2023

Mechanistic Insights into MftR-Dependent Regulation of the Redox Cofactor Mycofactocin

Aigera Mendauletova
University of Denver

Mechanistic Insights into MftR-Dependent Regulation of the Redox Cofactor Mycofactocin

Abstract

Organic redox cofactors are essential for life. While classic flavins and nicotinamides are widely distributed across all domains of life, nature has also evolved niche cofactors in subsets of life domains. For example, in *Actinobacteria*, coenzyme F420 is commonly used in place of flavin mononucleotide in enzymes associated with carbon fixation and oxidation of secondary alcohols. The importance of niche cofactors has long been recognized however, detailed understanding about their biosynthesis and physiological uses has been lagging. One class of niche cofactors is derived from ribosomally synthesized and posttranslationally modified peptides (RiPPs). To achieve their mature form, the genetically encoded RiPP precursor peptide undergoes significant post-translational modifications by diverse families of tailoring enzymes. Following synthesis by the ribosome, modifying enzymes process the precursor peptide into the mature redox cofactor. Currently, there are two known RiPP-derived redox cofactors, pyrroloquinoline quinone, which has been well-characterized, and mycofactocin (MFT), which was recently discovered.

While MFT biosynthetic proteins have been extensively characterized, the physiological conditions that require MFT is not well-understood. To gain insights into the regulation of MFT in *Mycobacterium smegmatis* mc2155, we investigated the DNA-binding and ligand-binding activity of the putative TetR-like regulator, MftR. Here, we demonstrated that MftR binds to the *mft* promoter region. We used DNase I footprinting to identify the 27 bp palindromic operator located 5' to *mftA* and found it to be highly conserved in *M. tuberculosis, M. bovis, M. ulcerans,* and *M. marinum*. To determine when the *mft* biosynthetic gene cluster (BGC) is induced, we screened for effectors of MftR. As a result, we found that MftR binds to long-chain acyl-CoA's with low micromolar affinities. To demonstrate that oleoyl-CoA induces the *mft* BGC *in vivo*, we re-engineered a fluorescent protein reporter system to express a MftA-mCherry fusion. Using the mCherry fluorescent readout, we show that the *mft* BGC is upregulated in *M. smegmatis* mc2155 when oleic acid is supplemented to the media. These results suggest that MftR controls expression of the *mft* BGC and that MFT production is induced by long chain acyl-CoA's. Since MFT-dependent dehydrogenases are known to colocalize with acyl carrier protein/CoA-modifying enzymes, these results suggest that MFT might be critical for fatty acid metabolism or cell wall reorganization.

Document Type
Dissertation

Degree Name
Ph.D.

Department
Chemistry and Biochemistry

First Advisor
Sandra S. Eaton

Second Advisor
Allegra Aron

Third Advisor

Michelle Knowles

Keywords

Organic redox cofactors, Ribosomally synthesized and posttranslationally modified peptides (RiPPs), Mycofactocin (MFT)

Subject Categories

Biochemistry | Biochemistry, Biophysics, and Structural Biology | Chemistry | Life Sciences | Physical Sciences and Mathematics

Publication Statement

Copyright is held by the author. User is responsible for all copyright compliance.

Mechanistic Insights into MftR-dependent Regulation of the Redox Cofactor

Mycofactocin

_____

A Dissertation

Presented to

the Faculty of the College of Natural Sciences and Mathematics

University of Denver

_____

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

_____

by

Aigera Mendauletova

June 2023

Advisor: Sandra S. Eaton

Author: Aigera Mendauletova
Title: Mechanistic Insights into MftR-dependent Regulation of the Redox Cofactor Mycofactocin
Advisor: Sandra S. Eaton
Degree Date: June 2023

Abstract

Organic redox cofactors are essential for life. While classic flavins and nicotinamides are widely distributed across all domains of life, nature has also evolved niche cofactors in subsets of life domains. For example, in *Actinobacteria*, coenzyme F_{420} is commonly used in place of flavin mononucleotide in enzymes associated with carbon fixation and oxidation of secondary alcohols. The importance of niche cofactors has long been recognized however, detailed understanding about their biosynthesis and physiological uses has been lagging. One class of niche cofactors is derived from ribosomally synthesized and posttranslationally modified peptides (RiPPs). To achieve their mature form, the genetically encoded RiPP precursor peptide undergoes significant post-translational modifications by diverse families of tailoring enzymes. Following synthesis by the ribosome, modifying enzymes process the precursor peptide into the mature redox cofactor. Currently, there are two known RiPP-derived redox cofactors, pyrroloquinoline quinone, which has been well-characterized, and mycofactocin (MFT), which was recently discovered.

While MFT biosynthetic proteins have been extensively characterized, the physiological conditions that require MFT is not well-understood. To gain insights into the regulation of MFT in *Mycobacterium smegmatis* mc^{2}155, we investigated the DNA-

binding and ligand-binding activity of the putative TetR-like regulator, MftR. Here, we demonstrated that MftR binds to the *mft* promoter region. We used DNase I footprinting to identify the 27 bp palindromic operator located 5' to *mftA* and found it to be highly conserved in *M. tuberculosis*, *M. bovis*, *M. ulcerans,* and *M. marinum.* To determine when the *mft* biosynthetic gene cluster (BGC) is induced, we screened for effectors of MftR. As a result, we found that MftR binds to long-chain acyl-CoA's with low micromolar affinities. To demonstrate that oleoyl-CoA induces the *mft* BGC *in vivo*, we re-engineered a fluorescent protein reporter system to express a MftA-mCherry fusion. Using the mCherry fluorescent readout, we show that the *mft* BGC is upregulated in *M. smegmatis* mc^{2}155 when oleic acid is supplemented to the media. These results suggest that MftR controls expression of the *mft* BGC and that MFT production is induced by long chain acyl-CoA's. Since MFT-dependent dehydrogenases are known to colocalize with acyl carrier protein/CoA-modifying enzymes, these results suggest that MFT might be critical for fatty acid metabolism or cell wall reorganization.

Acknowledgements

First, I would like to express my deepest appreciation to my advisor Dr. John Latham who has been and will continue to be an important role model in my scientific career. Having the opportunity to work with you and learn from you made me grow into a true scientist. Thank you for being a great mentor! I am grateful to Dr. Sandra Eaton, Dr. Allegra Aron, Dr. Michelle Knowles, and Dr. Scott Nichols for working with me as part of my Dissertation Committee. Dr. Sandra Eaton, I truly appreciate you and the time you spent helping me on many occasions in my last year of PhD. I am extremely grateful to Dr. Allegra Aron for your encouragement, kindness, and support. Special thanks to my fellow lab mates and peers. Richard and Vishnu, I thank you for making my transition into the lab a welcoming experience and for your guidance while teaching me. Anastasiia, Yi, Brad, Andres, and Jake, you made graduate school such a fun experience for me, I cherish all the memories we made together. Anastasiia, I couldn't have asked for a better friend, lab mate, volleyball partner, roommate, hiking buddy, and conference mate. I am so thankful that our paths have crossed. I could not have done this without you. You are truly the best! I would also like to thank my family and friends for always being there for me. I am deeply indebted to my parents for their unconditional love and constant support. Thank you for giving me freedom follow my own path – even if that means I had to move across the globe. And finally, to Sam, my husband! You shared this entire PhD journey with me, through my ups and downs. You are my most enthusiastic cheerleader and my constant advocate. You are amazing, and I love you! And to my dogs, Gency and Tzipi, who spent every day by my side as I was writing this dissertation.

Table of Contents

List of figures

ix

Abbreviations

5'-dA•	5' – deoxyadenosyl radical
AHDP	3-amino-5-[(p-hydroxyphenyl) methyl]-4,4-dimethyl-2-pyrrolidinone
AnSME	anaerobic sulfatase maturating enzyme
Aux	auxiliary
BGC	biosynthetic gene cluster
C_α	alpha – carbon
C_β	beta – carbon
C_γ	gamma – carbon
CFU	colony forming unit
CHAPS	3-[(3-cholamidopropyl) dimethylammonium)-1-propanesulfonate
CPG	cyclopropylglycine
DIC	N, N'- diisopropylcarbodiimide
DMSO	dimethyl sulfoxide
DMF	dimethylformamide
DTT	dithiothreitol
DTH	sodium dithionite
EDTA	ethylenediaminetetraacetic acid
EFI – EST	enzyme function initiative enzyme similarity tool
EMSA	electrophoretic mobility shift assay
FAM	fluorescein

FGly   α - formylglycine

FPLC   fast protein liquid chromatography

HEPES   2-[4-(2-hydroxyethyl) piperazin-1-yl] ethanesulfonic acid

HPLC   high performance liquid chromatography

HTH   helix – turn – helix

IPTG   isopropyl β-D-1-thiogalactopyranoside

ITC   isothermal titration calorimetry

mCPG   methyl – cyclopropylglycine

MFT   mycofactocin

Msmeg   *Mycobacterium smegmatis*

MTB   *Mycobacterium tuberculosis*

NMP   N-methyl-2-pyrrolidone

NMR   nuclear magnetic resonance

OA   oleic acid

OD   optical density

PCR   polymerase chain reaction

P_{mft} promoter region

PTM   post – translational modification

PMFT   premycofactocin

qRT-PCR   quantitative real-time polymerase chain reaction

QToF-LCMS   quadrupole time-of-flight liquid chromatography mass spectrometry

RiPPs   ribosomally synthesized post – translationally modified peptides

rSAM	radical S – adenosylmethionine
RRE	RiPP recognition element
SCIFF	six – cysteines – in – forty – five
SDS-PAGE	sodium dodecyl sulfate – polyacrylamide gel electrophoresis
SKF	sporulation killing factor
SPPS	solid phase peptide synthesis
SSN	sequence similarity network
TFA	trifluoroacetic acid
TIS	triisopropylsilane

Chapter One: Introduction

1.1 The significance of ribosomally synthesized and post-translationally modified peptides

Antibiotic resistance is a leading cause of mortality around the world, killing at least 1.27 million people worldwide in 2019.[1] It occurs when bacteria, viruses, fungi, and parasites develop resistance over time and no longer respond to the medicine which threatens our ability to treat common infections.[2] To combat this, new antibiotics are urgently needed. Historically, microbial natural products have been the major source of antibiotic compounds,[3–5] as a result, a lot of efforts have been applied to generate novel analogs with pharmacological properties.[6,7] The biosynthetic rigidity of non-ribosomal peptides and polyketides limits their ability to produce different analogs in large numbers.[8] However, innovation in genomics and techniques have led to the discovery of a rapidly growing superclass of natural products: ribosomally synthesized and post-translationally modified peptides (RiPPs).[9] RiPPs, unlike the other members of the natural products, possess an extraordinary biosynthetic manipulability.[10] Owing to this, their immense distribution in bacteria, and their engineerability, RiPPs have emerged as a potential source of new therapeutics.[11]

RiPPs are structurally diverse and possess a spectrum of activities including antibacterial,[12,13] antifungal,[14,15] antiviral,[16] and anticancer[17] (Figure 1.1). Currently,

1

several of the antimicrobial RiPPs are under the clinical trials but none yet approved for clinical use in humans.[18] Mostly, the research is hindered by the poor solubility and bioavailability of the RiPP products.[19]
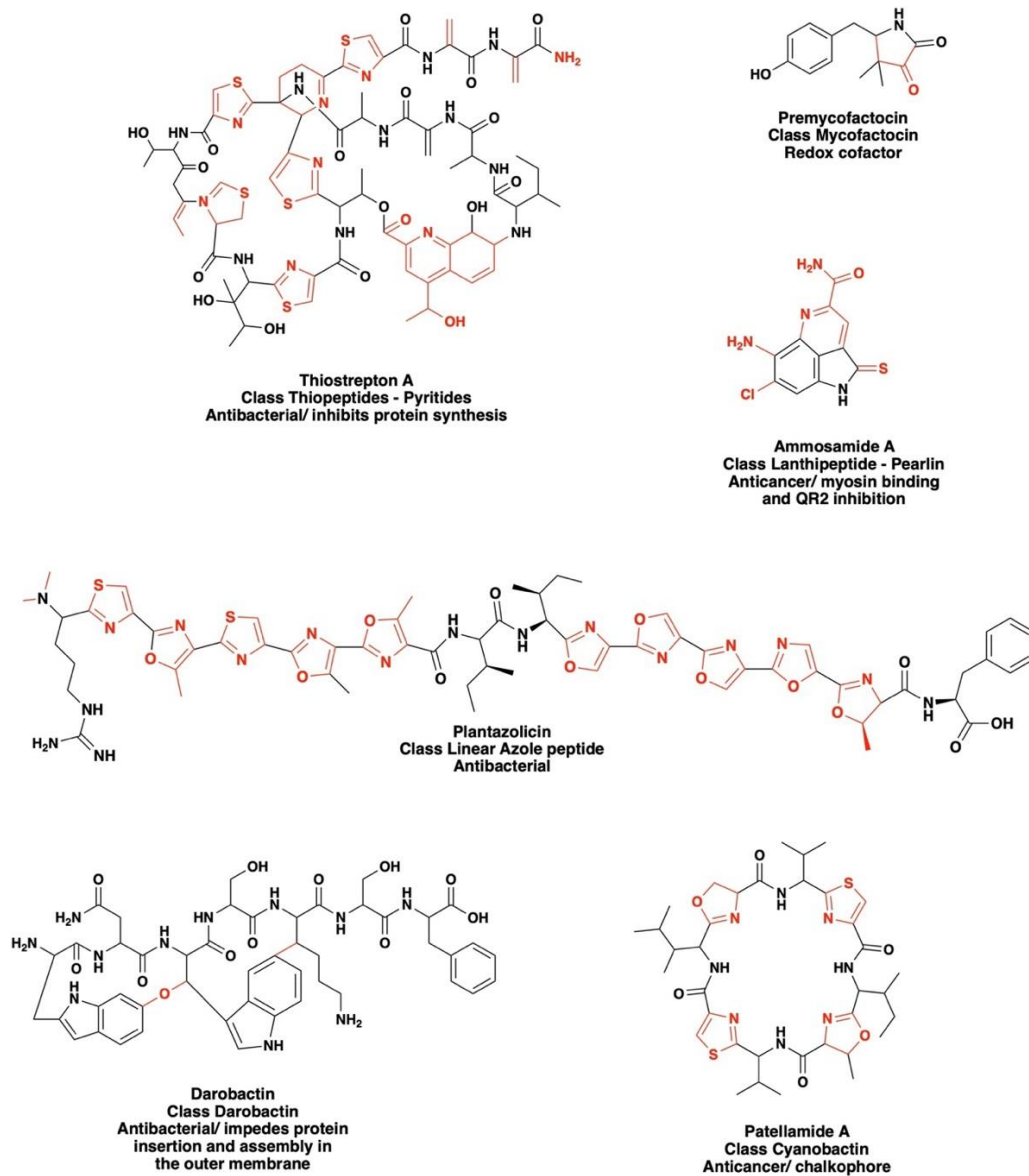
Figure 1.1 Representative members of several RiPP classes. RiPPs display wide array of structural diversity. PTMs are indicated in red.

1.1.1 Common features of RiPP biosynthesis

Despite being functionally and structurally diverse, the RiPP family contains a unifying biosynthetic logic. In bacteria, the genes that produce a RiPP are often clustered together, which makes them easily distinguishable. These genes include a short gene encoding for a precursor peptide and modifying enzymes that sequentially process it, yielding a mature product (Figure 1.2).[9] Precursor peptides are usually 20 – 110 residues in length and composed of multiple segments: leader, core, and follower. The leader and/or follower regions are recognized and bound by post-translationally modifying enzymes and/or a pathway associated RiPP recognition element. They serve as allosteric effectors to activate the biosynthetic enzymes. The core peptide is the segment where post - translational modifications (PTM) happen.[10] Usually, one of the tailoring enzymes in a gene cluster acts as a protease and cleaves the leader/follower sequence resulting in the final product.
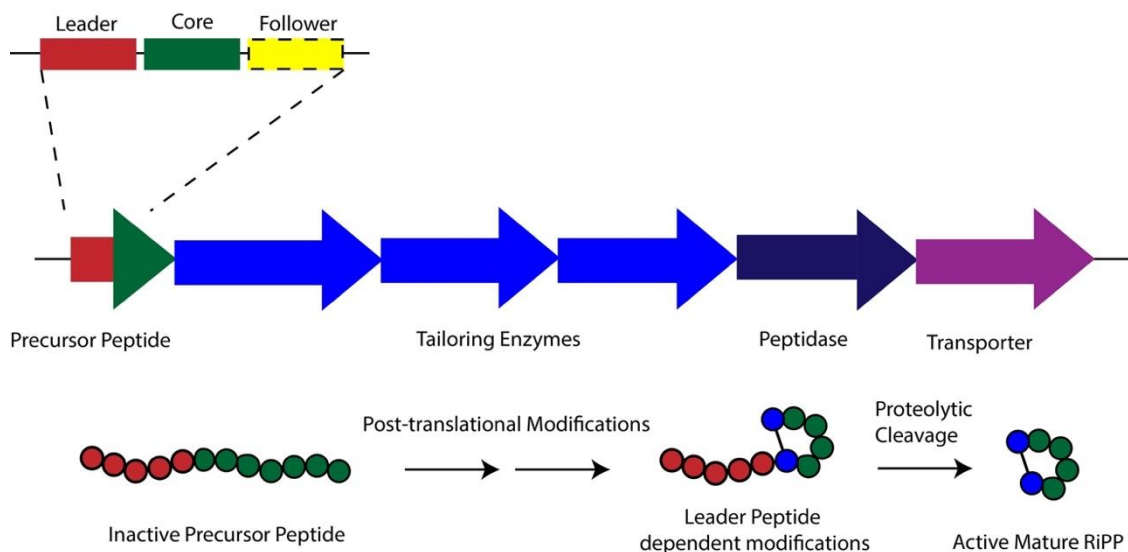
Figure 1.2 Biosynthesis of RiPPs. Generalized RiPP biosynthetic gene cluster with precursor peptide, tailoring enzymes, and proteolysis/export enzymes. The RiPP precursor peptide shows leader, core and follower regions separating the sites responsible for substrate binding and residues that are modified.

1.2 Radical S-adenosylmethionine enzymes

One family of enzymes that has become increasingly common in RiPP maturation is the radical S-adenosylmethionine (rSAM) superfamily. rSAM enzymes have been shown to catalyze a remarkable range of post-translational modifications on RiPP precursor peptides, oftentimes using unprecedented chemical strategies to modify peptides into complex natural products.[20]

The first studies on this type of enzymes date back to 1960s where the pyruvate formate - lyase activating enzymes[21,22] and lysine 2,3-aminomutase[23] were found to be dependent on SAM and Fe (II), but it was not fully classified as a novel superfamily until 2001.[24] At the time of the bioinformatic study it included over 600 members,[24] now this number

4

surpassed 720,000 known members across all domains of life making this superfamily the largest known while most of its members being currently uncharacterized.[25]

All rSAM enzymes have a conserved CX_3CX_2C motif where C is a cysteine residue. Those three cysteines coordinate three out of four iron atoms in the [4Fe-4S] cluster. The fourth iron atom, usually referred as unique iron, forms a coordination complex with the amino and carboxylate moieties of SAM and defines the catalytic mechanism (Figure 1.3).[26]
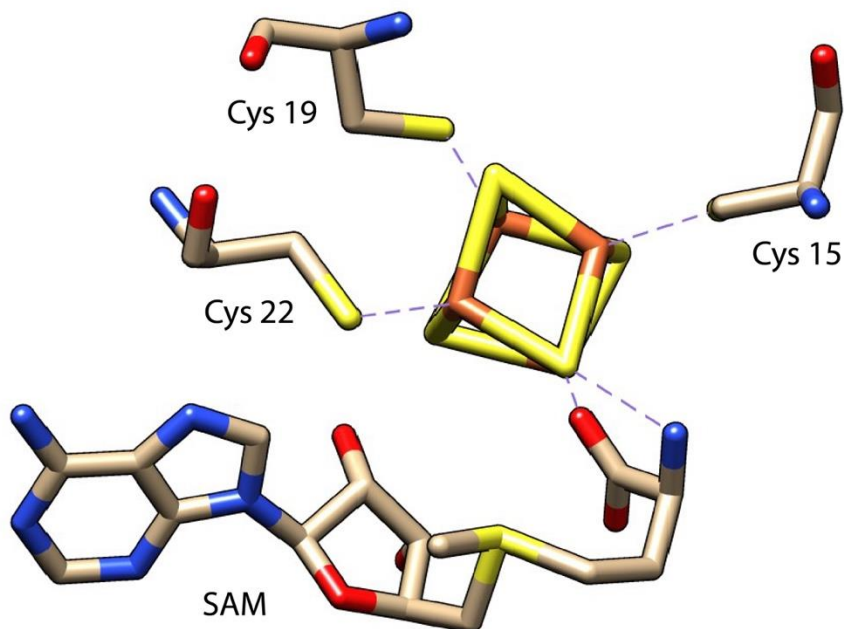
Figure 1.3 The rSAM cluster of anSME structure (PDB: 4K36). The conserved cysteine motif coordinates 3 iron atoms in a [4Fe-4S] cluster and the 4th iron atom is coordinated by SAM ligand.

The reaction begins with the reduction of the $[4Fe-4S]^{2+}$ state to the active $[4Fe-4S]^+$ oxidation state followed by inner-sphere electron to SAM resulting in the homolytic cleavage of the C5-S bond generating 5'-deoxyadenosyl radical and methionine (Figure 1.4 A). The 5'-deoxyadenosyl radical abstracts a hydrogen atom from the peptide substrate, leading to the formation of an alkyl radical (Figure 1.4 B). Subsequently, the alkyl radical recombines with an electron-rich functional group resulting in various outcomes. The types of post-translational modifications that have been observed include, but are not limited to, the formation of carbon-carbon bonds,[27–29] carbon-sulfur bonds,[28,30] carbon-oxygen bonds,[31] and epimerization,[32] amongst others, on the precursor peptide. As a result of these modifications, rSAM-dependent natural products have been shown to be important biological molecules such as antibiotics,[19,33] essential redox cofactors,[34,35] and quorum sensing molecules.[31,36]
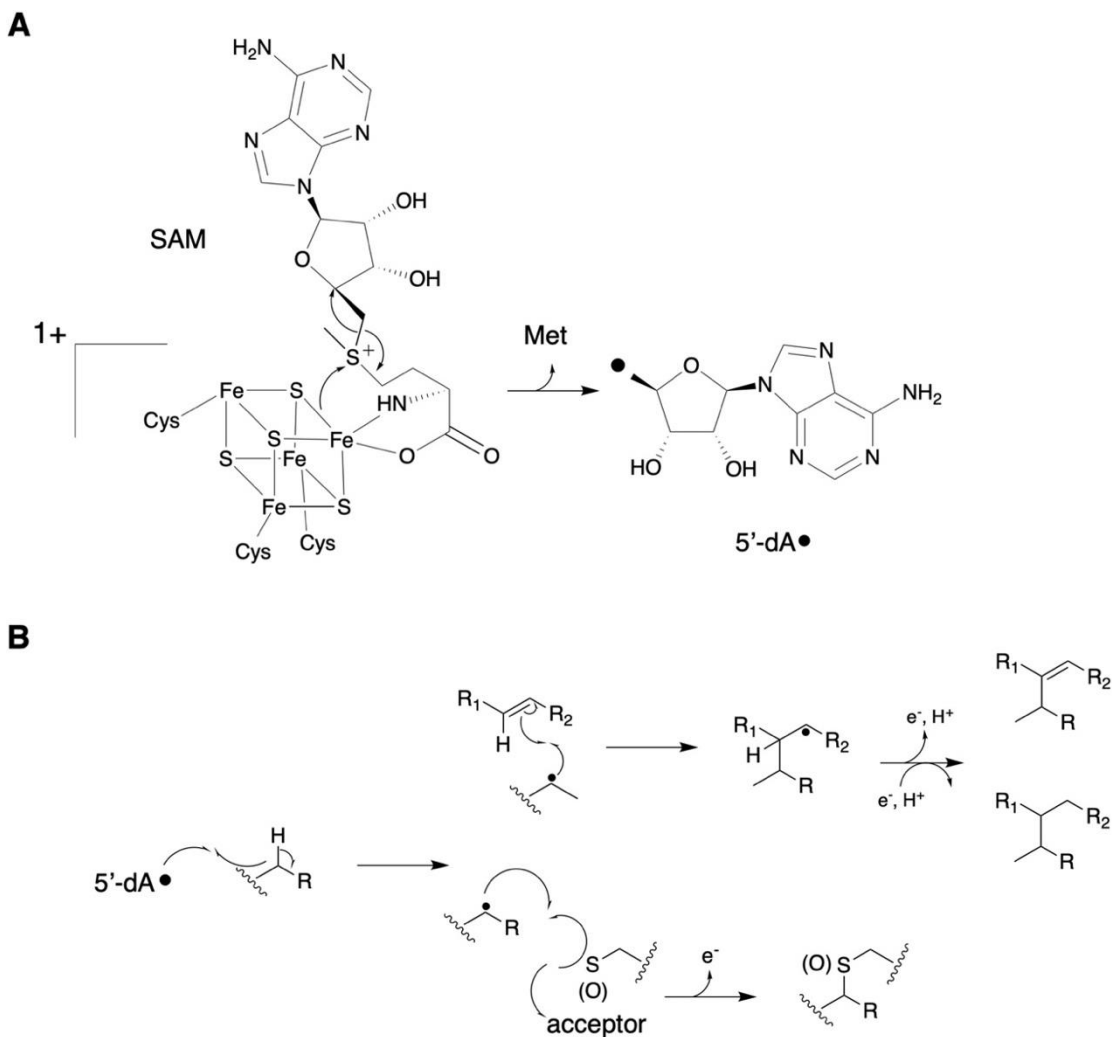
Figure 1.4 Mechanism of action of rSAM enzymes. (A) Schematic representation of the homolytic cleavage of SAM resulting in 5'-deoxyadenosyl radical (5'-dA•). (B) Mechanistic scheme for rSAM-enzyme-catalyzed bond formation.

1.3. Foundation of rSAM-SPASM enzymes in RiPP biosynthesis

Arguably, the most transformational set of studies in the rSAM-dependent RiPP field has involved a protein structure, a bioinformatic study, and a protein function.

To begin with, the rSAM protein structure referred to stemmed from the study of sulfatase maturation in bacteria. For anaerobic sulfatase to be functional, it must first undergo posttranslational oxidation of serine or cysteine to form the catalytic residue α-formylglycine (FGly).[37,38] In prokaryotes the enzyme performing this oxidation was found to be a member of the rSAM superfamily, AtsB.[39] AtsB was shown to oxidize serine through a 5'- deoxyadenosyl radical mediated abstraction of the C_β-hydrogen yielding an alkyl radical.[39] Loss of an electron and proton result in the formation of FGly (Figure 1.5).[40] From these pioneering studies, related enzymes with a similar function have been classified as the anaerobic sulfatase maturating enzymes (AnSMEs) subfamily.[41]
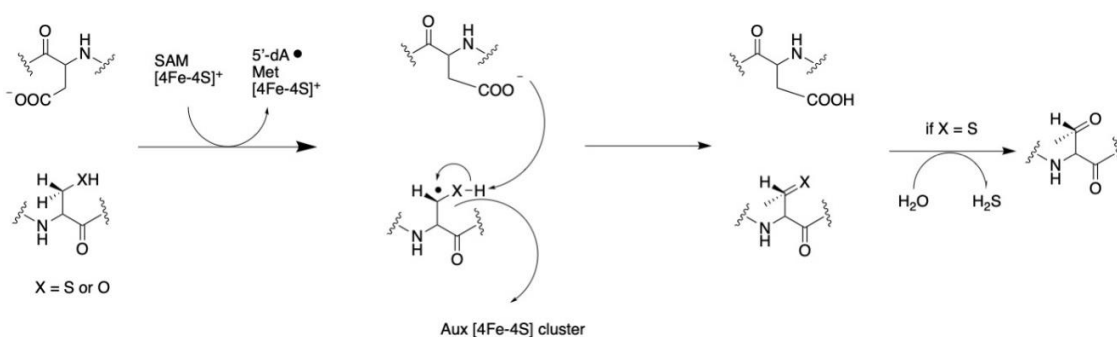
Figure 1.5 Formylglycine formation mechanism for cysteine (AnSME) and serine (AtsB) sulfatase maturing enzymes.

Structural characterization of AnSME revealed the protein to contain an elongated C-terminal domain that coordinates two auxiliary [4Fe-4S] clusters, in addition to the rSAM [4Fe-4S] cluster bound by the TIM barrel domain (Figure 1.6).[42] This extra C-terminal domain is annotated as a SPASM domain, named after the founding members subtilosin A, pyrroloquinoline quinone, anaerobic sulfatase, and mycofactocin.[43,44]

8
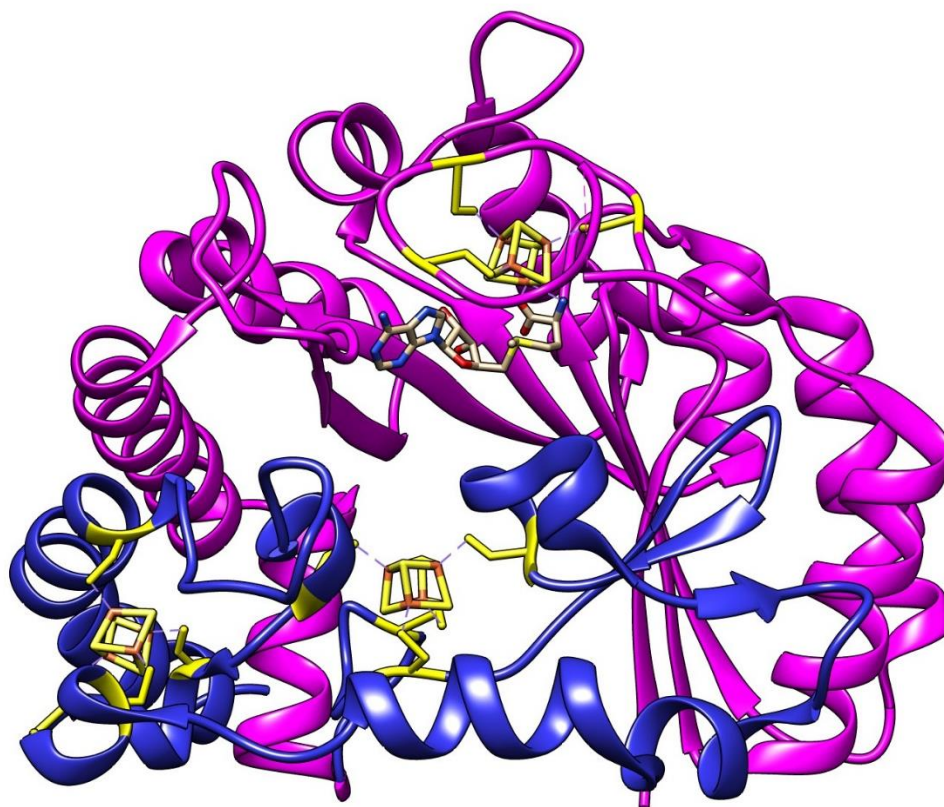
Figure 1.6 Structure of anSME (PDB ID 4K37). The RS domain contains the rSAM cluster and the $(\beta/\alpha)_6$ partial TIM barrel domain (magenta). The SPASM domain (blue) is the C-terminal extension containing 2 auxiliary [4Fe-4S] clusters. The cysteines involved in the binding of [4Fe-4S] clusters are shown in yellow.

While AnSME is not a RiPP maturase, its structure is homologous to RiPP modifying rSAM-SPASM enzymes and thus has been pertinent to the RiPP field. Interestingly, AnSME has structural homology with another C-terminal extension called a Twitch domain. The Twitch domain, found in BtrN and MoaA, is a truncated SPASM domain that binds a single auxiliary Fe-S cluster.[42,45] We will be focusing on rSAM-SPASM enzymes and the rSAM-Twitch enzymes SkfB, TqqB. We include SkfB and TqqB because of their
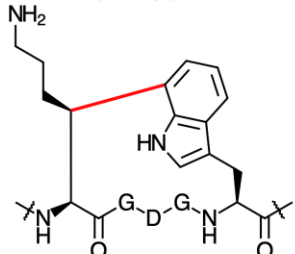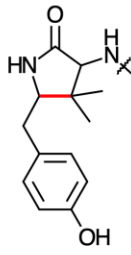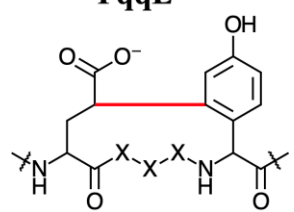
similarities in chemistry to rSAM-SPASM enzymes. Since publication of the AnSME structure, only three additional rSAM-SPASM proteins, out of ~60,000 annotated in the Interpro database (IPR023885), have had their structures solved (PqqE, SuiB, and CteB). Notably, the conventional thought is that the SPASM domain binds two auxiliary [4Fe-4S] clusters however, SPASM domains have also been reported to bind a [2Fe-2S] cluster in place of one auxiliary [4Fe-4S] cluster.[27]
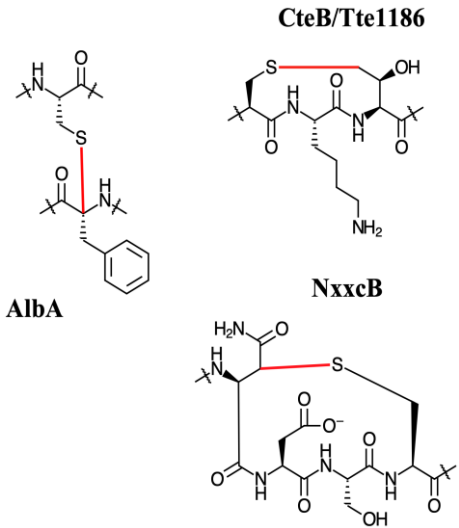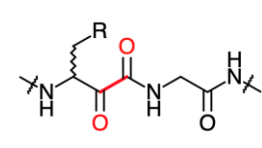
While the structure of AnSME was being elucidated, Haft and Basu reported on the association of rSAM-SPASM proteins to biosynthetic pathways that encoded for a peptide.[44] At the time, it was established that the redox cofactor, pyrroloquinoline quinone (PQQ), was synthetized from the precursor peptide PqqA.[46] In addition, it was known that the pqq biosynthetic gene cluster encoded for a rSAM-SPASM protein, PqqE, which was later shown to install a C-C crosslink between a conserved Tyr and Glu on the peptide PqqA.[27,47] The association of a peptide and a rSAM-SPASM protein, along with other bioinformatic evidence, led Haft to discover the mycofactocin (MFT) biosynthetic pathway.[48] Like PQQ, the redox cofactor MFT is synthesized from a precursor peptide and hinges on a critical transformation catalyzed by the rSAM-SPASM enzyme MftC.[49,50] Expanding on the association with peptides, Haft and Basu leveraged rSAM-SPASM proteins as molecular markers to discover new multicomponent (> 2 modifying enzymes) RiPP biosynthetic pathways.[44] Their approach used partial phylogenetic profiling and hidden Markov models (HMMs) to identify rSAM-SPASM proteins that were genetically clustered with putative precursor peptides.[44] From this study, Haft and Basu identified five

previously unknown RiPP natural product families that rely on rSAM-SPASM enzymes, some of which have been experimentally validated.[30,51] In addition, Haft and Basu built a collection of 68 protein subfamilies based on HMMs that are non-overlapping and functionally distinct.[44] The abundance of rSAM-SPASM proteins and their association with peptides led Haft and Basu to postulate that rSAM-SPASM proteins are commonplace in RiPP biosynthesis, a notion that is coming to fruition.

