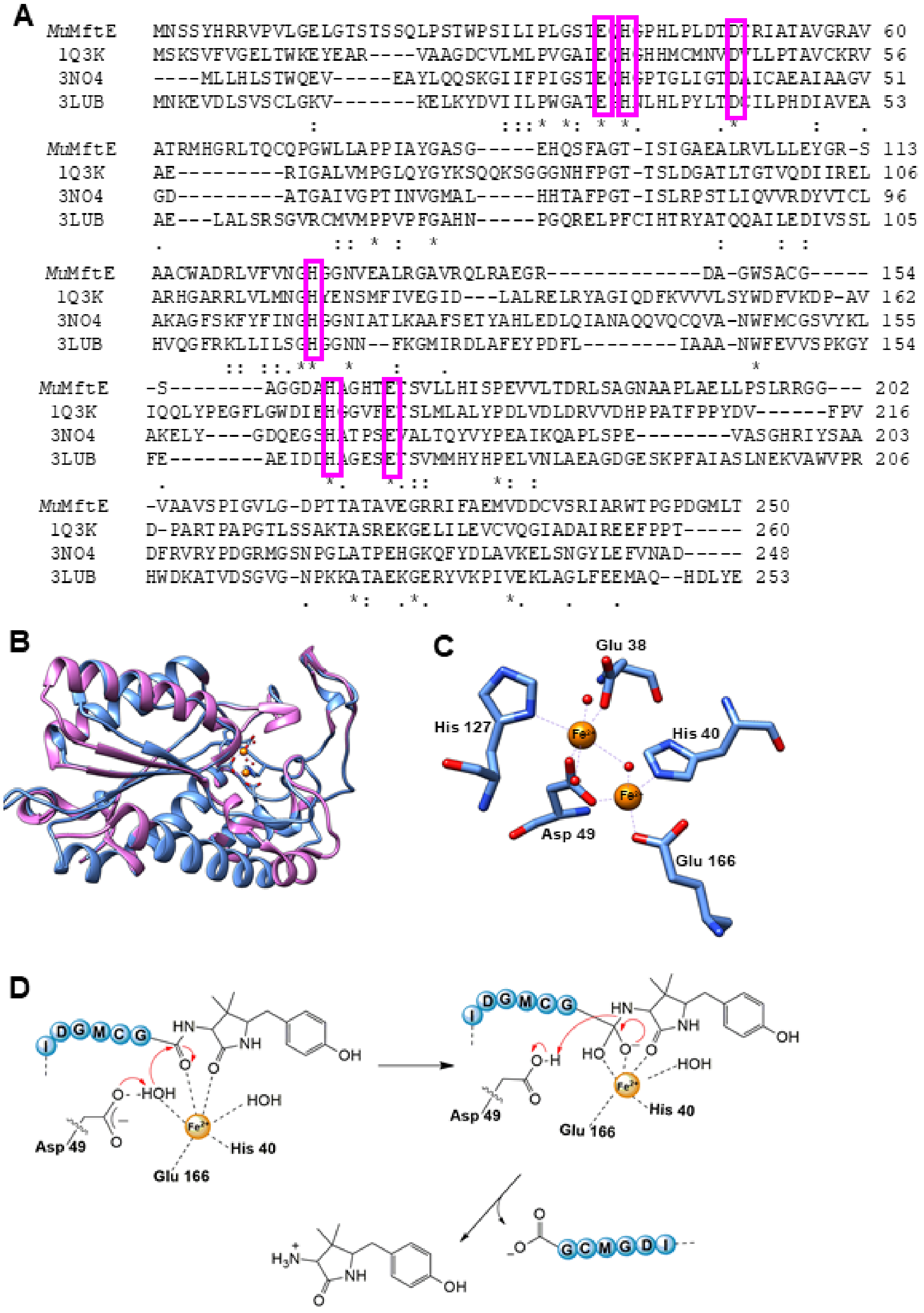


Figure 3.13 NMR spectral analysis of the reaction products of MftE. ^{1}H NMR spectra of the aromatic regions of (A) MftA* and (B) MftA (1-28) showing the loss of the aromatic peaks at δ = 7.12 and 6.77 ppm. (C) COSY NMR spectra of the aromatic regions showing cross couplings between the hydrogens on C9 and C9. (D) COSY NMR spectra of the lactam region showing important cross coupling peaks between the hydrogens on C7-C1 and C5-C1. (E) Structural representation of the results from COSY NMR. Cross-peak couplings are represented by thick bonds.

To decipher the reaction mechanism of MftE, we turned our attention to the members of creatininase amidohydrolase family. As described earlier, members of this family catalyze the hydrolysis of diverse array of substrates (197). They are typically binuclear zinc metalloenzyme that play key roles in bacterial creatininase degradation. A search of the PDB database with the sequence of MftE

revealed similarity to creatininase homologs whose structures are known. Sequence alignments (Figure 3.14A) and mapping of conserved residues of MftE onto the structures of a creatinine amidohydrolase (PDB ID 1Q3K) show that MftE retains the active site residues that bind the divalent cations (Figure 3.14B and C) (197). These divalent cation metals are used to activate water molecules and to stabilize the negatively charged oxyanion intermediate generated during catalysis reaction. Modeling studies suggest that the divalent cations interact with both the carbonyl of the peptide bond to be hydrolyzed and an active polar residue. This interaction results in the polarization of the carbonyl group for subsequent nucleophilic attack (198). As a member of the creatininase amidohydrolase family and on the basis of sequence conservation in the active site of MftE, a reaction mechanism similar to that of creatininase amidohydrolase has is proposed (Figure 3.14D). This mechanism involves the coordination Fe^{2+} by both the carbonyl on Gly20 and an active site polar residue (e.g. Glu or Asp). This polarization results in the susceptibility of carbonyl bond to nucleophilic attack and subsequent bond cleavage (Figure 3.14D). However, it should be noted that without additional guiding information regarding the crystal structure of MftE, it is problematic to accurately predict the mechanism of MftE hydrolysis.



```
MuMftE    MNSSYHRRVPVLGELGTSTSSQLPSTWPSILIPLGSTEQHGPHLPLDTDTRIATAVGRAV 60
1Q3K      MSKSVFVGELTWKEYEAR----VAAGDCVLMLPVGAIEQHGHHMCMNVDVLLPTAVCKRV 56
3NO4      ----MLLHLSTWQEV-----EAYLQQSKGIIFPIGSTEQHGPTGLIGTDAICAEAIAAGV 51
3LUB      MNKEVDLSVSCLGKV-------KELKYDVIILPWGATEQHNLHLPYLTDCILPHDIAVEA 53
                       :               :::* *: * *.      .*      :   .
MuMftE    ATRMHGRLTQCQPGWLLAPPIAYGASG-----EHQSFAGT-ISIGAEALRVLLLEYGR-S 113
1Q3K      AE---------RIGALVMPGLQYGYKSQQKSGGGNHFPGT-TSLDGATLTGTVQDIIREL 106
3NO4      GD---------ATGAIVGPTINVGMAL-----HHTAFPGT-ISLRPSTLIQVVRDYVTCL 96
3LUB      AE---LALSRSGVRCMVMPPVPFGAHN-----PGQRELPFCIHTRYATQQAILEDIVSSL 105
          .              :: * :  *                       :    : :
MuMftE    AACWADRLVFVNGHGGNVEALRGAVRQLRAEGR-------------DA-GWSACG----- 154
1Q3K      ARHGARRLVLMNGHYENSMFIVEGID---LALRELRYAGIQDFKVVVLSYWDFVKDP-AV 162
3NO4      AKAGFSKFYFINGHGGNIATLKAAFSETYAHLEDLQIANAQQVQCQVA-NWFMCGSVYKL 155
3LUB      HVQGFRKLLILSGHGGNN--FKGMIRDLAFEYPDFL--------IAAA-NWFEVVSPKGY 154
                ::  ::.**  *   :   .                         *
MuMftE    -S--------AGGDAHAGHTETSVLLHISPEVVLTDRLSAGNAAPLAELLPSLRRGG--- 202
1Q3K      IQQLYPEGFLGWDIEHGGVFETSLMLALYPDLVDLDRVVDHPPATFPPYDV------FPV 216
3NO4      AKELY----GDQEGSHATPSEVALTQYVYPEAIKQAPLSPE--------VASGHRIYSAA 203
3LUB      FE--------AEIDDHAGESETSVMMHYHPELVNLAEAGDGESKPFAIASLNEKVAWVPR 206
           .             *.   *.::     *: :
MuMftE    -VAAVSPIGVLG-DPTTATAVEGRRIFAEMVDDCVSRIARWTPGPDGMLT 250
1Q3K      D-PARTPAPGTLSSAKTASREKGELILEVCVQGIADAIREEFPPT----- 260
3NO4      DFRVRYPDGRMGSNPGLATPEHGKQFYDLAVKELSNGYLEFVNAD----- 248
3LUB      HWDKATVDSGVG-NPKKATAEKGERYVKPIVEKLAGLFEEMAQ--HDLYE 253
                       .   *:  .*.      *.   .   .
```

Figure 3.14 Sequence alignment and proposed mechanism of MftE catalysis (A) Sequence alignment of MftE with known crystal Structures of members of creatininase amidohydrolase family. Highlighted in magenta are conserved residues in the active sites for binding metal ions. (B) structure of MftE (blue) modeled onto the creatinine amidohydrolase (PDBID: 1Q3K) (magenta). (C) structure of MftE showing the catalytic residues used to coordinate two Fe^{2+}. (D) Proposed reaction mechanism for MftE hydrolysis.

3.11 *MftD is a member of the FMN-dependent alpha-hydroxy acid oxidase family (IPR012133)*

MftD is a member of the FMN- dependent alpha hydroxy acid oxidase family of proteins. This is a large protein family consisting of more than 25,000 protein sequences annotated as IPR012133. This family includes notable members such as FMN-dependent lactate dehydrogenase and mandelate dehydrogenase. To observe the relationship between the mycofactocin system oxidoreductase and other members of this protein family, a sequence similarity network (SSN) was generated for the IPR012133 family using EFI-EST. From the analysis of 34,115 Uniref90 sequences, a network consisting of 2,215 nodes where each node represents a cluster of protein sequences that are 50% identical were generated. Interestingly, the SSN results showed the divergence of the mycofactocin system oxidoreductase (MftD) from the central alpha-hydroxy acid oxidase cluster (Figure 3.15, yellow). Clusters of other notable members such as mandelate dehydrogenase were found to also stand out from the central cluster. Members of the alpha-hydroxy acid oxidase family are known to use an FMN-dependent ping-pong reaction mechanism for catalysis. In the first reductive half reaction, an alpha-hydroxy acid substrate is oxidized to an alpha-keto-acid resulting in the reduction of the enzyme-bound FMN. In the second oxidative half reaction, the flavin is re-oxidized by transferring an electron to an acceptor such us molecular oxygen to produce hydrogen peroxide (199). This commonality among the FMN-dependent dehydrogenases led us to speculate that if mycofactocin is indeed a redox cofactor, then MftD could catalyze the molecular oxygen dependent oxidation of the phenyl ring of AHDP to generate the corresponding catechol with a redox active center which is not apparent in AHDP.

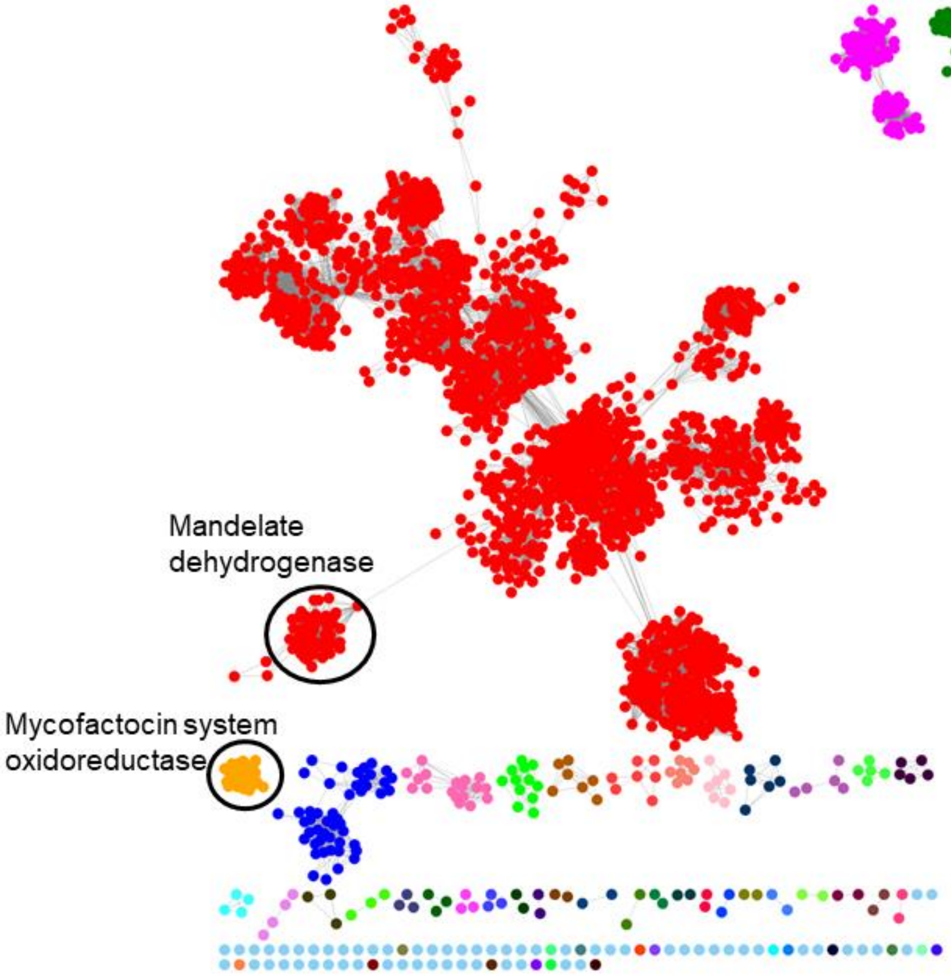


Figure 3.15 A sequence similarity network for MftD. A sequence similarity network for the FMN-dependent alpha-hydroxyacid oxidase family (IPR012133) shows the high sequence identity between members of this family (red) resulting clustering together of mycofactocin system flavin oxidoreductase (purple) together with other notable members such as lactate and mandelate dehydrogenase (green).

3.11 MftD Is an FMN Binding Protein.

To determine the role of MftD in the biosynthesis of mycofactocin, the *mftD* gene (*mul_0774*) was cloned from *M. ulcerans* Agy99, and the *Mu*MftD protein was heterologously expressed in and anaerobically purified from *E. coli.* The protein was purified to homogeneity and >95% purity as indicated by SDS-PAGE (Figure 3.16A lane 1). Characterization of the as-purified *Mu*MftD protein by UV−visible absorbance spectroscopy indicated the presence of a reduced flavin, which is characterized by the broad shoulder between 300 and 410 nm (Figure 3.16C red). Upon exposure of the anaerobically purified protein to air, we observed a color change from colorless to intense yellow (Figure 3.16B). The absorbance spectrum of the protein shifted from a reduced flavin to that of an oxidized flavin, characterized by the dual absorbance maxima at 360 and 445 nm (Figure 3.16C, green). To determine which species of flavin was bound, *Mu*MftD was heat precipitated and the soluble fraction was analyzed by HPLC. The retention time of the flavin in the soluble fraction was compared to authentic flavin standards of flavin adenine dinucleotide (FAD), flavin mononucleotide (FMN), and riboflavin (Figure 3.16D). From the HPLC analysis, it was evident that the absorbance features of the cofactor bound to *Mu*MftD originated from FMN, consistent with other members of the protein-fold family (200). Notably, when *Mu*MftD was purified aerobically, the protein appeared to lose FMN during the isolation process. Consistent with this observation, analysis of *Mu*MftD by analytical HPLC-SEC demonstrated that, upon oxidation, the protein releases the bound FMN (Figure

3.16E), in the absence of substrate. In our hands, we observed that the loss of FMN is irreversible.

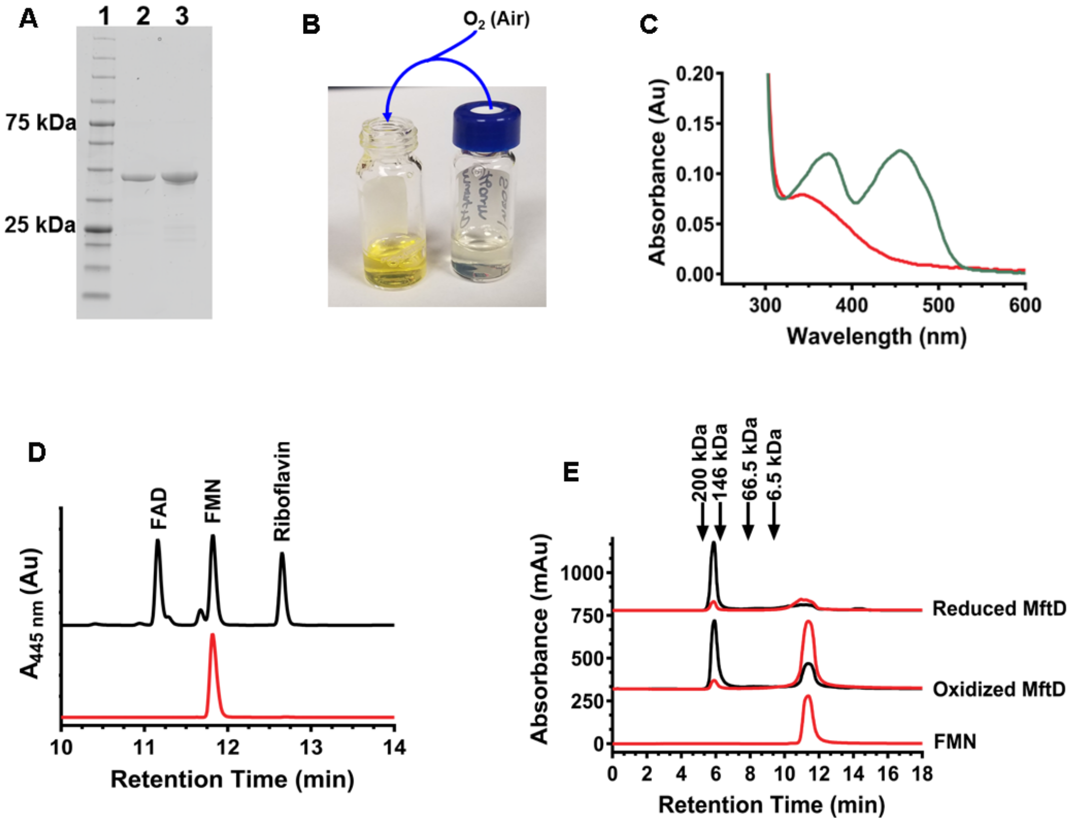


Figure 3.16 Characterization of MftD protein. (A) MftD was purified to homogeneity as determined by SDS-PAGE analysis. Lane 1 contains the Goldbio BLUEstain Protein Ladder standard, lane 2 contains His6-purified MuMftD, and lane 3 contains His6-purified *Ms*MftD. (B) Color change observed when the protein is exposed to oxygen indicating the oxidation of bound FMN upon oxygen exposure. (C) UV−vis spectral analysis of anaerobically prepared MftD indicates the presence of a reduced flavin (red). Upon exposure to air of the protein sample, the UV−vis spectrum shifts considerably, indicating that the flavin is oxidized. (D) High-pressure liquid chromatography analysis of the flavin-containing fraction of the protein precipitate (red) indicates that MftD likely binds FMN when compared to retention times of flavin standards (E) Analytical size exclusion chromatography of reduced and oxidized MftD shows that the purified protein migrates as a tetramer with a molecular weight of ~ 170 kDa. MftD loses > 50% of the bound FMN upon oxidation. Chromatograms were reported at 280 nm (black) for protein and 450 nm (red) for FMN.

