University of California, Irvine

Characterization of a Proteinaceous Nanocompartment in Mycobacteria and the Encapsulation of its Cargo Proteins

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

In Biological Sciences

By

Heidi Contreras

Dissertation Committee:
Professor Celia W. Goulding, Chair
Professor Thomas Poulos
Associate Professor Melissa B. Lodoen

2015

DEDICATION

To my parents, Leopoldo and Hortencia, in recognition of their sacrifices, bravery, and intelligence. Your stories inspired courage, so I too can bravely face all challenges in pursuit of my dreams and aspirations.

To my big sister, Gladys, for setting the example of the highest academic standard in the family that is unrivaled by most—you expect the best. We are now peers.

To my beautiful niece, Natalie Newton, for filling my life with magic, color, and hope.

To Adam, for providing me with so much love and beautiful music, and ensuring I make it safely to the other side of the hill.

To *all* of my friends: you too are my family. For everything you do for me and with me.

In Surviving:

I had not a dispute but a disquisition with Dilke, on various subjects; several things dovetailed in my mind, & at once it struck me, what quality went to form a Man of Achievement especially in literature & which Shakespeare possessed so enormously - I mean Negative Capability, that is when man is capable of being in uncertainties, Mysteries, doubts without any irritable reaching after fact & reason.

John Keats
Letter to George and Thomas Keats, Sunday, September 21, 1817

In Learning:

Everything not saved will be lost

Nintendo
"Quit Screen" Message

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I am incredibly grateful to the untiring support and guidance of my committee chair and Principle Invistigator, Dr. Celia W. Goulding—her enthusiasm supplied the energy to continue with this work. She is brilliant and incredibly innovative, encouraging her students to be creative while propelling scientific discovery. Thanks to her, I am now a biochemist.

Thank you very much to Dr. Nicholas P. Chim and Angelina Iniguez for their scientific expertise and friendship. The constant scientific discussions and ideas that arose from such topics will always be valuable to me. Wherever we go, may we continue to sing all the wrong lyrics and tell each other about it.

To my fellow and current lab members, thank you very much for all of the support and excellent memories. A very special thanks to Drs. Cedric P. Owens and Rodrigo Torres for providing excellent scientific/technical advice as well as scholastic advice as well as friendship. I would also like to thank Parker M. Johnson and Bryan Boubion for being great sports and great friends.

I would like to recognize all of Vincent Le's successful efforts in greatly contributing to this project. It was an excellent experience mentoring a brilliant student.

A special thanks to the Molecular Biology and Biochemistry department here at UCI—the wonderful people at the office have gone above and beyond for me. Thank you.

I would like to thank the National Institutes of Health (GM101945, H.C.). I also thank the Advance Light Source (ALS) at Berkeley National Laboratories and Standford Synchrotron Radiation Light source (SSRL) for their help in data collection. Special thanks to: Matthew S. Joens and Dr. James A. Fitzpatrick at the Waitt Advanced Biophotonics Center, Salk Institute for Biological Studies as well as Drs. Michael V. Tullius and Marcus A. Horwitz at the Division of Infectious Diseases, Department of Medicine, School of Medicine, UCLA.

CURRICULUM VITAE

Heidi Contreras

TRAINING AND PROFESSIONAL EXPERIENCE

Education
B.S. Microbiology, Arizona State University, 2007
B.A. English Literature, Arizona State University, 2007
Ph.D. Molecular Biology and Biochemistry, University of California, Irvine, March, 2015
(Advisor: Dr Celia W. Goulding)

Positions Held
Undergraduate Researcher (Dr. Valerie Stout), Arizona State University, Department of Microbiology/School of Life Sciences, 2005

Undergraduate Researcher (Dr. Josephine E. Clark-Curtiss), Arizona State University, Center for Infectious Diseases and Vaccinology/Biodesign Institute, 2006-2007

Post-baccalaureate Research Education Program (PREP) Scholar (Dr. Brenda G. Hogue, Dr. Josephine E. Clark-Curtiss), Arizona State University, Center for Infectious Diseases and Vaccinology/Biodesign Institute, 2007-2009

Graduate Student Researcher, Ph.D. Candidate (Dr. Celia W. Goulding), University of California, Irvine, Department of Molecular Biology and Biochemistry, 2009-present (Expected graduation date of February 2015).

Board of Directors at Cuties for a Cause (501(c)3-certified non-profit organization), Escondidio, CA. July 2010 – April 2013.

SCHOLARLY ACTIVITIES

Awards
2007 Dean's List, Arizona State University, Tempe, AZ
2007-2009 Post-baccalaureate Research Education Program (PREP) for Biomedical Research Scholar, Arizona State University, Tempe, AZ

2012-present Ruth L. Kirschstein (NIH F-31) Pre-Doctoral Fellow

2014 Dr. William F. Holcomb Scholarship

2014 Outstanding Student Poster at asm2014, American Society for Microbiology 114th General Meeting.

2014 Student Travel Award Recipient at asm2014, American Society for Microbiology 114th General Meeting.

2014 Honorable Mention for "Best Presentation" at University of California, Irvine Molecular Biology and Biochemistry Departmental Retreat. Lake Arrowhead, CA.

Previous Research Interests

Characterization of a proteinaceous microcompartment (Encapsulin, Enc) from *Mycobacterium tuberculosis* (Mtb), its cargo proteins, and implications in oxidative stress (present).

Elucidation of the mechanisms involved in the heme uptake pathway of Mtb (present).

Characterization of toxin:antitoxin *relBE*, *relFG*, and *relJK* genes in Mtb (Post-baccalaureate).

Characterization of the molecular mechanisms of plant pathogenesis by the causative agent *Pseudomonas tolasii* (Undergraduate).

Publications

Contreras H, Joens MS, McMath LM, Le VP[1], Tullius MV, Kimmey JM, Bionghi N, Horwitz MA, Fitzpatrick JAJ, Goulding CW. 2014. Characterization of a *Mycobacterium tuberculosis* (Mtb) nanocompartment and its potential cargo proteins. J. Biol. Chem. 289(26):18279-89.

Contreras H, Chim N, Credali N, Goulding CW. 2014. Heme uptake in bacterial pathogens. Curr. Opin. Chem. Biol. 19:34-41. (*Review Article*)

Deriu, E, Pezeshki, M, Edwards, RA, Ochoa, RO, **Contreras H**, Libby S, Fang F, Raffatellu M. 2013. Probiotic and Pathogenic Bacteria Compete for Iron in the Inflamed Intestine. Cell Host Microb. 14:3-4.

Owens CP, Chim N, Graves A, Harmston CA, Iniguez A, **Contreras H**, Liptak A, Goulding CW. 2013. The *Mycobacterium tuberculosis* Secreted Protein, Rv0203, Transfers Heme to Membrane Proteins, Mycobacterial membrane protein Large 3 (MmpL3) and MmpL11. J. Biol. Chem. 88:21714-21728.

McMath, LM, **Contreras H**, CP Owens, Goulding CW. 2013. The structural characterization of bacterioferritin, BfrA, from *Mycobacterium tuberculosis*. J. Porphyrins Phthalocyanines. 17: 229.

Chim N, Owens CP, **Contreras H**, Goulding CW. 2012. Advances in Mycobacterium Tuberculosis Therapeutics Discovery Utilizing Structural Biology. Infect Disord Drug Targets. Epub: 23167715. (*Review Article)*

Korch, S. B., **H. Contreras**, Clark-Curtiss J. E. 2009. Three *Mycobacterium tuberculosis* Rel toxinantitoxin modules inhibit mycobacterial growth and are expressed in infected human macrophages. J. Bacteriol. 191:1618-1630.

Posters and Presentations

Contreras H, Le VP, McMath LM, Goulding CW. Encapsulation of Mycobacterial Enzymes involved in oxidative stress response. West Coast Protein Crystallography Conference. Monterrey, CA (March 2013). Poster.

Contreras H, Goulding CW. Characterization of a mycobacterial nanocompartment and is potential cargo proteins. University of California, Irvine Molecular Biology and Biochemistry Departmental Retreat. Lake Arrowhead, CA (March 2014). Presentation.

Contreras H, Joens MS, McMath LM, Le VP[1], Tullius MV, Kimmey JM, Bionghi N, Horwitz MA, Fitzpatrick JAJ, Goulding CW. Evidence of Enzyme Compartmentalization in Mycobacteria. Gordon Research Seminar. Ventura, CA (January 2014). Poster.

Contreras H, Joens MS, McMath LM, Le VP[1], Tullius MV, Kimmey JM, Bionghi N, Horwitz MA, Fitzpatrick JAJ, Goulding CW. Characterization of Mycobacterial Encapsulation of Potential Cargo Enzymes. American Society of Microbiology 114th General Meeting. Boston, MA (May 2014). Poster.

TEACHING

Teaching Responsibilities

Arizona State University

Bacterial Genetics Laboratory—2008 (One semester)

University of California, Irvine

M112 General Microbiology—2012 (One quarter)
BioSci 98 Biochemistry—2012 (Two quarters)
M143 Human Parasitology—2011 (One quarter)
M118 Experimental Microbiology Lab—2011 (Two quarters)
BioSci 99 Molecular Biology—2011 (Two quarters)

SERVICE

Professional Activities

Member, American Society for Microbiology
Member, The Protein Society
Member, TB Structural Genomics Consortium

ABSTRACT OF THE DISSERTATION

Characterization of a Proteinaceous Nanocompartment in *Mycobacteria* and the Encapsulation of its Cargo Proteins

By Heidi Contreras

Doctor of Philosophy in Biological Sciences

University of California, Irvine

Professor Celia W. Goulding, Chair

Mycobacterium tuberculosis (Mtb), the etiological agent of tuberculosis, can evade host immune response. Peroxidases, ferritins, and folate aldolases are antioxidant enzymes, protecting Mtb from oxidative damage by reactive oxygen species in activated macrophages. Dye-decolorizing peroxidases (DyPs) are encapsulated by a bacterial nanocompartment, encapsulin (Enc), where packaged DyP interacts with Enc via a unique C-terminal extension. Genes encoding for Enc exist in an operon with DyPs. Mtb harbors an encapsulin homolog (Mt-Enc) adjacent to the gene that encodes for Mt-DyP. Our studies indicate that Mt-Enc self-assembles into a 60-subunit icosahedral shell independently of cargo enzymes. Preliminary characterization suggests that Mt-DyP is a heme-peroxidase with a C-terminal extension important for interaction with Mt-Enc—elimination of the C-terminal extension abolishes the interaction.

Mtb iron-storage ferritin protein (Mt-BfrB) and 7,8-dihydroneopterin aldolase (Mt-FolB) have C-terminal tails that also interact with Mt-Enc. Mt-DyP and Mt-BfrB are predicted to be involved in oxidative stress response. Mt-FolB may mitigate oxidative

attack, where Mt-FolB allows 7,8-dihydroneopterin to accumulate acting as an antioxidant. We observed encapsulation of Mt-DyP, Mt-BfrB, and Mt-FolB via co-purification and electron microscopy. Protein encapsulation has not been observed *in vivo* and its purpose is currently unknown, specifically encapsulation of enzymes involved in oxidative stress response. Functional studies of these three enzymes while encapsulated demonstrate that these enzymes retain enzymatic activity: Mt-DyP exhibits peroxidase activity with ABTS as a substrate, Mt-BfrB exhibits ferroxidase activity, and Mt-FolB exhibits aldolase activity; however, Mt-Enc alone does not appear to have enzymatic activity. This is the first study that illustrates mycobacterial encapsulation, providing structural and biochemical clues about mechanisms by which Mtb can detoxify the local environment to ensure long-term survival. We have also begun structural studies, probing Mt-Enc nanocompartment assembly in order to understand formation and utilize Mt-Enc as a therapeutic delivery platform.

In vivo analyses elucidating the role of Mt-Enc, and its involvement in survival during oxidative stress have started. Here, we aim to: i) characterize Mt-Enc; ii) preliminarily characterize DyP as a peroxidase; iii) begin structural analyses of both Mt-Enc and Mt-DyP; and iv) begin studies to determine the role of Mt-Enc and its implications in mycobacterial survival.

CHAPTER 1

INTRODUCTION: AN OVERVIEW OF TUBERCULOSIS, BACTERIAL COMPARTMENTALIZATION, AND A NOVEL FAMILY OF PEROXIDASES

Tuberculosis: impact on global health.

Tuberculosis (TB) is still one of the leading killers among infectious diseases resulting in 1.3 million deaths annually; 25% of TB patients are co-infected with human immunodeficiency virus (HIV) (**Figure 1.1A**) [1]. Multiple drug-resistant strains, HIV co-infection, length and cost of treatment, as well as ineffective vaccines against pulmonary TB exacerbate the prevalence of the disease [3]. The currently available vaccine is effective toward preventing disseminated Mtb when given to neonates and infants.

TB is caused by the slow-growing, pathogenic bacterium, *Mycobacterium tuberculosis* (Mtb). Transmission of the disease is caused by inhalation of aerosolized Mtb contained within droplet nuclei from a patient with an active infection to a healthy individual. Upon infection, the Mtb bacilli are phagocytized by alveolar macrophage [2]. (**Figure 1B**). A well-mounted immune response will promote phagosome-lysosome fusion, where the generation of reactive oxygen species, acidic pH, and an iron-limiting environment destroys the phagocytosed bacilli, thereby containing the spread of infection [4, 5]. However, the human host and Mtb have coexisted for many centuries, evolving an elegant mechanism by which Mtb can evade the host's immune defenses. For example, Mtb can prevent phagosome-lysosome fusion, establishing chronic infection by interfering with expression of late endocytic markers on the phagosome membrane that

kill phagocytosed bacilli [6]. The more permissive environment of the phagosome is conducive for Mtb replication and ultimately dissemination within the human host. Studies have suggested that a robust cellular immune response is required for protection against Mtb infection and disease, hence vaccines or adjuvants that induce production of helper T-cell (TH1) cytokines such as gamma interferon (IFN-γ) or tumor necrosis factor-alpha (TNF-α) from either $CD4^+$ or $CD8^+$ T cells are being studied [4, 7].

Anti-TB treatments currently available, also known as "first line drugs", include isoniazid, rifampicin, pyrazinomide, ethamobutol, and streptomycin; however, these are normally given to those with an active infection and who have not previously been infected [8]. Drug-resistant strains of TB, such as multiple-drug resistant (MDR) and extreme drug resistant (XDR) strains, are resistant to "first line" drugs [9], thus are treated with five groups of anti-TB drugs: group 1 (first line, oral agents), group 2 (injectable agents), group 3 (fluoroquinolones), group 4 (oral, bacteriostatic agents), and group 5 (anti-bacterial agents used for treatment of severe infections)[10]. Since the fight against TB is a battle on two fronts, treatment and prevention, current studies increasing the knowledge of the pathogen's physiology will help develop improved anti-TB therapies. On the other hand, an enormous amount of funding is being redirected to develop an improved TB vaccine, one that has yet to rival the efficacy and availability of the Bacillus Calmette-Guerin (BCG) vaccine.

BCG, developed in the 1920's, is often considered the "gold standard" of TB vaccines, as it is effective in newborns and infants. BCG vaccine is widespread in high TB-risk areas; however, BCG vaccine decreased in popularity as some children infected with HIV could become infected with *Mycobacterium bovis*—an attenuated

mycobacterial strain that is used as the BCG vaccine [11]. Although BCG works as early immunization/prevention against disseminated TB, its efficacy wanes in protection against pulmonary TB in adults [7, 12]. Studies are currently being conducted testing the various vaccine candidates that include recombinant BCG, viral vectors, DNA subunit vaccines, live attenuated bacterial strains, as well as an enhanced adjuvant that can be safe enough to use for those infected with HIV in addition to vaccines used in combination with drug therapy to prevent latent or reactivation of TB [7, 11]. Therefore, the challenge of creating a safe and efficient vaccine is great but necessary to combat this disease. In order to accomplish this, understanding mycobacterial physiology upon infection is important to identify new antigens/vaccine candidates as well as drug therapies.

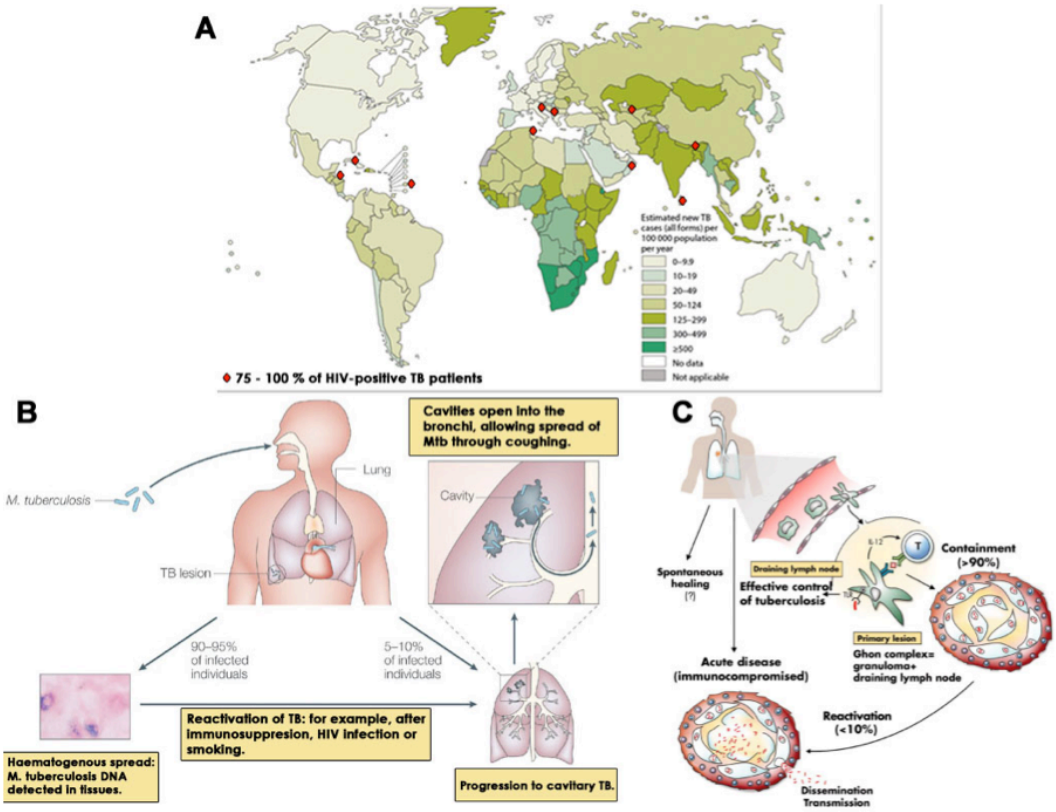

Figure 1.1: **TB-the disease.** A) Incidence map of TB cases reported in 2012/2013. Cases of TB co-infection with HIV are marked with a red diamond [1]; B) TB transmission via droplet nuclei [2] and C) bacilli fate within the human host after phagocytosis by macrophage [5].

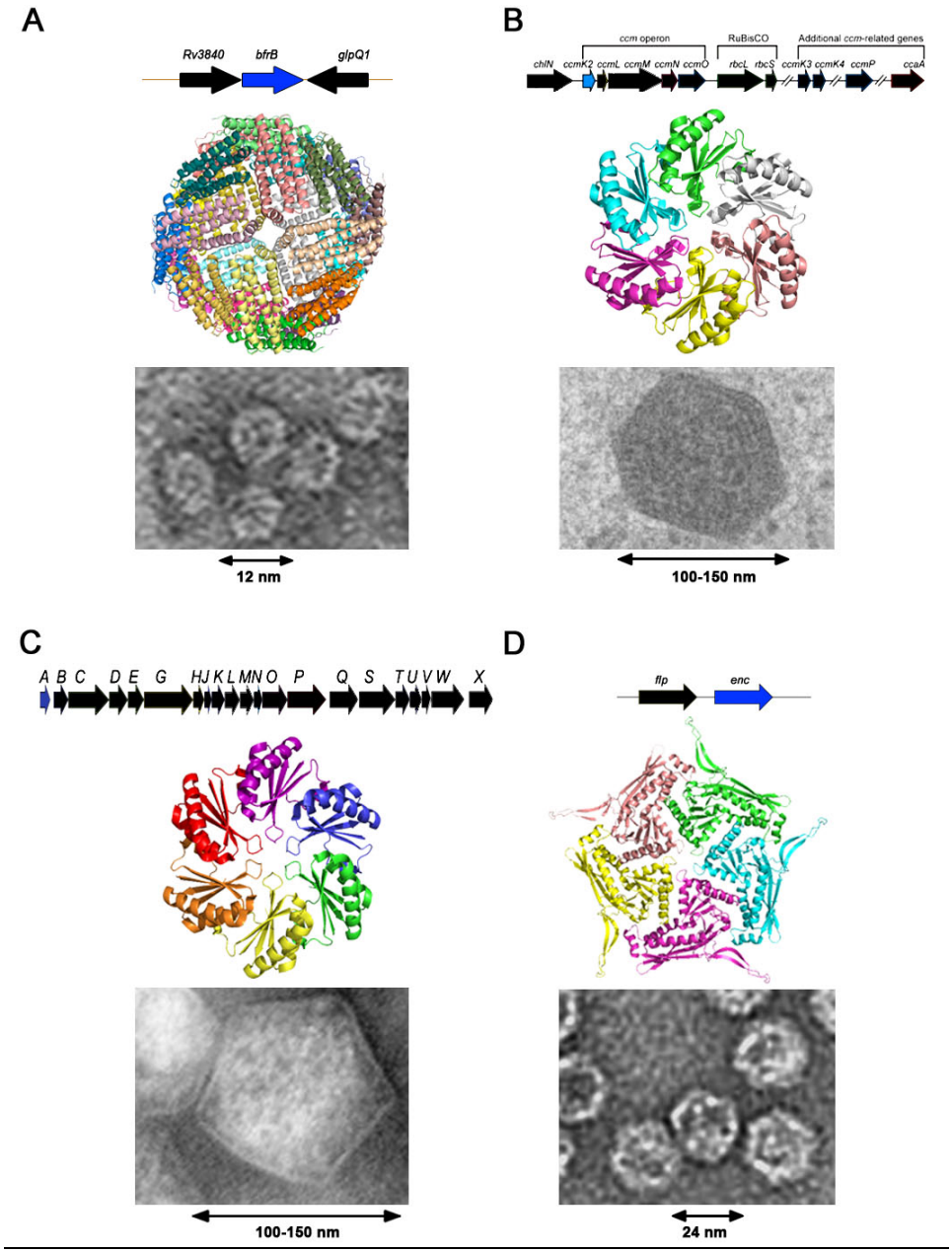

Figure 1.2: **Types of Bacterial Compartments.** Each panel consists of genomic context of shell proteins (top), where the gene colored in blue is shown as cartoon representation of shell protein structure (middle), and EM images of fully assembled compartment (bottom). A) Ferritin-like protein from Mtb, Mt-BfrB, its surrounding genes, and 24-subunit nanocage (PDB: 3QD8) [13]; B) Carboxysome and its genes, structure of CcmK2 (PDB: 2A1B) [14-17]; C) Pdu genes, structure of PduA (PDB: 3NGK) [14, 15, 17, 18] ; D) two-gene operon of Tm-Enc (PDB: 3DKT) and Tm-Flp [19]. Double-sided arrow at the bottom of electron micrograph represents the diameter of the compartment.

Bacterial Compartmentalization

Compartmentalization is important for organization, increasing the local concentration of functionally related enzymes, and confine unstable reaction intermediates against the complexity of the rest of the cell. A very important distinction between eukaryotes and prokaryotes are that prokaryotes lack membrane-bound organelles. Well-known examples of encapsulation are eukaryotic lipid bilayer-enclosed organelles, such as nuclei, mitochondria, or phagosomes. For examples, mitochondria compartments function in important cellular processes such as producing energy and protecting against oxidative stress [20]. Prokaryotes do compartmentalize within their cells [21], usually by self-assembly of proteinaneous cage-like structures [22]. Three examples of bacterial compartmentalization are iron-storing ferritins (**Figure 1.2A**) [23], the formation of proteinacious microcompartments (MCP) [21, 24] composed of proteins containing bacterial microcompartment (BMC) domains (**Figure 1.2B and 1.2C**) [25], and encapsulin (Enc) nanocompartments (**Figure 1.2D**) [19].

Ferritins

Ferritins are a family of iron-storing proteins that are found in animals and bacteria, assembling into 24-subunit nanocages with a diameter of ~120 Å (**Figure 1.2A**) [23]. Major functions among the various types of ferritins include: i) catalysis of ferrous (Fe^{2+}) iron to ferric (Fe^{3+}) iron via ferroxidase activity; ii) storage of Fe^{3+} iron for later release upon iron-starving conditions while sequestering Fe^{2+} from Fenton reactions; and iii) the consumption/detoxification of major cellular oxidants such as oxygen and hydrogen peroxide (H_2O_2) by utilizing iron [23]. Despite low sequence homology (~

15%), the quaternary and tertiary structures of many ferritins are highly similar, containing four-helical bundles assembling into 24-subunits, containing 4-3-2 symmetry (**Figure 1.2A**) [23, 26]. Furthermore, ferritins can accommodate approximately 4000 iron atoms within its cavity [26]. Eukaryotic ferritins have been widely studied; however one example of bacterial ferritin includes BfrB from Mtb (Mt-BfrB) [13]. The structure of Mt-BfrB is very similar to known ferritins; however, this ferritin harbors a C-terminal extension that is important for ferroxidase activity [13]. Of note, 3-fold pores influence Fe^{2+} access to catalytic sites [23] and the C-terminal extension lies along the 3-fold junction of Mt-BfrB [13].

Microcompartments

BMC domains are distributed to approximately 17% of sequenced bacterial genomes [15]. MCPs were first discovered via electron microscopy (EM), first mistaken as phages, but later generically termed polyhedral bodies based on their shape [24]. There are many types of MCPs, but three that have been predominately studied are: carboxysomes, 1,2-propanediol utilization (Pdu), and ethanolamine utilization (Eut) microcompartments [24]. MCPs have been reported to encapsulate several enzymes involved in a metabolic pathway, isolating that pathway from the rest of the cell. Genes encoding enzymes participating in MCP formation are usually found near genes that encode the metabolic pathway that the MCP encapsulates (**Figure 1.2B, top, and 2C, top**) [15]. The advantage of isolating enzymes may include: i) safeguarding cargo enzymes from competing reactions; ii) controlling of substrate access to cargo; iii) protecting unstable intermediates from the oxidative cytosol; or iv) protection of the cytosol from toxic

intermediates [14, 22, 27, 28]. Because of their complex assembly and ability to encapsulate enzymes involved in a metabolic process, some have argued their classification as bacterial organelles [15, 29].

Carboysomes, present in cyanobacteria and chemoautotrophs, display enhanced CO_2-fixation by maintaining CO_2 within the compartment [21, 24]. Ribulose 1,5 bisphosphate carboxylase/oxygenase (RuBisCO) requires CO_2 as a substrate but bicarbonate is abundant in the cytosol; RuBisCO cannot utilize bicarbonate as a substrate. In the case of carbon-fixation, compartmentalized RuBisCO cannot utilize inorganic carbon; therefore, carbonic anhydrase (CA) is also found within the carboxysome, thereby increasing CO_2 concentration and possibly stored for later use by RubisCO [30]. Bicarbonate can traverse the selective semi-permeable shell of the carboxysome where it is converted by CA, retaining and enhancing CO_2 fixation [31]. Carboxysome compartmentalization and carboxysome formation has been studied. Cameron, *et al* elegantly described the step-wise formation of the carboxysome, where the interior of the carboxysome compartment forms first, followed by shell proteins encapsulating the carboxysome nucleation proteins [17]. Ectopic expression of *Synechoccus elongates* PCC 7942 of the *ccm* operon (which includes shell proteins and encapsulated enzymes) and knockout variations were performed, following each fluorescently tagged protein during carboxysome assembly [17]. The carboxysome shell proteins are CcmK1, CcmK2, CcmK4, CcmO, and CcmL [32, 33]. Carboxysome formation begins with the interaction between CcmM and RuBisCo, forming the procarboxysome (PC); the N-terminal portion of CcmN (required for carboxysome biogenesis) interacts with the γ-CA domain of CcmM in the PC [17, 32,

34]. The C-terminal encapsulating peptide (EP) of CcmN interacts and localizes with CcmK2) shell protein (structure shown in Figure 2B(middle)—a localization event dependent upon the presence CcmO [17, 34]. Finally, CcmL shell protein caps the carboxysome, resulting in a fully assembled carboxysome, with a diameter ranging from 100-150 nm (Figure 2B, bottom) [17, 35]. The large range in carboxysome diameter is due to a dynamic number of subunits that form a compartment. The structure of CcmK2 shell protein is shown in Figure 2B (middle), with a pore at the center of the hexamer [16]. More studies are currently being performed, investigating the assembly and permeability of the carboxysomal shell to understand the passage of substrates or bi-products through the pores of the shell.

1,2-Propanediol/Ethanolamine Utilization (Pdu/Eut) MCPs are two other examples of bacterial microcompartments. *Salmonella enterica* Pdu MCPs are proposed to organize reactions as well as isolate 'unstable' reaction intermediates that may form within the compartment [14, 36] such as proteins involved in B_{12}-dependent 1,2-propanediol (1,2-PD) catabolism [27] and ethanolamine catabolism [24, 37]. 1,2-propanediol, a major anaerobic degradation product from the breakdown of rhamnose and fucose, is readily available to bacteria that are in proximity of plant breakdown matter—as is observed in enteric bacteria found in the large intestine of farm animals [24, 38]. Fully assembled Pdu MCPs also have a diameter that ranges from 100-150 nm [38]; Pdu MCPs are mostly polyhedral (Figure 1.2C, bottom), although not as regular as carboxysomes [39]. Pdu MCPs compartmentalize the enzymes involved in the sequential utilization of 1,2-PD, yielding ATP and propionyl-CoA while sequestering propionaldehyde (a toxic

intermediate) within the MCP [24, 39]. Recently, Cheng *et al* also reported a possible internal recycling mechanism performed by PduQ alcohol dehydrogenase, recycling NADH to NAD^+ to supply PduP (propionaldehyde dehydrogenase) for the conversion of propionaldehyde to propionyl-CoA [21]. Of the 23 genes in the *pdu* operon, 7 are similar to the shell proteins of the carboxysome [24]. It was reported that an N-terminal sequence of PduP (important for MCP assembly) interacts with the C-terminal region of PduA (a BMC shell protein [18]); authors indicated that the N-terminal sequence of enzymes is necessary for compartmentalization, as it interacts with the C-terminal sequence of shell proteins [39]. A common feature of MCP enzymes is a short N-terminal sequence proposed to be necessary for targeting enzymes to the MCPs, similar to the EPs observed in carboxysome formation [17, 30, 34, 39]. In this case, Pdu MCPs encapsulate enzymes involved in 1,2-PD degradation, but sequesters propionaldehyde and internally recycles for cofactor homeostasis to PduP [40]. *Salmonella* and some *Escherichia coli* strains have Eut MCPs, involved in ethanolamine utilization, where it is an important bacterial carbon source in the mammalian gut [24]. The roles of Eut MCPs are very similar to those of Pdu MCP, where 17 genes in the *eut* operon encode the proteins for Eut MCP [24]. Furthermore, due to the similarities between these three types of MCPs, it has been previously alluded to that these compartments are of particular interest as they could serve as important and novel vehicles for drug or vaccine delivery [36, 41].

Encapsulins (Enc) are a very recently discovered family of conserved prokaryotic compartments, first described by Sutter, *et al*, that resemble a virus-like particle [19].

Enc proteins were discovered in *Thermotoga maritima*, an anaerobic heterotroph, reported to be a multi-subunit protease, maritimacin, after observing that it shares 32% sequence identity and 53% amino acid similarity to *Brevibacterium linens* linocin, in addition to having similar size and shape to linocin as visualized by EM [42-44]. Enc are temperature and pH-stable, and self-assemble [45]. After the crystal structure for *T. maritima* maritimacin was published by Sutter, *et al*, the name of the multi-subunit protein was changed to Enc as it was not observed to have bacteriostatic nor bacteriocidal activities [19]. Herein, *T. maritima* Enc will be annotated as Tm-Enc. While MCPs noted above are compartments composed of hetero-subunits, Tm-Enc assembles into a 60-subunit icosahedral nanocompartment, with a diameter of ~ 240 Å (**Figure 1.2D**) [19]. Tm-Enc nanocompartment has 5-, 3-, and 2-fold symmetry, where at each point of symmetry uncharged, positively charged, and negatively charged pores are formed, respectively [19]. During refinement, extra electron density was observed along a hydrophobic channel within the interior of the Tm-Enc compartment, where it was discovered that amino acids from the C-terminal end of a ferritin-like protein (Flp) can be modeled into the density.

A genome-wide analysis search revealed that the gene encoding for Enc in bacteria lies in a two-gene operon with either Flps or dye-decolozing peroxidases (DyP). Tm-Enc and its putative cargos occur in a two-gene operon (Figure 2D, top), suggesting tight translational coupling (**Figure 1.3**) [19, 46]. Amino acid sequence analysis of Flps and DyPs revealed a hydrophobic C-terminal extension that was hypothesized to be important for cargo proteins to be targeted to the interior of the compartment, similar to the function of an EP [17, 34]. Over-expression of the co-operon Bl-Enc:Bl-DyP resulted

in co-purification of both proteins, where truncation of DyP C-terminus abolished the interaction, indicating that the C-terminus is important for interaction with Bl-Enc. It is possible that other enzymes with a similar C-terminal tail could interact with the interior of Enc when co-expressed. The addition of C-terminal extension to either teal fluorescent protein (TFP) and enhanced green florescent protein (GFP) allowed these to be encapsulated within Bl-Enc and *Rhodococcus erythropolis* N771 Enc (Re-Enc), respectively [47, 48]; removal of Bl-Enc C-terminal docking sequence failed to encapsulate TFP or non-specific binders, suggesting selective cargo loading [47], thus rendering the C-termini as an EP necessary for internalization of cargo.

Work regarding Enc nanocompartments as a drug delivery or antigen delivery platform has been occurring, where recently a hepatocellular carcinoma cell binding peptide was engineered to the exterior of Tm-Enc to be used as a targeting ligand to bind to HepG2 cells [49]. This work has further propelled the ability to use Enc nanocompartments as a therapeutic delivery system.

One interesting observation made by Sutter, *et al* was that Flps and DyPs are enzymes involved in oxidative stress [19]. This would implicate a possible role for Enc to be involved in mitigating oxidative stress in the local environment encountered by bacteria that harbor the Enc protein. This was further explored and corroborated by a recent report by McHugh, *et al* 2014, showing that *Myxococcus xanthus* Enc (EncA, Mx-Enc) encapsulates EncB (Mx-Flp), where isolation of Mx-Enc nanocompartments were five-fold higher in cells grown in nutrient-starved cells versus actively growing cells [50]. Furthermore, an Mx-Enc knockout strain displayed attenuated growth upon exposure to hydrogen peroxide (H_2O_2), where presumably nanocompartment formation protects the

cell from oxidative stress [50]. These results further implicate Enc proteins in their involvement in mitigating oxidative stress.

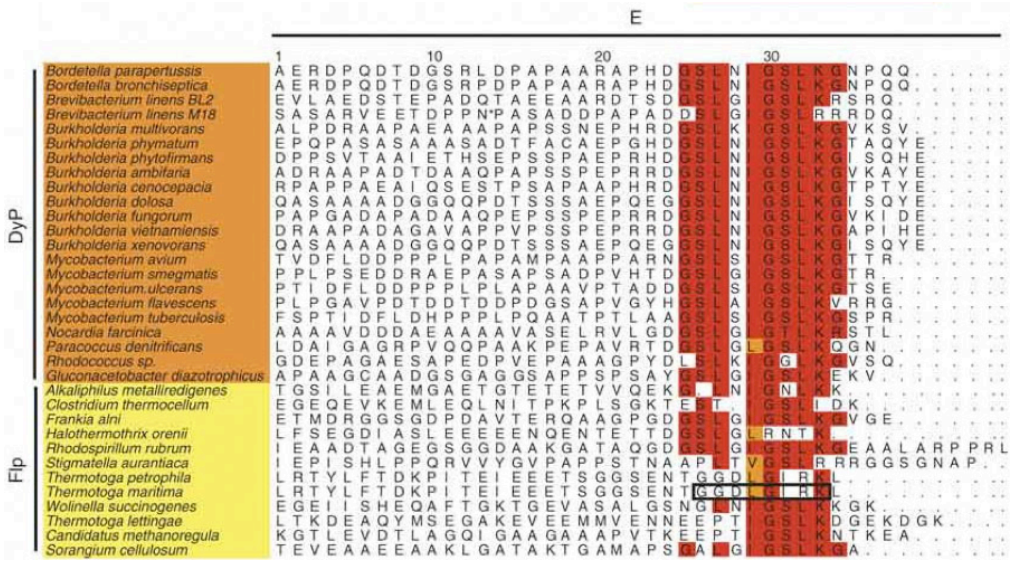

Figure 1.3: **Alignment of C-terminal extensions of putative Encapsulin cargo proteins** (Adapted from reference #15). Red indicates conserved anchor sequence.

Dye-Decolorizing Peroxidases (DyP)

Within the peroxidase superfamily, three distantly structurally unrelated classes were defined as plant, fungal, and bacterial [51]. However, since this classification, a novel class of dye-decolorizing peroxidases (DyP) has emerged when DyP from a fungus, *Thanatephorus cucumeris*, was discovered to contain an extracellular peroxidase-type enzyme able to decolorize anthraquinone and azo dyes [52-54]. DyPs are ubiquitous across bacterial species although they share low sequence identity (around 20% between DyPs from *Pseudomonas aeruginosa* (Pa-DyP), *Shewanella oneidensis* (So-DyP), and *Bacterioides thetaiotamicron* (Bt-DyP) (**Figure 1.4A**); yet these DyPs have structural homology (**Figure 1.4B**) [53]. One phylogenetic classification separates DyPs into four groups, each with a diverse function [55].

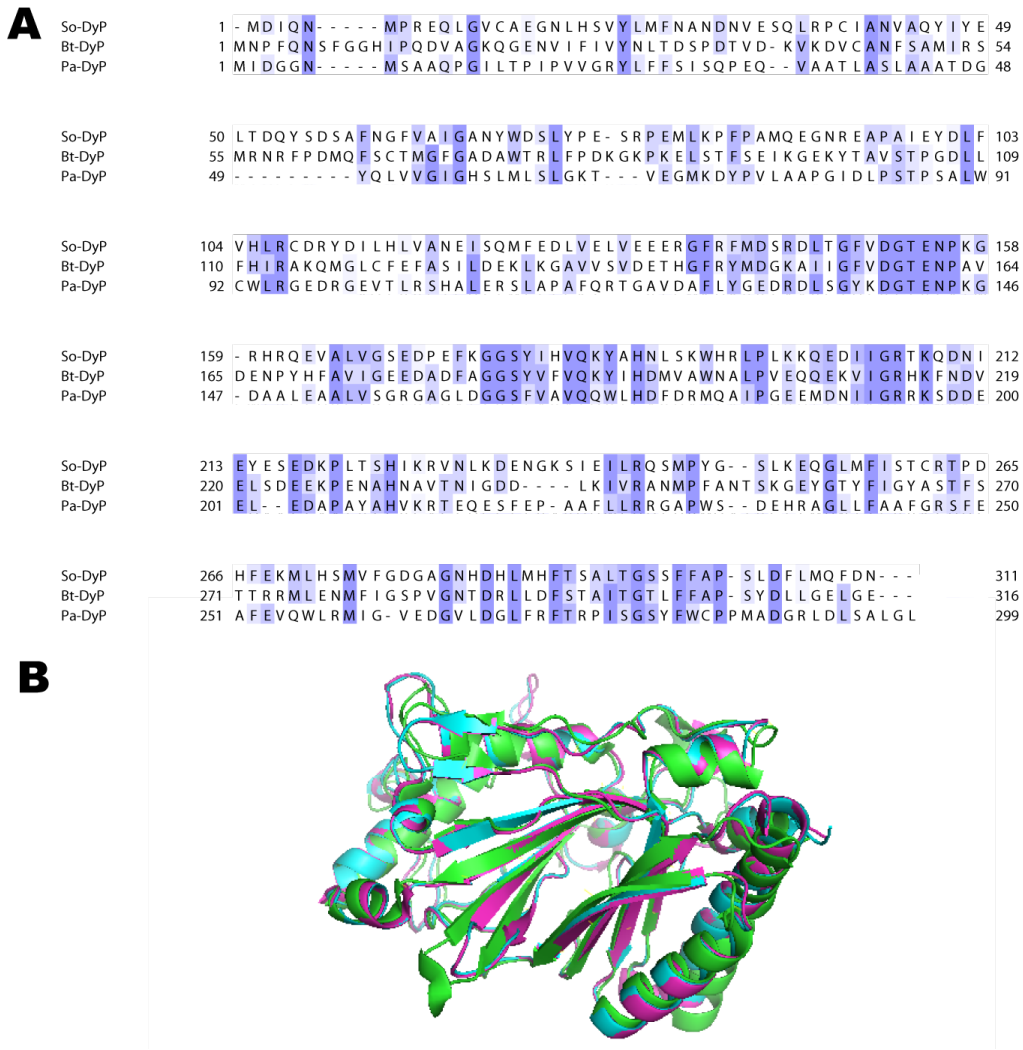

Figure 1.4: Homology between Dye-decolorizing Peroxidases. A) BLOSUM Alignment of three DyPs: *S. onedensis* (So-DyP), *B. thetaiotamicron* (Bt-DyP); and *P. aeruginosa* (Pa-DyP). B) Structural alignment of So-DyP (magenta), Bt-DyP (cyan), and Bl-DyP (green), with an estimated root mean square deviation (RMSD) of ~ 0.908

Notably the EfeB and YfeX subfamilies are of interest, as both have heme-binding abilities, as well as reported deferrochelation and peroxidase activities [56, 57]. *Escherichia coli* YfeX and EfeB were proposed to assist in extracting iron from heme while keeping the tetrapyrrole ring intact [56]. This was proposed based on protoporphyrin IX (PPIX) accumulation upon overexpression of YfeX in *E. coli* [56]. Moreover, when the media was supplemented with heme, the presence of PPIX increased

over time for *E. coli* cells overexpressing EfeB, suggesting that EfeB is producing PPIX from heme. Both YfeX and EfeB were observed to be functionally similar in heme utilization as both proteins could complement each other [56]. Heme demetallation by enzymes like YfeX and EfeB, as suggested by Goblirsch, *et al*, is important for organisms that do not have heme oxygenases, such as Mtb [57], although Mtb has a heme-degrading enzyme [58]. Contrary to this result, another study suggests that YfeX is a typical dye-decolorizing peroxidase that may be involved in the heme biosynthetic pathway in *E. coli* and does not possess deferrocheletase activity [59].