A third major influence in the rSAM-dependent field was the study of AlbA and its involvement in synthesis of Subtilosin A.[52–54] Subtilosin A is a bacteriocin that is comprised of a circular peptide containing three intramolecular thioether bridges.[52–54] Installation of the thioether bonds in Subtilosin A is carried out by the rSAM-SPASM enzyme AlbA. Like AnSME, AlbA was found to bind three [4Fe-4S] clusters and catalyzes the 5'-deoxyadenosyl radical mediated hydrogen abstraction from the precursor peptide, in this case SboA.[55] Unlike AnSME, AlbA was shown to abstract hydrogen from a C_α, forming the initial alkyl radical.[56] Subsequent radical recombination with the thiol/thiolate of cysteine and loss of an electron and proton forms the intramolecular thioether bond. RiPPs that contain C_α - S bonds are classified as sactipeptides (sulfur-to-alpha carbon thioether cross-linked peptides),[57] and include sporulation killing factor (SKF),[58] the thuricin family,[59,60] ruminococcin C,[61] six-cysteines-in-forty-five (SCIFF),[30] and streptosactin.[62] The pivotal study of AlbA, together with the bioinformatic study discussed above, effectively launched the growing field of rSAM-dependent RiPPs (Table 1).

Table 1.1 Current list of known rSAM-SPASM proteins and their respective products.

Protein	Product	PDB ID	Linkage	Example Structures
PqqE	Pyrroloquinoline Quinone	6C8V	C_γ-C3	
StrB/ SuiB	Streptide	5V1Q, 5V1T, 5V1S	C_β-C8	
MftC	Mycofactocin		C_α-C_β	
RrrB	Ryptides		C_δ-C3	
WgkB	Benzindole		C_α-C5 C_δ-C6	
XyeB	Cyclophane		C_β-C6 C_β-C4	
GrrB	Cyclophane		C_β-C4 C_β-C7	
FxsB	Cyclophane		C_β-C4	
AlbA	Subtilosin A		C_α-S	
SCIFF Maturase	Thermocellin	5WHY 5WGG	C_γ-S	

(CteB, Tte1186)				
SkfB (Twitch)	Sporulation Killing Factor	6EFN	C_α-S	
ThnB	Thuricin H		C_α-S	
QhpD	Mature QhpC		C_β-S C_γ-S	
RumC	Ruminococcin C		C_α-S	
QmpB	Sactipeptide		C_α-S	
NxxcB	Ranthipeptide		C_β-S	
GggB	Streptosactin		C_α-S	
PapB	Freyrasin		C_β-S	
TqqB (Twitch)	Rotapeptide		C_α-O	
PlpD	Unknown			

1.4 Expanding the realm of rSAM-SPASM-derived RiPPs

Following the publication of AlbA,[55] the Marahiel group continued to be prolific in the discovery of sactipeptide biosynthesis. First, they reported on SkfB, demonstrating that it catalyzed the formation of a thioether bond between a C_α - S bond on the precursor peptide to SKF.[58] Second, they showed that ThnB, installed a C_α - S bond on the precursor peptide of Thuricin H.[63] Soon after, the functions of Tte1186,[30] CteB,[64] and RumC,[61] amongst others, were discovered. While reports of C_α - S thioether bridges were predominant in the early stages of rSAM-dependent RiPP discovery, it should be noted that C_β-S and C_γ-S bonds, or ranthipeptides (radical non-α thioether peptides), have since become increasingly common. The first rSAM-SPASM enzyme to catalyze non-C_α-S thioether bonds was QhpD.[65] QhpD installs both a C_β-S and a C_γ-S thioether bond during the maturation of the γ - subunit (QhpC) of quinohemoprotein amine dehydrogenase.[65] The family of rSAM-SPASM enzymes involved in ranthipeptide biosynthesis has since expanded to include NxxcB[66] and PapB.[67] While sacti/ranthipeptides predominated early on, a second class of rSAM-dependent RiPPs has emerged.

Carbon-carbon bonds are not trivial to make yet rSAM-SPASM proteins have made them commonplace. The first reported in vitro C-C bond formation by a rSAM-SPASM in RiPP biosynthesis involved streptide.[29,68] Streptide is a quorum sensing molecule that is synthesized from the precursor peptide, StrA, by the str gene cluster.[68] Within the cluster, the rSAM-SPASM protein, StrB, was shown to install a C_β - C7 bond between Lys and Trp,

yielding a cyclophane.[29] Soon after, a PqqE was reported to catalyze the formation of a C_γ - C3 bond between Glu and Tyr on the precursor peptide PqqA, a critical step in the formation of PQQ.[27] More recently, XyeB, GrrB, FxsB, WgkB, and RrrB have all been reported to catalyze the formation of cyclophanes on their respective precursor peptide.[51,69–71] The linkages from non-aromatic carbons can occur from C_α, C_β, C_γ, or C_δ and typically attached to the C3 of Tyr, or less discriminately, to C4-C8 of Trp. From these studies, it is becoming more apparent that rSAM-SPASM proteins have been recruited to install cyclophanes in RiPP biosynthesis. An outlier in the formation of C-C bonds is MftC. As mentioned earlier, MftC is involved in the biosynthesis of the redox cofactor MFT.

Currently, MftC is the only rSAM-SPASM protein involved in C-C bond formation that does not yield a cyclophane. Instead, two independent studies demonstrated that MftC catalyzes the oxidative decarboxylation of the C-terminus of MftA, resulting in the α/β unsaturated tyramine (MftA**).[49,50] However, mechanistic studies later demonstrated that MftC catalyzes a two-step reaction. Following the decarboxylation, MftC installs a lactam moiety by catalyzing the C-C bond formation between the C_α of the penultimate Val and the C_α of the former Tyr.[34] While MftC is the outlier in the family of C-C bond forming rSAM-SPASM enzymes, it promises that potentially unusual and difficult posttranslational modifications are to come. Lastly, rSAM-SPASM proteins have been shown to install C-O bonds on RiPP precursor peptides. Recently, the Seyedsayamdost group reported on the formation of an aliphatic ether in a streptococcal quorum sensing molecule encoded by the tqq biosynthetic gene cluster.[31] TqqB was shown to catalyze the formation of a C-O bond

between a Thr derived alcohol and the C_α of an adjacent Gln on the precursor peptide TqqA.[31]

1.5 rSAM-SPASM story continues to evolve

We now know that rSAM-SPASM enzymes are widely used peptide modifying enzymes, with expanding functionality. The development of new bioinformatic tools has led genome-based discovery of rSAM dependent RiPPs and has been successfully employed in the discovery of sactipeptides,[64] ranthipeptides,[72] and ryptides,[71] to name a few. Different strategies have been implemented for RiPP mining, such as targeting conserved tailoring enzymes[73–75] or targeting the precursor peptide in their search.[76] For example, BAGEL performs rule-based strategy where it implements six frame translation which are used to search for the motifs and core peptides and classifies them into the RiPPs class.[77,78] The core peptide is detected by homology to the already known core peptides or the expected properties of the given class.[77] In contrast, RODEO implements profile HMM-based local genomic analysis and precursor peptide/structure prediction by implementing the heuristic scoring, motif analysis and machine learning to detect RiPPs.[76] Notably, the sactipeptide RiPP biosynthetic gene cluster was mapped via the rSAM-SPASM enzymes which catalyze the sactionine bond by employing the recent improved version of RODEO 2.0.[72] A search was performed to identify candidate proteins using rSAM-SPASM sequences of known sactipeptides (AlbA, SkfB, ThnB ThnC) and two SCIFFs (CteB, Tte1186).[72] Heuristics with support vector machine classification were employed in the putative precursor peptides search.[72] Then a sequence similarity network

16

(SSN) using the Enzyme Function Initiative Enzyme Similarity Tool (EFI-EST) was employed, to group the similar enzymes and distinguish them from novel RiPPs, for further precursor peptide analysis.[79] By doing so Hudson et al. were able to characterize the sactipeptide huazacin and identify hundreds of new sactipeptide BGCs. Interestingly, the rSAM-SPASM protein responsible for the biosynthesis of huazacin is related to QhpD.[72] This association led Hudson et al. to determine that the SCIFF family members freyrasin and thermocellin contain S-C$_\beta$ and S-C$_\gamma$ linkages, respectively. Subsequently, ranthipeptides have been designated as a new class of RiPPs since they are structurally distinct from sactipeptides.[72] Taken together, the work performed by Hudson et. al. indicate that the precursor peptide or peptide related tailoring enzymes could be used as markers in the mining of novel RiPPs.[72,78]

1.6 RiPP-derived redox cofactor mycofactocin

Mycofactocin is a RiPP derived redox cofactor found primarily in mycobacteria. The MFT biosynthetic gene cluster (BGC) consists of mftABCDEF (Figure 1.7A) and is highly conserved in mycobacteria, including pathogens, such as Mycobacterium tuberculosis (MTB), Mycobacterium ulcerans, Mycobacterium avium, and Mycobacterium bovis.[80]

As shown in Figure 1.7B, MFT biosynthesis starts with the MftC-catalyzed oxidative decarboxylation of the C-terminal Tyr forming MftA**[49,50] and the subsequent formation of a C–C bond resulting in a lactam derived from penultimate Val, MftA*.[34] Both reactions are dependent upon the RiPP recognition element (RRE) MftB, which binds MftA and delivers it to MftC.[49] Next, MftE hydrolyzes MftA*, forming 3-amino-5-[(p-

17

hydroxyphenyl) methyl]-4,4-dimethyl-2-pyrrolidinone.[81] Following cleavage, MftD catalyzes the FMN-dependent oxidation of the 3-amino group, resulting in an α-keto-amide moiety within the lactam, forming premycofactocin (PMFT).[82] Finally, a recent metabolomics analysis has indicated that MftF glycosylates PMFT with up to eight β1–β4 glucans, forming mature MFT.[83]

The figure below shows how precursor peptide MftA is modified by tailoring enzymes resulting in a natural product, mycofactocin, in general. However, each individual enzyme, will be discussed in greater detail in the following sub-chapters.
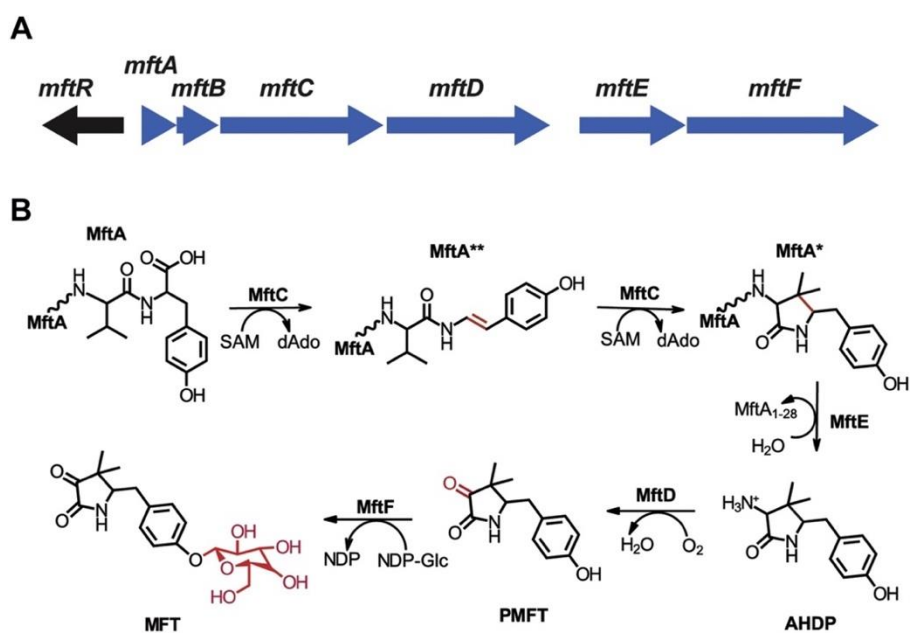
Figure 1.7 Mycofactocin biosynthetic gene cluster. (A) Schematic depiction of the gene organization of the MFT BGC. (B) MFT biosynthesis model; the enzyme modifications are highlighted in red.

1.6.1 rSAM-SPASM maturase defines the structure of mycofactocin

MftC is an rSAM-SPASM protein that performs the first step in the biosynthesis of mycofactocin. The reaction happens only in the presence of a discrete RRE - containing protein MftB.[84] MftB acts by bringing MftA and MftC to catalytically relevant proximity and orientation to form a ternary complex for the reaction to occur.[84] Two independent studies showed that MftC catalyzes the oxidative decarboxylation of the C-terminal MftA resulting in the α/β unsaturated tyramine MftA∗.[34,50] Further studies demonstrated that MftC converts MftA∗∗ to a lactam moiety MftA∗ by catalyzing the C-C bond formation between the C_β of the penultimate Val and the C_α of the former Tyr.[85]

The ability of MftC to catalyze two distinct reactions is fascinating. The biochemical, spectroscopic, and electrochemical studies were performed to shed light on the function of the [4Fe-4S] clusters.[85] It was shown that MftC contains three [4Fe-4S] clusters, one rSAM and two auxiliary clusters (Aux I, Aux II), all required for catalysis. Interestingly, the midpoint potential values for the Aux I and Aux II were within 50 mV allowing the clusters to undergo a reversible electron transfer between each other. Based on the similarity in the redox potential the mechanism for MftC catalysis has been proposed (Figure 1.8).[85] The crystal structure characterization of MftC is needed to confirm the proposed mechanism.
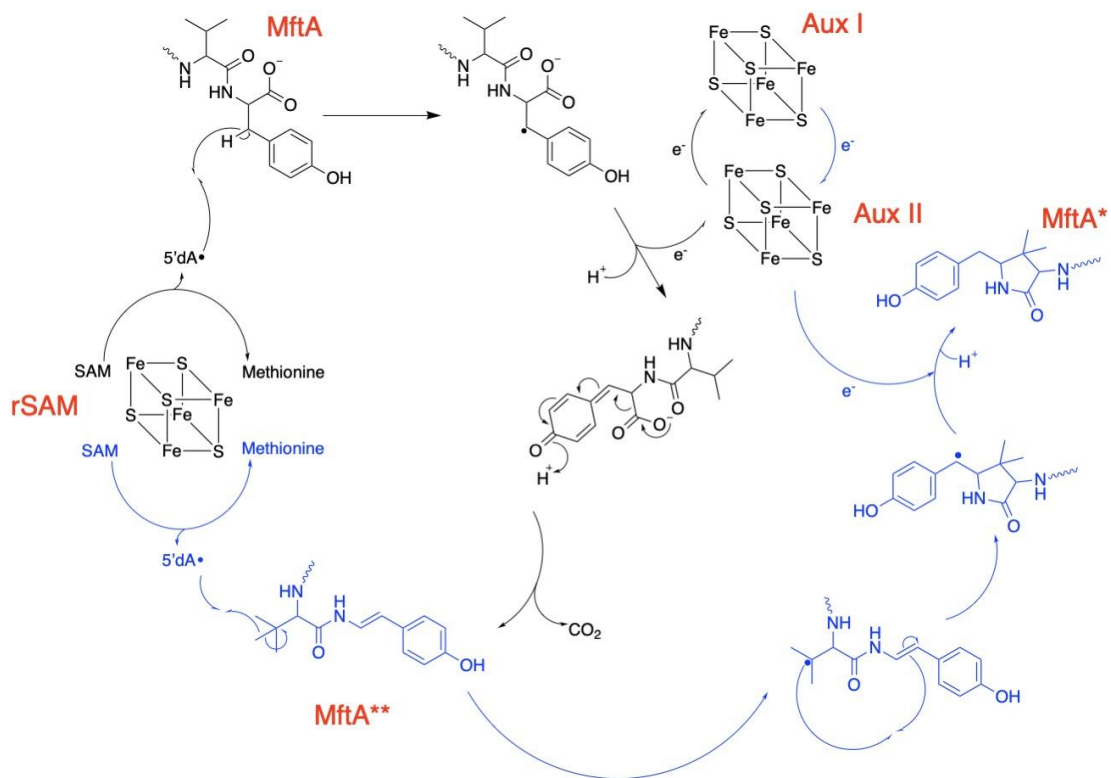
Figure 1.8 Proposed mechanism for MftC catalysis. All [4Fe-4S] clusters are required for catalysis. Decarboxylation step is shown in black; carbon-carbon crosslink is in blue.

The rSAM cluster, in the presence of sodium dithionite, catalyzes the cleavage of SAM and formation of 5′dA•. The 5′dA• then abstracts a H-atom from the substrate producing the substrate intermediate. An electron from the substrate is shuttled into the Aux I to form a product intermediate which in turn is shuttled to Aux II to catalyze the first step of the reaction, the oxidative decarboxylation.[86] The same electron is used in the second step of the reaction, the C-C bond formation, where it quenches the substrate radical to form MftA∗.[86]

1.6.2 MftE hydrolyzes MftA* to install AHDP moiety

MftE is an iron-dependent mycofactocin peptidase, a member of a large creatinine amidohydrolase family.[87] MftE hydrolyzes MftA* and results in the formation of 3-amino-5-[(p-hydroxyphenyl) methyl]-4,4-dimethyl-2-pyrrolidinone (AHDP).[81] AHDP was structurally characterized by 1D, 2D NMR spectroscopy and mass spectrometry (MS).[81] The active site of the MftE is conserved with other members of creatinine amidohydrolases and similar mechanism of action is suggested. However, without the actual crystal structure of MftE it is hard to propose the mechanism.

1.6.3 MftD catalyzes the formation of PMFT

MftD is an FMN-dependent mycofactocin oxidoreductase. MftD catalyzes the oxidative deamination of AHDP and subsequent formation of premycofactocin (PMFT) in two steps: in the presence of molecular oxygen, the amino group is oxidized to an imino group and then spontaneously hydrolyzed to produce an α-keto-amide releasing ammonium.[88] The structural elucidation of the PMFT was characterized by 1D, 2D NMR spectroscopy and HRMS.[88] In the same study, it was also shown that PMFT acts as a redox center *in vitro*. Using cyclic voltammetry, the overall midpoint potential of PMFT for $2e^-$/$2H^+$ was measured to be -255 mV. PMFT was then shown to be reduced to $PMFTH_2$ by an MFT-dependent short chain dehydrogenase.[88]

1.6.4 MftF catalyzes the oligoglycosylation of PMFT

The last enzyme in the mycofactocin biosynthesis is a glycosyltransferase, MftF. The metabolomics study including stable isotope labeling, metabolite induction, MS/MS networking, knockout studies, complementation was used to discover mature mycofactocin

21

(MFT) *in vivo*.[83] According to the study, MFT consists of up to nine β-1,4-linked glucose units. The mature MFTs were isolated in oxidized (mycofactocinones/MFT-*n*) and reduced (mycofactocinols/ MFT-*n*H$_2$) forms providing strong evidence that MFT is indeed a redox cofactor. Interestingly, it was shown that glycosylation appears on earlier stages by the presence of oligoglycosylated AHDP and glycyl-AHDP. In addition, MS/MS networking found a recurring mass difference corresponding to the 2-O-methylation of the second glucose moiety of oligoglycosylated MFT-*n*.[83]

The function of MftF has not been confirmed *in vitro* but some attempts have been made and will be discussed in the results section.

1.6.5 MftM catalyzes methylation of MFT

As mentioned above the finding of the methylmycofactocinones (MMFT-*n*) was quite remarkable since there are no known methyltransferase genes encoded within the MFT biosynthetic gene cluster. However, the researchers from Lackner lab applied reverse genetics study and were able to find a SAM-dependent methyltransferase, msmeg_6237 or MftM, responsible for the methylation of mycofactocinones in *M. smegmatis*.[89] The knockout analysis with the subsequent metabolome analysis showed complete disappearance of MMFT-*n* in the Δmsmeg_6237 without influencing MFT-*n*, whereas the re-introduction of MftM leads to the production of MMFT-*n*.

22
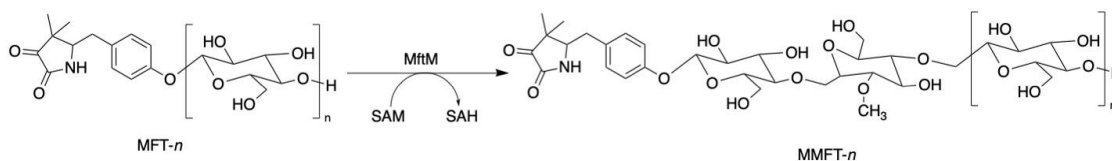
Figure 1.9 SAM-dependent methyltransferase MftM catalyzes reaction of mycofactocinones (MFT-*n*) to produce methylmycofactocinones (MMFT-*n*).

1.6.6 Mycofactocin-associated mycobacterial dehydrogenases

In addition to *mft* genes, three different dehydrogenase families (TIGR03971, TIGR03989, and TIGR04266) are found in genomes only when the *mft* BGC is present.[48] These MFT-dependent dehydrogenases have been shown to sequester NADH within their active sites and therefore require an additional electron acceptor, presumably MFT, to oxidize NADH for further catalytic turnover.[90] In support of this model, knockouts of the *mft* genes in *M. smegmatis* mc[2]155 (Msmeg) led to the inability of the organism to maintain homeostasis of cellular NAD$^+$/NADH pools and its inability to metabolize methanol and ethanol.[91,92] The failure of the knockouts to metabolize primary alcohols is likely due to the MFT-dependent alcohol dehydrogenase, Msmeg_6242, being trapped in a reduced state in the absence of MFT. More recently, a study demonstrated that *mftD*, and thus MFT, is required for Mtb survival *in vitro* and *in vivo* under hypoxic conditions.[93] However, until recently, direct evidence demonstrating that MFT is a redox cofactor was nonexistent. This changed when it was shown that both PMFT and MFT are capable of oxidizing MFT-dependent dehydrogenases.[83,88] Despite knowing the structure, biosynthesis, and redox attributes of MFT, information about physiological processes that require MFT has been lagging.

23

1.7 Cyclopropylglycine-forming synthases

1.7.1 TVG pathway

To discover new natural products, our lab used the sequence similarity network (SSN) of the IPR023867 as a bioinformatic marker targeting specifically rSAM-dependent RiPPs that have not been associated with previously known classes.[94] In doing so, we mapped >20 potentially new RiPP biosynthetic pathways. We experimentally validated the workflow by characterizing the rSAM enzyme TvgB, the precursor peptide TvgA with repeating TVGG motif from the TVG gene cluster and the product of their reaction (Figure 1.10 A).[94] The remaining enzymes of the TVG pathway consist of TvgC – an iron/α-ketoglutarate-dependent oxidase, TvgD – a putative transporter, TvgE – a putative peptidase, have not been characterized yet. We structurally characterized the TvgB product by MS, MS/MS, 1D/2D NMR and isotope labeling. TvgB installs the repeated formation of cyclopropylglycine (CPG) where a new bond is formed between the γ-carbons on the precursor valine.[94] To our knowledge, this is the first rSAM enzyme to catalyze the formation of a cyclopropane motif without being coupled to methyl addition in RiPP maturation.

Figure 1.10 TVG biosynthesis. (A) TVG biosynthetic gene cluster with the precursor peptide sequence with repeating TVGG motif. (B) Proposed mechanism for TvgB catalysis.

Our lab proposed a two-step mechanism for TvgB catalysis. To start with, the 5'-dA• abstracts a hydrogen from Val-C_β resulting in the β,γ-dehydro-Val.[94] In the second round, the 5'-dA• abstracts a hydrogen from the Val-C_γ resulting in the allylic radical which is then quenched by the addition of an electron and proton producing CPG.[94]

1.7.2 TIGSVS pathway

Discovery of the fascinating poly-CPG containing RiPP product catalyzed by rSAM-enzyme TvgB inspired our lab to expand the network of TvgB homolog by utilizing the radical-SAM.org tool.[94] We employed the SSN on cluster 1-1-201:AS65 that include TvgB

and 48 other proteins.[95] By doing the full work up, we found at least six conserved sequence motifs within the TVGG pathway, specifically the precursor peptides that had the conserved repeats of Val or Ile residues (Figure 1.11).
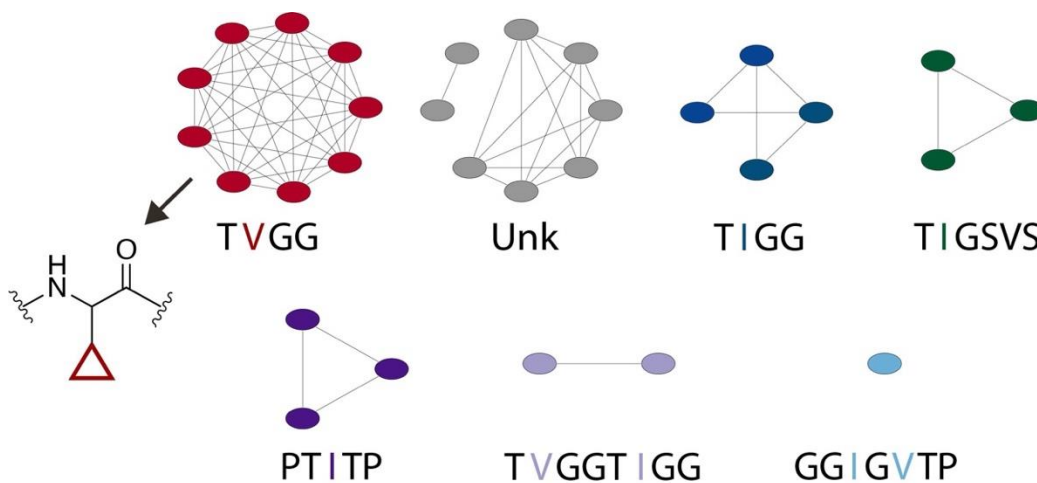
Figure 1.11 SSN of cluster 1-1-201:AS65 sorted according to the RiPP precursor peptides that require putative cyclopropyl synthases.

Out of the selected network, our lab chose TIGSVS pathway because of its unique gene cluster (Figure 1.12). The TIGSVS biosynthetic gene cluster consists of TigA – a distinct transporter, TigB – a precursor peptide with repeating TIGSVS sequence, TigC – an unknown non-homologous orphan protein, TigD – an rSAM enzyme, TigE – an rSAM enzyme, and TigF – a spdL-like protein.
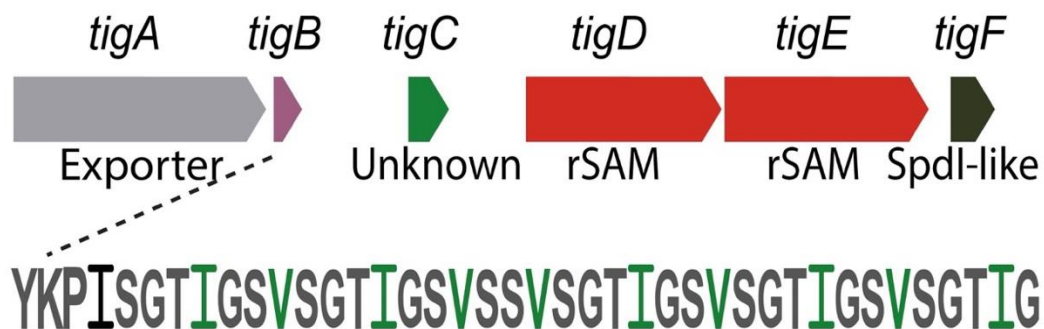
Figure 1.12 The TIGSVS biosynthetic gene cluster with the sequence of the precursor peptide with repeating TIGSVS sequence.

Our lab reconstituted enzymatic activity of TigE and characterized the product of the reaction from the precursor peptide TigB and rSAM enzyme TigE.[96] The MS analysis showed the loss of 2 Da per repeat and MS/MS localized the reaction to the Ile residue. 1D/2D NMR and isotope labeling experiments validated that TigE installs a carbon-carbon bond between the Ile-C_γ's on the TigB, forming a methyl – CPG moiety.[96]

1.8 Aims of the study

In this work, my advisor and I seek to understand how mycofactocin biosynthesis is regulated. Currently, it is thought that putative TetR- like protein MftR is a regulator of mycofactocin biosynthesis.[97,98] In general, TetR family of regulators (TFRs) are transcriptional repressors and implicated in the regulation of efflux pumps,[99] antibiotic biosynthesis,[100] the tricarboxylic acid cycle,[101] biofilm formation,[102] and quorum sensing molecules.[103] TFRs are functional as dimers that contain DNA-binding and regulatory-binding domains.[104] The DNA-binding domain consists of a helix-turn-helix (HTH) motif

27

that binds to a DNA operator sequence.[105] The regulatory domain consists of a binding pocket that specifically interacts with a variety of compounds, such as tetracycline,[106] biotin,[107] fatty acids,[108] flavonoids,[109] and cell-cell signaling molecules,[110] depending on the system.

MftR regulation of the mycofactocin biosynthetic gene cluster is supported by bioinformatics, which suggests the gene proximity of *mftR* and its arrangement to the *mft* BGC is indicative of its regulatory control of the pathway.[111] In addition, a transcriptomics study of macrophage samples infected with MTB showed that upregulation of the MftR homolog, Rv0691c, led to the repression of *mftB*, *mftC*, and *mftD*.[112] Currently, the specific DNA operator sequence that MftR recognizes, its regulatory role over *mft* BGC in Msmeg, and the conditions that MftR could regulate MFT biosynthesis are unknown.

1.9 Significance

Tuberculosis is caused by *Mycobacterium tuberculosis* (MTB) and remains one of the leading sources of mortality worldwide.[113] The success of MTB lies in its ability to persist for long periods of time and adapt to changing conditions within the human host. Typical treatments require multiple drugs taken over several months which highlights the efficiency of MTB as a pathogen.[114,115] Inability to complete the long antibiotic regimen leads to the emergence of antibiotic resistant MTB strains. Drug-resistant TB occurs in the form of multi-drug resistant (MDR) when it is resistant to two first-line drugs, rifampicin, and isoniazid; and extensively drug-resistance (XDR) when in addition to MDR it is also resistant to fluoroquinolones and at least on the three injectable second-line drugs.[116] The

combination of MTB physiology, human immune response, and antibiotic resistance are the major reasons why TB is so difficult to eradicate. To successfully develop improved and less toxic antimycobacterial drugs we need a better understanding of MTB metabolic pathways essential for growth, persistence, and survival.

Pan-genomic study of 36 MTB genomes identified genes encoded in mycofactocin biosynthetic pathway to be among the core genes essential for survival.[117] Mycofactocin is a RiPP-derived redox cofactor that consists of six conserved genes *mft*ABCDEF. It was first discovered by a bioinformatic study that show the correlation between nicotinoprotein families and mycofactocin, suggesting its role as an electron carrier for these NADH-dependent redox proteins for mycobacteria *in vivo*.[48] The products of MFT genes were suggested to be a part of the electron transport chain, which plays an important for MTB survival.[118] Additionally, *mft* genes were shown to be essential for ethanol assimilation,[91] methylotrophy,[92] hypoxia adaptation,[93] glucose metabolism,[119] redox homeostasis,[93] pyruvate metabolism.[120] Despite the importance of the mycofactocin pathway in MTB and progress that has been made so far, little information about the physiological conditions that lead to MFT production is known.

My project addresses this gap in knowledge by focusing on the putative transcriptional factor of mycofactocin. We demonstrate that MftR is indeed a transcriptional repressor of mycofactocin biosynthetic gene cluster.[121] Moreover, we show that the activation of MFT is dependent on the long chain acyl-CoAs. We suspect that MFT production and utilization is required for MFT - dependent dehydrogenases involved in fatty acid metabolism.

Chapter Two: Materials and methods

2.1 Materials and methods

All acyl-CoA's were purchased from Sigma Aldrich. Oligonucleotides used in this study were synthesized by Invitrogen. Double-stranded O_{mft}, and mutants thereof, were prepared from single-stranded oligonucleotides by mixing two oligos in equal molar amounts, heating the mixed oligos for 5 min at 95°C and then gradually cool the oligos to the room temperature. DNase I footprinting assays were contracted to Profacgen.