3.12 Reconstitution of MftD Activity

To determine if the product from the MftE reaction is the substrate for MftD, reactions containing AHDP and MftD were carried out aerobically, and analyzed by HPLC. The aerobic addition of AHDP to MftD and subsequent analysis by HPLC does in fact give rise to a new species at a retention time of ~15.8 min (Figure 3.17A, green) with the concomitant disappearance of AHDP at ~14.1 min. The associated k_{obs} for the single turnover reaction was measured to be 0.010 ± 0.002 min^{-1} (Figure 3.17B). Interestingly, the UV−visible absorbance spectrum for the new species is substantially different than that of AHDP, indicating that the resulting modification also impacted the electronics of the molecule (Figure 3.17 C and D). To ensure that this activity was not unique to *Mu*MftD, the gene *msmeg_1424*, encoding for *Ms*MftD, was cloned from *M. smegmatis* mc^2155 and the protein was purified from *E. coli* (Figure 3.16A). The purified protein was used to set reactions with AHDP and analyzed by HPLC. As expected, the HPLC chromatogram for the *Ms*MftD reaction showed similar results to that of *Mu*MftD (Figure 3.17A, blue), indicating that AHDP is most likely the substrate for both *Ms* and *Mu*MftD. Subsequent analyses were carried out solely with *Mu*MftD and will refer to the protein as simply MftD.

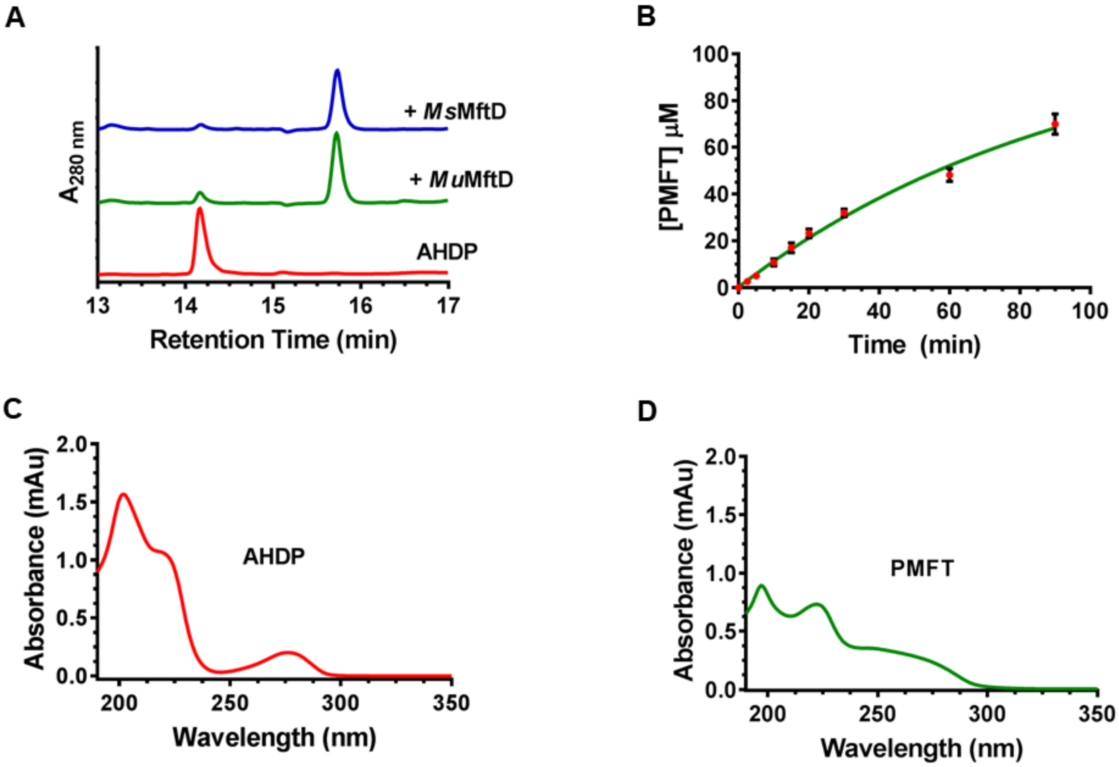


Figure 3.17 HPLC analysis of MftD reactions. (A) HPLC chromatograms of reactions containing AHDP (red), AHDP and *Ms*MftD (blue), AHDP and *Mu*MftD (green) indicate that AHDP is an active substrate for MftD. (B) Single turnover kinetic analysis of the MftD reaction monitored by HPLC. Triplicate reactions were carried out with 100 μM AHDP and 200 μM MftD. Data were fitted to a single exponential equation (green line) to determine the k_{obs}. UV-vis spectra of (C) the substrate AHDP, and (D) the product MftD reaction named PMFT.

3.13 HRMS analysis of MftD reactions

To determine the type of chemical modification imparted on AHDP by MftD, we analyzed the product of the reaction by high resolution mass spectrometry (HRMS). Consistent with the previous analysis of AHDP described above, the HRMS analysis of the AHDP $[M + H]^+$ ion, reanalyzed here, was found to have a mass-to-charge ratio (*m/z*) of 235.1445 (Figure 3.18A, black), within 4 ppm of the predicted mass ($C_{13}H_{18}N_2O_2$, theoretical *m/z* = 235.1441 Da). HRMS analysis of the ~15.8 min MftD product provided a *m/z* = 234.1107 (Figure 3.18A, green). This mass is consistent with the loss of NH_3 and the addition of a single oxygen atom to ADHP ($C_{13}H_{15}NO_3$, theoretical $[M + H]^+$ *m/z* = 234.1125). These results are inconsistent with our previous proposal suggesting that MftD could catalyze the formation of a catechol on the phenyl ring of AHDP. Instead, our HRMS analysis suggests that MftD might catalyze the oxidative deamination of AHDP, installing an α-keto moiety, resulting in the formation of 5-[(p-hydroxyphenyl)methyl]-4,4-dimethyl- 2,3-pyrrolidinedione (Figure 3.18B, green) or herein referred to as premycofactocin (PMFT). Precedence for flavin-dependent oxidative deamination comes from D-amino acid oxidase (DAAO). DAAO's function is to catalyze the flavin adenine dinucleotide (FAD)-dependent and stereospecific oxidation of D-amino acid to form α-keto acids (192, 201). However, it should be noted that DAAO and MftD belong to different protein-fold families (202, 203).

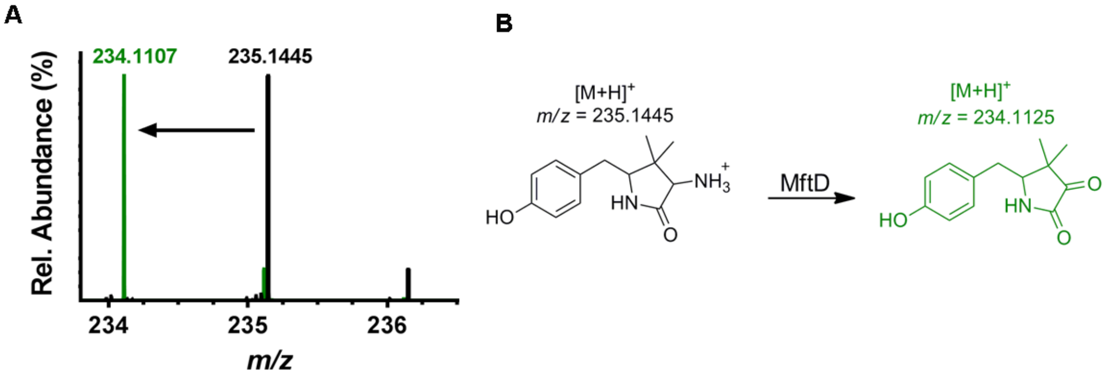


Figure 3.18 HRMS analysis of MftD modification reactions. (A) Analysis of the AHDP (black) and the MftD product (green) shows an ion with an m/z that is consistent with the loss of $-NH_3$ and the addition of O. (B) From the HPLC and HRMS analysis, MftD is proposed to catalyze the oxidative deamination of AHDP (black) to form premycofactocin (PMFT, green). The theoretical *m/z* for the molecules are indicated above their structure.

3.14 NMR structural elucidation of the MftD reaction Product.

To confirm that MftD indeed catalyzes the oxidative deamination of AHDP and the subsequent formation of PMFT, 1D and 2D NMR studies were carried out on the isolated product. Reactions containing AHDP and MftD were performed on a large scale (~2 mg of AHDP), and the HPLC purified MFT was analyzed by NMR. For ^{13}C NMR, we used synthesized MftA to contain uniformly ^{13}C labeled valine and uniformly ^{13}C labeled tyrosine at positions Val29 and Tyr30. The heavy peptide was used in reactions with MftB, MftC, MftE, and MftD to generate the uniformly ^{13}C-labeled PMFT. This was required to increase the ^{13}C NMR signal due to the limited quantities of isolated PMFT. Although we had already reported on the ^{1}H, ^{13}C, and COSY NMR characterization of AHDP, we further analyzed it by HSQC 2D NMR in this study. This allowed us to directly compare AHDP spectra to that of PMFT. Analysis of the COSY NMR spectra of the lactam region for both AHDP (Figure 3.19A, red) and PMFT (Figure 3.19A, green) showed some difference in the structure of the substrate and the product. Notably, the chemical shifts of the protons within the lactam region of PMFT are modestly shifted downfield by 0.1 - 0.4 ppm. This downfield shift is accompanied by the disappearance of the Hα in the 1H NMR spectrum of PMFT (Figure 3.19A), suggesting that the Cα of the lactam was the site of modification. Consistent with this observation, ^{13}C NMR indicated a drastic downfield shift of the Cα from ~61 ppm in AHDP (Figure S 2B) to ~210 ppm in PMFT (**Error! Reference source not found.**B). The chemical shift of the Cα is corroborated by HMBC spectra (Figure 3.19B), which showed long-range interaction between the protons on the geminal methyl and the Cα. Lastly, overlaid HSQC spectra of AHDP and PMFT showed the loss of the proton−carbon coupling associated with H−Cα (Figure 3.19C). Together, these NMR data provide strong evidence that MftD catalyzes the oxidative deamination of AHDP to form PMFT.

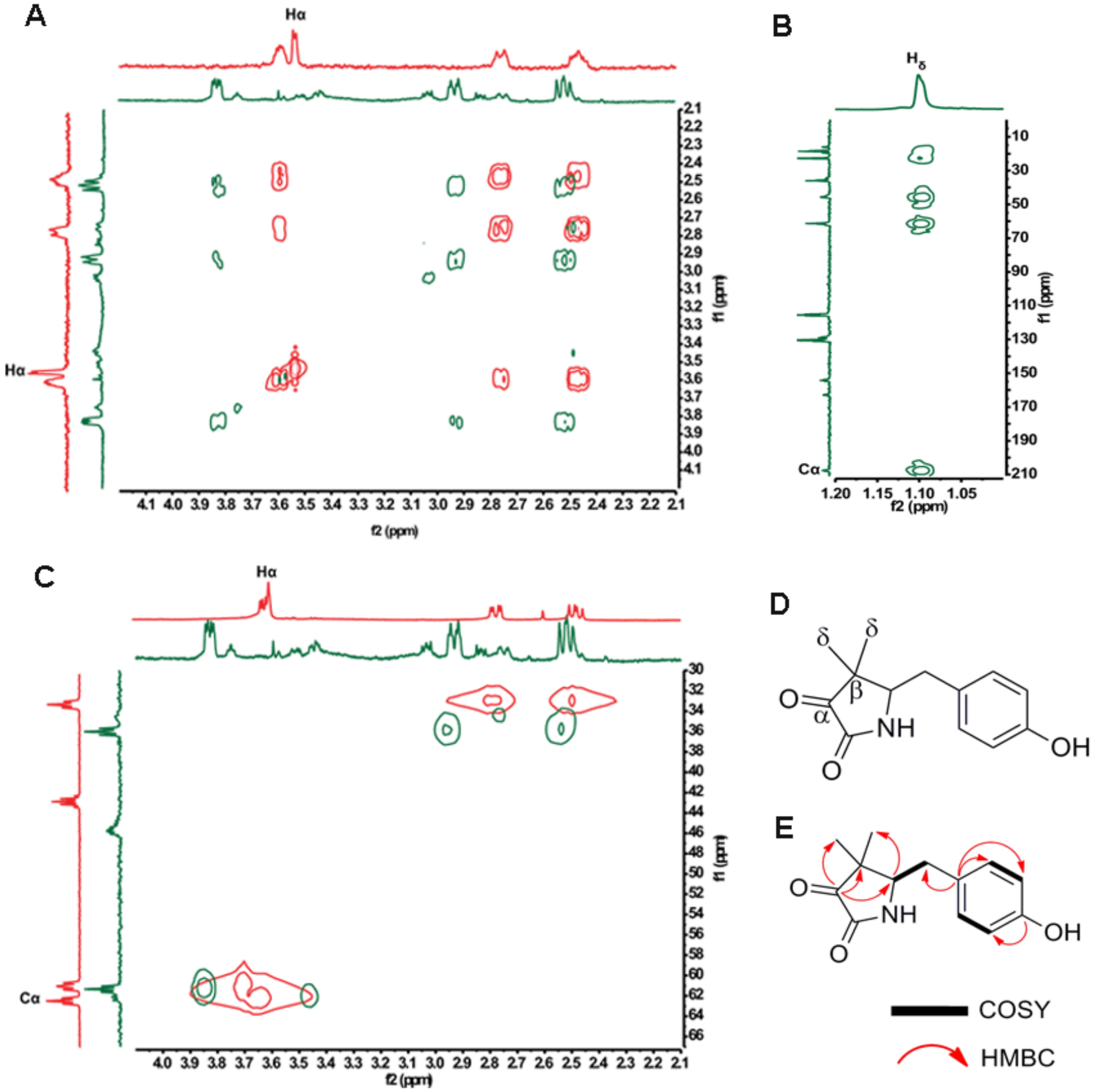


Figure 3.19 NMR analysis of MftD reaction product PMFT. (A) Overlaid COSY spectra of AHDP (red) and PMFT (green) showing the loss of the Hα signal. (B) HMBC spectrum of PMFT shows that the Cα has shifted downfield to ~210 ppm. (C) Overlaid HSQC spectra of ADHP (red) and PMFT (green) show the loss of the H−Cα heteronuclear interaction. (D) Reference structure of PMFT with the important carbon annotated. (E) Summary of relevant NMR correlations on PMFT.

3.15 Role of oxygen in MftD reaction.

Having solved the structure of PMFT, we next set out to determine the source of the incorporated oxygen atom. Since MftD catalyzed a similar reaction to DAAO, we turned to its well-characterized mechanism for guidance. The DAAO reaction is described by the two-step oxidative deamination of the D-amino group on the amino acid (Figure 3.20A). To begin with, DAAO oxidizes the amino acid to an α-imino acid through a hydride transfer from the Cα to FAD (202). This is followed by a nucleophilic substitution by an activated water and the subsequent oxidation to form the ketone (201). We rationalized that MftD could carry out a similar reaction mechanism, which we describe in Figure 3.20B.

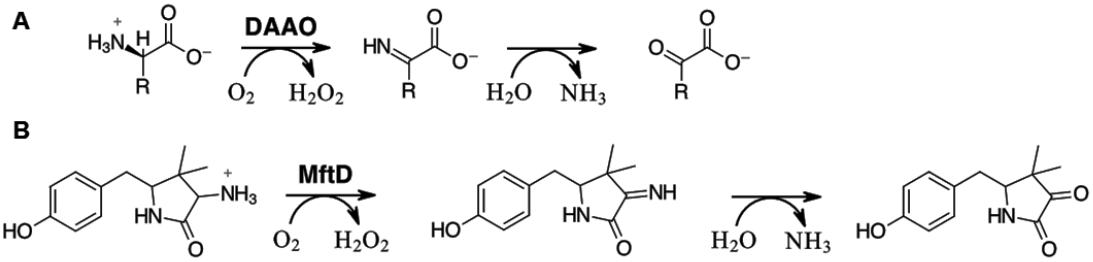


Figure 3.20 Reaction mechanism of MftD. (A) Condensed Reaction Scheme for D-Amino Acid Oxidase (DAAO) and (B) for MftD

To determine if water was the source of the inserted oxygen, we carried out MftD reactions in 99% enriched $^{18}OH_2$ and analyzed the resulting PMFT by HRMS. From this experiment, we expected to observe an enrichment of an PMFT ion that was consistent with a single ^{18}O being incorporated (+2 *m/z*). As expected, the mass spectrum of the reaction mixture carried out in $^{18}OH_2$ showed two predominant ions (Figure 3.21A), one at the expected *m/z* for unlabeled PMFT ([M

+ H]$^+$ *m/z* = 234.1125) and the other at a *m/z* = 236.1169, within 2 ppm of the expected mass for the ^{18}O incorporated PMFT ([M + H]$^+$ *m/z* = 236.1167). The percent incorporation was ~65%, which is near the expected theoretical maximum of 87%. Conversely, when the reaction was carried out in $^{18}O_2$, we did not observe any significant incorporation of ^{18}O into the product (Figure 3.18B). The rate of back exchange was found to not significantly affect the overall incorporation of ^{18}O into the product during the course of this experiment (Figure 3.21C). Therefore, these results are consistent with water being the source of the inserted oxygen atom in the MftD-catalyzed oxidative deamination of AHDP.

In the DAAO mechanism and the proposed mechanism for MftD, O_2 is the final electron acceptor (201). To verify that molecular oxygen is required for MftD turnover, reactions were carried out anaerobically. To achieve anoxic conditions, reactions were conducted in an anaerobic chamber and supplemented with glucose and glucose oxidase prior to the addition of MftD. Glucose oxidase catalyzes the oxidation of glucose to gluconate and hydrogen peroxide and is often added to reactions to achieve anaerobicity (204, 205). The subsequent HPLC chromatogram of the anaerobic reaction indicated only the presence of the starting material AHDP (Figure 3.21D). This suggests that molecular oxygen is required for the MftD activity. Since anaerobically purified MftD contains $FMNH_2$, and since we did not observe any turnover of MftD, it is likely that O_2 is required for the oxidation of $FMNH_2$. Notably, MftD was inactive in reactions carried out under the same conditions but with the addition of NAD$^+$ (Figure 3.21E), suggesting that the nicotinamide could not fulfill the role of O_2. Taken together, the results from this analysis is consistent with O_2 being the final electron acceptor in MftD catalysis.

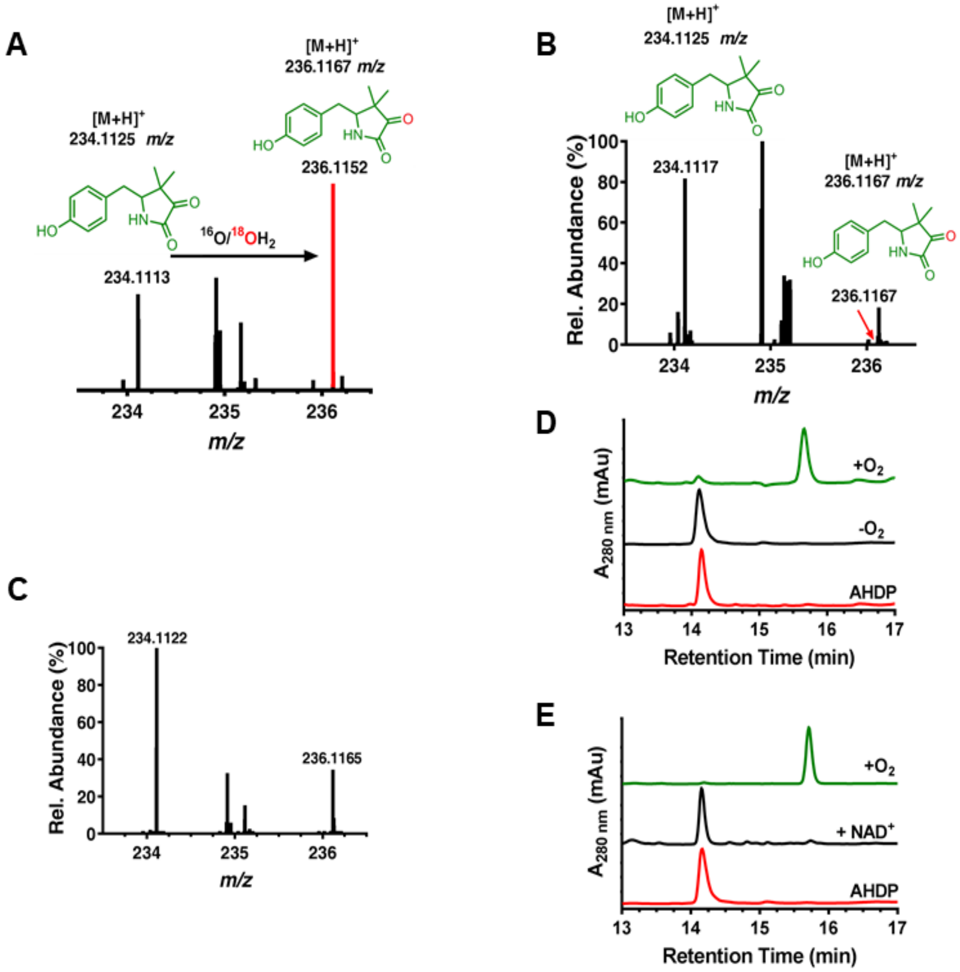


Figure 3.21 Assessing the requirement of oxygen for MftD catalysis. (A) High-resolution mass spectra of the MftD reaction carried out in $^{18}OH_2$ showing the enrichment of the ^{18}O-incorporated PMFT (red). The structures for natural abundance and enriched PMFT are shown above with their theoretical mass for the $[M + H]^+$ ion. (B) Control experiments for HRMS spectra monitoring the incorporation of ^{18}O into PMFT during MftD reaction in $^{16}OH_2/^{18}O_2$. (C) HRMS control spectra for a 60 min exchange experiment where ^{16}O-PMFT was incubated in $^{18}OH_2$ buffer. (D) HPLC analysis of MftD reactions under anaerobic (black) and aerobic (green) conditions suggests that molecular oxygen is required for catalytic turnover. The reference chromatogram for AHDP is shown in red. (E) Assessing MftD activity with NAD^+. HPLC chromatogram of the anerobic reaction of MftD with AHDP (red) in the presence of NAD+ (black) shows no PMFT formation at 15.8 min.