Rhodococcus jostii RHA1 harbors two homologs of DyP proteins [55]. DypB from *R. jostii* RHA1 has been identified as a lignin peroxidase [55]. DypA is hypothesized to function in capturing iron from heme as it has no lignin peroxidase activity and is a close homolog to YfeX [60]. Like the EfeB subfamily, other DyP enzymes have been shown to be bi-functional [57], such as DyP from the fungus *Bjerkandera adusta*, which is both a heme peroxidase and mediates hydrolysis of dyes with anthraquinone rings [46, 61].

Objectives of dissertation research: i) preliminary characterization of Mt-DyP and ii) Mt-Enc, as well as iii) identifying Mt-Enc cargo and its role in mycobacterial survival.

First, Mt-DyP has not been previously characterized, thus we sought preliminary characterization of Mt-DyP. Mt-DyP is a heme-dependent peroxidase and not a deferrocheletase as previously suggested. Furthermore, it is monomeric in its apo- form and tetrameric when heme-bound.

We have observed that Mtb contains a nanocompartment, harboring an *enc* gene

(Mt-Enc) that exists in a two-gene operon with a *dyp* gene (Mt-DyP). We show that Mt-Enc self-assembles into an icosahedral nanocompartment, similar to that observed for Tm-Enc, where Mt-DyP interacts with Mt-Enc via a C-terminal extension. Furthermore, we have identified two other putative cargo proteins: iron-storing Mt-BfrB (ferritin) [13, 62] and Mt-FolB (7,8-dihydroneopterin aldolase) [63]. Mt-BfrB and Mt-FolB harbor an aliphatic C-terminal extension and both enzymes are implicated in oxidative stress. All three of the aforementioned enzymes interact with Mt-Enc via their C-termini and remain active within the Mt-Enc nanocompartment.

Lastly, we sought to characterize the importance of mycobacterial Enc *in vivo*. We constructed an Enc knockout in the fast-growing, non-pathogenic strain of mycobacteria, *Mycobacterium smegmatis* mc^2155. We observed that *M.smegmatis* $mc^2155\Delta enc$ shows a delayed growth phenotype in the presence of H_2O_2, compared to the wild-type strain. It was also observed that both Mt-Enc and *M. smegmatis* Enc (Ms-Enc) have a type VII secretion signal (YXXXD/E) that may provide clues to Mt-Enc secretion as previously reported [64]. Attempts at identifying other putative cargo enzymes are also addressed.

REFERENCES

1. Organization, W.H., *Global Tuberculosis Report.* 2014.
2. Rook, G.A., K. Dheda, and A. Zumla, *Immune responses to tuberculosis in developing countries: implications for new vaccines.* Nat Rev Immunol, 2005. **5**(8): p. 661-7.
3. *Drug-Resistant TB*. 2014, Centers for Disease Control and Prevention.
4. Orme, I.M., R.T. Robinson, and A.M. Cooper, *The balance between protective and pathogenic immune responses in the TB-infected lung.* Nat Immunol, 2014. **16**(1): p. 57-63.
5. Kaufmann, S.H., *New issues in tuberculosis.* Ann Rheum Dis, 2004. **63 Suppl 2**: p. ii50-ii56.

6. Vergne, I., et al., *Mechanism of phagolysosome biogenesis block by viable Mycobacterium tuberculosis.* Proc Natl Acad Sci U S A, 2005. **102**(11): p. 4033-8.
7. Skeiky, Y.A. and J.C. Sadoff, *Advances in tuberculosis vaccine strategies.* Nat Rev Microbiol, 2006. **4**(6): p. 469-76.
8. Hafner, R. *NIAID TB Therapeutic Clinical Research Priorities*. in *Working Group on New TB Drugs 2011 Annual Meeting*.
9. van den Boogaard, J., et al., *New drugs against tuberculosis: problems, progress, and evaluation of agents in clinical development.* Antimicrob Agents Chemother, 2009. **53**(3): p. 849-62.
10. Organization, W.H. *Treatment of Tuberculosis: guidelines for national programmes*. 2010; 4th Edition:[Available from: http://www.who.int/tb/features_archive/new_treatment_guidelines_may2010/en/.
11. Thiry, G. *Vaccine development pipeline: progress and challenges*. in *WHO Product development for vaccines advisory committee meeting*. 2014. Geneva, Switzerland.
12. Fauci, A.S., *Multidrug-resistant and extensively drug-resistant tuberculosis: the National Institute of Allergy and Infectious Diseases Research agenda and recommendations for priority research.* J Infect Dis, 2008. **197**(11): p. 1493-8.
13. Khare, G., et al., *Ferritin structure from Mycobacterium tuberculosis: comparative study with homologues identifies extended C-terminus involved in ferroxidase activity.* PLoS One, 2011. **6**(4): p. e18570.
14. Yeates, T.O., M.C. Thompson, and T.A. Bobik, *The protein shells of bacterial microcompartment organelles.* Curr Opin Struct Biol, 2011. **21**(2): p. 223-31.
15. Jorda, J., et al., *Using comparative genomics to uncover new kinds of protein-based metabolic organelles in bacteria.* Protein Sci, 2013. **22**(2): p. 179-95.
16. Kerfeld, C.A., et al., *Protein structures forming the shell of primitive bacterial organelles.* Science, 2005. **309**(5736): p. 936-8.
17. Cameron, J.C., et al., *Biogenesis of a bacterial organelle: the carboxysome assembly pathway.* Cell, 2013. **155**(5): p. 1131-40.
18. Pang, A., et al., *Structural insights into higher order assembly and function of the bacterial microcompartment protein PduA.* J Biol Chem, 2014. **289**(32): p. 22377-84.
19. Sutter, M., et al., *Structural basis of enzyme encapsulation into a bacterial nanocompartment.* Nat Struct Mol Biol, 2008. **15**(9): p. 939-47.
20. Gregersen, N., J. Hansen, and J. Palmfeldt, *Mitochondrial proteomics--a tool for the study of metabolic disorders.* J Inherit Metab Dis, 2012. **35**(4): p. 715-26.
21. Chen, A.H. and P.A. Silver, *Designing biological compartmentalization.* Trends Cell Biol, 2012. **22**(12): p. 662-70.
22. Kang, S. and T. Douglas, *Biochemistry. Some enzymes just need a space of their own.* Science, 2010. **327**(5961): p. 42-3.
23. Theil, E.C., M. Matzapetakis, and X. Liu, *Ferritins: iron/oxygen biominerals in protein nanocages.* J Biol Inorg Chem, 2006. **11**(7): p. 803-10.
24. Kerfeld, C.A., S. Heinhorst, and G.C. Cannon, *Bacterial microcompartments.* Annu Rev Microbiol, 2010. **64**: p. 391-408.

25. English, R.S., et al., *Isolation and characterization of a carboxysome shell gene from Thiobacillus neapolitanus.* Mol Microbiol, 1994. **12**(4): p. 647-54.
26. Arosio, P., R. Ingrassia, and P. Cavadini, *Ferritins: a family of molecules for iron storage, antioxidation and more.* Biochim Biophys Acta, 2009. **1790**(7): p. 589-99.
27. Yeates, T.O., C.S. Crowley, and S. Tanaka, *Bacterial microcompartment organelles: protein shell structure and evolution.* Annu Rev Biophys, 2010. **39**: p. 185-205.
28. Corchero, J.L. and J. Cedano, *Self-assembling, protein-based intracellular bacterial organelles: emerging vehicles for encapsulating, targeting and delivering therapeutical cargoes.* Microb Cell Fact, 2011. **10**: p. 92.
29. Chowdhury, C., et al., *Diverse bacterial microcompartment organelles.* Microbiol Mol Biol Rev, 2014. **78**(3): p. 438-68.
30. Cannon, G.C., S. Heinhorst, and C.A. Kerfeld, *Carboxysomal carbonic anhydrases: Structure and role in microbial CO2 fixation.* Biochim Biophys Acta, 2010. **1804**(2): p. 382-92.
31. Cai, F., et al., *The pentameric vertex proteins are necessary for the icosahedral carboxysome shell to function as a CO2 leakage barrier.* PLoS One, 2009. **4**(10): p. e7521.
32. Kinney, J.N., S.D. Axen, and C.A. Kerfeld, *Comparative analysis of carboxysome shell proteins.* Photosynth Res, 2011. **109**(1-3): p. 21-32.
33. Sutter, M., et al., *Two new high-resolution crystal structures of carboxysome pentamer proteins reveal high structural conservation of CcmL orthologs among distantly related cyanobacterial species.* Photosynth Res, 2013. **118**(1-2): p. 9-16.
34. Kinney, J.N., et al., *Elucidating essential role of conserved carboxysomal protein CcmN reveals common feature of bacterial microcompartment assembly.* J Biol Chem, 2012. **287**(21): p. 17729-36.
35. Tanaka, S., et al., *Atomic-level models of the bacterial carboxysome shell.* Science, 2008. **319**(5866): p. 1083-6.
36. Choudhary, S., et al., *Engineered protein nano-compartments for targeted enzyme localization.* PLoS One, 2012. **7**(3): p. e33342.
37. Brinsmade, S.R., T. Paldon, and J.C. Escalante-Semerena, *Minimal functions and physiological conditions required for growth of salmonella enterica on ethanolamine in the absence of the metabolosome.* J Bacteriol, 2005. **187**(23): p. 8039-46.
38. Cheng, S., et al., *Bacterial microcompartments: their properties and paradoxes.* Bioessays, 2008. **30**(11-12): p. 1084-95.
39. Fan, C., et al., *Short N-terminal sequences package proteins into bacterial microcompartments.* Proc Natl Acad Sci U S A, 2010. **107**(16): p. 7509-14.
40. Cheng, S., et al., *Genetic analysis of the protein shell of the microcompartments involved in coenzyme B12-dependent 1,2-propanediol degradation by Salmonella.* J Bacteriol, 2011. **193**(6): p. 1385-92.
41. Kochut, A. and P. Dersch, *Bacterial invasion factors: tools for crossing biological barriers and drug delivery?* Eur J Pharm Biopharm, 2013. **84**(2): p. 242-50.

42. Hicks, P.M., et al., *Homomultimeric protease in the hyperthermophilic bacterium Thermotoga maritima has structural and amino acid sequence homology to bacteriocins in mesophilic bacteria.* Febs Letters, 1998. **440**(3): p. 393-8.
43. Valdes-Stauber, N. and S. Scherer, *Isolation and characterization of Linocin M18, a bacteriocin produced by Brevibacterium linens.* Appl Environ Microbiol, 1994. **60**(10): p. 3809-14.
44. Hicks, P.M., L.S. Chang, and R.M. Kelly, *Homomultimeric protease and putative bacteriocin homolog from Thermotoga maritima.* Methods Enzymol, 2001. **330**: p. 455-60.
45. Rahmanpour, R. and T.D. Bugg, *Assembly in vitro of Rhodococcus jostii RHA1 encapsulin and peroxidase DypB to form a nanocompartment.* FEBS J, 2013. **280**(9): p. 2097-104.
46. Roback, P., et al., *A predicted operon map for Mycobacterium tuberculosis.* Nucleic Acids Res, 2007. **35**(15): p. 5085-95.
47. Rurup, W.F., et al., *Self-sorting of foreign proteins in a bacterial nanocompartment.* J Am Chem Soc, 2014. **136**(10): p. 3828-32.
48. Tamura, A., et al., *Packaging guest proteins into the encapsulin nanocompartment from Rhodococcus erythropolis N771.* Biotechnol Bioeng, 2015. **112**(1): p. 13-20.
49. Moon, H., et al., *Developing genetically engineered encapsulin protein cage nanoparticles as a targeted delivery nanoplatform.* Biomacromolecules, 2014. **15**(10): p. 3794-801.
50. McHugh, C.A., et al., *A virus capsid-like nanocompartment that stores iron and protects bacteria from oxidative stress.* EMBO J, 2014. **33**(17): p. 1896-911.
51. Welinder, K.G., *Bacterial catalase-peroxidases are gene duplicated members of the plant peroxidase superfamily.* Biochim Biophys Acta, 1991. **1080**(3): p. 215-20.
52. Kim, S.J. and M. Shoda, *Purification and characterization of a novel peroxidase from Geotrichum candidum dec 1 involved in decolorization of dyes.* Appl Environ Microbiol, 1999. **65**(3): p. 1029-35.
53. Sugano, Y., *DyP-type peroxidases comprise a novel heme peroxidase family.* Cell Mol Life Sci, 2009. **66**(8): p. 1387-403.
54. Sugano, Y., et al., *DyP, a unique dye-decolorizing peroxidase, represents a novel heme peroxidase family: ASP171 replaces the distal histidine of classical peroxidases.* J Biol Chem, 2007. **282**(50): p. 36652-8.
55. Ahmad, M., et al., *Identification of DypB from Rhodococcus jostii RHA1 as a lignin peroxidase.* Biochemistry, 2011. **50**(23): p. 5096-107.
56. Letoffe, S., et al., *Bacteria capture iron from heme by keeping tetrapyrrol skeleton intact.* Proc Natl Acad Sci U S A, 2009. **106**(28): p. 11719-24.
57. Goblirsch, B., et al., *Chlorite dismutases, DyPs, and EfeB: 3 microbial heme enzyme families comprise the CDE structural superfamily.* J Mol Biol, 2011. **408**(3): p. 379-98.
58. Chim, N., et al., *Unusual diheme conformation of the heme-degrading protein from Mycobacterium tuberculosis.* Journal of Molecular Biology, 2010. **395**(3): p. 595-608.

59. Dailey, H.A., et al., *The Escherichia coli protein YfeX functions as a porphyrinogen oxidase, not a heme dechelatase.* MBio, 2011. **2**(6): p. e00248-11.
60. Roberts, J.N., et al., *Characterization of dye-decolorizing peroxidases from Rhodococcus jostii RHA1.* Biochemistry, 2011. **50**(23): p. 5108-19.
61. Sugano, Y., et al., *Degradation pathway of an anthraquinone dye catalyzed by a unique peroxidase DyP from Thanatephorus cucumeris Dec 1.* Biodegradation, 2009. **20**(3): p. 433-40.
62. McMath, L.M., et al., *Crystallization and preliminary X-ray crystallographic analysis of a Mycobacterium tuberculosis ferritin homolog, BfrB.* Acta Crystallogr Sect F Struct Biol Cryst Commun, 2010. **66**(Pt 12): p. 1657-61.
63. Goulding, C.W., et al., *Regulation by oligomerization in a mycobacterial folate biosynthetic enzyme.* Journal of Molecular Biology, 2005. **349**(1): p. 61-72.
64. Rosenkrands, I., et al., *Identification and characterization of a 29-kilodalton protein from Mycobacterium tuberculosis culture filtrate recognized by mouse memory effector cells.* Infect Immun, 1998. **66**(6): p. 2728-35.

CHAPTER 2

PRELIMINARY CHARACTERIZATION OF A DYE-DECOLORIZING PEROXIDASE from *Mycobacterium tuberculosis* H37Rv

The *M. tuberculosis* work was performed in collaboration with members of the Horowitz lab at UCLA.

ABSTRACT

Mycobacterium tuberculosis (Mtb) possesses a dye-decolorizing-like peroxidase (Mt-DyP)—part of a novel family of heme peroxidases. Characterization of other DyPs is currently underway. Although the first DyP was discovered in a fungus, DyPs are ubiquitous across bacteria and some fungi. Their precise roles in nature are unknown, but these enzymes have been implicated in the breakdown of high-redox-potential anthraquinone dyes and lignin degradation. Dyes and lignin are unlikely substrates for a DyP from a pathogen such as Mtb. Homologs of DyP found in *E. coli* were implicated in iron extraction from heme, thus presenting a possible role for Mtb DyP (Mt-DyP) in Mtb as part of the heme-uptake pathway. Although there are not other reported heme demetallation DyP-like enzymes, we sought to characterize Mt-DyP as a peroxidase or a deferrochelatase. This chapter demonstrates the preliminary characterization of Mt-DyP, where we show that Mt-DyP is a heme-peroxidase and not a deferrochelatase. Furthermore we show progress toward generating protein crystals that could ultimately be used to solve the structure of heme-bound Mt-DyP by X-ray crystallography.

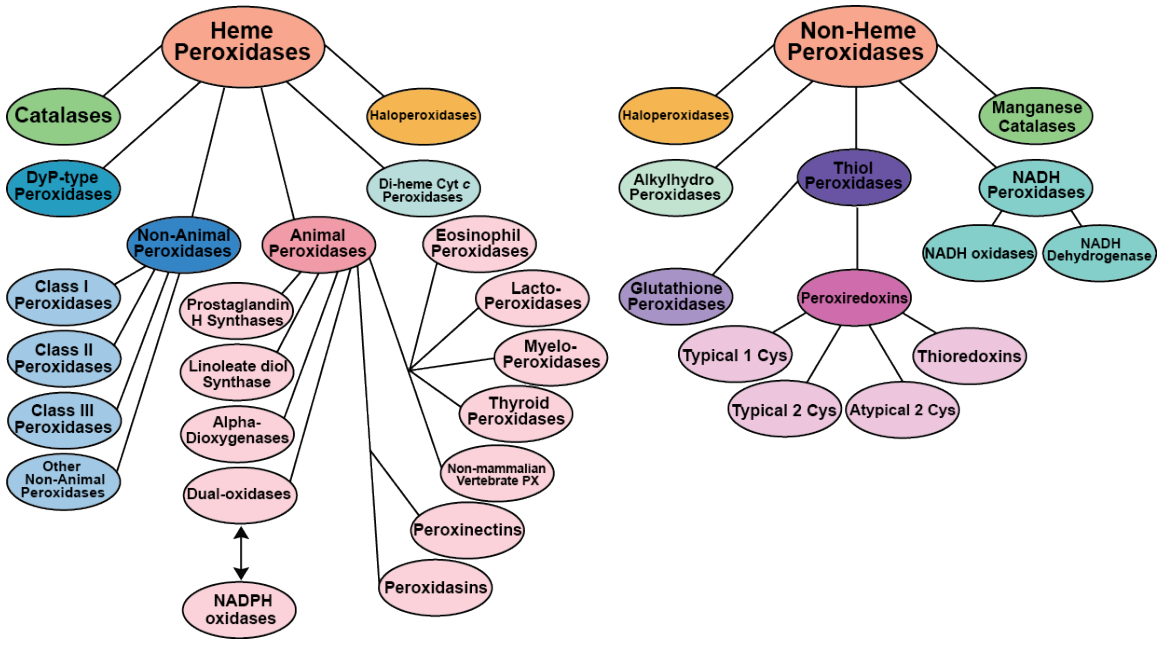

Figure 2.1: Schematic representation of the classes of peroxidases in the PeroxiBase database [1].

INTRODUCTION

Peroxidases are oxidoreductases that catalyze the oxidation of an electron-donating substrate and the reduction of peroxide, often hydrogen peroxide (H_2O_2). There are either heme- or non-heme peroxidases, and the various families have been grouped as shown in **Figure 2.1**, modified from the Peroxibase database [1]. Heme peroxidases are enzymes found in almost all organisms, and have been divided into the following classes: haloperoxidases, di-heme cytochrome *c* peroxidases, animal peroxidases, non-animal peroxidases, catalases, and dye-decolorizing-type peroxidases [1]. Heme peroxidases catalyze the following reaction:

$$ROOH + 2AH \rightarrow ROH + 2A\bullet + H_2O$$

In the above reaction, ROOH is a peroxide, AH is the substrate, and • is an unpaired electron. The general heme peroxidase mechanism is the following schema:

I. $(Fe^{3+}) + H_2O_2 \rightarrow (Fe^{4+} = O)\bullet^{+} + H_2O$

II. $(Fe^{4+} =O)\square^{+} + AH \rightarrow (Fe^{4+} =O) + A\square$

III. $(Fe^{4+} =O) + AH \rightarrow (Fe^{3+}) + A\square + H_2O$

In step I, H_2O_2 is heterolytically cleaved, releasing a water molecule while the second peroxide-derived oxygen coordinates to heme-iron, thus forming an oxoferryl intermediate called compound I. Compound I is then reduced to compound II in the presence of one substrate AH. Finally, the second substrate AH reduces compound I, thus returning the enzyme to the ferric resting state and releasing a water molecule [2, 3]. Although heme peroxidases catalyze the same general reaction as shown above, not all peroxidases catalyze in the same fashion. While classical heme peroxidases utilize a distal His residue in the active site as an acid-base catalyst to promote heterolytic cleavage of H_2O_2, it was recently discovered that a new family of peroxidases use Asp instead of His as the catalytic residue [4]; although this was also observed with other peroxidases [5, 6].

An emerging class of dye-decolorizing peroxidases (DyPs) are heme-peroxidases, ubiquitous among bacteria and some fungi, and are unique from classical peroxidases [7]. The first DyP discovered was from a fungus, *Bjerkandera adusta* (previously misidentified to have been isolated from *Thanatephorus cucumeris* Dec. 1, then *Geotrichm candidum)*, discovered to contain a secreted heme peroxidase-type enzyme able to decolorize anthraquinone dyes [7-11]. The discovery of a peroxidase with the ability to decolorize anthraquinone dyes efficiently at low pH had great industrial

implications in regards to alternatives in bioremediation, removing toxic waste from water without the use of toxic additives thus decreasing waste [12]. Currently, there are four classes of DyPs: A, B, C, and D [1]. Three important characteristics of DyPs are: i) an Asp catalytic residue, found in the conserved GXXDG motif (**Figure 2.2A**); ii) two-domain α+β protein, with four anti-parallel β-sheets sandwiched by α-helices in a ferredoxin-like fold (**Figure 2.2B**); and iii) classes A, C, and D DyPs have longer sequences compared to class B proteins [13]. One of the α-helices harbors an important proximal His residue as the fifth ligand to the heme-iron in DyP [14]. Classification of DyPs into a CDE superfamily of peroxidases, which include chlorite dismutases (Cld), DyP-, and EfeB-like proteins was proposed due to the detoxifying abilities, heme-binding, and structural homologies [15]. Recently, however, sequence-based structural differences observed between DyPs and Cld-like proteins are such that make this classification defunct; thus, reclassification of DyPs has been proposed to better fit different functional types of DyPs as new ones are being studied and reported [13]. As of now, according to PeroxiBase, the former classification still stands [1].

Possible bi-functionalities of bacterial and fungal DyPs include the breakdown of anthraquinone dyes [9] and lignin degradation [16]. A third proposed functionality of some of the members of DyP family include deferrochelatation [17]. Two DyP-like peroxidases from *Escherichia coli*, YfeX and EfeB, were proposed to assist in extracting iron from heme while keeping the tetrapyrrole ring intact, based on the accumulation of protoporphyrin IX (PPIX) while over-expressing YfeX [17]. This would implicate a unique bi-functionality of DyPs—peroxidase activity and iron acquisition from heme. Another study reports that YfeX is the penultimate enzyme in the heme biosynthetic

pathway, prior to ferrochelatase insertion of iron into PPIX, explaining the accumulation of PPIX [18]. Hence, the bi-functionality of DyPs in regards to iron acquisition from heme remains uncertain.

A

```
BlDyP          MNPFQNSFGGHIPQDVAGKQGENVIFIVYNLTDSPDTVDKVKDVCANFSAMIRSMRNRFP 60
MtbDyP         -----VAVPAVSPQPILAPLTPAAIFLVATIG--ADGEATVHDALSKISGLVRAIGFRDP 53
TyrA           ---MDIQNMPREQLGVCAEGNLHSVYLMFNAN--DNVESQLRPCIANVAQYIYELTDQYS 55
YfeX           --------MSQVQSGILPEHCRAAIWIEANVK--GEVD-ALRAASKTFADKLATFEAKFP 49
                               :        :::  .     :    ::    ..:  :  :  : .

BlDyP          DMQFSCTMGFGADAWTRLFPDKGKPKELSTFSEIKGEKYTAVSTPGDLLFHIRAKQMGLC 120
MtbDyP         TKHLSVVVSIGSDAWDRLFAGP-RPTELHPFVELTGPRHTAPATPGDLLFHIRAETMDVC 112
TyrA           DSAFNGFVAIGANYWDSLYPES-RPEMLKPFPAMQEGNREAPAIEYDLFVHLRCDRYDIL 114
YfeX           DAHLGAVVAFGNNTWRALSGGV-GAEELKDFPGYGKG--LAPTTQFDVLIHILSLRHDVN 106
                  :.  :.:* : *  *      .  *  *         * :   *::.*: .   .:

BlDyP          FEFASILDEKLKGAVVSVDETHGFRYMDGKAIIGFVDGTENPAVDENPYHFAVIGEEDAD 180
MtbDyP         FELAGRILKSMGDAVTVVDEVHGFRFFDNRDLLGFVDGTENPSG-PIAIKATTIGDEDRN 171
TyrA           HLVANEISQMFEDLVELVEEERGFRFMDSRDLTGFVDGTENPKGRH-RQEVALVGSEDPE 173
YfeX           FSVAQAAMEAFGDCIEVKEEIHGFRWVEERDLSGFVDGTENPAGEETRREVAVIK--DGV 164
               . .*    : : . :   :* :***:.: : : *********       . : :   *

BlDyP          FAGGSYVFVQKYIHDMVAWNALPVEQQEKVIGRHKFNDVELSDEEKPGNAHNAVTNIGDD 240
MtbDyP         FAGSCYVHVQKYVHDMASWESLSVTEQERVIGRTKLDDIELDDNAKPANSHVALNVITDD 231
TyrA           FKGGSYIHVQKYAHNLSKWHRLPLKKQEDIIGRTKQDNIEYESEDKPLTSHIKRVNLKDE 233
YfeX           DAGGSYVFVQRWEHNLKQLNRMSVHDQEMVIGRTKEANEEIDGDERPETSHLTRVDLK-E 223
                 *..*:.**:: *::   . :.: .** :*** *  : * ..: :* .:*     :  :

BlDyP          ----LKIVRANMPFANTSKGEYGTYFIGYASTFSTTRRMLENMFIGSPAGNTDRLLDFST 296
MtbDyP         DGTERKIVRHNMPFGEVGKGEYGTYFIGYSRTPTVTEQMLRNMFLGDPAGNTDRVLDFST 291
TyrA           NGKSIEILRQSMPYGSLK--EQGLMFISTCRTPDHFEKMLHSMVFGDGAGNHDHLMHFTS 291
YfeX           DGKGLKIVRQSLPYGTASG-THGLYFCAYCARLHNIEQQLLSM-FGDTDGKRDAMLRFTK 281
                    :*:* .:*:.       *  * . .      .: * .* :*.  *: * :: *:.

BlDyP          AITGTLFFVPSYDLLG------------------ELGE------ 316
MtbDyP         AVTGGLFFSPTIDFLDHPPPLPQAATPTLAAGSLSIGSLKGSPR 335
TyrA           ALTGSSFFAPSLDFLMQFDN------------------------ 311
YfeX           PVTGGYYFAPSLDKLMAL-------------------------- 299
               .:**  :* *: * *
```

B

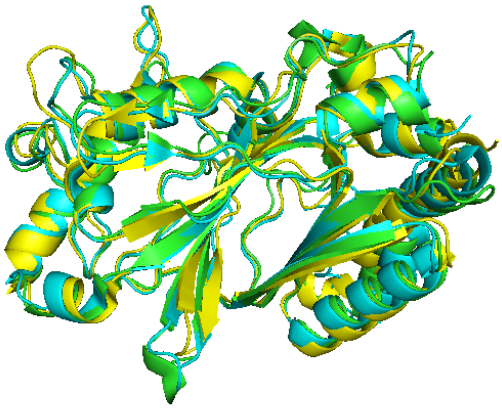

Figure 2.2: Similarities between prokaryotic DyPs. A) Clustal-W alignment of amino acid sequences from prokarytotic DyPs: *Brevibacterium linens* Bl-DyP, *Mycobacterium tuberculosis* Mt-DyP, *Shewanella onedensis* TyrA, and *Escherichia coli* YfeX. The conserved GXXDG motif and conserved fifth-ligand His are noted by a red box. B) Overlay of known bacterial DyP structures, with a Root Mean Square Deviation (RMSD) ~ 0.795 Å. Green is the Bt-DyP structure from *Bacteroides thetaiotamicron* [19]; cyan is TyrA [20], and yellow is Rh-DyPB from *Rhodococcus jostii* RHA1 [21].

For Mtb to utilize heme as an iron source, mycobacteria must have evolved a mechanism to extract iron from heme. We have identified a protein involved in Mtb heme degradation, whereby a cytosolic heme-degrading enzyme, MhuD, cleaves the tetrapyrrol ring to release iron [22, 23]. However, an Mtb MhuD deletion mutant does not completely abolish growth in limiting concentrations of heme or hemoglobin, although growth is attenuated, suggesting that another Mtb protein may also release iron from heme [24]. Mtb encodes a DyP-like peroxidase, *Rv0799c* (Mt-DyP), located in a two-gene operon with the gene encoding for a linocin-like encapsulating protein, similar to some DyPs [18, 25, 26]. Mt-DyP shares 39% sequence identity to *E. coli* YfeX, suggesting that Mt-DyP may possess a similar function as a deferrochelatase, thereby extracting iron from heme and contributing to heme uptake. We hypothesized that Mt-DyP may posess the capability of extracting iron from heme, thus revealing another Mtb protein involved in heme degradation. To date, no information regarding a DyP-like peroxidase has been reported in Mtb. We sought to preliminarily characterize Mt-DyP, including its possible function as either a deferrochelatase or a heme-dependent peroxidase, or both.

MATERIALS AND METHODS

Cloning.

The Mtb gene encoding protein Mt-DyP (*Rv0799c*) was PCR-amplified from Mtb H37Rv genomic DNA (obtained from BEI Resources) using the KOD HotStart Polymerase Kit

(Novagen) with 5′ and 3′ primers (MWG Operon) containing specific restriction sites outlined in **Table 2.1**, to either generate a C-terminal poly-histidine tag (His-tag), or a truncated C-terminus. PCR products were ligated into pCR-BluntII-TOPO (Invitrogen), and then transformed into *E. coli* OneShot TOP10 cells (Invitrogen). Double digestions with specific restriction enzymes were performed. Excised genes were ligated into the appropriate linearized *E. coli* expression vector and transformed into *E. coli* BL21-Gold (DE3) cells (Novagen). Each construct was verified by DNA sequencing using T7 promoter and reverse primers (Laguna Scientific).

Table 2.1: Primers for PCR-amplification and cloning.

Gene	Constructs	Primer Sequence
Rv0799c (DyP)	pET22::DyP	FWD: 5' GGCATATGGTGGCTGTGCCTGCTGTCTCTCCGCAGC REV: 5' GGCTCGAGTCGGGGGCTTCCTTTCAAGCTGCCGATCGATAGC
Rv0799c (DyP R240A)	pET22::DyP R240A	FWD: 5' CACCGAGCGCAAGATCGTGGCCCACAACATGCCGTTCGGC REV: 5' GCCGAACGGCATGTTGTGGGCCACGATCTTGCGCTCGGTG
Rv0799c (DyP H222A)	pET22::DyP H222A	FWD: 5' CGCGAAGCCAGCCAATTCAGCGGTAGCGCTCAACGTCATCAC REV: 5' GTGATGACGTTGAGCGCTACCGCTGAATTGGCTGGCTTCGCG
Rv0799c (DyP D149A)	pET22::DyP D149A	FWD: 5' ATCTGCTGGGCTTTGTCGCGGGCACCGAAAACCCAAGTG REV: 5' CACTTGGGTTTTCGGTGCCCGCGACAAAGCCCAGCAGAT

Overexpression and purification of Mtb proteins.

*Proteins were expressed from plasmids described in Table 2.2, using E. coli BL21 Gold (DE3) cells. T*he following concentrations of antibiotics to the media were added where appropriate: kanamycin (30 μg/ml) and ampicillin (50 μg/ml). Cells harboring appropriate expression vector was grown aerobically at 37 °C in LB media containing the either kanamycin or ampicillin. Protein expression was induced at OD_{600nm} ~0.4 by the

addition of isopropyl-β-d-thiogalactopyranoside (IPTG, 1 mM). Cells were harvested after four hours of induction, by centrifugation at 5100 x g for 20 min. Additionally, where indicated in the results, cultures expressing Mt-DyP_{His} were supplemented with 0.7 mM δ-aminolevulinic acid (δ-ALA) at the time of induction and left expression overnight at 18°C.

Harvested cell pellets were resuspended in 20 mL of Buffer A (50 mM Tris-HCl, pH 7.4, 350 mM NaCl, 10 mM imidazole and 10% glycerol), followed by the addition of phenylmethylsulfonyl fluoride (PMSF) and hen egg-white lysozyme. Cells were disrupted by sonication, clarified by centrifugation at 18000 x g for 30 min at 4° C and syringe-filtered (1.0 μm pore size) for removal of cell debris. The clarified cell lysate was then loaded onto a 5 mL Ni^{2+}-charged HisTrap column (GE Healthcare) pre-equilibrated with Buffer A. The protein(s) were eluted with a linear gradient of 10-500 mM imidazole (100 mL). Eluted fractions were collected, analyzed by SDS-PAGE, and concentrated using Amicon concentrators with the appropriate molecular-weight cutoff (Millipore, Bedford, MA). Mt-DyP_{His} was further purified on a Superdex 200 HiLoad 16/60 gel filtration chromatography column (GE Healthcare) utilizing 50 mM Tris-HCl, pH 7.4, and 150 mM NaCl. Apo-protein concentration was determined by UV/vis spectroscopy, the molar extinction coefficient at 280 nm (ε_{280nm} ~ 27,310 $M^{-1}cm^{-1}$) as predicted by the program Protein Calculator (Scripps), or determined by modified Lowry [27] assay for heme-bound protein.

Crystallization of apo- and heme-bound Mt-Dyp.

Mt-DyP was purified and reconstituted with heme as previously described. Purified

protein was concentrated to ~ 8 mg/mL. A Mosquito nanoliter dispensing robot was used to set up 600 sparse matrix crystallization conditions for each protein. Crystallization conditions were optimized and additive screens were utilized. Crystals were grown using hanging-drop vapor diffusion techniques. Diffraction data from diffracting, single crystals were collected at either Stanford Synchrotron Radiation Laboratory (SSRL) or with Mtb Structural Genomics Consortium at the Advanced Light Source at UC Berkeley (ALS). Diffracting crystals of heme-bound Mt-DyP were grown at 8 mg/mL in 0.2 M $MgCl_2$, 0.1 M Tris-HCl pH 8.5, and 13% polyethylene glycol.

Heme and protoporphyrin IX (PPIX) binding by Mt-DyP.

Solutions were made and titration experiments were preformed as previously described [28, 29]. Heme [30] and PPIX [31] solutions were freshly prepared prior to each experiment. Briefly, approximately 4 mg of heme was dissolved in 1mL of ice-cold 100 mM NaOH and vortexed periodically over 20 min period. 1 mL of 1 M Tris (pH 7.4) was added to the solution, centrifuged for 10 min at 4 °C at 13000 rpm. The supernatant was then diluted with 50 mM Tris (pH 7.4) and 150 mM NaCl, centrifuged again at 13000 rpm for 10 min to remove undissolved heme. Final concentrations were determined using an ε_{385} of 58.44 mM^{-1} cm^{-1}. Crystals of PPIX were dissolved in 150 mM NaCl, vortexed periodically over 20 min and centrifuged. Supernatant was collected and diluted in 2.7 N HCl, where concentration was determined using an ε_{408} of 262 mM^{-1} cm^{-1}. PPIX solutions were diluted in 50 mM Tris-HCl (pH 7.4) and 150 mM NaCl for experiments. Heme and PPIX solutions were protected from light and used within 12 h.

Micromolar increments of either heme or PPIX were titrated into either 5 μM or 2.5 μM of purified apo-Mt-DyP, respectively, in 50 mM Tris-HCl (pH 7.4) and 150 mM NaCl.

The rate of heme dissociation from heme-Mt-DyP was measured using H64Y/V68F apo-myoglobin as a heme scavenger protein. In brief, heme-Mt-DyP was incubated with 10-fold excess of H64Y/V68F apo-myoglobin, mixed in an SX.18MV stopped-flow spectrophotometer (Applied Photophysics) or by UV/vis spectrometry (DU-800 spectrophotometer, Beckman-Coulter) by equal volume and rapid mixing at room temperature. The reaction was monitored at 600 nm for either 1000 seconds or 200 minutes. The resulting time course for Mt-DyP was fitted to single- and double-exponential functions using Graph-pad Prism.

Spectroscopic analyses of Mt-DyP.

Spectroscopic analyses of apo-, PPIX-, and heme-bound Mt-DyP were performed on a DU-800 spectrophotometer (Beckman-Coulter). Briefly, UV/Vis spectra of ~ 2.5 μM Mt-DyP was monitored over 20 min upon the addition of 100 μM H_2O_2, where changes in Soret were observed.

Kinetics of Mt-DyP enzymatic activity.

All enzymatic reactions were performed in 100 μL reactions in a quartz cuvette. Peroxidase activity utilizing guaiacol as a substrate was monitored over 20 min as a change in absorbance at 470 nm, indicative of the production of tetraguaiacol. Two μM Mt-DyP (reconstituted with heme to obtain a 1:1 heme-to-protein molar ratio) and apo-Mt-DyP were assayed in the presence of 10 mM H_2O_2 (Sigma-Aldrich) and 10 mM

guaiacol (Sigma-Aldrich) [32] in 100 mM sodium citrate at pH 4. Reactions were initiated upon the addition of H_2O_2, and were monitored over a 12 min period observing the change in absorbance at 470 nm by UV/vis spectrometry (DU-800 spectrophotometer, Beckman-Coulter).

Further analysis of the activity of Mt-DyP was carried out by utilization of 2,2'-azino-bis(3-ethylbenzothiazoline-6-sulfonic acid (ABTS) as a substrate, as previously described [26]. Briefly, 250 mM ABTS was added to a reaction mixture containing 1.25 μM Mt-DyP in 50 mM HEPES at pH 5.5. Reaction was initiated upon the addition of 2.5 μM H_2O_2 and monitored for 10 min at 420 nm by UV/vis spectrometry (DU-800 spectrophotometer, Beckman-Coulter).

To assess deferrocheletase activity, 5 μg of *Mycobacterium smegmatis* mc^2155 cytosol and cell membrane was isolated as previously described [33, 34] and incubated with exogenous, recombinant Mt-DyP (5 μM). The change at 402 nm was monitored over 10 min by UV/vis spectrometry (DU-800 spectrophotometer, Beckman-Coulter).