2.2 General procedures

Nuclear Magnetic Resonance Spectroscopy

All NMR spectra were recorded using Bruker UltraShield 500/54 Plus spectrometer at the University of Denver. TopSpin v.2.1 program (Bruker) was used to process and analyze the spectra. The samples were prepared in 99.8 % deuterated DMSO.

Quadrupole Time-of-Flight Liquid Chromatography Mass Spectrometry

All mass spectrometry analysis were carried out using electrospray ionization in positive mode on a Shimadzu Prominence-I LC-2030 HPLC coupled to a Shimadzu LC-QTof-MS 9030 at University of Denver. Data were collected and processed by LabSolutions Insight.

2.3 Expression and purification of MftR

The MftR gene (Uniprot: A0QSB5) from *Mycobacterium smegmatis* mc^{2}155 was cloned into pET-28a (Novagen) using the Nde I and Xho I restrictions sites. The sequence verified plasmid was transformed into E. coli BL21(DE3) pLysS cells. An overnight culture was used to inoculate 1 L of terrific broth (TB). The cells were grown at 37°C and 220 rpm until OD$_{600}$ ~ 1.0 and at which point 1 mM IPTG was added to induce protein production. The temperature was dropped to 21 °C and the cells were grown overnight. The cells were then centrifuged at 5000 *g* for 10 min at 4°C. The resulting pellet was resuspended in 5 times volume of lysis buffer (50 mM HEPES, 300 mM NaCl, 40 mM Imidazole, pH 7.5). To the suspension, 1% w/v CHAPS, 0.1 mg/g of lysozyme and 0.05 mg/g of DNase were added, and the lysate was stirred for 20 min at room temperature. The lysate was disrupted by sonication at 50% output with a pulse of 3 s on and 3 s off for 5 min. The lysate was centrifuged at 20,000 *g* for 10 min and the supernatant was loaded onto a 5 mL His-Trap column (GE Healthcare) using an Akta Pure FPLC. The bound protein was washed with 25 mL lysis buffer, eluted with elution buffer (50 mM HEPES, 300 mM NaCl, 300 mM Imidazole, pH 7.5), and buffer-exchanged into storage buffer (50 mM HEPES, 300 mM NaCl, 10% glycerol, pH 7.5) over a HiPrep 26/10 desalting column. The protein was concentrated using 10 kDa spin concentrator (Millipore) at 5000 *g* for 20 min at 22 ºC. The purity of the protein was evaluated by SDS-PAGE.

2.4 Construction of P~mft~-*mftA*-pCherry

The *mft* promoter region from *M. smegmatis* mc^{2}155 was synthesized and cloned into the pCherry vector by Genscript. In short, the native sequence for the *mft* promotor region from -698 to 0 nt through the native sequence of *mftA* (+93 nt) was cloned into the Xba I and BamHI restriction sites of pCherry, replacing the existing *smyc* promoter. The existing pCherry fluorescent protein was fused to the C-terminus of MftA by removing the native stop codon in the *mftA* sequence. The full sequence of the plasmid can be found in the supplementary information.

2.5 Construction of pMftR+

The wild type *mftR* gene from *M. smegmatis* was PCR amplified from genomic DNA and cloned into the pCherry vector using the BamHI and Hind III restriction sites.

2.6 Electrophoretic mobility shift assays

His-tagged MftR was used to assess the protein binding to the P~mft~-MftA promoter fragment. DNA was mixed with increasing concentrations of the MftR protein in the reaction buffer (50 mM HEPES, 50 mM NaCl 0.1 mM EDTA, pH 7.5) and incubated at room temperature for 10 min. The reactions products were resolved by electrophoresis on a 5% (v/v) non-denaturing polyacrylamide gel in 1X TBE buffer at 200 V for 20 min on ice. Prior to the analysis, the gel was pre-run for 30 min at 150 V on ice. Results were visualized by GelRed and recorded using Azure Biosystems 600 imaging system.

2.7 Preparation of fluorescent FAM labeled probes

The promoter region was PCR amplified with 2x TOLO HIFI DNA polymerase premix from Pmft-pCherry. The FAM-labeled probes were purified by the Wizard SV Gel and PCR Clean-Up System (Promega, USA) and were quantified with NanoDrop 2000C (Thermo, USA).

2.8 DNase I footprinting assays

DNase I footprinting assays were performed similar to Wang et al [122]. Approximately 350 ng of DNA probes were incubated with 0 and 2 µg of MftR in a total volume of 40 µl of the same buffer as previously described in EMSA reactions. After incubation for 30 min at 30 °C, a 10 µl solution containing approximately 0.015 U DNase I (Promega) and 100 nmol freshly prepared $CaCl_2$ was added to the probe/MftR mixture and further incubated at 37 °C for 1 min. The reaction was stopped by adding 140 µl DNase I stop solution (200 mM unbuffered sodium acetate, 30 mM EDTA and 0.15% SDS). Samples were firstly extracted with phenol/chloroform, then precipitated with ethanol. Pellets were dissolved in 30 µL MilliQ water. The preparation of the DNA ladder, electrophoresis and data analysis were as described in Wang et al, except that the GeneScan-LIZ600 size standard (Applied Biosystems) was used.

2.9 Fluorescence polarization assays

Fluorescence polarization binding assays were carried out in binding buffer (50 mM HEPES, 50 mM NaCl, pH 7.5) containing 0.5 µM fluorescein labeled O_{mft} using Tecan Infinite M1000. The MftR protein was titrated into the binding solution. The fluorescein-

33

labeled O_{mft} was mixed with increasing concentrations of MftR in a 96-well plate (Corning), like the EMSA reactions, and were incubated for 10 min before the measurements were taken. Excitation and emission wavelengths of 470 nm and 525 nm were monitored. All the experiments were done in triplicate. The G-factor was calculated from a solution of free fluorophore. The data was plotted and analyzed using GraphPad Prism. The mutated fluorescein labeled O_{mft} were measured as described above as well.

2.10 Construction of MftR mutants

Mutant strand synthesis reactions were performed according to the Agilent QuikChange site-directed mutagenesis protocol. The primers used can be found in the supplementary information for each desired mutation. The *mftR*_pET28a plasmid was used as a template for site-directed mutagenesis studies. Sequence verified plasmids were transformed into E. coli BL21(DE3) pLysS cells for protein production and purified using the same protocol as wild-type MftR.

2.11 Isothermal titration calorimetry measurements

Isothermal titration calorimetry (ITC) was performed using TA Nano ITC. MftR was loaded in the sample cell at the concentration of 16 µM and 250 µM acyl-CoA's (C12-C18) were loaded in the syringe. To minimize the effect of buffer-mismatch the stock concentrations of acyl-CoAs were prepared using ITC buffer (50mM HEPES, 50mM NaCl, pH 7.5) and the pH was adjusted to within 0.05 units. In addition, MftR was buffer exchanged into ITC buffer using a PD-10 column (GE Life Sciences). The volume of the titrant added at each injection into the sample cell was 2.22 µL. The time interval between

the successive injections was 180 s. The temperature of the cell was kept at 25 ℃. The data obtained were fit by a independent one site model using NanoAnalyze Data Analysis Version 3.8.0 that was provided with the instrument.

2.12 Electroporation of *M. smegmatis*

1 liter of 7H9-OADC media was inoculated with Msmeg and incubated at 37 ℃ with shaking until an OD_{600} ~ 1.0 was reached. The cell culture was incubated on ice for 1.5 h and the cells were harvested by centrifugation at 2,000 *g* for 10 min. The pelleted cells were suspended in 500 mL ice cold 10% glycerol and centrifuged again. Following two additional wash processes, the cells were suspended in 25 mL of ice cold 10% glycerol and transferred to a 50 mL conical tube. The cells were pelleted by centrifugation as described above, suspended in 4 mL of ice cold 10% glycerol, aliquoted to 300 μL, and flash frozen or used immediately. To a thawed aliquot of competent cells, ~ 1 μg of plasmid DNA was added and incubated on ice for 5 min. The cells were transferred to a 0.2 cm electrode gap cuvette and pulsed with a Gene Pulser (Bio Rad) set to 2.50 kV, 25 μF, and 1000 Ω. Following electroporation, 1 mL of 7H9-OADC was added to the cells. The suspension was then incubated for 2 h at 37 ℃ and plated on 7H10-ADC plates supplemented with 50 μg/mL of hygromycin B.

2.13 Growth curve analysis

For fluorescence-based assays, a single Msmeg colony harboring Pmft-pCherry was used to inoculate 25 mL of Middlebrook 7H9 (7H9) media supplemented with 0.1% tyloxapol, 0.1% w/v glycerol, 50 μg/mL hygromycin, and 0.1% w/v glucose. For cultures

35

containing oleic acid (OA) only, 0.1% w/v OA was used. For cultures containing both OA and glucose, 0.05 % w/v glucose was used, OA was added 12h later to obtain a concentration of 0.05% w/v. For dose-dependent cultures containing oleic acid (0.01%, 0.02%, 0.05%, 0.1% w/v), glucose was supplemented to the media to 0.05% w/v. The starter culture was incubated at 37 °C with shaking overnight. The overnight culture was used to inoculate 7H9 media to an $OD_{600} = 0.05$ (1 cm pathlength) with carbon sources supplemented as described above. The cultures (30 mL) were incubated at 37 °C with shaking for 40 h. Aliquots were taken every 2 h and OD_{600} values were measured. To track the expression of the P_{mft} - pCherry fusion, fluorescence measurements were taken on a Cary Fluorimeter every 2 h using an excitation of 585 nm and an emission of 612 nm.

2.14 Spot plate dilution growth analysis

A single timepoint growth assay was standardized for the growth of Msmeg on 7H9 agar plates with various carbon sources supplemented. Briefly, a primary culture of Msmeg/Pmft-pCherry was grown in 7H9 supplemented with 0.1% w/v glucose overnight. The OD was measured and converted to colony forming units (cfu's) using the standard 3.13 x 10^7 cfu·mL^{-1}·OD^{-1} conversion factor [123]. Five-fold serial dilutions of cultures were made in 7H9 base media to yield 500, 100, 25, and 5 cfu/μL solutions. Two microliters of each dilution were transferred to 7H9 plates containing 0.1% w/v carbon source and the spots were allowed to dry. Following incubation at 37 °C for 3 days for the growth of individual colonies, plates were imaged using an Azure 600 fluorescent imager under identical conditions.

2.15 RNA isolation, cDNA synthesis, and qPCR

RNA was purified from *M. smegmatis* using a Zymo Fungal/Bacterial RNA MiniPrep Kit and contaminating DNA was removed using a Zymo RNA Clean and Concentrator Kit. The quality of RNA was checked by agarose electrophoresis prior to subsequent steps. cDNA was synthesized from RNA samples using the First Strand cDNA Synthesis Kit from NEB. qPCRs were performed using Luna Universal qPCR Master Mix (NEB) following the recommended protocols. Primer pairs were designed using the IDT Primer Quest Tool and were evaluated according to the standard curve method. mRNA expression data were normalized to SigA.

2.18 Generation of TigE mutants

Mutant strand synthesis reactions were performed according to the Agilent QuikChange site-directed mutagenesis protocol. The TigE_pET28a plasmid was used as a template for site-directed mutagenesis studies. The following mutants were made TigE C339Y, TigE C339C, TigE C339A. Sequence verified plasmids were co-transformed into E. coli BL21(DE3) cells with the pPH151 plasmid containing the *suf* operon for protein production. 10mL of overnight terrific broth (TB) medium was used to inoculate 1L of TB medium containing 50mg/mL of kanamycin and 50 mg/mL chloramphenicol. The culture was incubated at 37℃ with shaking at 200 rpm until $OD_{600} \sim 1.0$ was attained, at which point the culture was cooled on ice for 30 min to reach room temperature. Then it was induced by 1mM isopropyl-β-thiogalactopyranoside (IPTG), 1.5 g/L of sodium fumarate, and 1x trace metals. Following a 12-hour induction at 21℃ with shaking at 200 rpm, the

37

cells were harvested by centrifugation at 7000 rpm for 10 min. The protein purification and reconstitution were performed anaerobically (Coy Lab anaerobic chamber). The harvested cells were resuspended in five times cell weight of lysis buffer (50mM HEPES, 300 mM NaCl, pH 8.0). CHAPS (1% w/v), lysozyme (1mg/g), 0.1 mg/g of DNase were added and stirred on ice for 20 min. Cell lysis was completed by sonication and the lysate was centrifuged at 13000 rpm for 10 min at 4°C. The supernatant was loaded onto a 5mL penta-HisTrap FF Ni-NTA column (GE Healthcare) pre-equilibrated with lysis buffer using AKTA Pure FPLC (GE Life Sciences). The column was washed with lysis buffer and the bound protein was eluted using elution buffer (50 mM HEPES, 300mM NaCl, 100 mM Imidazole, pH 8.0). It was buffer exchanged into storage buffer (50 mM HEPES, 300 mM NaCl, 10% glycerol, 10mM DTT, pH 8.0) using a HiPrep 26/10 desalting column.

2.19 Protein reconstitution of TigE mutants

10mM DTT and 12 molar equivalents of $FeCl_3$ and Na_2S were added to the buffer-exchanged protein and stirred on ice for 30 min. The reconstituted protein solution was passed through 0.45 μm filter to remove precipitated particles. It was further buffer exchanged into storage buffer (50 mM HEPES, 300 mM NaCl, 10% glycerol, 10mM DTT, pH 8.0) using a HiPrep 26/10 desalting column. The protein was concentrated using 50 kDa spin concentrator (Millipore) at 5000 g for 20 min at 4 ºC. The purity of the protein was evaluated by SDS-PAGE.

2.20 Synthesis and purification of the precursor peptide TigB-3R by SPPS

TigB-3R (WDLVYKPISGTIGSVSGTIGSVSSVSGTIGSVSG) was prepared by Fmoc-based solid phase peptide synthesis (SPPS) on a preloaded Rink Amide ProTide (LL) resin (CEM Liberty Blue automated microwave peptide synthesizer). The deprotection solution consisted of 10% piperazine (w/v) in 10:90 solution of EtOH:NMP. The activator was 0.5M DIC in DMF, the activator base solution was 0.5M Oxyma in DMF. The Fmoc-protected amino acids were at dissolved in DMF and made at 2M concentration. The synthesis was performed on a 0.5mmol scale. Upon completion of the synthesis, the resin was washed several times with methanol and chloroform and dried thoroughly under vacuum. The peptide was cleaved from the resin by cleavage cocktail (95% TFA: 2.5% TIS: 2.5% H_2O) for 30 min at 38℃. The mixture was filtered into cold diethyl ether and left at room temperature for 10 min. Once precipitated it was centrifuged for 10 min at 5000 rpm at 4℃. The pellet was redissolved in DMSO and checked by LC-MS. The dissolved pellet was purified by HPLC. Injections were performed on a semi-preparative 10×250mm C4 15-20 μm reverse column (Vydac) equilibrated in 0.1% formic acid. The peptide was eluted with gradient of acetonitrile 5-40% over 15 min. The purity of the TigB-3R was confirmed by QToF-LCMS.

2.21 Enzymatic activity assays for TigE mutants and TigB-3R

TigB-3R and TigE mutant reactions were performed in an anaerobic chamber. The TigB-3R peptide was transported into the glovebox as a lyophilized powder and resuspended in DMSO to a final concentration of ~5mM (at 280 nm, $\varepsilon = 6.9$ mM^{-1}cm^{-1}).

39

Reactions were carried out in reaction buffer (50mM HEPES, pH 8.0) with the final concentrations of 10mM DTT, 0.2mM TigB-3R, 0.2mM TigE mutant, 1mM NADPH, 10μM Fld, 10μM FldR, 10mM SAM overnight. At the end of the incubation period, the reaction tubes were taken out of the anaerobic chamber and quenched by 1% TFA. After centrifugation and filtration, the reaction was uploaded onto QTof-LCMS to check for product formation.

2.22 Expression, purification of TigD protein

The TigD_pET28a plasmid (Genscript) was co-transformed into E. coli BL21(DE3) cells with the pPH151 plasmid containing the *suf* operon for protein production. An overnight culture was grown into 10mL of terrific broth (TB) medium that was used to inoculate 1L of TB medium containing 50mg/mL of kanamycin and 50 mg/mL chloramphenicol. The culture was incubated at 37℃ with shaking at 200 rpm until OD_{600} ~ 1.0 was attained, at which point the culture was cooled on ice for 30 min to reach room temperature. Then it was induced by 1mM isopropyl-β-thiogalactopyranoside (IPTG), 1.5 g/L of sodium fumarate, and 1x trace metals. Following a 12-hour induction at 21℃ with shaking at 200 rpm, the cells were harvested by centrifugation at 7000 rpm for 10 min. The protein purification and reconstitution were performed anaerobically (Coy Lab anaerobic chamber). The harvested cells were resuspended in five times cell weight of lysis buffer (50mM HEPES, 300 mM NaCl, pH 8.0). CHAPS (1% w/v), lysozyme (1mg/g), 0.1 mg/g of DNase were added and stirred on ice for 20 min. Cell lysis was completed by sonication and the lysate was centrifuged at 13000 rpm for 10 min at 4℃. The supernatant was loaded

onto a 5mL penta-HisTrap FF Ni-NTA column (GE Healthcare) pre-equilibrated with lysis buffer using AKTA Pure FPLC (GE Life Sciences). The column was washed with lysis buffer and the bound protein was eluted using elution buffer (50 mM HEPES, 300mM NaCl, 100 mM Imidazole, pH 8.0). It was buffer exchanged into storage buffer (50 mM HEPES, 300 mM NaCl, 10% glycerol, 10mM DTT, pH 8.0) using a HiPrep 26/10 desalting column.

2.23 Protein reconstitution of TigD protein

10mM DTT and 12 molar equivalents of $FeCl_3$ and Na_2S were added to the buffer-exchanged protein and stirred on ice for 30 min. The reconstituted protein solution was passed through 0.45 μm filter to remove precipitated particles. It was further buffer exchanged into storage buffer (50 mM HEPES, 300 mM NaCl, 10% glycerol, 10mM DTT, pH 8.0) using a HiPrep 26/10 desalting column. The protein was concentrated using 50 kDa spin concentrator (Millipore) at 5000 g for 20 min at 4 °C. The purity of the protein was evaluated by SDS-PAGE.

2.24 Iron and sulfur quantification of TigD protein

Protein concentration of the as purified TigD was determined by the Bradford Assay. To start with iron quantification, 25 μM TigD was mixed with 10 μL of 3M TCA. Upon protein precipitation the mixture was centrifuged at 14,000 rpm for 10 min. The supernatant was transferred into separate centrifuge tube. 330 μL of diH_2O was added to the supernatant. 20 μL of 75mM sodium ascorbate, 20 μL of 10mM ferrozine, 20 μL of saturated sodium acetate were added to the supernatant. Each reagent was added and mixed

properly before the addition of the next. The obtained solution was transferred to a cuvette and absorbance at 562 nm was recorded ($\varepsilon = 27.9$ mM^{-1}cm^{-1}).

For sulfur quantification, 200 μL of 12μM TigD was mixed with 600 μL of 1% zinc acetate and 50 μL of 7% (w/v) sodium hydroxide and left for 15 min at room temperature. 150 μL of 0.1 % (w/v) DMPD made in 5M HCl was added to the mixture. As soon as 150 μL of 10mM FeCl$_3$ made in 1M HCl was added, the lid was closed and solution mixed. The lid was closed to avoid the oxidation of H$_2$S by air. The solution was vortexed and incubated at room temperature for 15 min. The absorbance was measured at 670 nm ($\varepsilon = 34.5$ mM^{-1}cm^{-1}).

2.25 Preparation of the Fmoc - mCPG

For FMOC - mCPG preparation 2-amino-2-(2-methylcyclopropyl) acetic acid, mixture of diastereomers (mCPG) and 9-fluorenylmethyl chloroformate, CAS RN:28920-43-6 (Fmoc-Cl) was used. A solution of Fmoc-Cl (1 equiv., 10 mmol) in dioxane (30mL) at 0℃ was added to a solution of mCPG (1 equiv., 10 mmol) in 10% aqueous Na$_2$CO$_3$ (26 mL) and dioxane (10mL). mixture was stirred for 1 hour at 0℃. The mixture was warmed to room temperature and the solution was poured into H$_2$O (100mL). The solution was extracted with Et$_2$O ($3 \times 50 mL$) and the aqueous layer was cooled in an ice bath. The reaction mixture was acidified with HCl and extracted with EtOAc ($3 \times 50 mL$). The combined extract was dried over MgSO$_4$ and filtered. The product was obtained by evaporating the solvent under reduced pressure.

2.26 Synthesis and purification of the TigB-3R mCPG by SPPS

TigB-3R mCPG is the reaction product of TigE and Tig3R where TigE installs a carbon-carbon bond between Ile-C_γ. To synthesize it we used Fmoc- mCPG instead of Ile (in red). TigB-3R mCPG (WDLVYKPISGTIGSVSGTIGSVSSVSGTIGSVSG) was prepared by Fmoc-based solid phase peptide synthesis (SPPS) on a preloaded Rink Amide ProTide (LL) resin (CEM Liberty Blue automated microwave peptide synthesizer). The deprotection solution consisted of 10% piperazine (w/v) in 10:90 solution of EtOH:NMP. The activator was 0.5M DIC in DMF, the activator base solution was 0.5M Oxyma in DMF. The Fmoc-protected amino acids were dissolved in DMF and made at 2M concentration. The synthesis was performed on a 0.5mmol scale. Upon completion of the synthesis, the resin was washed several times with methanol and chloroform and dried thoroughly under vacuum. The peptide was cleaved from the resin by cleavage cocktail (95% TFA: 2.5% TIS: 2.5% H_2O) for 30 min at 38℃. The mixture was filtered into cold diethyl ether and left at room temperature for 10 min. Once precipitated it was centrifuged for 10 min at 5000 rpm at 4℃. The pellet was redissolved in DMSO and checked by LC-MS. The dissolved pellet was purified by HPLC. Injections were performed on a semi-preparative 10×250mm C4 15-20 μm reverse column (Vydac) equilibrated in 0.1% formic acid. The peptide was eluted with gradient 5-40% over 15 min. The purity of the TigB-3R was confirmed by QToF-LCMS.

2.27 Activity assays of TigD

TigB-3R, TigB-3R mCPG and TigD reactions were performed in an anaerobic chamber. The TigB-3R and TigB-3R mCPG were transported into the glovebox as a lyophilized powder and resuspended in DMSO to a final concentration of ~5mM (at 280 nm, $\varepsilon = 1.49$ mM^{-1}cm^{-1}). Reactions were carried out in reaction buffer (50mM HEPES, pH 8.0) using different reducing systems/ reductants. First the reaction was carried out with the final concentrations of 10mM DTT, 0.2mM TigB-3R/ TigB-3R mCPG, 0.2mM TigD, 1mM NADPH, 10μM Fld, 10μM FldR, 10mM SAM overnight. Second, 10mM DTT, 0.2mM TigB-3R/ TigB-3R mCPG, 0.2mM TigD, 10mM SAM, 4mM DTH overnight. And lastly, 10mM DTT, 0.2mM TigB-3R/ TigB-3R mCPG, 0.2mM TigD, 10mM SAM, 10μM Frd, 1mM NADPH overnight. At the end of the incubation period, the reaction tubes were taken out of the anaerobic chamber and quenched by 1% TFA. After centrifugation and filtration, the reaction was uploaded onto Qtof-LCMS to check for product formation.

Chapter Three: Results

3.1 Identifying and sequencing the mycofactocin operator

To provide evidence that MftR is a regulator of MFT biosynthesis, we ran a series of electrophoretic mobility shift assays (EMSA) to demonstrate that MftR binds to the MFT promoter region. To begin with, recombinant his-tagged Msmeg MftR was purified from *E. coli* (Figure 3.1) and the 565 bp promoter region (P_{mft}) between -470 to +95 relative to MftA was PCR amplified.
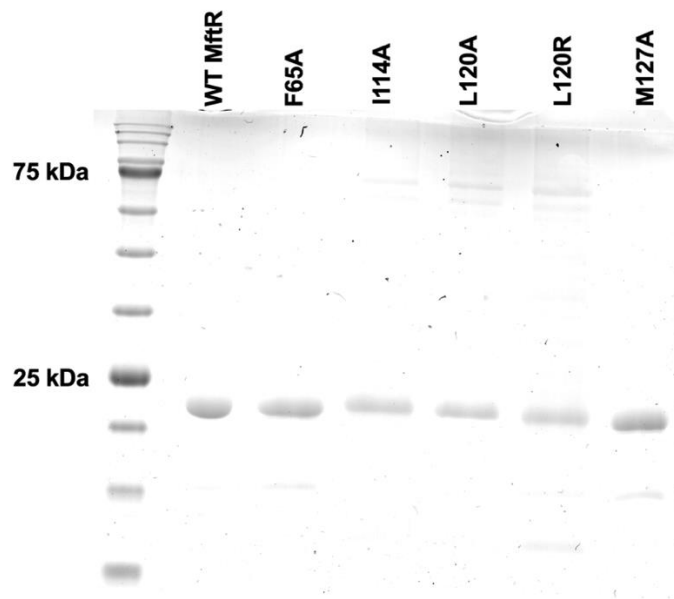

Figure 3.1 Sodium dodecyl sulfate-polyacrylamide gel electrophoresis analysis of WT MftR and mutants. The purified protein and variants (4 mg) were analyzed on 15% polyacrylamide gel, stained with Coomassie Brilliant Blue G-250. Lane 1 is the blue stain protein marker.

Next, EMSAs were carried out in triplicate with a fixed concentration of the PCR amplified P_{mft} and varying concentrations of MftR. As shown in Figure 3.2A, the addition of MftR to unlabeled P_{mft} resulted in a single impeded band in a concentration-dependent manner with an estimated $K_d \sim 0.6$ mM. This result indicates that at least a single binding site of MftR with at least one binding affinity is present in the P_{mft} regulatory region.

To determine the exact location of the MftR-binding site in the *mft* regulatory region, Dnase I protection assays were performed using a P_{mft} DNA probe labeled with 6-FAM, in the presence and absence of MftR. As shown in Figure 3.2B, MftR protected a single region extending from -79 to -53 from Dnase I digestion. The shift in Dnase I hypersensitivity by three nucleotides when MftR is present suggests the establishment of new contacts being made to and/or a modification of the DNA structure. To validate this finding, the MftR protected sequence was synthesized with a 6-FAM label and used in a subsequent EMSA. As shown in Figure 3.2C, increasing concentrations of MftR resulted in a single concentration-dependent impeded band, consistent with the original EMSA with P_{mft}. Consistent with EMSA assays, increasing concentrations of MftR resulted in a concentration-dependent change in polarization, which upon fitting three independent experiments to a single site binding model, led to an observed dissociation constant (K_d) of 1.3 ± 0.6 µM (Figure 3.2D). As a result, we propose that the MFT operator (O_{mft}) sequence includes at least one MftR binding motif within the sequence 5'-TCCATTCTGGCACTCGATGCCATATAT (Figure 3.2E).
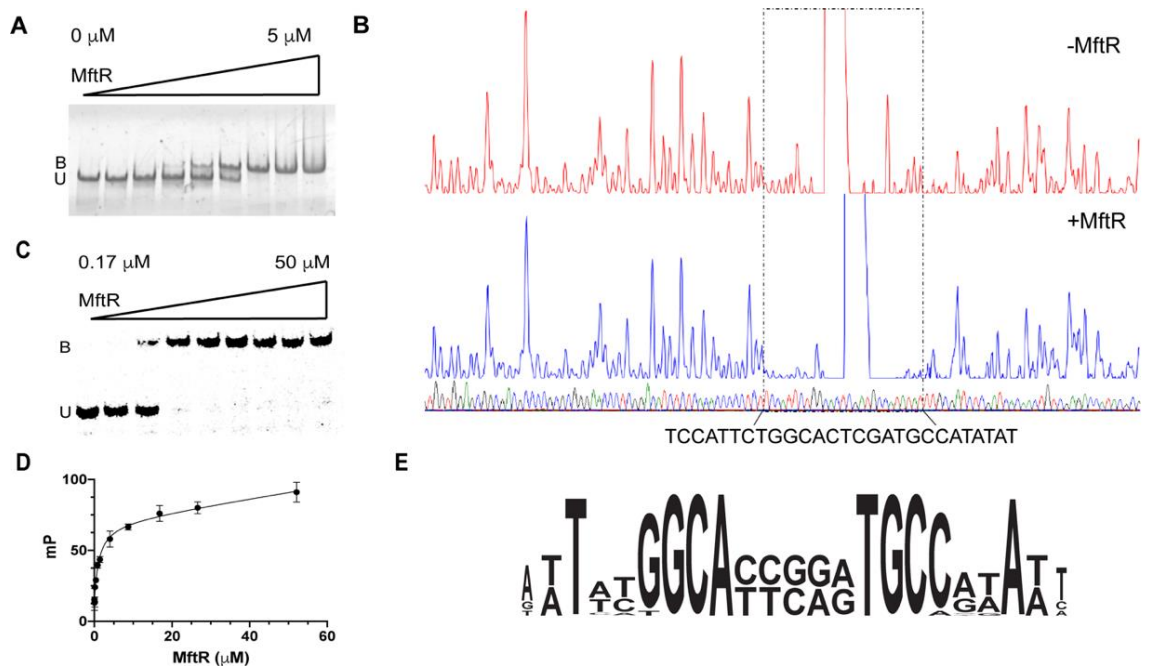
Figure 3.2 Identification of the mycofactocin operator. A) EMSA of the P_{mft}-MftR complex. The P_{mft} region (0.5 mM) was mixed with increasing concentrations of MftR. The U and B represent unbound and bound fractions, respectively. The assays were performed in triplicate, producing similar results. B) Dnase I footprinting assay to determine the binding site of P_{mft} to MftR. Top fluorogram shows control reaction of P_{mft} 350 ng with no protein added. Upon the addition of MftR (2 mg) the distinct binding area was determined and is shown in the bottom fluorogram. The 27 base pairs sequence (O_{mft}) on the P_{mft}-MftA was confirmed to be the binding region responsible for the interaction with MftR. C) An EMSA validating that MftR binds to O_{mft}. Fluorescein-labeled O_{mft} (0.5 mM) was mixed with increasing concentrations of MftR. The U and B represent unbound and bound fractions respectively. D) A representative fluorescence polarization binding assay showing the change in polarization of FAM-O_{mft} as a function of MftR concentration. All assays were performed in triplicate and the average (filled circles), and standard deviations (error bars) are shown with the non-linear fit (line). E) A WebLogo representation of the Omft region showing the palindromic sequence found in 83 species of Mycobacterium and Mycolicibacterium.

47

Controls with a "cold" competitive ligand and with a non-specific DNA sequence (Figure 3.3) suggests that the interaction between MftR and the identified region is specific. In addition, fluorescence anisotropy experiments were carried out to estimate the binding affinity between MftR and the 6-FAM-labeled 27 bp sequence.
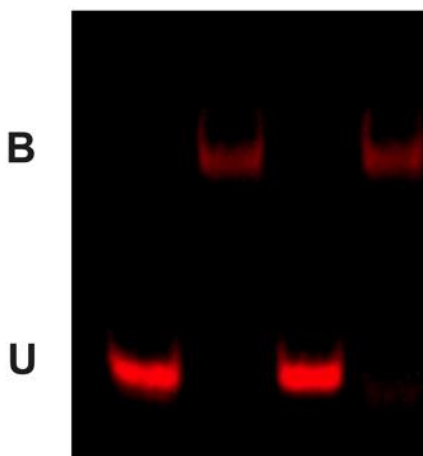

Figure 3.3 O_{mft} control with "cold" competitive ligand and a non-specific DNA sequence. Unlabeled "cold" probe (4μM) was mixed with WT MftR (2μM) left at room temperature to react for 30 min. FAM-labeled O_{mft} (20 nM) was then added to the reaction and left for additional 30 minutes (lane 3). Unlabeled non-specific DNA sequence (4μM) was mixed WT MftR (2μM) left at room temperature to react for 30 min. FAM-labeled O_{mft} (20 nM) was added and left to react for additional 30 minutes (lane 4). Protein-DNA complex was separated by electrophoresis on 5% polyacrylamide gel and imaged using FAM excitation and emission wavelengths. The U and B represent unbound and bound fractions respectively.

48

Next, we employed quantitative real-time PCR (qRT-PCR) analysis to demonstrate that MftR regulates the *mft* BGC *in vivo*. We measured the transcript levels of *mftA-F* and *mftR* in wild-type Msmeg and compared them to the transcript levels in Msmeg harboring a mycobacterial expression vector consisting of *mftR* under the control of the constitutive expression promoter P_{smyc} (pMftR+). Over production of *mftR* in the expression strain, as compared to wild-type Msmeg, was confirmed by qRT-PCR analysis which revealed that the *mftR* transcript abundance was increased by approximately 5-fold (Figure 3.4). Conversely, overproduction of *mftR* led to reduced transcript levels in all *mft* biosynthetic genes. Notably, the transcript levels of *mftA* and *mftC* were reduced by ~15-fold and ~20-fold, respectively. However, the most remarkable change in transcript levels was that of *mftD which* was reduced nearly 80-fold. The transcript levels for *mftB*, *mftE*, and *mftF* were also decreased however, to a lesser extent (<10-fold). Taken together with the EMSA's, the DNA footprinting, and the fluorescence anisotropy experiments, MftR is a regulator of MFT biosynthesis in Msmeg.
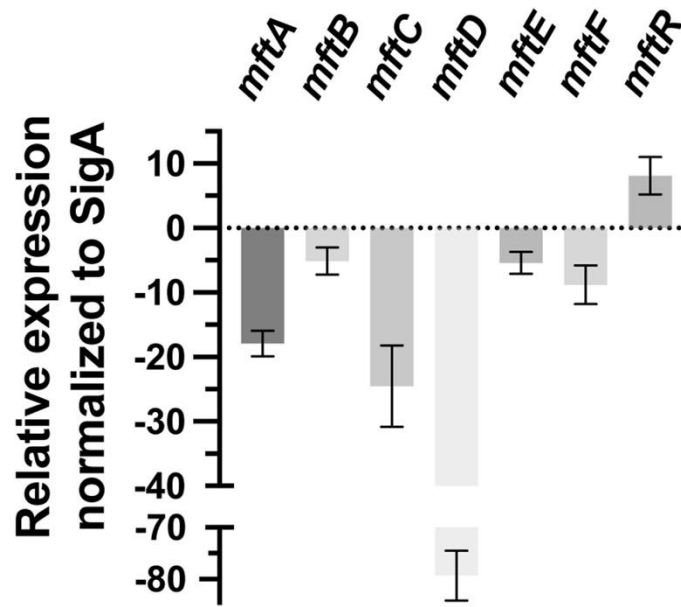
Figure 3.4 Overexpression of MftR downregulates the mycofactocin (MFT) biosynthetic gene cluster (BGC). A qRT-PCR based quantification of the MFT BGC in M. smegmatis grown in 7H9 ADC. The relative expression levels of the differentially expressed genes were compared between the bacterial strains wild-type M. smegmatis and M. smegmatis transformed with pMftR+.