3.16 PMFT contains redox active center.

As stated previously, the structure of AHDP did not have an apparent physiologically relevant redox center. However, following oxidative deamination by MftD, we rationalized that the resulting α-keto-amide moiety could be the active site of PMFT. Indeed α-keto acids are a common $2e^{-}/2H^{+}$ redox moiety and can be found in biologically important molecules such as pyruvate, α-ketoglutarate, and oxaloacetate. To provide insight about the possible redox behavior of PMFT, cyclic voltammetry (CV) was used to directly measure the midpoint potential of PMFT. Three-electrode CV experiments were carried out using a glassy carbon working electrode, an Ag/AgCl reference electrode, and a platinum counter electrode. PMFT was noncovalently adsorbed onto the working electrode using single-walled carbon nanotubes (SWCNTs) as a coabsorbent. CV measurements yielded a single signal that was observed between −100 and −400 mV as compared to the buffer/SWCNT control (Figure 3.22). From this measurement, the reductive and oxidative potentials were found to be −370 and −140 mV, respectively, and the overall midpoint potential to be −255 mV. We did not observe significantly differentiated peaks for two single electron transfer events; thus, we define the midpoint potential for the $2e^{-}/2H^{+}$ transfer as −255 mV.

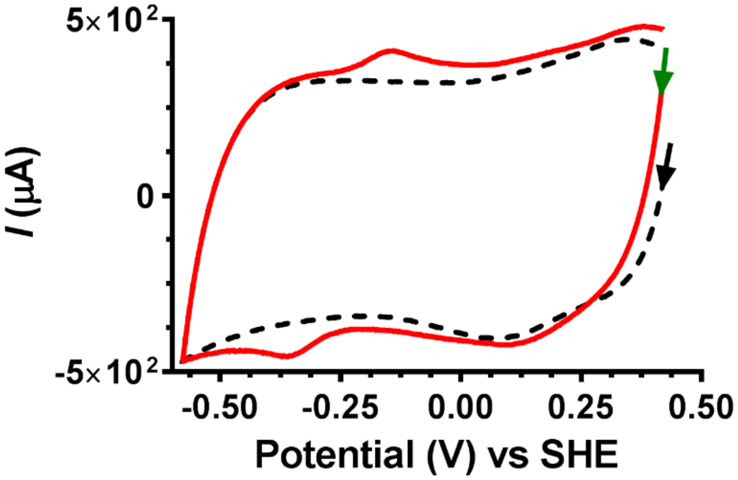


Figure 3.22 Cyclic voltammetry of PMFT. Overlaid cyclic voltammogram of PMFT/SWCNT (red) and buffer/SWCNT (black). Voltammetry of PMFT was measured at pH 7.0 and at 22 °C with a scan rate of 50 mV/s.

3.17 PMFT is redox and functionally active in vitro.

Having shown that PMFT is redox active, we next postulated that it could be used to oxidize NADH in mycofactocin dependent dehydrogenases. *M. smegmatis* mc^{2}155 carveol dehydrogenase (*Ms*CDH, *Msmeg_1410*) was chosen as the model mycofactocin-dependent dehydrogenase to test this hypothesis since the *R. erythropolis* DCL14 homologue has been shown to catalyze the oxidation of carveol to carvone using 2,6-dichlorophenolindophenol (DCPIP) as an electron acceptor (206) and since it was known to copurify with NADH (182). This latter property is significant because it could allow us to monitor the PMFT-dependent oxidation of NADH to NAD$^+$ by monitoring UV−visible absorbance at 340 nm (ΔεNADH 340 nm = 6.22 mM^{-1} cm^{-1}). To demonstrate that PMFT could be used as an oxidant, scanning UV−visible spectroscopic measurements were carried out with purified *Ms*CDH (see Figure 3.23A for an SDS-PAGE). As shown in Figure 3.23D, the incremental addition of PMFT resulted in a stoichiometric decrease in absorbance at 340 nm, suggesting that PMFT is mediating the oxidation of NADH to NAD+ on *Ms*CDH. It should be noted that we did not observe the same oxidative activity toward free NADH (Figure 3.23C). Consequently, this implies that PMFT is biologically active with *Ms*CDH and potentially with other mycofactocin-dependent dehydrogenases.

After finding that PMFT was capable of oxidizing NADH bound to a mycofactocin-dependent dehydrogenase, we next sought to couple PMFT to the *Ms*CDH-catalyzed oxidation of carveol. To do so, *Ms*CDH was incubated with excess amounts of carveol and PMFT in a stoichiometric ratio. The subsequent reaction mixture was analyzed by HPLC (Figure 3.23E). We observed a change in the retention time of PMFT in the chromatogram, from ~15.8 to ~14.9 min, with an accompanying change in the UV−visible spectrum (Figure 3.24 A and B). In addition, we observed a concomitant appearance of new species at a retention time of ~19.5 min corresponding to carvone (Figure 3.23F). We expected that the species at ~14.9 min is the reduced form of PMFT (PMFTH$_2$). To confirm this hypothesis, we analyzed the HPLC-purified species by HRMS. As expected, HRMS analysis (Figure 3.24C) of the ~14.9 min species provided an *m/z* = 236.1267, which is consistent with the *m/z* of PMFTH$_2$ ($C_{13}H_{17}NO_3$, theoretical $[M + H]^+$ *m/z* = 236.1281). Taken together, the observation that PMFT oxidizes the

NADH on *Ms*CDH, that PMFT facilitates *Ms*CDH carveol oxidation activity, and a mass consistent with $PMFTH_2$ provides compelling evidence that PMFT is a redox cofactor.

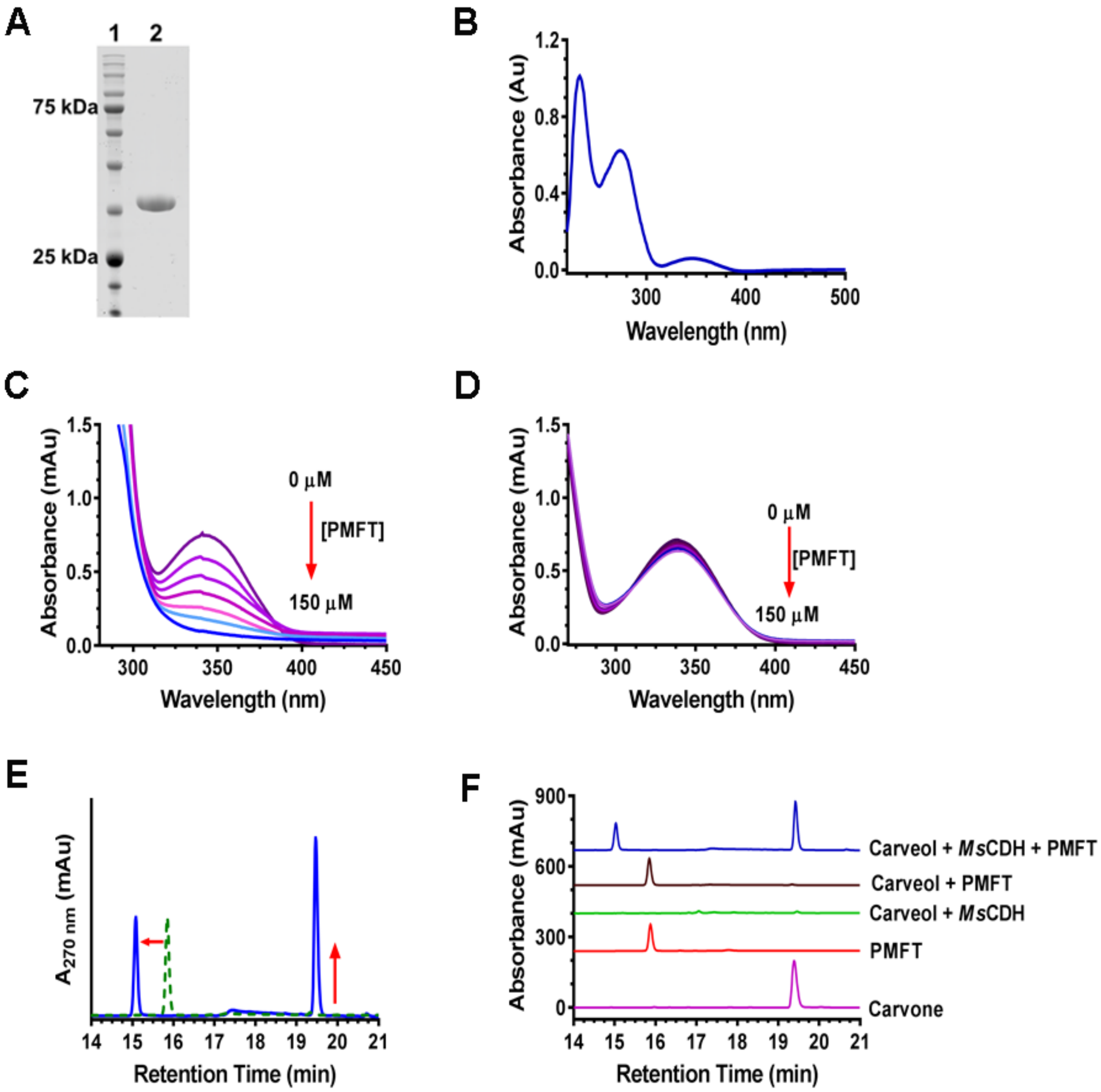


Figure 3.23 Analysis of MsCDH activity with carveol in the presence of PMFT. (A) An SDS-PAGE of *Ms*CDH shows the protein was purified to > 95% purity with a migration molecular weight of ~36 kDa (Lane 1 contains GoldBio BLUEstain Protein Ladder and Lane 2 contains *Ms*CDH protein). (B) UV-Vis spectrum of *Ms*CDH shows an absorbance at 340 nm suggesting that the protein contains NADH. (C) Overlaid UV−visible spectra showing the oxidation of *Ms*CDH-bound NADH by PMFT. (D) Control scanning spectroscopic experiment showing that PMFT does not oxidize free NADH in solution. (E) HPLC analysis demonstrates that *Ms*CDH is active towards carveol in the presence of PMFT (green). The chromatogram for PMFT is shown in green. (F) HPLC chromatograms of *Ms*CDH reactions with carveol(brown) in the presence (blue) and absence (green) of PMFT. Carvone chromatogram was reported at 260 nm while PMFT and PMFTH$_2$ chromatograms were reported at 280 nm.

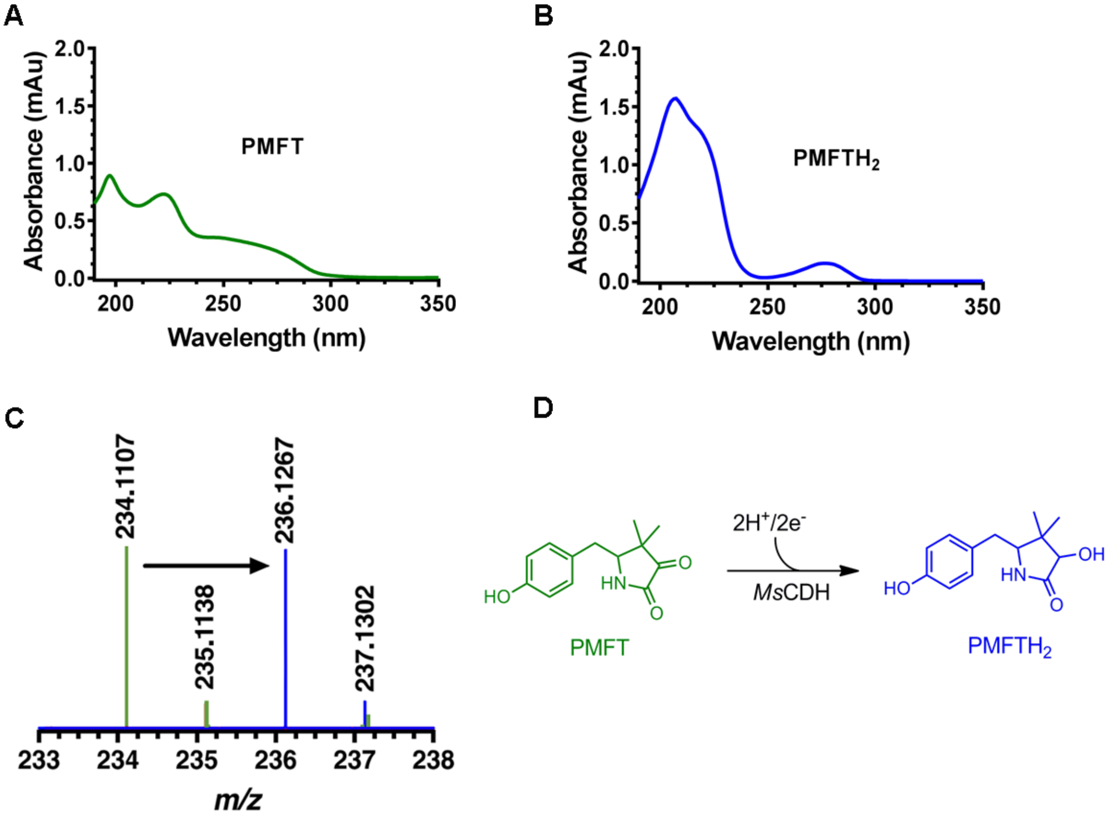


Figure 3.24 Analysis of the reaction product of MsCDH with PMFT. UV-visible spectra of (A) PMFT before *Ms*CDH reaction with carveol and (B) PMFTH$_2$ after carveol transformation reaction. (C) HRMS analysis of the ~14.9 min peak shows an ion with an m/z that is consistent with the mass of PMFTH$_2$ (theoretical $[M + H]^+$ *m/z* = 236.1281). (D) The proposed structure of PMFT following a 2e$^-$/2H$^+$ reduction.

3.18 PMFT is a physiological redox cofactor

To determine the efficacy of PMFT as a redox mediator, the apparent rate constant (k_{obs}) governing the oxidation of *Ms*CDH-bound NADH by PMFT was determined. To do so, time-dependent stopped-flow UV−vis spectrometry experiments were carried out at two PMFT concentrations, in triplicate each, monitoring the oxidation of *Ms*CDH-bound NADH at 340 nm (Figure 3.25A). Data were fitted to a single exponential decay, which yielded a $k_{obs} = 0.8 \pm 0.1\ \text{s}^{-1}$. To directly compare this rate to the rate of oxidation of *Ms*CDH by DCPIP, we carried out similar stopped-flow reactions monitoring the absorbance of DCPIP at 600 nm ($\Delta\varepsilon_{\text{DCPIP 600 nm}} = 20.7\ \text{mM}^{-1}\ \text{cm}^{-1}$, Figure 3.25B) due to spectral overlap with NADH. These reactions were fitted to a linear equation, which provided a $k_{obs} < 0.01\ \text{s}^{-1}$, substantially slower than that observed for PMFT. Control stopped-flow experiments were carried out to ensure that DCPIP was not oxidizing *Ms*CDH on the millisecond time scale (Figure 3.25C). While there is a slim possibility that DCPIP could oxidize *Ms*CDH faster than the deadtime of the instrument (3 ms), the more likely scenario is that DCPIP is nonspecifically interacting with *Ms*CDH, which leads to the slow rate of oxidation. Overall, the >100-fold rate enhancement of PMFT over DCPIP further indicates that PMFT is a physiological mediator for *Ms*CDH.

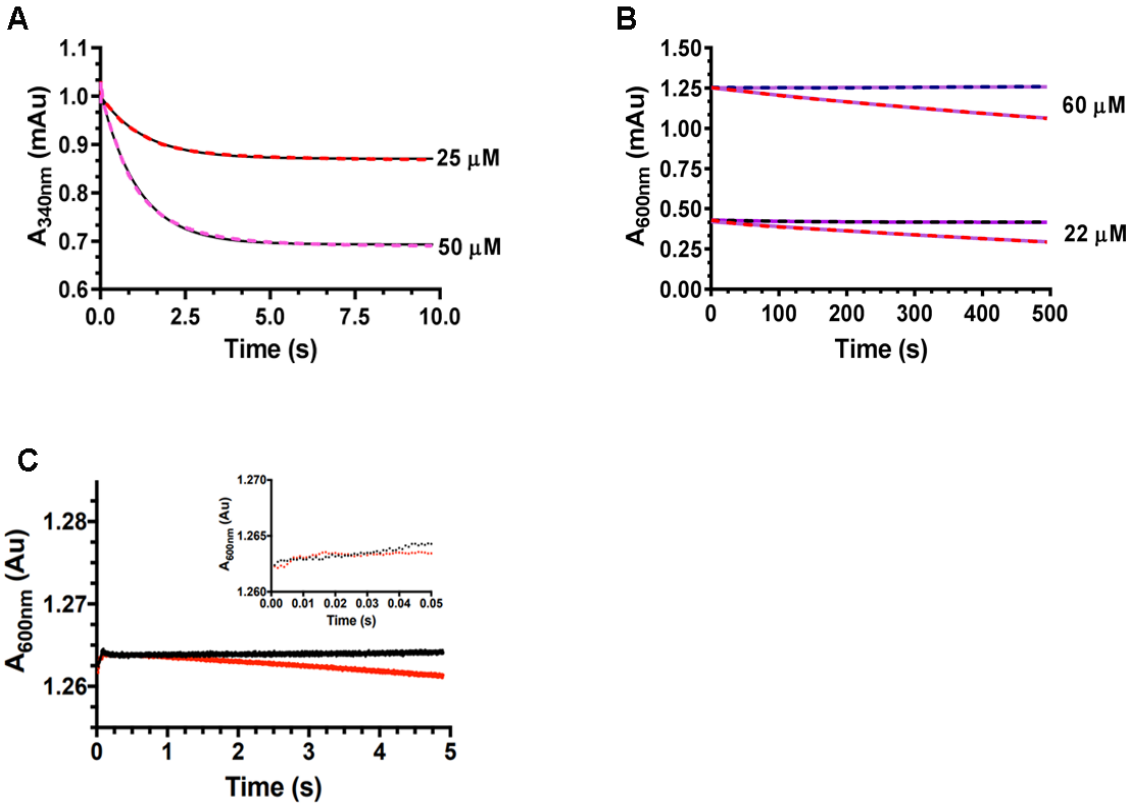


Figure 3.25 Stop flow kinetic analysis of the oxidation of MsCDH bound NADH. (A) Analysis of a single turnover oxidation reaction with 160 μM *Ms*CDH and 25 μM (red dashes) or 50 μM (magenta dashes) PMFT. (B) Analysis of single turnover oxidation reaction with 160 μM *Ms*CDH and 60 μM (cyan) or 22 μM DCPIP (grey). Controls with DCPIP and buffer are shown in blue and black. The reduction of DCPIP was monitored at 600 nm during the reaction, sampling every 100 ms. Each kinetic trace is an average of 3 individual experiments. The averages were fitted to a single exponential decay (magenta and red) to determine the rate constant. (C) Stopped-flow kinetic analysis of a single turnover oxidation reaction with 160 μM *Ms*CDH (red) and 60 μM DCPIP and buffer (black) and 60 μM DCPIP. Data acquisition occurred every 1 ms. The inset of the same reaction shows that steep decrease in absorbance due to rapid oxidation of *Ms*CDH was not present.