Generation and characterization of Mtb*ΔmbtB strains deficient in Mt-DyP, MhuD, or both Mt-DyP and MhuD.*

Mtb*ΔmbtB*, a mycobactin-deficient mutant of Mtb, was used as the parental strain to construct the following double and triple mutants, Mtb*ΔmbtBΔdyP*, Mtb*ΔmbtBΔmhuD*, and Mtb*ΔmbtBΔmhuDΔdyP,* via specialized transduction as previously described [35]. To generate the Mtb*ΔmbtBΔmhuD* mutant, we constructed an allelic exchange substrate by cloning a 1.9 kb PCR product including the entire *mhuD* gene and flanking regions and then replacing 222 nucleotides of the *mhuD* gene (encoding amino acids 10-85 of the

105 amino acid (aa) MhuD protein) with an apramycin resistance cassette, using essentially the same strategy described previously for construction of Mtb*ΔmbtBΔmmpL11* and Mtb*ΔmbtBΔRv0203* [35]. To generate the Δ*dyP* mutants, Mtb*ΔmbtBΔdyP* and Mtb*ΔmbtBΔmhuDΔdyP*, we first inserted an 849 nt in-frame, unmarked deletion of *dyP* (encoding aa 21-303 of the 335 aa mt-DyP protein) flanked by a hygromycin resistance gene and sacB cassette (hyg-sacB) into the *dyP* region of the chromosome by specialized transduction. In a second step, we passaged hygromycin resistant clones that were sensitive to sucrose (due to expression of sacB) in the absence of hygromycin to allow for recombination to occur between homologous regions flanking the hyg-sacB cassette (eliminating the hyg-sacB cassette from the chromosome and leaving just the unmarked Δ*dyP* mutation), and then plated on 7H10 plates containing 2% sucrose. Sucrose resistant clones were confirmed to have lost hygromycin resistance by plating on 7H10 plates with and without hygromycin. All mutants were confirmed to have the correct genotype by PCR.

Heme utilization experiments were carried out as previously described [35], except the 7H9 growth medium was supplemented with 10% OADC and 0.01% Tyloxapol (7H9-OADC-TLX). Log-phase bacteria were inoculated at an initial OD_{750} of 0.0005 into 30 mL 7H9-OADC-TLX containing no supplement, 0.2 μM heme, or 10 ng/mL mycobactin J and grown for 21 days until the growth of all strains had plateaued in the presence of mycobactin J. Growth of the double and triple mutants (Mtb*ΔmbtBΔdyP*, Mtb*ΔmbtBΔmhuD*, and Mtb*ΔmbtBΔmhuDΔdyP*) was compared with the growth of the Mtb*ΔmbtB* parental strain under identical conditions to determine whether the mutations impacted the ability of the bacteria to acquire iron from heme. For

comparisons, all OD_{750} measurements were normalized to the OD_{750} measured for the MtbΔ*mbtB* parental strain in the presence of mycobactin J at the same time point.

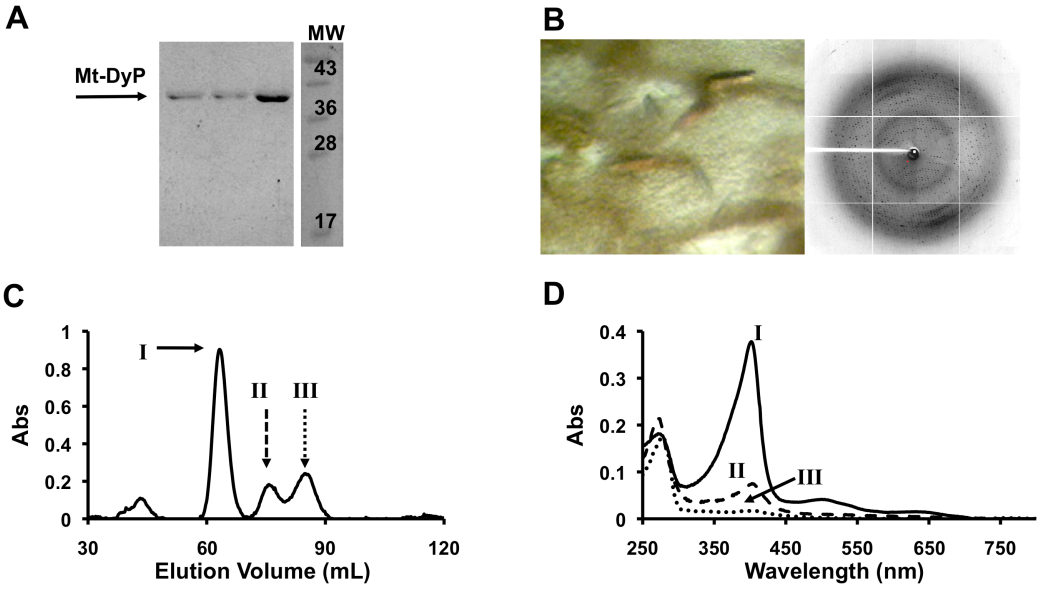

Figure 2.3: Purified Mt-DyP. A) Elution fractions from nickel-affinity purification of Mt-DyP. B) Mt-DyP protein crystals (left) and its appropriate diffraction pattern. Image obtained from Advanced Light Source, beamline 8-2-1. C) Representation of three distinct elution peaks from gel-filtration experiment performed on a S-200 16/600 Superdex column, yielding an Mt-DyP tetramer (I, 116 kDa), dimer (II, 58 kDa), and a monomer (III, 29 kDa). D) UV/vis spectra of elutions resulting from three elution peaks post-gel filtration conveys unique heme concentrations for tetramer (solid line), dimer (dashed line) and monomer (dotted line).

RESULTS

Over-expression and purification of Mt-DyP.

Mt-DyP was purified to homogeneity by nickel-affinity chromatography (**Figure 2.3A**). Purified Mt-DyP was colored pink suggesting heme-bound Mt-DyP. Size exclusion chromatography (SEC) was utilized to separate apo- and heme-bound species, where three distinct oligomeric states of Mt-DyP were separated and analyzed at 402 nm to determine heme bound fractions (**Figures 2.3C and 2.3D**). Monomeric Mt-DyP was in its apo form, similar to other bacterial DyPs (DyP2 from *Amycolatopsis* sp. 75iv2 and *E. coli* apo-EfeB [5, 36]), and a minor dimeric Mt-DyP species with a sub-stoichimetric

quantity of heme was also observed. Crystals of heme-Mt-DyP diffracted to ~ 3.2 Å (**Figure 2.3B**); crystals are currently being screened and optimized to improve diffraction data.

PPIX and Heme-binding of Mt-DyP.

From purification, tetrameric Mt-DyP had a near 1:1 molar stoichiometry of heme bound (**Figure 2.3D**). To obtain the protein:heme stoichiometry, apo-Mt-DyP was reconstituted with either PPIX or heme, and a UV/vis spectra was recorded (**Figure 2.4A**) . The change in absorbance at 402 nm was plotted against heme concentration, yielding approximately a 1:1 protein:heme or protein:PPIX molar ratio (**Figure 2.4B and C**). The change in absorbance was calculated at 407 nm and 402 nm because that was the absorbance maxima for PPIX- and heme-Mt-DyP, respectively **Figure 2.4B and C**). This evidence suggests that Mt-DyP has the ability to bind both heme and PPIX.

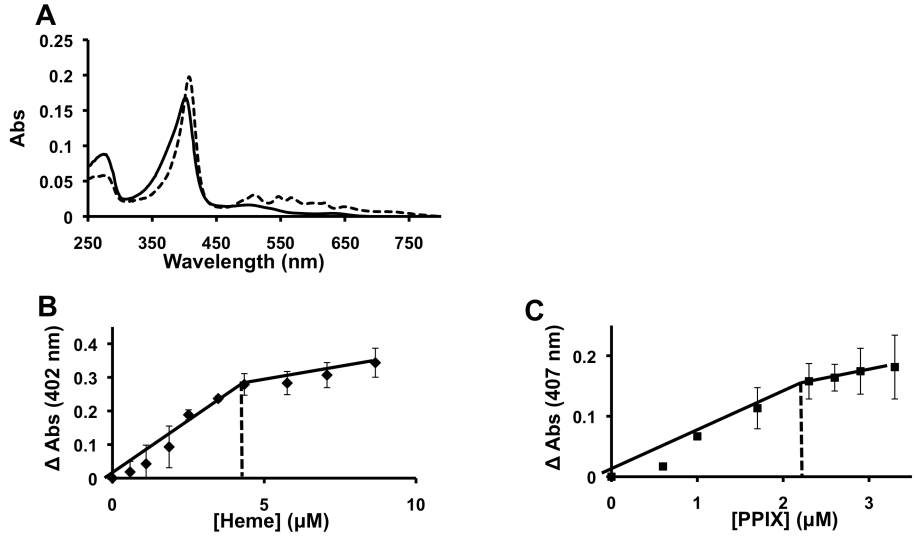

Figure 2.4: Heme- and PPIX-binding of Mt-DyP. A) UV/vis spectra of heme- (solid line) or PPIX- (dashed line) bound Mt-DyP to a 1:1 molar ratio. B and C) Heme or PPIX titrated into 5 μM or 2.5 μM apo-Mt-DyP, respectively. Measuring the difference spectra between apo-Mt-DyP and heme- or PPIX-only sample to find the wavelength of maximum absorbance difference at which change remains constant generated saturation curve.

The off rate was measured using the apo-H64Y/V68F-myoglobin assay whereby the apomyoglobin mutant acts as a high-affinity heme scavenger [37]. The rate of the reaction depends solely on the unimolecular rate of heme release, k_{-heme}. The release of heme from ferric Mt-DyP was recorded over time and the time courses fit to both single- and double-exponential functions (**Figure 2.5A and D**). Both fitting choices produced statistically good results (R^2 > 96%), although the double-exponential fit produced slightly more normal residuals (**Figure 2.5B and E, C and F**). Fitting a single-exponential function to the data produces a k_{-heme} of 2.3 X 10^{-3} ± 2.4 X 10^{-4} sec^{-1} when monitored over 1000 seconds whereas double-exponential fitting yields a fast phase, k_f, equal to 5.1 X 10^{-2} ± 1.5 X 10^{-2} sec^{-1} and a slow phase, k_s, equal to 1.0 X 10^{-3} ± 2.4 X 10^{-4} sec^{-1}. The first 10 seconds of the experiment were omitted as fluctuation skewed the fit for the two-phase association calculation. For monitoring the reaction over 200 minutes, fitting a single-exponential function to the data produces a k_{-heme} of 4.7 X 10^{-5} ± 1.7 X 10^{-6} sec^{-1} when monitored over 1000 seconds whereas double-exponential fitting yields a fast phase, k_f, equal to 3.2 X 10^{-4} ± 4.9 X 10^{-5} sec^{-1} and a slow phase, k_s, equal to 1.5 X 10^{-5} ± 3.5 X 10^{-6} sec^{-1}. While literature searches have not revealed published heme-off rates for DyPs. One recent report published the heme-off rate of *Listeria monogcytogenes* HemQ (chlorite dismutase-like protein [15]) to be k_{-off}= 8.1 sec^{-1} X 10^{-3} [38], a lower heme affinity compared to Mt-DyP.

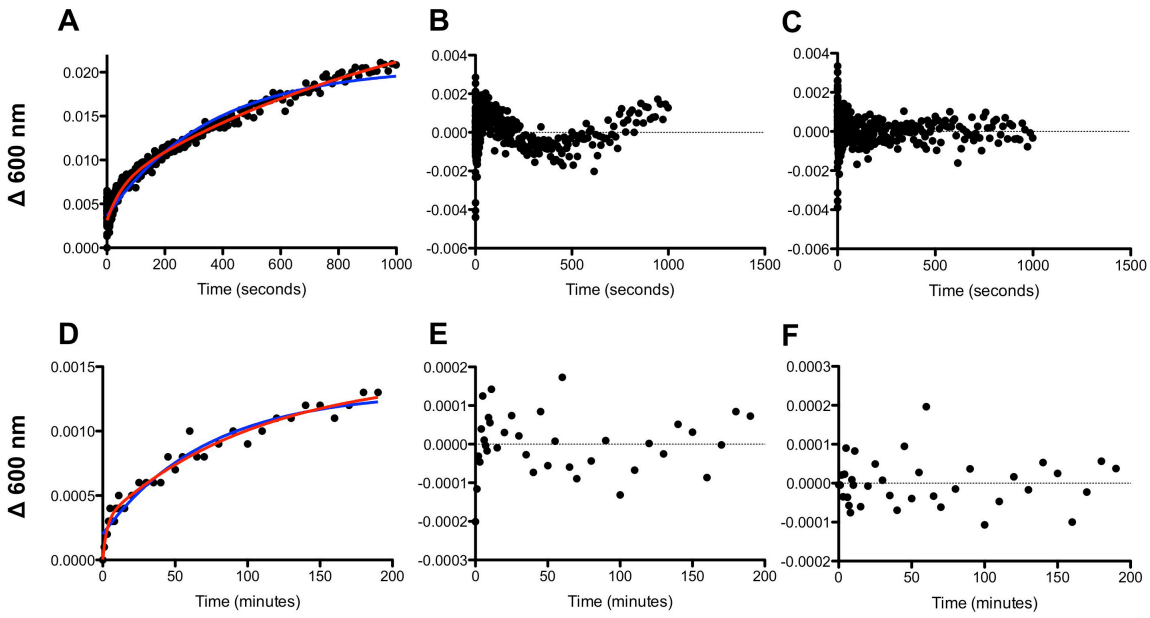

Figure 2.5: Heme off-rate measurements of heme-bound Mt-DyP. Experimental data fitted to either single (blue) or double (red) exponential curves between (A)10 – 1000 seconds or (D) 0 – 200 minutes for binding of ferric heme to apo-Mb from heme-bound Mt-DyP. Residual plots for the single (B and E) and double (C and F) curve fit of the reaction of heme dissociation from Mt-DyP to apo-Mb. The random distribution of residuals is representative point distance of data point from the curve fit.

Mt-DyP is a heme-dependent peroxidase.

To determine whether Mt-DyP possesses heme peroxidase activity, we investigated the ability of heme-Mt-DyP to form a central peroxidase intermediate, compound I, in the presence of H_2O_2. Previously, Rj-DypB has been observed to form a stable compound I [32] in the presence of H_2O_2. Addition of H_2O_2 to heme-Mt-DyP resulted in a color change of brown to green (**Figure 2.6A**), indicative of compound I formation. Further, the UV/vis spectrum of heme-Mt-DyP after addition of H_2O_2 at pH 7.4 was measured, displaying a decrease and blue-shifted, hypochromatic Soret peak from 402 nm to 398 nm, with the emergence of a slight shoulder at around 345 nm (**Figure 2.6B**) and a broad hyperchromaticity between 576 and 648 (**Figure 2.6B, inset**) similar to that observed for Rj-DypB [21] and horse radish peroxidase [39] (**Table 2.3**).

Interestingly, this peroxidase compound I intermediate was stable for at least 20 mins, suggesting that Mt-DyP is able to form an intermediate comparable to compound I upon addition of H_2O_2. This suggested that Mt-DyP is a heme peroxidase.

Mt-DyP was tested for peroxidase activity using either guaiacol or ABTS as substrates. First, peroxidase activities of heme-bound and apo-Mt-DyP were compared with heme-Mt-DyP utilizing guaiacol as a substrate (**Figure 2.6C**). Apo-Mt-DyP had no detectable peroxidase activity compared with heme-Mt-DyP, suggesting that Mt-DyP is a heme peroxidase that can utilize guaiacol as a substrate. Second, utilizing ABTS as a chromagen, heme-Mt-DyP produced a radical cation in the presence of H_2O_2 and active peroxidase, while apo-Mt-DyP displayed background peroxidase activity, respectively (**Figure 2.6D**). These results further suggest that Mt-Dyp is a heme peroxidase.

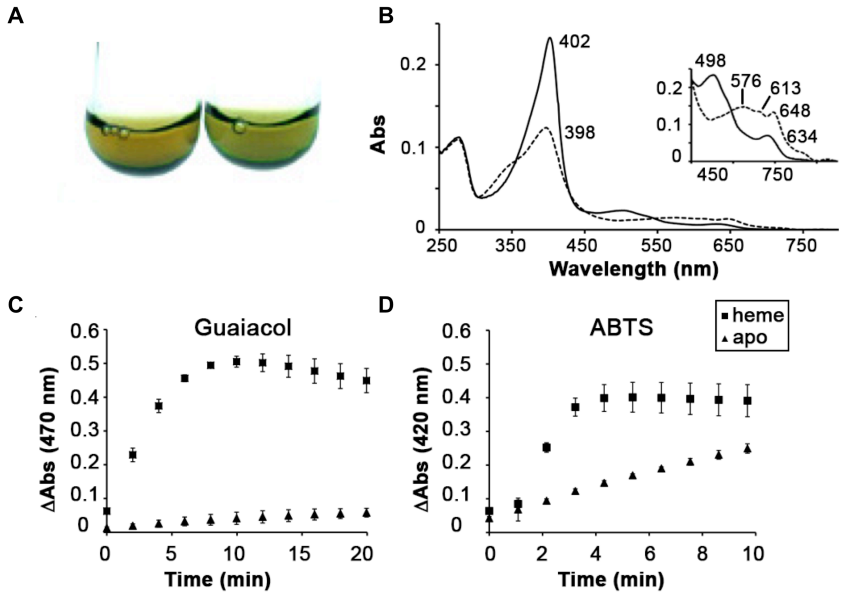

Figure 2.6: Peroxidase activity of Mt-DyP. A) Resting enzyme (left), Mt-DyP turns green (right) upon the addition of H_2O_2. **B)** UV/Vis spectra of ~ 2.5 µM Mt-DyP was monitored over 20 min upon the addition of 100 µM H_2O_2, where changes in Soret were observed. Solid line is spectrum resting enzyme; dashed line is spectrum of enzyme in the presence of H_2O_2. Inset: the visible region magnified 10-fold. **C)** Production of tetraguaiacol was monitored over 20 min as a change in absorbance at 470 nm. The reaction was performed in the presence of 2 µM Mt-DyP with 10 mM guaiacol and 10 mM H_2O_2 in 100 mM sodium citrate pH 4. **D)** To further monitor peroxidase activity, 1.25 µM Mt-DyP or encapsulated Mt-DyP in the presence of 250 mM ABTS at pH 5.5. Reaction was initiated upon the addition of 2.5 µM H_2O_2. Apo-Mt-DyP (triangles), and Mt-DyP (squares) are shown here.

Table 2.2: Absorption spectral comparisons of resting and compound I forms of heme-dependent peroxidases.

Name	Resting (nm)	Compound I (nm)	Ref
Mt-DyP	402, 498, 634	398, 576, 613, 648	This Study
Rj-DypB	404, 503, 634	400, 580, 613, 648	[21]
HRP	403, 498, 640	400, 525 (sh), 577, 622 (sh), 651	[39]

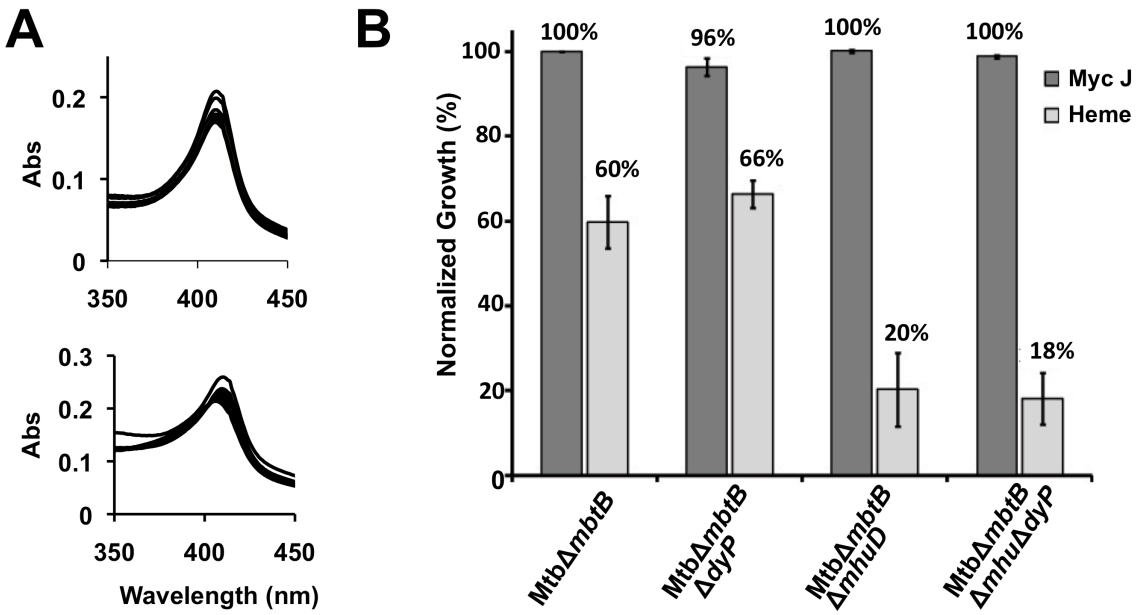

Figure 2.7: Examining Mt-DyP deferrochelatase activity *in vitro* and *in vivo*. A) 5 μg of *M. smegmatis* mc^{2}155 cytosolic fraction (A, top) or (A, bottom) 5 μg of *M. smegmatis* mc^{2}155 TritonX-114-extracted membrane incubated with 5 μg Mt-DyP, respectively. Possible deferrochelatase activity was monitored over 20 minites, where a decrease in Soret intensity would be observed over time. B) Mtb*ΔmbtB*, Mtb*ΔmbtBΔdyP*, Mtb*ΔmbtBΔmhuD* and Mtb*ΔmbtBΔmhuDΔdyP* were grown in 7H9-OADC containing 0.01% tyloxapol media supplemented with either 0.2 μM heme or mycobactin J (myc) (10 ng/mL). Growth was monitored by measuring absorbance at 750 nm over 21 days and the results shown are taken from the time when growth of the parental Mtb*ΔmbtB* strain had reached a plateau. The experiment was performed twice with similar results.

Mt-DyP is not a deferrochelatase.

Even though DyP has heme-dependent peroxidase activity, deferrochelatase activity of Mt-DyP was also assessed. Heme-bound Mt-DyP was incubated with either *M. smegmatis* cytosolic or membrane fractions, as previously described for Ec-YfeX

[17]. No significant change in the Soret region was observed under the reaction conditions tested, indicating no PPIX product was formed as a result of iron extraction from heme (**Figures 2.7A, top and bottom**), thus suggesting that Mt-DyP does not possess deferrochelatase activity. However, deferrochelatase activity will require an electron source for iron extraction from heme, and the *in vitro* assays tested may not contain the necessary requirements for activity. Thus, we sought to investigate the function of Mt-DyP *in vivo* by observing the effects of Mt-DyP upon mycobacterial heme-iron acquisition.

Mycobacteria possess a heme uptake pathway [35, 40] along with a cytosolic heme-degrading enzyme, MhuD, which has been shown to degrade heme by cleavage of the tetrapyrrole ring to release iron and the products, mycobiilins [22, 23]. To test for Mt-DyP deferrochelatase activity, we compared growth of Mtb deletion mutants lacking MhuD, Mt-DyP, or both MhuD and Mt-DyP, in medium containing a limiting heme concentration (0.2 μM) as a source of iron (**Figure 2.7B**). Under these conditions, Mtb∆*mbtB* supplemented with 0.2 μM heme grows to 60% of the growth achieved in the presence of exogenous siderophore (mycobactin). Of note, this parental strain, Mtb∆*mbtB*, has an interrupted mycobactin biosynthetic pathway to prevent siderophore-mediated iron acquisition [35, 40] and cannot utilize Fe(III) in the medium in the absence of added siderophore. The Mtb∆*mbtB*∆*dyP* mutant in the presence of 0.2 μM heme grew at a rate similar to Mtb∆*mbtB*. In contrast, the growth rate of the Mtb∆*mbtB*∆*mhuD* mutant was significantly attenuated in the presence of 0.2 μM heme compared with Mtb∆*mbtB* (67% reduction; 20% vs. 60%) (**Figure 2.7B**). These results suggest that MhuD but not Mt-DyP is required for heme degradation when heme is the sole iron

source, at least when MhuD is present. To determine whether Mt-DyP is required for iron sequestration in the absence of MhuD, we generated the triple mutant, MtbΔ*mbtB*Δ*mhuD*Δ*dyP*. Growth of MtbΔ*mbtB*Δ*mhuD*Δ*dyP* in the presence of 0.2 μM heme was significantly attenuated compared with MtbΔ*mbtB* (18% vs. 60%), but no more so than MtbΔ*mbtB*Δ*mhuD*, indicating that Mt-DyP is not required for iron sequestration from heme in either the presence or absence of MhuD. Taken together, these results indicate that under the conditions tested, Mt-DyP does not participate in iron extraction from heme and probably does not possess deferrochelatase activity.

DISCUSSION

The positive implications of discovering a novel family of peroxidases that are able to breakdown high-redox dyes could have a great industrial impact, since contaminated runoff water from the textile industry results in dye-polluted water [41]. It was reported that a position-specific iterative BLAST search (PSI-BLAST) reveals more than 400 proteins sharing homology with *B. adusta* DyP; however, seven of those are annotated as DyP-like [7, 9]. We have identified a DyP-like, Class B peroxidase in Mtb—whose classification is likely to change to new class P peroxidase [13]. Mostly found in bacteria and lower eukaryotes, class B/P peroxidases have shorter amino acid sequences as well as lie downstream of the gene encoding for a linocin-like, encapsulating protein, Enc [25, 39]. There are no obvious differences among DyPs that demonstrate a need for reclassification; although new proposed guidelines for classification are important for DyPs within the former A, C, and D classes [13].

The active heme-bound form of Mt-DyP is tetrameric, approximately 60% heme-loaded. Heme-bound Mt-DyP can be obtained from heterologously over-expressed protein in *E. coli*, contrary to previous reports that prokaryotic DyPs are usually monomeric and purified in their apo-form [7]. Our observation was that apo-Mt-DyP is inactive, thus indicative of the need for heme for activity. In the beginning of our studies a small percentage of a heme-bound species could be purified, thus indicating to us that Mt-DyP binds heme. To increase heme incorporation for Mt-DyP, the addition of γ-ALA improved heme-bound fractions post-purification. UV/vis scans immediately post-Mt-DyP with heme bound would shift the Soret peak to ~ 410 nm (data not shown); post gel-filtration and buffer exchange the Soret peak shifted to 402 nm. The shift of the Soret to the right is likely to the effect of imidazole in the purification buffer, coordinating the heme iron.

In some peroxidases, peroxide binds to heme iron, protonating a distal His thus promoting heterolytic cleavage of H_2O_2, where a distal Arg is believed to help in the formation of compound I intermediate [3]. Distal Arg stabilizes the compound I intermediate for heterolytic cleavage of O-O [2]. Chloroperoxidase (CPO) is a catalase, performs dismutation of H_2O_2 and oxidation of alcohol previously thought to be performed via an essential thiolate ligand; however, it was Glu distal to the heme that was found to be important in the oxidation of chloride and formation of compound I [42]. Glu serves as the general acid to facilitate heterolytic cleavage by providing electrostatic destabilization to the oxoferryl intermediate [42]. DyPs do not follow the peroxidase mechanisms of classical peroxidases. The catalytic His in non-DyPs is replaced by an Asp, as in *B. adusta* DyP [7]. It was believed that the breakdown of anthraquinone dyes

is efficient in DyPs because of the lower pKa of Asp (pKa_{Asp} ~3.8) rather than His (pKa_{His} ~ 6.0) in HRP, for example (Sugano, *et al* 2009). Later reports have indicated that the distal Asp in *Auricularia auricula-judae* DyP (AauDyPI) is not as important in catalysis, but rather as a "gate" to a channel connecting the solvent with the heme for substrate access—thus upon the addition of H_2O_2 induces the distal Asp to open up the channel for substrate access [43]. In fact, the distal Arg was determined to be essential for catalytic activity in Rh-DyPB [44]. In the event that larger, biological substrates are too large for heme access via the channel, a surface-exposed Tyr could serve as a substrate interaction site [43]. Efforts for solving the crystal structures of both apo- and heme-Mt-DyP are currently underway. Until then, the mechanisms heme-binding and peroxidase activity by Mt-DyP are currently unknown.

Mt-DyP is a heme peroxidase that can utilize guaiacol and ABTS as a substrate; however deferrochelatase activity was not observed, unlike that reported in regards to YfeX/EfeB [17]. The addition of *M. smegmatis* lysates was not able to provide an appropriate environment conducive of heme-degradation by Mt-DyP. Furthermore, in an Mtb heme-degrading knock mutant, MtbΔMhuD, growth is attenuated in the presence of heme as the sole iron source, suggesting the presence of another enzyme that facilitates heme demetallation for Mtb survival. This may be because Mt-Dyp does not belong to the same class of DyPs as EfeB [13], although unlikely.

Here is the first report regarding preliminary characterization of a dye-decolorizing peroxidase in Mtb. Although it is unlikely that Mt-DyP would encounter anthraquinone dyes, ABTS, guaiacol, or lignin, as Mtb is a facultative intracellular pathogen, a few possibilities in regards to pathogenesis and survival within the human

host could be inferred. One could be that Mt-DyP can help detoxify the Mtb environment while residing within activated macrophage. Another could be that its unknown, preferred substrate is larger than is expected. The possibility of an expanded substrate range for Mt-DyP has increased with the surface-exposed substrate interaction site identified in AauDyPI, allowing other substrates to access the heme in the peroxidase active site. Until the structure of Mt-DyP is solved, however, the potential for Mt-DyP is still undiscovered.

REFERENCES

1. Fawal, N., et al., *PeroxiBase: a database for large-scale evolutionary analysis of peroxidases.* Nucleic Acids Res, 2013. **41**(Database issue): p. D441-4.
2. Poulos, T.L. and J. Kraut, *A hypothetical model of the cytochrome c peroxidase . cytochrome c electron transfer complex.* J Biol Chem, 1980. **255**(21): p. 10322-30.
3. Hiner, A.N., et al., *Mechanisms of compound I formation in heme peroxidases.* J Inorg Biochem, 2002. **91**(1): p. 27-34.
4. Yoshida, T., et al., *The catalytic mechanism of dye-decolorizing peroxidase DyP may require the swinging movement of an aspartic acid residue.* FEBS J, 2011. **278**(13): p. 2387-94.
5. Liu, X., et al., *Crystal structure and biochemical features of EfeB/YcdB from Escherichia coli O157: ASP235 plays divergent roles in different enzyme-catalyzed processes.* J Biol Chem, 2011. **286**(17): p. 14922-31.
6. Sundaramoorthy, M., J. Terner, and T.L. Poulos, *The crystal structure of chloroperoxidase: a heme peroxidase--cytochrome P450 functional hybrid.* Structure, 1995. **3**(12): p. 1367-77.
7. Sugano, Y., *DyP-type peroxidases comprise a novel heme peroxidase family.* Cell Mol Life Sci, 2009. **66**(8): p. 1387-403.
8. Kim, S.J. and M. Shoda, *Purification and characterization of a novel peroxidase from Geotrichum candidum dec 1 involved in decolorization of dyes.* Appl Environ Microbiol, 1999. **65**(3): p. 1029-35.
9. Sugano, Y., et al., *Degradation pathway of an anthraquinone dye catalyzed by a unique peroxidase DyP from Thanatephorus cucumeris Dec 1.* Biodegradation, 2009. **20**(3): p. 433-40.
10. Sugano, Y., et al., *DyP, a unique dye-decolorizing peroxidase, represents a novel heme peroxidase family: ASP171 replaces the distal histidine of classical peroxidases.* J Biol Chem, 2007. **282**(50): p. 36652-8.

11. Ruiz-Duenas, F.J., et al., *Pleurotus ostreatus heme peroxidases: an in silico analysis from the genome sequence to the enzyme molecular structure.* C R Biol, 2011. **334**(11): p. 795-805.
12. Robinson, T., et al., *Remediation of dyes in textile effluent: a critical review on current treatment technologies with a proposed alternative.* Bioresour Technol, 2001. **77**(3): p. 247-55.
13. Yoshida, T. and Y. Sugano, *A structural and functional perspective of DyP-type peroxidase family.* Arch Biochem Biophys, 2015.
14. Sugano, Y., Y. Ishii, and M. Shoda, *Role of H164 in a unique dye-decolorizing heme peroxidase DyP.* Biochem Biophys Res Commun, 2004. **322**(1): p. 126-32.
15. Goblirsch, B., et al., *Chlorite dismutases, DyPs, and EfeB: 3 microbial heme enzyme families comprise the CDE structural superfamily.* J Mol Biol, 2011. **408**(3): p. 379-98.
16. Bholay, A.D.B., Bhavna V.; Jadhav, Priyanka U.; Palekar, Kaveri S.; Dhalkari, Mayuri V.; Nalawade, P. M., *Bacterial Lignin Peroxidase: A Tool for Biobleaching and Biodegradation of Industrial Effluents.* Universal Journal of Environmental Research & Technology, 2012. **2**(1): p. 7.
17. Letoffe, S., et al., *Bacteria capture iron from heme by keeping tetrapyrrol skeleton intact.* Proc Natl Acad Sci U S A, 2009. **106**(28): p. 11719-24.
18. Dailey, H.A., et al., *The Escherichia coli protein YfeX functions as a porphyrinogen oxidase, not a heme dechelatase.* MBio, 2011. **2**(6): p. e00248-11.
19. Zubieta, C., et al., *Crystal structures of two novel dye-decolorizing peroxidases reveal a beta-barrel fold with a conserved heme-binding motif.* Proteins, 2007. **69**(2): p. 223-33.
20. Zubieta, C., et al., *Identification and structural characterization of heme binding in a novel dye-decolorizing peroxidase, TyrA.* Proteins, 2007. **69**(2): p. 234-43.
21. Roberts, J.N., et al., *Characterization of dye-decolorizing peroxidases from Rhodococcus jostii RHA1.* Biochemistry, 2011. **50**(23): p. 5108-19.
22. Chim, N., et al., *Unusual diheme conformation of the heme-degrading protein from Mycobacterium tuberculosis.* Journal of Molecular Biology, 2010. **395**(3): p. 595-608.
23. Nambu, S., et al., *A new way to degrade heme: the Mycobacterium tuberculosis enzyme MhuD catalyzes heme degradation without generating CO.* J Biol Chem, 2013. **288**(14): p. 10101-9.
24. Contreras, H., et al., *Characterization of a Mycobacterium tuberculosis nanocompartment and its potential cargo proteins.* J Biol Chem, 2014. **289**(26): p. 18279-89.
25. Sutter, M., et al., *Structural basis of enzyme encapsulation into a bacterial nanocompartment.* Nat Struct Mol Biol, 2008. **15**(9): p. 939-47.
26. Rahmanpour, R. and T.D. Bugg, *Assembly in vitro of Rhodococcus jostii RHA1 encapsulin and peroxidase DypB to form a nanocompartment.* FEBS J, 2013. **280**(9): p. 2097-104.
27. Lowry, O.H., et al., *Protein measurement with the Folin phenol reagent.* J Biol Chem, 1951. **193**(1): p. 265-75.

28. Owens, C.P., et al., *Characterization of heme ligation properties of Rv0203, a secreted heme binding protein involved in Mycobacterium tuberculosis heme uptake.* Biochemistry, 2012. **51**(7): p. 1518-31.
29. Turlin, E., et al., *Staphylococcus aureus FepA and FepB proteins drive heme iron utilization in Escherichia coli.* PLoS One, 2013. **8**(2): p. e56529.
30. Dawson, R.M.C., *Data for biochemical research*. 3rd ed. 1986, Oxford: Clarendon Press. xii, 580 p.
31. Sil, S., et al., *Protoporphyrin IX-induced structural and functional changes in human red blood cells, haemoglobin and myoglobin.* J Biosci, 2004. **29**(3): p. 281-91.
32. Ogola, H.J., et al., *Molecular characterization of a novel peroxidase from the cyanobacterium Anabaena sp. strain PCC 7120.* Appl Environ Microbiol, 2009. **75**(23): p. 7509-18.
33. Bannantine, J.P. and J.R. Stabel, *Identification of two Mycobacterium avium subspecies paratuberculosis gene products differentially recognised by sera from rabbits immunised with live mycobacteria but not heat-killed mycobacteria.* J Med Microbiol, 2001. **50**(9): p. 795-804.
34. Lee, B.Y., S.A. Hefta, and P.J. Brennan, *Characterization of the major membrane protein of virulent Mycobacterium tuberculosis.* Infect Immun, 1992. **60**(5): p. 2066-74.
35. Tullius, M.V., et al., *Discovery and characterization of a unique mycobacterial heme acquisition system.* Proc Natl Acad Sci U S A, 2011. **108**(12): p. 5051-6.
36. Brown, M.E., T. Barros, and M.C. Chang, *Identification and characterization of a multifunctional dye peroxidase from a lignin-reactive bacterium.* ACS Chem Biol, 2012. **7**(12): p. 2074-81.
37. Hargrove, M.S., et al., *His64(E7)-->Tyr apomyoglobin as a reagent for measuring rates of hemin dissociation.* J Biol Chem, 1994. **269**(6): p. 4207-14.
38. Hofbauer, S., et al., *Structure and heme-binding properties of HemQ (chlorite dismutase-like protein) from Listeria monocytogenes.* Arch Biochem Biophys, 2015.
39. Dunford, H.B., *Heme peroxidases*. 1999: John Wiley.
40. Jones, C.M. and M. Niederweis, *Mycobacterium tuberculosis can utilize heme as an iron source.* J Bacteriol, 2011. **193**(7): p. 1767-70.
41. Phugare, S.S., et al., *A study on significant microbial interaction leading to decolorization and degradation of textile dye Rubine 3GP.* J Basic Microbiol, 2011. **51**(5): p. 499-514.
42. Yi, X., et al., *Replacement of the proximal heme thiolate ligand in chloroperoxidase with a histidine residue.* Proc Natl Acad Sci U S A, 1999. **96**(22): p. 12412-7.
43. Strittmatter, E., et al., *First crystal structure of a fungal high-redox potential dye-decolorizing peroxidase: substrate interaction sites and long-range electron transfer.* J Biol Chem, 2013. **288**(6): p. 4095-102.
44. Singh, R., et al., *Distal heme pocket residues of B-type dye-decolorizing peroxidase: arginine but not aspartate is essential for peroxidase activity.* J Biol Chem, 2012. **287**(13): p. 10623-30.

CHAPTER 3

CHARACTERIZATION OF A *Mycobacterium tuberculosis* NANOCOMPARTMENT AND ITS POTENTIAL CARGO PROTEINS

This project was performed in collaboration with members of the Fitzpatrick laboratory at the Salk Institute.

ABSTRACT

Mycobacterium tuberculosis (Mtb) has evolved various mechanisms by which the bacterium can maintain homeostasis under numerous environmental assaults generated by the host immune response. Mtb harbors enzymes involved in the oxidative stress response that aid in survival during the production of reactive oxygen species in activated macrophages. Previous studies have shown that a dye-decolorizing peroxidase (DyP) is encapsulated by a bacterial nanocompartment, encapsulin (Enc), whereby packaged DyP interacts with Enc via a unique C-terminal extension. Mtb also harbors an encapsulin homolog (CFP-29, Mt-Enc), within an operon with Mtb DyP (Mt-DyP), that contains a C-terminal extension. Together these observations suggest that Mt-DyP (a heme-peroxidase) interacts with Mt-Enc. Like Mt-DyP, Mtb iron-storage ferritin protein, Mt-BfrB, and an Mtb protein involved in folate biosynthesis, 7,8-dihydroneopterin aldolase (Mt-FolB), have C-terminal tails that could also interact with Mt-Enc. For the first time, we show by co-purification and electron microscopy that mycobacteria via Mt-Enc can

encapsulate Mt-DyP, Mt-BfrB, and Mt-FolB. Functional studies of free or encapsulated enzyme demonstrate that cargo proteins retain enzymatic activity within the Mt-Enc nanocompartment. Mt-DyP, Mt-FolB and Mt-BfrB all have antioxidant properties, suggesting that if these proteins are encapsulated by Mt-Enc, then this nanocage may play a role in the Mtb oxidative stress response. This report provides initial structural and biochemical clues regarding the molecular mechanisms that utilize compartmentalization, which may aid the mycobacterial cell in detoxification of the local environment to ensure long-term survival.

INTRODUCTION

The etiologic agent of tuberculosis, *Mycobacterium tuberculosis*, infects ~ 11.1 million people per year, resulting in over 1.3 million deaths worldwide [4]. With the emergence of *M. tuberculosis* strains resistant to the major anti-tuberculosis therapies, the discovery of novel drug targets is imperative to triumph over tuberculosis infection. Deciphering the complicated biology of *M. tuberculosis*, which has both active and latent forms, will aid in the discovery of new therapeutics to combat this highly successful pathogen.