3.2 Expanding MftR role in mycobacteria

We examined if the position and sequence of O_{mft} in the *mft* promoter region is similar in MTB since the organism encodes for the *mft* operon (*rv0691a-rv0696*) and a MftR homologue (*rv0691c,* 69% identity). To answer this question, we carried out a BLAST analysis of the identified O_{mft} sequence and the ~500 bp *mft* promoter from MTB. Accordingly, we found a single well-aligned sequence with 85% conservation. Similar, to Msmeg, the putative O_{mft} in MTB extends from -78 to -52 relative to the *mftA* start codon, with the assumption that the operator sequence is the same length. We expanded our search to the *Mycobacterium* and *Mycolicibacterium* genera using a customized BLAST analysis

(Expect = 1,000; Match/Mismatch = 1, -1; Gap Costs = 0,2). Under these conditions, 83 sequences were identified with sequence identities >84%. As shown in the Weblogo depiction of the multiple sequence alignment of all sequences (Figure 3.2E),[124] we found that putative O_{mft} regions are highly conserved. In addition, our analysis identified a palindromic region consisting of the residues T-N_2-GGCA-N_5-TGCC-N_2-A. Despite the apparent conservation of the palindrome, single nucleotide replacements within the sequence did not impede the ability of MftR to bind O_{mft} in fluorescence polarization or EMSA assays (Figure 3.5).
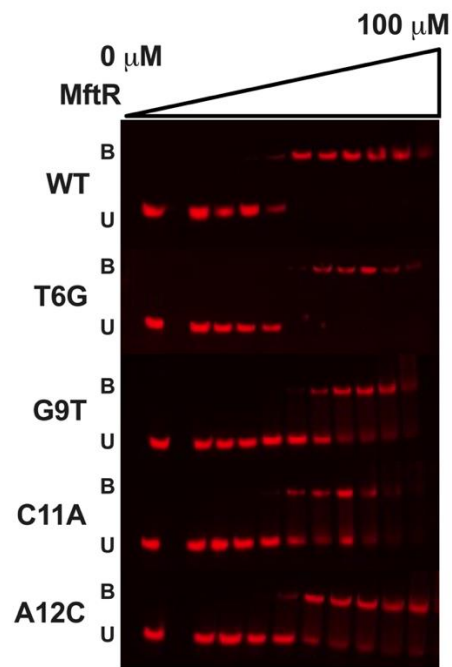
Figure 3.5 Single nucleotide replacements in O_{mft} didn't interrupt the MftR binding.

3.3 Long-chain acyl-CoA's are effectors of MftR

Next, we assessed which metabolites activate MftR and thus could induce MFT production. To do so, we carried out competitive EMSAs where FAM-labeled O_{mft} and MftR were incubated in the presence of potential effectors. We initially chose our effectors based on cholesterol catabolism, a process that putatively includes MFT biosynthetic genes.[125] Despite the loose association of MFT to cholesterol catabolism, we did not observe DNA release by MftR in the presence of cholesterol (Figure 3.6A, lane 3), propionyl-CoA (Figure 3.6A, lane 6), succinyl-CoA (not shown), and acetoacetyl-CoA (not shown). Knowing that TFRs have a propensity to be activated by fatty acyl-CoA's,[126] we expanded our effector screening to include short, medium, and long-chain acyl-CoA's. Subsequently, we observed that the addition of myristoyl- and oleoyl-CoA's disrupted the MftR-O_{mft} complex (Figure 3.6A, lanes 8 and 9) and resulted in both bound and unbound O_{mft}. Conversely, the addition of fatty acid carboxylates did not result in the same disruption of the MftR-O_{mft} complex (Figure 3.6A, lane 10), suggesting that CoA is a requisite for acyl-CoA binding. However, CoA alone did not disrupt the MftR-O_{mft} complex either (Figure 3.6A, lane 4), suggesting that protein contacts with the ligand rely on both the fatty acid and the CoA.

Isothermal titration calorimetry (ITC) experiments were performed to validate the competitive EMSA findings and to determine the specificity and affinity of MftR towards acyl-CoA's. A typical thermogram was obtained when oleoyl-CoA was titrated into MftR (Figure 3.6B). The K_d value (1.4 ± 0.1 μM, Figure 3.6C), and the number of binding sites

(~0.6 sites per monomer MftR) were obtained from the nonlinear one-site model to the normalized fitting curve. The K_d value is comparable to known mycobacterial TFRs that are activated by oleoyl-CoA.[127,128] To establish the specific acyl-CoA(s) that activate MftR, we measured the K_d's for myristoyl-, palmitoyl-, and steroyl-CoA's and found the values to be within the 2-4 μM range (Figure 3.6C). Of note, we observed a ~10-fold decrease in binding affinity with lauroyl-CoA as compared to oleoyl-CoA. This drop in affinity is consistent with EMSAs that indicated medium and short chain acyl-CoA's do not disrupt the MftR-O_{mft} complex. Taken together with the EMSAs, our ITC data suggests that MftR, and thus likely MFT biosynthesis, is activated by long chain acyl-CoA's.
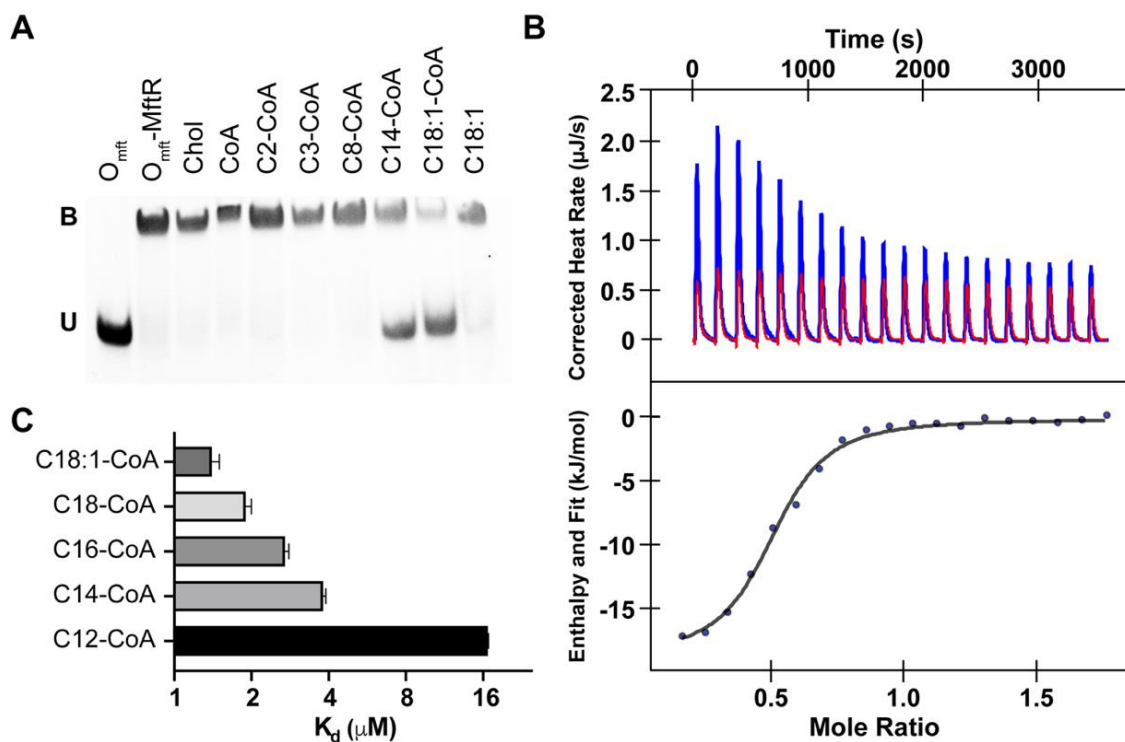
Figure 3.6 Effectors of MftR. A) An electrophoretic mobility shift assay used to screen effectors of MftR. The 6-FAM labeled Omft region (0.5 mM) was incubated with the MftR (5 μM) in the absence or the presence of effectors (100 μM). Protein-DNA complexes were separated by electrophoresis on 5% polyacrylamide gel. The assays were performed in triplicate, producing similar results. B) A representative isothermal titration calorimetry thermogram for the binding of the oleoyl-CoA to MftR regulator (blue) and the control of oleoyl-CoA into buffer (red). Each peak corresponds to the injection of 2.22 mL of 250 μM oleoyl-CoA into the cell containing 16 μM MftR or buffer. The integrated thermogram (blue dots) was fitted to a single site binding model (black line) to determine the K_d. The experiment was done in triplicate producing the similar results. C) A bar graph depicting the ITC measured K_d's for various acyl-CoA's. Bars represent the mean of three independent experiments and the error bars represent the standard deviation of the experiments.

3.4 Structural contributions to ligand binding

Next, we examined which amino acid residues contribute to the interaction between MftR and oleoyl-CoA. The crystal structure of Msmeg MftR is currently unavailable however, an unpublished structure of *Rhodococcus jostii* RHA1 MftR has been deposited in the Protein Data Bank (PDB ID: 2RAE, 54% identical, Figure 3.7).

```
2RAE    LRSRRKPSSRIGRRPSTTQDRISTVGIELFTEQGFDATSVDEVAEASGIARRTLFRYFPS   60
MftR    ----MSEGSRAGRRRSTTQDHIAGVAIDLFAARGFDAVSVDDVAAAAGISRRTLFRYYAS   56
            . .** *** *****:*: *.*:**: :****.***:** *:**:*******: *

2RAE    KNAIPWGDFDAHLAEMRAQLAAQPDDIPIVDGLTAALLQFNAFPASEEINHRKRMGLILR   120
MftR    KSAIPWGDFDSHLQHLQNLLRTLSSEVSLGDALREALLTFNTYADHEMAEHRQRMRVILE   116
        *.********:** .::   * :  .:: : *.*   *** **::     *  :**:** :**.

2RAE    VPALQAYSVVMYEGWRNVIAEYVASRLGTSPTDHVPRTVGYLLLGVAMSAYEQWLDDDSL   180
MftR    TEELQAYSMTMYAGWRTVIAEYVADRLGMSTTDLMPQTVAWLMLGVALSAYENWLADESV   176
        .  *****:.** ***.*******.*** * ** :*:**.:*:****:****:** *:*:

2RAE    ELNELLASGMQSLYDGLSSLGEPDTRT    207
MftR    SLRGAIAAAFEVARPGLDALQPAH---    200
        .*.  :*:.::    **.:*    .
```

Figure 3.7 A sequence alignment between 2RAE and Msmeg MftR.

Using this structure, we employed Swiss Dock to model oleoyl-CoA bound to MftR (Figure 3.8).[129,130] From the docked structure, we identified eight conserved residues on MftR that were expected to create the acyl binding pocket or bind CoA through electrostatic interactions.
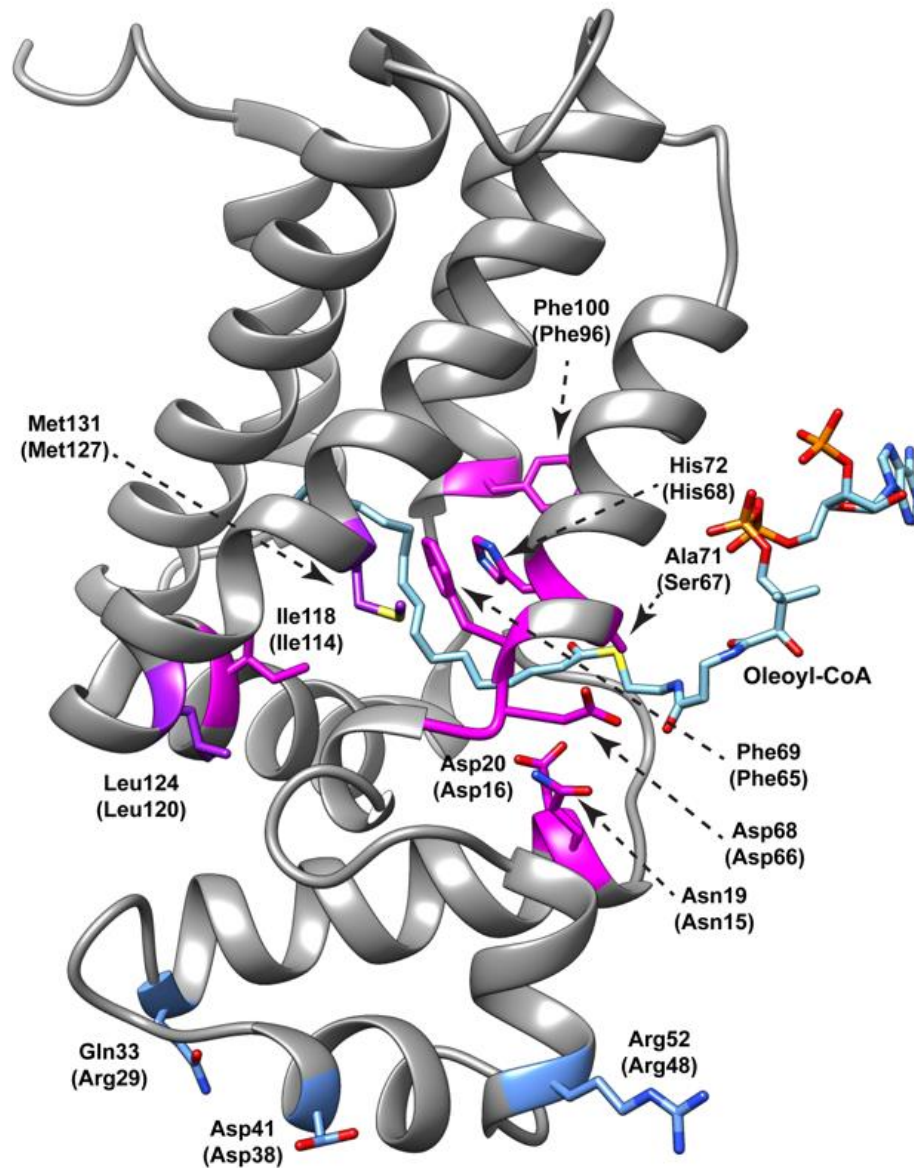
Figure 3.8 A docked structure of oleoyl-CoA bound to MftR. The Rhodococcus jostii RHA1 MftR (2RAE, grey) was docked with oleoyl-CoA (blue) using Swiss Dock. Shown is the lowest energy model (-10 kJ/mol) calculated. Residues associated with the oleoyl-CoA binding pocket (magenta), DNA binding motif (blue), and at the interface (purple) are annotated for both 2RAE and the corresponding M. smegmatis mc^{2}155 MftR (parenthesis).

56

Following single amino acid replacements of the residues, we measured the K_d values of the mutant proteins to oleoyl-CoA using ITC (Figure 3.8). For the putative acyl binding pocket residues Phe65, Phe96, and Ile114, mutations to alanine led to no or modest change to the dissociation constant for oleoyl-CoA (Table 3.1).

Table 3.1 Dissociation constants for MftR

MftR variant	[1]K_d (µM)
WT	1.4 ± 0.1
Q15A	2.0 ± 0.1
D16R	[2]n.d
D16W	4.8 ± 0.2
F65A	1.9 ± 0.1
D66A	1.6 ± 0.1
S67A	5.4 ± 0.2
S67W	8.3 ± 0.2
H68A	[2]n.d
F96A	1.4 ± 0.1
I114A	2.4 ± 0.1
L120A	5.2 ± 0.2
L120R	6.2 ± 0.2
M127A	4.8 ± 0.1

[1] The K_d values are reported as the average of at least three independent experiments and the standard deviation. [2] Not detected.

This is consistent with other TFR proteins where single residue changes to the acyl binding pocket resulted in little to no change in their dissociation constant to acyl-CoA's.[125] However, we cannot rule out the possibility that Phe65, Phe96, and Ile114 do not participate in acyl-CoA binding. Conversely, when His68 was mutated to alanine, binding of oleoyl-CoA by the protein was undetectable. To validate this observation, competitive EMSAs were carried out using the H68A mutant. Here, it was observed that the H68A mutant remained bound to O_{mft} even in the presence of 100 µM oleoyl-CoA (Figure 3.9A). Thus, it is highly likely that His68 is an important residue for oleoyl-CoA binding. To evaluate residues that were expected to interact with CoA, Gln15, Asp16, Ser67, and Asp66 were targeted. Of these, mutants of Asp16 and Ser67 had the greatest effect (Table 3.1). For instance, the S67A mutation led to a 5-fold increase in the K_d value for oleoyl-CoA and the addition of bulk, by the mutation S67W, led to an 8-fold increase in the K_d value. Likewise, the D16W mutant, which increased bulk and removed the negative charge, led to a 5-fold increase in the K_d for oleoyl-CoA. Moreover, when the charge was changed from negative to positive by the mutant D16R, binding of oleoyl-CoA was no longer detectable. This latter observation was validated using competitive EMSA's (Figure 3.9B). Even in the presence of 100 µM oleoyl-CoA, the D16R mutant remained bound to O_{mft}. Taken together, it is likely that Ser67 and Asp16 participate in oleoyl-CoA binding, likely through electrostatic interactions.
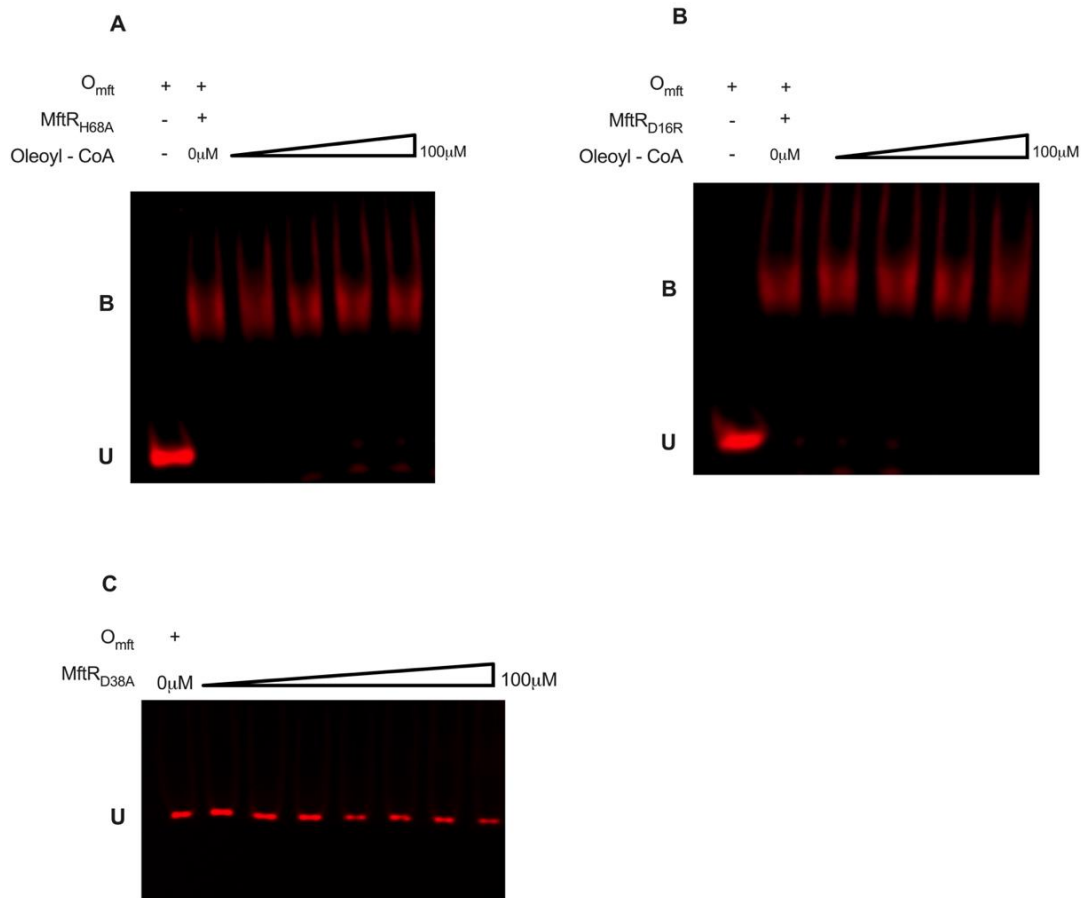
Figure 3.9 EMSA validating the importance of MftRH68A, MftRD16R, and MftRD38A mutants. A) FAM-labeled O_{mft} (100nM) was mixed with MftRH68A ($2\mu M$) and left to react for 10 min at room temperature. Increased concentrations of oleoyl-CoA were added into the reactions (1, 10, 50, 100 μM). The U and B represent unbound and bound fractions respectively. Protein-DNA complex was separated by electrophoresis on 5% polyacrylamide gel and imaged using FAM excitation and emission wavelengths. B) FAM-labeled O_{mft} (100nM) was mixed with MftRD16R ($2\mu M$) and left to react for 10 min at room temperature. Increased concentrations of oleoyl-CoA were added into the reactions (1, 10, 50, 100 μM). C) FAM-labeled O_{mft} (100nM) was mixed with increased concentrations of MftRD38A (up to 100μM) and left to react for 20 min at room temperature. Protein-DNA complex was separated by electrophoresis on 5% polyacrylamide gel and imaged using FAM excitation and emission wavelengths.

59

Next, to determine which residues are important for the MftR-DNA interaction we targeted three residues (Figure 3.8, Arg29, Asp38, and Arg48) on the helix-turn-helix domain for site directed mutagenesis. Similar residues have been shown to be important for DNA-protein contacts in other TFRs.[131–133] The dissociation constants of the MftR mutants and O_{mft} were measured by fluorescence polarization (Table 3.2). As expected, the alanine mutants for Arg29 and Arg48 had a significant impact on the ability of MftR to bind O_{mft}, with a 10- and 20-fold increase in the K_d values, respectively (Table 3.2). It is possible that Arg29 and Arg48 interact with the phosphate backbone on O_{mft}, thus making them important for DNA binding. Interestingly, the D38A mutant abolished the ability of MftR to binding O_{mft} (Table 3.2).

Table 3.2 Dissociation constants for MftR variants and Omft measured by fluorescence polarization

MftR variant	^1K_d (μM)
WT	1.3 ± 0.6
R29A	20.5 ± 2.0
D38A	^{2}n.d
R48A	13.0 ± 1.1
L120R	28 ± 6.0
M127A	13 ± 6.0

[1] The K_d values are reported as the average of at least three independent experiments and the standard deviation. [2] Not detected.

To corroborate this observation, an EMSA was carried out with D38A mutant and O_{mft} (Figure 3.9C). The addition of up to 100 µM of the D38A mutant to O_{mft} resulted in a single unbound species, confirming that Asp38 is important for binding to O_{mft}. The importance of Asp38 is likely explained by Asp38 forming of a salt bridge with Arg47 upon binding DNA. A similar salt bridge has been shown to be important for the master regulator Msmeg_6564.[134] In Msmeg_6564, the hydrogen bond between Glu37-Lys47 (~2.5 Å, PDBID: 4JL3) stabilizes the HTH domain in the major groove of DNA and was shown to be important for DNA interaction.[134] Nevertheless, until a DNA bound structure of MftR becomes available, the role of Asp38 in O_{mft} binding remains speculative.

Lastly, we examined the effects of mutations at the interface of the DNA binding and effector binding domains. Here residues Leu120 and Met127 were mutated to arginine and alanine, respectively, and the dissociation constants for oleoyl-CoA and O_{mft} were measured as described above. While we did not observe significant impact on the dissociation constant for oleoyl-CoA (Table 3.1), we did find that the L120R and M127A mutants impaired the ability of MftR to bind O_{mft}, with an observed increase in K_d values by 10- and 25-fold, respectively (Table 3.2). This suggests that disruption of the packing between the DNA binding domain and the regulatory domain impacts binding of DNA substantially more than oleoyl-CoA.

3.5 MFT biosynthesis is induced by oleoyl-CoA *in vivo*

To validate that oleoyl-CoA is an effector of MFT *in vivo*, we turned to a well-established far-red reporting system designed to detect gene expression in mycobacteria.[135] We repurposed the pCherry3 vector by replacing the existing constitutive promoter, P_{smyc}, with *mftR*-P_{mft}-*mftA* and in frame with mCherry (Figure 3.10) yielding P_{mft}-pCherry. As a result, the plasmid encodes *mftR* under its native promoter, the *mft* promoter, and a *mftA-mCherry* gene fusion. The addition of *mftR* is expected to suppress background fluorescence that may arise by having increased copies of P_{mft} and insufficient MftR to repress expression. Transcription and translation of *mftA* is expected to result in a fluorescent MftA-mCherry reporter, the intensity of which can be measured as a function of time. Together, the P_{mft}-pCherry reporter system was expected to provide information about the timing and relative abundance of MftA production.
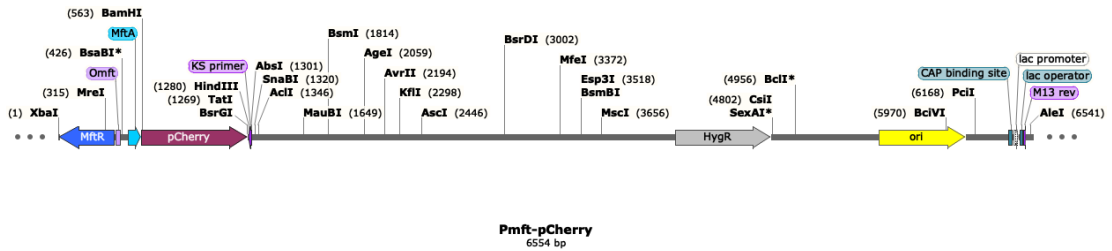
Figure 3.10 A schematic representation of the engineered plasmid encoding for the promoter region of mft and mftA-mCherry gene fusion. The pCherry 3 vector was remodeled by replacing Psmyc with Pmft-mftA and in frame with mCherry.

To directly establish that oleoyl-CoA induces MFT biosynthesis *in vivo*, a colony dilution experiment was carried out using Msmeg transformed with P_{mft}-pCherry grown on

Middlebrook 7H9 (7H9) supplemented with glucose or oleic acid. We used oleic acid in the growth media since it is well known that bacterial fatty acid shuttle systems convert free fatty acids to the corresponding acyl-CoAs and since Msmeg encodes for at least two known fatty acid transporters.[136] As shown in Figure 3.11, the fluorescence intensity of Msmeg colonies containing the reporter system is starkly increased in the presence of oleic acid as compared to glucose alone.
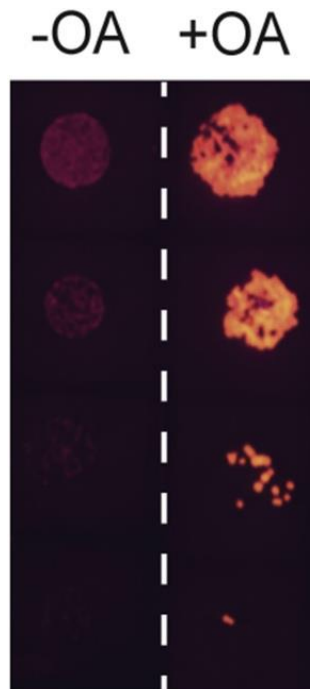

Figure 3.11 Induction of MFT biosynthesis by oleoyl- CoA *in vivo*. A spot test growth analysis of *M. smegmatis* mc^{2}155 transformed with the plasmid P$_{mft}$-pCherry encoding for the mftR, the regulatory region of the *mft* BGC, and a *mftA-mCherry* gene fusion. Approximately 500, 100, 25, 5 cells were grown on 7H9 supplemented with 0.1% w/v glucose or 0.05% w/v each glucose and oleic acid (OA). Fluorescence images were obtained simultaneously from two different petri dishes using 585 nm excitation and 636 nm emission wavelengths.

To assess when this occurs in real time, growth curve assays with Msmeg/P_{mft}-pCherry were carried out. When cell cultures were grown to nearly the same optical density (Figure 3.12A), the fluorescence intensity of Msmeg/P_{mft}-pCherry supplemented with 0.1% oleic acid only is nearly three times in magnitude as compared to the control (Figure 3.12B). We observed the same effect when oleic acid was added to the media after 12 h of incubation.
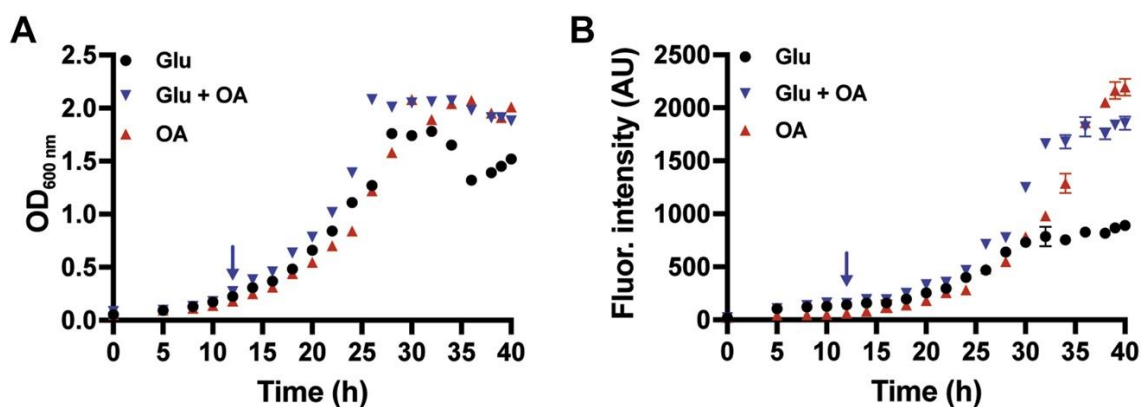
Figure 3.12 Growth curve assays with Msmeg/P_{mft} - pCherry. A) OD600nm measurements for the 7H9 media supplemented with 0.1% tyloxapol, 50μg/mL hygromycin, and glucose and/or oleic acid (OA) as carbons sources. Concentrations of carbon sources were glucose only (0.1% w/v), OA only (0.1%wv), and glucose/OA (0.05% w/v each) where OA was added after 12 hours (blue arrow) B) Corresponding fluorescence intensities (ex. 585 nm/em. 612 nm) of Msmeg/ Pmft – pCherry cultures described in A.

Next, we determined if MFT induction is dependent on the concentration of oleic acid. We carried out a titration series of growth curves where 7H9 media was supplemented with 0.05% w/v glucose and 0.01, 0.02, 0.05, or 0.1% w/v oleic acid. As shown in Figure 3.13A

and Figure 3.13B, the addition of 0.05% w/v oleic acid produced the highest fluorescence intensity within a 40-h timeframe.
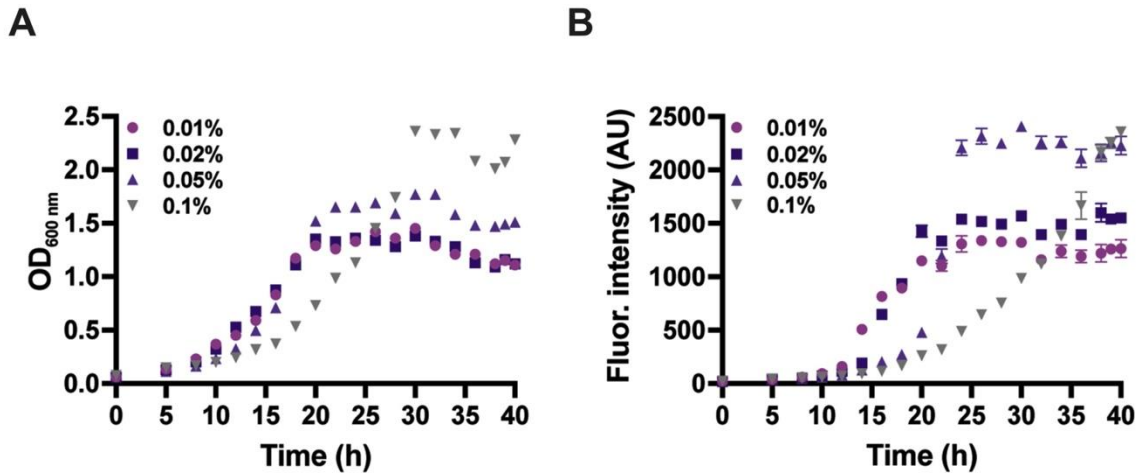
Figure 3.13 Concentration – dependent growth curve assays with Msmeg/Pmft -pCherry. A) OD600 measurements for dose-dependent addition of OA to the 7H9 media supplemented with 0.05 % w/v glucose and 0.01, 0.02, 0.05, 0.1% w/v of OA. D) Corresponding fluorescence intensities of Msmeg/ Pmft – pCherry cultures described in C.