CHAPTER FOUR: DISCUSSION

Mycofactocin is a ribosomally synthesized and post translationally modified peptide natural product whose biosynthesis, structure and function, prior to this work, was unknown. Six genes were identified to be responsible for the biosynthesis of mycofactocin and the putative product was postulated to be the second member of the RiPP-derived redox cofactor(169). This assumption was solely based on bioinformatic studies that concluded that mycofactocin associated dehydrogenases might require mycofactocin for catalytic turnover and that the pathway appeared reminiscent to that of PQQ (169). To validate this hypothesis, research efforts were aimed at in vitro reconstitution of each enzyme in the pathway to gain insights into their role and functions during mycofactocin biosynthesis. Within a short period of time, two independent studies, one by our lab demonstrated that MftC catalyzed the oxidative decarboxylation of the C-terminal tyrosine of MftA as the first step in the biosynthesis of mycofactocin (74, 98). A subsequent study of the gene products of the pathway concluded that MftE hydrolyzed the last two residues of the decarboxylated MftA (185). However, a further mechanistic study by our lab demonstrated that MftC catalyzed an additional redox neutral cyclization step by converting the decarboxylated product to a lactam moiety (99).

MftC is a member of the RS-SPASM protein family. Proteins within this family are known to bind two auxiliary [Fe-4S] clusters. However, there is no clear understanding of the roles of these clusters in radical SAM chemistry. While recent deposition of crystal structures of some RS-SPASM enzymes (e.g., anSME, PqqE, CteB, and SuiB) into the protein data base has provided some insights into the possible function of these auxiliary clusters, the rules by which to assign their function have not been established. For instance, in the biosynthesis of thermocellum, the RS-SPASM enzyme CteB was shown to install a thioether bond on the precursor peptide CteA. Both auxiliary clusters of CteB are in [4Fe-4S] state however, the Aux I cluster was found to be ligated by three cysteines with an open coordination site for substrate binding and orientation during catalysis (83). Another example is the RS-SPASM enzyme PqqE which is involved in the biosynthesis of PQQ. PqqE was shown to install a carbon-carbon crosslink on the precursor peptide PqqA. Whereas both auxiliary clusters of PqqE are ligated by eight conserved cysteine residues, the Aux I and Aux II clusters were shown to be in different configurations of [2Fe-2S] and [4Fe-4S] respectively and their exact role in catalysis has yet to determined. A third scenario can be found in anSME, a RS-SPASM protein involved in the maturation of Cys-type or Ser-type sulfatase. Unlike CteB and PqqE, both auxiliary clusters in anSME are fully ligated and in the [4Fe-4S] configuration. However, the precise catalytic role of these clusters in anSME is yet to be fully understood (117).

In this dissertation, we have provided additional information regarding the roles of the Aux I and Aux II clusters in the RS-SPASM family. Similar to QhpD, PqqE, SCIFF, and AlbA (74, 76, 119, 191), the Aux I and Aux II clusters in MftC are required for modification of MftA but not for SAM cleavage. This may suggest that the Aux I and II clusters participate directly in the decarboxylation and crosslinking reactions on MftA by either binding the peptide, transferring electron(s) during catalysis, or a combination of both. Our bioinformatic and mutational analysis indicated that the Aux I cluster may be coordinated by at least three Cys residues (251, 269, and 323), potentially leaving an open coordination site. An open coordination site on the Aux I cluster would be indicative of a substrate binding role similar to that observed in CteB (83). However, since Cys258 is conserved, despite the little catalytic impact observed for the mutant and since guiding structural information for MftC is lacking, this proposal is limited.

The ability of MftC to catalyze two distinct chemistries: an oxidative and redox neutral reaction, in the same active site is very interesting. This chemistry is unique to the RS-SPASM family and suggests that MftC has the means to modulate the electrochemical environment within its active site for each catalytic step. The resting state reduction potential for the Aux I cluster is similar to what was measured for the same cluster in SCIFF maturase (193). This similarity suggests that the midpoint potential of the Aux I cluster in MftC may be modulated through substrate interaction or active site rearrangement to accommodate both oxidative and redox neutral chemistries. Unlike SCIFF, MftC catalyzes two distinct chemistries on two electrostatically different compounds. In the first step, MftC catalyzes the oxidative decarboxylation of the MftA peptide, resulting in the α/β unsaturated bond on MftA** (Figure 1.21). Since the C-terminus of MftA is presumably deprotonated, the loss of the carboxylate would also result in the loss of a negative charge in the active site, which could influence the reduction potential of the Aux I cluster. Further electrochemical investigations of MftC in the presence of each substrate will be required to confirm this hypothesis. Although the potential role of the Aux I cluster in MftC catalysis has been provided, the role of the Aux II is yet to be determined. Despite behaving as two redox active centers, the redox potentials of the Aux I and Aux II clusters are very similar: within ~ -50 mV apart. This suggests that there is not a steep barrier for the reversible electron transfer between the two clusters. Indeed, the calculated difference in Gibbs free energy between the two clusters was found to be ~4.8 kJ/mol, which could allow for the reversible electron transfer between these clusters during MftC catalysis. These results together with the evidence that the Aux II cluster is required for catalysis suggests that the Aux II cluster may participate in electron shuttling in and out of the active site.

A common feature in RiPP biosynthesis is the processing of the functionalized peptide by peptidases after modifications have been imparted. This step generally occurs as a mechanism to cleave the leader peptide and liberate the biologically functional RiPP product. An example of this mechanism can be found in subtilosin biosynthesis where AlbE or AlbF hydrolyze the leader peptide off the precursor peptide, SboA after it has undergone chemical transformations by AlbA (76, 207). An alternate scenario can be found in the biosynthesis of PQQ (Figure 1.16), where the two component protease, PqqF/G work in tandem to hydrolyze off the leader peptide on PqqA after chemical transformations by PqqE (154, 157, 159). This mechanism liberates the biologically relevant small molecule for further processing. When looking at the mycofactocin biosynthetic pathway through the lens of subtilosin or PQQ biosynthesis, one would expect MftE to hydrolyze the biologically relevant moiety from the precursor peptide. Indeed, previous work demonstrated that MftE hydrolyzes the last two residues from MftA* (185). This finding was solely based UHPLC, LCMS and derivatization assays and it was until later that our lab demonstrated that two products are formed from MftC catalysis (99). The two reaction products of MftC, MftA** and MftA* are isomers and so are the hydrolyzed products of MftE reaction.

In this dissertation, we clarified this conundrum by providing additional information regarding the hydrolysis of MftE. Similar to PqqF/G and AlbE/F, our experimental findings demonstrate that MftE is a metalloprotease that selectively hydrolyzes MftA* to form MftA (1-28) and the biologically relevant small molecule AHDP. However, unlike other metalloproteases that are capable of multiple turnovers, MftE catalyzes a single turnover reaction in our hands, suggesting that we a missing an important cofactor in our reaction. It should also be noted that MftE belongs to a protein fold family that is different from that of PqqF/G and AlbE/F. Whereas PqqF/G and AlbE/F are both Zn^{2+} metalloproteases belonging to the MEROPs peptidase M16 family, MftE on the other hand belongs to the peptidase amidohydrolase family of proteins which possess a binuclear metal binding center in their active site (159, 185). MftE was found to bind one mol equivalent of Fe^{2+} and mol equivalent of Zn^{2+} however, Fe^{2+} was observed to have the most catalytic impact on MftE hydrolysis, suggesting that zinc may be playing an auxiliary role. Nevertheless, guiding structural information for MftE and additional kinetic assays are required to adequately determine the role of both Fe^{2+} and Zn^{2+} in MftE hydrolysis reaction.

A bioinformatic study in 2011 suggested that mycofactocin could function as a redox cofactor, just like its brethren PQQ (169). A subsequent saturation transfer difference (STD) NMR study also concluded that mycofactocin is a putative redox cofactor utilized by mycofactocin-associated short chain dehydrogenases for multiple catalytic turnovers (182). This proposal is further corroborated by a recent *in vivo* expression studies that demonstrated that mycofactocin genes are required for assimilation of primary alcohols in the model organism *M. smegmatis* mc^{2}155 (179, 180). While these findings have combined to suggest that mycofactocin is a redox cofactor, no direct evidence has been provided to date. In this dissertation, we have demonstrated that AHDP is an active substrate for MftD. We combined *in vitro* activity assays with HRMS and NMR structural workup to provide conclusive evidence that MftD catalyzes the oxidative deamination of AHDP to form the α-keto moiety called premycofactocin (PMFT). The formation of an α-keto moiety on PMFT led us to hypothesize that mycofactocin could be redox active. To probe for redox activity, we used cyclic voltammetry to measure the midpoint potential of PMFT. Indeed, the cyclic voltammetry measurement showed a single signal with a midpoint potential of ~255 mV. This midpoint potential is comparable to that of PQQ (-240 mV), a peptide-derived redox cofactor utilized by methanol and glucose dehydrogenases for multiple catalytic turnover (142, 145). This finding led us to further investigate whether mycofactocin-associated short chain dehydrogenases could directly recruit mycofactocin for catalytic turnover as postulated by Daniel Haft (169, 182). By combining activity assays with spectroscopic experiments, we provided direct evidence that PMFT oxidizes the NADH bound to *Ms*CDH, a mycofactocin-associated short chain dehydrogenase. *Ms*CDH was found to cycle PMFT for multiple catalytic turnovers by transferring a $2H^+/2e^-$ onto PMFT to form $PMFTH_2$, demonstrating that PMFT undergoes a 2e−/2H+ reduction by *Ms*CDH.

These findings have coalesced to provide the first direct evidence that mycofactocin is a biologically active RiPP-derived redox cofactor utilized by mycofactocin-associated short chain dehydrogenases for multiple catalytic turnover. In addition, these findings continue to validate Haft's hypothesis in 2011 that mycofactocin is the second member of RiPP-derived redox cofactors.

CHAPTER FIVE: SUMMARY

In summary, the work presented in this dissertation has shed light on the structure, function and chemistries of each enzyme involved in the biosynthesis of mycofactocin. To begin with, we provided additional information regarding the roles of the auxiliary clusters in MftC, a unique RS-SPASM protein capable of catalyzing oxidative and redox neutral reactions in the same active site. In addition, we have clarified how mycofactocin is made by conclusively demonstrating that MftA*, the final product of MftC reaction is the precursor for mycofactocin biosynthesis. This specificity is dictated by the peptidase MftE. Lastly, we demonstrated for the first time that MftD catalyzes the oxidative deamination of AHDP, forming the redox-active center in PMFT. We measured the midpoint potential of PMFT and provided the first direct evidence it is a biologically active redox cofactor utilized by mycofactocin-associated short chain dehydrogenases. Taken together, we provide the most direct evidence that mycofactocin is the second member of the RiPP-derived redox cofactor family in addition to the sole quintessential member, PQQ.

REFERENCES

1. Baker, D. D., Chu, M., Oza, U., and Rajgarhia, V. (2007) The value of natural products to future pharmaceutical discovery. *Nat. Prod. Rep.* **24**, 1225
2. Newman, D. J., and Cragg, G. M. (2012) Natural Products As Sources of New Drugs over the 30 Years from 1981 to 2010. *J. Nat. Prod.* **75**, 311–335
3. Fischbach, M. A., Walsh, C. T., and Clardy, J. (2008) The evolution of gene collectives: How natural selection drives chemical innovation. *Proc. Natl. Acad. Sci.* **105**, 4601-4608.
4. Shen, B. (2015) A New Golden Age of Natural Products Drug Discovery. *Cell*. **163**, 1297–1300
5. Li, G., and Lou, H. X. (2018) Strategies to diversify natural products for drug discovery. *Med. Res. Rev.* **38**, 1255–1294
6. Newman, D. J., and Cragg, G. M. (2016) Natural Products as Sources of New Drugs from 1981 to 2014. *J. Nat. Prod.* **79**, 629–661
7. Rodrigues, T., Reker, D., Schneider, P., and Schneider, G. (2016) Counting on natural products for drug design. *Nat. Chem.* **8**, 531–541
8. Chen, Y., Garcia De Lomana, M., Friedrich, N. O., and Kirchmair, J. (2018) Characterization of the Chemical Space of Known and Readily Obtainable Natural Products. *J. Chem. Inf. Model.* **58**, 1518–1532
9. Zhang, X., and Li, S. (2017) Expansion of chemical space for natural products by uncommon P450 reactions. *Nat. Prod. Rep.* **34**, 1061–1089
10. Rosén, J., Gottfries, J., Muresan, S., Backlund, A., and Oprea, T. I. (2009)

Novel Chemical Space Exploration via Natural Products. *J. Med. Chem.* **52**, 1953–1962

11. Harvey, A. L., Edrada-Ebel, R., and Quinn, R. J. (2015) The re-emergence of natural products for drug discovery in the genomics era. *Nat. Rev. Drug Discov.* **14**, 111–129
12. James C. Sacchettini, C. D. P. (1997) Creating Isoprenoid diversity. *Science*. **277**, 1788–1789
13. Tarkowská, D., and Strnad, M. (123AD) Isoprenoid-derived plant signaling molecules: biosynthesis and biological importance. *Planta*. **247**, 1051–1066
14. Schrader, J., and Bohlmann, J. (2015) *Biotechnology of isoprenoids*, (Eds), Springer, Switzerland
15. Withers, S. T., and Keasling, J. D. (2006) Biosynthesis and engineering of isoprenoid small molecules. *Appl. Microbiol. Biotechnol.* **73**, 980–990
16. Peñuelas, J., and Munné-Bosch, S. (2005) Isoprenoids: An evolutionary pool for photoprotection. *Trends Plant Sci.* **10**, 166–169
17. Goldstein, J. L., & Brown, M. S. (1990) Regulation of the mevalonate pathway. *nature.com*. **342**, 425
18. Tholl, D. (2015) Biosynthesis and Biological Functions of Terpenoids in Plants, pp. 63–106, Springer, Cham, 10.1007/10_2014_295
19. Huang, M., Lu, J., Huang, M., Bao, J., Chen, X., and Wang, Y. (2012) Terpenoids : Natural products for cancer therapy Terpenoids : natural products for cancer therapy. *Expert Opin. Investig. Drugs*. **21**, 1801-1818.

20. Wang, Q., Quan, S., and Xiao, H. (2019) Towards efficient terpenoid biosynthesis: manipulating IPP and DMAPP supply. *Bioresour. Bioprocess.* **6**, 6

21. Goto, T., Takahashi, N., Hirai, S., and Kawada, T. (2010) Various Terpenoids Derived from Herbal and Dietary Plants Function as PPAR Modulators and Regulate Carbohydrate and Lipid Metabolism. *PPAR Res.* 10.1155/2010/483958

22. Cushnie, T. P. T., Cushnie, B., and Lamb, A. J. (2014) Alkaloids: An overview of their antibacterial, antibiotic-enhancing and antivirulence activities. *Int. J. Antimicrob. Agents*. **44**, 377–386

23. Dewick, P. (2002) *Medicinal natural products: a biosynthetic approach*, John Wiley & Sons

24. Roy, A. (2017) A Review on the Alkaloids an Important Therapeutic Compound from Plants. *Int. J. Plant Biotechnol.* **3**, 1–9

25. Kaur, R., and Rev, S. A. (2015) Alkaloids-important therapeutic secondary metabolites of plant origin. *J. Crit. Rev.* **2**, 1–8

26. O'Connor, S. E. (2012) Alkaloids. in *Natural Products in Chemical Biology*, pp. 209–237, John Wiley & Sons, Inc., Hoboken, NJ, USA, 10.1002/9781118391815.ch9

27. Roberts, M. (2013) *Alkaloids: biochemistry, ecology, and medicinal applications*, Ed, Springer Science & Business Media

28. Lahlou, M. (2013) The Success of Natural Products in Drug Discovery.

Pharmacol. Pharm. **4**, 17–31

29. Kuščer, E., Raspor, P., and Petković, H. (2005) Rational design of polyketide natural products. *Food Technol. Biotechnol.* **43**, 403–410
30. Mcdaniel, R., Welch, M., and Hutchinson, C. R. (2005) Genetic Approaches to Polyketide Antibiotics. 1. *Chem. Rev.* 10.1021/cr0301189
31. Hopwood, D. A. (1997) Genetic Contributions to Understanding Polyketide Synthases. *Chem. Rev.* **97**, 2465–2498
32. Staunton, J., and Weissman, K. J. (2001) Polyketide biosynthesis: A millennium review. *Nat. Prod. Rep.* **18**, 380–416
33. Khosla, C. (1997) Harnessing the biosynthetic potential of modular polyketide synthases. *Chem. Rev.* **97**, 2577–2590
34. Kittendorf, J. D., and Sherman, D. H. (2006) Developing tools for engineering hybrid polyketide synthetic pathways. *Curr. Opin. Biotechnol.* **17**, 597–605
35. Hopwood, D. A., and Sherman, D. H. (1990) Molecular Genetics of Polyketides and its Comparison to Fatty Acid Biosynthesis. *Annu. Rev. Genet.* **24**, 37–62
36. Malpartida, F., Nature, D. H.-, and 1984, undefined (1984) Molecular cloning of the whole biosynthetic pathway of a Streptomyces antibiotic and its expression in a heterologous host. *Nature*. **309**, 462
37. Cheng, Y. Q., Tang, G. L., and Shen, B. (2003) Type I polyketide synthase requiring a discrete acyltransferase for polyketide biosynthesis. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 3149–3154

38. Mootz, H. D., Schwarzer, D., and Marahiel, M. A. (2002) Ways of assembling complex natural products on modular nonribosomal peptide synthetases. *ChemBioChem*. **3**, 490–504

39. von Döhren, H., Keller, U., Vater, J., and Zocher, R. (1997) Multifunctional Peptide Synthetases. *Chem. Rev.* **97**, 2675–2706

40. Marahiel, M. A., Stachelhaus, T., & Mootz, H. D. (1997) Modular peptide synthetases involved in nonribosomal peptide synthesis. *Chem. Rev.* **79**, 2651–2674

41. van Wageningen, A.A., Kirkpatrick, P.N., Williams, D.H., Harris, B.R., Kershaw, J.K., Lennard, N.J., Jones, M., Jones, S.J. and Solenberg, P. J. (1988) Sequencing and analysis of genes involved in the biosynthesis of a vancomycin group antibiotic. *Chem. Biol.* **5**, 155–162

42. Weber, G., Schorgendorfer, K., Schneider-Scherzer, E., and Leitner, E. (1994) The peptide synthetase catalyzing cyclosporine production in Tolypocladium niveum is encoded by a giant 45.8-kilobase open reading frame. *Curr. Genet.* **26**, 120–125

43. Healy, F. G., Wach, M., Krasnoff, S. B., Gibson, D. M., and Loria, R. (2000) The txtAB genes of the plant pathogen Streptomyces acidiscabies encode a peptide synthetase required for phytotoxin thaxtomin A production and pathogenicity. *Mol. Microbiol.* **38**, 794–804

44. Keating, T. A., Marshall, C. G., and Walsh, C. T. (2000) Reconstitution and Characterization of the Vibrio cholerae Vibrobactin Synthetase from VibB,

VibE, VibE, and VibH. *Biochemistry*. **39**, 15522–15530

45. Schwarzer, D., and Marahiel, M. A. (2001) Multimodular biocatalysts for natural product assembly. *Naturwissenschaften*. **88**, 93–101

46. Hetrick, K. J., and van der Donk, W. A. (2017) Ribosomally synthesized and post-translationally modified peptide natural product discovery in the genomic era. *Curr. Opin. Chem. Biol.* **38**, 36–44

47. Arnison, P. G., Bibb, M. J., Bierbaum, G., Bowers, A. A., Bugni, T. S., Bulaj, G., Camarero, J. A., Campopiano, D. J., Challis, G. L., Clardy, J., Cotter, P. D., Craik, D. J., Dawson, M., Dittmann, E., Donadio, S., Dorrestein, P. C., Entian, K. D., Fischbach, M. A., Garavelli, J. S., Göransson, U., Gruber, C. W., Haft, D. H., Hemscheidt, T. K., Hertweck, C., Hill, C., Horswill, A. R., Jaspars, M., Kelly, W. L., Klinman, J. P., Kuipers, O. P., Link, A. J., Liu, W., Marahiel, M. A., Mitchell, D. A., Moll, G. N., Moore, B. S., Müller, R., Nair, S. K., Nes, I. F., Norris, G. E., Olivera, B. M., Onaka, H., Patchett, M. L., Piel, J., Reaney, M. J. T., Rebuffat, S., Ross, R. P., Sahl, H. G., Schmidt, E. W., Selsted, M. E., Severinov, K., Shen, B., Sivonen, K., Smith, L., Stein, T., Süssmuth, R. D., Tagg, J. R., Tang, G. L., Truman, A. W., Vederas, J. C., Walsh, C. T., Walton, J. D., Wenzel, S. C., Willey, J. M., and Van Der Donk, W. A. (2013) Ribosomally synthesized and post-translationally modified peptide natural products: Overview and recommendations for a universal nomenclature. *Nat. Prod. Rep.* **30**, 108–160