In mammalian cells, exosomes are secreted vesicles proposed to be involved in mediating the adaptive immune response [5, 6]. Mycobacterial proteins have been discovered in exosomes of cells infected with *M. tuberculosis*. One mycobacterial protein identified within exosomes was a 29-kDa culture filtrate protein (CFP-29, Rv0798c) [7], reported to be secreted [8], although lacking a signal peptide. CFP-29 has 58% amino acid identity to *Brevibacterium linens* M18 linocin protein (also known as encapsulin, Bl-

Enc) [8], implying that CFP-29 is an *M. tuberculosis* encapsulin homolog (Mt-Enc). Throughout the remainder of this work, CFP-29 (Rv0798c) will be referred to as Mt-Enc.

Additionally, the gene encoding Bl-Enc is within a two-gene operon containing a gene encoding for a dye-decolorizing peroxidase (Bl-DyP) with a ~ 25-amino acid C-terminal extension (**Figure 3.1A**). The crystal structure of a close homolog of Bl-Enc, *Thermotoga maritima* encapsulin (Tm-Enc), has been determined, revealing a 60-subunit icosahedral nanocompartment [1]. Electron microscopy (EM) shows that co-expression of Bl-DyP and Bl-Enc results in encapsulation of Bl-DyP within the Bl-Enc nanocage, where encapsulation of Bl-DyP is abolished upon truncation of its C-terminal extension [1]. Finally, a search across bacterial genomes demonstrates that genes encoding for Enc are usually in two-gene operons with a gene encoding for either DyP or ferritin-like proteins, all with predominately hydrophobic C-terminal extensions (**Figure 3.1A**) [1]. These observations suggest that Enc proteins compartmentalize DyP or ferritin-like proteins via their C-terminal tails. Notably, both DyPs and ferritin proteins possess antioxidant properties [9, 10].

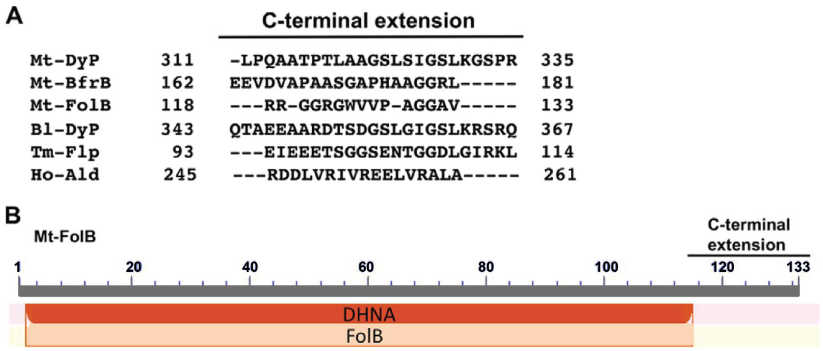

Figure 3.1: C-terminal extensions of potential Mt-Enc protein cargos. A) C-terminal extension residue sequences of Mt-DyP, Mt-BfrB, and Mt-FolB beyond what is seen in homologs that are not encapsulated as well as *B. linens* DyP (Bl-DyP [1]), *T. maritima* ferritin-like protein (Tm-Flp [1]), and a putative *H. ochraceum* aldolase (Ho-Ald [2]) thought to be encapsulated by its C-terminal tail to BMCs. B*)* the *top numbered gray line* represents Mt-FolB residues, and the *brown* and *light orange boxes* represent conserved 7,8-dihydroneopterin aldolase (*DHNA*) and FolB domains, respectively, from various organisms generated by the Conserved Domain Database [3].

Mt-Enc, the only homolog of Enc in *M. tuberculosis*, is part of a two-gene operon with the gene for a DyP-like peroxidase (Mt-DyP, Rv0799c), where Mt-DyP contains a C-terminal extension similar to Bl-DyP (**Figure 3.1A**). DyP-type peroxidases are a novel family of fungal and bacterial heme-dependent peroxidases that can oxidize lignin and anthraquinone dyes in the presence of H_2O_2 [11, 12]. However, the *Escherichia coli* DyP ortholog, Ec-YfeX, was proposed to possess deferrochelatase activity, where Ec-YfeX can catalyze the extraction of iron from heme without cleavage of the tetrapyrrole ring [13], although in a parallel study, Ec-YfeX was demonstrated to be a heme-dependent peroxidase with no detectable deferrochelatase activity [14]; however we have identified Mt-DyP as a peroxidase (see Chapter 2). As a heme-dependent peroxidase, Mt-DyP may protect *M. tuberculosis* against host oxidative assault by H_2O_2.

Notably, two other *M. tuberculosis* proteins have aliphatic C-terminal extensions beyond what is seen in other bacterial homologs and similar to Mt-DyP, suggesting that these proteins may also be cargo proteins for Mt-Enc. Mt-BfrB, one of the two iron storage ferritin proteins within *M. tuberculosis*, has a C-terminal tail (**Figure 3.1A**) and has displayed antioxidant properties [15, 16]. In addition, an enzyme involved in folate metabolism, 7,8-dihydroneopterin aldolase (Mt-FolB), also has an aliphatic C-terminal extension (**Figure 3.1B**). In a previous study, Mt-FolB structure determination shows that this C-terminal extension is disordered, and Mt-FolB enzymatic activity is independent of the C-terminal tail. Furthermore, its substrate has also been implicated in *M. tuberculosis* resistance to oxidative stress [17-19].

Herein we describe the Mt-Enc encapsulation of three different *M. tuberculosis* enzymes, Mt-DyP, Mt-BfrB, and Mt-FolB, all of which possess antioxidant properties.

We show by EM that Mt-Enc can encapsulate *M. tuberculosis* cargo enzymes and show by biochemical analyses that the enzymes remain active within the Mt-Enc nanocompartment. To our knowledge, this study provides the initial report into mycobacterial compartmentalization of enzymes, whereby protein encapsulation by Mt-Enc may function to combat oxidative stress within the human host.

MATERIALS AND METHODS

Cloning.

The *M. tuberculosis* genes encoding proteins Mt-BfrA (*Rv1876*), Mt-BfrB (*Rv3841*), Mt-DyP (*Rv0799c*), Mt-Enc (*Rv0798c*), and Mt-FolB (*Rv3607c*) were PCR-amplified from *M. tuberculosis* H37Rv genomic DNA using the KOD HotStart Polymerase Kit (Novagen) with 5′ and 3′ primers (MWG Operon, **Table 3.1**) containing specific restriction sites to generate either a C-terminal polyhistidine tag (His tag) or a tagless protein. PCR products were ligated into pCR-BluntII-TOPO (Invitrogen) and then transformed into *E. coli* OneShot TOP10 cells (Invitrogen). Double digestions with specific restriction enzymes were performed from each vector of choice. Excised genes were ligated into the appropriate linearized pET vector and transformed into *E. coli* BL21-Gold (DE3) cells (Novagen). Each final construct was verified by DNA sequencing using T7 promoter and reverse primers (Laguna Scientific).

Table 3.1: Primers used in this study.

Gene	Construct	Primer Sequence
Rv0798c (Enc)	pET22::Enc	FWD: 5' GGCATATGAACAATCTCTACCGCGATTTGGCACCGGTCACCGAA G REV: 5' GGCTCGAGGTGGCTGAGCGCGACCGACGCCTCG
Rv0799c (DyP)	pET22::DyP	FWD: 5' GGCATATGGTGGCTGTGCCTGCTGTCTCTCCGCAGC REV: 5' GGCTCGAGTCGGGGGCTTCCTTTCAAGCTGCCGATCGATAGC
*Rv0799c*ΔC -term (DyPΔC-term)	pET28::DyP DyPΔC-term	FWD: 5' CCACCGCCCCTAACGCAGGCGGCGAC REV: 5' GTCGCCGCCTGCGTTAGGGGCGGTGG
Rv0798c-*Rv0799c* (Enc-DyP)	pET22::Enc-DyP	DyP FWD and Enc REV
Rv3607 (FolB)	pET28::FolB	FWD: 5' GCATATGGCTGACCGAATCGAACTGCGCGGCCTGACCGTGCATG REV: 5' GCTCGAGTCATCATACCGCGCCGCCCGCCGGG
Rv3841 (BfrB)	pET28::BfrB	FWD: 5' CCATATGACAGAATACGAAGGGCCTAAG REV: 5' GCTCGAGCTACTAGAGGCGGCCCCCG
Rv3841 (BfrBΔ161-187)	pET28:: BfrBΔ161-187	FWD: 5' GCCATATGACAGAATACGAAGGGCCTAAGACAAAATTCC REV: 5' GCCTCGAGCACATCCACTTCACGTGCG
Rv1876 (BfrA)	pET28::BfrA	FWD: 5' GCAAGCTTTCATCAGGTCGGTGGGCGAGAGACGCACTGC REV: 5' GCCTCGAGTCATCACACATCCAGTTCACGTGCG

Overexpression and purificaton of Mtb Proteins.

Proteins were expressed from plasmids alone or in combination with one another, using *E. coli* BL21 Gold (DE3) cells. The following concentrations of antibiotics to the media were added where appropriate: kanamycin (30 μg/ml) and ampicillin (50 μg/ml). Cells harboring expression vector(s) were grown aerobically at 37 °C in LB medium containing the appropriate antibiotic(s). Protein expression was induced at A_{600} ~ 0.8 by the addition of isopropyl-β-d-thiogalactopyranoside (1 mM). Cells were harvested after 4 h of induction (apart from Mt-Enc and proteins in complex with Mt-Enc, which were induced overnight at 18 °C) by centrifugation at 5100 × *g* for 20 min. Additionally, only

where indicated cultures expressing Mt-Enc·DyP were supplemented with 0.7 mM δ-aminolevulinic acid (δ-ALA) at the time of induction.

Harvested cell pellets were resuspended in 20 ml of Buffer A (50 mM Tris-HCl, pH 7.4, 350 mM NaCl, 10 mM imidazole, and 10% glycerol), followed by the addition of PMSF and hen egg white lysozyme. Pellets containing Mt-Enc and proteins in complex with Mt-Enc were resuspended in Buffer A containing 0.2% lauryldimethylamine oxide for solubilization. Cells were disrupted by sonication, clarified by centrifugation at 18,000 × *g* for 30 min at 4 °C, and syringe-filtered (1μm pore size) for removal of cell debris. The clarified cell lysate was then loaded onto a 5-ml Ni^{2+}-charged HisTrap column (GE Healthcare) pre-equilibrated with Buffer A. The protein(s) was eluted with a linear gradient of 10–500 mM imidazole (100 ml). Proteins that co-purify with Mt-Enc elute at a higher imidazole concentration (>250 mM imidazole). Eluted fractions were collected, analyzed by SDS-PAGE, and concentrated using Amicon concentrators with the appropriate molecular weight cut-off (Millipore, Bedford, MA). Mt-DyP and Mt-BfrB were further purified on a Superdex 200 HiLoad 16/60 gel filtration chromatography column (GE Healthcare) utilizing 50 mM Tris-HCl, pH 7.4, and 150 mM NaCl, and their respective molecular weights were calculated using molecular weight standards (Bio-Rad). Mt-BfrB was dialyzed into 20 mM Tris-HCl, pH 8.0, and 10 mm NaCl and purified by ion exchange (5 ml; HiTrap™ Q HP, GE Healthcare). The protein was eluted with a linear gradient of 0.01–1.0 M NaCl (100 ml); the purified protein eluted between 400 and 600 mM NaCl. Protein concentration was determined by UV-visible spectroscopy, utilizing molar extinction coefficients at 280 nm as predicted

by the program Protein Calculator (Scripps) or as determined by either a modified Lowry [20] or Bradford [21] assay using bovine serum albumin as a standard.

Electron microscopy.

Samples were negatively stained as follows. Ultrathin carbon and silicon monoxide grids (Ted Pella catalogue nos. 01824 and 01829, respectively) were prepared by removing the Formvar backing by dipping the grids in 100% chloroform for 10 s and glow-discharging them for 30 s (Leica SCD500, Leica, Vienna). Samples were diluted 1:10 in buffer, and 5 µl of this was placed on a grid and immediately washed twice in double-distilled H_2O, followed by a 1-min incubation in 1% uranyl acetate. Excess stain was then aspirated off, and the grid was allowed to dry at room temperature. Following negative staining, the grids were imaged on a Zeiss Libra 120 PLUS EF-TEM (Carl Zeiss).

Kinetics of enzyme activity within Mt-Enc.

Mt-DyP: All enzymatic reactions were performed in 100 µl reactions. Peroxidase activity utilizing guaiacol as a substrate was monitored over 20 min as a change in absorbance at 470 nm, indicative of the production of tetraguaiacol. Two µM holo-Mt-DyP (reconstituted with heme to obtain a 1:1 heme/protein molar ratio), apo-Mt-DyP, Mt-Enc, and Mt-Enc·DyP (grown in 0.7 mM δ-ALA, where the heme bound Mt-DyP concentration is estimated from Soret peak intensity utilizing UV-visible spectroscopy, $\varepsilon_{411} \sim 125{,}213 \pm 24717\ M^{-1}\ cm^{-1}$) were ass ayed in the presence of 10 mM H_2O_2 (Sigma-Aldrich) and 10 mM guaiacol (Sigma-Aldrich) [22] in 100 mM sodium citrate at pH 4. Reactions were initiated upon the addition of H_2O_2 and were monitored over a 12

min period, observing the change in absorbance at 470 nm by UV-visible spectrometry (DU-800 spectrophotometer, Beckman-Coulter) at 25 °C. The rate of product formation was calculated using the extinction coefficient of 26,600 M^{-1} cm^{-1} of tetraguaiacol at 470 nm [23].

Further analysis of the activity of Mt-DyP was carried out by utilization of 2,2′-azino-bis(3-ethylbenzothiazoline-6-sulfonic acid (ABTS) as a substrate, as described previously [24]. Briefly, 250 μM ABTS was added to a reaction mixture containing 1.25 μM Mt-DyP, Mt-Enc, or Mt-Enc·DyP (grown in 0.7 mM δ-ALA, where the heme-bound Mt-DyP concentration is estimated from Soret peak intensity) in 50 mM HEPES at pH 5.5. The reaction was initiated upon the addition of 10 μM H_2O_2 and monitored for 10 min at 420 nm by UV-visible spectrometry (DU-800 spectrophotometer, Beckman-Coulter) at 25 °C. The rate of product formation was calculated using the extinction coefficient of 36,000 M^{-1} cm^{-1} of the ABTS radical cation at 420 nm [25].

Mt-BfrB

Ferroxidase activity was measured following an amended protocol described by Khare *et al.* [26], whereby activity was monitored at 320 nm with some differences. Briefly, purified Mt-BfrB apo-Mt-BfrB, apo-Mt-BfrBΔ167–181, and Mt-Enc-BfrB proteins were dialyzed against 100 mM MES, pH 6.5, and 100 mM KCl. Ferrous ammonium sulfate (FAS) was dissolved in degassed 0.1 m HCl for a 5 mM stock solution of FAS diluted in 100 mM MES, pH 6.5, and 100 mM KCl. FAS was added to a 1 μM concentration of a 24-subunit assembly of Mt-BfrB at 25 °C to a final FAS concentration of 100 μM.

Ferrous iron oxidation was monitored by observing the change in absorbance at 320 nm over 10 min, using a DU-800 spectrophotometer (Beckman-Coulter).

Mt-FolB

Mt-FolB aldolase activity was monitored by fluorimetry, as described previously [18]. Briefly, Mt-FolB, Mt-Enc·FolB, or Mt-Enc (1 μM) was assayed in the presence of 80 μM 7,8-dihydroneopterin (Sigma-Aldrich) in 50 mM Tris-HCl, pH 8.0, and 150 mM NaCl. The production of 6-hydroxymethyl-7,8-dihydropterin was determined by fluorimetry (Hitachi F-4500 fluorescence spectrometer), where the λ_{ex} was 430 nm and the change in λ_{em} at 524 nm was monitored. Data were acquired over 35 min.

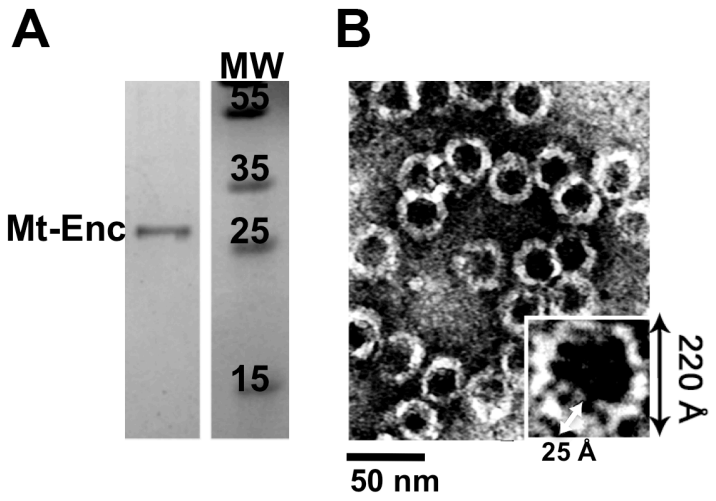

Figure 3.2: Mt-Enc assembles into a nanoparticle. A) Elution fractions of Mt-Enc$_{His}$ purification. B) EM image of purified Mt-EncHis; the **inset** shows a single particle with a double-sided arrow indicating ~220 Å Mt-Enc particle diameter and thickness of ~ 25 Å.

RESULTS

Mt-Enc is a Nanocompartment.

Mt-Enc is the only Enc homolog in the *M. tuberculosis* proteome. Recombinant Mt-Enc was purified to near homogeneity, as shown by SDS-PAGE (**Figure 3.2A**). EM analysis revealed that purified Mt-Enc forms a shell-like nanocompartment with a diameter of ~220 Å and shell thickness of ~25 Å, as determined using ImageJ [27] (**Figure 3.2B**). The shell diameter and thickness of Mt-Enc are in good agreement with the previously solved structure of the 60-subunit Tm-Enc nanocompartment [1], suggesting that Mt-Enc also forms a 60-subunit assembly.

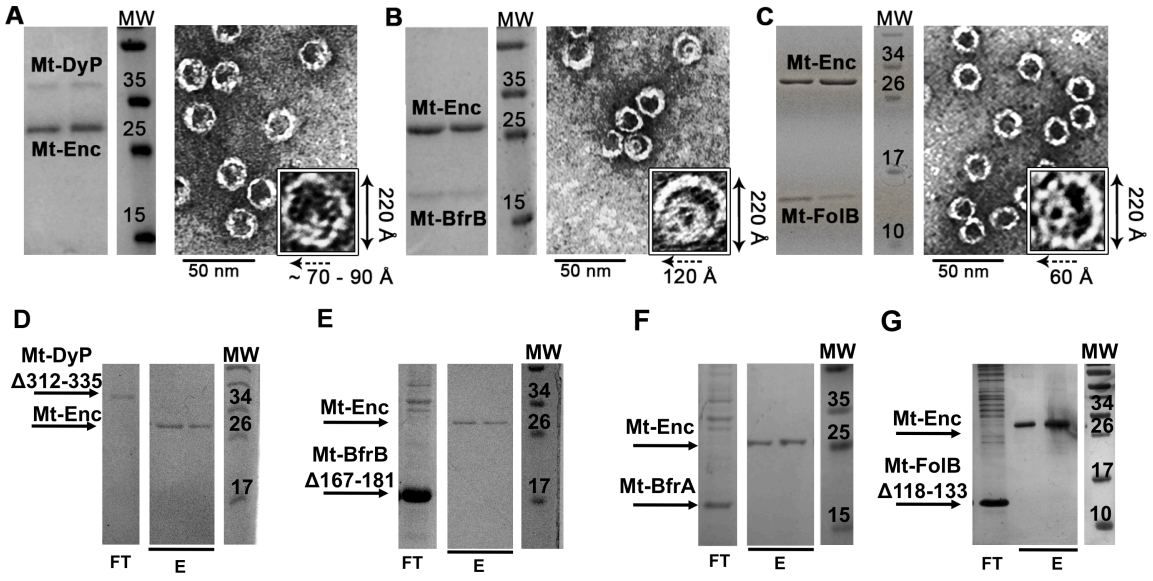

Figure 3.3: Characterization of Mt-Enc with cargo proteins. Mt-Enc and co-elution of putative interacting *M. tuberculosis* protein: Mt-DyP (~37 kDa) A); Mt-BfrB (~20 kDa) B); or Mt-FolB (~14 kDa). C) *Left*, SDS-PAGE of eluted fractions with a molecular weight marker (MW). *Right*, EM images with an *inset* of a single magnified Mt-Enc particle with cargo protein with a *double-sided arrow* indicating 220-Å Mt-Enc particle diameter. *Dashed arrow*, diameter of cargo protein. D–G) demonstrate that proteins lacking a C-terminal extension do not co-purify with Mt-Enc. Mt-Enc does not co-elute with Mt-DyPΔ312–335 (D), Mt-BfrBΔ167–181 (E), Mt-BfrA (F), or Mt-FolBΔ118–133 G). Shown are flow-through (*FT*) and elution (*E*) fractions analyzed by SDS-PAGE.

Co-purification of Mt-Enc and Cargo Proteins.

Mt-DyP is found upstream of Mt-Enc and harbors a C-terminal extension, similar to that observed in other organisms that also contain Enc and DyP in a two-gene operon. Thus, we propose that Mt-DyP interacts with Mt-Enc [1, 24]. The two-gene operon, whose gene products are Mt-Enc and Mt-DyP, was cloned into an expression vector so that Mt-Enc has a C-terminal His tag and Mt-DyP does not. Purification of overexpressed Mt-DyP and Mt-Enc by affinity chromatography demonstrated that Mt-DyP and Mt-Enc co-elute (**Figure 3.3A**) suggesting that an Mt-Enc·DyP protein complex has formed. EM analysis was utilized to visualize the Mt-Enc·DyP complex, revealing a large sphere similar to Mt-Enc (**Figure 3.2B**) containing a density that we presume to be Mt-DyP (**Figure 3.3A)** because no discernable density was observed within Mt-Enc alone (**Figure 3.2B**). The diameter of the Mt-Enc nanocompartment is ~220 Å, and its cargo, Mt-DyP, has a diameter ranging from 70 to 90 Å, as measured by ImageJ [27]. The oligomeric state of encapsulated Mt-DyP is unknown. To test whether the C-terminal extension of Mt-DyP is required to target Mt-DyP to be encapsulated by Mt-Enc, we deleted the last 24 C-terminal amino acids of Mt-DyP (Mt-DyPΔ312–335). Mt-Enc and Mt-DyPΔ312–335 were then co-expressed, and upon purification, Mt-DyPΔ312–335 was observed in the flow-through, and only Mt-Enc was eluted, as seen by SDS-PAGE (**Figure 3.3D**). These results suggest that Mt-Enc encapsulates Mt-DyP via its C-terminal tail.

We previously carried out studies of Mt-BfrB (ferritin [28]) and Mt-FolB (second enzyme in folate biosynthesis [18]) and noted that both mycobacterial proteins have extra ~20-amino acid C-terminal extensions (**Figure 3.1**) in comparison with most of their other bacterial homologs. Both Mt-BfrB [15] and Mt-FolB [18] are known to have

antioxidant properties, along with DyPs [29, 30]. Due to the observation that ferritin-like proteins with C-terminal extensions are contained within operons with Enc and are proposed to interact [1], we hypothesized that Mt-BfrB and Mt-FolB may also interact with Mt-Enc through their C-terminal extensions.

Mt-Enc encapsulates fully assembled Mt-BfrB, where Mt-BfrB forms a 24-subunit icosahedral shell [28], in contrast to Mt-Enc, which forms a 60-subunit icosahedral shell. We co-expressed Mt-Enc with a C-terminal His tag together with full-length Mt-BfrB. Mt-Enc and Mt-BfrB co-eluted from an affinity column (**Figure 3.3B**). Visualization of the Mt-Enc·BfrB complex by EM clearly shows a smaller spherical shell within the interior of a larger shell. The larger shell has a diameter of ~220 Å, similar to Mt-Enc alone, and the inner shell has a diameter of ~120 Å, similar to that of Mt-BfrB (**Figure 3.3B**) [28], suggesting that Mt-Enc encapsulates fully assembled 24-subunit Mt-BfrB (**Figure 3.3B**). Because it is proposed that Mt-Enc encapsulates its target protein via a C-terminal extension, Mt-BfrB was truncated to remove 15 C-terminal residues (Mt-BfrBΔ167–181). First, we showed by analytical gel filtration analysis that Mt-BfrBΔ167–181 has an elution profile similar to that of Mt-BfrB, suggesting that it retains its 24-subunit assembly in the absence of its C-terminal extension (data not shown). Second, upon co-expression of Mt-Enc and Mt-BfrBΔ167–181 followed by affinity chromatography purification, Mt-BfrBΔ167–181 was observed in the flow-through and did not co-elute with Mt-Enc (**Figure 3.3E**). The other *M. tuberculosis* ferritin homolog, Mt-BfrA, also self-assembles into a 24-subunit nanocage [31, 32]; however, Mt-BfrA does not possess a C-terminal extension. When Mt-BfrA was co-expressed with Mt-Enc followed by affinity chromatography purification, it did not co-elute with Mt-Enc

(**Figure 3.3F**). These observations validate the importance of the C-terminal extension in potentiating Mt-BfrB binding to the interior of the Mt-Enc compartment.

Similarly, Mt-FolB contains an extended C terminus composed of aliphatic amino acids (**Figure 3.1B)** thus, it is possible that Mt-FolB might interact with Mt-Enc via its C-terminal tail. Co-expression of Mt-FolB and Mt-Enc resulted in co-purification of Mt-FolB with Mt-Enc. EM images of the Mt-Enc·FolB complex showed fully assembled Mt-Enc that contained a potential protein cargo (**Figure 3.3C**). The diameter of the cargo is ~ 60 Å, similar to the diameter of either the tetrameric or octameric forms of Mt-FolB [18], suggesting that Mt-Enc also encapsulates Mt-FolB. Upon deletion of the C-terminal tail of Mt-FolB, where Mt-FolBΔ118–133 retains its enzyme activity [18], co-expression with Mt-Enc followed by purification resulted in the elution of Mt-Enc alone (**Figure 3.3C**). These results suggest that Mt-FolB can be encapsulated by Mt-Enc via its C-terminal extension.

EM images show that nearly all Mt-Enc nanocompartments are occupied by Mt-DyP (~ 80%), which was co-expressed from a single plasmid. Fewer Mt-Enc assemblies were occupied with Mt-BfrB or Mt-FolB (<30%); however, Mt-Enc and Mt-BfrB or Mt-FolB was co-expressed from separate plasmids. Reduced encapsulation is not surprising because heterologous overexpression of mycobacterial proteins in *E. coli* is not representative of the coupled translation or tight regulation in *M. tuberculosis*.

Analysis of Enzymatic Activity of Cargo Proteins within Mt-Enc.

We have demonstrated that Mt-DyP has heme-dependent peroxidase activity that may utilize guaiacol or ABTS as substrates. Previously, it has been shown that Mt-BfrB has

ferroxidase activity [26], and Mt-FolB has aldolase activity even in its C-terminal truncated form [18]. Furthermore, we have shown that co-expression of each of these proteins with Mt-Enc results in encapsulation of each cargo protein via its C-terminal extension (**Figure3.3**). Next, we examined whether Mt-DyP, Mt-BfrB, and Mt-FolB retain their activities when encapsulated by Mt-Enc.

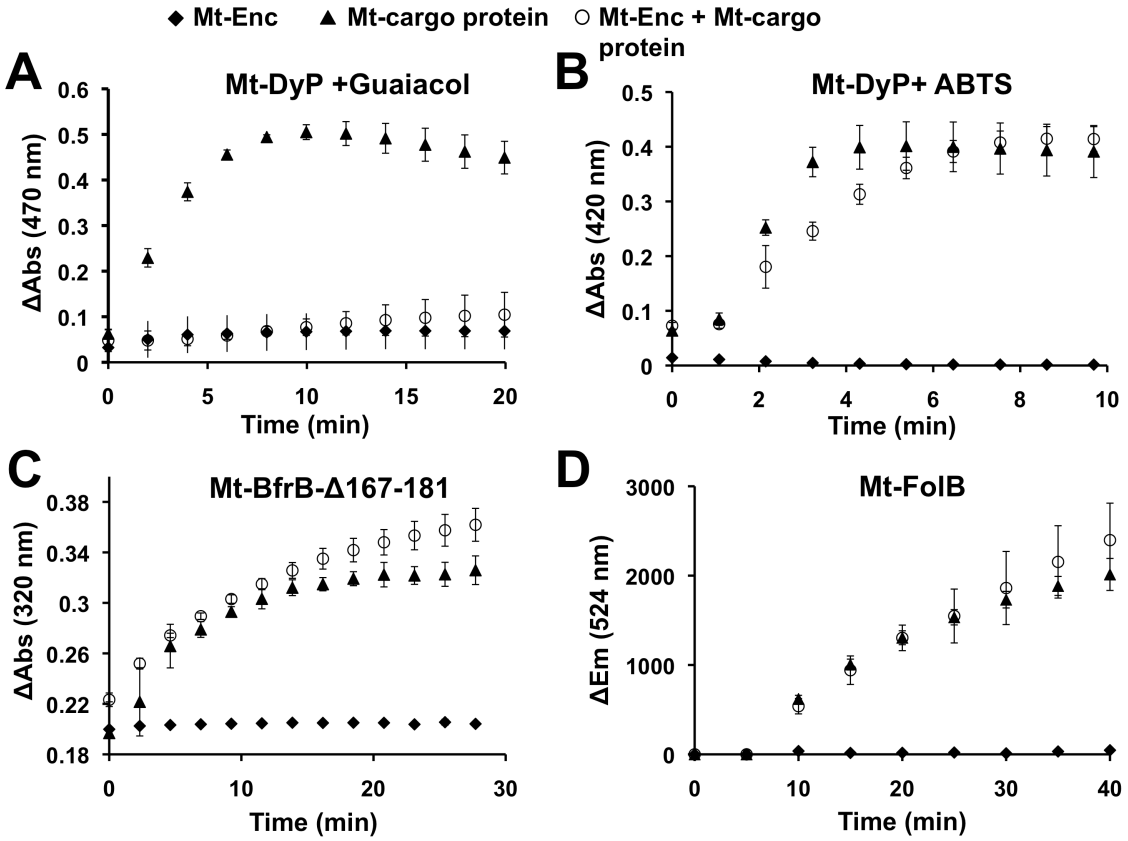

Figure 3.4: Analysis of enzymatic activity of cargo protein within Mt-Enc. A*)* Mt-DyP peroxidase activity utilizing guaiacol as a substrate in the presence of H_2O_2 was monitored by observing the production of tetraguaiacol at 470 nm. B), Mt-DyP peroxidase activity utilizing ABTS as a substrate in the presence of H_2O_2 was monitored by observing the change in absorbance at 420 nm. C) Mt-BfrB ferroxidase activity was monitored by observing the oxidation of ferrous iron to ferric iron at 320 nm. The encapsulated BfrB was compared with the rate of the Mt-BfrBΔ167–181. *D*, Mt-FolB enzyme activity was monitored, utilizing 7,8-dihydroneopterin as a substrate, by observing the change in emission at 524 nm. *Error bars*, S.E.

Mt-DyP.

To ensure that C-terminal truncation of Mt-DyP did not result in loss of activity, we demonstrated that heme-bound Mt-DyPΔ312–335 retained its peroxidase activity with both guaiacol and ABTS as substrates (data not shown). In an attempt to produce encapsulated holo-Mt-DyP, Mt-Enc and Mt-DyP were co-expressed in the presence of δ-ALA. After purification of the Mt-Enc·DyP complex, the heme concentration within the complex (and presumably bound to Mt-DyP) was quantified using UV-visible spectroscopy (to determine heme concentration) and SDS-polyacrylamide gel band intensity analysis (to determine Mt-DyP concentration) using ImageJ [27]. Utilizing either guaiacol or ABTS as substrates in the presence of H_2O_2, we assessed the peroxidase activity of holo-Mt-DyP within Mt-Enc. With guaiacol as a substrate, Mt-Enc·DyP showed minimal peroxidase activity compared with unencapsulated holo-Mt-DyP (**Figure 3.4A**). In contrast, encapsulated Mt-DyP retained peroxidase activity with ABTS as a substrate similar to unencapsulated holo-Mt-DyP (**Figure 3.4B**), although we observed a slightly decreased rate. Within Mt-Enc, we could not monitor the oligomeric state of holo-Mt-DyP, and we have shown that holo-Mt-DyP was active as a tetramer (**Figures 2.3C and 2.6C and D**); thus, the observed decrease in rate of activity of encapsulated Mt-DyP may be due to the presence of various heme-bound Mt-DyP oligomeric states. In light of these results, we have demonstrated that Mt-DyP retains heme peroxidase activity while encapsulated; however, it appears to have a more limited substrate set than unencapsulated holo-Mt-DyP.

Mt-BfrB

To determine the activity of encapsulated Mt-BfrB, ferroxidase activity of the Mt-Enc·BfrB complex was measured at 320 nm to monitor iron core formation [33]. To determine if the Mt-Enc·BfrB complex maintained ferroxidase activity, we quantified the amount of encapsulated Mt-BfrB by densitometry measurements of SDS-polyacrylamide gel bands using ImageJ [27]. Upon the addition of ferrous iron to Mt-BfrB alone, we observed ferroxidase activity. The specific activity of Mt-BfrB oxidizing Fe(II) is 0.073 ± 0.006 μmol min^{-1} mg^{-1}, where the activity is similar to a previous study of Mt-BfrB [26]. Truncation of the C-terminal tail of Mt-BfrB, Mt-BfrBΔ167–181, resulted in fully assembled ferritin (data not shown) with a slightly reduced oxidation rate of Fe(II) (data not shown), as also observed by Khare *et al.[26]*. In contrast, Mt-Enc alone was not capable of oxidizing Fe(II) (**Figure 3.4C)**. Mt-BfrB encapsulated in Mt-Enc retained its ferroxidase activity with a rate similar to that of Mt-BfrBΔ167–181 (**Figure 3.4C**). Reduction in the rate of Fe(II) oxidation by encapsulated Mt-BfrB similar to that of Mt-BfrBΔ167–181 suggests that the C-terminal tail of Mt-BfrB is probably involved in binding to the interior of Mt-Enc, rendering it unavailable to enhance ferroxidase activity to the level observed for unencapsulated full-length Mt-BfrB.

Mt-FolB

Aldolase activity of Mt-FolB within Mt-Enc was assayed by measuring the production of 6-hydroxymethyl-7,8-dihydropterin as a change in fluorescence emission at 524 nm. The concentration of encapsulated Mt-FolB was estimated by densitometry measurements from SDS-PAGE analysis using ImageJ [27]. The level of aldolase activity achieved

upon the addition of a 1.5-fold molar excess of encapsulated Mt-FolB was similar to that of unencapsulated Mt-FolB (**Figure 3.4D**). The specific activity of Mt-FolB is 0.18 ± 0.02 μmol min^{-1} mg^{-1}, similar to our previous study of Mt-FolB [18]. Mt-Enc alone did not exhibit aldolase activity. These results suggest that Mt-FolB retains aldolase activity while encapsulated.

DISCUSSION

Encapsulins are a recently discovered family of conserved prokaryotic proteinaceous nanocompartments, first described by Sutter *et al.* [1]. Enc cargo proteins have C-terminal tails required for their encapsulation, where Enc and cargo proteins are usually found in a two-gene operon, suggesting tight translational coupling. We have demonstrated that *M. tuberculosis* has an Enc nanocompartment similar in diameter to that of Tm-Enc [1], implying that it is also a 60-subunit icosahedral shell. Mt-Enc exists in a two-gene operon, where the upstream Mt-DyP harbors a C-terminal extension. Co-expression of Mt-DyP and Mt-Enc results in encapsulation of Mt-DyP by Mt-Enc, an interaction mediated by the C-terminal tail of Mt-DyP. Although it is possible that Mt-Enc and Mt-DyP are expressed as a polycistronic mRNA transcript, the transcriptional and translational regulation of these proteins in *M. tuberculosis* is unknown. Additionally, we have demonstrated that Mt-BfrB and Mt-FolB are also encapsulated by Mt-Enc, via an interaction mediated by the C-terminal extension of each protein cargo, where loss of this C-terminal tail abolishes the interaction with Mt-Enc and their subsequent encapsulation. These results show that Mt-Enc could potentially compartmentalize other mycobacterial enzymes that bear a C-terminal tail.

Compartmentalization is important for organization, increasing the local concentration of functionally related enzymes and confining unstable reaction intermediates within the cell. Many bacteria produce protein-based microcompartments (MCPs) that are larger than encapsulin nanocages but also encapsulate various cargo enzyme [34, 35]. In *Salmonella enterica*, MCPs are proposed to organize reactions as well as isolate "unstable" reaction intermediates that may form within the compartment [34], such as proteins involved in 1,2-propanediol utilization (Pdu) [36] and ethanolamine utilization (Eut) [34]. Indeed, a recent comparative genomics study predicts that in some organisms, MCPs may encapsulate bacterioferritin [37], similar to Mt-Enc. A common feature of MCP-compartmentalized enzymes is an additional short N-terminal sequence found to be necessary for targeting enzymes to the MCP interior [34, 38], akin to the cargo protein C-terminal extensions required for their encapsulation in Enc [1]. It could be surmised that engineering a C-terminal tail on an enzyme could potentiate encapsulation by Mt-Enc and possibly deliver and/or release cargo enzymes under conditions where the compartment disassembles.

This study and previous ones [1, 24] demonstrate that encapsulins from four different organisms self-assemble into 60-subunit nanocompartments with a diameter of ~220–240 Å and can accommodate multiple copies of a single cargo protein. Distant sequence homologs of Tm-Enc, *Pyrococcus furiosus* PfV, and the major caspid protein gp5 of the HK97 virus (18 and 9% sequence identity, respectively) are close structural homologs of Tm-Enc (root mean square deviations of 2.39 and 2.65 Å, respectively) [39]. All three proteins form icosahedral cages; however, Tm-Enc has $T = 1$ symmetry with 60 subunits/cage [1], PfV has $T = 3$ symmetry with 180 subunits/cage [39], and gp5 has $T =$

7 symmetry with 420 subunits/cage [40]. Thus, one can postulate that Enc homologs, such as Mt-Enc, may form higher symmetry icosahedral cages *in vivo* to allow for more than one cargo protein or even an entire reaction pathway to be encapsulated, as observed for MCPs [35].

Passage of substrates through encapsulin is currently unknown. Charged pores observed for Tm-Enc may mediate substrate specificity [1], regulating the flux of substrates in and out of encapsulin that are available to the protein cargo. This may also be the case with Mt-Enc and the differential activity of encapsulated Mt-DyP with various electron-donor substrates, where guaiacol is neutral and ABTS is negatively charged. Rhamanpour and Bugg [24] reported peroxidase activity of Rj-DypB when encapsulated within Rj-Enc with lignin as a substrate. This was an unexpected result because the expected pore size of Enc (<5 Å wide) is not large enough to accommodate the passage of lignin [24], suggesting that the Enc pore is dynamic. Indeed, utilizing dynamic light scattering, it was observed that native Rj-Enc has a diameter of 220 Å; however, after acidic disassembly followed by reassembly with Rj-DyP, the complex has a diameter of 310 Å [24]. The flexibility of the Enc monomer may allow Enc to form compartments consisting of more than 60 subunits or to expand its pore size to modulate substrate entry [1, 24].

Mycobacteria are resistant to assault by host oxidative metabolites [30], where the presence of iron storage proteins and peroxidases aids in mycobacterial viability. A recent study demonstrated that an *M. tuberculosis* double deletion mutant of *bfrA* and *bfrB* exhibited vulnerability to oxidative stress [16], and the Mtb∆*bfrB* mutant has greater susceptibility than the Mtb∆*bfrA* mutant [15, 16], highlighting the importance of Mt-

BfrB in *M. tuberculosis* survival during oxidative assault. Mt-BfrB has antioxidant properties by storing iron, thereby preventing the generation of hydroxyl radicals upon oxygen exposure. Furthermore, *in vivo*, it has been shown that Mt-BfrB is up-regulated under hypoxic conditions [41], during assault by NO donors [42], and during adaptation to the stationary phase [43]. We have demonstrated that Mt-DyP can function as a heme-dependent peroxidase; thus, Mt-DyP may aid in mycobacteria resistance to host oxidative assault [30]. This could potentially hold true for Mt-FolB. Mt-FolB is a 7,8-dihydroneopterin aldolase that is involved in folate biosynthesis, converting folate precursor 7,8-dihydroneopterin to 6-hydroxymethyl-7,8-dihydropterin. Mt-FolB displays cooperative dependence on 7,8-dihydroneopterin binding from an apotetrameric form to active substrate-bound octameric form [18]. Furthermore, its substrate possesses antioxidant properties [19, 44]. Thus, if *M. tuberculosis* encounters host oxidative stress, then 7,8-dihydroneopterin may function as an antioxidant maintaining a relatively low cellular concentration. However, when oxidative stress diminishes, the concentration of 7,8-dihydroneopterin would increase to the critical level to form active, octameric Mt-FolB, and *M. tuberculosis* would continue with its normal folate biosynthetic pathway [18]. Recently, an *in vitro* study showed that there were five Mt-FolB reaction products [45]. Another potential benefit of encapsulation of Mt-FolB could be to (*a*) prevent side products of the reaction forming and (*b*) allow 7,8-dihydroneopterin to remain "free" within the cytosol to act as an antioxidant. Thus, *M. tuberculosis* encapsulation of proteins may provide the bacterium multiple mechanisms for protecting itself from oxidative damage arising from host immune defenses.