However, in general, increasing concentrations of oleic acid had little effect on the overall production of the MftA-mCherry fusion. The lack of immediate production of MftA-mCherry after the addition of oleic acid and the concentration independence is perplexing. Since we did not observe a diauxic growth curve, we do not suspect that Msmeg is displaying a prototypical substrate bias. Rather, we expect a subtler and unknown mechanism is at play.

3.6 Attempts to overexpress glycosyltransferase MftF from *E. coli*

The analysis began with the attempts to overexpress and isolate an integral membrane protein MftF from species *Mycobacterium smegmatis* mc^2 155 (MsMftF)*, Mycobacterium ulcerans Agy 99* (MuMftF), and *Thermomonospora curvata* (TcMftF) using *E. coli* expression system summarized in Table 3.3.

Table 3.3 Strategies used for overexpression of MsMftF, MuMftF, and TcMftF genes in *E. coli*

Approach	Methodology	Reasoning	Examples
Modifications to culture conditions	Vary the concentration of inducer	Reduces the protein expression rate	Range from 0.1mM to 1mM of inducer
	Lowering the temperature of induced cultures	Reduces the protein expression rate, promotes proper folding and solubility	15-20 ℃ during the induction
Screening and considerations for expression hosts	Co-expressing chaperones	Aid in protein folding and solubility	PG-KJE8 (dnaK-dnaJ-grpE-groES-groEL); ArticExpress (DE3) use Cpn10 and Cpn60 chaperonins

Use of strains for membrane and toxic proteins	Reduces transcription rate of the derivatives strains improves membrane protein overexpression yields	*E. coli* strains BL21(DE3) pLysS BL21(DE3) Gold C43(DE3), Lemo21(DE3),
Use of *E. coli* optimized genes and use of strains for rare codons	Aims to improve protein production by optimizing translation elongation	Rosetta (DE3) for and *E. coli* optimized MsMftF/pET21a (+)
Use of soluble fusion tags at N- or C- terminus	Stabilizes the expression and improves protein solubility	Pet28HMT (MBP,His-tagged,TEV), Pet28GTH (GST, His-tagged, TEV), pGEX6p-1 (GST) pet24mCTH (mCherry, His-tagged)
Use of pMy vectors for protein expression in mycobacteria	Addresses the mismatch in codon usage and the absence of cofactors and PTMs in *E. coli*	pMyNT, pMyC, pMyBadC vectors

We used *E. coli* as a host because its relatively simple, inexpensive, and fast-growing. Moreover, in the study performed by Korepanova A. et al.,[137] they cloned and overexpressed around seventy integral membrane proteins from *Mycobacterium tuberculosis* genome using T7 expression system. The molecular weight of the expressed proteins in the study ranged from 8.1 to 71.3 kDa with transmembrane helices from 1 to 14.[137] MftF is an integral membrane protein with a molecular weight of around 50.7 kDa and proposed to have 1-3 transmembrane helices. Table 3.3 has several expression conditions that we tested for MftF protein in attempt to maximize the efficient production of soluble protein in bacteria. We evaluated the efficiency through small-scale expression screening by SDS-PAGE stained with Coomassie blue or analyzed by Western blotting with anti-His antibody. However, overexpression of MftR was not detected.

While the strategies used here worked for many targets, including membrane proteins, every protein has unique biophysical features that often requires optimization steps for successful expression.

Chapter Four: Contributions to the class of cyclopropylglycyl synthases

4.1 Input towards the structural elucidation of TvgB product

To structurally characterize the TvgB product we employed 1D/2D NMR techniques. When we compare the ^{13}C NMR spectra of the product of TvgB reaction to the isotopically labeled peptide ^{13}C-TvgA-4R we observed three shifts on Val carbons (Figure 4.1A). Since not all the four repeats are being modified by TvgB, the product spectrum contains peaks for unmodified Val. We assigned the new peaks: at δ_C 55.57 ppm to the C_α of ^{13}C-Val from the 57.29 ppm in the unmodified peptide; at δ_C 12.80 ppm to the C_β of ^{13}C-Val from 29.77 ppm in the substrate; at δ_C 2.26 ppm to the C_γ of ^{13}C-Val from 17.44/18.62 ppm.

Next, we performed ^{13}C HSQC on ^{13}C-TvgA-4R (Figure 4.1B, C). The difference in the spectra of starting material and the product was consistent with the peaks observed in the 1D ^{13}C NMR spectra. We detected new coupled features at δ_H 0.25 ppm, δ_C 2.26 ppm; and δ_H 0.39 ppm, δ_C 2.26 ppm; which we assigned to saturated methylene groups that are bonded ($-CH_2-CH_2-$). Another set of new features at δ_H 1.04 ppm, δ_C 12.80 ppm was assigned to C_β; at δ_H 3.74 ppm, δ_C 55.57 ppm was assigned to C_α.

From the data described above, we believed that TvgB catalyzed the formation of cyclopropylglycine (CPG) from Val. To support this, we ran time-constant HSQC of modified ^{13}C-TvgA-4R peptide (Figure 4.1D). The spectra showed C_α and C_β have odd

number of neighbors appearing in red (bonding to the carbonyl carbon is excluded due to C-O decoupling) and C_γ carbons are directly bonded to two other nuclei showing in blue (Figure 4.1D, E).

Taken together, our analysis of 1D/2D NMR experiments on the product indicates that TvgB installs a C-C bond in each of the TVGG repeats between the geminal methyl groups on Val side chains.
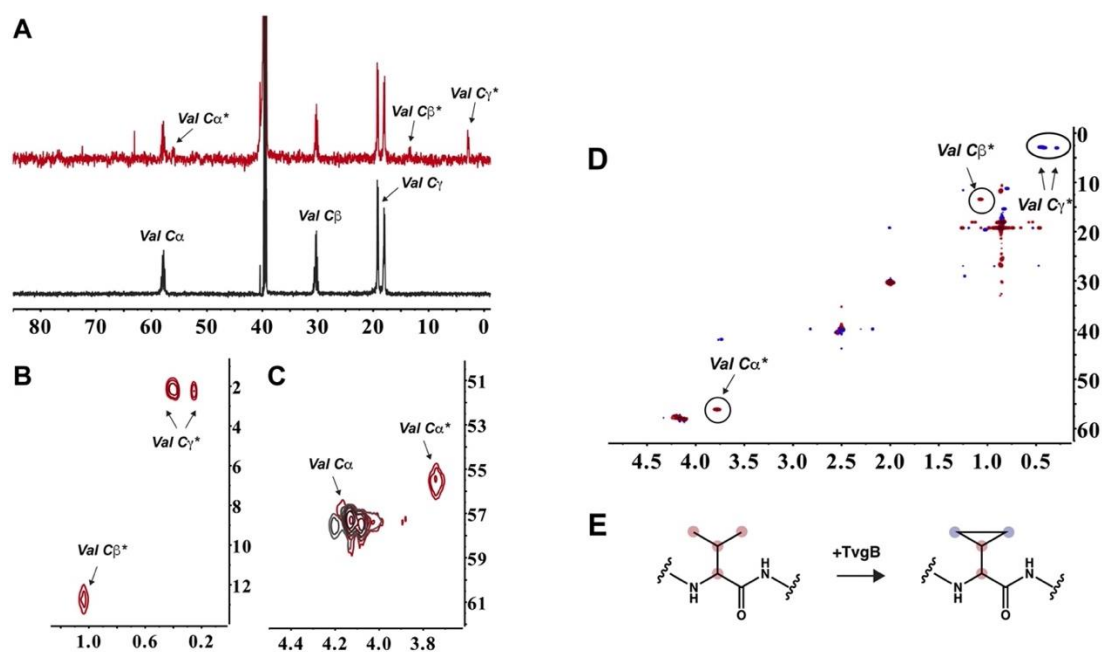
Figure 4.1 Structural elucidation of the modified TvgA-4R. A) Stacked ^{13}C NMR spectra of the TvgA-4R peptide (black) and modified TvgA-4R (red). B) Stacked ^{13}C HSQC spectra of the TvgA-4R peptide (black) and modified TvgA-4R (red) featuring new peaks. C) Stacked ^{13}C HSQC spectra of the TvgA-4R peptide (black) and modified TvgA-4R (red) highlighting new features. D) CT-HSQC of the modified TvgA-4R. E) A schematic representation of TvgB-catalyzed modification.
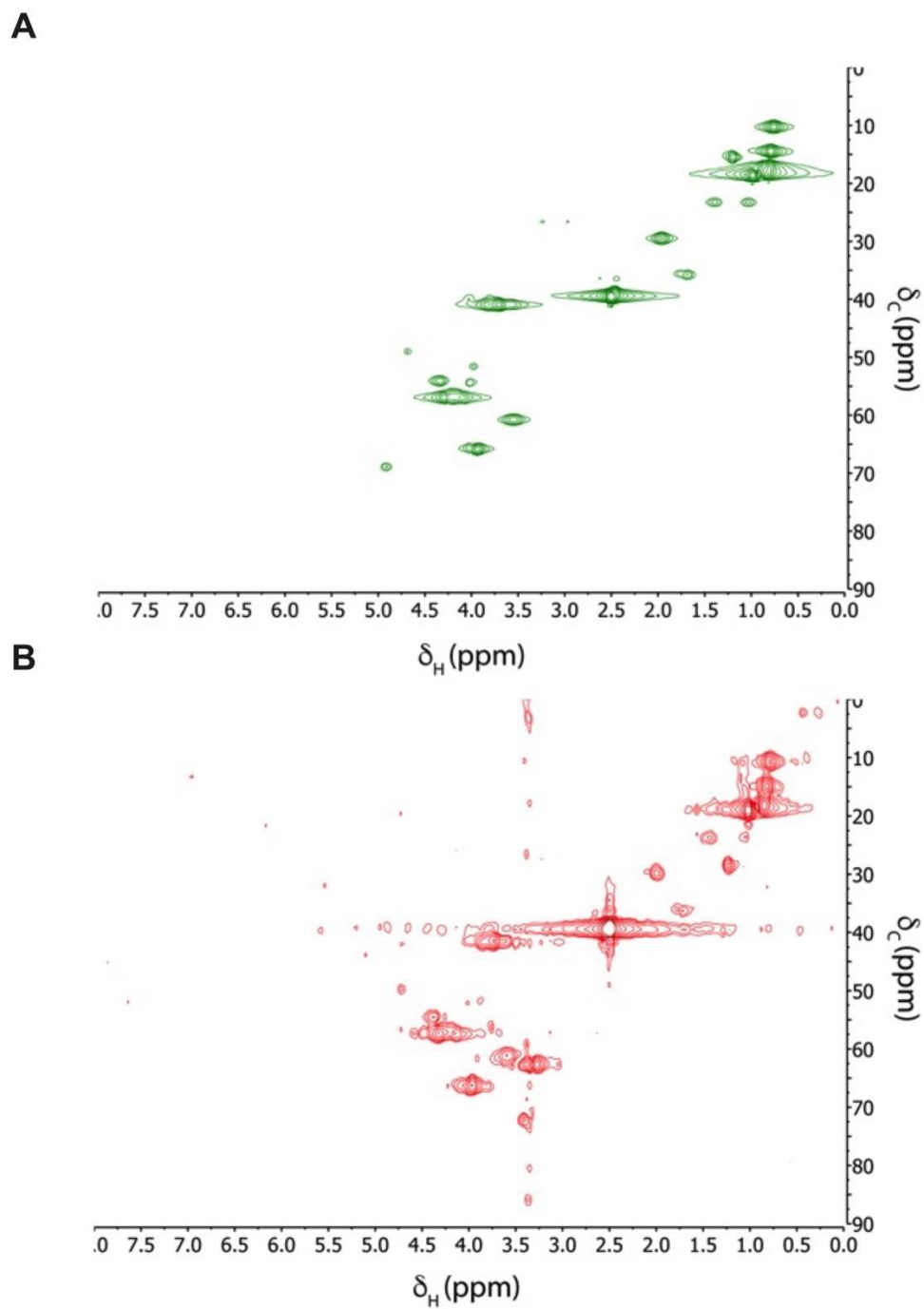
Figure 4.2 ¹³C HSQC NMR spectra. A) substrate peptide TvgA-4R. B) modified
TvgA-4R.
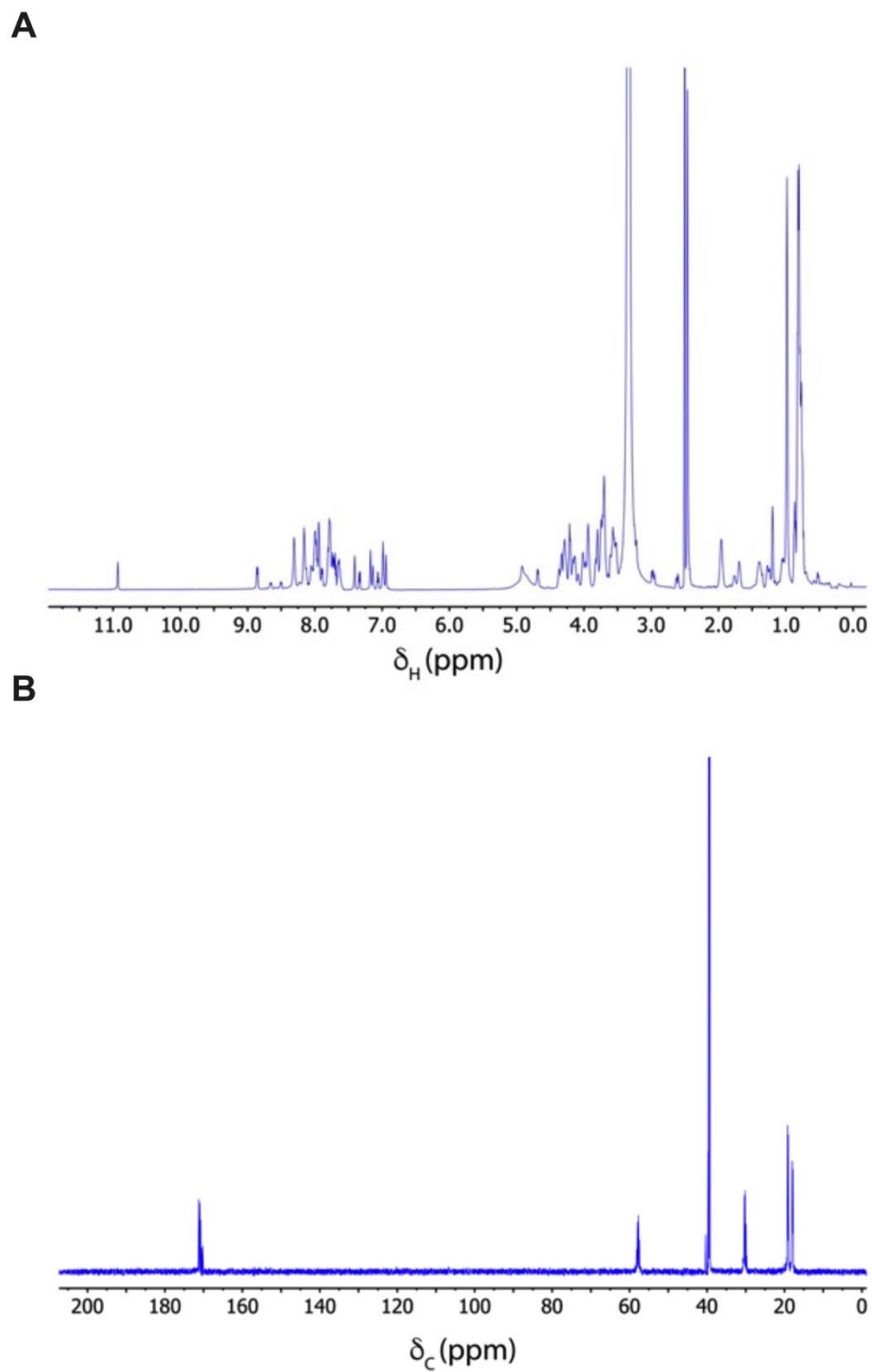
Figure 4.3 NMR spectra of the unmodified ^{13}C ^{15}N TvgA-4R peptide. A) ^{1}H NMR. B) ^{13}C NMR.
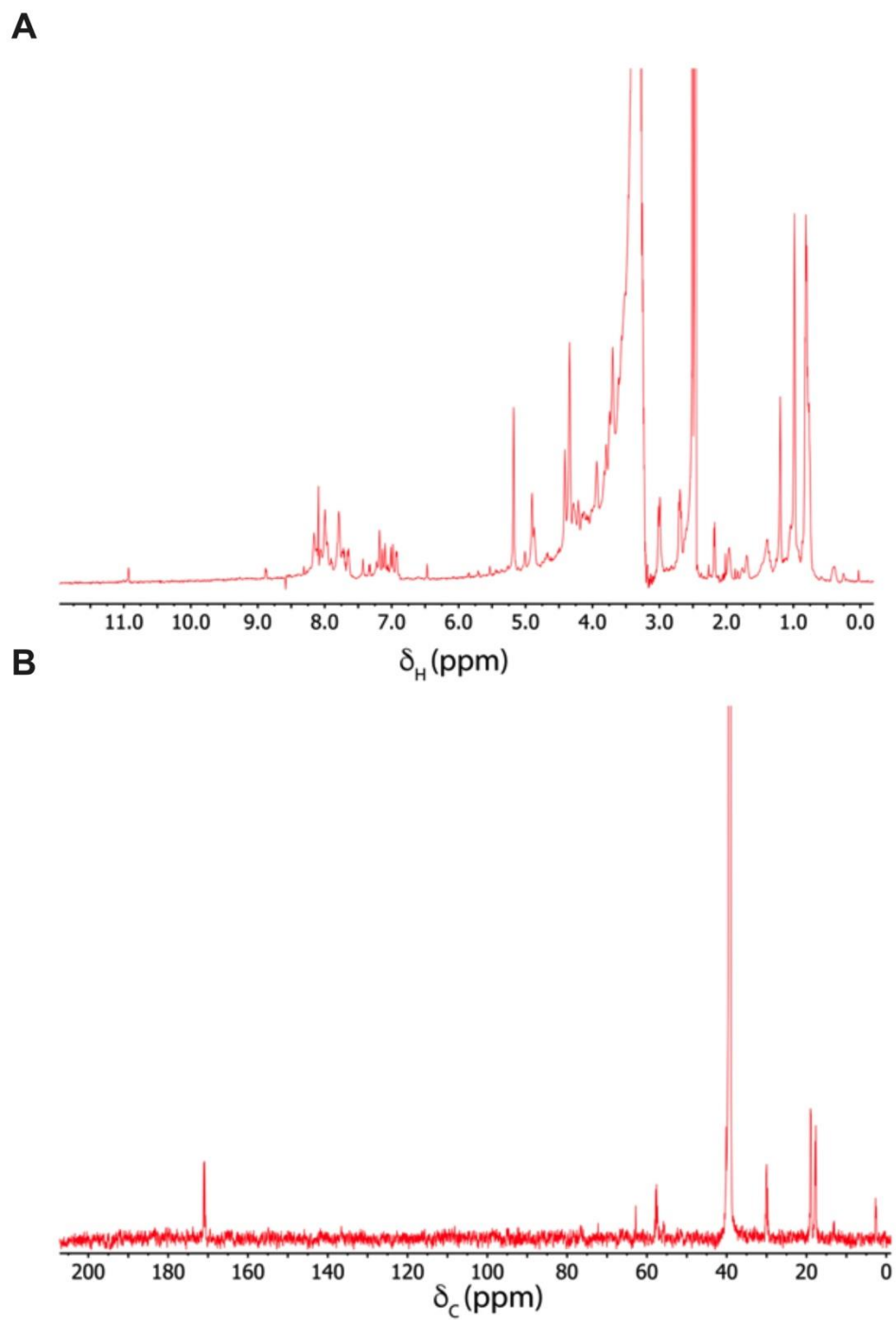
Figure 4.4 NMR spectra of the modified ^{13}C ^{15}N TvgA-4R peptide. A) ^{1}H NMR. B) ^{13}C NMR.
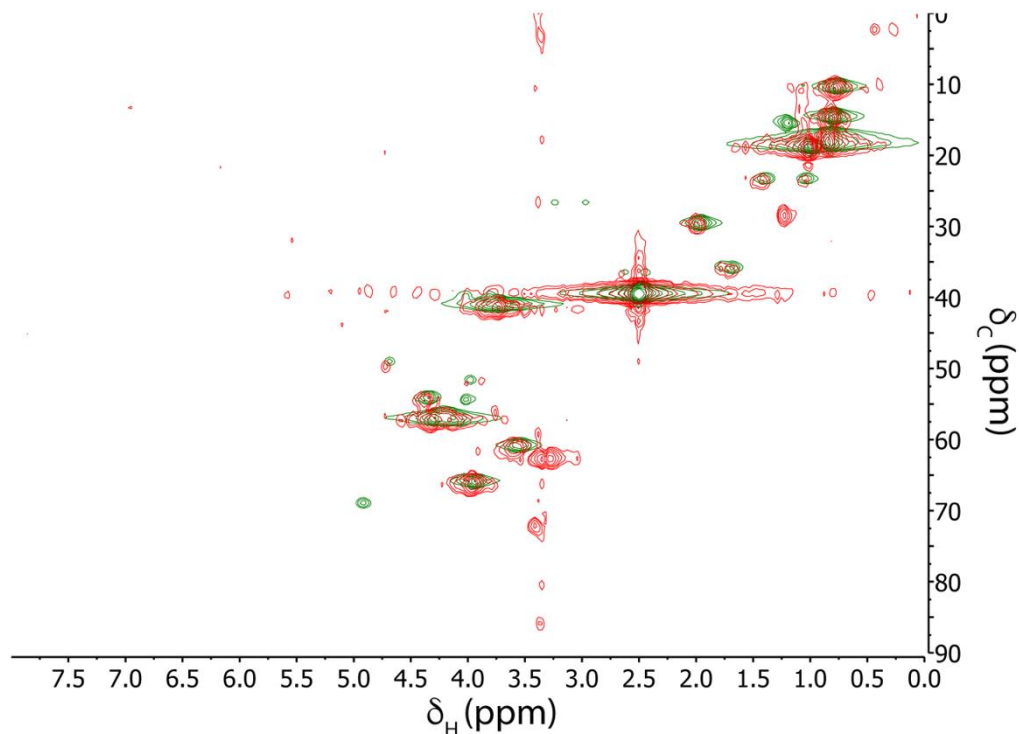
Figure 4.5 Overlaid ^{13}C HSQC of the substrate TvgA-4R (green) and the modified TvgA-4R (red).

4.2 Structural characterization of TigE protein

Collaborating with the Grove Lab from Albert Einstein College of Medicine we obtained a crystal structure of TigE protein. The substrate-free partial crystal structure was solved to 1.74Å resolution. Even though this structure is not fully solved, there is still a lot of information that could be extracted from it. The overall structure of TigE protein contain three functionally distinct domains: the N-terminal RRE domain, the radical SAM domain, and the C-terminal SPASM domain (Figure 4.6.

74
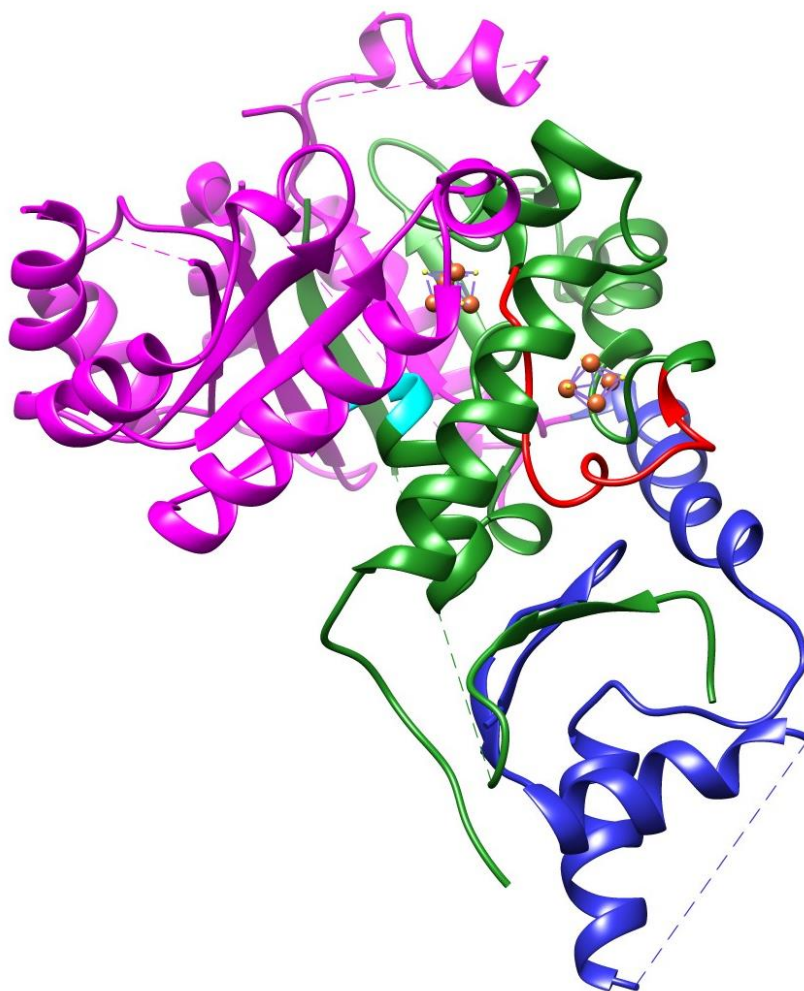
Figure 4.6 TigE partial structure solved at 1.74Å resolution. RRE domain is colored in blue; rSAM domain is shown in magenta; SPASM domain in green, consists of two auxiliary [4Fe-4S] clusters, $CX_2CX_5CX_3C$ motif (red), and GYW motif (cyan).

The N-terminal domain of rSAM enzymes adopt a typical winged-helix-turn (wHTH) motif which was founded to be conserved among prokaryotic RiPP classes: either fused with the enzyme or a separate small protein acting as a chaperone. The partial structure of TigE with the sequence homology to other rSAM enzymes suggests the existence of the

N-terminal RiPP precursor recognition element (RRE) domain fused to the TigE. This domain is composed of two antiparallel β-strands and three consecutive α-helices.

The rSAM cluster of TigE has a partial β/α triose partially solved (TIM) barrel. Cys 120, Cys 124, Cys 127 from the conserved $CX_3CX\Phi C$ motif (where Φ is an aromatic residue, Tyr in this case) ligate three out of four irons in the SAM cluster. The fourth iron is coordinated by the amino and carboxylate groups of the methionine on SAM.

The SPASM domain of TigE houses two auxiliary [4Fe-4S] clusters. Cys 360, Cys 377, Cys 427, and Tyr 339 coordinate auxiliary cluster I. Most [4Fe-4S] clusters are ligated by four Cys residues, or they may have one open site or non-Cys coordinating residue. For example, Asp ligation of auxiliary [4Fe-4S] cluster II in PqqE was shown to contribute to a more positive redox potential resulting in a thermodynamically favorable electron sink. To the best of our knowledge, amino acid Tyr has never been shown to participate as the ligand in coordinating iron from a canonical [4Fe-4S] cluster. However, for a nitrogenase when [8Fe-7S] cluster goes through two-electron oxidation it changes conformation in which a highly conserved oxygen-based residue (Ser/or, in some cases, Tyr) coordinates Fe-6 and a backbone amide of Cys – Fe-5 of the cluster. Ser ligation plays a role in protecting the structural and functional integrity of the [8Fe-7S] cluster under environmental stress. [138]
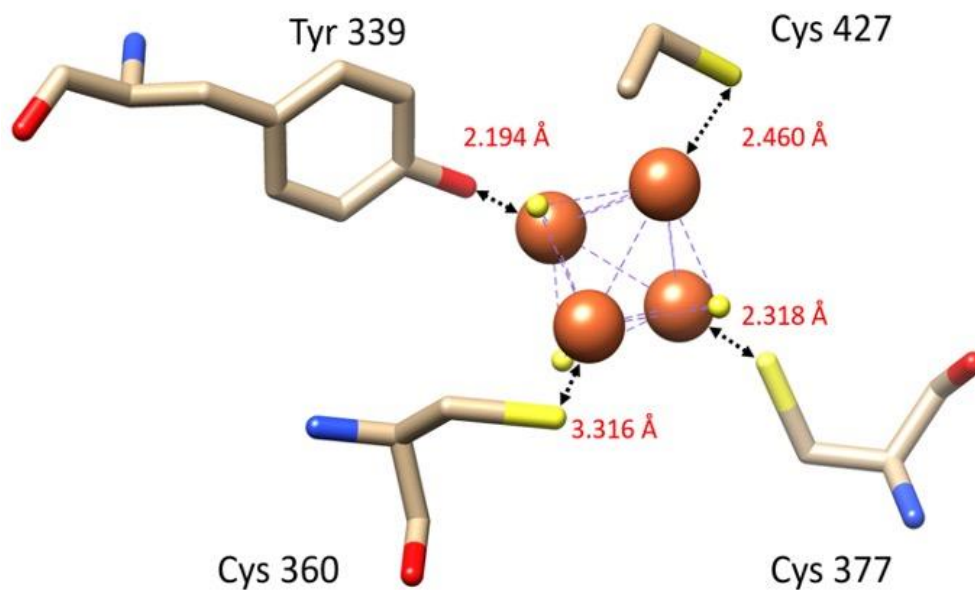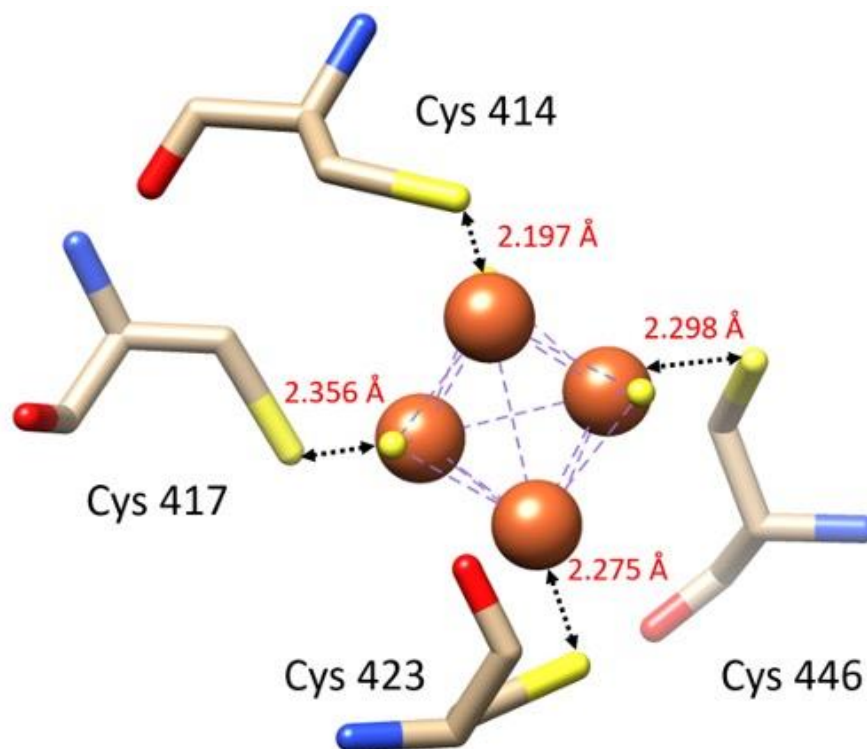
Figure 4.7 Auxiliary cluster I of the C-terminal SPASM domain. [4Fe-4S] cluster is shown in ball and stick representation; iron atoms are shown in orange and sulfur atoms are in yellow.

Cys 414, Cys 417, Cys 423, Cys 446 coordinate the second auxiliary cluster. The auxiliary clusters of C-terminal SPASM domain have a highly conserved $CX_2CX_5CX_3C$ motif where the first three cysteines coordinate the auxiliary cluster II, and the last cysteine coordinates auxiliary cluster I.

Figure 4.8 Auxiliary cluster II of the C-terminal SPASM domain. [4Fe-4S] cluster is shown in ball and stick representation; iron atoms are shown in orange and sulfur atoms are in yellow.

4.2.1 The role of Tyr 339 in the class of CPG synthases

Interestingly, a GYW motif, which houses the Tyr 339 residue, is conserved in the CP synthase, TvgB. Upon sequence analysis of the remaining 47 members belonging to the 1-1-201:AS65 cluster, we found the GYW motif to be strictly conserved despite sequence similarities ranging from 25-98% (Figure 4.9). We searched the Uniprot database for members of the rSAM-SPASM family that contain the GYW motif and only found members belonging to the 1-1-201:AS65 cluster. Simply removing either the Gly or Trp from the motif led to ambiguous results. Therefore, we suggest that GYW motif is specific to members of the 1-1-201:AS65 cluster, a CP-synthases motif.

Figure 4.9 Highly conserved CP-synthase motif for the members of the rSAM-SPASM family containing GYW motif belonging to the 1-1-201:AS65 cluster only.

The most notable feature from the TigE structure is Tyr 339 which lies within the CP synthase motif GYW and coordinates the open iron site from Aux I [4Fe-4S] cluster. To determine if Tyr 339 is required for catalysis, we generated the variants Y339A, Y339F, Y339C of TigE. We designed the primers with the desired mutation and used the TigE/ pET 28a as a template. Sequence verified plasmids and plasmids containing suf operon were co-transformed into E. coli BL21(DE3) cells for protein production The mutants were purified anaerobically using immobilized metal affinity chromatography, and reconstituted. The mutants show UV-Vis absorption spectra that are typical for an Fe-S cluster containing protein and similar to the wild type TigE (Figure 4.10).
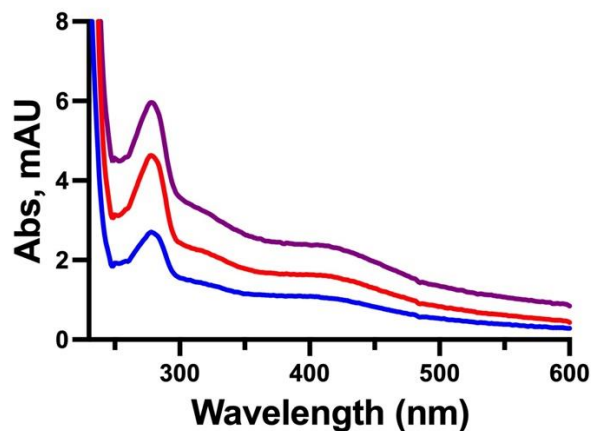
Figure 4.10 UV-visible spectrum of TigE mutants, the appearance of 410 nm peak suggests the presence of [4Fe-4S] clusters. TigE Y339C is shown in purple, TigE Y339A is shown in red, and TigE Y339F is colored blue.