48. Hudson, G. A., and Mitchell, D. A. (2018) RiPP antibiotics: biosynthesis and

engineering potential. *Curr. Opin. Microbiol.* **45**, 61–69

49. Ortega, M. A., and van der Donk, W. A. (2016) New Insights into the Biosynthetic Logic of Ribosomally Synthesized and Post-translationally Modified Peptide Natural Products. *Cell Chem. Biol.* **23**, 31–44

50. Goto, Y., and Suga, H. (2018) Engineering of RiPP pathways for the production of artificial peptides bearing various non-proteinogenic structures. *Curr. Opin. Chem. Biol.* **46**, 82–90

51. Strieker, M., Tanović, A., and Marahiel, M. A. (2010) Nonribosomal peptide synthetases: structures and dynamics. *Curr. Opin. Struct. Biol.* **20**, 234–240

52. Santos-Aberturas, J., Chandra, G., Frattaruolo, L., Lacret, R., Pham, T. H., Vior, N. M., Eyles, T. H., and Truman, A. W. (2019) Uncovering the unexplored diversity of thioamidated ribosomal peptides in Actinobacteria using the RiPPER genome mining tool. *Nucleic Acids Res.* **47**, 4624–4637

53. Dang, T., and Süssmuth, R. D. (2017) Bioactive Peptide Natural Products as Lead Structures for Medicinal Use. *Acc. Chem. Res.* **50**, 1566–1576

54. Férir, G., Petrova, M. I., Andrei, G., Huskens, D., Hoorelbeke, B., Snoeck, R., Vanderleyden, J., Balzarini, J., Bartoschek, S., Brönstrup, M., Süssmuth, R. D., and Schols, D. (2013) The Lantibiotic Peptide Labyrinthopeptin A1 Demonstrates Broad Anti-HIV and Anti-HSV Activity with Potential for Microbicidal Applications. *PLoS One*. **8**, e64010

55. Hoffmann, J., Wray, V., Müller, R., Wink, J., Gerth, K., Volz, C., Jansen, R., Stadler, M., Bernecker, S., and Mohr, K. I. (2015) Pinensins: The First

Antifungal Lantibiotics. *Angew. Chemie Int. Ed.* **54**, 11254–11258

56. Zheng, Q., Wang, Q., Wang, S., Wu, J., Gao, Q., and Liu, W. (2015) Thiopeptide Antibiotics Exhibit a Dual Mode of Action against Intracellular Pathogens by Affecting Both Host and Microbe. *Chem. Biol.* **22**, 1002–1007
57. Kjaerulff, L., Sikandar, A., Zaburannyi, N., Adam, S., Herrmann, J., Koehnke, J., and Müller, R. (2017) Thioholgamides: Thioamide-Containing Cytotoxic RiPP Natural Products. *ACS Chem. Biol.* **12**, 2837–2841
58. Ibrahim, M., Guillot, A., Wessner, F., Algaron, F., Besset, C., Courtin, P., Gardan, R., and Monnet, V. (2007) Control of the transcription of a short gene encoding a cyclic peptide in Streptococcus thermophilus: A new quorum-sensing system? *J. Bacteriol.* **189**, 8845–8854
59. Gardan, R., Besset, C., Guillot, A., Gitton, C., and Monnet, V. (2009) The oligopeptide transport system is essential for the development of natural competence in Streptococcus thermophilus strain LMD-9. *J. Bacteriol.* **191**, 4647–4655
60. Davidson, V. L. (2005) Structure and mechanism of tryptophylquinone enzymes. *Bioorg. Chem.* **33**, 159–170
61. Davidson, V. L. (2011) Generation of protein-derived redox cofactors by posttranslational modification. *Mol. Biosyst.* **7**, 29–37
62. Klinman, J. P., and Bonnot, F. (2014) Intrigues and intricacies of the biosynthetic pathways for the enzymatic quinocofactors: PQQ, TTQ, CTQ, TPQ, and LTQ. *Chem. Rev.* **114**, 4343–4365

63. Yukl, E. T., and Wilmot, C. M. (2012) Cofactor biosynthesis through protein post-translational modification. *Curr. Opin. Chem. Biol.* **16**, 54–59

64. Oman, T. J., and Van Der Donk, W. A. (2010) Follow the leader: the use of leader peptides to guide natural product biosynthesis. *Nat. Chem. Biol.* **6**, 9

65. Flühe, L., and Marahiel, M. A. (2013) Radical S-adenosylmethionine enzyme catalyzed thioether bond formation in sactipeptide biosynthesis. *Curr. Opin. Chem. Biol.* **17**, 605–612

66. Rea, M. C., Sit, C. S., Clayton, E., O'connor, P. M., Whittal, R. M., Zheng, J., Vederas, J. C., Ross, R. P., and Hill, C. (2010) Thuricin CD, a posttranslationally modified bacteriocin with a narrow spectrum of activity against Clostridium difficile. *Proc. Natl. Acad. Sci.* **107**, 9352–9357

67. Rea, M. C., Dobson, A., O'sullivan, O., Crispie, F., Fouhy, F., Cotter, P. D., Shanahan, F., Kiely, B., Hill, C., and Ross, R. P. (2010) Effect of broad-and narrow-spectrum antimicrobials on Clostridium difficile and microbial diversity in a model of the distal colon. *Proc. Natl. Acad. Sci.* **108**, 4639–4644

68. Engelberg-Kulka, H., and Hazan, R. (2003) Cannibals defy starvation and avoid sporulation. *science.* **301**, 467–468

69. González-Pastor, J., Hobbs, E., and Losick, R. (2003) Cannibalism by sporulating bacteria. *sciencemag*. **301**, 510–513

70. Sutyak, K. E., Anderson, R. A., Dover, S. E., Feathergill, K. A., Aroutcheva, A. A., Faro, S., and Chikindas, M. L. (2008) Spermicidal Activity of the Safe

Natural Antimicrobial Peptide Subtilosin. *Infect. Dis. Obstet. Gynecol.* **2008**, 1–6

71. Silkin, L., Hamza, S., Kaufman, S., Cobb, S. L., and Vederas, J. C. (2008) Spermicidal bacteriocins: Lacticin 3147 and subtilosin A. *Bioorg. Med. Chem. Lett.* **18**, 3103–3106
72. Benjdia, A., Guillot, A., Lefranc, B., Vaudry, H., Leprince, J., and Berteau, O. (2016) Thioether bond formation by SPASM domain radical SAM enzymes: CαH-atom abstraction in subtilosin A biosynthesis. *Chem. Commun.* **52**, 6249–6252
73. Repka, L. M., Chekan, J. R., Nair, S. K., and Van Der Donk, W. A. (2017) Mechanistic Understanding of Lanthipeptide Biosynthetic Enzymes. *Chem. Rev.* **117**, 5457–5520
74. Bruender, N. A., Wilcoxen, J., Britt, R. D., and Bandarian, V. (2016) Biochemical and Spectroscopic Characterization of a Radical S-Adenosyl- l -methionine Enzyme Involved in the Formation of a Peptide Thioether Cross-Link. *Biochemistry*. **55**, 2122–2134
75. Wieckowski, B. M., Hegemann, J. D., Mielcarek, A., Boss, L., Burghaus, O., and Marahiel, M. A. (2015) The PqqD homologous domain of the radical SAM enzyme ThnB is required for thioether bond formation during thurincin H maturation. *FEBS Lett.* **589**, 1802–1806
76. Flühe, L., Knappe, T. A., Gattner, M. J., Schäfer, A., Burghaus, O., Linne, U., and Marahiel, M. A. (2012) The radical SAM enzyme AlbA catalyzes

thioether bond formation in subtilosin A. *Nat. Chem. Biol.* **8**, 350–357

77. Flühe, L., Burghaus, O., Wieckowski, B. M., Giessen, T. W., Linne, U., and Marahiel, M. A. (2013) Two [4Fe-4S] clusters containing radical SAM enzyme SkfB catalyze thioether bond formation during the maturation of the sporulation killing factor. *J. Am. Chem. Soc.* **135**, 959–962
78. Bruender, N. A., and Bandarian, V. (2016) SkfB Abstracts a Hydrogen Atom from C α on SkfA To Initiate Thioether Cross-Link Formation. *Biochemistry*. **55**, 4131–4134
79. Sit, C. S., McKay, R. T., Hill, C., Ross, R. P., and Vederas, J. C. (2011) The 3D Structure of Thuricin CD, a Two-Component Bacteriocin with Cysteine Sulfur to α-Carbon Cross-links. *J. Am. Chem. Soc.* **133**, 7680–7683
80. Murphy, K., O'sullivan, O., Rea, M. C., Cotter, P. D., and Ross, R. P. (2011) Genome Mining for Radical SAM Protein Determinants Reveals Multiple Sactibiotic-Like Gene Clusters. *PLoS One*. **6**, 20852
81. Hudson, G. A., Burkhart, B. J., Dicaprio, A. J., Schwalen, C. J., Kille, B., Pogorelov, T. V, and Mitchell, D. A. (2019) Bioinformatic Mapping of Radical S-Adenosylmethionine-Dependent Ribosomally Synthesized and Post-Translationally Modified Peptides Identifies New Cα, Cβ, and Cγ-Linked Thioether-Containing Peptides. *ournal Am. Chem. Soc.* **141**, 8228–8238
82. Haft, D. H., and Kumar Basu, M. (2011) Biological Systems Discovery In Silico: Radical S-Adenosylmethionine Protein Families and Their Target Peptides for Posttranslational Modification. *J. Bacteriol.* **193**, 2745–2755

83. Grove, T. L., Himes, P. M., Hwang, S., Yumerefendi, H., Bonanno, J. B., Kuhlman, B., Almo, S. C., and Bowers, A. A. (2017) Structural Insights into Thioether Bond Formation in the Biosynthesis of Sactipeptides. *J. Am. Chem. Soc.* **139**, 11734–11744

84. Precord, T. W., Mahanta, N., and Mitchell, D. A. (2019) Reconstitution and Substrate Specificity of the Thioether-Forming Radical S -Adenosylmethionine Enzyme in Freyrasin Biosynthesis . *ACS Chem. Biol.* 10.1021/acschembio.9b00457

85. Grell, T. A. J., Kincannon, W. M., Bruender, N. A., Blaesi, E. J., Krebs, C., Bandarian, V., and Drennan, C. L. (2018) Structural and spectroscopic analyses of the sporulation killing factor biosynthetic enzyme SkfB, a bacterial AdoMet radical sactisynthase. *J. Biol. Chem.* **293**, 17349–17361

86. Schramma, K. R., and Seyedsayamdost, M. R. (2017) Lysine-Tryptophan-Crosslinked Peptides Produced by Radical SAM Enzymes in Pathogenic Streptococci. *ACS Chem. Biol.* **12**, 922–927

87. Davis, K. M., Schramma, K. R., Hansen, W. A., Bacik, J. P., Khare, S. D., Seyedsayamdost, M. R., and Ando, N. (2017) Structures of the peptide-modifying radical SAM enzyme SuiB elucidate the basis of substrate recognition. *Proc. Natl. Acad. Sci.* **114**, 10420–10425

88. Schramma, K. R., and Seyedsayamdost, M. R. (2017) Lysine-Tryptophan-Crosslinked Peptides Produced by Radical SAM Enzymes in Pathogenic Streptococci. *ACS Chem. Biol*. **12**, 922–927

89. Schramma, K. R., Forneris, C. C., Caruso, A., and Seyedsayamdost, M. R. (2018) Mechanistic Investigations of Lysine−Tryptophan Cross-Link Formation Catalyzed by Streptococcal Radical S-Adenosylmethionine Enzymes. *Biochemistry*. **57**, 461-468.

90. Schramma, K. R., Bushin, L. B., and Seyedsayamdost, M. R. (2015) Structure and biosynthesis of a macrocyclic peptide containing an unprecedented lysine-to-tryptophan crosslink. *Nat. Chem.* **7**, 431–437

91. Broderick, J. B., Duffus, B. R., Duschene, K. S., and Shepard, E. M. (2014) Radical S-Adenosylmethionine Enzymes. *Chem. Rev.* **114**, 4229–4317

92. Mehta, A. P., Abdelwahed, S. H., Mahanta, N., Fedoseyenko, D., Philmus, B., Cooper, L. E., Liu, Y., Jhulki, I., Ealick, S. E., and Begley, T. P. (2015) Radical S-Adenosylmethionine (SAM) Enzymes in Cofactor Biosynthesis: A Treasure Trove of Complex Organic Radical Rearrangement Reactions. *J. Biol. Chem.* **290**, 3980–3986

93. Ortega, M. A., Hao, Y., Zhang, Q., Walker, M. C., Van Der Donk, W. A., and Nair, S. K. (2014) Structure and mechanism of the tRNA-dependent lantibiotic dehydratase NisB. *Nature*. **517(7535)**, 509

94. Koehnke, J., Mann, G., Bent, A.F., Ludewig, H., Shirran, S., Botting, C., Lebl, T., Houssen, W.E., Jaspars, M. and Naismith, J. H. (2015) Structural analysis of leader peptide binding enables leader-free cyanobactin processing. *Nat. Chem. Biol.* **11(8)**, 558

95. Burkhart, B. J., Hudson, G. A., Dunbar, K. L., and Mitchell, D. A. (2015) A

Prevalent Peptide-Binding Domain Guides Ribosomal Natural Product Biosynthesis. *Nat Chem Biol*. **11**, 564–570

96. Latham, J. A., Iavarone, A. T., Barr, I., Juthani, P. V., and Klinman, J. P. (2015) PqqD is a novel peptide chaperone that forms a ternary complex with the radical S-adenosylmethionine protein PqqE in the pyrroloquinoline quinone biosynthetic pathway. *J. Biol. Chem.* **290**, 12908–12918
97. Bruender, N. A., and Bandarian, V. (2016) The Radical S-Adenosyl- l -methionine Enzyme MftC Catalyzes an Oxidative Decarboxylation of the C-Terminus of the MftA Peptide. *Biochemistry*. **55**, 2813–2816
98. Khaliullin, B., Aggarwal, P., Bubas, M., Eaton, G. R., Eaton, S. S., and Latham, J. A. (2016) Mycofactocin biosynthesis: modification of the peptide MftA by the radical S-adenosylmethionine protein MftC. *FEBS Lett.* **590**, 2538–2548
99. Khaliullin, B., Ayikpoe, R., Tuttle, M., and Latham, X. J. A. (2017) Mechanistic elucidation of the mycofactocin-biosynthetic radical S -adenosylmethionine protein , MftC. *J. Biol. Chem.* **292**, 13022–13033
100. Vey, J. L., and Drennan, C. L. (2011) Structural Insights into Radical Generation by the Radical SAM Superfamily. *Chem. Rev.* **111.4**, 2487–2506
101. Grell, T. A. J., Goldman, P. J., and Drennan, C. L. (2014) Function of SPASM/Twitch Subfamily. *J. Biol. Chem.* **290**, 3964–3971
102. Dowling, D. P., Vey, J. L., Croft, A. K., and Drennan, C. L. (2012) Structural diversity in the AdoMet radical enzyme superfamily. *Biochim. Biophys. Acta*

- *Proteins Proteomics*. **1824**, 1178–1195

103. Nicolet, Y., Rubach, J. K., Posewitz, M. C., Amara, P., Mathevon, C., Atta, M., Fontecave, M., and Fontecilla-Camps, J. C. (2008) X-ray Structure of the [FeFe]-Hydrogenase Maturase HydE from Thermotoga maritima * □ S The atomic coordinates and structure factors (codes 3CIW and 3CIX) have been deposited in the Protein Data Downloaded from. *J. Biol. Chem.* **283**, 18861

104. Berkovitch, F., Nicolet, Y., Wan, J. T., Jarrett, J. T., and Drennan, C. L. (2004) Crystal structure of biotin synthase, an S-adenosylmethionine-dependent radical enzyme. *Science*. **303**, 76–9

105. Jarrett, J. T. (2005) The novel structure and chemistry of iron–sulfur clusters in the adenosylmethionine-dependent radical enzyme biotin synthase. *Arch. Biochem. Biophys.* **433**, 312–321

106. Chatterjee, A., Li, Y., Zhang, Y., Grove, T. L., Lee, M., Krebs, C., Booker, S. J., Begley, T. P., and Ealick, S. E. (2008) Reconstitution of ThiC in thiamine pyrimidine biosynthesis expands the radical SAM superfamily. *Nat. Chem. Biol.* **4**, 758–765

107. Farrar, C. E., and Jarrett, J. T. (2009) Protein Residues That Control the Reaction Trajectory in S-Adenosylmethionine Radical Enzymes: Mutagenesis of Asparagine 153 and Aspartate 155 in Escherichia coli Biotin Synthase. *Biochemistry*. **48**, 55

108. Layer, G., Heinz, D. W., Jahn, D., and Schubert, W. D. (2004) Structure and function of radical SAM enzymes. *Curr. Opin. Chem. Biol.* **8**, 468–476

109. Nicolet, Y., Amara, P., Mouesca, J. M., & Fontecilla-Camps, J. C. (2009) Unexpected electron transfer mechanism upon AdoMet cleavage in radical SAM proteins. *Proc. Natl. Acad. Sci.* **106(35)**, 14867–14871

110. Davis, K. M., Schramma, K. R., Hansen, W. A., Bacik, J. P., Khare, S. D., Seyedsayamdost, M. R., and Ando, N. (2017) Structures of the peptide-modifying radical SAM enzyme SuiB elucidate the basis of substrate recognition. *Proc. Natl. Acad. Sci.* **114**, 201703663

111. Grell, T. A. J., Goldman, P. J., and Drennan, C. L. (2014) Function of SPASM/Twitch Subfamily. *J. Biol. Chem.* **290**, 3964–3971

112. Goldman, P. J., Grove, T. L., Booker, S. J., and Drennan, C. L. (2013) X-ray analysis of butirosin biosynthetic enzyme BtrN redefines structural motifs for AdoMet radical chemistry. *Proc. Natl. Acad. Sci.* **110**, 15949–15954

113. Grove, T. L., Lee, K.-H., St Clair, J., Krebs, C., and Booker, S. J. (2008) In Vitro Characterization of AtsB, a Radical SAM Formylglycine-Generating Enzyme That Contains Three [4Fe-4S] Clusters. *Biochemistry*. **47.28**, 7523–7538

114. Lanz, N. D., and Booker, S. J. (2015) Auxiliary iron–sulfur cofactors in radical SAM enzymes. *Biochim. Biophys. Acta - Mol. Cell Res.* **1853**, 1316–1334

115. Hänzelmann, P., and Schindelin, H. (2006) Binding of 5-GTP to the C-terminal FeS cluster of the radical S-adenosylmethionine enzyme MoaA provides insights into its mechanism. *Proc. Natl. Acad. Sci.* . **103.18**, 6829–6834

116. Hänzelmann, Petra, and H. S. (2004) Crystal structure of the S-adenosylmethionine-dependent enzyme MoaA and its implications for molybdenum cofactor deficiency in humans. *Proc. Natl. Acad. Sci.* **101**, 12870-12875.