Encapsulation of DyP and ferritin homologs has previously been observed in other organisms [1, 24]; however, there is no evidence that Enc interacts with FolB in other organisms. A recent study suggests that a potential *Haliangium ochraceum* aldolase is compartmentalized by a bacterial BMC by its C-terminal extension (**Figure 3.1A**) [2]. Furthermore, *M. tuberculosis* i*n vivo* microarray data suggest that Mt-Enc and Mt-FolB both are up-regulated under hypoxic conditions [46] and exposure to cephalexin [47], further suggesting that they indeed may interact.

This study provides the first evidence of a mycobacterial nanocompartment that has the ability to encapsulate cargo proteins, where encapsulated cargo proteins retain their enzymatic activities. To date, there are no reports illustrating the formation of a proteinaceous nanocompartment in mycobacteria. It is unknown which cargo protein(s) Mt-Enc encapsulates *in vivo*, how it is regulated, and whether it can accommodate multiple cargo proteins. As has been suggested for other bacterial compartments [48], the potential for carrying cargo and apparent protein stability make Mt-Enc an appealing vehicle to be exploited as a drug delivery system.

REFERENCES

1. Sutter, M., et al., *Structural basis of enzyme encapsulation into a bacterial nanocompartment.* Nat Struct Mol Biol, 2008. **15**(9): p. 939-47.
2. Lassila, J.K., et al., *Assembly of Robust Bacterial Microcompartment Shells Using Building Blocks from an Organelle of Unknown Function.* Journal of Molecular Biology, 2014. **426**(11): p. 2217-2228.
3. Marchler-Bauer, A., et al., *CDD: a Conserved Domain Database for the functional annotation of proteins.* Nucleic Acids Research, 2011. **39**: p. D225-D229.
4. Organization, W.H., *Global Tuberculosis Report.* 2014.
5. Tetta, C., et al., *Extracellular vesicles as an emerging mechanism of cell-to-cell communication.* Endocrine, 2013. **44**(1): p. 11-9.
6. Ohno, S., A. Ishikawa, and M. Kuroda, *Roles of exosomes and microvesicles in disease pathogenesis.* Adv Drug Deliv Rev, 2013. **65**(3): p. 398-401.

7. Giri, P.K., et al., *Proteomic analysis identifies highly antigenic proteins in exosomes from M. tuberculosis-infected and culture filtrate protein-treated macrophages.* Proteomics, 2010. **10**(17): p. 3190-202.
8. Rosenkrands, I., et al., *Identification and characterization of a 29-kilodalton protein from Mycobacterium tuberculosis culture filtrate recognized by mouse memory effector cells.* Infect Immun, 1998. **66**(6): p. 2728-35.
9. Chiancone, E., et al., *Iron and proteins for iron storage and detoxification.* Biometals, 2004. **17**(3): p. 197-202.
10. Kaur, A., et al., *Coordination of frontline defense mechanisms under severe oxidative stress.* Mol Syst Biol, 2010. **6**: p. 393.
11. Kim, S.J. and M. Shoda, *Purification and characterization of a novel peroxidase from Geotrichum candidum dec 1 involved in decolorization of dyes.* Appl Environ Microbiol, 1999. **65**(3): p. 1029-35.
12. Ahmad, M., et al., *Identification of DypB from Rhodococcus jostii RHA1 as a lignin peroxidase.* Biochemistry, 2011. **50**(23): p. 5096-107.
13. Letoffe, S., et al., *Bacteria capture iron from heme by keeping tetrapyrrol skeleton intact.* Proc Natl Acad Sci U S A, 2009. **106**(28): p. 11719-24.
14. Dailey, H.A., et al., *The Escherichia coli protein YfeX functions as a porphyrinogen oxidase, not a heme dechelatase.* MBio, 2011. **2**(6): p. e00248-11.
15. Pandey, R. and G.M. Rodriguez, *A ferritin mutant of Mycobacterium tuberculosis is highly susceptible to killing by antibiotics and is unable to establish a chronic infection in mice.* Infect Immun, 2012. **80**(10): p. 3650-9.
16. Reddy, P.V., et al., *Iron storage proteins are essential for the survival and pathogenesis of Mycobacterium tuberculosis in THP-1 macrophages and the guinea pig model of infection.* J Bacteriol, 2012. **194**(3): p. 567-75.
17. Dantola, M.L., et al., *Stability of 7,8-dihydropterins in air-equilibrated aqueous solutions.* Helvetica Chimica Acta, 2008. **91**(3): p. 411-425.
18. Goulding, C.W., et al., *Regulation by oligomerization in a mycobacterial folate biosynthetic enzyme.* Journal of Molecular Biology, 2005. **349**(1): p. 61-72.
19. Schobersberger, W., et al., *Neopterin and 7,8-dihydroneopterin induce apoptosis in the rat alveolar epithelial cell line L2.* Febs Letters, 1996. **397**(2-3): p. 263-268.
20. Lowry, O.H., et al., *Protein measurement with the Folin phenol reagent.* J Biol Chem, 1951. **193**(1): p. 265-75.
21. Bradford, M.M., *A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding.* Anal Biochem, 1976. **72**: p. 248-54.
22. Ogola, H.J., et al., *Molecular characterization of a novel peroxidase from the cyanobacterium Anabaena sp. strain PCC 7120.* Appl Environ Microbiol, 2009. **75**(23): p. 7509-18.
23. Chance, B. and A.C. Maehly, *Assay of Catalases and Peroxidases.* Methods in Enzymology, 1955. **2**: p. 764-775.
24. Rahmanpour, R. and T.D. Bugg, *Assembly in vitro of Rhodococcus jostii RHA1 encapsulin and peroxidase DypB to form a nanocompartment.* FEBS J, 2013. **280**(9): p. 2097-104.

25. Shin, K.S. and Y.J. Lee, *Purification and characterization of a new member of the laccase family from the white-rot basidiomycete Coriolus hirsutus.* Arch Biochem Biophys, 2000. **384**(1): p. 109-15.
26. Khare, G., et al., *Ferritin structure from Mycobacterium tuberculosis: comparative study with homologues identifies extended C-terminus involved in ferroxidase activity.* PLoS One, 2011. **6**(4): p. e18570.
27. Schneider, C.A., W.S. Rasband, and K.W. Eliceiri, *NIH Image to ImageJ: 25 years of image analysis.* Nat Methods, 2012. **9**(7): p. 671-5.
28. McMath, L.M., et al., *Crystallization and preliminary X-ray crystallographic analysis of a Mycobacterium tuberculosis ferritin homolog, BfrB.* Acta Crystallogr Sect F Struct Biol Cryst Commun, 2010. **66**(Pt 12): p. 1657-61.
29. Mishra, S. and J. Imlay, *Why do bacteria use so many enzymes to scavenge hydrogen peroxide?* Arch Biochem Biophys, 2012. **525**(2): p. 145-60.
30. Manca, C., et al., *Mycobacterium tuberculosis catalase and peroxidase activities and resistance to oxidative killing in human monocytes in vitro.* Infect Immun, 1999. **67**(1): p. 74-9.
31. Gupta, V., et al., *Crystal Structure of Bfr A from Mycobacterium tuberculosis: Incorporation of Selenomethionine Results in Cleavage and Demetallation of Haem.* Plos One, 2009. **4**(11).
32. McMath, L.M., et al., *The structural characterization of bacterioferritin, BfrA, from Mycobacterium tuberculosis.* Journal of Porphyrins and Phthalocyanines, 2013. **17**(3): p. 229-239.
33. Liu, X. and E.C. Theil, *Ferritin reactions: direct identification of the site for the diferric peroxide reaction intermediate.* Proc Natl Acad Sci U S A, 2004. **101**(23): p. 8557-62.
34. Choudhary, S., et al., *Engineered protein nano-compartments for targeted enzyme localization.* PLoS One, 2012. **7**(3): p. e33342.
35. Kerfeld, C.A., S. Heinhorst, and G.C. Cannon, *Bacterial microcompartments.* Annu Rev Microbiol, 2010. **64**: p. 391-408.
36. Yeates, T.O., C.S. Crowley, and S. Tanaka, *Bacterial microcompartment organelles: protein shell structure and evolution.* Annu Rev Biophys, 2010. **39**: p. 185-205.
37. Jorda, J., et al., *Using comparative genomics to uncover new kinds of protein-based metabolic organelles in bacteria.* Protein Sci, 2013. **22**(2): p. 179-95.
38. Fan, C., et al., *Short N-terminal sequences package proteins into bacterial microcompartments.* Proc Natl Acad Sci U S A, 2010. **107**(16): p. 7509-14.
39. Akita, F., et al., *The crystal structure of a virus-like particle from the hyperthermophilic archaeon Pyrococcus furiosus provides insight into the evolution of viruses.* J Mol Biol, 2007. **368**(5): p. 1469-83.
40. Wikoff, W.R., et al., *Topologically linked protein rings in the bacteriophage HK97 capsid.* Science, 2000. **289**(5487): p. 2129-33.
41. Sherman, D.R., et al., *Regulation of the Mycobacterium tuberculosis hypoxic response gene encoding alpha -crystallin.* Proc Natl Acad Sci U S A, 2001. **98**(13): p. 7534-9.
42. Garbe, T.R., N.S. Hibler, and V. Deretic, *Response to reactive nitrogen intermediates in Mycobacterium tuberculosis: induction of the 16-kilodalton*

alpha-crystallin homolog by exposure to nitric oxide donors. Infect Immun, 1999. **67**(1): p. 460-5.
43. Voskuil, M.I., K.C. Visconti, and G.K. Schoolnik, *Mycobacterium tuberculosis gene expression during adaptation to stationary phase and low-oxygen dormancy.* Tuberculosis (Edinb), 2004. **84**(3-4): p. 218-27.
44. Dántola, M.L., et al., *Stability of 7,8-Dihydropterins in Air-Equilibrated Aqueous Solutions.* Helvetica Chimica Acta, 2008. **91**(3): p. 14.
45. Czekster, C.M. and J.S. Blanchard, *One substrate, five products: reactions catalyzed by the dihydroneopterin aldolase from Mycobacterium tuberculosis.* J Am Chem Soc, 2012. **134**(48): p. 19758-71.
46. Rustad, T.R., et al., *The enduring hypoxic response of Mycobacterium tuberculosis.* PLoS One, 2008. **3**(1): p. e1502.
47. Boshoff, H.I. and C.E. Barry, 3rd, *Tuberculosis - metabolism and respiration in the absence of growth.* Nat Rev Microbiol, 2005. **3**(1): p. 70-80.
48. Corchero, J.L. and J. Cedano, *Self-assembling, protein-based intracellular bacterial organelles: emerging vehicles for encapsulating, targeting and delivering therapeutical cargoes.* Microb Cell Fact, 2011. **10**: p. 92.

CHAPTER 4

Investigation of a Mycobacterial Encapsulin Nanocompartment and its Effects on Survival

ABSTRACT

Mycobacterium tuberculosis (Mtb) possesses a proteinaceous nanocompartment (Mt-Enc)—part of a novel family of Encapsulin (Enc) proteins. Structural and physiological characterization of other Enc proteins is currently underway. Although bacteria do not have organelles, bacterial microcompartments observed in other bacteria have been reported to function similarly to organelles. Their precise roles in nature are unknown, but these enzymes have been implicated in oxidative stress response by virtue of the antioxidant cargo enzymes they carry. Research on if and how bacteria utilize compartmentalization is currently underway; however; the possibility of using them as a therapeutic or vaccine delivery system is the most intriguing. Enc proteins are conserved across bacterial species, however in Mtb they have been described as being highly antigenic and secreted. Although Mt-Enc lacks a signal peptide, it has a recognition sequence in the C-terminus for secretion via the type VII secretion system (TVIISS). In this chapter, we begin structural characterization studies, where we show progress toward obtaining the crystal structure of Mt-Enc. Furthermore, we show the generation of a Δ*enc* strain in the fast-growing, non-pathogenic model organism *Mycobacterium smegmatis*. Lastly, we preliminarily probe the *in vivo* functions of mycobacterial Enc by utilizing *M. smegmatis*Δ*enc.*

INTRODUCTION

A family of conserved bacterial proteins called Encapsulins (Enc) has recently emerged. Enc form a family of conserved bacterial proteins that form proteinaceous nanocompartments [1]. These large molecular assemblies were previously referred to as linocin-like proteins based on their abundant presence in the supernatant of *Brevibacterium linens* (Bl-Enc), attributing to their role as having bacteriostatic activity against various bacterial species [2] . Homologs of Bl-Enc are found in several bacterial organisms, including *Thermotoga maritima*, *Myxococcus xanthus*, *Rhodococcus jostii* RHA1, *Mycobacterium smegmtais*, and *Mycobacterium tuberculosis* (Mtb) (**Figure 4.1**). Although bacteriostatic activity was reported for *B. linens*, none was reported for *T. maritima* and Mtb Enc homologs [1-3].

Mtb encapsulin (Mt-Enc) is a highly stable nanocompartment that induces a T-cell-mediated immune response [3], with a 58% amino acid identity to Bl-Enc. Encapsulins from other organisms have been shown to be self-assembling, thermostable compartments that can protect bacteria from oxidative stress [4]. Sutter, *et al* reported the genomic context in which the genes for encapsulins are found in the bacterial chromosome: 1) they are often in a two-gene operon; 2) the gene encoding encapsulin lies downstream in the operon; and 3) the upstream gene encodes for either a ferritin-like protein or a dye-decolorizing peroxidase-family of proteins [1]. *R. jostii* Enc (Rj-Enc) was shown to disassemble after treatment with acetate buffer pH 3.0 although reassembles upon treatment of neutral pH [4]. Reassembly of Rj-Enc was observed with and without cargo protein DyPB, and Rj-Enc can form expanded structures thus changing the number of subunits per compartment [4].

Although Enc alone has not been characterized to have an enzymatic function, reports have suggested these proteins can play an important role in oxidative stress and iron starvation response by virtue of the cargo enzymes that they carry [4, 5]. For example, *M. xanthus* Enc (Mx-Enc) has an iron-rich core, upon which assembly is induced under starvation conditions and is important for survival under oxidative stress [5]. To address if the Mx-Enc iron-dense core is due to the isolation of free iron to protect the bacteria from oxidative damage via the Fenton reaction, a *M. xanthus*Δ*enc* mutant was generated to examine its sensitivity to hydrogen peroxide (H_2O_2), in which the mutant resulted in greater sensitivity to H_2O_2; these results suggested that Mx-Enc is important for protection of *M. xanthus* from oxidative damage under starvation conditions [5].

Mt-Enc, also known as CFP-29, was proposed to be secreted, although it lacks a signal peptide. Rosenkrands *et al* reported that Mt-Enc was membrane-enriched [3]. Although no discernable signal peptide was found, a predicted general secretion signal for the mycobacterial type VII secretion system (TVIISS) was discovered for specialized secretion of mycobacterial proteins lacking a signal peptide [6, 7]. Within the TVIISS, there are 5 types of systems: ESX-1 (Mtb virulence and conjugation for *M. smegmatis*); ESX-3 (iron and zinc homeostasis); ESX-5 (secretion of Pro-Glu/Pro-Pro-Glu proteins) and the largely uncharacterized ESX-2 and ESX-4 [8]. Both Mt-Enc and *M. smegamtis* Enc homolog, MSMEG_5830 (Ms-Enc), harbor a five C-terminal amino acid residues (YASHD and YLSHD, respectively) consistent with the motif described for TVII-mediated secretion substrates, YxxxD/E [7]; Mt-Enc and Ms-Enc have 83% identity. Other proteins that utilize TVIISS are T-cell antigens ESAT-6 and CFP-10 [9, 10]. It is

possible that Mt-Enc may not always be secreted into the extracellular space but may remain in the perisplasmic space or may be sloughed off of the mycobacterial capsule layer during *in vitro* growth conditions hence mistakenly identified as secreted as was observed for ESX-1-secreted proteins from *M. marinum* [11]. Furthermore, mycobacterial Enc may carry cargo *in vivo* that aids in alleviating the extracellular assault against mycobacterial infection imposed by the host.

To utilize Mt-Enc as a delivery system, we sought to understand the structural components of this mycobacterial nanocompartment. The concept of compartmentalization in bacterial cells is relatively new, and preliminary studies shown here will reveal clues regarding this unique prokaryotic compartment in terms of mycobacterial physiology. Therefore, we sought to further investigate Mt-Enc by attempting to solve its crystal structure as well as its role in mycobacteria to ultimately tailor it as a specific vaccine delivery vehicle to combat disease.

MATERIALS AND METHODS

Cloning

The Mtb genes encoding proteins Mt-Enc (*Rv0798c* were PCR-amplified from Mtb H37Rv genomic DNA using the KOD HotStart Polymerase Kit (Novagen) with 5′ and 3′ primers (MWG Operon) into the appropriate restriction sites, to generate a C-terminal poly-histidine tag (His-tag). PCR products were ligated into pCR-BluntII-TOPO (Invitrogen), and then transformed into *Escherichia coli* OneShot TOP10 cells (Invitrogen). Double digestions with specific restriction enzymes were performed from each vector of choice (**Table 4.1**). Excised genes were ligated into the appropriate

linearized pET22, pYA1611, or pKM402$_{AflII}$ vector and transformed into *E. coli* BL21-Gold (DE3) cells (Novagen). Plasmid verification was performed via DNA sequencing using T7 promoter and reverse primers as well as custom sequencing primers for pYA1611, pKM343, and pKM402 primers (Laguna Scientific). Mycobacterial tetracycline-induduible expression plasmids pYA1611 [12] and pKM402$_{SDM}$ [13] were electroporated into electrocompetent *M. smegmatis* mc^{2}155 [14].

To permit cloning of Mt-Enc into the PstI site of pYA1611, site-directed mutagenesis (SDM) primers were generated to mutagenize the PstI site found within *Mt-Enc* gene: F-5` CGTGCGCCTCTACCTTCAGGAGACGCTGACG, R-5 CGTCAGCGTCTCCTGAAGGTAGAGGCGCACG; pKM420$_{AflII}$ was mutagenized to allow the removal of the Che9c mycobacteriophage recombinase by adding an AflII site and inserting of anhydrotetracycline-inducible genes noted below.

Table 4.1: Primers used in this study.

Gene	Construct	Primer Sequence
Rv0798c-Rv0799c (Enc)	pET22::Enc-DyP; Kan^R, IPTG-inducible	FWD: 5`GGCATATGAACAATCTCTACCGCGATTTGGCACCGGTCACC GAAG REV: 5` GGCTCGAGTCGGGGGCTTCCTTTCAAGCTGCCGATCGATAGC
Rv0798c (Enc)	pYA1611::Enc$_{His}$; Hyg^R; ATC-inducible	FWD: 5`CGCTGCAGATGAACAATCTCTACCGCGATTTGGCACCGGTC AC REV: 5`CGCTGCAGTCAGTGGTGGTGGTGGTGGTG
Rv0798c (Enc)	pKM402$_{SDM}$::Enc$_{His}$; Kan^R; ATC-inducible	
Rv0798c Y234A(Enc)	pYA1611::EncY234A$_{Enc}$/ pKM402$_{SDM}$::EncY234A$_{His}$; Hyg^R; ATC-inducible	FWD: 5`GACGTTGCAATCGGGGCGGCCAGCCACGACACG REV: 5`CGTGTCGTGGCTGGCCGCCCCGATTGCAACGTC
*Rv0798c*D238A (Enc)	pYA1611::EncD2348$_{Enc}$/ pKM402$_{SDM}$::EncD238A$_{His;}$ Kan^R; ATC-inducible	FWD: 5`GTACGCCAGCCACGCTACGGACACCGTGCG REV: 5`CGCACGGTGTCCGTAGCGTGGCTGGCGTAC
*Rv0798c*Y234A/D238A (Enc)	pYA1611::EncY234A/D238A$_{Enc}$/ pKM402$_{SDM}$::EncY234A/D238A$_{His}$	*Generated Y234A mutation first, then used D238A primers to generate second mutation
800 bp upstream of *MSMEG_5830*	pKM343::FL1-Enc-FL2; Hyg^R/Cam^R	FWD: 5`CGGCGGCCGCTCGAACTCGCGACCAAGCTCGT REV: 5`CGACTAGTTCCGGCGATGTGACGCTTGAACGTGCGGGTC
800 bp downstream of *MSMEG_5830*	pKM343::FL1-Enc-FL2; Hyg^R/Cam^R	FWD: 5`CGGGGGCCCGCTACCTGTCCCATGACGCC REV: 5`GCGGGGCCCTGGTGCCCGGCGGTTTCT
Hyg probe	N/A	FWD: 5`GCCTTAAGGTGACACAAGAATCCCTGTTACTTCTCGACCG REV: 5`GCCTTAAGTCAGGCGCCGGGGGCGGTGTCCGG

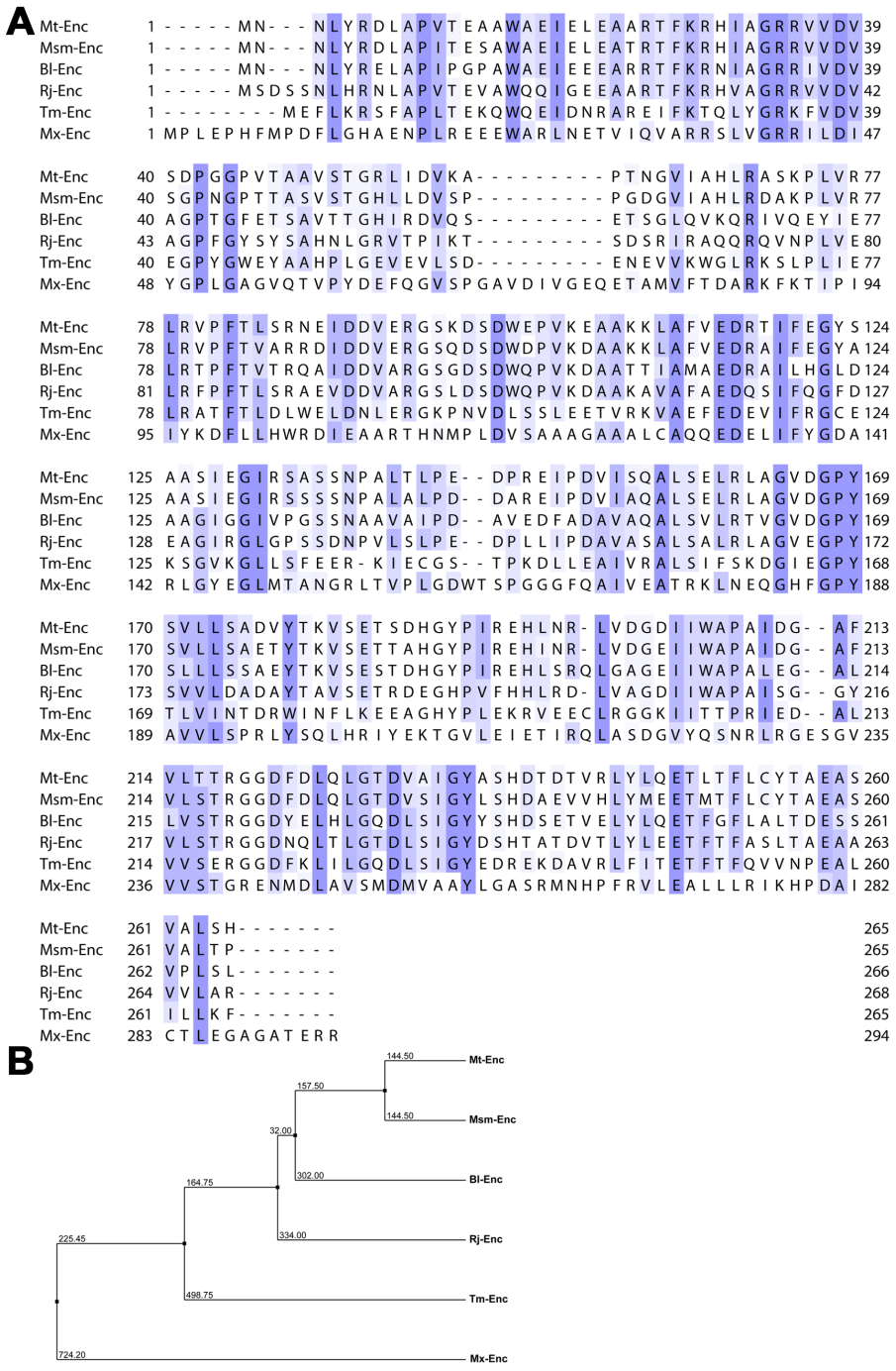

Figure 4.1: Sequence alignment and phylogenetic tree of Enc amino acid sequences from six bacterial organisms. Sequence alignment of Enc amino acid sequence from 6 organisms: Mtb (Mt-); *M. smegmatis* (Ms-); *B. linens* (Bl-); *R. jostii* RHA1 (Rj-); *T. maritimia* (Tm-); and *M. xanthus* (Mx-). A) BLOSUM alignment of Enc proteins from five bacterial species. Shades of purple are indicative degree of conservation. B) Phylogenetic tree using PAM 250 distance calculation. Alignment and tree were generated using JalView.

Mutant generation of M. smegmatis mc^{2}155Δ*MSMEG_5830* (Δ*enc*).

Generation of *M. smegmatis* mc^{2}155Δ*enc*::*hyg* was generated by following the van Kessel and Hatfull method of mycobacterial allelic exchange [13]. Briefly, we electroporated pJV53 (plasmid harboring acetamide-inducible mycobacteriophage Che9c recombinase) into electrocompetent *M. smegmatis* mc^{2}155 as previously described [14]. Che9c recombinase includes gp60 and gp61, homologous to RecE and RecT recombinases in *E. coli*, respectively [13]. Once electroporated, mycobacterial cells harboring the recombinase plasmid were subcultured in Middlebrook 7H9 (Becton Dickinson, New Jersey) containing 0.2% succinate, 10% ADS (albumin-dextrose-saline), and 0.05% Tyloxapol to an OD_{600} ~ 0.02. 0.2% acetamide was added to cultures when OD_{600} ~ 0.5 was reached, and cells were made electrocompetent after 3 hours of acetamide induction of recombinase. Electrocompetent cells of *M. smegmatis* mc^{2}155 + pJV53 were then electroporated with linear DNA containing: ~ 800 bp upstream and downstream of *MSMEG_5830* (*Ms-Enc*), chloramphenicol (Cam^R/ *cat* gene) and hygromycin (Hyg^R/*hyg* gene) resistance cassettes; flanking regions were subcloned into plasmid pKM343.

Growth curves comparing M. smegmatis mc^{2}155 and *M. smegmatis*mc^{2}155Δ*enc under various conditions.*

Cultures were grown overnight; cells are pelleted at 2000 RPM for 10 min at 4°C.

For H_2O_2 experiments: Cell pellet was then resuspended in 500 µL of 7H9 + 5% ADS media, and OD_{750} was taken. Resuspension was back diluted to starting OD_{750}~ 0.04 in media. Cultures are grown to OD_{750}~ 0.1 and then oxidative stress was induced upon the addition of 1 mM, 2 mM and 3 mM H_2O_2. Various time points were taken over 50 hours.

For low-iron experiments: Cell pellet was then washed three times in 10 mL of iron-deplete media (IDM) [15]. Cells were resuspended in 500 μL IDM, and an OD_{750} was taken. Resuspension was back-diluted to starting OD_{750}~ 0.05, upon which IDM was supplemented with 2 μM ferric ammonium citrate (FAC). Various time points were taken over 100 hours.

For low pH experiments: Cell pellet was then resuspended in 500 μL of 7H9 + 5% ADS media, and OD_{750} was taken. Resuspension was back diluted to starting OD_{750}~ 0.1 in one of three kinds of media—"normal", pH 4 media, and pH5. Various time points were taken over 50 hours.

Over-expression of Mt-Enc proteins in E. coli and M. smegmatis.

Mt-Enc was expressed from plasmids described in ***Table 4.1*** *using either E. coli BL21 Gold (DE3) cells, M. smegmatis mc²155, or M. smegmatis mc²155Δenc::hyg. T*he following concentrations of antibiotics to the media were added where appropriate: ampicillin (Amp 50 μg/ml), kanamycin (Kan 30 μg/ml), hygromycin (Hyg 150 μg/ml), or chloramphenicol (Cam 20 μg/ml). Cells harboring appropriate expression vector were grown aerobically at 37 °C in LB media (for *E. coli*) or in 7H9 media (for *M. smegmatis*) containing the appropriate antibiotic(s). Heterologous protein expression in *E. coli* was induced at OD_{600nm} ~ 0.8 by the addition of isopropyl-β-d-thiogalactopyranoside (IPTG, 1 mM). Cells were harvested after overnight induction at 18 °C by centrifugation at 5100 x g for 20 min. For *M. smegmatis* protein expression, cells were induced at an OD_{600} ~ 0.1 with 200 ng/mL of anhydrotetracycline as previously described [12]. Importantly, in all assays, ATc was used to induce gene expression since it has low toxicity and a high affinity to TetR [16]. Aliquots of cells were taken to monitor expression in *M. smegmatis*

and cells harvested after 24 hour induction at 37 °C by centrifugation as mentioned above.

Purification of Mt-Enc$_{His}$, Mt-Enc$_{His}$-DyP, and Mt-Enc$_{His}$-BfrB from E. coli and Mt-Enc$_{His}$ from M. smegmatis.

Harvested *E. coli* cell pellets were resuspended in 20 mL of Resuspension buffer (50 mM Tris-HCl, pH 7.4, 350 mM NaCl, 10 mM imidazole and 10% glycerol, 0.2% lauryldimethylamine-oxide (LDAO)), followed by the addition of phenylmethylsulfonyl fluoride (PMSF) and hen egg-white lysozyme. *E. coli* cells were disrupted by sonication, clarified by centrifugation at 18000 x g for 30 min at 4° C and syringe-filtered (1 µm pore size) for removal of cell debris.

Harvested *M. smegmatis* cell pellets were resuspended in 10 mL of mycobacterial lysis buffer (50 mM Tris pH 7.4, 10 mM $CaCl_2$, 0.02% sodium azide, 0.2% LDAO). Cells were disrupted by sonication, facilitated by the addition of glass beads (Sigma-Aldrich, St. Louis, MO). Mycobacterial lysate was clarified by first centrifuging cells at 40000 x g for 10 min at 4° C to remove unbroken cells, followed by a 27000 x g for 45 min at 4° C. Lysate was syringe-filtered prior to loading onto Ni^{2+}-charged column.

Clarified cell lysate was then loaded onto a 5 mL Ni^{2+}-charged HisTrap column (GE Healthcare) pre-equilibrated with Wash buffer (50 mM Tris-HCl, pH 7.4, 350 mM NaCl, 10 mM imidazole and 10% glycerol). Protein was eluted with a linear gradient of 10-500 mM imidazole (100 mL), Mt-Enc elutes at an imidazole concentration >250 mM. Eluted fractions were collected, analyzed by SDS-PAGE, and concentrated using Amicon concentrators with the appropriate molecular-weight cutoff (Millipore, Bedford, MA).

Extinction coefficients at 280 nm as predicted by the program Protein Calculator (Scripps), or determined by either modified Lowry [17] or Bradford [18] assay using bovine serum albumin as a standard.

Crystallization of Mt-Enc$_{His}$-DyP and Mt-Enc$_{His}$-BfrB.

Mt-Enc$_{His}$-DyP and Mt-Enc$_{His}$-BfrB were purified and concentrated to ~ 11 mg/mL. A Mosquito nanoliter dispensing robot was used to set up 600 sparse matrix crystallization conditions for each protein. Crystallization conditions were optimized and additive screens were utilized. Crystals were grown using hanging-drop vapor diffusion techniques. Diffraction data from diffracting, single crystals were collected at either Stanford Synchrotron Radiation Laboratory (SSRL) or with Mtb Structural Genomics Consortium at the Advanced Light Source at UC Berkeley (ALS). Diffracting crystals of Mt-Enc$_{His}$-DyP and Mt-Enc$_{His}$-BfrB were grown at 11 mg/mL in 0.1 M Trimethlyamine *N*-oxide (TMN), 0.1 M Tris-HCl pH 8.5, and 20% polyethylene glycol. No cryoprotectant was used.

Mt-Enc$_{His}$ Polyclonal Antibody Generation.

Antibody generation was performed as previously described, although with slight modifications [12]. Purified Mt-Enc$_{His}$ was emulsified in aluminium hydroxide and used to immunize two New Zealand White rabbits by subcutaneous injections. Rabbits were given three booster injections of Mt-Enc$_{His}$ antigens at 2-week intervals over a 63-day period, and anti-sera were collected after three booster injections (Thermo Fischer, Rockford, IL). Antibody titer was determined via ELISA (Thermo Fischer, Rockford,

IL). Enrichment of IgG antibodies was performed using a 5 mL Protein G column, following manufacturers instructions (GE Life Sciences, Pittsburg, PA).

Mt-Enc$_{His}$ Immunoblot Analysis.

M. smegmatis was grown in 7H9-ADS-Tylox medium at 37°C to an OD_{600} ~ 0.7. The culture was pelleted by centrifugation, and cells were washed with PBS, pelleted, and then resuspended in mycobacterial lysis buffer and cells were lysed with glass beads by using a sonicator. For Western analysis, 500 ng of purified Mt-Enc$_{His}$ or 10 μg of H37Rv whole-cell lysate or cellular fractions were resolved on an SDS–12% PAGE gel and then transferred onto a nitrocellulose membrane (Millipore). Membranes were blocked in 3% skim milk in PBST buffer, incubated with αMt-Enc$_{His}$ (1:10,000) rabbit polyclonal antiserum, and developed with a horseradish peroxidase-labeled anti-rabbit immunoglobulin G (IgG) antibody (Millipore, Darmstadt, Germany) and Luminata Classico chemiluminescent substrate system (Millipore, Darmstadt, Germany). Membranes were exposed to X-ray film at room temperature.

Fractionation of Mycobacterial Cells.

Fractionation of mycobacterial fractions was performed as previously described [19], although with modifications described herein. Mycobacteria were grown to desired density in appropriate media. Mycobacterial cells were harvested by centrifugation at 4000 x *g* for 15 min at 4°C. Supernatant was discarded and cells were washed with 1/5 of the original culture volume with pre-chilled phosphate-buffered saline + 0.05% Tween 80, pH 8.0 (PBST), and PMSF was added. Mycobacterial cells were disrupted as

described above. Unbroken cells and debris was removed by centrifuging sonicate 4000 x *g* for 10 min at 4°C. Centrifugation steps above 4000 x *g* were performed on a Beckman Ultracentrifuge using a TLS55 rotor. Supernatant was saved, then centrifuged at 27,000 x *g* for 45 min. at 4°C. The resulting pellet is the cell wall fraction (contains proteins and non-protein compounds such as arabinogalactan-peptidoglycan). The cell wall fraction/pellet was resuspended the pellet in 10 mM ammonium bicarbonate and quantified and supernatant was further processed.

The supernatant was centrifuged at 150,000 x *g* for 1 hour at 4°C. The supernatant is the soluble cytosolic fraction. The resulting membrane pellet was washed with PBS and resuspended in 10 mM ammonium bicarbonate and quantified.

For extraction of *all* the hydrophobic membrane proteins, 32% of pre-treated TritonX-114 (TX-114) was added to cell resuspension (prior to sonication) to a final concentration of 2% and extraction methods were followed. Supernatant was saved and incubated for 10 min at 37°C to allow phase partitioning, then centrifuged at 1000 x *g* for 10 min at room temperature. The upper aqueous phase, which contains hydrophilic proteins, was collected. The lower TX-114 detergent phase enriched for hydrophobic proteins was collected. The detergent phase was washed to minimize contamination by soluble proteins by adding a volume of mycobacterial lysis buffer corresponding with the volume of the aqueous phase previously removed. Incubated the resuspension for 10 min at 37°C, then centrifuge at 1000 x *g* for 10 min at room temperature to allow phase partitioning. The upper aqueous supernatant was carefully removed, then repeated three times. Proteins were precipitated from the detergent phase by adding 5 volumes of chilled 100% acetone, and incubated overnight at – 20°C. The next morning, preparation

was centrifuged at 14,000 x *g* for 30 min at 4°C to sediment the proteins. The pellet was air-dried, then resuspended in 10 mM ammonium bicarbonate and quantified.

Southern blot analysis of M. smegmatis mc^{2}155Δ*enc*::*hyg*^{r}.

Hyg probes were PCR-amplified and labeled using DIG High Prime DNA Labeling and Detection Starter Kit II, according to manufacturer's instructions (Roche, Indianapolis, IN). Genomic DNA was extracted using *N*-acetyl-*N*,*N*,*N*-trimethyl ammonium bromide (CTAB) method [20]. Southern hybridization was performed as previously described [21].

RESULTS

Crystallization of Mt-Enc$_{His}$-DyP and Mt-Enc$_{His}$-BfrB

Because we were unable to obtain crystals of Mt-Enc$_{His}$, Mt-Enc$_{His}$-DyP and Mt-Enc$_{His}$-BfrB were purified to homogeneity by nickel-affinity chromatography as previously described [22]. Purified Mt-Enc$_{His}$DyP and Mt-Enc$_{His}$-BfrB were concentrated and utilized for screening crystallization conditions. Crystals of Mt-Enc$_{His}$-DyP diffracted to ~ 3.0 Å (**Figure 4.2A and B**), indexed to P 2 3 with the following unit cell dimensions: 313.7 Å, 313.7 Å, 313.7 Å, 90°, 90°, 90°. Crystals of Mt-Enc$_{His}$-BfrB diffracted to ~ 8.0 Å (**Figure 4.2C and D**). Crystals are currently being screened and optimized to improve diffraction data.

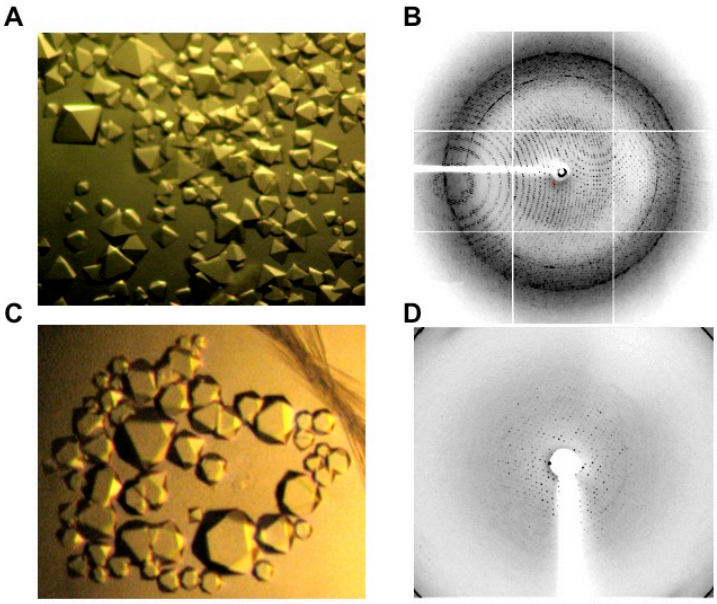

Figure 4.2: Attempts at structurally characterizing Mt-Enc. A) Crystals of Mt-Enc$_{His}$-DyP grown in 0.1 M Tris pH 8.5, 0.1 M TMN, 20% PEG MME 2000, at 11 mg/mL protein concentration. Crystals grown in this condition were harvested and screened. B) Diffraction obtained for a crystal harvested in (A). Resolution reached ~ 3.0 Å. C) Crystals of Mt-Enc$_{His}$-BfrB grown in 0.1 M Tris pH 8.5, 0.1 M TMN, 20% PEG MME 2000. D) Diffraction obtained for a crystal harvested in (B). Resolution reached ~ 8.8 Å.

Detection of over-expression of Mt-Enc$_{His}$ in M. smegmatis mc^{2}155.