In all cases, the TigE variants maintained near wild-type iron and sulfide content. The values are 10.0 ± 0.1 Fe and 10.0 ± 0.1 S per TigE monomer; 9.6 ± 0.1 Fe and 9.6 ± 0.3 S per TigE Y339A monomer; 12.2 ± 0.1 Fe and 13.0 ± 0.1 S per TigE Y339F monomer; 11.9 ± 0.9 Fe and 13.0 ± 0.1 S per TigE Y339C monomer. However, in reactions with TigB3R, none of the variants could modify the peptide substrate, even at longer incubation times. This suggests that Tyr 339 may be playing a role beyond coordination, such as redox tuning of the Aux I cluster.

This proposed mechanism, similarly to TvgB, requires the cleavage of two SAM molecules per product produced (Figure 4.11). In the first step, 5'-dA• abstracts a hydrogen from Ile-C_β. The loss of an electron followed by a deprotonation (formally a hydrogen atom) from ethyl $C\gamma$ results in β,γ-dehydro-Ile.[96] After release of 5'-dA and methionine and a second round of SAM cleavage, the 5'-dA• abstracts a hydrogen from an allylic Ile-$C\gamma$. The resonance stabilized radical combines with an electron from Ile C_β-C_γ unsaturated

80

bond with a concomitant quenching of the remaining free electron by a hydrogen forming

mCPG.[96]
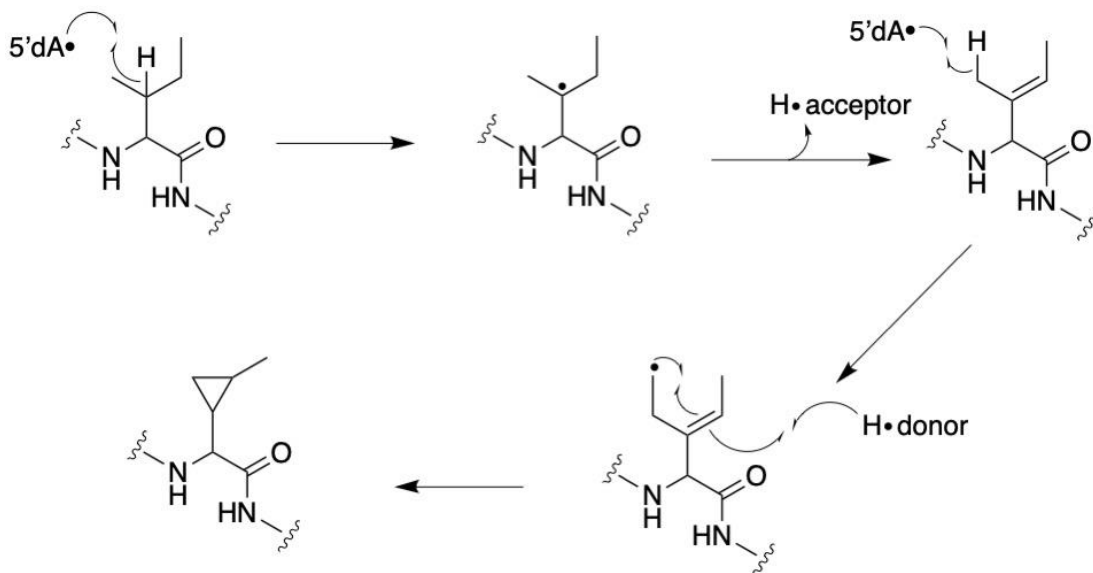
Figure 4.11 Proposed mechanism for TigE catalysis.

For this proposed mechanism, a hydrogen acceptor is required in the first oxidative

step, and a hydrogen donor is required for the second step. In the case of TigE, our

hypothesis is that Tyr 339 is participating directly in the catalysis as this hydrogen

donor/acceptor. This line of reasoning should be tested in the future studies. Evidence of

our mechanistic proposal is currently limited. In addition, deuterium labeling studies of the

Ile would allow us to understand which hydrogen atoms are removed or retained in the

final product.

4.3 Isolation and characterization of TigD

To expand on the exciting modification performed by CP synthase TigE we focused our attention on another rSAM-SPASM enzyme in the TIGSVS BGC, TigD protein. Like TigE, TigD houses a CP synthase motif GYW which was shown to be important for catalysis. N-terminally His-tagged TigD was recombinantly expressed in *E. coli* using plasmid containing suf operon to help with the incorporation of the iron-sulfur clusters into TigD. It was purified anaerobically using immobilized metal affinity chromatography and reconstituted with iron and sulfur (Figure 4.12A).
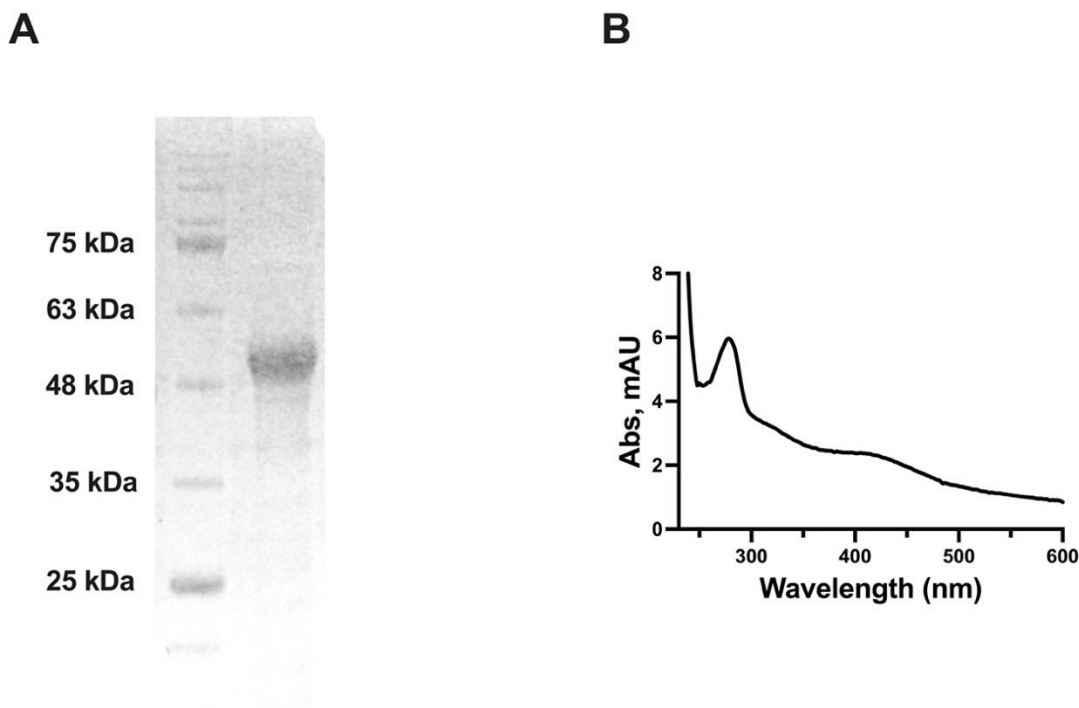
Figure 4.12 Characterization of TigD protein. A) SDS-PAGE gel analysis of TigD protein showing it was purified to near homogeneity and the molecular weight is approximately 52 kDa. Lane 1 is the blue stain protein marker. B) UV-visible spectrum of TigD protein, the 410 nm peak suggests the presence of [4Fe-4S] clusters.

Isolated protein subjected to colorimetric ferrozine and sulfide quantification assays lead to an average Fe/S content of 15.0 ± 0.1 Fe and 12.0 ± 0.1 S per TigD monomer. This result suggests the presence of at least three [4Fe-4S] clusters, analogous to anSME enzyme and other CP-synthases TvgB, and TigE. The UV-Vis spectrum of TigD protein shows the presence of [4Fe-4S] clusters in the enzyme with a peak absorbance at 410 nm (Figure 4.12B) in addition to the protein absorbance at 280 nm.

As other rSAM-SPASM enzymes, TigD can produce L-Met and 5'-deoxydenosine in the absence of the substrate. Reaction of sodium dithionite reduced TigD with SAM leads to the formation of 5'-deoxyadenosine.

Next, we set out to find the substrate for TigD protein. The precursor peptide TigB consists of TIGSVS motif that is repeated up to five times, suggesting that five identical products are formed from the peptide. We recently discovered that TigE installs a carbon-carbon bond between the Ile-C_γ's on TigB, forming methyl – CPG moiety. We expected that TigD would catalyze the formation of CPG from Val residues, like TvgB. Using microwave assisted solid peptide synthesis, we synthesized the truncated TigB peptide with three repeating TIGSVS motifs and a N-terminal Trp to increase absorbance and aid in purification of TigB3R (Figure 4.13A). To evaluate the enzymatic activity of TigD, reactions were set up containing TigD, SAM, dithiothreitol, and dithionite. A change in the retention time and mass of TigB3R was not observed by HPLC or LCMS (Figure 4.13B).
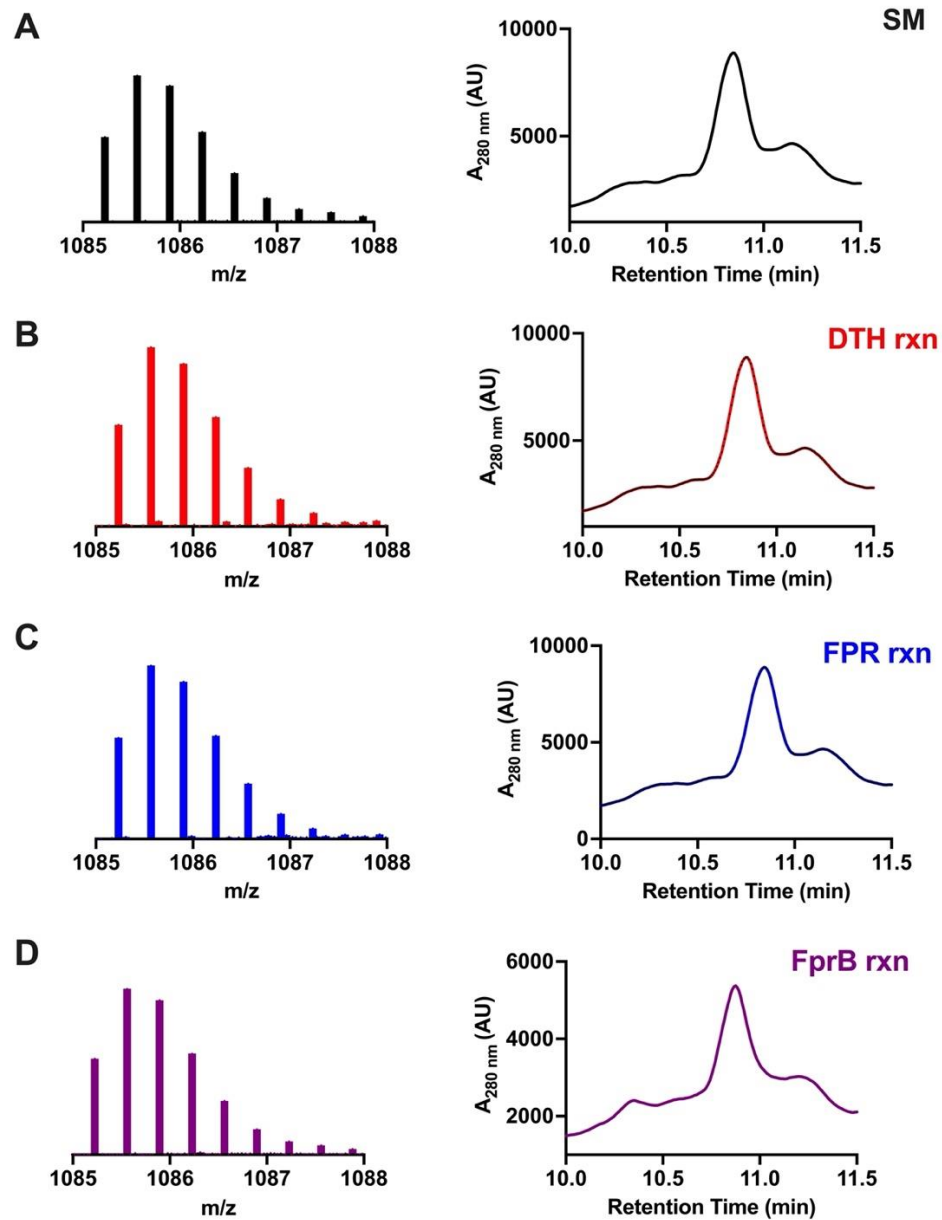
Figure 4.13 Analysis of TigD reaction with TigB3R. A) HRMS spectra of $[M+2H]^{3+}$ ions of TigB3R precursor peptide and LC chromatogram. B) HRMS spectra of $[M+2H]^{3+}$ ions of TigB3R precursor peptide from DTH reaction and LC chromatogram (red). C) HRMS spectra of $[M+2H]^{3+}$ ions of TigB3R precursor peptide from FPR reaction and LC chromatogram (blue). D) HRMS spectra of $[M+2H]^{3+}$ ions of TigB3R precursor peptide from FprB reaction and LC chromatogram (purple).

84

It is known that some rSAM enzymes, including TigE protein, are active in vitro only with the flavodoxin (FldA)/ flavodoxin (FPR) system.[27] Hence, we set up overnight reaction of TigD with TigB3R and SAM substituting DTH with FldA, FPR, and NADPH (Figure 4.13C). The reaction was quenched and evaluated by HPLC and LCMS. TigD is inactive in the presence of flavodoxin, encouraging us to find a homologous reducing system. One such enzyme used in redox metabolism is ferredoxin – $NADP^+$ reductase (FprB). To check the ability of FprB to support TigD activity we set up overnight reaction of TigD protein with TigB3R precursor peptide, SAM, FprB and NADPH (Figure 4.13D). Still, no reaction was observed.

Next, we hypothesized that TigD reacts with the product of the TigE catalyzed reaction, TigB3R-mCPG. Using the microwave assisted solid peptide synthesis, we synthesized the truncated TigB3R-mCPG with three repeating TIGSVS motifs (Fmoc - mCPG was used instead of Ile, shown in red) and a N-terminal Trp to increase absorbance and aid in purification of TigB3R-mCPG (Figure 4.14A). To evaluate the enzymatic activity of TigD, three sets of reactions were set up first - containing TigD, TigB3R-mCPG, SAM, dithiothreitol, and dithionite; second - TigD, TigB3R-mCPG, SAM, FldA, FPR, and NADPH; third - TigD, TigB3R-mCPG, SAM, FprB and NADPH Surprisingly, still no change in the retention time and mass of TigB3R-mCPG was observed by HPLC or LCMS.
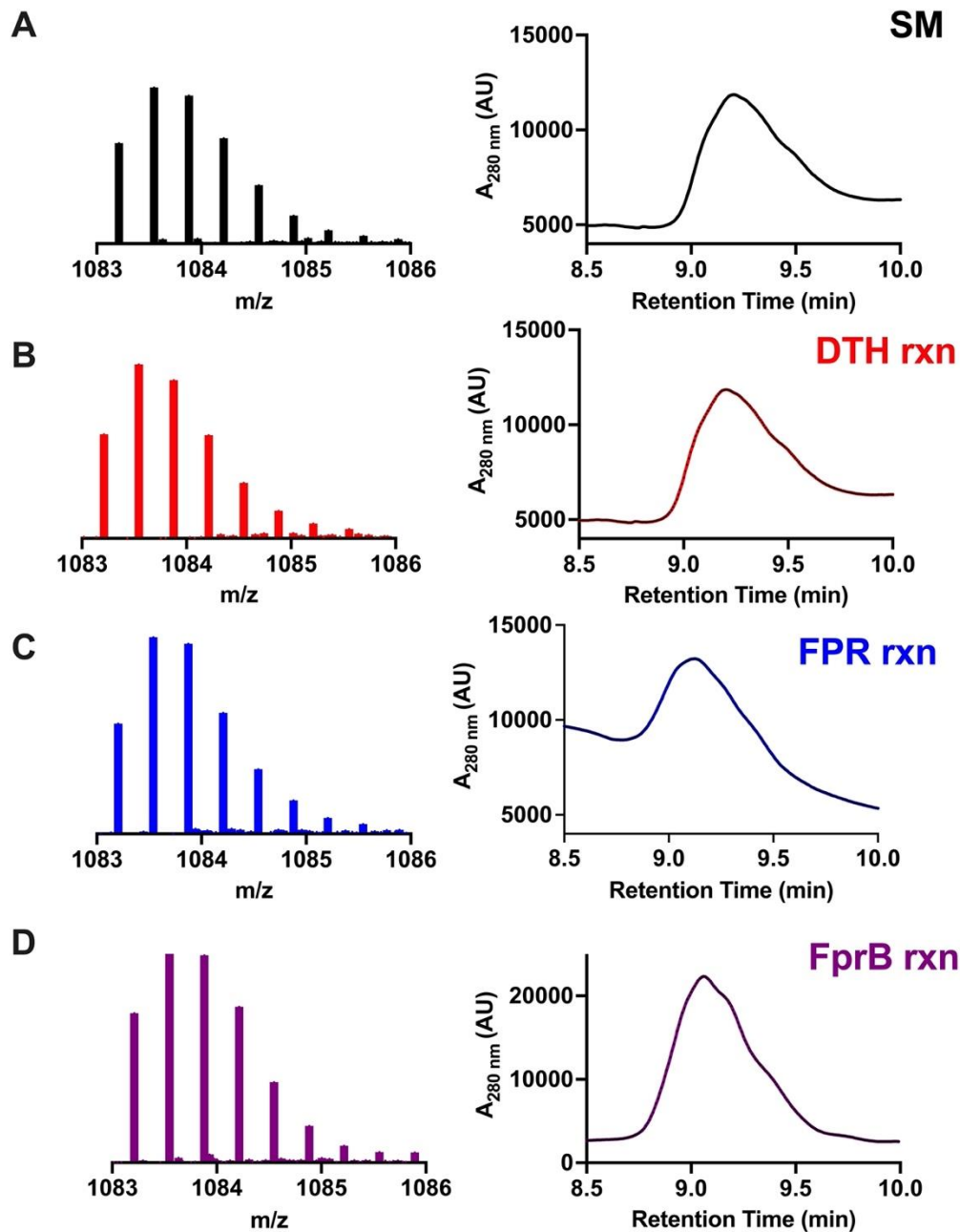
Figure 4.14 Analysis of TigD reaction with TigB3R-mCPG. A) HRMS spectra of [M+2H]$^{3+}$ ions of TigB3R-mCPG and LC chromatogram. B) HRMS spectra of [M+2H]$^{3+}$ ions of TigB3R-mCPG from DTH reaction and LC chromatogram (red). C) HRMS spectra of [M+2H]$^{3+}$ ions of TigB3R-mCPG from FPR reaction and LC chromatogram (blue). D) HRMS spectra of [M+2H]$^{3+}$ ions of TigB3R-mCPG from FprB reaction and LC chromatogram (purple).

All rSAM enzymes require one electron reduction of the $[4Fe-4S]^{2+}$ cluster to the $+1$-oxidation state. This usually is done using an artificial chemical reductant dithionite or a physiological reducing system as flavodoxin/flavodoxin reductase or ferredoxin $NADP^+$ reductase. However, the proposed reducing systems were not able to activate TigD to perform substrate – based catalysis. Instead, they led to the formation of uncoupled reaction to generate methionine and 5' – dA with no modification of the substrate. Perhaps, a use of a different suitable reducing system is needed to support TigD activation. Some RiPP BGCs carry out reactions with a well-defined order of the post-translational modifications. So, there might be a possibility that TigB3R precursor peptide or TigB3R- mCPG must be further modified before the reacting with TigD.

Chapter Five: Discussion

The *mft* BGC is one of the most widely distributed RiPP biosynthetic pathways and is highly concentrated in mycobacteria genomes. Of the ~625 unique species that encode for MFT biosynthesis, 300 are found in the mycobacteria genus including many pathogenic mycobacteria species.[80] Despite the wide distribution of the *mft* BGC and its frequent occurrence in mycobacterial pathogens, little is known about the physiological conditions that lead to MFT production. To address this gap in knowledge we focused our efforts on the putative regulator MftR, which has remained relatively unexplored. As a point of clarification, it should be noted that the gene designation of MftR used throughout this dissertation is derived from automated gene annotation used by Uniprot and NCBI and should not be confused with the MarR-family regulatory protein that has been named major facilitator transport regulator (Uniprot: Q2SVY7) and is associated with urate response in *Burkholderia* sp.[139,140]

In this study, we demonstrated that MftR binds to the promoter region of the *mft* gene cluster, suggesting that it acts as a cis regulatory element in the transcription of the *mft* BGC. Notably, we used DNase I footprinting to sequence the DNA binding region of MftR and identified a stretch of 27 bp residing -79 bp upstream to the start of the *mft* operon. We confirmed that MftR binds the 27 bp region by measuring the K_d of the MftR-DNA complex by fluorescence anisotropy. In addition, we used qRT-PCR and growth assays to

demonstrate that overexpression of *mftR* reduces transcript levels of the *mft* BGC. Additionally, we found that the *mft* operator is conserved in both sequence and location in the other mycobacterial genomes. This suggests that MftR homologues likely regulate the *mft* BGC using a similar operator sequence. Our findings are consistent with pathogen-sequencing data from *in vivo* samples of macrophages infected with MTB. In this study, Peterson *et al.*[141] found that Rv0691c is highly expressed and affects nearly 50 gene targets. Importantly, they found that upregulation of Rv0691c resulted in the repression of *mftB, mftC,* and *mftD* transcription. Additionally, our findings are consistent with regulatory networks that were built for transcription factors in MTB.[142] Using chromatin immunoprecipitation sequencing (CHiP-Seq) experiments, Minch *et al.*[142] found that Rv0691c binds to a region 5' to the *mftA* homologue Rv0691a. The motif consensus described by Minch *et al* is conserved in the O_{mft} sequence we show in Figure 3.2E. Taken together with our sequencing, binding affinity data, qRT-PCR data, and bioinformatic data, it is logical to conclude that MftR is a *mft* BGC transcriptional repressor in mycobacteria that harbor both the *mft* BGC and *mftR*.

Significantly, we found that MftR is activated by long-chain acyl-CoA's. Using competitive EMSAs and ITC we showed that acyl-CoA's ranging from C12-C18 are effectors of MftR *in vitro*. Additionally, we repurposed a fluorescent reporter system to show that the *mft* BGC is upregulated in Msmeg cultures when oleate is supplemented to growth media as compared to glucose alone. These findings suggest that MFT production and utilization is required for some aspect of fatty-acid metabolism. Both bioinformatic and direct evidence support this view. In Msmeg, putative MFT-dependent dehydrogenases

belonging to the TIGR03989 family are co-localized with fatty-acid modifying enzymes. For instance, *msmeg_4801* is co-localized with a putative 3-oxo-acyl carrier protein reductase and *msmeg_2204* is in a gene cluster that contains an acyl-CoA dehydrogenase. More directly, *in vivo* studies have shown that the *mymA* operon, consisting of the genes *rv3083-rv3089*, is required for cell wall maintenance[143,144] and for maintaining the mycolic acid composition in MTB when exposed to acidic pH.[145] Encoded in the *mymA* operon is Rv3086, a putative MFT-dependent dehydrogenase. While enzymatic activity has yet to be established for Rv3086, it has been proposed to carry out the conversion of terminal methyl groups of fatty acids to carboxylic groups for condensation.[144] Therefore, it is likely that the activation of the *mft* BGC by acyl-CoA's is due to MFT-dependent dehydrogenases that are associated with fatty acid metabolism.

We have found that MftR is at least one regulator of MFT biosynthesis. While our findings significantly progress current knowledge about the physiological conditions that induce MFT biosynthesis, we recognize that the regulatory network of MFT biosynthesis is likely incomplete. For instance, it is known that the MFT-dependent dehydrogenases Msmeg_6242 and Msmeg_1410 are required for primary alcohol and carveol catabolism, respectively,[80,146] yet neither catabolic pathway incorporates long-chain acyl-CoA's. Therefore, we expect other regulators influence the timing of MFT production. Supporting this notion, the MTB regulator, Rv0678, controls the expression level of *mmpS5-mmpL5* genes which are associated with azole efflux.[147] Additionally, CHiP-Seq experiments show that Rv0678 also binds to regions of DNA that encode for MftB and MftD, potentially making Rv0678 a secondary regulator of MFT biosynthesis in MTB.[148] While the

regulatory network for MFT biosynthesis may not be complete, our findings here provide the new insight to the physiological conditions that lead to MFT production.

Despite the rapid progress in solving the biosynthesis, structure, and function of MFT, little is known about the physiological processes that require the molecule. Previously, we speculated that MftR is a regulator of MFT biosynthesis.[80] The present study confirms this to be true, at least in Msmeg.

Overall, this study underpins the importance of MFT in mycobacteria. However, to fully grasp the physiological roles that MFT participates in, future studies should be dedicated to investigating chemistries of MFT-dependent dehydrogenases.

Chapter Six: Summary

Here, we report that *msmeg_1420*, annotated as MftR, is a transcriptional repressor of the mycofactocin biosynthetic gene cluster (BGC) in Msmeg. We found that MftR binds a DNA sequence in the promoter region of the mycofactocin BGC. We mapped the 27 bp *mft* operator (O_{mft}) by DNAse I footprinting and measured dissociation constant (K_d) of the MftR – O_{mft} complex by fluorescence anisotropy. We employed relative RT-quantitative PCR (qRT-PCR) to demonstrate that overexpression of MftR results in the repression of *mft* genes in Msmeg.

To determine under what conditions the biosynthesis of mycofactocin might be induced, we employed electrophoretic mobility shift assays and isothermal titration calorimetry to identify effectors of MftR. We identified that MftR is activated by long – chain acyl CoAs ranging from C12 to C18.

To demonstrate that the identified effectors translate *in vivo*, we use an engineered fluorescence reporter system to show that when it is supplemented to growth media it induces the expression of mycofactocin BGC in Msmeg.

In summary, we provide here mechanistic insights into the MftR dependent regulation of MFT biosynthesis and demonstrate that the induction of MFT biosynthesis is acyl-CoA dependent.

References

(1)   Murray, C. J.; Ikuta, K. S.; Sharara, F.; Swetschinski, L.; Robles Aguilar, G.; Gray, A.; Han, C.; Bisignano, C.; Rao, P.; Wool, E.; Johnson, S. C.; Browne, A. J.; Chipeta, M. G.; Fell, F.; Hackett, S.; Haines-Woodhouse, G.; Kashef Hamadani, B. H.; Kumaran, E. A. P.; McManigal, B.; Agarwal, R.; Akech, S.; Albertson, S.; Amuasi, J.; Andrews, J.; Aravkin, A.; Ashley, E.; Bailey, F.; Baker, S.; Basnyat, B.; Bekker, A.; Bender, R.; Bethou, A.; Bielicki, J.; Boonkasidecha, S.; Bukosia, J.; Carvalheiro, C.; Castañeda-Orjuela, C.; Chansamouth, V.; Chaurasia, S.; Chiurchiù, S.; Chowdhury, F.; Cook, A. J.; Cooper, B.; Cressey, T. R.; Criollo-Mora, E.; Cunningham, M.; Darboe, S.; Day, N. P. J.; De Luca, M.; Dokova, K.; Dramowski, A.; Dunachie, S. J.; Eckmanns, T.; Eibach, D.; Emami, A.; Feasey, N.; Fisher-Pearson, N.; Forrest, K.; Garrett, D.; Gastmeier, P.; Giref, A. Z.; Greer, R. C.; Gupta, V.; Haller, S.; Haselbeck, A.; Hay, S. I.; Holm, M.; Hopkins, S.; Iregbu, K. C.; Jacobs, J.; Jarovsky, D.; Javanmardi, F.; Khorana, M.; Kissoon, N.; Kobeissi, E.; Kostyanev, T.; Krapp, F.; Krumkamp, R.; Kumar, A.; Kyu, H. H.; Lim, C.; Limmathurotsakul, D.; Loftus, M. J.; Lunn, M.; Ma, J.; Mturi, N.; Munera-Huertas, T.; Musicha, P.; Mussi-Pinhata, M. M.; Nakamura, T.; Nanavati, R.; Nangia, S.; Newton, P.; Ngoun, C.; Novotney, A.; Nwakanma, D.; Obiero, C. W.; Olivas-Martinez, A.; Olliaro, P.; Ooko, E.; Ortiz-Brizuela, E.; Peleg, A. Y.; Perrone, C.; Plakkal, N.; Ponce-de-Leon, A.; Raad, M.; Ramdin, T.; Riddell, A.; Roberts, T.; Robotham, J. V.; Roca, A.; Rudd, K. E.; Russell, N.; Schnall, J.; Scott,

J. A. G.; Shivamallappa, M.; Sifuentes-Osornio, J.; Steenkeste, N.; Stewardson, A. J.; Stoeva, T.; Tasak, N.; Thaiprakong, A.; Thwaites, G.; Turner, C.; Turner, P.; van Doorn, H. R.; Velaphi, S.; Vongpradith, A.; Vu, H.; Walsh, T.; Waner, S.; Wangrangsimakul, T.; Wozniak, T.; Zheng, P.; Sartorius, B.; Lopez, A. D.; Stergachis, A.; Moore, C.; Dolecek, C.; Naghavi, M. Global Burden of Bacterial Antimicrobial Resistance in 2019: A Systematic Analysis. *Lancet* **2022**, *399* (10325), 629–655. https://doi.org/10.1016/S0140-6736(21)02724-0.

(2)   World Health Organization. Antibiotic resistance https://www.who.int/news-room/fact-sheets/detail/antibiotic-resistance.

(3)   Newman, D. J.; Cragg, G. M. Natural Products as Sources of New Drugs from 1981 to 2014. *J. Nat. Prod.* **2016**, *79* (3), 629–661. https://doi.org/10.1021/acs.jnatprod.5b01055.

(4)   Katz, L.; Baltz, R. H. Natural Product Discovery: Past, Present, and Future. *J. Ind. Microbiol. Biotechnol.* **2016**, *43* (2–3), 155–176. https://doi.org/10.1007/s10295-015-1723-5.

(5)   Ganesan, A. The Impact of Natural Products upon Modern Drug Discovery. *Curr. Opin. Chem. Biol.* **2008**, *12* (3), 306–317. https://doi.org/10.1016/j.cbpa.2008.03.016.

(6)   Adnani, N.; Rajski, S. R.; Bugni, T. S. Symbiosis-Inspired Approaches to Antibiotic Discovery. *Nat. Prod. Rep.* **2017**, *34* (7), 784–814. https://doi.org/10.1039/C7NP00009J.

(7)   Masschelein, J.; Jenner, M.; Challis, G. L. Antibiotics from Gram-Negative

Bacteria: A Comprehensive Overview and Selected Biosynthetic Highlights. *Nat. Prod. Rep.* **2017**, *34* (7), 712–783. https://doi.org/10.1039/C7NP00010C.

(8)   Winn, M.; Fyans, J. K.; Zhuo, Y.; Micklefield, J. Recent Advances in Engineering Nonribosomal Peptide Assembly Lines. *Nat. Prod. Rep.* **2016**, *33* (2), 317–347. https://doi.org/10.1039/c5np00099h.

(9)   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.; van der Donk, W. A. Ribosomally Synthesized and Post-Translationally Modified Peptide Natural Products: Overview and Recommendations for a Universal Nomenclature. *Nat. Prod. Rep.* **2013**, *30* (1), 108–160. https://doi.org/10.1039/C2NP20085F.

(10)   Montalbán-López, M.; Scott, T. A.; Ramesh, S.; Rahman, I. R.; van Heel, A. J.; Viel, J. H.; Bandarian, V.; Dittmann, E.; Genilloud, O.; Goto, Y.; Grande Burgos,
95

M. J.; Hill, C.; Kim, S.; Koehnke, J.; Latham, J. A.; Link, A. J.; Martínez, B.; Nair, S. K.; Nicolet, Y.; Rebuffat, S.; Sahl, H.-G.; Sareen, D.; Schmidt, E. W.; Schmitt, L.; Severinov, K.; Süssmuth, R. D.; Truman, A. W.; Wang, H.; Weng, J.-K.; van Wezel, G. P.; Zhang, Q.; Zhong, J.; Piel, J.; Mitchell, D. A.; Kuipers, O. P.; van der Donk, W. A. New Developments in RiPP Discovery, Enzymology and Engineering. *Nat. Prod. Rep.* **2021**, *38* (1), 130–239. https://doi.org/10.1039/D0NP00027B.

(11) Mendauletova, A.; Kostenko, A.; Lien, Y.; Latham, J. How a Subfamily of Radical S-Adenosylmethionine Enzymes Became a Mainstay of Ribosomally Synthesized and Post-Translationally Modified Peptide Discovery. *ACS Bio Med Chem Au* **2022**, *2* (1), 53–59. https://doi.org/10.1021/acsbiomedchemau.1c00045.