117. Goldman, P. J., Grove, T. L., Sites, L. A., Mclaughlin, M. I., Booker, S. J., and Drennan, C. L. (2013) X-ray structure of an AdoMet radical activase reveals an anaerobic solution for formylglycine posttranslational modification. *Proc. Natl. Acad. Sci.* **110.21**, 8519–8524

118. Benjdia, A., Subramanian, S., Leprince, J., Vaudry, H., Johnson, M. K., and Berteau, O. (2010) Anaerobic sulfatase-maturating enzyme - A mechanistic link with glycyl radical-activating enzymes? *FEBS J.* **277**, 1906–1920

119. Barr, I., Stich, T. A., Gizzi, A. S., Grove, T. L., Bonanno, J. B., Latham, J. A., Chung, T., Wilmot, C. M., Britt, R. D., Almo, S. C., and Klinman, J. P. (2018) X-ray and EPR Characterization of the Auxiliary Fe-S Clusters in the Radical SAM Enzyme PqqE. *Biochemistry*. **57**, 1306–1315

120. Saichana, N., Tanizawa, K., Ueno, H., Pechoušek, J., Novák, P., and Frébortová, J. (2017) Characterization of auxiliary iron–sulfur clusters in a radical S-adenosylmethionine enzyme PqqE from Methylobacterium extorquens AM1. *FEBS Open Bio*. **7**, 1864–1879

121. Grove, T. L., Ahlum, J. H., Qin, R. M., Lanz, N. D., Radle, M. I., Krebs, C., and Booker, S. J. (2013) Further Characterization of Cys-Type and Ser-Type Anaerobic Sulfatase Maturating Enzymes Suggests a Commonality in the

Mechanism of Catalysis. *Biochemistry*. **52**, 2874–2887

122. Benjdia, Alhosna, Jérôme Leprince, Corine Sandstrom, Hubert Vaudry, and O. B. (2009) Mechanistic Investigations of Anaerobic Sulfatase-Maturating Enzyme: Direct Cβ H-Atom Abstraction Catalyzed by a Radical AdoMet Enzyme. *J. Am. Chem. Soc.* **131**, 8348–8349

123. Benjdia, A., Leprince, J., Guillot, A., Vaudry, H., Rabot, S., and Berteau, O. (2007) Anaerobic sulfatase-maturating enzymes: Radical SAM enzymes able to catalyze in vitro sulfatase post-translational modification. *J. Am. Chem. Soc.* **129**, 3462–3463

124. Barr, I., Latham, J. A., Iavarone, A. T., Chantarojsiri, T., Hwang, J. D., and Klinman, J. P. (2016) Demonstration that the radical s-adenosylmethionine (SAM) Enzyme PqqE catalyzes de novo carbon-carbon cross-linking within a peptide substrate PqqA in the presence of the peptide chaperone PqqD. *J. Biol. Chem.* **291**, 8877–8884

125. Davidson, V. L. (2011) Generation of protein-derived redoxcofactors by posttranslational modification. *Mol. BioSyst.* **7**, 29–37

126. Yukl, E. T., and Wilmot, C. M. (2012) Cofactor biosynthesis through protein post-translational modification. *Curr. Opin. Chem. Biol.* **16**, 54–59

127. Mure, M. (1991) Tyrosine-Derived Quinone Cofactors. *Acc. Chem. Res.* **37**, 131–139

128. Davidson, V. L. (2005) Structure and mechanism of tryptophylquinone enzymes. *Bioorg. Chem.* **33**, 159–170

129. Davidson, V. L. (2018) Protein-Derived Cofactors Revisited: Empowering Amino Acid Residues with New Functions. *Biochemistry*. **57**, 3115–3125

130. Kagan, H. M., and Trackman, P. (1991) Properties and Function of Lysyl Oxidase. *Am J Respir Cell Mol Biol*. **5.3**, 206–210

131. Satoh, A., Kim, J.-K., Miyahara, I., Devreese, B., Vandenberghe, I., Hacisalihoglu, A., Okajima, T., Kuroda, I., Adachi, O., Duine, J. A., Van Beeumen, J., Tanizawa, K., and Hirotsu, K. (2002) Crystal structure of quinohemoprotein amine dehydrogenase from Pseudomonas putida Identification of a novel quinone cofactor encaged by multiple thioether cross-bridges. *J. Biol. Chem.* **277**, 2830–2834

132. Datta, S., Mori, Y., Takagi, K., Kawaguchi, K., Chen, Z.-W., Okajima, T., Kuroda, ichi, Ikeda, T., Kano, K., Tanizawa, K., and Scott Mathews, F. (2001) Structure of a quinohemoprotein amine dehydrogenase with an uncommon redox cofactor and highly unusual crosslinking. *Proc. Natl. Acad. Sci.* **98(25)**, 14268–14273

133. Klinman, J. P., & Bonnot, F. (2013) Intrigues and intricacies of the biosynthetic pathways for the enzymatic quinocofactors: PQQ, TTQ, CTQ, TPQ, and LTQ. *Chem. Rev.* **114**, 4343–4365

134. Anthony, C. (2001) Pyrroloquinoline quinone (PQQ) and quinoprotein enzymes. *Antioxidants Redox Signal.* **3**, 757–774

135. Anthony, C., and Zatman, L. J. (1967) The microbial oxidation of methanol: the prosthetic group of the alcohol dehydrogenase of Pseudomonas sp.

M27: a new oxidoreductase prosthetic group. *Biochem. J.* **104**, 960

136. Duine, J. A., Jzn, J. F., & Van Zeeland, J. K. (1979) Glucose dehydrogenase from Acinetobacter calcoaceticus: a “quinoprotein.” *FEBS Lett.* **108**, 443–446

137. Goodwin, P. M., and Anthony, C. (1998) The Biochemistry, Physiology and Genetics of PQQ and PQQ-containing Enzymes. *Adv. Microb. Physiol.* **40**, 1–80

138. Janes, S. M., Mu, D., Wemmer, D., Smith, A. J., Kaur, S., Maltby, D., ... & Klinman, J. P. (1990) A new redox cofactor in eukaryotic enzymes: 6-hydroxydopa at the active site of bovine serum amine oxidase. *sciencemag*. **248**, 981–987

139. Mcintire, W. S., Wemmer, D. E., Chistoserdov, A., and Lidstrom, M. E. (1991) A new cofactor in a prokaryotic enzyme: Tryptophan tryptophylquinone as the redox prosthetic group in methylamine dehydrogenase. *sciencemag*. **252**, 817–824

140. Wang, S.X., Mure, M., Medzihradszky, K.F., Burlingame, A.L., Brown, D.E., Dooley, D.M., Smith, A.J., Kagan, H.M. and Klinman, J. P. (1996) A Crosslinked Cofactor in Lysyl Oxidase : Redox Function for Amino Acid Side Chains. *sciencemag*. **273**, 1078–1084

141. Shen, Y. Q., Bonnot, F., Imsand, E. M., RoseFigura, J. M., Sjölander, K., & Klinman, J. P. (2012) Distribution and Properties of the Genes Encoding the Biosynthesis of the Bacterial Cofactor, Pyrroloquinoline Quinone.

Biochemistry. **51**, 2265–2275

142. Duine, J. A. (1999) The PQQ story. *J. Biosci. Bioeng.* **88**, 231–236

143. Itoh, S., Ohshiro, Y., and Agawa, T. (1986) Reaction of reduced PQQ (PQQH2) and molecular oxygen. *Bull. Chem. Soc. Jpn.* **59**, 1911–1914

144. Saleh, F. S., Rahman, M. R., Okajima, T., Mao, L., and Ohsaka, T. (2011) Determination of formal potential of NADH/NAD+ redox couple and catalytic oxidation of NADH using poly(phenosafranin)-modified carbon electrodes. *Bioelectrochemistry*. **80**, 121–127

145. Tan, S. L. J., Kan, J. M., and Webster, R. D. (2013) Differences in proton-coupled electron-transfer reactions of flavin mononucleotide (FMN) and flavin adenine dinucleotide (FAD) between buffered and unbuffered aqueous solutions. *J. Phys. Chem. B*. **117**, 13755–13766

146. Urban, P. F., and Klinqenberg, M. (1969) Redox Potentials of Ubiquinone and Cytochrome. *Eur. J. Biochem.* **9**, 519–525

147. Emahi, I., Mitchell, M. P., and Baum, D. A. (2017) Electrochemistry of Pyrroloquinoline Quinone (PQQ) on Multi-Walled Carbon Nanotube-Modified Glassy Carbon Electrodes in Biological Buffers. *J. Electrochem. Soc.* **164**, H3097–H3102

148. Yokoyama, K., and Lilla, E. A. (2018) C–C bond forming radical SAM enzymes involved in the construction of carbon skeletons of cofactors and natural products. *Nat. Prod. Rep.* **35**, 660–694

149. Chowanadisai, W., Bauerly, K. A., Tchaparian, E., Wong, A., Cortopassi, G.

A., and Rucker, R. B. (2010) Pyrroloquinoline quinone stimulates mitochondrial biogenesis through cAMP response element-binding protein phosphorylation and increased PGC-1alpha expression. *J. Biol. Chem.* **285**, 142–52

150. Velterop, J. S., Sellink, E., Meulenberg, J. J., David, S., Bulder, I., and Postma, P. W. (1995) Synthesis of pyrroloquinoline quinone in vivo and in vitro and detection of an intermediate in the biosynthetic pathway. *J. Bacteriol.* **177**, 5088–5098

151. Houck, D. R., Hanners, J. L., and Unkefer, C. J. (1988) Biosynthesis of pyrroloquinoline quinone. 1. Identification of biosynthetic precursors using carbon-13 labeling and NMR spectroscopy. *J. Am. Chem. Soc.* **110**, 6920–6921

152. Kleef, M. van, and letters, J. D. (1988) L-tyrosine is the precursor of PQQ biosynthesis in Hyphomicrobium X. *FEBS Lett.* **237**, 91–97

153. Magnusson, O. T., Toyama, H., Saeki, M., Schwarzenbacher, R., and Klinman, J. P. (2004) The Structure of a Biosynthetic Intermediate of Pyrroloquinoline Quinone (PQQ) and Elucidation of the Final Step of PQQ Biosynthesis. *J. AM. CHEM. SOC*. **126**, 5342–5343

154. Puehringer, S., Metlitzky, M., and Schwarzenbacher, R. (2008) The pyrroloquinoline quinone biosynthesis pathway revisited: A structural approach. *BMC Biochem.* **9**, 8

155. Magnusson, O. T., Toyama, H., Saeki, M., Rojas, A., Reed, J. C., Liddington,

R. C., Klinman, J. P., and Schwarzenbacher, R. (2004) Quinone biogenesis: Structure and mechanism of PqqC, the final catalyst in the production of pyrroloquinoline quinone. *Proc. Natl. Acad. Sci.* **101**, 7913–7918

156. Koehn, E. M., Latham, J. A., Armand, T., Evans, R. L., Tu, X., Wilmot, C. M., Iavarone, A. T., and Klinman, J. P. (2019) Discovery of Hydroxylase Activity for PqqB Provides a Missing Link in the Pyrroloquinoline Quinone Biosynthetic Pathway. *J. Am. Chem. Soc.* **141**, 4398–4405
157. Wei, Q., Ran, T., Ma, C., He, J., Xu, D., and Wang, W. (2016) Crystal Structure and Function of PqqF Protein in the Pyrroloquinoline Quinone Biosynthetic Pathway. *J. Biol. Chem.* **291**, 15575–87
158. Barr, I., Latham, J. A., Iavarone, A. T., Chantarojsiri, T., Hwang, J. D., and Klinman, J. P. (2016) The pyrroloquinoline quinone (PQQ) biosynthetic pathway: Demonstration of de novo carbon-carbon cross-linking within the peptide substrate (PqqA) in the presence of the Radical SAM enzyme (PqqE) and its peptide chaperone (PqqD). *J. Biol. Chem.* **291**, 8877–8884
159. Martins, A. M., Latham, J. A., Martel, P. J., Barr, I., Iavarone, A. T., & Klinman, J. P. (2019) A two-component protease in Methylorubrum extorquens with high activity toward the peptide precursor of the redox cofactor pyrroloquinoline quinone. *J. Biol. Chem.* 10.1074/jbc.RA119.009684
160. Bonnot, F., Iavarone, A. T., and Klinman, J. P. (2013) Multistep, eight-electron oxidation catalyzed by the cofactorless oxidase, PqqC: identification

of chemical intermediates and their dependence on molecular oxygen. *Biochem.* . **52**, 4667–4675

161. Rastogi, N., Legrand, E., and Sola, C. (2001) The mycobacteria: an introduction to nomenclature and pathogenesis. *Rev. Sci. Tech. Int. Des Epizoot.* **20**, 21–54

162. Prasanna, A. N., and Mehra, S. (2013) Comparative Phylogenomics of Pathogenic and Non-Pathogenic Mycobacterium. *PLoS One*. **8**, 71248

163. Castañeda-Hernández, D. M., & Rodriguez-Morales, A. J. (2013) Epidemiological Burden of Tuberculosis in Developing Countries. *Curr. Top. Public Heal.*

164. Macneil, A., Glaziou, P., Sismanidis, C., Maloney, S., and Floyd, K. (2017) Morbidity and Mortality Weekly Report Global Epidemiology of Tuberculosis and Progress Toward Achieving Global Targets-2017. *Morb. Mortal. Wkly. Rep.* **68**, 263

165. Pienaar, E., Linderman, J. J., and Kirschner, D. E. (2018) Emergence and selection of isoniazid and rifampin resistance in tuberculosis granulomas. *PLoS One*. **13**, e0196322

166. Seung, K. J., Keshavjee, S., and Rich, M. L. Multidrug-Resistant Tuberculosis and Extensively Drug-Resistant Tuberculosis. *Cold Spring Harb. Perspect. Med.* **5**, a017863

167. Somoskovi, A., Parsons, L. M., & Salfinger, M. (2001) The molecular basis of resistance to isoniazid, rifampin, and pyrazinamide in Mycobacterium

tuberculosis. *Respir. Res.* **2**, 164

168. Yang, T., Zhong, J., Zhang, J., Li, C., Yu, X., Xiao, J., Jia, X., Ding, N., Ma, G., Wang, G., Yue, L., Liang, Q., Sheng, Y., Sun, Y., Huang, H., and Chen, F. (2018) Pan-genomic study of Mycobacterium tuberculosis reflecting the primary/secondary genes, generality/individuality, and the interconversion through copy number variations. *Front. Microbiol.* **9**, 1886

169. Haft, D. H. (2011) Bioinformatic evidence for a widely distributed, ribosomally produced electron carrier precursor, its maturation proteins, and its nicotinoprotein redox partners. *BMC Genomics*. **12**, 21

170. Sassetti, C. M., and Rubin, E. J. (2003) Genetic requirements for mycobacterial survival during infection. *Proc. Natl. Acad. Sci.* **100**, 12989–12994

171. Dejesus, M. A., Gerrick, E. R., Xu, W., Park, S. W., Long, J. E., Boutte, C. C., Rubin, E. J., Schnappinger, D., Ehrt, S., Fortune, S. M., Sassetti, C. M., and Ioerger, T. R. (2017) Comprehensive Essentiality Analysis of the Mycobacterium tuberculosis Genome via Saturating Transposon Mutagenesis. *MBio*. **8.1**, e02133-16

172. Brzostek, A., Pawelczyk, J., Rumijowska-Galewicz, A., Dziadek, B., and Dziadek, J. (2009) Mycobacterium tuberculosis Is Able To Accumulate and Utilize Cholesterol. *J. Bacteriol.* **191**, 6584–6591

173. Pandey, A. K., and Sassetti, C. M. (2008) Mycobacterial persistence requires the utilization of host cholesterol. *Proc. Natl. Acad. Sci.* . **105.11**, 4376–4380

174. Griffin, J. E., Pandey, A. K., Gilmore, S. A., Mizrahi, V., Mckinney, J. D., Bertozzi, C. R., and Sassetti, C. M. (2012) Cholesterol Catabolism by Mycobacterium tuberculosis Requires Transcriptional and Metabolic Adaptations. *Chem. Biol.* **19**, 218–227

175. Wilburn, K. M., Fieweger, R. A., and Vanderven, B. C. (2018) Cholesterol and fatty acids grease the wheels of Mycobacterium tuberculosis pathogenesis. *Pathog. Dis.* **76.2**, fty021

176. Pandey, A. K., and Sassetti, C. M. (2008) Mycobacterial persistence requires the utilization of host cholesterol. *Proc. Natl. Acad. Sci.* **105**, 4376–4380

177. Griffin, J. E., Gawronski, J. D., Dejesus, M. A., Ioerger, T. R., and Akerley, B. J. (2011) High-Resolution Phenotypic Profiling Defines Genes Essential for Mycobacterial Growth and Cholesterol Catabolism. *PLoS Pathog*. **7**, 1002251

178. Wipperman, M. F., Sampson, N. S., and T. Thomas, S. (2014) Pathogen roid rage: cholesterol utilization by Mycobacterium tuberculosis. *Crit. Rev. Biochem. Mol. Biol.* **49**, 269–293

179. Krishnamoorthy, G., Kaiser, P., Lozza, L., Hahnke, K., Mollenkopf, H.-J., and Kaufmann, S. H. E. (2019) Mycofactocin Is Associated with Ethanol Metabolism in Mycobacteria. *MBio*. **10**, 1–14

180. Dubey, A. A., and Jain, V. (2019) Mycofactocin is essential for the establishment of methylotrophy in Mycobacterium smegmatis. *Biochem. Biophys. Res. Commun.* **516**, 1073–1077

181. Wecksler, S. R., Stoll, S., Tran, H., Magnusson, O. T., Wu, S.-P., King, D., David Britt, R., and Klinman, J. P. (2009) Pyrroloquinoline Quinone Biogenesis: Demonstration That PqqE from Klebsiella pneumoniae Is a Radical S-Adenosyl-L-methionine Enzyme. *Biochemistry*. **48**, 10151–10161

182. Haft, D. H., Pierce, P. G., Mayclin, S. J., Sullivan, A., Gardberg, A. S., Abendroth, J., Begley, D. W., Phan, I. Q., Staker, B. L., Myler, P. J., Marathias, V. M., Lorimer, D. D., and Edwards, T. E. (2017) Mycofactocin-associated mycobacterial dehydrogenases with non-exchangeable NAD cofactors. *Sci. Rep.* **7**, 41074

183. Minch, K. J., Rustad, T. R., Peterson, E. J. R., Winkler, J., Reiss, D. J., Ma, S., Hickey, M., Brabant, W., Morrison, B., Turkarslan, S., Mawhinney, C., Galagan, J. E., Price, N. D., Baliga, N. S., and Sherman, D. R. (2015) The DNA-binding network of Mycobacterium tuberculosis. *Nat. Commun.* **6**, 5829

184. Stahl', D. A., and Urbance', J. W. (1990) The Division between Fast-and Slow-Growing Species Corresponds to Natural Relationships among the Mycobacteria. *J. Bacteriol.* **172**, 116–124

185. Bruender, N. A., and Bandarian, V. (2017) The creatininase homolog MftE from Mycobacterium smegmatis catalyzes a peptide cleavage reaction in the biosynthesis of a novel RiPP. *J. Biol. Chem.* **292**, 4371–4381

186. Ruszczycky, M. W., Zhong, A., and Liu, H.-W. (2018) Following the electrons: peculiarities in the catalytic cycles of radical SAM enzymes. *Nat. Prod. Rep.* **35.7**, 615–621

187. Crooks, G. E., Hon, G., Chandonia, J. M., and Brenner, S. E. (2004) WebLogo: A sequence logo generator. *Genome Res.* **14**, 1188–1190

188. Fourmond, V. (2016) QSoas: A Versatile Software for Data Analysis. *Anal. Chem.* **88**, 5050–5052

189. Léger, C., and Bertrand, P. (2008) Direct Electrochemistry of Redox Enzymes as a Tool for Mechanistic Studies. *Chem. Rev.* **108**, 2379–2438

190. Gerlt, J. A., Bouvier, J. T., Davidson, D. B., Imker, H. J., Sadkhin, B., Slater, D. R., and Whalen, K. L. (2015) Enzyme function initiative-enzyme similarity tool (EFI-EST): A web tool for generating protein sequence similarity networks. *Biochim. Biophys. Acta (BBA)-Proteins Proteomics*. **1854**, 1019–1037

191. Nakai, T., Ito, H., Kobayashi, K., Takahashi, Y., Hori, H., Tsubaki, M., Tanizawa, K., and Okajima, T. (2015) The radical S-Adenosyl-L-methionine enzyme QhpD catalyzes sequential formation of intra-protein sulfur-to-methylene carbon thioether bonds. *J. Biol. Chem.* **290**, 11144–11166

192. Walsh, Christopher T., Agnes Schonbrunn, and R. H. A. (1971) Studies on the Mechanism of Action of D-Amino Acid Oxidase Evidence for Removal of Substrate α-Hydrogen As a Proton. *J. Biol. Chem.* **246**, 6855–6866

193. Walker, L. M., Kincannon, W. M., Bandarian, V., and Elliott, S. J. (2018) Deconvoluting the reduction potentials for the three [4Fe-4S] clusters in an AdoMet radical SCIFF maturase. *Biochemistry*. **57**, 6050–6053

194. Tsuru, D., Oka, I., and Yoshimoto, T. (1976) Creatinine Decomposing

Enzymes in *Pseudomonas putida*. *Agric. Biol. Chem.* **40**, 1011–1018

195. Yoshimoto, T., Oka, I., and Biophysics, D. T. (1976) Creatine amidinohydrolase of Pseudomonas putida: crystallization and some properties. *Arch. Biochem. Biophys.* **177**, 508–515

196. Grochowski, L. L., Xu, H., and White, R. H. (2009) An iron(II) dependent formamide hydrolase catalyzes the second step in the archaeal biosynthetic pathway to riboflavin and 7,8-didemethyl-8-hydroxy-5- deazariboflavin. *Biochemistry*. **48**, 4181–4188

197. Beuth, B., Niefind, K., and Schomburg, D. (2003) Crystal Structure of Creatininase from Pseudomonas putida : A Novel Fold and a Case of Convergent Evolution. *J. Mol. Biol.* **2836**, 287–301

198. Gilboa, R., Spungin-Bialik, A., Wohlfahrt, G., Schomburg, D., Blumberg, S., and Shoham, G. (2001) Interactions of Streptomyces griseus aminopeptidase with amino acid reaction products and their implications toward a catalytic mechanism. *Proteins Struct. Funct. Bioinformatics.* **44**, 490–504

199. Lockridge, O., Massey, V., and Sullivan, P. A. (1972) Mechanism of Action of the Flavoenzyme Lactate Oxidase. *J. Biol. Chem.* **247**, 6097–8106

200. Sukumar, N., Dewanti, A., Merli, A., Rossi, G. L., Mitra, B., and Mathews, F. S. (2009) Structures of the G81A mutant form of the active chimera of (S)-mandelate dehydrogenase and its complex with two of its substrates. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **65**, 543–552

201. Molla, G., Sacchi, S., Bernasconi, M., Pilone, M. S., Fukui, K., and Pollegioni, L. (2006) Characterization of human d-amino acid oxidase. *FEBS Lett.* **580**, 2358–2364

202. Lindqvist, Y. (1989) Refined structure of spinach glycolate oxidase at 2 Å resolution. *J. Mol. Biol.* **209**, 151–166

203. Umhau, S., Pollegioni, L., Molla, G., Diederichs, K., Welte, W., Pilone, M. S., and Ghisla, S. (2000) The x-ray structure of D-amino acid oxidase at very high resolution identifies the chemical mechanism of flavin-dependent substrate dehydrogenation. *Proc. Natl. Acad. Sci.* **97**, 12463-12468.