The endogenous levels of Ms-Enc are unknown, and may be so low that it may be difficult to detect. It has been mentioned that although Mt-Enc was discovered in the culture filtrate, it is possible that Mt-Enc may be membrane-enriched [3]. Because of the high level of identity between Mt-Enc and Ms-Enc, we attempted to probe for Ms-Enc in *M. smegmatis* mc^{2}155 in the various cell fractions utilizing the α-Mt-Enc$_{His}$ antibody; however, we were unable to detect it presumably because of the low levels of expressed Ms-Enc under normal growth conditions. One other possibility was that antibodies against Mt-Enc were generated from *E. coli*-expressed, recombinant Mt-Enc$_{His}$. It is possible that the polyclonal α-Mt-Enc$_{His}$ antibody more readily recognizes the His-tag as opposed to Mt-Enc epitopes. Therefore, we sought to overexpress Mt-Enc in *M.*

smegmatis mc^{2}155 for the following three reasons: 1) to detect mycobacterial Enc to probe for protein localization; 2) express Mt-Enc$_{His}$ and utilize the His tag to potentially co-purify encapsulated proteins; 3) to grow *M. smegmatis* mc^{2}155Δ*enc* with Mt-Enc$_{His}$ and mutants. Tetracycline (Tc)-regulated gene expression system was used to control endogenous expression of *Mt-Enc$_{His}$ and mutants* [23]. In this system, in the absence of Tc, the Tc repressor (TetR) binds to *tetO* and represses gene expression. Conversely, in the presence of Tc, Tc binds to TetR and induces a conformational change that results in dissociation of TetR from *tetO*, thus inducing the expression of TetR-controlled genes [23]. Importantly, all overexpression experiments were performed in *M. smegmatis* mc^{2}155 first, then ultimately in *M. smegmatis* mc^{2}155Δ*enc*. As seen in **Figure 4.3**, overexpression of *Mt-Enc$_{His}$* is sufficient for detection of *Mt-Enc$_{His}$ and mutants by Western blot.*

Generation of M. smegmatis mc^{2}155Δ*enc and preliminary mutant characterization.*

The importance of Enc proteins is currently unknown; however, it has been hypothesized that Enc proteins may be involved in bacterial survival in the presence of oxidative stress [4, 5]. We generated *M. smegmatis* mc^{2}155Δ*enc*::*hyg-cat* via allelic exchange as described above, suggesting that Ms-Enc is not an essential gene (**Figure 4.4**). We first conducted experiments to compare growth of wild-type *M. smegmatis* mc^{2}155 (WT) and *M. smegmatis* mc^{2}155Δ*enc::hyg-cat* in (Δ*enc*) 7H9 + ADS + Tylox$^{0.0\%}$ + Hyg^{100} media. Under regular growth media, Δ*enc* (red) does not observe a growth defect compared to WT (blue) (**Figure 4.5A**).

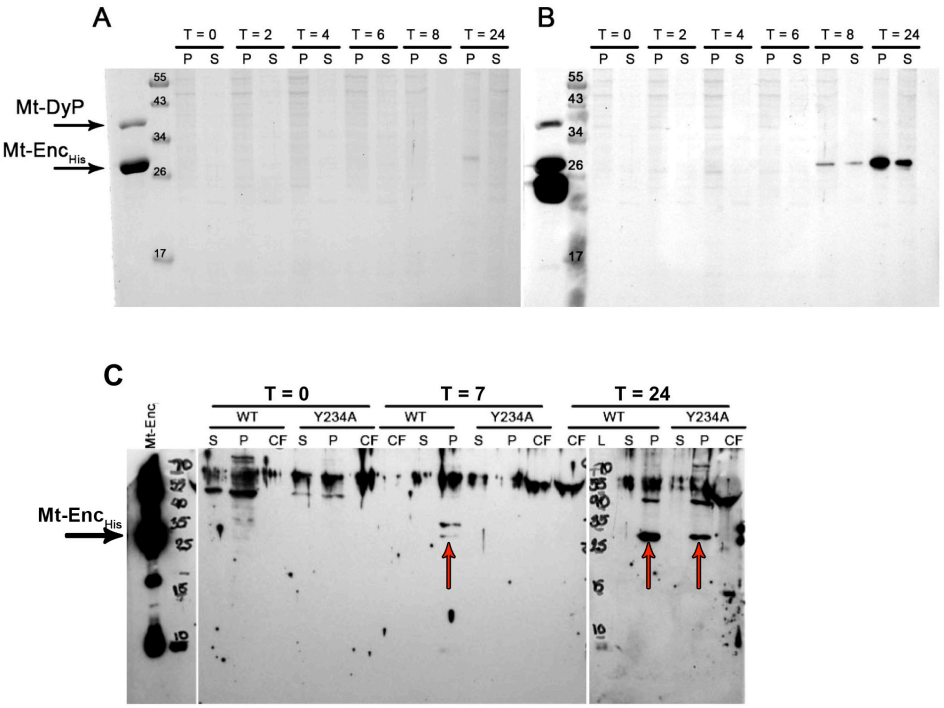

Figure 4.3: **Overexpression of Mt-Enc$_{His}$ and Mt-EncY234A$_{His}$ in *M. smegmatis* mc^{2}155.** A and B) Overexpression of Mt-Enc$_{His}$ with pYA1611 in the far left purified Mt-Enc$_{His}$-DyP are used as a control. A) 12% SDS-PAGE of mycobacterial cells expressing Mt-Enc$_{His}$. B) Immunoblot of the same samples in A). Greatest expression is observed after 24 hours post ATc-induction, although some protein is still found in the insoluble pellet. C) Immunoblot analysis of Mt-Enc$_{His}$ and Mt-EncY234A$_{His}$ overexpressed from pKM402$_{AflII}$. Expression for both proteins is best observed after 24 hours post ATc-induction; however some Mt-Enc$_{His}$ is observed after 7 hours post induction. T = hours; P = insoluble pellet; CF = culture filtrate; S = soluble fraction. Immunoblot analyses were done with anti-Mt-Enc antibody (1:10,000 dilution).

To examine the effect of Ms-Enc in oxidative stress, growth curves in the presence in H_2O_2, low iron, or low pH. First, we observed the effects of H_2O_2 between WT and Δ*enc*. It appears that Δ*enc* shows an increased sensitivity to 2 mM H_2O_2 (red) compared to WT at the same concentration (blue), while it appears that everything died at 3 mM H_2O_2 (**Figure 4.5B**). This is reminiscent of the observation made in *Myxococcus xanthus*, where under low-iron conditions, ΔMx-Enc is more sensitive to H_2O_2 [5].

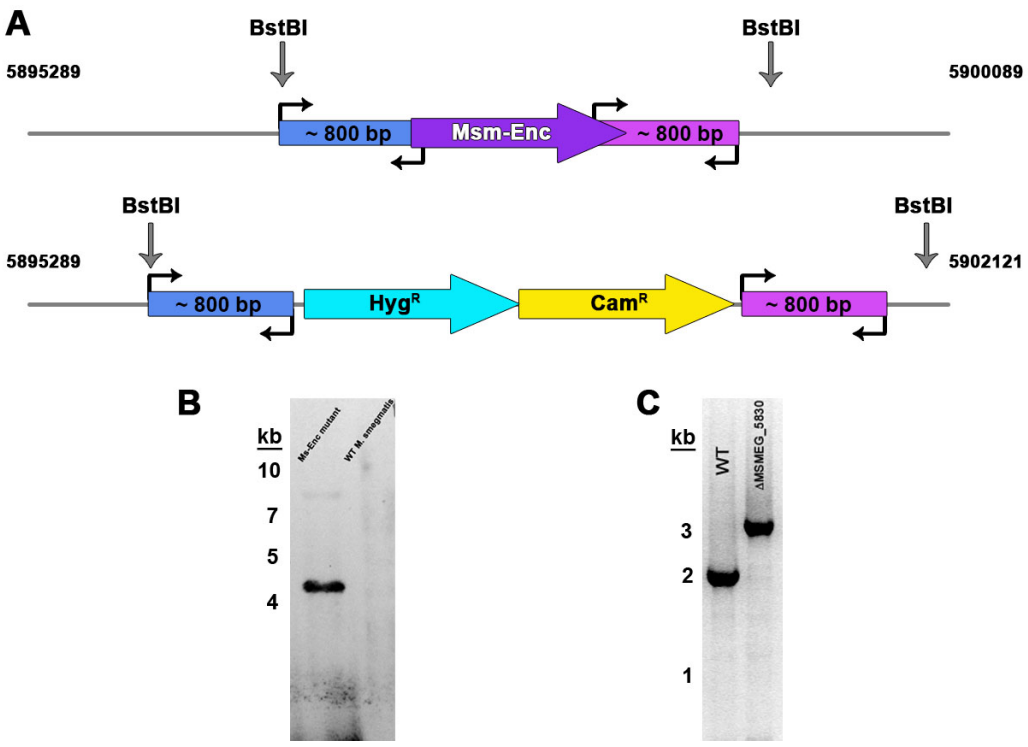

Figure 4.4: Generation and identification of *M. smegmatis* mc^{2}155Δ*enc*. A) General schematic of organization of wild-type (top) and mutant (bottom). Numbers on either side of the schematic represent the genomic vicinity of where allelic exchange was designed to occur. BstBI is the enzyme used to digest CTAB-extracted mycobacterial DNA for Southern hybridization. B) Southern blot analysis comparing mutant and wild-type genomic DNA for the insertion of the *hyg*^{R} cassette. The *hyg*^{R} probe hybridized to mutant genomic DNA but not wild-type. C) PCR-amplification of the genomic region where allelic exchange occurred. Because of the insertion of *hyg*^{R} and *cat,* the amplicon for the mutant would increase ~ 1.5 kb. Cam^R is encoded by the *cat* gene (chloramphenicol acetyltransferase).

To test the effect of low iron growth conditions, WT (blue) and Δ*enc* (red) were depleted of iron and grown in iron-deficient media. There is no observable difference between WT (blue) and Δ*enc* (red) (**Figure 4.5C**). The decrease in OD_{750} is due to the lack of iron in the media, thus inability for cells to grow at a higher density. Upon supplementation with 2 μM FAC, both WT (purple) and Δ*enc* (orange) grew to similar levels. This suggests there is no growth defect upon the loss of Ms-Enc in low-iron conditions. Based on these growth curves it does not seem that Ms-Enc is involved in

iron uptake independently of other well-characterized iron-uptake mechanisms in mycobacteria [24, 25].

Maintaining intracellular pH is important to the mycobacterial, as it has been hypothesized that mycobacteria can encounter acidic pH while residing within activated macrophage [26]. Therefore, we tested WT and Δ*enc* in media at pH 4 and pH 5 (**Figure 4.5D**). Normal cells are able to tolerate pH 5, while pH 4 cells will begin to die. A comparison between WT and Δ*enc* revealed no differences between the two. It is possible that Ms-Enc is not involved in regulating intracellular pH of mycobacterial cells.

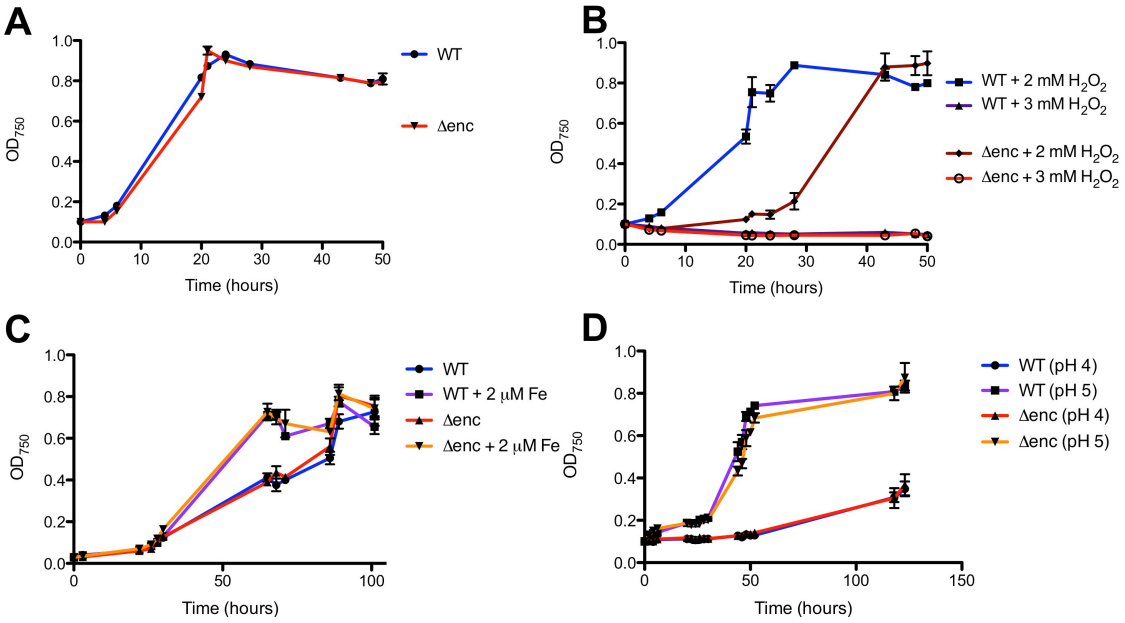

Figure 4.5: Preliminary characterization of an Encapsulin knockout mutant in *M. smegmatis* mc^{2}155. Growth curves comparing WT and Δ*enc* was performed in A) "normal" 7H9+ADS+Tylox$^{0.01\%}$+Hyg^{100} media; B) 7H9+ADS+Tylox$^{0.01\%}$+Hyg^{100} containing either 2 mM or 3 mM H_2O_2; C) IDM media supplemented with 2 μM FAC; and D) media at a pH of either 4 or 5. WT = blue and purple; Δ*enc* and orange

DISCUSSION

Although work has been carried out in attempts to understand the role of Mt-Enc in mycobacterial physiology, there is still plenty that remains unknown. This protein has been elusive, as we have yet to consistently observe its presence if we rely on endogenous expression levels. However, this may be due to the α-Mt-Enc antibody not being specific enough for Ms-Enc. Mt-Enc expression is hypothesized to be low under normal growth conditions, although it is possible that increased Mt-Enc mRNA levels can be detected when mycobacteria experience oxidative stress *if* Mt-Enc is in fact involved in mycobacterial survival under such duress. Based on our studies, we have shown that Mt-Enc is not essential, as the Δ*enc* knock out strain is still able to survive. Under normal growth conditions, we are unable to detect mycobacterial Enc in *M. smegmatis* mc^2155, possibly due to the α-Mt-Enc unable to detect Ms-Enc. Very preliminary co-immunoprecipitation experiments with polyclonal αMt-Enc and *M. smegmatis* mc^2155 lysate revealed the presence of Ms-Enc in cultures grown at pH 4.0 (data not shown). We assumed that at pH 4.0, Ms-Enc was able to be unfolded, therefore making it trypsin digestion sites accessible to trypsin protease. The stability of Ms-Enc at pH 4.0 is comparable to what was observed for Rj-Enc where at pH 3.0, the nanocompartment disassembled [4]

To study localization of Mt-Enc, it is not possible to rely on endogenous expression until we determine the appropriate conditions upon which the mycboacterial cell would need to produce more Enc. Mt-Enc was first identified as a culture filtrate protein with antigenic properties [3]. Examination of the amino acid sequence of Mt-Enc as well as cross-reactivity of anti-Mt-Enc antibody with ESAT6/CFP-10 revealed a

unique sequence, reminiscent of the TVII secretion signal [7]. If Mt-Enc is secreted, we hypothesized it would be via TVIISS. Since the mycobacterial capsule layer was reported [11], it was suggested that ESX-1 proteins might be associated with the capsule so Mtb can interact with the macrophage membrane. It can be argued that, under normal growth conditions (media containing Tyloxapol to prevent clumping, with gentle agitation), the capsule layer is sloughed off, therefore releasing proteins normally found in the capsule layer but have been misidentified as secreted, culture filtrate proteins. Misidentification of culture filtrate proteins is a concern in regards to studying TVIISS, as described my Mehra, *et al* when attempting to identify proteins that are legitimately secreted and were not just found in the culture filtrate; one problem they identify is that cytoplasmic proteins found in the culture filtrate were due to bacterial lysis [27]. It is possible that Mt-Enc may just be associated with the cell membrane and not actually secreted to the extracellular milieu, as it is difficult to conceive that a fully assembled nanocompartment is secreted. Alternatively, because Rj-Enc is able to fall apart at pH 3 and then reassemble into different subunit composition [4], it is possible that different oligomers of Mt-Enc can be secreted and then reassembled in the extracellular space allowing for other Mtb secreted proteins to be compartmentalized. To date, however, there is no data in the literature suggesting the later.

One of Mt-Enc cargos is and iron-storing Mtb ferritin (Mt-BfrB), has also been reported to be secreted and it also lacks a signal peptide [28, 29]. Recent reports show that ESX3-type secretion, one of the five types of TVIISS, is important for *M. tuberculosis* and *M. smegmatis* adaptation to low-iron conditions [30, 31]. It is enticing to hypothesize that Mt-BfrB is secreted via Mt-Enc thereby facilitating iron acquisition

and storage on an extracellular level; however the main role for Mt-BfrB has been characterized as an iron storage protein. The antioxidant activities of Mt-BfrB lies in isolating iron to protect the cell from peroxides that can be generated from the Fenton reaction. We have previously shown that Mt-BfrB maintains activity within the Mt-Enc compartment [22]. We purified Mt-BfrB and Mt-Enc$_{His}$-BfrB and iron-loaded by monitoring the ferroxidase activity as previously described [22] and observed proteins via electron microscopy (EM) to observe the iron core within the protein (**Figure 4.6, +Fe vs - Fe**), as iron can be visualized via EM [32]. It was not immediately evident based on EM images that neither Mt-BfrB nor Mt-Enc$_{His}$-BfrB was iron-loaded. We attributed this to growing cells and expressing them in LB iron-rich media has loaded these proteins with iron prior to adding more iron via ferroxidase assay. To verify if there was difference in iron content between Mt-BfrB and Mt-Enc$_{His}$-BfrB (- Fe vs + Fe), samples were analyzed by laser-induced fluorescence (LIF, exciting the samples at λ_{ex}= 224 nm). LIF confirmed that both Mt-BfrB (+Fe) and Mt-Enc-BfrB$_{His}$ (+Fe) were iron-loaded, with Mt-BfrB containing more iron than Mt-Enc$_{His}$-BfrB; the iron-treated protein samples had higher iron content than the non-iron-treated protein (data not shown). One other possibility offered as to why it has been difficult to observe iron via EM was that visualization of the iron in the samples my not be possible within the limits of the systems available with our collaborators at the SALK institute.

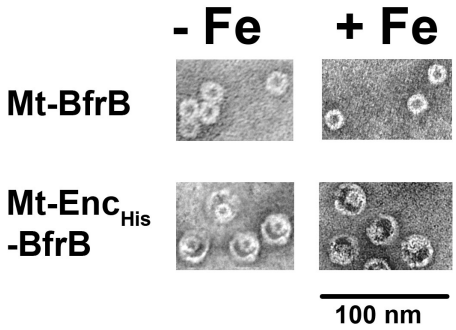

Figure 4.6: EM images of "apo-" protein (left, - Fe) and iron-loaded protein (right, +Fe). Top, Mt-BfrB and bottom, Mt-Enc$_{His}$-BfrB. Scale bar is representative of 100 nm (1000 Å).

To address Mt-Enc secretion we turned to a mycobacterial Tc-inducble system. Even though it is unlikely that overexpression of Mt-Enc would allow us to observe encapsulation of cargo enzymes, it would help us in determining its localization as well as the ability to complement Δ*enc* in *M. smegmatis* mc^{2}155. This work is in the very early stages. Preliminary work in regards to solving the crystal structure of Mt-Enc has begun via crystallization attempts with Mt-Enc$_{His}$-DyP and Mt-Enc$_{His}$-BfrB, as we were unable to obtain crystals of Mt-Enc$_{His}$ (**Figure 4.2A and C**). It may be that cargo proteins help stabilize the compartment. It is possible that the nanocompartment assembles with a different number of subunits—different assembly capabilities can help accommodate multiple cargos and substrates to enter pores. It could be argued that different number of Enc nanocompartment assemblies exist when purifying Mt-Enc$_{His}$ without cargo. Obtaining the structure of Mt-Enc (in complex with either Mt-Enc$_{His}$-DyP and Mt-Enc$_{His}$-BfrB) will help us exploit its properties similar to what has been observed [33], where targeting of Tm-Enc nanocompartment to cancer cells by utilizing a surface-exposed loop and engineering a targeting ligand. The therapeutic possibilities of Mt-Enc have only begun to be explored.

REFERENCES

1. Sutter, M., et al., *Structural basis of enzyme encapsulation into a bacterial nanocompartment.* Nat Struct Mol Biol, 2008. **15**(9): p. 939-47.
2. Valdes-Stauber, N. and S. Scherer, *Isolation and characterization of Linocin M18, a bacteriocin produced by Brevibacterium linens.* Appl Environ Microbiol, 1994. **60**(10): p. 3809-14.
3. Rosenkrands, I., et al., *Identification and characterization of a 29-kilodalton protein from Mycobacterium tuberculosis culture filtrate recognized by mouse memory effector cells.* Infect Immun, 1998. **66**(6): p. 2728-35.
4. Rahmanpour, R. and T.D. Bugg, *Assembly in vitro of Rhodococcus jostii RHA1 encapsulin and peroxidase DypB to form a nanocompartment.* FEBS J, 2013. **280**(9): p. 2097-104.
5. McHugh, C.A., et al., *A virus capsid-like nanocompartment that stores iron and protects bacteria from oxidative stress.* EMBO J, 2014. **33**(17): p. 1896-911.
6. Stanley, S.A., et al., *Acute infection and macrophage subversion by Mycobacterium tuberculosis require a specialized secretion system.* Proc Natl Acad Sci U S A, 2003. **100**(22): p. 13001-6.
7. Daleke, M.H., et al., *General secretion signal for the mycobacterial type VII secretion pathway.* Proc Natl Acad Sci U S A, 2012. **109**(28): p. 11342-7.
8. Abdallah, A.M., et al., *Type VII secretion--mycobacteria show the way.* Nat Rev Microbiol, 2007. **5**(11): p. 883-91.
9. Sorensen, A.L., et al., *Purification and characterization of a low-molecular-mass T-cell antigen secreted by Mycobacterium tuberculosis.* Infect Immun, 1995. **63**(5): p. 1710-7.
10. Berthet, F.X., et al., *A Mycobacterium tuberculosis operon encoding ESAT-6 and a novel low-molecular-mass culture filtrate protein (CFP-10).* Microbiology, 1998. **144 (Pt 11)**: p. 3195-203.
11. Sani, M., et al., *Direct visualization by cryo-EM of the mycobacterial capsular layer: a labile structure containing ESX-1-secreted proteins.* PLoS Pathog, 2010. **6**(3): p. e1000794.
12. Korch, S.B., H. Contreras, and J.E. Clark-Curtiss, *Three Mycobacterium tuberculosis Rel toxin-antitoxin modules inhibit mycobacterial growth and are expressed in infected human macrophages.* J Bacteriol, 2009. **191**(5): p. 1618-30.
13. van Kessel, J.C. and G.F. Hatfull, *Recombineering in Mycobacterium tuberculosis.* Nat Methods, 2007. **4**(2): p. 147-52.
14. Goude, R. and T. Parish, *Electroporation of mycobacteria.* J Vis Exp, 2008(15).
15. Jones, C.M. and M. Niederweis, *Role of porins in iron uptake by Mycobacterium smegmatis.* J Bacteriol, 2010. **192**(24): p. 6411-7.
16. Gossen, M. and H. Bujard, *Anhydrotetracycline, a novel effector for tetracycline controlled gene expression systems in eukaryotic cells.* Nucleic Acids Res, 1993. **21**(18): p. 4411-2.
17. Lowry, O.H., et al., *Protein measurement with the Folin phenol reagent.* J Biol Chem, 1951. **193**(1): p. 265-75.
18. Bradford, M.M., *A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding.* Anal Biochem, 1976. **72**: p. 248-54.

19. Besra, G.S., *Preparation of cell-wall fractions from mycobacteria.* Methods Mol Biol, 1998. **101**: p. 91-107.
20. Somerville, W., et al., *Extraction of Mycobacterium tuberculosis DNA: a question of containment.* J Clin Microbiol, 2005. **43**(6): p. 2996-7.
21. Sambrook, J.F., Russell, D. W., *Molecular Cloning: A Laboratory Manual*. 3rd ed. 2001, Cold Spring Harbor: Cold Spring Harbor Press.
22. Contreras, H., et al., *Characterization of a Mycobacterium tuberculosis nanocompartment and its potential cargo proteins.* J Biol Chem, 2014. **289**(26): p. 18279-89.
23. Ehrt, S., et al., *Controlling gene expression in mycobacteria with anhydrotetracycline and Tet repressor.* Nucleic Acids Res, 2005. **33**(2): p. e21.
24. Fiss, E.H., S. Yu, and W.R. Jacobs, Jr., *Identification of genes involved in the sequestration of iron in mycobacteria: the ferric exochelin biosynthetic and uptake pathways.* Mol Microbiol, 1994. **14**(3): p. 557-69.
25. Hall, R.M., et al., *Iron transport in Mycobacterium smegmatis: occurrence of iron-regulated envelope proteins as potential receptors for iron uptake.* J Gen Microbiol, 1987. **133**(8): p. 2107-14.
26. Rao, M., et al., *Intracellular pH regulation by Mycobacterium smegmatis and Mycobacterium bovis BCG.* Microbiology, 2001. **147**(Pt 4): p. 1017-24.
27. Mehra, A., et al., *Mycobacterium tuberculosis type VII secreted effector EsxH targets host ESCRT to impair trafficking.* PLoS Pathog, 2013. **9**(10): p. e1003734.
28. Weldingh, K. and P. Andersen, *Immunological evaluation of novel Mycobacterium tuberculosis culture filtrate proteins.* FEMS Immunol Med Microbiol, 1999. **23**(2): p. 159-64.
29. Rosenkrands, I., et al., *Hypoxic response of Mycobacterium tuberculosis studied by metabolic labeling and proteome analysis of cellular and extracellular proteins.* J Bacteriol, 2002. **184**(13): p. 3485-91.
30. Serafini, A., et al., *The ESX-3 secretion system is necessary for iron and zinc homeostasis in Mycobacterium tuberculosis.* PLoS One, 2013. **8**(10): p. e78351.
31. Serafini, A., et al., *Characterization of a Mycobacterium tuberculosis ESX-3 conditional mutant: essentiality and rescue by iron and zinc.* J Bacteriol, 2009. **191**(20): p. 6340-4.
32. Khare, G., P. Nangpal, and A.K. Tyagi, *Unique Residues at the 3-Fold and 4-Fold Axis of Mycobacterial Ferritin Are Involved in Oligomer Switching.* Biochemistry, 2013. **52**(10): p. 1694-1704.
33. Moon, H., et al., *Developing genetically engineered encapsulin protein cage nanoparticles as a targeted delivery nanoplatform.* Biomacromolecules, 2014. **15**(10): p. 3794-801.

CHAPTER 5

Conclusions and Future Directions: Understanding and Exploiting The Therapeutic and Immunization Capabilities of the Mycobacterial Proteinaceous Nanocompartment, Encapsulin

SUMMARY

Bacterial microcompartments have been observed in across bacteria. Carboxysomes as well as Pdu/Eut microcompartments are reported to function by both concentrating substrates for enzymes involved in metabolic processes as well as protecting the enzymes from the environment. We have presented the initial characterization a new bacterial compartment, a nanocompartment called Encapsulin (Enc) from *Mycobacterium tuberculosis* (Mt-Enc). We also carried out functional characterization of one of the putative Enc cargo enzymes, a dye-decolorizing peroxidase (DyP), as a heme-peroxidase; we have also started crystallization attempts for Mt-DyP. Mt-Enc self-assembles into a 60-subunit nanocompartment, capable of carrying protein cargo. The gene encoding for Mt-Enc lies directly downstream from the gene encoding for Mt-DyP in a two-gene operon, similarly to what is observed in other bacteria that harbor genes for Enc. Pull-down experiments reveal that Mt-Enc internalizes Mt-DyP via its C-terminal extension. The C-terminal extension is important for protein internalization within the Mt-Enc nanocompartment. Other putative cargo antioxidant enzymes, Mtb ferritin (Mt-BfrB) and Mtb 7,8-dihydroneopterin aldolase (Mt-FolB), also interact with Mt-Enc via their C-terminal extension; upon deletion of the C-terminal extension, interaction with Enc is

abolished. Furthermore, cargo enzymes remain active within the compartment. Additionally, we constructed a knockout Enc mutant in the fast-growing, non-pathogenic strain of mycobacteria, *Mycobacterium smegmatis* to determine the physiological role of mycobacterial Enc. Lastly, we are attempting to structurally characterize Mt-Enc (using the complexes of Mt-Enc-DyP and Mt-Enc-BfrB) by structure determination utilizing X-ray crystallography. Structural characterization will provide structural elements of Mt-Enc to use as either part of an immunization program or drug delivery system. Furthermore, characterization of Mt-Enc has been *in vitro,* thus *in vivo* studies to characterization mycobacterial Enc is necessary to begin to understand the role Mt-Enc may have in pathogenesis.

MYCOBACTERIAL COMPARTMENTALIZATION BY Mt-Enc

Bacterial compartmentalization can increase the local concentrations of functionally related enzymes and cofactors while confining unstable or harmful reactants against the complexity of a cell or the extracellular milieu [1, 2]. The protein-based compartments have been considered bacterial organelles for their ability to confine metabolic processes, similar to mitochondria. There are four types of bacterial compartments: carboxysomes, Pdu/Eut microcompartments, iron-storing ferritins, and encapsulins, [1-4]. Encapsulins were originally identified in *Thermatoga maritima*, Tm-Enc, and whose structure was solved [4]. Tm-Enc assembles into a 60-subunit icosahedral compartment, reminiscent of a viral capsid. The gene encoding for Tm-Enc extists in a two-gene operon with the gene encoding for a ferritin-like protein (Flp), located directly upstream of Tm-Enc, suggesting tight translational coupling [4]. A genome-wide analysis search across several bacterial

species reveals that genes encoding for Enc proteins also exist in a two-gene operon with either Flps or DyPs [4]. Mtb also harbors a gene encoding for both Enc and DyP [5]. Prior to Mt-Enc identified as a nanocompartment, one previous study identified it as a secreted culture filtrate protein with T-cell stimulatory properties [6]. We were able to heterologously express and purified Mt-Enc as a fully assembled nanocompartment [5]. We observed by electron microscopy (EM) that Mt-Enc assembles into an icosahedral shell similar in size to that of Tm-Enc [4]. Furthermore, we were able to show that Mt-DyP interacts with Mt-Enc—an interaction that is abolished upon truncation of the C-terminus [5]. We were also able to identify other interacting proteins based on their extended C-termini. For example, although Mt-Enc lies in a two-gene operon with Mt-DyP, Mtb also encodes a ferritin-like protein, Mt-BfrB, that is important for iron storage as well as pathogenesis [7, 8]. Mt-BfrB also harbors a C-terminal extension that is necessary for interaction with Mt-Enc [5]. Mt-FolB also has a C-terminal extension that is important for interaction with Mt-Enc [5].

We were also able to show that Mt-Enc self-assembles into an icosahedral shell with and without cargo, similar to what has previously been observed with Tm-Enc, *Brevibacterium linens* Enc (Bl-Enc), and *Rhodococcus jostii* Enc (Rj-Enc) [4, 5, 9]. EM analysis shows that fully assembled cargo enzymes reside within Mt-Enc; biochemical analyses show that cargo enzymes remain active within Mt-Enc [5]. Cargo enzymes within Rj-Enc have also been shown to retain peroxidase activity [9]. Taken together, these results suggested that Mt-Enc can contain functional cargo enzymes, but are not necessary for Mt-Enc nanocompartment assembly. Our study and others report that proteins can be targeted to the interior of the compartment if cargo enzymes have a C-

terminal extension [5, 10]. Based on these findings, Mt-Enc has become an attractive alternative for biotherapeutic delivery.

STRUCTURAL CHARACTERIZATION OF Mt-Enc: PROBING ASSEMBLY AND CARGO INTERACTIONS

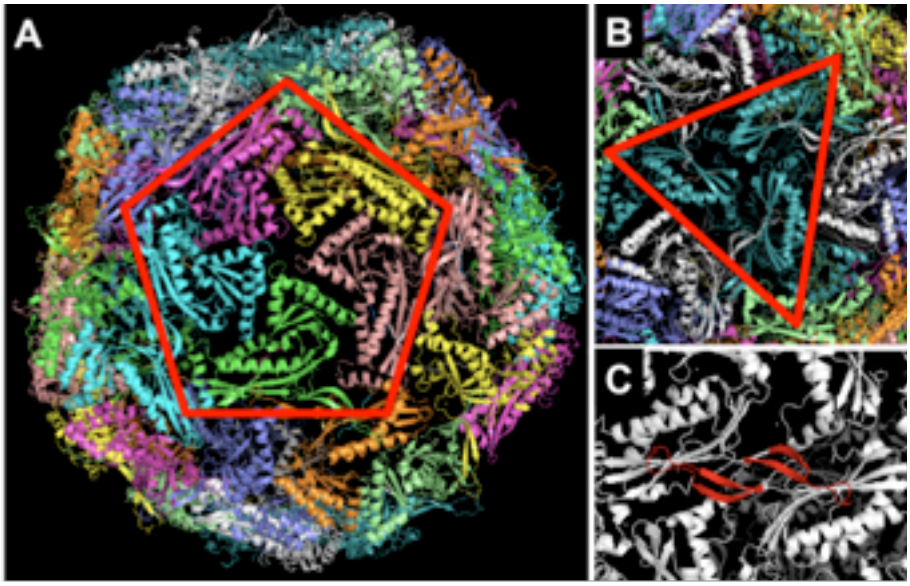

Figure 5.1: Structure of Tm-Enc (PDB: 3DKT). **A** 60-subunit icosahedral Tm-Enc nanocage with 5-fold pore (red pentagon). **B** close-up of the 3-fold pore (red triangle). **C** Two E-loops colored red from separate monomers

Structural contributions to Mt-Enc assembly and cargo packaging are poorly understood. To address this, we have initialized crystallization attempts of Mt-Enc. Because we were unable to obtain Mt-Enc crystals, we began Mt-Enc-DyP and Mt-Enc-BfrB crystallization. Diffracting protein crystals of both Mt-Enc-DyP and Mt-Enc-BfrB have been obtained. In order to begin critical analysis of nanocompartment assembly, the previously published structure of Tm-Enc [4] could be used to further probe assembly. It is possible that self-assembly may be mediated at the points of symmetry of Tm-Enc (2-

fold, 3-fold, or 5-fold (**Fig. 1)**), similar to what has been observed in regards to the self assembly at the 4-fold axis of symmetry for the Mt-BfrB nanocage [11]

ENCAPSULINS AS BIOTHERAPUTICS

Enzyme encapsulation for treatment is not entirely novel. Liposomes have been studied as a way to deliver enzymes to treat biofilms by encapsulating enzymes that help generate bactericidal agents [12]. One problem encountered with liposome encapsulation was the low encapsulation efficiency, where best observed liposome encapsulation was ~ 40% [13]. Because sometimes enzymes that are to be encapsulated may be susceptible to damage, it is important to understand both nanocompartment assembly and the region to which cargo enzymes are targeted to ultimately increase encapsulation efficiency.

Enc nanocompartments have been observed to be very stable, thus making them attractive biotherapeutic tools [9, 14]. Their unique immunogenicity and bioavailability make them excellent candidates for antigen delivery for novel immunization strategies as well as antimicrobial therapeutics. Furthermore, the durability/stability observed in both Rj-Enc [9] and Tm-Enc [15] also make them attractive candidates for Trojan-horse delivery of therapies. For example, a recent study re-engineered Tm-Enc for delivery of a small tracking probe and anticancer molecule covalently linked to the interior of the compartment, where the cargo is released upon pH-associated changes upon cellular internalization [15]. Tm-Enc was specifically targeted to carcinomas via engineering a targeting ligand on the surface-exposed "E-loop" of Enc monomer (**Fig. 1C**). The structural modifications are tolerated by Enc and thus, may be a versatile drug delivery vehicle platform.

Novel nanoparticle associated strategies represent key research areas in the production of next-generation vaccines, thus antigen display by Enc provides a promising alternative to preparation within adjuvant liposomes [15, 16]. Fusing a small antigenic proteins or fragments of a highly antigenic protein, such as ESAT6 early secreted protein [17, 18], to the surface of Mt-Enc via the "E-loop" could generate an immune response superior to current vaccine strategies. Mt-Enc has great potential for use as a repetitive display system for the development of anti-TB vaccine. Given the general stability [14] and hollow structure of nanocompartment, fusion of Mt-Enc with antigens secreted by the pathogen can be genetically engineered into the peptide sequence of the nanocompartment to be displayed on its exterior.

IN VIVO CHARACTERIZATON OF Mt-Enc AND ITS ROLE IN MYCOBACTERIAL SURVIVAL WITHIN IN THE HOST

Mt-Enc can reside within macrophage, preventing phagosome maturation. Mt-Enc was proposed to be a secreted, antigenic protein [6, 19]. Mt-Enc was also reported to be membrane-enriched, where it may be slowly released from the membrane to become recognized by the immune system [6]. The antigenic properties lie in its T-cell stimulatory properties by eliciting an increase in interferon-γ, which in turn activates macrophage [6]. It is possible that Mt-Enc being predominantly found in the membrane may be due to making Mt-Enc available for immune recognition in early infection.

The optimum expression conditions of Mt-Enc are currently unknown. Preliminary studies suggest that Mt-Enc is predominantly found in the membrane and is important for early adaptation to host assault (see Chapter 4). Mt-Enc was reported to be

secreted, although lacking a signal peptide [6]. The possibility of Mt-Enc secretion via the newly discovered mycobacterial Type VII secretion system during times of environmental stress may be important for mycobacterial survival. We have also observed that in order to avoid the generation of peroxides via the Fenton reaction, *if* Mt-Enc is secreted, it may be secreted with Mt-BfrB to sequester the iron and preventing the generation of oxidants. A Δ*enc* in *Myxococcus xanthus* displayed hypersensitivity to cells treated with H_2O_2 compared to the untreated cells in iron-starving conditions [20]. It could be that the *M. smegmatis*Δ*enc* would display a similar phenotype. Although it may be that mycobacterial Enc is secreted with different cargo enzymes, depending on the environmental stress encountered. Studies have begun, addressing the identification of *in vivo* cargo while cells grow under various growth conditions.

CONCLUDING REMARKS

This dissertation presents an analysis of two new enzyme families. First, we began characterization of a mycobacterial protein that belongs to a novel family of bacterial compartments, Encapsulins. Mt-Enc can carry protein cargo involved in oxidative stress, while cargo enzymes maintain functionality. Second, we were also able to characterize one of the putative Mt-Enc cargo enzymes, Mt-DyP—part of a novel family of peroxidases—as a heme-dependent peroxidase that is not a deferrochelatase. Although we have made exciting new discoveries in regards to two previously uncharacterized mycobacterial proteins, much work still remains to be done to characterize both Mt-Enc and Mt-DyP. The discovery of Mt-Enc and its ability to carry functional enzymes may be important in further understanding the mechanisms by which

Mtb can survive within the human host. In a recent studies, Enc nanocompartments are being probed for promiscuous packaging of enzymes that contain the extended C-terminal extension [10, 14] as well as exploiting structural elements to target Enc compartments to cancer cells [15]. These studies are an indication of the important functions that Enc proteins can have in terms of disease treatment and prevention.