(12) Molohon, K. J.; Saint-Vincent, P. M. B.; Park, S.; Doroghazi, J. R.; Maxson, T.; Hershfield, J. R.; Flatt, K. M.; Schroeder, N. E.; Ha, T.; Mitchell, D. A. Plantazolicin Is an Ultranarrow-Spectrum Antibiotic That Targets the Bacillus Anthracis Membrane. *ACS Infect. Dis.* **2016**, *2* (3), 207–220. https://doi.org/10.1021/acsinfecdis.5b00115.

(13) Kelly, W. L.; Pan, L.; Li, C. Thiostrepton Biosynthesis: Prototype for a New Family of Bacteriocins. *J. Am. Chem. Soc.* **2009**, *131* (12), 4327–4334. https://doi.org/10.1021/ja807890a.

(14) Mohr, K. I.; Volz, C.; Jansen, R.; Wray, V.; Hoffmann, J.; Bernecker, S.; Wink, J.; Gerth, K.; Stadler, M.; Müller, R. Pinensins: The First Antifungal Lantibiotics. *Angew. Chemie Int. Ed.* **2015**, *54* (38), 11254–11258.

https://doi.org/10.1002/anie.201500927.

(15)   Sánchez-Hidalgo, M.; Martín, J.; Genilloud, O. Identification and Heterologous

Expression of the Biosynthetic Gene Cluster Encoding the Lasso Peptide

Humidimycin, a Caspofungin Activity Potentiator. *Antibiotics* **2020**, *9* (2), 67.

https://doi.org/10.3390/antibiotics9020067.

(16)   Torres, N. I.; Noll, K. S.; Xu, S.; Li, J.; Huang, Q.; Sinko, P. J.; Wachsman, M. B.;

Chikindas, M. L. Safety, Formulation and In Vitro Antiviral Activity of the

Antimicrobial Peptide Subtilosin Against Herpes Simplex Virus Type 1.

Probiotics Antimicrob. Proteins **2013**, *5* (1), 26–35.

https://doi.org/10.1007/s12602-012-9123-x.

(17)   Frattaruolo, L.; Lacret, R.; Cappello, A. R.; Truman, A. W. A Genomics-Based

Approach Identifies a Thioviridamide-Like Compound with Selective Anticancer

Activity. *ACS Chem. Biol.* **2017**, *12* (11), 2815–2822.

https://doi.org/10.1021/acschembio.7b00677.

(18)   Rowe, S. M.; Spring, D. R. The Role of Chemical Synthesis in Developing RiPP

Antibiotics. *Chem. Soc. Rev.* **2021**, *50* (7), 4245–4258.

https://doi.org/10.1039/D0CS01386B.

(19)   Hudson, G. A.; Mitchell, D. A. RiPP Antibiotics: Biosynthesis and Engineering

Potential. *Curr. Opin. Microbiol.* **2018**, *45*, 61–69.

https://doi.org/10.1016/j.mib.2018.02.010.

(20)   Broderick, J. B.; Duffus, B. R.; Duschene, K. S.; Shepard, E. M. Radical S -

Adenosylmethionine Enzymes. *Chem. Rev.* **2014**, *114* (8), 4229–4317.

https://doi.org/10.1021/cr4004709.

(21)   Knappe, J.; Bohnert, E.; Brummer, W. S-Adenosyl-L-Methionine, a Component of

the Clastic Dissimilation of Pyruvate in Escherichia Coli. *Biochim. Biophys. Acta*

1965, *107* (3), 603–605. https://doi.org/10.1016/0304-4165(65)90205-9.

(22)   Knappe, J.; Schacht, J.; Möckel, W.; Höpner, T.; Vetter, H.; Edenharder, R.

Pyruvate Formate-Lyase Reaction in Escherichia Coli. The Enzymatic System

Converting an Inactive Form of the Lyase into the Catalytically Active Enzyme.

Eur. J. Biochem. **1969**, *11* (2), 316–327. https://doi.org/10.1111/j.1432-

1033.1969.tb00775.x.

(23)   Chirpich, T. P.; Zappia, V.; Costilow, R. N.; Barker, H. A. Lysine 2,3-

Aminomutase. Purification and Properties of a Pyridoxal Phosphate and S-

Adenosylmethionine-Activated Enzyme. *J. Biol. Chem.* **1970**, *245* (7), 1778–1789.

(24)   Sofia, H. J.; Chen, G.; Hetzler, B. G.; Reyes-Spindola, J. F.; Miller, N. E. Radical

SAM, a Novel Protein Superfamily Linking Unresolved Steps in Familiar

Biosynthetic Pathways with Radical Mechanisms: Functional Characterization

Using New Analysis and Information Visualization Methods. *Nucleic Acids Res.*

2001, *29* (5), 1097–1106. https://doi.org/10.1093/nar/29.5.1097.

(25)   Balo, A. R.; Tao, L.; Britt, R. D. Characterizing SPASM/Twitch Domain-

Containing Radical SAM Enzymes by EPR Spectroscopy. *Appl. Magn. Reson.*

2022, *53* (3–5), 809–820. https://doi.org/10.1007/s00723-021-01406-2.

(26)   Broderick, W. E.; Broderick, J. B. Radical SAM Enzymes: Surprises along the

Path to Understanding Mechanism. *JBIC J. Biol. Inorg. Chem.* **2019**, *24* (6), 769–

776. https://doi.org/10.1007/s00775-019-01706-w.

(27) Barr, I.; Latham, J. A.; Iavarone, A. T.; Chantarojsiri, T.; Hwang, J. D.; Klinman, J. P. 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.* **2016**, *291* (17), 8877–8884. https://doi.org/10.1074/jbc.C115.699918.

(28) Benjdia, A.; Decamps, L.; Guillot, A.; Kubiak, X.; Ruffié, P.; Sandström, C.; Berteau, O. Insights into the Catalysis of a Lysine-Tryptophan Bond in Bacterial Peptides by a SPASM Domain Radical S-Adenosylmethionine (SAM) Peptide Cyclase. *J. Biol. Chem.* **2017**, *292* (26), 10835–10844. https://doi.org/10.1074/jbc.M117.783464.

(29) Schramma, K. R.; Bushin, L. B.; Seyedsayamdost, M. R. Structure and Biosynthesis of a Macrocyclic Peptide Containing an Unprecedented Lysine-to-Tryptophan Crosslink. *Nat. Chem.* **2015**, *7* (5), 431–437. https://doi.org/10.1038/nchem.2237.

(30) Bruender, N. A.; Wilcoxen, J.; Britt, R. D.; Bandarian, V. Biochemical and Spectroscopic Characterization of a Radical S -Adenosyl- l -Methionine Enzyme Involved in the Formation of a Peptide Thioether Cross-Link. *Biochemistry* **2016**, *55* (14), 2122–2134. https://doi.org/10.1021/acs.biochem.6b00145.

(31) Clark, K. A.; Bushin, L. B.; Seyedsayamdost, M. R. Aliphatic Ether Bond Formation Expands the Scope of Radical SAM Enzymes in Natural Product Biosynthesis. *J. Am. Chem. Soc.* **2019**, *141* (27), 10610–10615.

https://doi.org/10.1021/jacs.9b05151.

(32) Morinaka, B. I.; Vagstad, A. L.; Helf, M. J.; Gugger, M.; Kegler, C.; Freeman, M. F.; Bode, H. B.; Piel, J. Radical S -Adenosyl Methionine Epimerases: Regioselective Introduction of Diverse D -Amino Acid Patterns into Peptide Natural Products. *Angew. Chemie Int. Ed.* **2014**, *53* (32), 8503–8507. https://doi.org/10.1002/anie.201400478.

(33) Rea, M. C.; Dobson, A.; O'Sullivan, O.; Crispie, F.; Fouhy, F.; Cotter, P. D.; Shanahan, F.; Kiely, B.; Hill, C.; Ross, R. P. Effect of Broad- and Narrow-Spectrum Antimicrobials on Clostridium Difficile and Microbial Diversity in a Model of the Distal Colon. *Proc. Natl. Acad. Sci.* **2011**, *108* (supplement_1), 4639–4644. https://doi.org/10.1073/pnas.1001224107.

(34) Khaliullin, B.; Ayikpoe, R.; Tuttle, M.; Latham, J. A. Mechanistic Elucidation of the Mycofactocin-Biosynthetic Radical S-Adenosylmethionine Protein, MftC. *J. Biol. Chem.* **2017**, *292* (31), 13022–13033. https://doi.org/10.1074/jbc.M117.795682.

(35) Shen, Y.-Q.; Bonnot, F.; Imsand, E. M.; RoseFigura, J. M.; Sjölander, K.; Klinman, J. P. Distribution and Properties of the Genes Encoding the Biosynthesis of the Bacterial Cofactor, Pyrroloquinoline Quinone. *Biochemistry* **2012**, *51* (11), 2265–2275. https://doi.org/10.1021/bi201763d.

(36) Schramma, K. R.; Seyedsayamdost, M. R. Lysine-Tryptophan-Crosslinked Peptides Produced by Radical SAM Enzymes in Pathogenic Streptococci. *ACS Chem. Biol.* **2017**, *12* (4), 922–927. https://doi.org/10.1021/acschembio.6b01069.
100

(37)   Fang, Q.; Peng, J.; Dierks, T. Post-Translational Formylglycine Modification of
        Bacterial Sulfatases by the Radical S-Adenosylmethionine Protein AtsB. *J. Biol.
        Chem.* **2004**, *279* (15), 14570–14578. https://doi.org/10.1074/jbc.M313855200.

(38)   Berteau, O.; Guillot, A.; Benjdia, A.; Rabot, S. A New Type of Bacterial Sulfatase
        Reveals a Novel Maturation Pathway in Prokaryotes. *J. Biol. Chem.* **2006**, *281*
        (32), 22464–22470. https://doi.org/10.1074/jbc.M602504200.

(39)   Szameit, C.; Miech, C.; Balleininger, M.; Schmidt, B.; von Figura, K.; Dierks, T.
        The Iron Sulfur Protein AtsB Is Required for Posttranslational Formation of
        Formylglycine in the Klebsiella Sulfatase. *J. Biol. Chem.* **1999**, *274* (22), 15375–
        15381. https://doi.org/10.1074/jbc.274.22.15375.

(40)   Grove, T. L.; Lee, K.-H.; St. Clair, J.; Krebs, C.; Booker, S. J. In Vitro
        Characterization of AtsB, a Radical SAM Formylglycine-Generating Enzyme That
        Contains Three [4Fe-4S] Clusters. *Biochemistry* **2008**, *47* (28), 7523–7538.
        https://doi.org/10.1021/bi8004297.

(41)   Grove, T. L.; Ahlum, J. H.; Qin, R. M.; Lanz, N. D.; Radle, M. I.; Krebs, C.;
        Booker, S. J. Further Characterization of Cys-Type and Ser-Type Anaerobic
        Sulfatase Maturating Enzymes Suggests a Commonality in the Mechanism of
        Catalysis. *Biochemistry* **2013**, *52* (17), 2874–2887.
        https://doi.org/10.1021/bi400136u.

(42)   Goldman, P. J.; Grove, T. L.; Booker, S. J.; Drennan, C. L. X-Ray Analysis of
        Butirosin Biosynthetic Enzyme BtrN Redefines Structural Motifs for AdoMet
        Radical Chemistry. *Proc. Natl. Acad. Sci.* **2013**, *110* (40), 15949–15954.

https://doi.org/10.1073/pnas.1312228110.

(43)   Haft, D. H.; Selengut, J. D.; Richter, R. A.; Harkins, D.; Basu, M. K.; Beck, E. TIGRFAMs and Genome Properties in 2013. *Nucleic Acids Res.* **2012**, *41* (D1), D387–D395. https://doi.org/10.1093/nar/gks1234.

(44)   Haft, D. H.; Basu, M. K. Biological Systems Discovery In Silico: Radical S - Adenosylmethionine Protein Families and Their Target Peptides for Posttranslational Modification. *J. Bacteriol.* **2011**, *193* (11), 2745–2755. https://doi.org/10.1128/JB.00040-11.

(45)   Grell, T. A. J.; Goldman, P. J.; Drennan, C. L. SPASM and Twitch Domains in S-Adenosylmethionine (SAM) Radical Enzymes. *J. Biol. Chem.* **2015**, *290* (7), 3964–3971. https://doi.org/10.1074/jbc.R114.581249.

(46)   Goosen, N.; Huinen, R. G.; van de Putte, P. A 24-Amino-Acid Polypeptide Is Essential for the Biosynthesis of the Coenzyme Pyrrolo-Quinoline-Quinone. *J. Bacteriol.* **1992**, *174* (4), 1426–1427. https://doi.org/10.1128/jb.174.4.1426-1427.1992.

(47)   Felder, M.; Gupta, A.; Verma, V.; Kumar, A.; Qazi, G. N.; Cullum, J. The Pyrroloquinoline Quinone Synthesis Genes of Gluconobacter Oxydans. *FEMS Microbiol. Lett.* **2000**, *193* (2), 231–236. https://doi.org/10.1016/S0378-1097(00)00487-0.

(48)   Haft, D. H. Bioinformatic Evidence for a Widely Distributed, Ribosomally Produced Electron Carrier Precursor, Its Maturation Proteins, and Its Nicotinoprotein Redox Partners. *BMC Genomics* **2011**, *12* (1), 1–21.

https://doi.org/10.1186/1471-2164-12-21.

(49)   Khaliullin, B.; Aggarwal, P.; Bubas, M.; Eaton, G. R.; Eaton, S. S.; Latham, J. A. Mycofactocin Biosynthesis: Modification of the Peptide MftA by the Radical S-Adenosylmethionine Protein MftC. *FEBS Lett.* **2016**, *590* (16), 2538–2548. https://doi.org/10.1002/1873-3468.12249.

(50)   Bruender, N. A.; Bandarian, V. The Radical S -Adenosyl- l -Methionine Enzyme MftC Catalyzes an Oxidative Decarboxylation of the C-Terminus of the MftA Peptide. *Biochemistry* **2016**, *55* (20), 2813–2816. https://doi.org/10.1021/acs.biochem.6b00355.

(51)   Nguyen, T. Q. N. P. D.; Tooh, Y. W.; Sugiyama, R.; Nguyen, T. Q. N. P. D.; Purushothaman, M.; Leow, L. C.; Hanif, K.; Yong, R. H. S.; Agatha, I.; Winnerdy, F. R.; Gugger, M.; Phan, A. T.; Morinaka, B. I. Post-Translational Formation of Strained Cyclophanes in Bacteria. *Nat. Chem.* **2020**, *12* (11), 1042–1053. https://doi.org/10.1038/s41557-020-0519-z.

(52)   Babasaki, K.; Toshifumi, T.; Shimonishi, Y.; Kurahashi, K. Subtilosin A, a New Antibiotic Peptide Produced by Bacillus Subtilis 168: Isolation, Structural Analysis, and Biogenesis1. *J. Biochem.* **1985**, *98* (3), 585–603. https://doi.org/10.1093/oxfordjournals.jbchem.a135315.

(53)   Zheng, G.; Hehn, R.; Zuber, P. Mutational Analysis of the Sbo-Alb Locus of Bacillus Subtilis : Identification of Genes Required for Subtilosin Production and Immunity. *J. Bacteriol.* **2000**, *182* (11), 3266–3273. https://doi.org/10.1128/JB.182.11.3266-3273.2000.

(54) Kawulka, K. E.; Sprules, T.; Diaper, C. M.; Whittal, R. M.; McKay, R. T.; Mercier, P.; Zuber, P.; Vederas, J. C. Structure of Subtilosin A, a Cyclic Antimicrobial Peptide from Bacillus Subtilis with Unusual Sulfur to α-Carbon Cross-Links: Formation and Reduction of α-Thio-α-Amino Acid Derivatives ,. *Biochemistry* **2004**, *43* (12), 3385–3395. https://doi.org/10.1021/bi0359527.

(55) Flühe, L., Knappe, T., Gattner, M. et al.; Leif, F.; Thomas A., K.; Michael J., G.; Antje, S.; Olaf, B.; Uwe, L.; Mohamed A., M.; Flühe, L.; Knappe, T. A.; Gattner, M. J.; Schäfer, A.; Burghaus, O.; Linne, U.; Marahiel, M. A. The Radical SAM Enzyme AlbA Catalyzes Thioether Bond Formation in Subtilosin A. *Nat. Chem. Biol.* **2012**, *8* (4), 350–357. https://doi.org/10.1038/nchembio.798.

(56) Benjdia, A.; Guillot, A.; Lefranc, B.; Vaudry, H.; Leprince, J.; Berteau, O. Thioether Bond Formation by SPASM Domain Radical SAM Enzymes: C α H-Atom Abstraction in Subtilosin A Biosynthesis. *Chem. Commun.* **2016**, *52* (37), 6249–6252. https://doi.org/10.1039/C6CC01317A.

(57) Murphy, K.; O'Sullivan, O.; Rea, M. C.; Cotter, P. D.; Ross, R. P.; Hill, C. Genome Mining for Radical SAM Protein Determinants Reveals Multiple Sactibiotic-Like Gene Clusters. *PLoS One* **2011**, *6* (7), e20852. https://doi.org/10.1371/journal.pone.0020852.

(58) Flühe, L.; Burghaus, O.; Wieckowski, B. M.; Giessen, T. W.; Linne, U.; Marahiel, M. A. 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.* **2013**, *135* (3), 959–962.

https://doi.org/10.1021/ja310542g.

(59) Rea, M. C.; Sit, C. S.; Clayton, E.; O'Connor, P. M.; Whittal, R. M.; Zheng, J.; Vederas, J. C.; Ross, R. P.; Hill, C. Thuricin CD, a Posttranslationally Modified Bacteriocin with a Narrow Spectrum of Activity against Clostridium Difficile. *Proc. Natl. Acad. Sci.* **2010**, *107* (20), 9352–9357. https://doi.org/10.1073/pnas.0913554107.

(60) Mo, T.; Ji, X.; Yuan, W.; Mandalapu, D.; Wang, F.; Zhong, Y.; Li, F.; Chen, Q.; Ding, W.; Deng, Z.; Yu, S.; Zhang, Q. Thuricin Z: A Narrow-Spectrum Sactibiotic That Targets the Cell Membrane. *Angew. Chemie Int. Ed.* **2019**, *58* (52), 18793–18797. https://doi.org/10.1002/anie.201908490.

(61) Balty, C.; Guillot, A.; Fradale, L.; Brewee, C.; Boulay, M.; Kubiak, X.; Benjdia, A.; Berteau, O. Ruminococcin C, an Anti-Clostridial Sactipeptide Produced by a Prominent Member of the Human Microbiota Ruminococcus Gnavus. *J. Biol. Chem.* **2019**, *294* (40), 14512–14525. https://doi.org/10.1074/jbc.RA119.009416.

(62) Bushin, L. B.; Covington, B. C.; Rued, B. E.; Federle, M. J.; Seyedsayamdost, M. R. Discovery and Biosynthesis of Streptosactin, a Sactipeptide with an Alternative Topology Encoded by Commensal Bacteria in the Human Microbiome. *J. Am. Chem. Soc.* **2020**, *142* (38), 16265–16275. https://doi.org/10.1021/jacs.0c05546.

(63) Wieckowski, B. M.; Hegemann, J. D.; Mielcarek, A.; Boss, L.; Burghaus, O.; Marahiel, M. A. The PqqD Homologous Domain of the Radical SAM Enzyme ThnB Is Required for Thioether Bond Formation during Thurincin H Maturation. *FEBS Lett.* **2015**, *589* (15), 1802–1806.

https://doi.org/10.1016/j.febslet.2015.05.032.

(64)   Grove, T. L.; Himes, P. M.; Hwang, S.; Yumerefendi, H.; Bonanno, J. B.;

Kuhlman, B.; Almo, S. C.; Bowers, A. A. Structural Insights into Thioether Bond

Formation in the Biosynthesis of Sactipeptides. *J. Am. Chem. Soc.* **2017**, *139* (34),

11734–11744. https://doi.org/10.1021/jacs.7b01283.

(65)   Nakai, T.; Ito, H.; Kobayashi, K.; Takahashi, Y.; Hori, H.; Tsubaki, M.; Tanizawa,

K.; Okajima, T. The Radical S-Adenosyl-l-Methionine Enzyme QhpD Catalyzes

Sequential Formation of Intra-Protein Sulfur-to-Methylene Carbon Thioether

Bonds. *J. Biol. Chem.* **2015**, *290* (17), 11144–11166.

https://doi.org/10.1074/jbc.M115.638320.

(66)   Caruso, A.; Bushin, L. B.; Clark, K. A.; Martinie, R. J.; Seyedsayamdost, M. R.;

Alessio, C.; Leah B., B.; Kenzie A., C.; Ryan J., M.; Mohammad R., S.; Caruso,

A.; Bushin, L. B.; Clark, K. A.; Martinie, R. J.; Seyedsayamdost, M. R. Radical

Approach to Enzymatic β-Thioether Bond Formation. *J. Am. Chem. Soc.* **2019**,

141 (2), 990–997. https://doi.org/10.1021/jacs.8b11060.

(67)   Precord, T. W.; Mahanta, N.; Mitchell, D. A. Reconstitution and Substrate

Specificity of the Thioether-Forming Radical S -Adenosylmethionine Enzyme in

Freyrasin Biosynthesis. *ACS Chem. Biol.* **2019**, *14* (9), 1981–1989.

https://doi.org/10.1021/acschembio.9b00457.

(68)   Ibrahim, M.; Guillot, A.; Wessner, F.; Algaron, F.; Besset, C.; Courtin, P.; Gardan,

R.; Monnet, V. V. Control of the Transcription of a Short Gene Encoding a Cyclic

Peptide in Streptococcus Thermophilus : A New Quorum-Sensing System? *J.*

Bacteriol. **2007**, *189* (24), 8844–8854. https://doi.org/10.1128/JB.01057-07.

(69)   Ma, S.; Chen, H.; Li, H.; Ji, X.; Deng, Z.; Ding, W.; Zhang, Q. Post-Translational

Formation of Aminomalonate by a Promiscuous Peptide-Modifying Radical SAM

Enzyme. *Angew. Chemie Int. Ed.* **2021**, *60* (36), 19957–19964.

https://doi.org/10.1002/anie.202107192.

(70)   Bushin, L. B.; Clark, K. A.; Pelczer, I.; Seyedsayamdost, M. R. Charting an

Unexplored Streptococcal Biosynthetic Landscape Reveals a Unique Peptide

Cyclization Motif. *J. Am. Chem. Soc.* **2018**, *140* (50), 17674–17684.

https://doi.org/10.1021/jacs.8b10266.

(71)   Caruso, A.; Martinie, R. J.; Bushin, L. B.; Seyedsayamdost, M. R.

Macrocyclization via an Arginine-Tyrosine Crosslink Broadens the Reaction

Scope of Radical S -Adenosylmethionine Enzymes. *J. Am. Chem. Soc.* **2019**, *141*

(42), 16610–16614. https://doi.org/10.1021/jacs.9b09210.

(72)   Hudson, G. A.; Burkhart, B. J.; DiCaprio, A. J.; Schwalen, C. J.; Kille, B.;

Pogorelov, T. V.; Mitchell, D. A. Bioinformatic Mapping of Radical S -

Adenosylmethionine-Dependent Ribosomally Synthesized and Post-

Translationally Modified Peptides Identifies New Cα, Cβ, and Cγ-Linked

Thioether-Containing Peptides. *J. Am. Chem. Soc.* **2019**, *141* (20), 8228–8238.

https://doi.org/10.1021/jacs.9b01519.

(73)   Blin, K.; Medema, M. H.; Kazempour, D.; Fischbach, M. A.; Breitling, R.;

Takano, E.; Weber, T. AntiSMASH 2.0—a Versatile Platform for Genome Mining

of Secondary Metabolite Producers. *Nucleic Acids Res.* **2013**, *41* (W1), W204–

W212. https://doi.org/10.1093/nar/gkt449.

(74)   Blin, K.; Shaw, S.; Kloosterman, A. M.; Charlop-Powers, Z.; van Wezel, G. P.;
        Medema, M. H.; Weber, T. AntiSMASH 6.0: Improving Cluster Detection and
        Comparison Capabilities. *Nucleic Acids Res.* **2021**, *49* (W1), W29–W35.
        https://doi.org/10.1093/nar/gkab335.

(75)   Medema, M. H.; Blin, K.; Cimermancic, P.; de Jager, V.; Zakrzewski, P.;
        Fischbach, M. A.; Weber, T.; Takano, E.; Breitling, R. AntiSMASH: Rapid
        Identification, Annotation and Analysis of Secondary Metabolite Biosynthesis
        Gene Clusters in Bacterial and Fungal Genome Sequences. *Nucleic Acids Res.*
        2011, *39* (suppl_2), W339–W346. https://doi.org/10.1093/nar/gkr466.

(76)   Tietz, J. I.; Schwalen, C. J.; Patel, P. S.; Maxson, T.; Blair, P. M.; Tai, H.-C.;
        Zakai, U. I.; Mitchell, D. A. A New Genome-Mining Tool Redefines the Lasso
        Peptide Biosynthetic Landscape. *Nat. Chem. Biol.* **2017**, *13* (5), 470–478.
        https://doi.org/10.1038/nchembio.2319.

(77)   van Heel, A. J.; de Jong, A.; Song, C.; Viel, J. H.; Kok, J.; Kuipers, O. P.
        BAGEL4: A User-Friendly Web Server to Thoroughly Mine RiPPs and
        Bacteriocins. *Nucleic Acids Res.* **2018**, *46* (W1), W278–W281.
        https://doi.org/10.1093/nar/gky383.

(78)   Zhong, Z.; He, B.; Li, J.; Li, Y.-X. Challenges and Advances in Genome Mining
        of Ribosomally Synthesized and Post-Translationally Modified Peptides (RiPPs).
        Synth. Syst. Biotechnol. **2020**, *5* (3), 155–172.
        https://doi.org/10.1016/j.synbio.2020.06.002.

(79)   Gerlt, J. A.; Bouvier, J. T.; Davidson, D. B.; Imker, H. J.; Sadkhin, B.; Slater, D. R.; Whalen, K. L. Enzyme Function Initiative-Enzyme Similarity Tool (EFI-EST): A Web Tool for Generating Protein Sequence Similarity Networks. *Biochim. Biophys. Acta - Proteins Proteomics* **2015**, *1854* (8), 1019–1037. https://doi.org/10.1016/j.bbapap.2015.04.015.

(80)   Ayikpoe, R.; Govindarajan, V.; Latham, J. A. Occurrence, Function, and Biosynthesis of Mycofactocin. *Appl. Microbiol. Biotechnol.* **2019**, *103* (7), 2903–2912. https://doi.org/10.1007/s00253-019-09684-4.

(81)   Ayikpoe, R.; Salazar, J.; Majestic, B.; Latham, J. A. Mycofactocin Biosynthesis Proceeds through 3-Amino-5-[(p -Hydroxyphenyl)Methyl]-4,4-Dimethyl-2-Pyrrolidinone (AHDP); Direct Observation of MftE Specificity toward MftA*. *Biochemistry* **2018**, *57* (37), 5379–5383. https://doi.org/10.1021/acs.biochem.8b00816.

(82)   Ayikpoe, R. S.; Latham, J. A. 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.* **2019**, *141* (34), 13582–13591. https://doi.org/10.1021/jacs.9b06102.

(83)   Peña-Ortiz, L.; Graça, A. P.; Guo, H.; Braga, D.; Köllner, T. G.; Regestein, L.; Beemelmanns, C.; Lackner, G. Structure Elucidation of the Redox Cofactor Mycofactocin Reveals Oligo-Glycosylation by MftF. *Chem. Sci.* **2020**, *11* (20), 5182–5190. https://doi.org/10.1039/D0SC01172J.

(84)   Khaliullin, B.; Aggarwal, P.; Bubas, M.; Eaton, G. R.; Eaton, S. S.; Latham, J. A.

Mycofactocin Biosynthesis: Modification of the Peptide MftA by the Radical S-Adenosylmethionine Protein MftC. *FEBS Lett.* **2016**, *590* (16), 2538–2548. https://doi.org/10.1002/1873-3468.12249.

(85) Ayikpoe, R.; Ngendahimana, T.; Langton, M.; Bonitatibus, S.; Walker, L. M.; Eaton, S. S.; Eaton, G. R.; Pandelia, M.-E.; Elliott, S. J.; Latham, J. A. Spectroscopic and Electrochemical Characterization of the Mycofactocin Biosynthetic Protein, MftC, Provides Insight into Its Redox Flipping Mechanism. *Biochemistry* **2019**, *58* (7), 940–950. https://doi.org/10.1021/acs.biochem.8b01082.

(86) Ayikpoe, R. S. New Insights into Mycofactocin Biosynthesis, Structure and Function, 2019.

(87) Bruender, N. A.; Bandarian, V. The Creatininase Homolog MftE from Mycobacterium Smegmatis Catalyzes a Peptide Cleavage Reaction in the Biosynthesis of a Novel Ribosomally Synthesized Post-Translationally Modified Peptide (RiPP). *J. Biol. Chem.* **2017**, *292* (10), 4371–4381. https://doi.org/10.1074/jbc.M116.762062.

(88) Ayikpoe, R. S.; Latham, J. A. 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.* **2019**, *141* (34), 13582–13591. https://doi.org/10.1021/jacs.9b06102.

(89) Ellerhorst, M.; Barth, S. A.; Graça, A. P.; Al-Jammal, W. K.; Peña-Ortiz, L.; Vilotijevic, I.; Lackner, G. S -Adenosylmethionine (SAM)-Dependent

Methyltransferase MftM Is Responsible for Methylation of the Redox Cofactor Mycofactocin. *ACS Chem. Biol.* **2022**, *17* (11), 3207–3217. https://doi.org/10.1021/acschembio.2c00659.

(90)   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.; Edwards, T. E. Mycofactocin-Associated Mycobacterial Dehydrogenases with Non-Exchangeable NAD Cofactors. *Sci. Rep.* **2017**, *7* (1), 41074. https://doi.org/10.1038/srep41074.

(91)   Krishnamoorthy, G.; Kaiser, P.; Lozza, L.; Hahnke, K.; Mollenkopf, H.-J.; Kaufmann, S. H. E. Mycofactocin Is Associated with Ethanol Metabolism in Mycobacteria. *MBio* **2019**, *10* (3). https://doi.org/10.1128/mBio.00190-19.

(92)   Dubey, A. A.; Jain, V. Mycofactocin Is Essential for the Establishment of Methylotrophy in Mycobacterium Smegmatis. *Biochem. Biophys. Res. Commun.* **2019**, *516* (4), 1073–1077. https://doi.org/10.1016/j.bbrc.2019.07.008.

(93)   Krishnamoorthy, G.; Kaiser, P.; Constant, P.; Abu Abed, U.; Schmid, M.; Frese, C. K.; Brinkmann, V.; Daffé, M.; Kaufmann, S. H. E. Role of Premycofactocin Synthase in Growth, Microaerophilic Adaptation, and Metabolism of Mycobacterium Tuberculosis. *MBio* **2021**, *12* (4). https://doi.org/10.1128/mBio.01665-21.

(94)   Kostenko, A.; Lien, Y.; Mendauletova, A.; Ngendahimana, T.; Novitskiy, I. M.; Eaton, S. S.; Latham, J. A. Identification of a Poly-Cyclopropylglycine–Containing Peptide via Bioinformatic Mapping of Radical S-Adenosylmethionine

Enzymes. *J. Biol. Chem.* **2022**, *298* (5), 101881.

https://doi.org/10.1016/j.jbc.2022.101881.

(95) Zallot, R.; Oberg, N.; Gerlt, J. A. The EFI Web Resource for Genomic

Enzymology Tools: Leveraging Protein, Genome, and Metagenome Databases to

Discover Novel Enzymes and Metabolic Pathways. *Biochemistry* **2019**, *58* (41),

4169–4182. https://doi.org/10.1021/acs.biochem.9b00735.

(96) Lien, Y. Characterization of Cyclopropyl Synthases Involved in Maturation of

Ribosomally Synthesized and Posttranslationally Modified Peptides., 2022.

(97) Cuthbertson, L.; Nodwell, J. R. The TetR Family of Regulators. *Microbiol. Mol.*

Biol. Rev. **2013**, *77* (3), 440–475. https://doi.org/10.1128/MMBR.00018-13.

(98) 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.; Sherman, D. R. The DNA-Binding Network of

Mycobacterium Tuberculosi S. *Nat. Commun.* **2015**, *6* (1), 5829.

https://doi.org/10.1038/ncomms6829.

(99) Grkovic, S.; Brown, M. H.; Roberts, N. J.; Paulsen, I. T.; Skurray, R. a. QacR Is a

Repressor Protein That Regulates Expression of the Staphylococcus Aureus

Multidrug Efflux Pump QacA. *J. Biol. Chem.* **1998**, *273* (29), 18665–18673.

https://doi.org/10.1074/jbc.273.29.18665.

(100) Uguru, G. C.; Stephens, K. E.; Stead, J. A.; Towle, J. E.; Baumberg, S.;

McDowall, K. J. Transcriptional Activation of the Pathway-Specific Regulator of

the Actinorhodin Biosynthetic Genes in Streptomyces Coelicolor. *Mol. Microbiol.*

2005, *58* (1), 131–150. https://doi.org/10.1111/j.1365-2958.2005.04817.x.

(101) Krug, A.; Wendisch, V. F.; Bott, M. Identification of AcnR, a TetR-Type
Repressor of the Aconitase Gene Acn in Corynebacterium Glutamicum. *J. Biol.
Chem.* **2005**, *280* (1), 585–595. https://doi.org/10.1074/jbc.M408271200.

(102) Croxatto, A.; Chalker, V. J.; Lauritz, J.; Jass, J.; Hardman, A.; Williams, P.;
Cámara, M.; Milton, D. L. VanT, a Homologue of Vibrio Harveyi LuxR,
Regulates Serine, Metalloprotease, Pigment, and Biofilm Production in Vibrio
Anguillarum. *J. Bacteriol.* **2002**, *184* (6), 1617–1629.
https://doi.org/10.1128/JB.184.6.1617-1629.2002.