204. Hartmann, C., Brzovic, P., and Klinman, J. P. (2234) Spectroscopic detection of chemical intermediates in the reaction of para-substituted benzylamines with bovine serum amine oxidase. *Biochemistry*. **32**, 2234–2241

205. Jonsson, Thorlakur, Michael H. Glickman, Shujun Sun, and J. P. K. (1993) Experimental evidence for extensive tunneling of hydrogen in the lipoxygenase reaction: implications for enzyme catalysis. *ournal Am. Chem. Soc.* **31**, 10319–10320

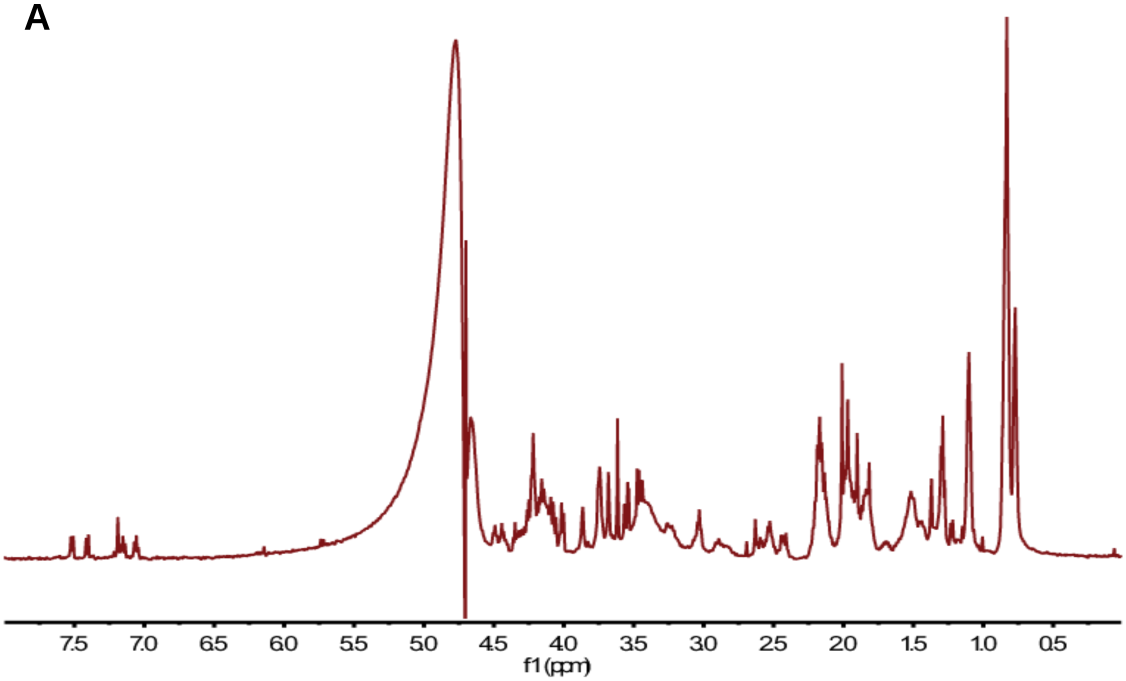
206. van der Werf, M. J., van der Ven, C., Barbirato, F., M Eppink, M. H., M de Bont, J. A., and H van Berkel, W. J. (1999) Stereoselective Carveol Dehydrogenase from Rhodococcus erythropolis DCL14 a novel nicotinoprotein belonging to the short chain dehydrogenase/reductase superfamily. *J. Biol. Chem.* **274**, 26296–26304

207. Zheng, G., Hehn, R., and Zuber, P. (2000) Mutational Analysis of the sbo-

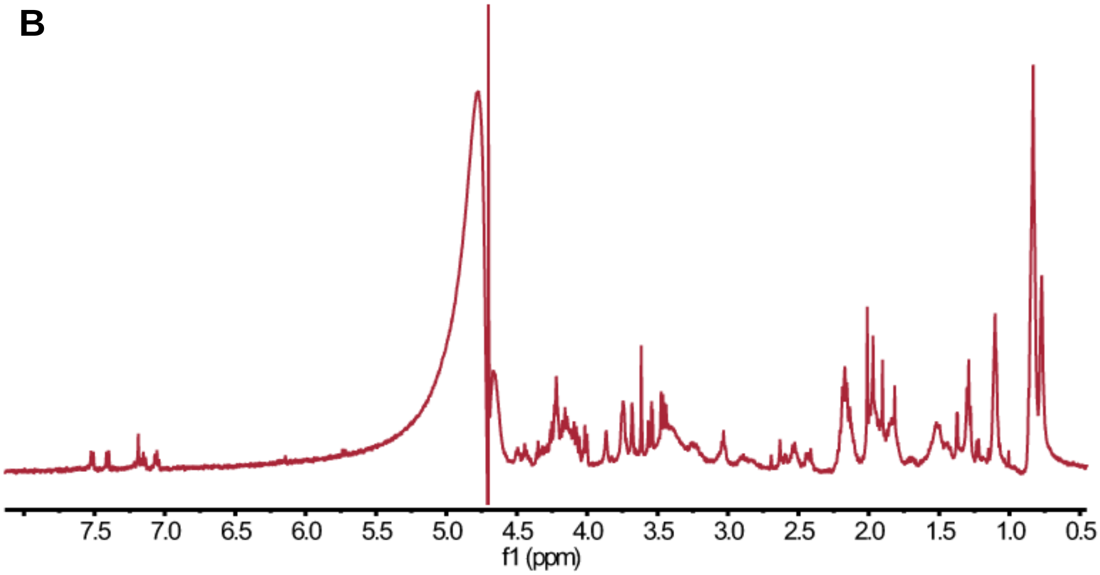
alb Locus of Bacillus subtilis : Identification of Genes Required for Subtilosin Production and Immunity Mutational Analysis of the sbo-alb Locus of Bacillus subtilis : Identification of Genes Required for Subtilosin Product. **182**, 3266–3273

APPENDIX A: SUPPLEMENTAL DATA

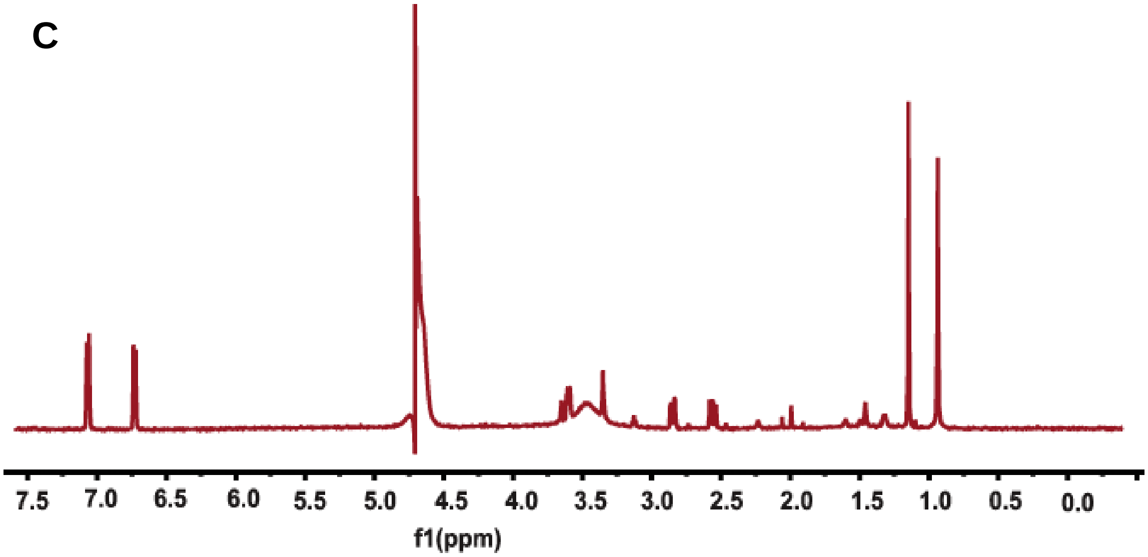
A



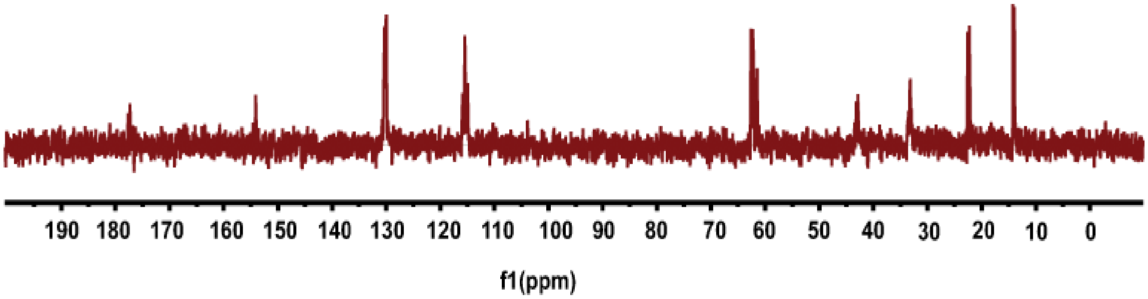
B



C


7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1(ppm)

D


190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1(ppm)

E

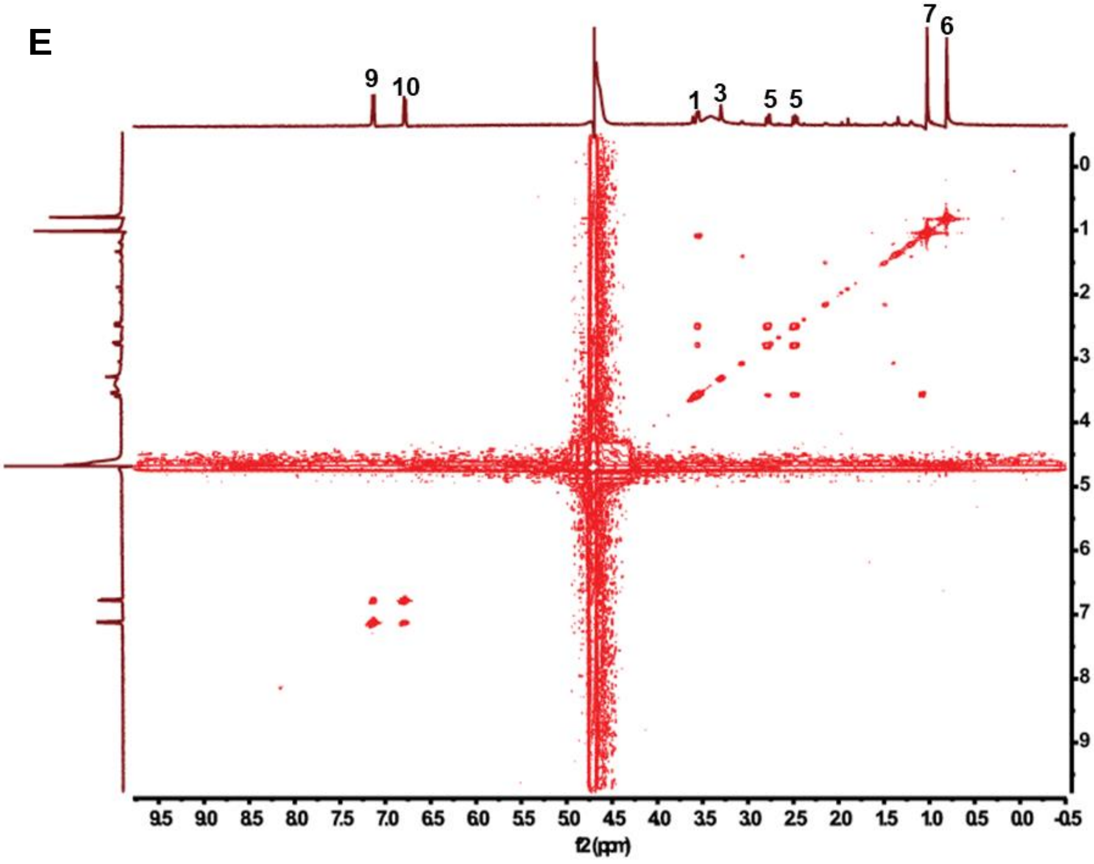
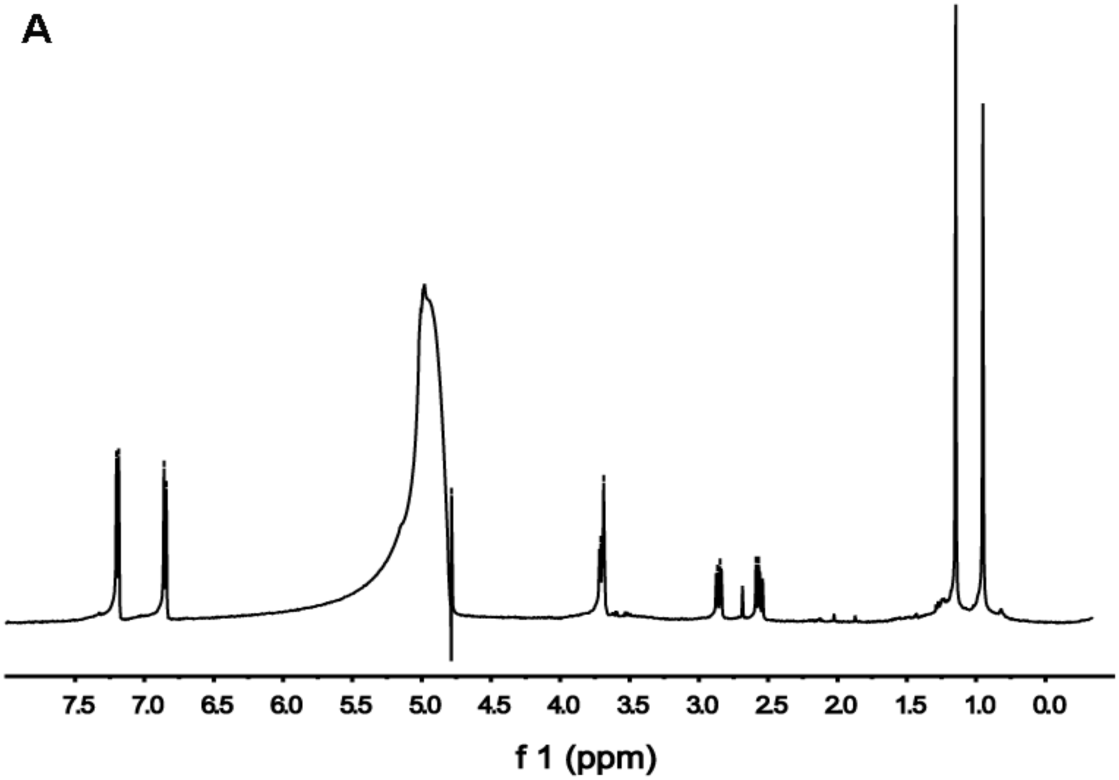
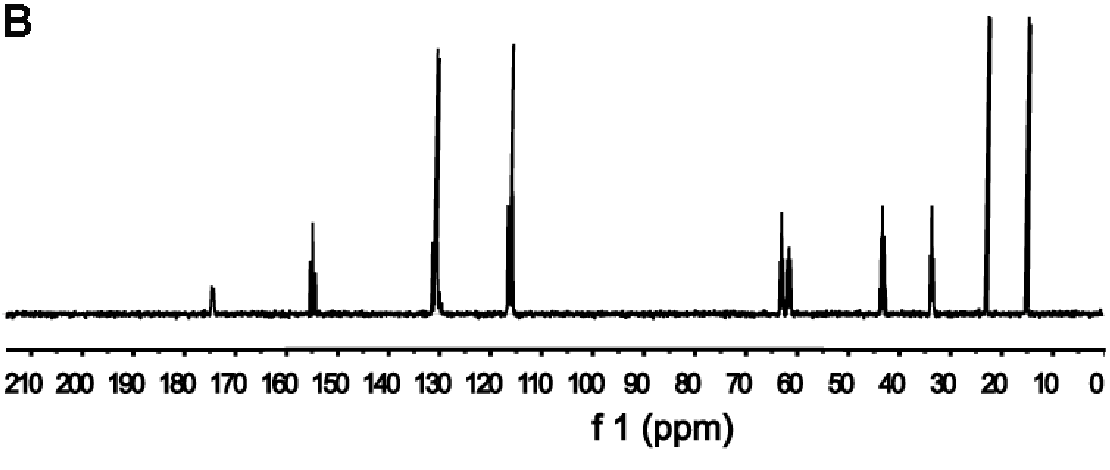


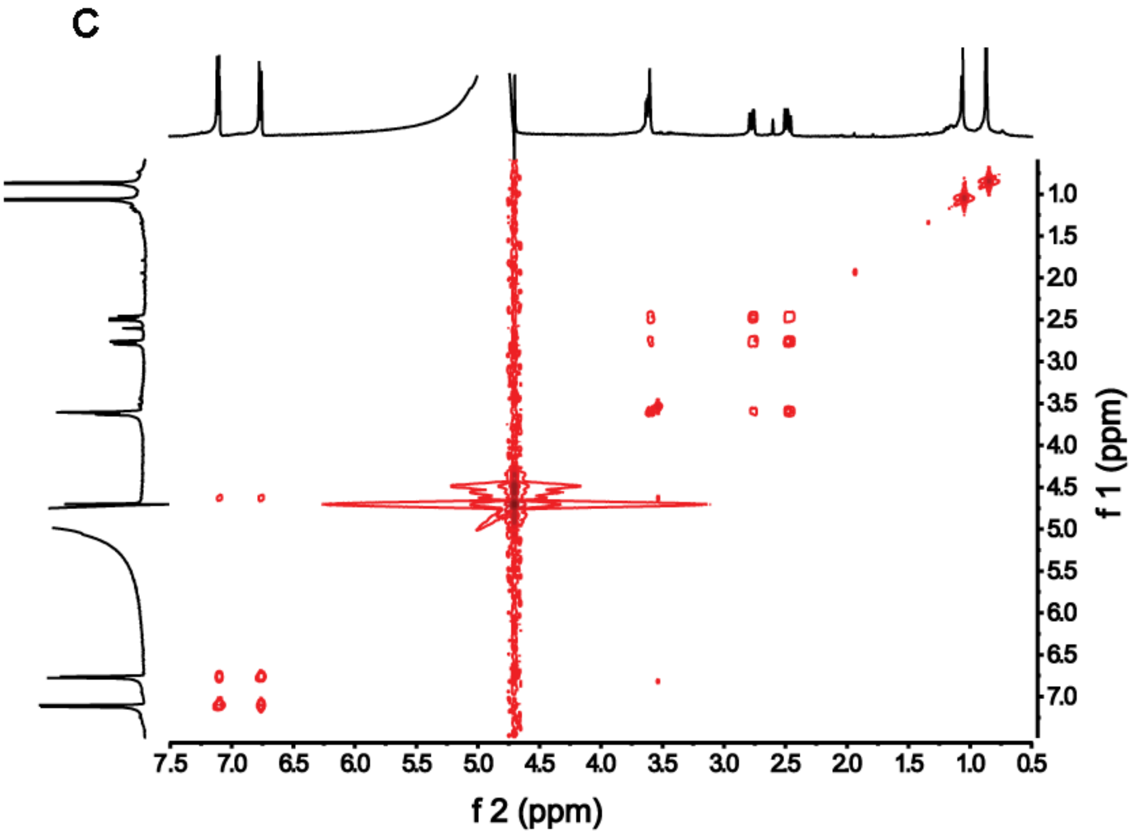
Figure S 1 NMR spectra of the substrate and products of MftE reaction in D2O. Shown are (A) ^{1}H NMR spectrum of MftA*, (B)^{1}H NMR spectrum of MftA (1-28), (C) ^{1}H NMR spectrum of AHDP, (D)^{13}C NMR spectrum of AHDP and (E) COSY NMR spectrum of AHDP.