REFERENCES

1. Cannon, G.C., S. Heinhorst, and C.A. Kerfeld, *Carboxysomal carbonic anhydrases: Structure and role in microbial CO2 fixation.* Biochim Biophys Acta, 2010. **1804**(2): p. 382-92.
2. Kerfeld, C.A., S. Heinhorst, and G.C. Cannon, *Bacterial microcompartments.* Annu Rev Microbiol, 2010. **64**: p. 391-408.
3. Theil, E.C., M. Matzapetakis, and X. Liu, *Ferritins: iron/oxygen biominerals in protein nanocages.* J Biol Inorg Chem, 2006. **11**(7): p. 803-10.
4. Sutter, M., et al., *Structural basis of enzyme encapsulation into a bacterial nanocompartment.* Nat Struct Mol Biol, 2008. **15**(9): p. 939-47.
5. Contreras, H., et al., *Characterization of a Mycobacterium tuberculosis nanocompartment and its potential cargo proteins.* J Biol Chem, 2014. **289**(26): p. 18279-89.
6. Rosenkrands, I., et al., *Identification and characterization of a 29-kilodalton protein from Mycobacterium tuberculosis culture filtrate recognized by mouse memory effector cells.* Infect Immun, 1998. **66**(6): p. 2728-35.
7. Khare, G., et al., *Ferritin structure from Mycobacterium tuberculosis: comparative study with homologues identifies extended C-terminus involved in ferroxidase activity.* PLoS One, 2011. **6**(4): p. e18570.
8. Reddy, P.V., et al., *Iron storage proteins are essential for the survival and pathogenesis of Mycobacterium tuberculosis in THP-1 macrophages and the guinea pig model of infection.* J Bacteriol, 2012. **194**(3): p. 567-75.
9. Rahmanpour, R. and T.D. Bugg, *Assembly in vitro of Rhodococcus jostii RHA1 encapsulin and peroxidase DypB to form a nanocompartment.* FEBS J, 2013. **280**(9): p. 2097-104.
10. Tamura, A., et al., *Packaging guest proteins into the encapsulin nanocompartment from Rhodococcus erythropolis N771.* Biotechnol Bioeng, 2015. **112**(1): p. 13-20.
11. Khare, G., P. Nangpal, and A.K. Tyagi, *Unique Residues at the 3-Fold and 4-Fold Axis of Mycobacterial Ferritin Are Involved in Oligomer Switching.* Biochemistry, 2013. **52**(10): p. 1694-1704.

12. Kaszuba, M. and M.N. Jones, *Hydrogen peroxide production from reactive liposomes encapsulating enzymes.* Biochim Biophys Acta, 1999. **1419**(2): p. 221-8.
13. Colletier, J.P., et al., *Protein encapsulation in liposomes: efficiency depends on interactions between protein and phospholipid bilayer.* BMC Biotechnol, 2002. **2**: p. 9.
14. Rurup, W.F., et al., *Self-sorting of foreign proteins in a bacterial nanocompartment.* J Am Chem Soc, 2014. **136**(10): p. 3828-32.
15. Moon, H., et al., *Developing genetically engineered encapsulin protein cage nanoparticles as a targeted delivery nanoplatform.* Biomacromolecules, 2014. **15**(10): p. 3794-801.
16. Hsu, T., et al., *The primary mechanism of attenuation of bacillus Calmette-Guerin is a loss of secreted lytic function required for invasion of lung interstitial tissue.* Proc Natl Acad Sci U S A, 2003. **100**(21): p. 12420-5.
17. Brodin, P., et al., *ESAT-6 proteins: protective antigens and virulence factors?* Trends Microbiol, 2004. **12**(11): p. 500-8.
18. Boggaram, V., et al., *Early secreted antigenic target of 6 kDa (ESAT-6) protein of Mycobacterium tuberculosis induces interleukin-8 (IL-8) expression in lung epithelial cells via protein kinase signaling and reactive oxygen species.* J Biol Chem, 2013. **288**(35): p. 25500-11.
19. Giri, P.K., et al., *Proteomic analysis identifies highly antigenic proteins in exosomes from M. tuberculosis-infected and culture filtrate protein-treated macrophages.* Proteomics, 2010. **10**(17): p. 3190-202.
20. McHugh, C.A., et al., *A virus capsid-like nanocompartment that stores iron and protects bacteria from oxidative stress.* EMBO J, 2014. **33**(17): p. 1896-911.

APPENDIX:

Characterization of aRv0203 homolog in *Mycobacterium smegmatis*: examination of mycobacterial heme acquisition from the host.

ABSTRACT

Mycobacterium smegmatis (Msm) has a heme uptake system homologous to that of Mtb, hypothesized to be facilitated by heme-binding protein Rv0203. Amino acid sequence alignment shows that Y59, shown to be critical for heme binding in Mtb Rv0203 [1, 2] is conserved among the homologs. The sequence of three Msm homologs (MSMEG_5111, MSMEG_5498 and MSMEG_0243) terminate at the C-terminus just after Cys114 (Rv0203 numbering and is involved in a disulfide bond wi th Cys40 near the start of the mature protein) thus, missing α-helix 5 present in Rv0203. Preliminary studies show that α-helix 5 is not essential for Rv0203 heme binding. However, α-helix 5 may act as an anchor to the mycobacterial cell surface facilitated through a direct interaction or indirectly via small molecules or modifications. The goal of this research is to unravel the process by which Rv0203 facilitates heme transport. Heme is important to establishing infection in the human host, thus the study of proteins involved the mycobacterial heme uptake will lead to a better understanding of this deadly pathogen.

INTRODUCTION

Tuberculosis (TB) is still one of the leading killers among infectious diseases. New tools to prevent and treat this disease are urgently needed. Iron is an essential element for most pathogens, including Mtb, where it must acquire iron from its human host environment. In humans, transferrin iron accounts for less than 1% of the body's total iron whereas heme iron can represent greater than 80% [3]. One may speculate that mycobacteria are capable of acquiring iron from human heme sources. Hence iron acquisition pathways are well studied in

mycobacteria as their components are essential to mycobacterial viability. For several decades, studies have focused on Mtb iron uptake through siderophore-mediated avenues from host transferrin (Ratledge 2004), with the hope of finding a novel drug target; however the most abundant host iron source is heme from hemoglobin rather than iron-bound transferrin. Recently, a unique mycobacterial heme uptake system has been discovered [1] (**Figure A.1**). Mtb secreted heme-binding protein, Rv0203 (**Figure A.2B**), is proposed to bind heme from the macrophage environment [1, 2], shuttling heme back to the mycobacterial cell whereby putative membrane-associated heme transporters, MmpL3 and MmpL11 [4], transport heme into the mycobacterial cytosol. Once in the cytosol, heme is then degraded by heme-degrader MhuD by cleavage of the tetrapyrrol ring, thereby releasing iron for storage or cellular utilization [5, 6].

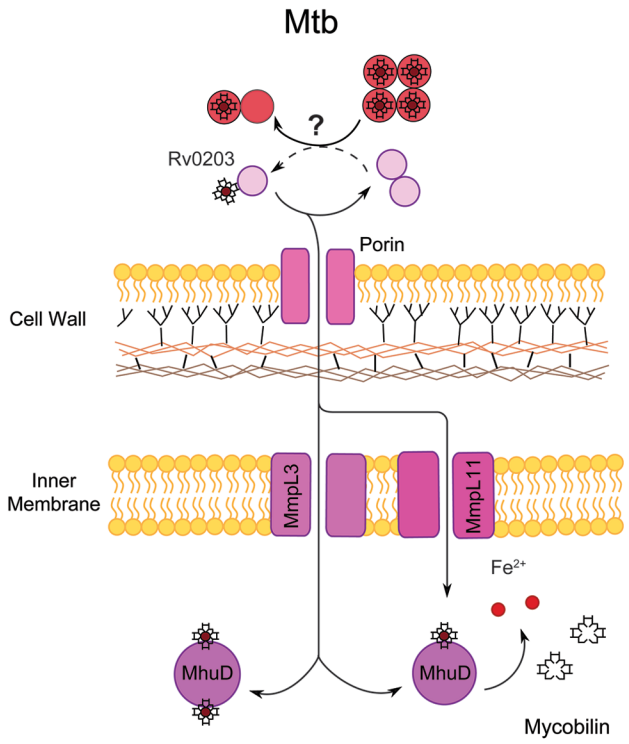

Figure A.1: Schematic representation of how host heme is acquired for Mtb survival. Light pink circles represent mycobacterial hemophore, Rv0203. Extracellular, host hemoglobin is represented by dark red circles, with with or without heme. Heme molecules are abstract symbols with smaller red circles signifying bound iron.

It has been shown that *M. smegmatis* has a heme uptake system homologous to that of *M. tuberculosis* [1, 7]. *M. smegmatis* has six homologs of Rv0203 (**Figure A.2A**), including a predicted transmembrane protein with a domain similar to Rv0203 (MSMEG2099, nitrate/sulfonate ABC transporter). Amino acid sequence alignment shows that Y59, shown to be critical for heme binding in *M. tuberculosis* Rv0203, is conserved among the homologs [1, 2]. The sequence of three *M. smegmatis* homologs (MSMEG_5111, MSMEG_5498 and MSMEG_0243) terminate at the C-terminus just after Cys114 (numbering from Rv0203 and is involved in a disulfide bond with Cys40 near the start of the mature protein) and thus, missing α-helix 5 present in Rv0203 (**Figure A.2A**). It is hypothesized that if all *M. smegmatis* homologs function as extracellular heme-binding proteins, the α-helix 5 is not essential for Rv0203 function (*i.e.*, heme binding). Furthermore, α-helix 5 may act as an anchor to the mycobacterial cell surface facilitated through a direct interaction or indirectly via small molecules. Thus, 'free' in solution (as in its recombinant purified form), α-helix 5 may interfere with heme binding, creating mixed heme-bound species or the lack of α-helix 5 may prevent heme binding.

Heme uptake is important for establishing infection within the human host, thus studies of proteins involved the mycobacterial heme uptake system will lead to a better understanding of this deadly pathogen. The crystal structure of apo-Rv0203 has already been sovled (**Figure A.2B**) [1]; however, heme-bound Rv0203 has not. Obtaining the heme-bound structure of Rv0203, or one of its homologs, may elucidate into the mechanism by which heme is acquired from host proteins. To address this, we turned to the *M. smegmatis* homologs. This work focused on the characterization of proteins involved in Mtb iron acquisition from heme by: i) biochemically characterizing the Rv0203 homolog, MSMEG_0243 from *M. smegmatis* and ii)

attempts to solve the heme-bound structure of this novel mycobacterial hemophore to begin understanding heme import to the mycobacterial cell.

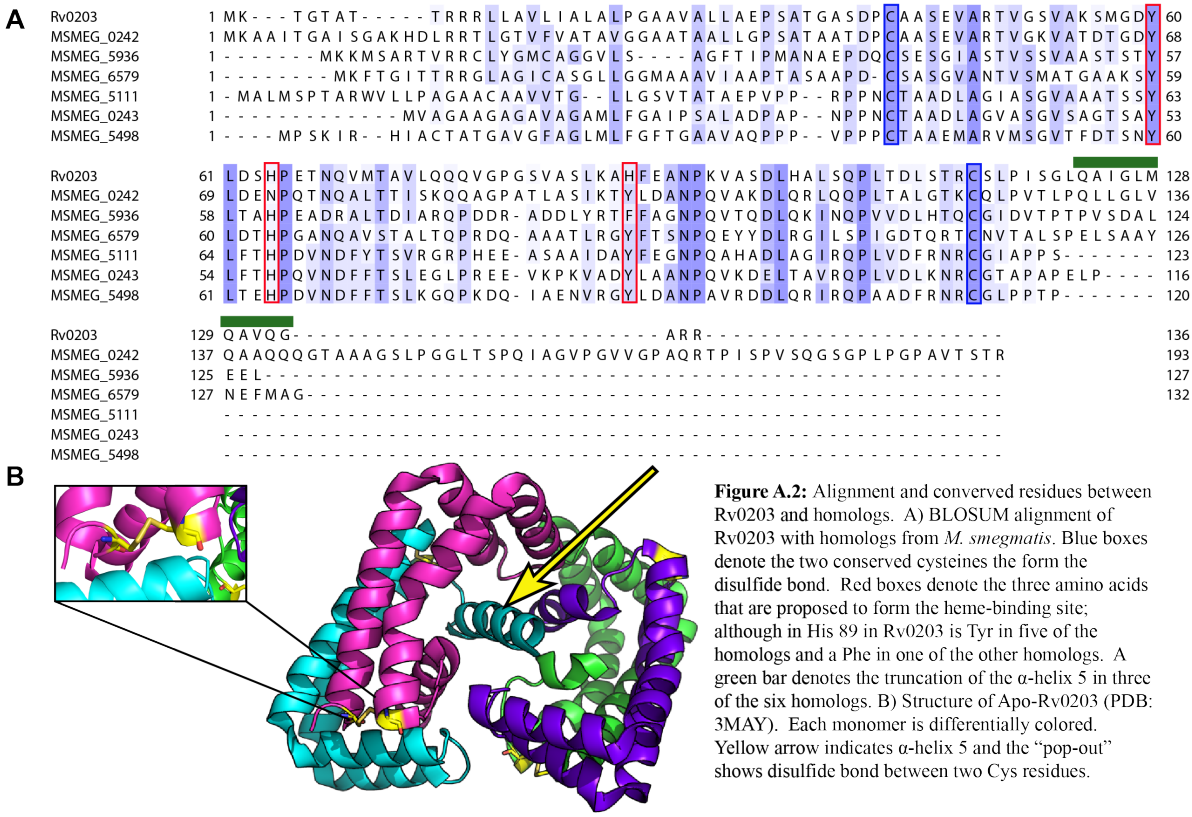

Figure A.2: Alignment and converved residues between Rv0203 and homologs. A) BLOSUM alignment of Rv0203 with homologs from *M. smegmatis*. Blue boxes denote the two conserved cysteines the form the disulfide bond. Red boxes denote the three amino acids that are proposed to form the heme-binding site; although in His 89 in Rv0203 is Tyr in five of the homologs and a Phe in one of the other homologs. A green bar denotes the truncation of the α-helix 5 in three of the six homologs. B) Structure of Apo-Rv0203 (PDB: 3MAY). Each monomer is differentially colored. Yellow arrow indicates α-helix 5 and the "pop-out" shows disulfide bond between two Cys residues.

MATERIALS AND METHODS

Cloning.

The Mtb gene encoding protein Rv0203 (*Rv0203*) was PCR-amplified from Mtb H37Rv genomic DNA (obtained from BEI Resources) using the KOD HotStart Polymerase Kit (Novagen) with 5′ and 3′ primers (MWG Operon) containing specific restriction sites outlined in **Table A.1**, to either generate a C-terminal poly-histidine tag (His-tag), or a C-terminal glutathione-S-transferase (GST) tag. PCR products were ligated into pCR-BluntII-TOPO (Invitrogen), and then transformed into *E. coli* OneShot TOP10 cells (Invitrogen). Double digestions with specific restriction enzymes were performed. Excised genes were ligated into the appropriate linearized *E. coli* expression vector and transformed into *E. coli* BL21-Gold (DE3)

cells (Novagen). Each construct was verified by DNA sequencing using T7 promoter and reverse primers (Laguna Scientific).

Table A.1: Primers for PCR-amplification and cloning.

Gene	Constructs	Primer Sequence
Rv0203	pGEX4T2::Rv0203	5`CCGGATCCGAACCATCAGCGACCGGCGCGTCGGACCCG 5'GGGCGGCCGCCTACTACCGGCGGGCGCCCTGCA
*Rv0203*αΔ5	pET28:: Rv0203αΔ5	5`CCCATATGGACCCGTGCGCGGCCAGCGAAGTGGCGAGGA 5`GGGCGGCCGCCTACTAGCTGATCGGCAGCGAGCACCGAGTCG
	pGEX4T2:: *Rv0203*αΔ5	5`CCCATATGGACCCGTGCGCGGCCAGCGAAGTGGCGAGGA 5`GGGCGGCCGCCTACTAGCTGATCGGCAGCGAGCACCGAGTCG
MSMEG_0243	pET28::MSMEG_0243	5`CCATATGCCGAACCCGCCCAACTGCACGGCCGCAGAC
	pGEX4T2::MSMEG_0243	5` CCGGATCCCCGAACCCGCCCAACTGCACGGCCGCAGACCT 5` GGGCGGCCGCCTACTACGGCAATTCAGGTGCCGGCGCGGT
*MSMEG_0243*Y89H	pET28::MSMEG_0243Y82H	5` CCAAGGTTGCCGACCATCTCGCAGCCAACCC 5`GGGTTGGCTGCGAGATGGTCGGCAACCTTGG
	pGEX4T2:MSMEG_0243Y82H	5` CCAAGGTTGCCGACCATCTCGCAGCCAACCC 5`GGGTTGGCTGCGAGATGGTCGGCAACCTTGG
*MSMEG_0243*Y89F	pET28::MSMEG_0243Y82F	5`CAAGGTTGCCGACTTTCTCGCAGCCAACCC 5'GGGTTGGCTGCGAGAAAGTCGGCAACCTTG
MSMEG_5111	pET28::MSMEG_5111	5`CCATATGCGGCCTCCGAACTGCACGGCCGCCGAC 5`GGGCGGCCGCCTACTAACTGGGCGGAGCGATACCGCAGCG
	pGEX4T2::5111	5` CCGGATCCCGGCCTCCGAACTGCACGGCCGCCGACCTC 5`GGGCGGCCGCCTACTAACTGGGCGGAGCGATACCGCAGCG
MSMEG_5498	pET28::MSMEG_5498	5`GCGATATGCCGCCGTGCACGGCAGCGGAAATGGCGCGGG 5`CGCGGCCGCTCATCAGGGCGTCGGGGGCAG
	pGEX4T2::MSMEG_5498	5` CCGGATCCCCGCCGCCGTGCACGGCAGCGGAAATGGCG 5` GGGCGGCCGCCGGTTGCGAAAATCGGCCGCGGGTTGGCGGA

Heterologous over-expression and purification of mycobacterial proteins.

Proteins were expressed from plasmids described in Table A.1, using E. coli BL21 Gold (DE3) cells. The following concentrations of antibiotics to the media were added where appropriate: kanamycin (30 μg/ml) and ampicillin (50 μg/ml). Cells harboring appropriate expression vector was grown aerobically at 37 °C in LB media containing the either kanamycin or ampicillin. Protein expression was induced at OD_{600nm} ~0.8 by the addition of isopropyl-β-d-thiogalactopyranoside (IPTG, 1 mM). Cells were harvested after four hours of induction, by centrifugation at 5100 x g for 20 min.

Harvested cell pellets were resuspended in 20 mL of Buffer A (50 mM Tris-HCl, pH 7.4, 350 mM NaCl, 10 mM imidazole and 10% glycerol), followed by the addition of phenylmethylsulfonyl fluoride (PMSF) and hen egg-white lysozyme. Cells were disrupted by sonication, clarified by centrifugation at 18000 x g for 30 min at 4° C and syringe-filtered (1.0

μm pore size) for removal of cell debris. The clarified cell lysate was then loaded onto a 5 mL Ni^{2+}-charged HisTrap column (GE Healthcare) pre-equilibrated with Buffer A. The protein(s) were eluted with a linear gradient of 10-500 mM imidazole (100 mL). Eluted fractions were collected, analyzed by SDS-PAGE, silver-stained, and concentrated using Amicon concentrators with the appropriate molecular-weight cutoff (Millipore, Bedford, MA). Proteins were further purified on a Superdex 200 HiLoad 10/30 gel filtration chromatography column (GE Healthcare) utilizing 50 mM Tris-HCl, pH 7.4, and 150 mM NaCl. Cleavage of the His_6 tag was conducted in cleavage buffer (50 mm Tris, pH 7.4, 150 mm NaCl, 10 mm $CaCl_2$) by adding 1 ml of thrombin-agarose suspension (Sigma) to the protein. After an overnight incubation at 4 °C, the thrombin-agarose was removed on a glass frit. Proteins were then run over an S75 gel filtration column equilibrated with 50 mm Tris, pH 7.4, 150 mm NaCl to separate proteins from the His_6 tag. Apo-protein concentration was determined by UV/vis spectroscopy, the molar extinction coefficient (**Table A.2**) as predicted by the program Protein Calculator (Scripps), or determined by modified Lowry [8] assay for heme-bound protein.

Table A.2: Estimated molecular weight and molar extincition coefficients.

Protein	Molecular Weight (kDa)	Extinction coefficent ($M^{-1}cm-1$)
Rv0203Δα5$_{His}$	10.4	1400
MSMEG_0243 $_{His}$	12.4	2680
MSMEG_5111 $_{His}$	10.8	3960
MSMEG_5498 $_{His}$	13.7	2680
GST_ Rv0203Δα5	35.4	41960
GST_MSMEG0243	39.0	43600
GST_MSMEG_5111	39.8	44520
GST_MSMEG_5498	40.2	48930

Heme binding of Rv0203Δα5 and M. smegmatis homologs.

Solutions were made and titration experiments were preformed as previously described [2, 9]. Heme [10] solutions were freshly prepared prior to each experiment. Briefly, approximately 4 mg of heme was dissolved in 1mL of ice-cold 100 mM NaOH and vortexed periodically over 20 min period. 1 mL of 1 M Tris (pH 7.4) was added to the solution, centrifuged for 10 min at 4 °C

at 13000 rpm. The supernatant was then diluted with 50 mM Tris (pH 7.4) and 150 mM NaCl, centrifuged again at 13000 rpm for 10 min to remove undissolved heme. Final concentrations were determined using an ε_{385} of 58.44 mM^{-1} cm^{-1}. Heme solutions were protected from light and used within 12 h.

The rate of heme dissociation from proteins was measured using H64Y/V68F apo-myoglobin as a heme scavenger protein. In brief, a dilute solution (approximately 10 μM) of H64Y/V68F-Mb [11] was acidified to pH 3.0 on ice, and heme was extracted using two equal-volume washes of ice-cold methyl ethyl ketone. Apo-H64Y/V68F-Mb was then extensively dialyzed at 4 °C against 50 mM Tris (pH 7.4) and 150 mM NaCl to remove methyl ethyl ketone and the protein concentration measured using an ε_{280} of 17 mM^{-1} cm^{-1}. Heme-bound proteins were incubated with 10-fold excess of H64Y/V68F apo-myoglobin, mixed in an SX.18MV stopped-flow spectrophotometer (Applied Photophysics) by equal volume and rapid mixing at room temperature. The reaction was monitored at 600 nm for a specified amount of time. The resulting time course was fitted to single- and double-exponential functions using Graph-pad Prism.

To measure the on-rate, stocks of ferrous CO-heme were made by anaerobically reducing fresh heme solutions at various concentrations with 2 mM sodium dithionite and then bubbling CO through the ferrous heme solutions. The heme concentrations of the ferrous CO-heme stock were verified using an ε_{407} of 147 mM^{-1} cm^{-1} [12]. Holoprotein samples were degassed, reduced with 2 mM sodium dithionite, and saturated with CO. Stopped-flow measurements for on-rate experiments were taken as previously described [2].

For low-iron experiments: Cell pellet was then washed three times in 10 mL of iron-deplete media (IDM) Cells were resuspended in 500 μL IDM, and an OD_{750} was taken. Resuspension was back-diluted to starting OD_{750}~ 0.05, upon which IDM was supplemented with either 1 μM ferric ammonium citrate (FAC) or 1 μM heme . Various time points were taken over 24 hours.

Crystallization conditions of heme-MSMEG_0243.

Proteins were purified and reconstituted with heme as described above. Purified protein was concentrated to ~ 18 mg/mL. A Mosquito nanoliter dispensing robot was used to set up 600 sparse matrix crystallization conditions for each protein. Crystallization conditions were optimized and additive screens were utilized. Crystals were grown using hanging-drop vapor diffusion techniques. Diffraction data from diffracting, single crystals were collected at either Stanford Synchrotron Radiation Laboratory (SSRL) or with Mtb Structural Genomics Consortium at the Advanced Light Source at UC Berkeley (ALS). Diffracting crystals of heme-bound MSMEG_0243 were grown at 18 mg/mL in 2 M $AmSO_2$ (ammonium sulfate), 0.1 M Tris-HCl pH 8.5, and 0.04% PPG (propylene glycol) 400 .

Fractionation of Mycobacterial Cells.

Fractionation of mycobacterial fractions was performed as previously described [13], although with modifications described herein. Mycobacteria were grown to desired density in appropriate media. Mycobacterial cells were harvested by centrifugation at 4000 x *g* for 15 min at 4°C. Supernatant was discarded and cells were washed with 1/5 of the original culture volume with pre-chilled phosphate-buffered saline + 0.05% Tween 80, pH 8.0 (PBST), and PMSF was added. Mycobacterial cells were disrupted as described above. Unbroken cells and debris was removed

by centrifuging sonicate 4000 x *g* for 10 min at 4°C. Centrifugation steps above 4000 x *g* were performed on a Beckman Ultracentrifuge using a TLS55 rotor. Supernatant was saved, then centrifuged at 27,000 x *g* for 45 min. at 4°C. The resulting pellet is the cell wall fraction (contains proteins and non-protein compounds such as arabinogalactan-peptidoglycan). The cell wall fraction/pellet was resuspended the pellet in 10 mM ammonium bicarbonate and quantified and supernatant was further processed.

The supernatant was centrifuged at 150,000 x *g* for 1 hour at 4°C. The supernatant is the soluble cytosolic fraction. The resulting membrane pellet was washed with PBS and resuspended in 10 mM ammonium bicarbonate and quantified.

For extraction of *all* the hydrophobic membrane proteins, 32% of pre-treated TritonX-114 (TX-114) was added to cell resuspension (prior to sonication) to a final concentration of 2% and extraction methods were followed. Supernatant was saved and incubated for 10 min at 37°C to allow phase partitioning, then centrifuged at 1000 x *g* for 10 min at room temperature. The upper aqueous phase, which contains hydrophilic proteins, was collected. The lower TX-114 detergent phase enriched for hydrophobic proteins was collected. The detergent phase was washed to minimize contamination by soluble proteins by adding a volume of mycobacterial lysis buffer corresponding with the volume of the aqueous phase previously removed. Incubated the resuspension for 10 min at 37°C, then centrifuge at 1000 x *g* for 10 min at room temperature to allow phase partitioning. The upper aqueous supernatant was carefully removed, then repeated three times. Proteins were precipitated from the detergent phase by adding 5 volumes of chilled 100% acetone, and incubated overnight at – 20°C. The next morning, preparation was centrifuged at 14,000 x *g* for 30 min at 4°C to sediment the proteins. The pellet was air-dried, then resuspended in 10 mM ammonium bicarbonate and quantified.

2-Dimensional Gel Electrophoresis (2-DGE)

2-DGE was performed as previously described [14]. Mycobacterial cultures that were analyzed were grown in either iron- or heme supplemented conditions as described above. Cytoplasmic and culture filtrate proteins used for 2-DGE analysis were pepared via TCA-precipitation [15]. Membrane proteins were isolated as described above, then TCA-precipitated in preparation for analysis. Isoelectric focusing (IEF) was performed with 150 μg of protein, on an IPG strip, pH 4-7 (BioRad, Hercules, CA), within a porcelain strip holder. IPGPhorII was run for approximately 9 hours, at 20°C (50 μA/strip, 1 step at 35,000 V-hrs with a ramp-up gradient to 8000 volts. After IPG strip was re-equilibrated with BioRad Equilibration buffers, strip was carefully placed inside an elongated well of a BioRad Criterion gel. IEF proteins were resolved by molecular weight at 150 V for 1 hour in MOPS buffer. In-gel trypsin digestion, mass-spectrometry, and identification was performed as previously described [16].

RESULTS

Purification of Rv0203Δα5 and *M. smegmatis* homologs.

First, to determine the affect of Δα5 on heme-binding of Rv0203, we first deleted the amino acids that form the alpha-helix. We originally cloned, expressed, and purified proteins with a cleavable, N-terminal glutathione-S-transferase (GST) tag; however continued protein characterization with GST-recombinant proteins was stopped because of difficulties arising from the GST tag. Therefore, we cloned, over-expressed and purified $Rv0203\Delta\alpha5_{His}$, $MSMEG\_0243_{His}$, $MSMEG\_5111_{His}$, $MSMEG\_5498_{His}$ (**Figure A.3**). Proteins did not easily stain with Coomassie blue, with the exception of $MSMEG\_0243_{His}$; thus gels analyzing proteins were silver-stained. Due to the high quantity obtained from purification, characterization was

focused on MSMEG_0243$_{His}$.

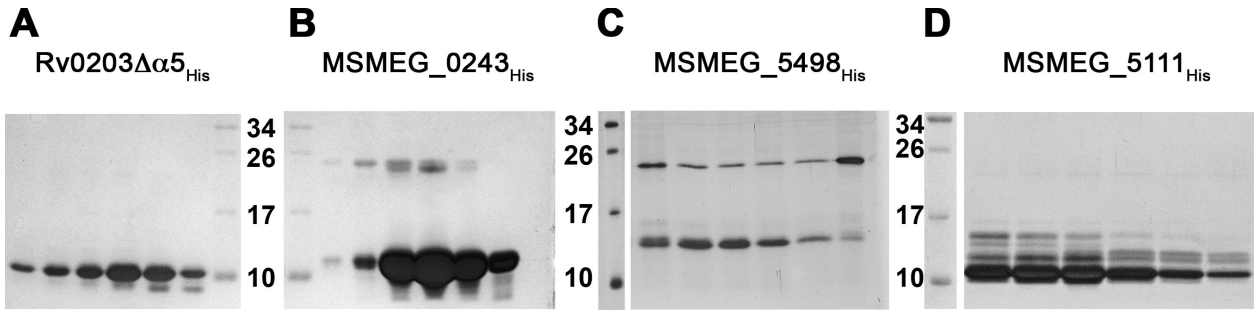

Figure A.3: Purification of Rv0203Δα5$_{His}$ and *M. smegmatis* homologs. Silver-stained SDS-PAGE of post-nickel affinity purification, with corresponding protein names above gel images. Expected sizes for purified proteins, in kDa, are: (A) Rv0203Δα5$_{His}$ ~10.4; (B) MSMEG_0243$_{His}$ ~12.4; (C) MSMEG_5498$_{His}$ ~13.7; (D) Rv0203Δα5$_{His}$ ~10.8. Bands above expected sizes are either dimmers or contaminating proteins for (B) and (C), respectively.

Solve the heme-bound structure of Rv0203 *M. smegmatis* homolog, MSMEG_0243. To date, the structure of heme-bound Rv0203 has not yet been solved. Preliminary data suggests that loss of α-helix 5 in MSMEG_0243 does not adversely affect heme-binding; however, the mode by which Rv0203 or MSMEG_0243 aids in heme transport into the mycobacterial cell is unknown. Therefore, attempts at obtaining the heme-bound structure utilizing heme-bound MSMEG_0243 are currently underway. Heme-MSMEG_0243 crystals were obtained with heme (Figure A.4A); although crystals diffracted (Figure A.4B) we were still unable to obtain high-resolution data. Furthermore, we also attempted to crystallize MSMEG_0243 with cyanoheme to stabilize the heme:protein complex. Diffracting crystals were obtained, although we were still unable to solve the structure, possibly due to mixed conformations of the heme molecule.

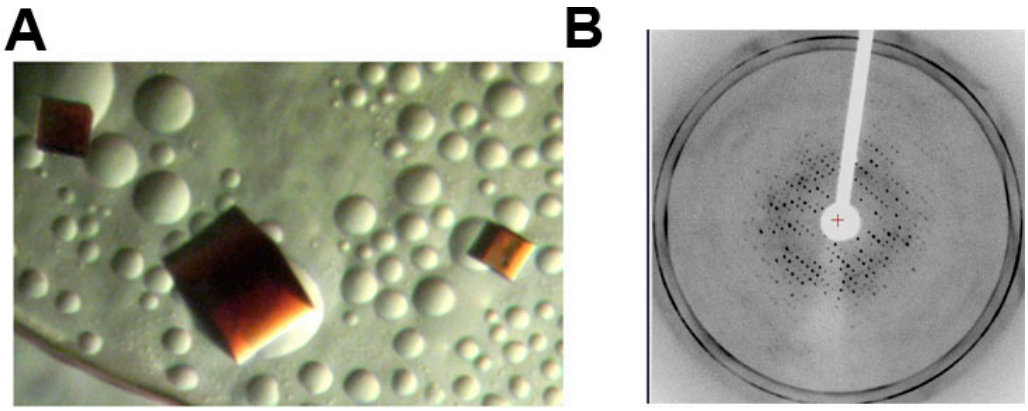

Figure A.4: Crystallization and diffraction of heme-MSMEG_0243. A) Cube-like heme crystals of MSMEG_0243 (2M $AmSO_4$, 0.1 M Tris, pH 8.5, 0.04% PPG 400) at 18 mg/mL. B) Diffraction of heme-MSMEG_0243 crystals, to resolution of 4.1 Å.

Biochemical characterization of Rv0203 *M. smegmatis* homolog, MSMEG_0243. MSMEG_0243 and Rv0203 have similar spectroscopic properties; however, the *M. smegmatis* homolog is missing the α-helix 5. Untagged MSMEG_0243 was reconstituted with heme at a 1:1 ratio. Loss of the α-helix 5 has not affected heme binding, but spectroscopic analysis indicates a difference in heme binding between Rv0203Δα5 and MSMEG_0243 (**Figure A.5**). Two peaks in the Soret region are observed in MSMEG_0243 (**Figure A.5A**), and indistinguishable features in the Q-band region compared to wild-type Rv0203 and Rv0203Δα5 (**Figure A.5B**) [2].

Heme-binding affinities of MSMEG_0243 and Rv0203Δα5 were measured using stopped flow techniques where a 10-fold excess of H64Y/V68F-Mb [11] were mixed in a stopped flow spectrophotometer by rapid, equal-volume mixing at room temperature. The reaction was monitored at 600 nm over 100 seconds for Rv0203Δα5 (**Figure A.6A and B**) or 200 seconds (**Figure A.6C-F**). Kinetic parameters are shown in **Table A.3**. MSMEG_0243 has a slightly higher heme-binding affinity to heme compared to Rv0203 and Rv0203Δα5 [2]. Loss of the α-helix 5 did not greatly affect heme binding of Rv0203Δα5 (**Table A.3**). These results suggest that the Rv0203 homolog, MSMEG_0243, can still bind heme without the α-helix 5, although

truncated Rv0203 does not bind as tightly. A preliminary heme on-rate determination of MSMEG_0243 was measured at k_{obs}= 42.47±12.45 $\mu M^{-1}s^{-1}$ (**Figure A.7**), three-times lower than k_{obs}= 133±19 $\mu M^{-1}s^{1}$ for wild-type Rv0203 [2]; however, this would need to be respeated as the calculated k_{obs} resulted in confidence value of R^2 ~ 88%.

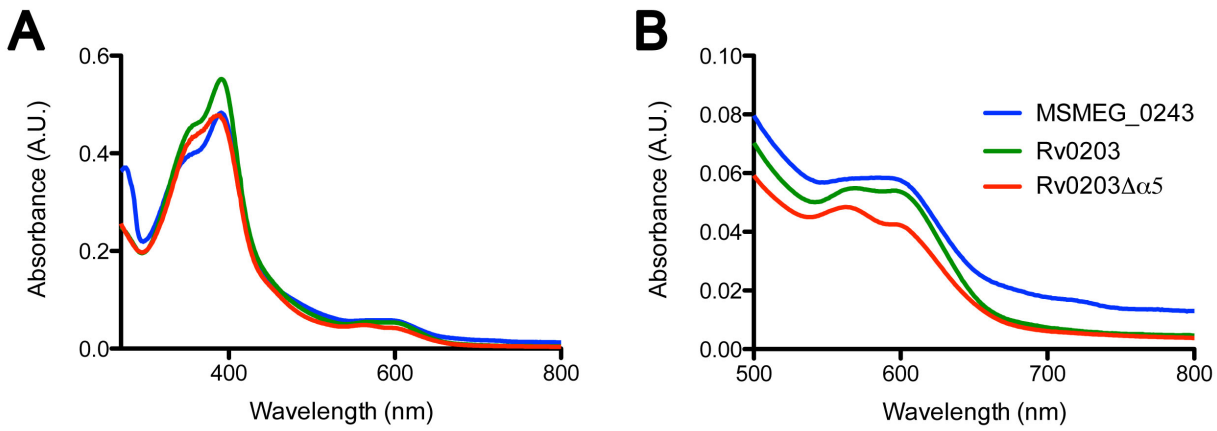

Figure A.5: Spectroscopic analyses between Rv0203Δα5 and MSMEG_0243. A) Spectra of ~ 10 μM heme-bound protein. B) Zoom-in of the spectra observed in (A), showing different Q-band features. Green trace is wild-type Rv0203, red is Rv0203Δα5 and blue is MSMEG_0243. The Soret of MSMEG_0243 is 391 nm with a distinct shoulder that 353 nm. Rv0203Δα5 has a Soret of 388 nm while wild-type has a Soret of 391 nm.

Table A.3: Kinetic parameters for binding of heme.

Protein	k_{-heme} (s^{-1})	Fit	Reference
Rv0203	8.2 ± 1.6 x 10^{-2}	Single	[2]
	k_f= 7.7 x 10^{-1} ± 5.7 x 10^{-2} k_s= 5 ± 1 x 10^{-1}	Double	[2]
Rv0203Δα5 (410 nm)	1.2 x 10^{-1} ± 3.0 x 10^{-3}	Single	This study
Figure A.6A	k_f= 4.7 x 10^{-2} ± 1.7 x 10^{-2} k_s= 3.4 x 10^{-2} ± 1.4 x 10^{-3}	Double	This study
Rv0203Δα5 (600 nm)	1.3 x 10^{-1} ± 3.3 x 10^{-3}	Single	This study
Figure A.6B	k_f= 2.9 x 10^{-2} ± 1.6 x 10^{-2} k_s= 2.1 x 10^{-2} ± 3.0 x 10^{-3}	Double	This study
MSMEG_0243 (410 nm)	2.9 x 10^{-1} ± 3.6 x 10^{-4}	Single	This study
Figure A.6C	k_f= 8.4 x 10^{-2} ± 4.7 x 10^{-3} k_s= 1.5 x 10^{-2} ± 7.4 x 10^{-4}	Double	This study
MSMEG_0243 (600 nm)	3.2 x 10^{-2} ± 1.0 x 10^{-3}	Single	This study
Figure A.6D	k_f= 1.9 x 10^{-1} ± 3.6 x 10^{-2} k_s= 1.9 x 10^{-2} ± 1.6 x 10^{-3}	Double	This study
MSMEG_0243 Y82H (600 nm)	1.3 x 10^{-2} ± 4.0 x 10^{-4}	Single	This study
N/A	k_f= 6.0 x 10^{-1} ± 2.1 x 10^{-2} k_s= 7.8 x 10^{-3} ± 1.6 x 10^{-4}	Double	This study
MSMEG_0243 Y82F (410 nm)	1.8 x 10^{-2} ± 3.0 x 10^{-4}	Single	This study
Figure A.6E	k_f= 1.4 x 10^{-1} ± 1.0 x 10^{-2} k_s= 1.0 x 10^{-2} ± 3.7 x 10^{-4}	Double	This study
MSMEG_0243 Y82F (600 nm)	1.3 x 10^{-2} ± 8.5 x 10^{-4}	Single	This study

Figure A.6F	$k_f = 1.7 \times 10^{-1} \pm 3.6 \times 10^{-2}$ $k_s = 3.6 \times 10^{-3} \pm 9.1 \times 10^{-4}$	Double	This study

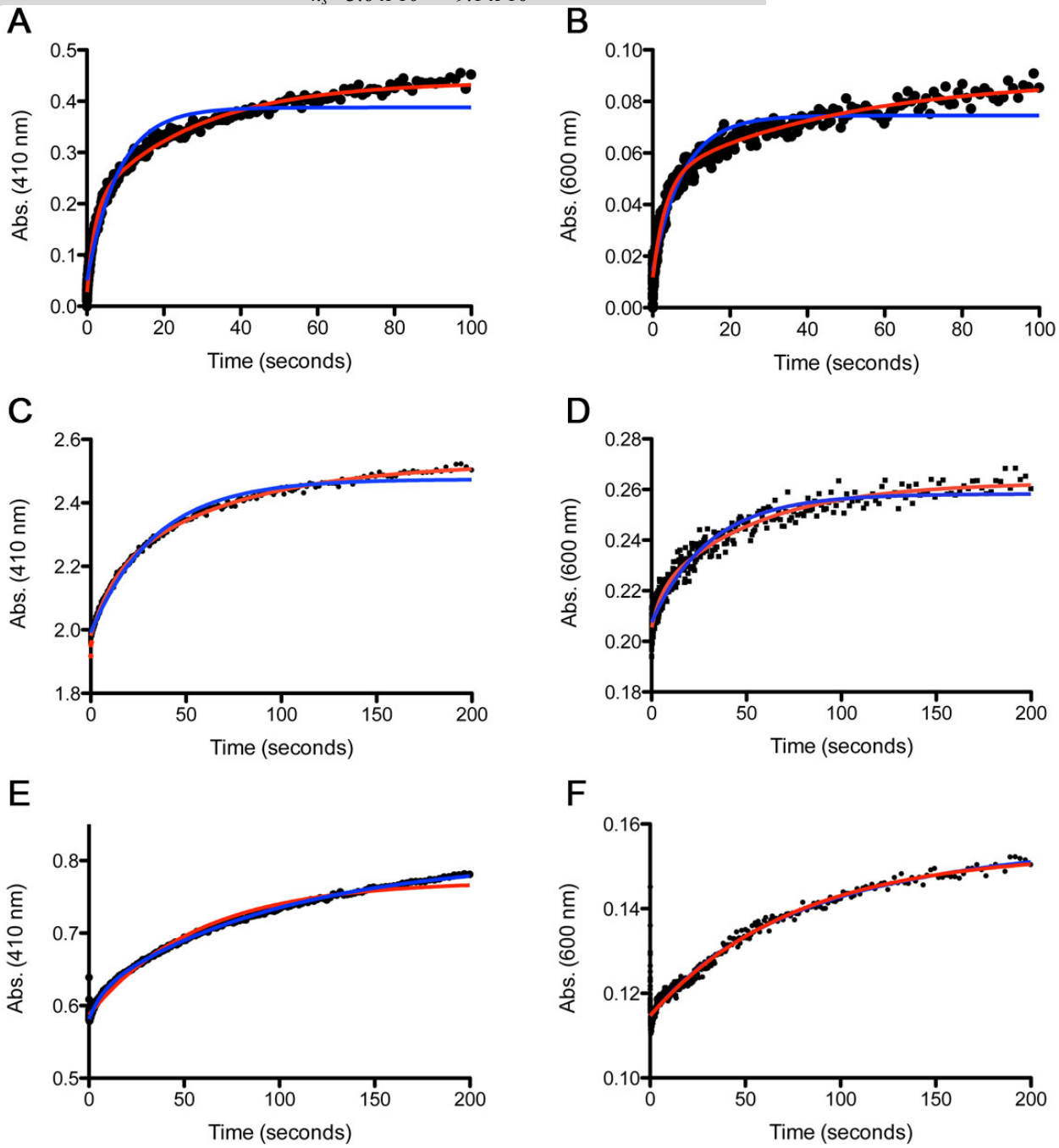

Figure A.6: Heme-off rate of Rv00203Δα5, MSMEG_0243, and mutants. Time courses of transfer of heme from holo-protein (A, C, E) measured at 410 nm and (B, D, F) measured at 600 nm, to a 10-fold excess amount of apo-H64Y/V68F-Mb. The exponential fitting results are noted in red (single exponential) or blue (double exponential). Experiments were conducted twice, and a representative change in absorbance vs. time is shown.