(103) Pompeani, A. J.; Irgon, J. J.; Berger, M. F.; Bulyk, M. L.; Wingreen, N. S.;
Bassler, B. L. The Vibrio Harveyi Master Quorum-Sensing Regulator, LuxR, a
TetR-Type Protein Is Both an Activator and a Repressor: DNA Recognition and
Binding Specificity at Target Promoters. *Mol. Microbiol.* **2008**, *70* (1), 76–88.
https://doi.org/10.1111/j.1365-2958.2008.06389.x.

(104) Hinrichs, W.; Kisker, C.; Duvel, M.; Muller, A.; Tovar, K.; Hillen, W.; Saenger,
W. Structure of the Tet Repressor-Tetracycline Complex and Regulation of
Antibiotic Resistance. *Science* **1994**, *264* (5157), 418–420.

(105) Orth, P.; Hillen, W.; Saenger, W.; Hinrichs, W. Structural Basis of Gene
Regulation by the Tetracycline Inducible TetR Repressor-Operator System. *Nat.
Struct. Biol.* **2000**, *7* (3), 215–219. https://doi.org/10.1038/73324.

(106) Rodríguez-García, A.; Combes, P.; Pérez-Redondo, R.; Smith, M. C. A. M. C. M.;
Smith, M. C. A. M. C. M. Natural and Synthetic Tetracycline-Inducible Promoters

for Use in the Antibiotic-Producing Bacteria Streptomyces. *Nucleic Acids Res. Rodríguez-García, A., Combes, P., Pérez-Redondo, R., Smith, M. C. A. M. C. M. Smith, M. C. A. M. C. M. Nat. Synth. tetracycline-inducible Promot. use Antibiot. Bact. Streptomyces. Nucleic Aci* **2005**, *33* (9), 1–8. https://doi.org/10.1093/nar/gni086.

(107) Brune, I.; Götker, S.; Schneider, J.; Rodionov, D. A.; Tauch, A. Negative Transcriptional Control of Biotin Metabolism Genes by the TetR-Type Regulator BioQ in Biotin-Auxotrophic Corynebacterium Glutamicum ATCC 13032. *J. Biotechnol.* **2012**, *159* (3), 225–234. https://doi.org/10.1016/j.jbiotec.2011.12.001.

(108) Anand, S.; Singh, V.; Singh, A. K.; Mittal, M.; Datt, M.; Subramani, B.; Kumaran, S. Equilibrium Binding and Kinetic Characterization of Putative Tetracycline Repressor Family Transcription Regulator Fad35R from Mycobacterium Tuberculosis. *FEBS J.* **2012**, *279* (17), 3214–3228. https://doi.org/10.1111/j.1742-4658.2012.08707.x.

(109) HIROOKA, K.; FUJITA, Y. Identification of Aromatic Residues Critical to the DNA Binding and Ligand Response of the *Bacillus Subtilis* QdoR (YxaF) Repressor Antagonized by Flavonoids. *Biosci. Biotechnol. Biochem.* **2011**, *75* (7), 1325–1334. https://doi.org/10.1271/bbb.110098.

(110) Onaka, H.; Ando, N.; Nihira, T.; Yamada, Y.; Beppu, T.; Horinouchi, S. Cloning and Characterization of the A-Factor Receptor Gene from Streptomyces Griseus. These Include : Cloning and Characterization of the A-Factor Receptor Gene from Streptomyces Griseus. **1995**, *177* (21), 6083–6092.

(111) Ahn, S. K.; Cuthbertson, L.; Nodwell, J. R. Genome Context as a Predictive Tool for Identifying Regulatory Targets of the TetR Family Transcriptional Regulators. *PLoS One* **2012**, *7* (11), e50562. https://doi.org/10.1371/journal.pone.0050562.

(112) Peterson, E. J.; Bailo, R.; Rothchild, A. C.; Arrieta-Ortiz, M. L.; Kaur, A.; Pan, M.; Mai, D.; Abidi, A. A.; Cooper, C.; Aderem, A.; Bhatt, A.; Baliga, N. S. Path-seq Identifies an Essential Mycolate Remodeling Program for Mycobacterial Host Adaptation. *Mol. Syst. Biol.* **2019**, *15* (3), 1–19. https://doi.org/10.15252/msb.20188584.

(113) *Global Tuberculosis Report*; 2022.

(114) Li, K.; Schurig-Briccio, L. A.; Feng, X.; Upadhyay, A.; Pujari, V.; Lechartier, B.; Fontes, F. L.; Yang, H.; Rao, G.; Zhu, W.; Gulati, A.; No, J. H.; Cintra, G.; Bogue, S.; Liu, Y.-L.; Molohon, K.; Orlean, P.; Mitchell, D. A.; Freitas-Junior, L.; Ren, F.; Sun, H.; Jiang, T.; Li, Y.; Guo, R.-T.; Cole, S. T.; Gennis, R. B.; Crick, D. C.; Oldfield, E. Multitarget Drug Discovery for Tuberculosis and Other Infectious Diseases. *J. Med. Chem.* **2014**, *57* (7), 3126–3139. https://doi.org/10.1021/jm500131s.

(115) Li, W.; Upadhyay, A.; Fontes, F. L.; North, E. J.; Wang, Y.; Crans, D. C.; Grzegorzewicz, A. E.; Jones, V.; Franzblau, S. G.; Lee, R. E.; Crick, D. C.; Jackson, M. Novel Insights into the Mechanism of Inhibition of MmpL3, a Target of Multiple Pharmacophores in Mycobacterium Tuberculosis. *Antimicrob. Agents Chemother.* **2014**, *58* (11), 6413–6423. https://doi.org/10.1128/AAC.03229-14.

(116) Seung, K. J.; Keshavjee, S.; Rich, M. L. Multidrug-Resistant Tuberculosis and

Extensively Drug-Resistant Tuberculosis. *Cold Spring Harb. Perspect. Med.* **2015**, *5* (9), a017863. https://doi.org/10.1101/cshperspect.a017863.

(117) 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.; Chen, F. Pan-Genomic Study of Mycobacterium Tuberculosis Reflecting the Primary/Secondary Genes, Generality/Individuality, and the Interconversion Through Copy Number Variations. *Front. Microbiol.* **2018**, *9*. https://doi.org/10.3389/fmicb.2018.01886.

(118) Cook, G. M.; Hards, K.; Vilchèze, C.; Hartman, T.; Berney, M. Energetics of Respiration and Oxidative Phosphorylation in Mycobacteria. *Microbiol. Spectr.* **2014**, *2* (3). https://doi.org/10.1128/microbiolspec.MGM2-0015-2013.

(119) Billig, S.; Schneefeld, M.; Huber, C.; Grassl, G. A.; Eisenreich, W.; Bange, F.-C. Lactate Oxidation Facilitates Growth of Mycobacterium Tuberculosis in Human Macrophages. *Sci. Rep.* **2017**, *7* (1), 6484. https://doi.org/10.1038/s41598-017-05916-7.

(120) Serafini, A.; Tan, L.; Horswell, S.; Howell, S.; Greenwood, D. J.; Hunt, D. M.; Phan, M.; Schembri, M.; Monteleone, M.; Montague, C. R.; Britton, W.; Garza-Garcia, A.; Snijders, A. P.; VanderVen, B.; Gutierrez, M. G.; West, N. P.; Carvalho, L. P. S. Mycobacterium Tuberculosis Requires Glyoxylate Shunt and Reverse Methylcitrate Cycle for Lactate and Pyruvate Metabolism. *Mol. Microbiol.* **2019**, *112* (4), 1284–1307. https://doi.org/10.1111/mmi.14362.

(121) Mendauletova, A.; Latham, J. A. Biosynthesis of the Redox Cofactor Mycofactocin Is Controlled by the Transcriptional Regulator MftR and Induced by

Long Chain Acyl-CoA Species. *J. Biol. Chem.* **2021**, 101474.
https://doi.org/10.1016/j.jbc.2021.101474.

(122) Wang, Y.; Cen, X. F.; Zhao, G. P.; Wang, J. Characterization of a New GlnR
Binding Box in the Promoter of AmtB in Streptomyces Coelicolor Inferred a
PhoP/GlnR Competitive Binding Mechanism for Transcriptional Regulation of
AmtB. *J. Bacteriol.* **2012**, *194* (19), 5237–5244. https://doi.org/10.1128/JB.00989-
12.

(123) Peñuelas-Urquides, K.; Villarreal-Treviño, L.; Silva-Ramírez, B.; Rivadeneyra-
Espinoza, L.; Said-Fernández, S.; de León, M. B. Measuring of Mycobacterium
Tuberculosis Growth. A Correlation of the Optical Measurements with Colony
Forming Units. *Brazilian J. Microbiol.* **2013**, *44* (1), 287–290.
https://doi.org/10.1590/S1517-83822013000100042.

(124) Crooks, G. E.; Hon, G.; Chandonia, J.-M.; Brenner, S. E. WebLogo: A Sequence
Logo Generator: Figure 1. *Genome Res.* **2004**, *14* (6), 1188–1190.
https://doi.org/10.1101/gr.849004.

(125) Pandey, A. K.; Sassetti, C. M. Mycobacterial Persistence Requires the Utilization
of Host Cholesterol. *Proc. Natl. Acad. Sci.* **2008**, *105* (11), 4376–4380.
https://doi.org/10.1073/pnas.0711159105.

(126) Cuthbertson, L.; Nodwell, J. R. The TetR Family of Regulators. *Microbiol. Mol.
Biol. Rev.* **2013**, *77* (3), 440–475. https://doi.org/10.1128/mmbr.00018-13.

(127) Lara, J.; Diacovich, L.; Trajtenberg, F.; Larrieux, N.; Malchiodi, E. L.; Fernández,
M. M.; Gago, G.; Gramajo, H.; Buschiazzo, A. Mycobacterium Tuberculosis FasR

Senses Long Fatty Acyl-CoA through a Tunnel and a Hydrophobic Transmission Spine. *Nat. Commun.* **2020**, *11* (1), 1–13. https://doi.org/10.1038/s41467-020-17504-x.

(128) Biosynthesis, L. Crossm TetR-Type Regulator SLCG _ 2919 Is a Negative Regulator Of. *Proc. Natl. Acad. Sci.* **2018**, *118* (December), 1–14. https://doi.org/10.1073/pnas.2019305118.

(129) Grosdidier, A.; Zoete, V.; Michielin, O. SwissDock, a Protein-Small Molecule Docking Web Service Based on EADock DSS. *Nucleic Acids Res.* **2011**, *39* (suppl), W270–W277. https://doi.org/10.1093/nar/gkr366.

(130) Grosdidier, A.; Zoete, V.; Michielin, O. Fast Docking Using the CHARMM Force Field with EADock DSS. *J. Comput. Chem.* **2011**, *32* (10), 2149–2159. https://doi.org/10.1002/jcc.21797.

(131) Le, T. B. K.; Schumacher, M. A.; Lawson, D. M.; Brennan, R. G.; Buttner, M. J. The Crystal Structure of the TetR Family Transcriptional Repressor SimR Bound to DNA and the Role of a Flexible N-Terminal Extension in Minor Groove Binding. *Nucleic Acids Res.* **2011**, *39* (21), 9433–9447. https://doi.org/10.1093/nar/gkr640.

(132) Bhukya, H.; Jana, A. K.; Sengupta, N.; Anand, R. Structural and Dynamics Studies of the TetR Family Protein, CprB from Streptomyces Coelicolor in Complex with Its Biological Operator Sequence. *J. Struct. Biol.* **2017**, *198* (2), 134–146. https://doi.org/10.1016/j.jsb.2017.03.006.

(133) Palanca, C.; Rubio, V. Structure of AmtR, the Global Nitrogen Regulator of

Corynebacterium Glutamicum , in Free and DNA-Bound Forms. *FEBS J.* **2016**, *283* (6), 1039–1059. https://doi.org/10.1111/febs.13643.

(134) Yang, S.; Gao, Z.; Li, T.; Yang, M.; Zhang, T.; Dong, Y.; He, Z.-G. Structural Basis for Interaction between Mycobacterium Smegmatis Ms6564, a TetR Family Master Regulator, and Its Target DNA. *J. Biol. Chem.* **2013**, *288* (33), 23687–23695. https://doi.org/10.1074/jbc.M113.468694.

(135) Carroll, P.; Schreuder, L. J.; Muwanguzi-Karugaba, J.; Wiles, S.; Robertson, B. D.; Ripoll, J.; Ward, T. H.; Bancroft, G. J.; Schaible, U. E.; Parish, T. Sensitive Detection of Gene Expression in Mycobacteria under Replicating and Non-Replicating Conditions Using Optimized Far-Red Reporters. *PLoS One* **2010**, *5* (3), e9823. https://doi.org/10.1371/journal.pone.0009823.

(136) Yao, J.; Rock, C. O. Exogenous Fatty Acid Metabolism in Bacteria. *Biochimie* **2017**, *141*, 30–39. https://doi.org/10.1016/j.biochi.2017.06.015.

(137) Korepanova, A.; Gao, F. P.; Hua, Y.; Qin, H.; Nakamoto, R. K.; Cross, T. A. Cloning and Expression of Multiple Integral Membrane Proteins from Mycobacterium Tuberculosis in Escherichia Coli. *Protein Sci.* **2004**, *14* (1), 148–158. https://doi.org/10.1110/ps.041022305.

(138) Rutledge, H. L.; Field, M. J.; Rittle, J.; Green, M. T.; Tezcan, F. A. Role of Serine Coordination in the Structural and Functional Protection of the Nitrogenase P-Cluster. *J. Am. Chem. Soc.* **2022**, *144* (48), 22101–22112. https://doi.org/10.1021/jacs.2c09480.

(139) Grove, A. Urate-Responsive MarR Homologs from Burkholderia. *Mol. Biosyst.*

2010, *6* (11), 2133. https://doi.org/10.1039/c0mb00086h.

(140)   Gupta, A.; Grove, A. Ligand-Binding Pocket Bridges DNA-Binding and

Dimerization Domains of the Urate-Responsive MarR Homologue MftR from

Burkholderia Thailandensis. *Biochemistry* **2014**, *53* (27), 4368–4380.

https://doi.org/10.1021/bi500219t.

(141)   Peterson, E. J.; Bailo, R.; Rothchild, A. C.; Arrieta-Ortiz, M. L.; Kaur, A.; Pan,

M.; Mai, D.; Abidi, A. A.; Cooper, C.; Aderem, A.; Bhatt, A.; Baliga, N. S. Path-

seq Identifies an Essential Mycolate Remodeling Program for Mycobacterial Host

Adaptation. *Mol. Syst. Biol.* **2019**, *15* (3). https://doi.org/10.15252/msb.20188584.

(142)   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.; Sherman, D. R. The DNA-Binding Network of

Mycobacterium Tuberculosi S. *Nat. Commun.* **2015**, *6* (1), 5829.

https://doi.org/10.1038/ncomms6829.

(143)   Turkarslan, S.; Peterson, E. J. R.; Rustad, T. R.; Minch, K. J.; Reiss, D. J.;

Morrison, R.; Ma, S.; Price, N. D.; Sherman, D. R.; Baliga, N. S. A

Comprehensive Map of Genome-Wide Gene Regulation in Mycobacterium

Tuberculosis. *Sci. Data* **2015**, *2* (1), 150010. https://doi.org/10.1038/sdata.2015.10.

(144)   Singh, A.; Gupta, R.; Vishwakarma, R. A.; Narayanan, P. R.; Paramasivan, C. N.;

Ramanathan, V. D.; Tyagi, A. K. Requirement of the MymA Operon for

Appropriate Cell Wall Ultrastructure and Persistence of Mycobacterium

Tuberculosis in the Spleens of Guinea Pigs. *J. Bacteriol.* **2005**, *187* (12), 4173–

4186. https://doi.org/10.1128/JB.187.12.4173-4186.2005.

(145) Singh, A.; Jain, S.; Gupta, S.; Das, T.; Tyagi, A. K. MymA Operon of
Mycobacterium Tuberculosis: Its Regulation and Importance in the Cell Envelope.
FEMS Microbiol. Lett. **2003**, *227* (1), 53–63. https://doi.org/10.1016/S0378-
1097(03)00648-7.

(146) Dubey, A. A.; Jain, V. Mycofactocin Is Essential for the Establishment of
Methylotrophy in Mycobacterium Smegmatis. *Biochem. Biophys. Res. Commun.*
2019, *516* (4), 1073–1077. https://doi.org/10.1016/j.bbrc.2019.07.008.

(147) Milano, A.; Pasca, M. R.; Provvedi, R.; Lucarelli, A. P.; Manina, G.; Luisa de
Jesus Lopes Ribeiro, A.; Manganelli, R.; Riccardi, G. Azole Resistance in
Mycobacterium Tuberculosis Is Mediated by the MmpS5–MmpL5 Efflux System.
Tuberculosis **2009**, *89* (1), 84–90. https://doi.org/10.1016/j.tube.2008.08.003.

(148) 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.; Sherman, D. R. The DNA-Binding Network of
Mycobacterium Tuberculosis. *Nat. Commun.* **2015**, *6*, 1–10.
https://doi.org/10.1038/ncomms6829.

Appendix A: Supplementary information

DNA sequences

Pmft Sequence

TACGCCTGCAGCTCTTCGGTTTCCAGGATCACGCGCATACGCTGCCGGTG
CTCGGCCATCTCGTGGTCGGCGTAGGTGTTGAAGGTCAGCAGTGCCTCGCGC
AGGGCGTCGCCGAGTGACACCTCGGAGCTGAGGGTCCTCAGCAGGTTCTGCA
GGTGCTGCAGGTGTGAGTCGAAATCACCCCACGGGATGGCGCTCTTGGAGGC
GTAGTACCGGAACAGGGTCCGTCGCGAGATGCCCGCGGCGGCCGCGACGTCG
TCGACGCTGACCGCGTCGAACCCTCGGGCCGCGAACAGATCGATCGCCACGC
CGGCGATGTGGTCCTGCGTGGTGGAGCGGCGACGCCCCGCGCGGGAACCCTC
GGACATCTCTCACACCCCCTCTTCCATTCTGGCACTCGATGCCATATATTTGC
GATCTCGATCACAACTGTCGAGACCATACGCGACGAAAGGGAGTCCACATGG
AACCGAATCAGCACGTCGAGGCCGAGACCGAACTCGTCACCGAGACTCTCGT
GGAAGAGGTCTCGATCGACGGTATGTGCGGGGTCTAC

Pmft-pCherry

TCTAGATACGCCTGCAGCTCTTCGGTTTCCAGGATCACGCGCATACGCTGCCG
GTGCTCGGCCATCTCGTGGTCGGCGTAGGTGTTGAAGGTCAGCAGTGCCTCG
CGCAGGGCGTCGCCGAGTGACACCTCGGAGCTGAGGGTCCTCAGCAGGTTCT
GCAGGTGCTGCAGGTGTGAGTCGAAATCACCCCACGGGATGGCGCTCTTGGA
GGCGTAGTACCGGAACAGGGTCCGTCGCGAGATGCCCGCGGCGGCCGCGAC
GTCGTCGACGCTGACCGCGTCGAACCCTCGGGCCGCGAACAGATCGATCGCC
ACGCCGGCGATGTGGTCCTGCGTGGTGGAGCGGCGACGCCCCGCGCGGGAAC

CCTCGGACATCTCTCACACCCCCTCTTCCATTCTGGCACTCGATGCCATATAT

TTGCGATCTCGATCACAACTGTCGAGACCATACGCGACGAAAGGGAGTCCAC

ATGGAACCGAATCAGCACGTCGAGGCCGAGACCGAACTCGTCACCGAGACTC

TCGTGGAAGAGGTCTCGATCGACGGTATGTGCGGGGTCTACGGATCCATGGT

CTCGAAGGGCGAGGAGGACAACATGGCGATCATCAAGGAGTTCATGCGCTTC

AAGGTCCACATGGAGGGCTCGGTCAACGGCCACGAGTTCGAGATCGAGGGC

GAGGGCGAGGGCCGCCCGTACGAGGGCACCCAGACCGCCAAGCTGAAGGTC

ACCAAGGGCGGCCCGCTGCCGTTCGCCTGGGACATCCTGTCGCCGCAGTTCA

TGTACGGCAGCAAGGCCTACGTCAAGCACCCGGCCGACATCCCGGACTACCT

GAAGCTGTCGTTCCCGGAGGGCTTCAAGTGGGAGCGCGTCATGAACTTCGAG

GACGGCGGCGTCGTCACCGTCACCCAGGACTCGTCGCTGCAGGACGGCGAGT

TCATCTACAAGGTCAAGCTGCGGGGCACCAACTTCCCGTCGGACGGCCCGGT

CATGCAGAAGAAGACCATGGGCTGGGAGGCCTCGTCGGAGCGCATGTACCC

GGAGGACGGCGCCCTGAAGGGCGAGATCAAGCAGCGGCTGAAGCTGAAGGA

CGGCGGCCACTACGACGCCGAGGTCAAGACCACCTACAAGGCCAAGAAGCC

GGTCCAGCTGCCGGGCGCCTACAACGTGAACATCAAGCTGGACATCACCAGC

CACAACGAGGACTACACCATCGTCGAGCAGTACGAGCGCGCCGAGGGCCGC

CACAGCACCGGCGGCATGGACGAGCTGTACAAGTGAAAGCTTATCGATACCG

TCGACCTCGAGGGGGGGCCCGGTACGTACCCGGGGATCATCGAGCCGAGAA

CGTTATCGAAGTTGGTCATGTGTAATCCCCTCGTTTGAACTTTGGATTAAGCG

TAGATACACCCTTGGACAAGCCAGTTGGATTCGGAGACAAGCAAATTCAGCC

TTAAAAAGGGCGAGGCCTGCGGTGGTGGAACACCGCAGGGCCTCTAACCGCT

123

CGACGCGCTGCACCAACCAGCCCGCGAACGGCTGGCAGCCAGCGTAAGGCG

CGGCTCATCGGGCGGCGTTCGCCACGATGTCCTGCACTTCGAGCCAAGCCTC

GAACACCTGCTGGTGTGCACGACTCACCCGGTTGTTGACACCGCGCGCGGCC

GTGCGGGCTCGGTGGGGCGGCTGTGTCGCCCTTGCCAGCGTGAGTAGCGCGT

ACCTCACCTCGCCCAACAGGTCGCACACAGCCGATTCGTACGCCATAAAGCC

AGGTGAGCCCACCAGCTCCGTAAGTTCGGGCGCTGTGTGGCTCGTACCCGCG

CATTCAGGCGGCAGGGGGTCTAACGGGTCTAAGGCGGCGTGTACGGCCGCCA

CAGCGGCTCTCAGCGGCCCGGAAACGTCCTCGAAACGACGCATGTGTTCCTC

CTGGTTGGTACAGGTGGTTGGGGGTGCTCGGCTGTCGCTGGTGTTCCACCACC

AGGGCTCGACGGGAGAGCGGGGGAGTGTGCAGTTGTGGGGTGGCCCCTCAG

CGAAATATCTGACTTGGAGCTCGTGTCGGACCATACACCGGTGATTAATCGT

GGTCTACTACCAAGCGTGAGCCACGTCGCCGACGAATTTGAGCAGCTCTGGC

TGCCGTACTGGCCGCTGGCAAGCGACGATCTGCTCGAGGGGATCTACCGCCA

AAGCCGCGCGTCGGCCCTAGGCCGCCGGTACATCGAGGCGAACCCAACAGC

GCTGGCAAACCTGCTGGTCGTGGACGTAGACCATCCAGACGCAGCGCTCCGA

GCGCTCAGCGCCCGGGGGTCCCATCCGCTGCCCAACGCGATCGTGGGCAATC

GCGCCAACGGCCACGCACACGCAGTGTGGGCACTCAACGCCCCTGTTCCACG

CACCGAATACGCGCGGCGTAAGCCGCTCGCATACATGGCGGCGTGCGCCGAA

GGCCTTCGGCGCGCCGTCGACGGCGACCGCAGTTACTCAGGCCTCATGACCA

AAAACCCCGGCCACATCGCCTGGGAAACGGAATGGCTCCACTCAGATCTCTA

CACACTCAGCCACATCGAGGCCGAGCTCGGCGCGAACATGCCACCGCCGCGC

TGGCGTCAGCAGACCACGTACAAAGCGGCTCCGACGCCGCTAGGGCGGAATT

GCGCACTGTTCGATTCCGTCAGGTTGTGGGCCTATCGTCCCGCCCTCATGCGG

ATCTACCTGCCGACCCGGAACGTGGACGGACTCGGCCGCGCGATCTATGCCG

AGTGCCACGCGCGAAACGCCGAATTCCCGTGCAACGACGTGTGTCCCGGACC

GCTACCGGACAGCGAGGTCCGCGCCATCGCCAACAGCATTTGGCGTTGGATC

ACAACCAAGTCGCGCATTTGGGCGGACGGGATCGTGGTCTACGAGGCCACAC

TCAGTGCGCGCCAGTCGGCCATCTCGCGGAAGGGAGCAGCGCGCACGGCGG

CGAGCACAGTTGCGCGGCGCGCAAAGTCCGCGTCAGCCATGGAGGCATTGCT

ATGAGCGACGGCTACAGCGACGGCTACAGCGACGGCTACAACCGGCAGCCG

ACTGTCCGCAAAAAGCGGCGCGTGACCGCCGCCGAAGGCGCTCGAATCACCG

GACTATCCGAACGCCACGTCGTCCGGCTCGTGGCGCAGGAACGCAGCGAGTG

GCTCGCCGAGCAGGCTGCACGCCGCGAACGCATCCGCGCCTATCACGACGAC

GAGGGCCACTCTTGGCCGCAAACGGCCAAACATTTCGGGCTGCATCTGGACA

CCGTTAAGCGACTCGGCTATCGGGCGAGGAAAGAGCGTGCGGCAGAACAGG

AAGCGGCTCAAAAGGCCCACAACGAAGCCGACAATCCACCGCTGTTCTAACG

CAATTGGGGAGCGGGTGTCGCGGGGGTTCCGTGGGGGGTTCCGTTGCAACGG

GTCGGACAGGTAAAAGTCCTGGTAGACGCTAGTTTTCTGGTTTGGGCCATGCC

TGTCTCGTTGCGTGTTTCGTTGCGCCCGTTTTGAATACCAGCCAGACGAGACG

GGGTTCTACGAATCTTGGTCGATACCAAGCCATTTCCGCTGAATATCGGGGA

GCTCACCGCCAGAATCGGTGGTTGTGGTGATGTACGTGGCGAACTCCGTTGT

AGTGCCTGTGGTGGCATCCGTGGCCACTCTCGTTGCACGGTTCGTTGTGCCGT

TACAGGCCCCGTTGACAGCTCACCGAACGTAGTTAAAACATGCTGGTCAAAC

TAGGTTTACCAACGATACGAGTCAGCTCATCTAGGGCCAGTTCTAGGCGTTGT

TCGTTGCGCGGTTCGTTGCGCATGTTTCGTGTGGTTGCTAGATGGCTCCGCAA

CCACACGCTTCGAGGTTGAGTGCTTCCAGCACGGGCGCGATCCAGAAGAACT

TCGTCGTGCGACTGTCCTCGTTGATCCTTGCCGAGCTGGGATGGAAGCTCGGC

CGACCACCCTGGAGGAGATGATCGAGGATGCCAGGGCCTTTCACGCCCGCCG

CTGCTGAGCGTCCGCCGCCGGGCCCGCACCGCCGTCGGCCGGCCCGCTCCGG

GCTCGCAGCAGCGGGCTTCGGCGCGGGCCCGGGGCTCCCGAGCGCGGGCGG

GGCTCCGGGCGGCCGCCGGGGGCCGGGGGCGGCGCCGGGCGGCCCGGGGCG

TCAGGCGCCGGGGGCGGTGTCCGGCGGCCCCCAGAGGAACTGCGCCAGTTCC

TCCGGATCGGTGAAGCCGGAGAGATCCAGCGGGGTCTCCTCGAACACCTCGA

AGTCGTGCAGGAAGGTGAAGGCGAGCAGTTCGCGGGCGAAGTCCTCGGTCCG

CTTCCACTGCGCCCCGTCGAGCAGCGCGGCCAGGATCTCGCGGTCGCCCCGG

AAGGCGTTGAGATGCAGTTGCACCAGGCTGTAGCGGGAGTCTCCCGCATAGA

CGTCGGTGAAGTCGACGATCCCGGTGACCTCGGTCGCGGCCAGGTCCACGAA

GATGTTGGTCCCGTGCAGGTCGCCGTGGACGAACCGGGGTTCGCGGCCGGCC

AGCAGCGTGTCCACGTCCGGCAGCCAGTCCTCCAGGCGGTCCAGCAGCCGGG

GCGAGAGGTAGCCCCACCCGCGGTGGTCCTCGACGGTCGCCGCGCGGCGTTC

CCGCAGCAGTTCCGGGAAGACCTCGGAATGGGGGGTGAGCACGGTGTTCCCG

GTCAGCGGCACCCTGTGCAGCCGGCCGAGCACCCGGCCGAGTTCGCGGGCCA

GGGCGAGCAGCGCGTTCCGGTCGGTCGTGCCGTCCATCGCGGACCGCCAGGT

GGTGCCGGTCATCCGGCTCATCACCAGGTAGGGCCACGGCCAGGCTCCGGTG

CCGGGCCGCAGCTCGCCGCGGCCGAGGAGGCGGGGCACCGGCACCGGGGCG

TCCGCCAGGACCGCGTACGCCTCCGACTCCGACGCGAGGCTCTCCGGACCGC

ACCAGTGCTCGCCGAACAGCTTGATCACCGGGTCGGGCTCGCCGACCAGTAC

GGGGTTGGTGCTCTCGCCGGGCACCCGCAGCACCGGCGGCACCGGCAGCCCG

AGCTCCTCCAGGGCTCGGCGGGCCAGCGGCTCCCAGAATTCCTGGTCGTTCC

GCAGGCTCGCGTAGGAATCATCCGAATCAATACGGTCGAGAAGTAACAGGG

ATTCTTGTGTCACAGCGGACCTCTATTCACAGGGTACGGGCCGGCTTAATTCC

GCACGGCCGGTCGCGACACGGCCTGTCCGCACCGCGGATCAGGCGTTGACGA

TGACGGGCTGGTCGGCCACGTCGGGGACGACGGGGAGTCAGGCAACTATGG

ATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTG

GTAACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAAAACTTC

ATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTTTTGATAATCTCATGACC

AAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAA

AGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGC

AAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCT

ACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAAT

ACTGTCCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGC

ACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTG

GCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAA

GGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGA

GCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAG

CGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAG

GGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTA

TCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTG

127

ATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTT

TTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTA

TCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATACCGC

TCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGA

AGAGCGCCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAA

TGCAGCTGGCACGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAAC

GCAATTAATGTGAGTTAGCTCACTCATTAGGCACCCCAGGCTTTACACTTTAT

GCTTCCGGCTCGTATGTTGTGTGGAATTGTGAGCGGATAACAATTTCACACAG

GAAACAGCTATGACCATGATTACCAGATCTGGCTCGCACCGCGGTGGCGGCC

GC

Table S1 primers used in this study

Purpose	Primer	Sequence $(5' - 3')$
EMSAs;	FAM-O$_{mft}$	TCCATTCTGGCACTCGATGCCATATAT
MftR	Q15A	CGCCGCTCCACCACGGCGGACCACATCGCCGGC
mutants,	D16R	TCCACCACGCAGCGACACATCGCCGGC
ITC	D16W	TCCACCACGCAGTGGCACATCGCCGGC
	R29A	GATCTGTTCGCGGCCGCGGGGTTCGACGCGGTC
	D38A	GCGGTCAGCGTCGACGCGGTCGCGGCCGCCGCG
	R48A	GCGGGCATCTCGCGAGCGACCCTGTTCCGGTAC
	F65A	GCAGGTGTGAGTCGGCATCACCCCACGGGATG
	D66A	CCGTGGGGTGATTTCGCGTCACACCTGCAGCAC
	S67A	TGGGGTGATTTCGACGCGCACCTGCAGCACCTG
	S67W	GGTGATTTCGACTGGCACCTGCAGCAC
	H68A	GGTGATTTCGACTCAGCGCTGCAGCACCTGCAG
	F96A	GAGGCACTGCTGACCGCGAACACCTACGCCGAC
	I114A	GCAGCGTATGCGCGTGGCCCTGGAAACCGAAGA
	L120A	CCTGGAAACCGAAGAGGCGCAGGCGTATTCGATG
	L120R	CTGGAAACCGAAGAGCGGCAGGCGTATTCGATG
	M127A	GCGTATTCGATGACCGCGTACGCCGGCTGGCG

Appendix B: List of publications

Mendauletova A., Kostenko A., Lien Y., and Latham J.A. How a subfamily of radical S-adenosylmethionine enzymes became a mainstay of ribosomally synthesized and post-translationally modified peptide discovery. *ACS Bio Med Chem Au* 2022, 2, 1, 53-59

Mendauletova A., Latham J.A. Biosynthesis of the redox cofactor mycofactocin is controlled by the transcriptional regulator MftR and induced by long-chain acyl CoA species. *J. Biol. Chem* 2022, 298, 1, 101474

Kostenko A., Lien Y., **Mendauletova A.,** Ngendahimana T., Novitskiy I.M., Eaton S.S., Latham J.A. Identification of poly-cyclopropylglycine-containing peptide via bioinformatic mapping of radical S-adenosylmethionine enzymes. *J. Biol. Chem* 2022, 298, 5, 101881

Lien Y., Zizola C., **Mendauletova A.,** Ngendamina T., Kostenko A., Eaton S.S., Latham J.A., Grove T. L. A structural, biochemical, and bioinformatic basis for identifying radical-SAM cyclopropyl synthases. *Manuscript in prep.*