A


7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.0
f 1 (ppm)

B


210
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f 1 (ppm)


C
f 2 (ppm)
f 1 (ppm)
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5

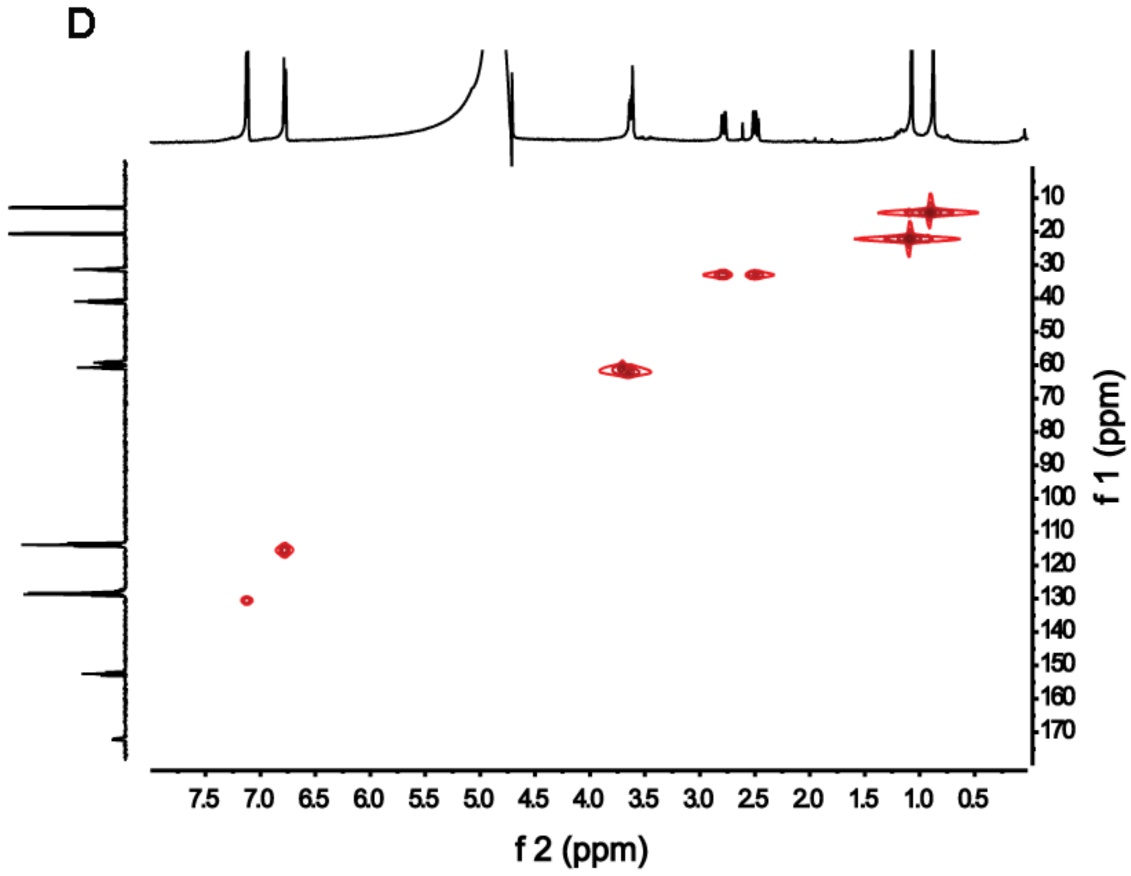
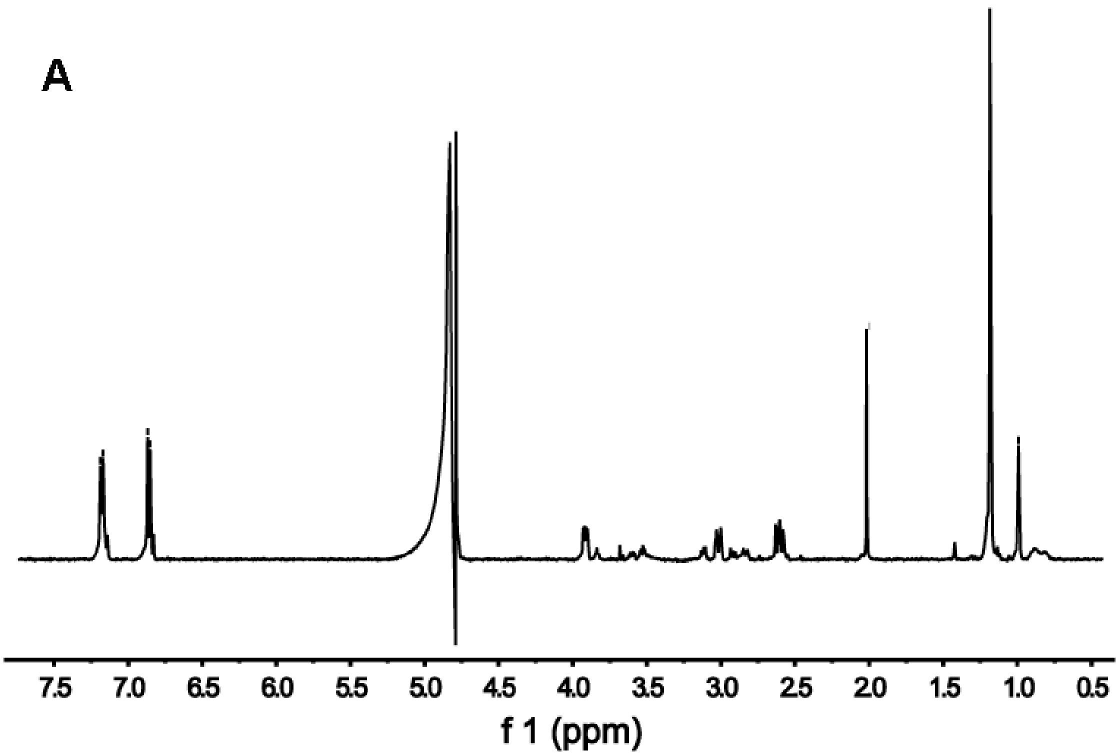
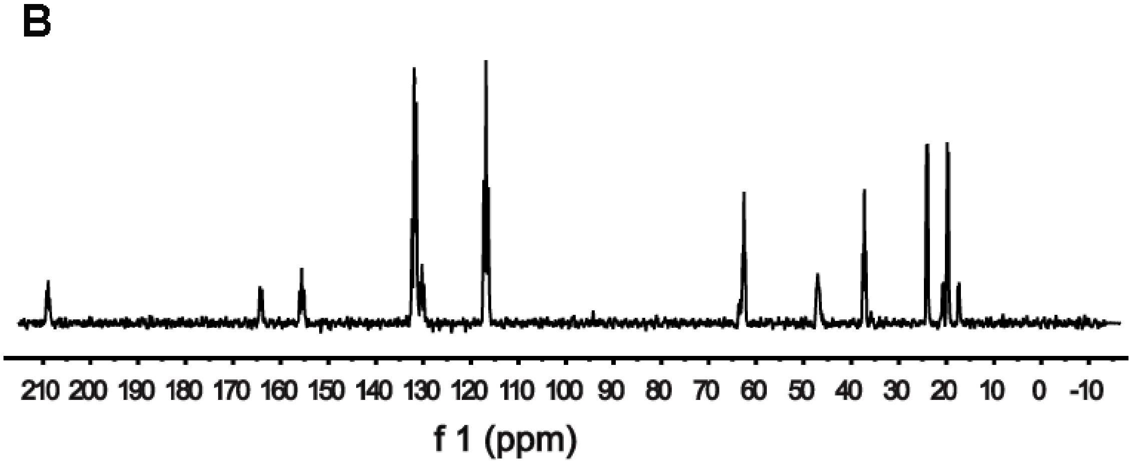
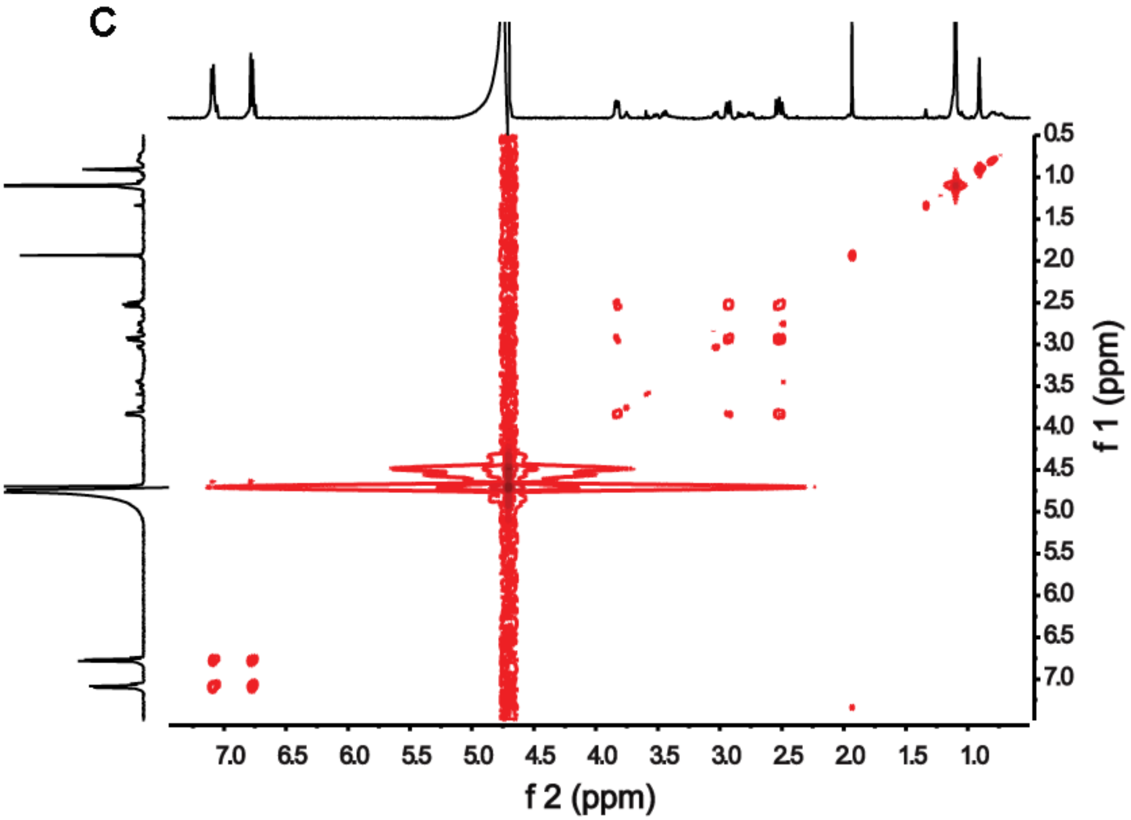


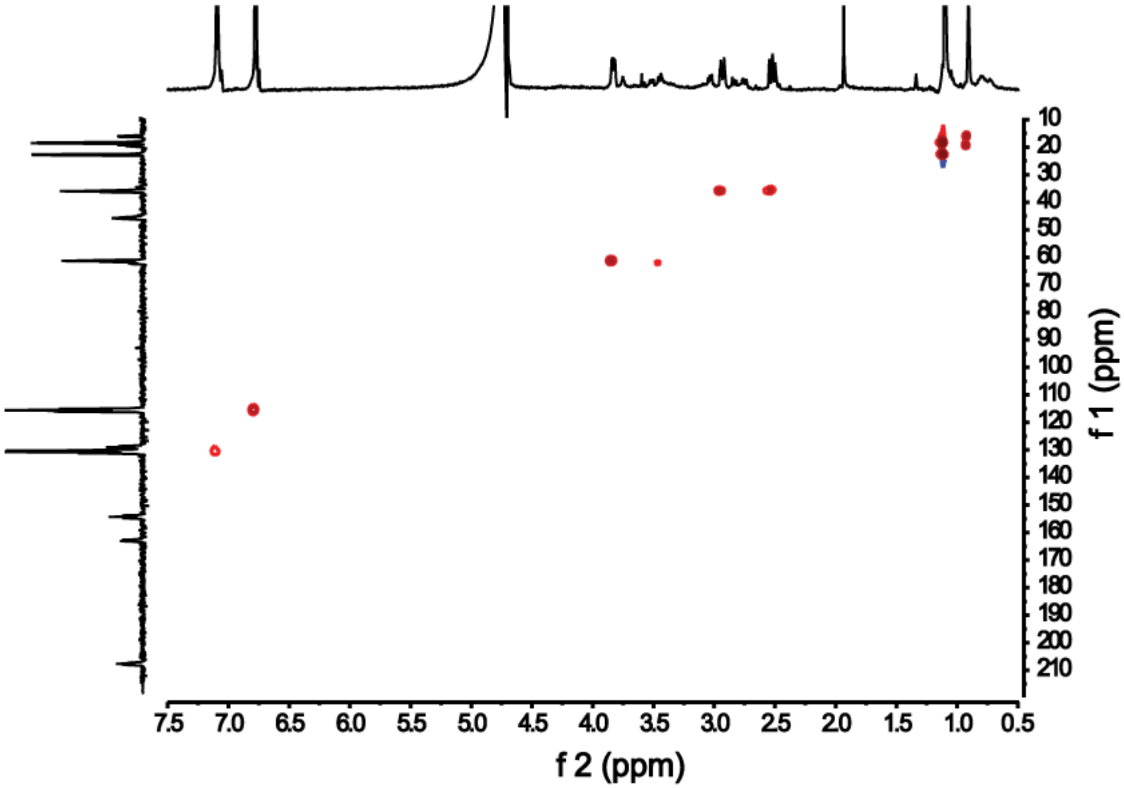
Figure S 2. NMR spectra of MftD reaction substrate AHDP, in D2O. NMR spectra of the substrate AHDP in D2O. Shown are ^{1}H (A), ^{13}C (B), COSY (C) and HSQC (D) NMR spectra.


A
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
f 1 (ppm)


B
210
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
-10
f 1 (ppm)


C
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0
5.5
6.0
6.5
7.0
f 1 (ppm)
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
f 2 (ppm)

D



7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
f 2 (ppm)
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190
200
210
f 1 (ppm)

E

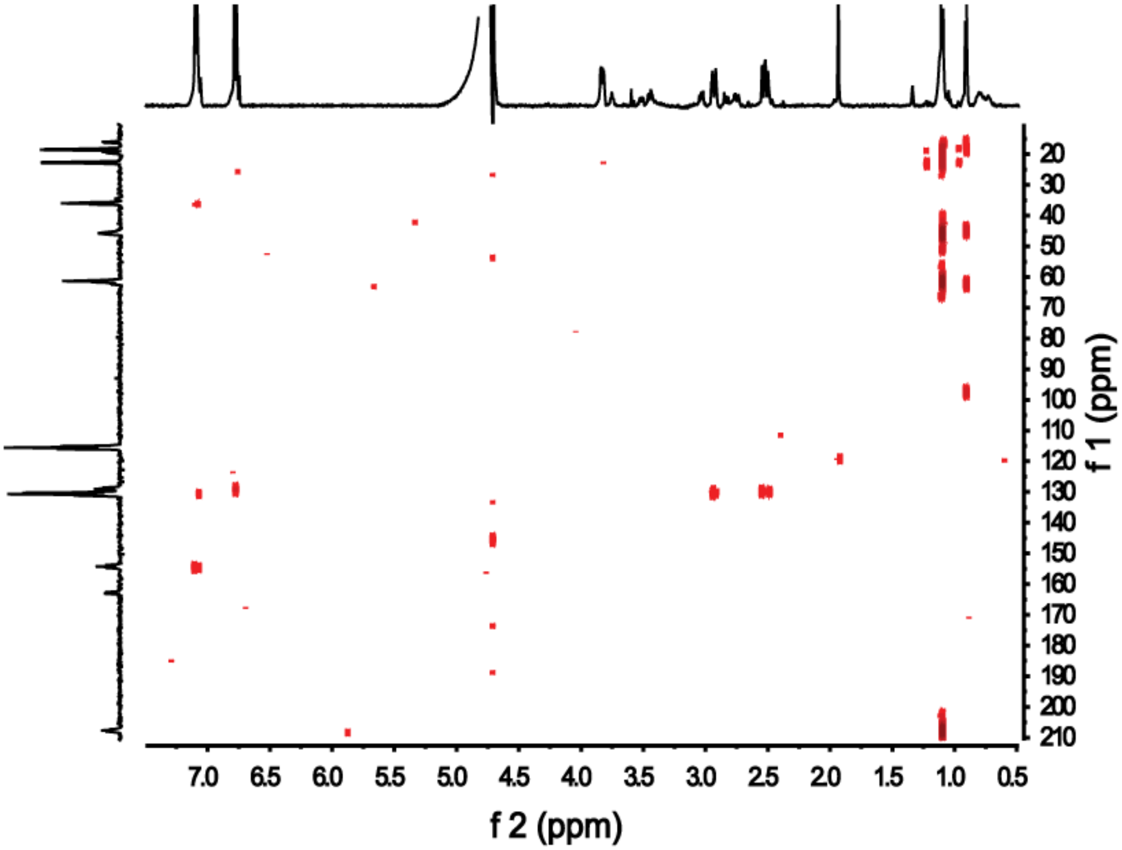


Figure S3 NMR spectra of MftD reaction Product, PMFT in D2O. Shown are ^{1}H (A), ^{13}C (B), COSY (C) and HSQC (D) and HMBC (E) NMR spectra.

APPENDIX B: LIST OF PUBLICATIONS

Khaliullin, B., **Ayikpoe, R**., Tuttle, M., and Latham, X. J. A. (2017) Mechanistic elucidation of the mycofactocin-biosynthetic radical S -adenosylmethionine protein, MftC. *J. Biol. Chem.* **292**, 13022–13033.

Ayikpoe, R., Salazar, J., Majestic, B., and Latham, J. A. (2018) Mycofactocin Biosynthesis Proceeds through 3-Amino-5-[(p-hydroxyphenyl)methyl]-4,4-dimethyl-2-pyrrolidinone (AHDP); Direct Observation of MftE Specificity toward MftA. *Biochemistry*. **57**, 5379–5383.

Ayikpoe, R., Ngendahimana, T., Langton, M., Bonitatibus, S., Walker, L. M., Eaton, S. S., Eaton, G. R., Pandelia, M.-E., Elliott, S. J., and Latham, J. A. (2019) Spectroscopic and Electrochemical Characterization of the Mycofactocin Biosynthetic Protein, MftC, Provides Insight into Its Redox Flipping Mechanism. *Biochemistry*. **58**, 940–950.

Ayikpoe, R., Govindarajan, V., and Latham, J. A. (2019) Occurrence, function, and biosynthesis of mycofactocin. *Appl. Microbiol. Biotechnol.* **103**, 2903–2912.

Ngendahimana, T., **Ayikpoe, R.,** Latham, J. A., Eaton, G. R., & Eaton, S. S. (2019) Structural insights for vanadium catecholates and iron-sulfur clusters obtained from multiple data analysis methods applied to electron spin relaxation data. *J. Inorg. Biochem.*

Ayikpoe, R. S., and Latham, J. A. (2019) MftD Catalyzes the Formation of a Biologically Active Redox Center in the Biosynthesis of the Ribosomally Synthesized and Post-translationally Modified Redox Cofactor, Mycofactocin. *J. Am. Chem. Soc.* ***141***(34), 13582-13591.