The proposed heme-binding site on Rv0203 consists of Tyr59, His63, and His89 [1, 2]. The hydroxyl group on Try59 may be a coordinating ligand to heme-iron and His63 and His89

do not directly coordinate heme; however, these His residues are important for heme binding as they may non-covalently interact with heme [2]. MSMEG_0243 has two Tyr residues and one His in the proposed heme biding site (**Figure A.2**). While the spectra of MSMEG_0243 seems to have non-descript features compared to Rv0203Δα5, it is still able to bind heme; however, because there is no single peak in the Soret region, it may be that MSMEG_0243 does not bind heme in a single conformation. It is possible that the Tyr substitution in MSMEG_0243 from His89 or another aromatic substitution, such as Phe, can further stabilize heme binding or result in a single heme-binding conformation. To elucidate the role of Tyr82 in MSMEG_0243, Tyr82 was mutated to His as in Rv0203. Tyr82 was also mutated to Phe as it may be that pi-mediated interactions arising from aromatic side-chains may help stabilize heme. Both Y82H and Y82F mutations had similar k_{-heme}, except when the data was fit to double exponential, results varied widely between Y82H and Y82F. Obvious spectral changes in heme binding were not observed via spectroscopic methods between mutants (data not shown). Therefore, we determined the heme-off rates of native MSMEG_0243 as well as the Y82H and Y82F mutants using stopped-flow techniques to compare relative heme binding affinities. MSMEG_0243 Y82F binds slightly tighter than Y82H; however, Y82H has similar binding to wild-type Rv0203 based on double exponential calculations [2]. The second Tyr (Y82) in the proposed heme-binding site of MSMEG_0243 exhibits similar binding affinities to Y82F mutant. Thus, it appears that Y82F and Y82 have slightly higher binding affinities compared to Y82H and wild-type Rv0203. These experiments do not further implicate the α-helix 5, suggesting that the second Tyr in the proposed heme-binding site may help stabilize the heme. Further analysis of Y82F still needs to be performed to further verify the stabilizing effects of Y82F and Y82.

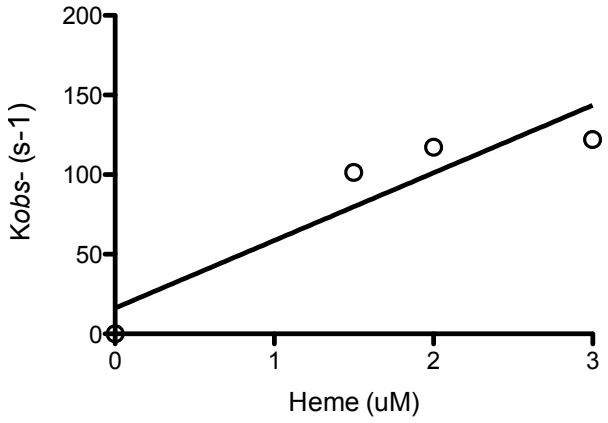

Figure A.7: Preliminary heme on-rate calculation for MSMEG_0243. Plot of k_{obs} of ferrous CO-heme binding vs protein concentration. Protein:heme ratio was 1:1to obtain an estimated rate constant for heme binding.

Preliminary analysis of differential protein expression patterns of *M. smegmatis* grown in either heme or iron.

Iron metabolism and uptake is elegantly regulated, as iron uptake in Mtb has been greatly examined [3, 17] in mycobacteria. In light of the discovery of heme-uptake in mycobacteria [1, 18], the mechanism by which heme is ushered back to the mycobacterial cell is unknown. We sought to perform a proteomic-wide study comparing *M. smegmatis* grown in iron or heme, alone, in hopes that it may elucidate key regulatory networks and pathways that are involved in either heme or iron uptake pathways. Similar proteomic experiments were carried out in *Staphylococcus aureus*, where a re-direction of metabolism was observed when bacteria were treated with either heme or iron [19]. We hypothesized a difference in metabolic and regulatory pathways in response to the mode of iron or heme-iron acquisition in *M. smegmatis*.

To observe differential protein expression patterns in iron-depleted cultures grown in the presence of either exogenous heme or iron, we performed cell fractionation in which culture filtrate and cytosolic proteins were analyzed by 2DGE (**Figure A.8A-C**); however, membrane proteins were isolated and analyzed by SDS-PAGE (**Figure A.8D**). Although there were no dramatic differences, we were able to observe minute differences between protein spots between

"iron" and "heme" samples. These minute differences could indicate the intricate orchestration of expression patterns within the cell. Spots were excised from the whole gel, performed in-gel trypsin digestion, and peptides were separated by μLC-MS/MS. Peptides were identified via MASCOT database (**Table A.4**).

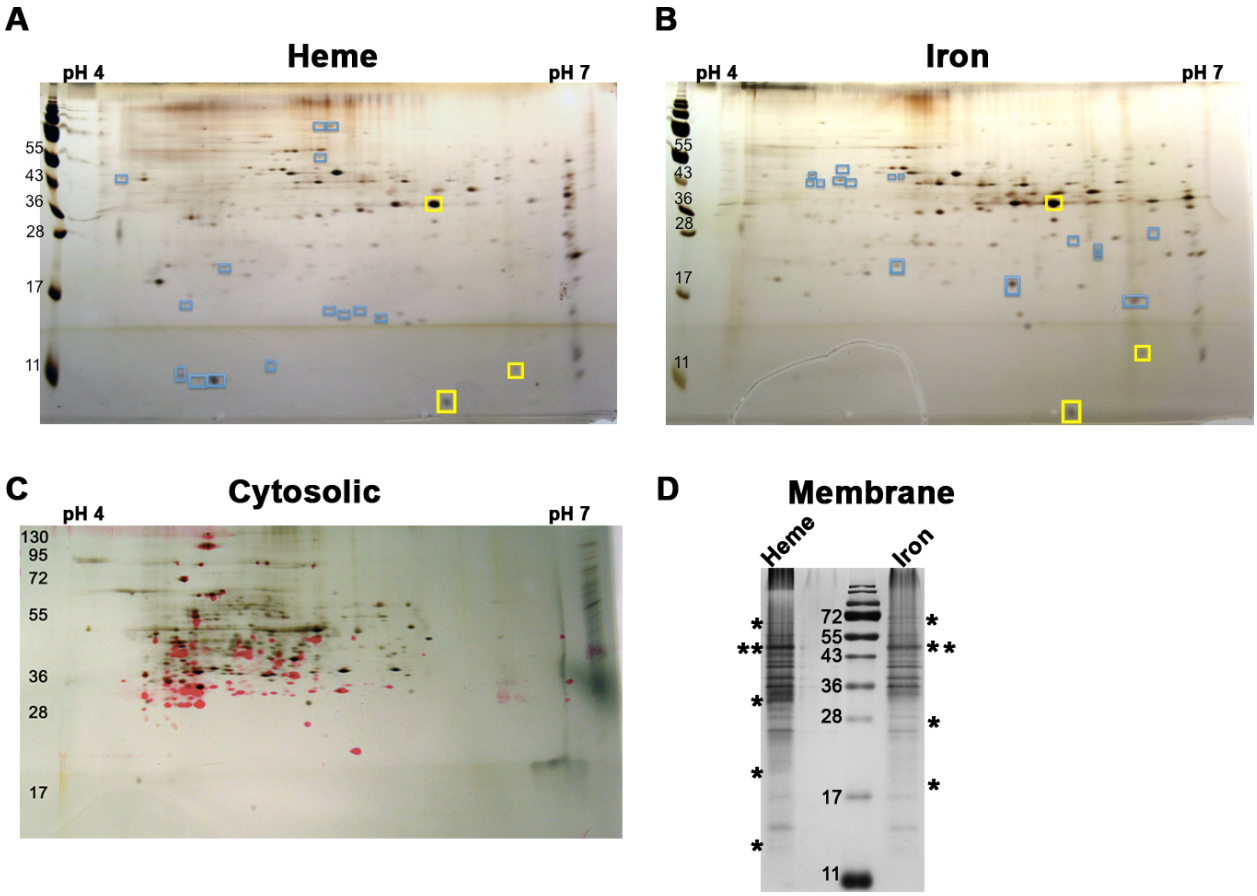

Figure A.8: Proteome-wide analysis of *M. smegmatis* mc^{2}155 cultures grown in either heme or iron. 2DGE of culture filtrate from cultures grown in either heme (A) or iron (B). Blue box spots are unique proteins, yellow boxes indicate similar proteins. C) Overlay of 2DGE of iron (magenta) or heme (black) cytosolic proteins. D) Membrane proteins isolated and analyzed via SDS-PAGE. Overlay better indicates differences between the two cultures. Two asterisks indicate "control" proteins, while single asterisk indicate possible unique proteins.

DISCUSSION AND CONCLUSION

This project had two main goals: 1) crystallize heme-bound homolog of Rv0203 and 2) identify differentially expressed cellular proteins when cultured in the presence of either iron or heme.

We were unable to obtain high-resolution data of a heme-bound Rv0203 homolog, MSMEG_0243. One of the reasons we believe this was not possible was probably due to lack of a single heme-binding conformation. This was likely the case was heme-Rv0203. It may be that we would need to identify a stabilizing protein for heme-Rv0203, and this partner may not only be important for stabilizing Rv0203, but this elusive binding partner may assist Rv0203 in transporting the heme molecule back to the mycobacterial cell. This is one of the reasons why we began 2DGE analysis.

Another possibility was the hemophore-like protein from *Porphyromonas gingivalis* exhibits a variety of isoforms, proposed to be a consequence of glycosylation or other posttranslational modifications to the expressed protein [20]. It is possible that Rv0203 has a variety of isoforms that we have not yet identified. Rv0203 would have to be extracted by hemin-agarose chromatography from Mtb culture filtrate and then 2DGE can be carried out on the eluted fraction. If a variety of Rv0203 isoforms are observed, then each spot on the 2DGE gel can be excised, in-gel trypsin digested, and then the peptides will be extracted and analyzed by μLC-MS/MS in an attempt to determine the modifications and the modified residues.

Mycobacterial heme uptake has already been identified in mycobacteria, but the orchestration by which heme is imported into the mycobacterial cell is still unknown. Work presented here are attempts to address the mechanisms by which heme is transported back to the cell. Although work has been initiated, a lot of work still needs to be done to fully understand mycobacterial heme uptake.

Table A.4: Peptide identification from proteome analysis. Cov = coverage; I.D. = identity

HEME MEMBRANE

Protein Name	ID	% Cov.	MW (kDa)	pI	Rv Homolog	I.D.(%)
eptc-inducible aldehyde dehydrogenase [Mycobacterium smegmatis str. MC2 155]	YP_885303	5	55.32	4.98	Aldehyde dehydrogenase (Rv0458)	87
extracellular solute-binding protein, family protein 5, putative [Mycobacterium smegmatis str. MC2 155]	YP_885056	15	60.637	4.53	ProW (Rv3757c)	25
catalase/peroxidase HPI [Mycobacterium smegmatis str. MC2 155]	YP_887765	17	81.948	4.99	KatG (Rv1908c)	66
dihydrolipoamide acetyltransferase [Mycobacterium smegmatis str. MC2 155]	YP_888560	14	61.049	4.65	DlaT (Rv2215)	74
hypothetical protein MSMEG_6143 [Mycobacterium smegmatis str. MC2 155]	YP_890364	2	63.746	8.99	Rv3635	72
catalase/peroxidase HPI [Mycobacterium smegmatis str. MC2 155]	YP_890597	5	81.121	4.73	KatG (Rv1908c)	71
NAD synthetase [Mycobacterium smegmatis str. MC2 155]	YP_888881	4	75.396	5.29	NadE (Rv2438c)	83
serine esterase, cutinase family protein [Mycobacterium smegmatis str. MC2 155]	YP_886454	10	31.084	5.08	Cut3 (Rv3451)	61
SPFH domain-containing protein/band 7 family protein [Mycobacterium smegmatis str. MC2 155]	YP_885889	19	54.572	4.91	GcvH (Rv1826)	35
eptc-inducible aldehyde dehydrogenase [Mycobacterium smegmatis str. MC2 155]	YP_885925	7	56.172	5.01	Rv0458	84
piperideine-6-carboxylic acid dehydrogenase [Mycobacterium smegmatis str. MC2 155]	YP_886134	3	55.123	5.05	Pcd (Rv3293)	80
TPR repeat-containing protein [Mycobacterium smegmatis str. MC2 155]	YP_889338	8	72.965	5.44	Rv1204c	47
phosphate/sulphate permease [Mycobacterium smegmatis str. MC2 155]	YP_885460	6	43.126	7.78	PitA (Rv0545c)	60
aldo/keto reductase [Mycobacterium smegmatis str. MC2 155]	YP_888917	6	39.102	4.98	Rv2298	31
sugar ABC transporter, substrate-binding protein, putative [Mycobacterium smegmatis str. MC2 155]	YP_884917	7	48.138	4.41	UspC (Rv2318)	25
low temperature requirement protein LtrA [Mycobacterium smegmatis str. MC2 155]	YP_889024	7	47.43	6.29	PstC2 (Rv0929)	33
short chain dehydrogenase [Mycobacterium smegmatis str. MC2 155]	YP_889138	13	29.511	5.15	Rv3224	51
sodium/proline symporter [Mycobacterium smegmatis str. MC2 155]	YP_889549	4	53.771	5.42	PutP (Rv1528c)	47

NAD-dependent epimerase/dehydratase [Mycobacterium smegmatis str. MC2 155]	YP_887774	15	31.063	5.14	GalE1 (Rv3634c)	63
short chain dehydrogenase [Mycobacterium smegmatis str. MC2 155]	YP_888467	9	31.223	8.52	Rv2129c	68
oligopeptide transport ATP-binding protein OppD [Mycobacterium smegmatis str. MC2 155]	YP_885050	10	36.264	5.05	OppD (Rv1281c)	50
fumarate reductase iron-sulfur subunit [Mycobacterium smegmatis str. MC2 155]	YP_884830	8	29.476	5.71	Rv0247c	84
DoxX subfamily protein, putative [Mycobacterium smegmatis str. MC2 155]	YP_887566	5	30.312	9.73	Fas (Rv2524c)	28
50S ribosomal protein L25/general stress protein Ctc [Mycobacterium smegmatis str. MC2 155]	YP_889670	8	22.537	4.56	RplY (Rv1015c)	75
carbon monoxide dehydrogenase subunit G (CoxG) family protein [Mycobacterium smegmatis str. MC2 155]	YP_885155	14	24.376	5.64	Rv0369c	64
ftsk/SpoIIIE family protein [Mycobacterium smegmatis str. MC2 155]	YP_885028	4	144.723	5.8	EccC3 (Rv0284)	75
ABC transporter ATP-binding protein [Mycobacterium smegmatis str. MC2 155]	YP_889894	5	67.987	6.11	Rv1272c	35
3-hydroxyacyl-CoA dehydrogenase [Mycobacterium smegmatis str. MC2 155]	YP_884631	9	27.066	5.2	Rv1144	49
replicative DNA helicase [Mycobacterium smegmatis str. MC2 155]	YP_891092	7	112.616	6.11	DnaB (Rv0058)	70
short-chain dehydrogenase/reductase SDR [Mycobacterium smegmatis str. MC2 155]	YP_886382	4	26.212	6.37	Rv1050	41
peptidase S8/S53 subtilisin kexin sedolisin [Mycobacterium smegmatis str. MC2 155]	YP_888481	3	66.051	6.42	MycP3 (Rv0291)	27
[Mn] superoxide dismutase [Mycobacterium smegmatis str. MC2 155]	YP_890640	10	22.921	5.45	SodA (Rv3846)	81
hypothetical protein MSMEG_6518 [Mycobacterium smegmatis str. MC2 155]	YP_890331	26	13.354	5.03	Hpt (Rv3624c)	73
large-conductance mechanosensitive channel [Mycobacterium smegmatis str. MC2 155]	YP_889729	18	16.616	4.88	FadE13 (Rv0975c)	85
hypothetical protein MSMEG_1981 [Mycobacterium smegmatis str. MC2 155]	YP_886345	16	18.165	11.36	Rv1558	38
aldehyde dehydrogenase [Mycobacterium smegmatis str. MC2 155]	YP_891076	3	53.747	5.25	Rv0223c	38
hypothetical protein MSMEG_0183	YP_884600	6	19.539	10.74	ThrC (Rv1295c)	48

[Mycobacterium smegmatis str. MC2 155] glycosyl transferase, family protein 2 [Mycobacterium smegmatis str. MC2 155]	YP_890189	6	77.216	8.57	Rv3631	34
gas vesicle protein [Mycobacterium smegmatis str. MC2 155]	YP_889979	12	26.266	5.46	LpqL (Rv0418)	47
transmembrane protein [Mycobacterium smegmatis str. MC2 155]	YP_888495	11	10.901	11.25	Rv2146c	87
stas domain-containing protein [Mycobacterium smegmatis str. MC2 155]	YP_889787	5	28.233	8.53	Rv0941c	48
hypothetical protein MSMEG_6143 [Mycobacterium smegmatis str. MC2 155]	YP_890364	5	63.763	8.99	Rv3635	72
dihydrolipoamide acetyltransferase [Mycobacterium smegmatis str. MC2 155]	YP_888560	1	61.049	4.65	DlaT (Rv2215)	74
hypothetical protein MSMEG_6739 [Mycobacterium smegmatis str. MC2 155]	YP_890947	15	16.471	9.98	sigI (Rv1189)	30
hypothetical protein MSMEG_1281 [Mycobacterium smegmatis str. MC2 155]	YP_885671	3	122.788	5.59	mce2D (Rv0592)	31
DNA-binding protein [Mycobacterium smegmatis str. MC2 155]	YP_884675	4	29.656	6.24	Rv0474	28

HEME CULTURE FILTRATE

Protein Name	ID	% Cov.	MW (kDa)	pI	Rv Homolog	I.D. (%)
precorrin-6Y C5,15-methyltransferase (decarboxylating) [Mycobacterium smegmatis str. MC2 155]	YP_888169	3	41.98	5.59	cobL (Rv2072c)	71
LamB/YcsF family protein [Mycobacterium smegmatis str. MC2 155]	YP_890277	2	27.446	5.07	ephB (Rv1938)	31
hypothetical protein MSMEG_0243 [Mycobacterium smegmatis str. MC2 155]	YP_884658	10	11.777	4.68	Rv0203	31
immunogenic protein MPB64/MPT64 [Mycobacterium smegmatis str. MC2 155]	YP_885448	35	24.837	4.53	TB22.2 (Rv3036c)	55
endoribonuclease L-PSP superfamily protein [Mycobacterium smegmatis str. MC2 155]	YP_886057	25	16.264	4.889	Rv2704	41

HEME CYTOSOLIC PROTEINS

Protein Name	ID	% Cov.	MW (kDa)	pI	Rv Homolog	I.D. (%)
universal stress protein family protein, putative	YP_888104.1	65	15.1	4.95	Rv1636	78
serine 3-dehydrogenase, MSMEG_2092	YP_885333.1	4	26.4	5.69	Rv0484c	82
hypothetical protein MSMEG_1513	YP_885895.1	7	30.1	5.65	Rv3463	67
3-ketoacyl-(acyl-carrier-protein) reductase, MSMEG_0372	YP_884785.1	11	46.5	5.58	FabG (Rv0242c)	84
inosine 5'-monophosphate dehydrogenase, MSMEG_1602	YP_885981.1	6	53.527	5.61	GuaB2 (Rv3411c)	90

IRON MEMBRANE

Protein Name	ID	% Cov.	MW (kDa)	pI	Rv Homolog	I.D. (%)
inner membrane permease YgbN [Mycobacterium smegmatis str. MC2 155]	YP_884865	14	48.809	5.71	lipL (Rv1497)	37
DNA-directed RNA polymerase subunit beta' [Mycobacterium smegmatis str. MC2 155]	YP_885754	6	147.106	6.05	RpoC (Rv0688)	91
hydrolase, alpha/beta fold family protein [Mycobacterium smegmatis str. MC2 155]	YP_886397	10	33.553	9.54	Hpx (Rv3171c)	59
glyoxylate carboligase [Mycobacterium smegmatis str. MC2 155]	YP_889714	12	65.136	5.68	IlvB1 (Rv3003c)	32
clavaldehyde dehydrogenase [Mycobacterium smegmatis str. MC2 155]	YP_889803	10	26.883	5.73	Rv0765	34
glutamate permease [Mycobacterium smegmatis str. MC2 155]	YP_887060	9	24.344	9.09	PstA2 (Rv0936)	34
succinate dehydrogenase flavoprotein subunit [Mycobacterium smegmatis str. MC2 155]	YP_884831	7	70.354	5.48	SdhA (Rv0248c)	80
amidohydrolase 2 [Mycobacterium smegmatis str. MC2 155]	YP_888329	6	34.32	4.8	Rv2303c	26
TROVE domain-containing protein [Mycobacterium smegmatis str. MC2 155]	YP_885585	6	61.373	9.29	Rv1760	41
SPFH domain-containing protein/band 7 family protein [Mycobacterium smegmatis str. MC2 155]	YP_885899	17	54.572	4.91	Rv1488	25
hydrolase [Mycobacterium smegmatis str. MC2 155]	YP_885210	22	53.037	4.44	LpqL (Rv0418)	68
penicillin binding protein transpeptidase domain-containing protein [Mycobacterium smegmatis str. MC2 155]	YP_884451	20	51.528	5.32	PbpA (Rv0016c)	80
AMP-binding enzyme [Mycobacterium smegmatis str. MC2 155]	YP_888530	10	64.732	5.36	Fad15 (Rv2187)	73
hydrogenase-2, large subunit [Mycobacterium smegmatis str. MC2 155]	YP_886615	12	60.15	5.01	Rv1722	24
L-sorbosone dehydrogenase [Mycobacterium smegmatis str. MC2 155]	YP_884629	12	44.694	4.44	Rv0195	29
acetyl-/propionyl-coenzyme A carboxylase alpha chain [Mycobacterium smegmatis str. MC2 155]	YP_886179	19	63.27	5.26	AccA3 (Rv3285)	86
propionyl-CoA carboxylase beta chain [Mycobacterium smegmatis str. MC2 155]	YP_886185	7	58.676	5	AccD5 (Rv3280)	87
extracellular solute-binding protein, family protein 5 [Mycobacterium smegmatis str. MC2 155]	YP_889249	8	61.266	4.92	OppA (Rv1280c)	69
hypothetical protein MSMEG_6346 [Mycobacterium smegmatis str. MC2 155]	YP_890563	15	15.957	10.63	Rv3519	35
fatty-acid--CoA ligase [Mycobacterium smegmatis str. MC2 155]	YP_887769	4	65.933	5.39	RVBD_1925	70
NAD(P) transhydrogenase subunit alpha [Mycobacterium smegmatis str. MC2 155]	YP_884526	11	52.731	5.48	PntAa (Rv0155)	44
pyruvate dehydrogenase [Mycobacterium smegmatis str. MC2 155]	YP_886632	8	62.346	5.84	IlvB1 (Rv3003c)	31
transmembrane protein [Mycobacterium smegmatis str. MC2 155]	YP_888756	3	71.631	5.18	Rv2345	58
ABC transporter binding protein [Mycobacterium smegmatis str. MC2 155]	YP_889019	3	43.951	4.41	PPE28 (Rv1800)	33
DNA-directed RNA polymerase subunit alpha [Mycobacterium smegmatis str. MC2 155]	YP_885906	19	37.954	4.62	RpoA (RvRv3457c)	91
sugar ABC transporter, substrate-binding protein, putative [Mycobacterium smegmatis str. MC2 155]	YP_884916	21	48.138	4.41	PkfB (Rv2029c)	38
gamma-glutamyltransferase [Mycobacterium smegmatis str. MC2 155]	YP_887392	6	66.656	5.13	GgtB (Rv2394)	71
ABC transporter, ATP-binding protein SugC [Mycobacterium smegmatis str. MC2 155]	YP_889308	19	43.687	5.02	SugC (Rv1238)	84
inner membrane permease YgbN [Mycobacterium smegmatis str. MC2 155]	YP_884865	17	48.809	5.71	lipL (Rv1497)	37
periplasmic binding protein [Mycobacterium smegmatis str. MC2 155]	YP_885437	10	37.052	4.39	FecB (Rv3044)	62

hypothetical protein MSMEG_0923 [Mycobacterium smegmatis str. MC2 155]	YP_885426	9	32.687	4.93	Rv0044c	32
glutamine-binding periplasmic protein/glutamine transport system permease protein [Mycobacterium smegmatis str. MC2 155]	YP_890525	8	50.957	4.91	GlnH (Rv0411c)	28
hypothetical protein MSMEG_1980 [Mycobacterium smegmatis str. MC2 155]	YP_886344	22	42.996	5.43	Rv0575c	57
3-ketoacyl-(acyl-carrier-protein) reductase [Mycobacterium smegmatis str. MC2 155]	YP_884785	10	46.584	5.58	FabG (Rv0242c)	84
LppZ protein [Mycobacterium smegmatis str. MC2 155]	YP_886711	16	38.031	4.78	LppZ (Rv3006)	80
universal stress protein family protein [Mycobacterium smegmatis str. MC2 155]	YP_889712	7	31.353	6.08	Rv2319c	31
hypothetical protein MSMEG_5268 [Mycobacterium smegmatis str. MC2 155]	YP_889514	9	37.826	4.61	TB39.8 (Rv0020c)	47
alcohol dehydrogenase, iron-containing [Mycobacterium smegmatis str. MC2 155]	YP_890461	7	46.49	5.36	Rv1619	21
sugar ABC transporter substrate-binding protein [Mycobacterium smegmatis str. MC2 155]	YP_888921	17	34.272	4.81	Pks9 (Rv1664)	59
D-ribose-binding periplasmic protein [Mycobacterium smegmatis str. MC2 155]	YP_887411	20	32.148	4.39	EchA10 (1142c)	42
cytochrome c oxidase subunit 2 [Mycobacterium smegmatis str. MC2 155]	YP_888545	7	38.186	5.31	CtaC (Rv2200c)	76
metallo-beta-lactamase superfamily protein [Mycobacterium smegmatis str. MC2 155]	YP_890293	9	33.888	6.24	Rv3577	68
short chain dehydrogenase [Mycobacterium smegmatis str. MC2 155]	YP_885266	5	33.189	8.67	Rv0439c	70
hypothetical protein MSMEG_3107 [Mycobacterium smegmatis str. MC2 155]	YP_887422	5	33.922	4.77	Rv1455	50
oligopeptide transport ATP-binding protein AppF [Mycobacterium smegmatis str. MC2 155]	YP_885049	5	37.3	6.5	OppD (Rv1281c)	54
3-isopropylmalate dehydrogenase [Mycobacterium smegmatis str. MC2 155]	YP_887596	9	37.849	5.47	LeuB (Rv2995c)	36
hypothetical protein MSMEG_0365 [Mycobacterium smegmatis str. MC2 155]	YP_884778	6	66.236	4.86	Rv2079	22
sugar ABC transporter ATP-binding protein [Mycobacterium smegmatis str. MC2 155]	YP_888919	16	29.263	6.09	CysA1 (Rv2397c)	33
dimethyladenosine transferase [Mycobacterium smegmatis str. MC2 155]	YP_889677	4	32.813	7.12	KsgA (Rv1010)	76
secreted protein [Mycobacterium smegmatis str. MC2 155]	YP_888487	10	34.601	4.61	Rv3224B	35
beta-glucosidase A [Mycobacterium smegmatis str. MC2 155]	YP_889388	3	52.46	5	TrcS (Rv1032c)	30
hypothetical protein MSMEG_1332 [Mycobacterium smegmatis str. MC2 155]	YP_885722	15	32.942	4.79	Rv0633c	51
integral membrane protein [Mycobacterium smegmatis str. MC2 155]	YP_884664	9	39.457	5.46	Rv0205	70
hypothetical protein MSMEG_2396 [Mycobacterium smegmatis str. MC2 155]	YP_886736	22	20.463	4.6	Rv2980	53
heat shock protein HtpX [Mycobacterium smegmatis str. MC2 155]	YP_885528	5	31.98	9.43	HptX (Rv0563)	86
LprG protein [Mycobacterium smegmatis str. MC2 155]	YP_887386	21	24.15	4.46	LprG (Rv1411c)	53
hypothetical protein MSMEG_5238 [Mycobacterium smegmatis str. MC2 155]	YP_889484	9	21.718	4.73	Rv1100	60
secreted protein [Mycobacterium smegmatis str. MC2 155]	YP_886470	15	22.04	6.06	ValS (Rv2448c)	39
metallo-beta-lactamase superfamily protein [Mycobacterium smegmatis str. MC2 155]	YP_885181	16	26.802	6.08	Rv0406c	57
cytosine/purines/uracil/thiamine/allantoin permease family protein [Mycobacterium smegmatis str. MC2 155]	YP_885569	10	50.393	7.77	Rv3454	25
macrocin-O-methyltransferase [Mycobacterium smegmatis str. MC2 155]	YP_884801	12	27.258	5.3	Rv3191c	30
translation initiation factor IF-1 [Mycobacterium smegmatis str. MC2 155]	YP_885901	17	8.498	9.45	InfA (Rv3462c(	99
MmpL11 protein [Mycobacterium	YP_884656	3	102.969	5.98	Rv0202c	69

smegmatis str. MC2 155]						
secreted protein [Mycobacterium smegmatis str. MC2 155]	YP_886470	24	22.04	6.06	ValS (Rv2448c)	39
hypothetical protein MSMEG_2261 [Mycobacterium smegmatis str. MC2 155]	YP_886613	10	19.374	5.13	Rv3074	34
Rieske iron-sulfur protein [Mycobacterium smegmatis str. MC2 155]	YP_890760	8	13.938	4.98	QcrA (Rv2195)	34
hypothetical protein MSMEG_0823 [Mycobacterium smegmatis str. MC2 155]	YP_885227	14	19.705	5.18	Rv2091c	27
transmembrane protein [Mycobacterium smegmatis str. MC2 155]	YP_885362	6	17.709	9.15	Rv0513	65
hypothetical protein MSMEG_6228 [Mycobacterium smegmatis str. MC2 155]	YP_890447	3	23.913	5.92	Rv2896c	50
YceI like family protein [Mycobacterium smegmatis str. MC2 155]	YP_886426	5	17.974	6.84	Rv1890c	27
membrane flanked domain-containing protein [Mycobacterium smegmatis str. MC2 155]	YP_886188	5	19.474	5.94	Rv3278c	75
large-conductance mechanosensitive channel [Mycobacterium smegmatis str. MC2 155]	YP_889720	28	16.616	4.88	Rv0985c	54
hypothetical protein MSMEG_6518 [Mycobacterium smegmatis str. MC2 155]	YP_890731	26	13.354	5.03	Rv3863	52
hypothetical protein MSMEG_0366 [Mycobacterium smegmatis str. MC2 155]	YP_884779	40	18.747	4.44	PlsC (Rv2483c)	31
hypothetical protein MSMEG_1781 [Mycobacterium smegmatis str. MC2 155]	YP_886153	23	16.6	5.54	FadD13 (Rv3089)	28
sec-independent translocase [Mycobacterium smegmatis str. MC2 155]	YP_889319	15	14.525	4.85	TatB (Rv1224)	69
FHA domain-containing protein [Mycobacterium smegmatis str. MC2 155]	YP_884454	7	17.115	11.19	Rv0019c	88
molybdopterin biosynthesis protein [Mycobacterium smegmatis str. MC2 155]	YP_889723	20	17.933	4.28	MoaB2 (Rv0984)	82
MmpS5 protein [Mycobacterium smegmatis str. MC2 155]	YP_887799	20	17.298	5.09	MmpS5 (Rv0677)	65
MmpS4 protein [Mycobacterium smegmatis str. MC2 155]	YP_884793	26	15.285	5.63	mmpS4 (Rv0451c)	60
glycerol kinase [Mycobacterium smegmatis str. MC2 155]	YP_890448	11	55.423	4.81	glpK (Rv3696c)	83
hypothetical protein MSMEG_2698 [Mycobacterium smegmatis str. MC2 155]	YP_887073	8	15.462	7.9	lexA (Rv2720)	89
enoyl-CoA hydratase [Mycobacterium smegmatis str. MC2 155]	YP_890141	4	28.725	5.42	echA19 (Rv3516)	84
hypothetical protein MSMEG_6722 [Mycobacterium smegmatis str. MC2 155]	YP_890930	2	33.863	9.57	Rv2237	46
hypothetical protein MSMEG_2668 [Mycobacterium smegmatis str. MC2 155]	YP_887004	6	17.057	7.93	RVBD_1762Ac	42
preprotein translocase subunit SecA [Mycobacterium smegmatis str. MC2 155]	YP_886246	1	106.479	5.34	secA1 (Rv3240c)	82
KanY protein [Mycobacterium smegmatis str. MC2 155]	YP_890501	7	15.258	9.26	Rv3718c	79
30S ribosomal protein S10 [Mycobacterium smegmatis str. MC2 155]	YP_885818	14	11.483	9.41	rpsJ Rv0700	98
acyl carrier protein [Mycobacterium smegmatis str. MC2 155]	YP_888601	44	10.73	3.91	acpP Rv2244	93
MmpL4 protein [Mycobacterium smegmatis str. MC2 155]	YP_884640	8	105.686	5.31	mmpL5 Rv0676c	71
inner membrane protein [Mycobacterium smegmatis str. MC2 155]	YP_884590	9	12.994	11.08	Rv2272	53
epoxide hydrolase [Mycobacterium smegmatis str. MC2 155]	YP_890916	7	33.17	5.27	ephF (Rv0134)	30
hypothetical protein MSMEG_2801 [Mycobacterium smegmatis str. MC2 155]	YP_887132	56	4.06	11.54	rplE (Rv0716)	36
hemerythrin HHE cation binding region [Mycobacterium smegmatis str. MC2 155]	YP_886755	4	21.248	4.82	lhr (Rv3296)	31
acyl-CoA-dehydrogenase [Mycobacterium smegmatis str. MC2 155]	YP_884819	9	66.547	5	fadE5 (Rv0244c)	81

IRON CULTURE FILTRATE

Protein Name	ID	% Cov.	MW (kDa)	pI	Rv Homolog	I.D. (%)

integration host factor [Mycobacterium smegmatis str. MC2 155]	YP_887366	22	11.628	9.91	mihF (Rv1388)	96
hypothetical protein MSMEG_3417 [Mycobacterium smegmatis str. MC2 155]	YP_887721	4	18.297	5.06	Rv0088	29
hypothetical protein MSMEG_6251 [Mycobacterium smegmatis str. MC2 155]	YP_890470	11	22.795	5.27	Rv3705c	72
molybdopterin biosynthesis protein [Mycobacterium smegmatis str. MC2 155]	YP_889723	13	17.933	4.28	moaB2 (Rv0984)	82

IRON CYTOSOLIC PROTEINS

Protein Name	ID	% Cov.	MW (kDa)	pI	Rv Homolog	I.D. (%)
carnitinyl-CoA dehydratase	YP_889444.1	10	28	5.39	EchA1 (Rv0222)	47
peptidyl-prolyl cis-trans isomerase B	YP_884444.1	33	18.7	5.42	PpiA (Rv0009)	90

REFERENCES

1. Tullius, M.V., et al., *Discovery and characterization of a unique mycobacterial heme acquisition system.* Proc Natl Acad Sci U S A, 2011. **108**(12): p. 5051-6.
2. Owens, C.P., et al., *Characterization of heme ligation properties of Rv0203, a secreted heme binding protein involved in Mycobacterium tuberculosis heme uptake.* Biochemistry, 2012. **51**(7): p. 1518-31.
3. Ratledge, C., *Iron, mycobacteria and tuberculosis.* Tuberculosis (Edinb), 2004. **84**(1-2): p. 110-30.
4. Owens, C.P., et al., *The Mycobacterium tuberculosis secreted protein Rv0203 transfers heme to membrane proteins MmpL3 and MmpL11.* J Biol Chem, 2013. **288**(30): p. 21714-28.
5. Chim, N., et al., *Unusual diheme conformation of the heme-degrading protein from Mycobacterium tuberculosis.* Journal of Molecular Biology, 2010. **395**(3): p. 595-608.
6. Nambu, S., et al., *A new way to degrade heme: the Mycobacterium tuberculosis enzyme MhuD catalyzes heme degradation without generating CO.* J Biol Chem, 2013. **288**(14): p. 10101-9.
7. Jones, C.M. and M. Niederweis, *Mycobacterium tuberculosis Can Utilize Heme as an Iron Source.* Journal of Bacteriology, 2011. **193**(7): p. 1767-1770.
8. Lowry, O.H., et al., *Protein measurement with the Folin phenol reagent.* J Biol Chem, 1951. **193**(1): p. 265-75.
9. Turlin, E., et al., *Staphylococcus aureus FepA and FepB proteins drive heme iron utilization in Escherichia coli.* PLoS One, 2013. **8**(2): p. e56529.
10. Dawson, R.M.C., *Data for biochemical research*. 3rd ed. 1986, Oxford: Clarendon Press. xii, 580 p.
11. Hargrove, M.S., et al., *His64(E7)-->Tyr apomyoglobin as a reagent for measuring rates of hemin dissociation.* J Biol Chem, 1994. **269**(6): p. 4207-14.
12. Hargrove, M.S., D. Barrick, and J.S. Olson, *The association rate constant for heme binding to globin is independent of protein structure.* Biochemistry, 1996. **35**(35): p. 11293-9.
13. Besra, G.S., *Preparation of cell-wall fractions from mycobacteria.* Methods Mol Biol, 1998. **101**: p. 91-107.

14. Mattow, J., et al., *Two-dimensional gel electrophoresis-based proteomics of mycobacteria.* Methods Mol Biol, 2009. **465**: p. 111-42.
15. Carpentier, S.C., et al., *Preparation of protein extracts from recalcitrant plant tissues: an evaluation of different methods for two-dimensional gel electrophoresis analysis.* Proteomics, 2005. **5**(10): p. 2497-507.
16. Whitelegge, J.P., et al., *Subtle modification of isotope ratio proteomics; an integrated strategy for expression proteomics.* Phytochemistry, 2004. **65**(11): p. 1507-15.
17. Rodriguez, G.M., *Control of iron metabolism in Mycobacterium tuberculosis.* Trends Microbiol, 2006. **14**(7): p. 320-7.
18. Jones, C.M. and M. Niederweis, *Mycobacterium tuberculosis can utilize heme as an iron source.* J Bacteriol, 2011. **193**(7): p. 1767-70.
19. Friedman, D.B., et al., *Staphylococcus aureus redirects central metabolism to increase iron availability.* PLoS Pathog, 2006. **2**(8): p. e87.
20. Gao, J.L., K.A. Nguyen, and N. Hunter, *Characterization of a hemophore-like protein from Porphyromonas gingivalis.* J Biol Chem, 2010. **285**(51): p. 40028-38.