Purification and Characterization of Two Heme Induced Proteins of *Mycobacterium tuberculosis*

Genesis Terrazas Valero

Departments of Microbiology
Faculty Mentor: Dr. Avishek Mitra
Thesis Second Reader: Dr. Erika Lutter

ABSTRACT

Mycobacterium tuberculosis (Mtb), the causative agent of human tuberculosis (TB), is transmitted through aerosols and successfully replicates within alveolar macrophages by inhibiting macrophage maturation. Before the SARS-CoV-2 pandemic, Mtb had surpassed HIV/AIDS to become the leading cause of death worldwide from a single infectious agent. In 2019, 11 million people were infected with Mtb resulting in 1.3 million deaths. The WHO estimates that globally 3.3% of new cases and 18% of previously treated cases were multidrug-resistant TB (MDR-TB) in 2019. Thus, there is an imminent need to develop chemotherapeutic strategies to prevent TB disease progression. Mtb is entirely dependent on iron acquisition to colonize the human host successfully. Greater than 75% of host iron is stored in heme, making it the primary source of iron for Mtb in the human host. An overarching goal of our lab is to identify novel components of Mtb that are required for Hm iron acquisition so that we can develop novel strategies to block heme iron acquisition and starve this pathogen of an essential nutrient. Previous transcriptomic studies had identified that Mtb significantly increases expression of *ppe64* and *rv0125* in response to heme iron in the growth medium. The goal of this project was to clone these genes into an expression vector and attempt to purify the encoded recombinant proteins from *E. coli*.

1. INTRODUCTION

1.1. Global Impact of *Mycobacterium tuberculosis*

The etiological agent of human tuberculosis (TB) is the acid-fast, slow-growing, bacterial pathogen *Mycobacterium tuberculosis* (Mtb). Transmission of tuberculosis occurs through the inhalation of aerosols expelled by contagious individuals. Upon inhalation, Mtb enters the alveolar space inside the lungs, where alveolar macrophages engulf the bacteria by phagocytosis. After that, the pathogen inhibits macrophage maturation leading to the formation of granulomas (Figure 1.1) [1]. As a result, human tuberculosis has been identified as one of the top two causes of death worldwide. Before the SARS-CoV-2 pandemic, Mtb had surpassed HIV/ AIDS, making it the number one cause of death resulting from a single infectious agent. In 2019, an estimated 10 million people became infected with MTB, which resulted in the death of 1.3 million people. Furthermore, it is believed that approximately 1.7 billion people, about 25% of the world's population, are latently infected with *Mycobacterium tuberculosis* [2]. Although treatments for TB are available such as the two-month RHZE regimen consisting of treatment with rifampicin (R), isoniazid (H), pyrazinamide (Z), and ethambutol (E), the rise in drug-resistant strains has made the development of novel strategies for treatment and the prevention of Mtb infections an imminent need in recent years [3]. The World Health Organization reported that in 2019, the incidence of rifampicin-resistant TB (RR-TB) and multidrug-resistant TB (MDR-TB) was 3.3% in new and previously treated cases, 17.7% [4].

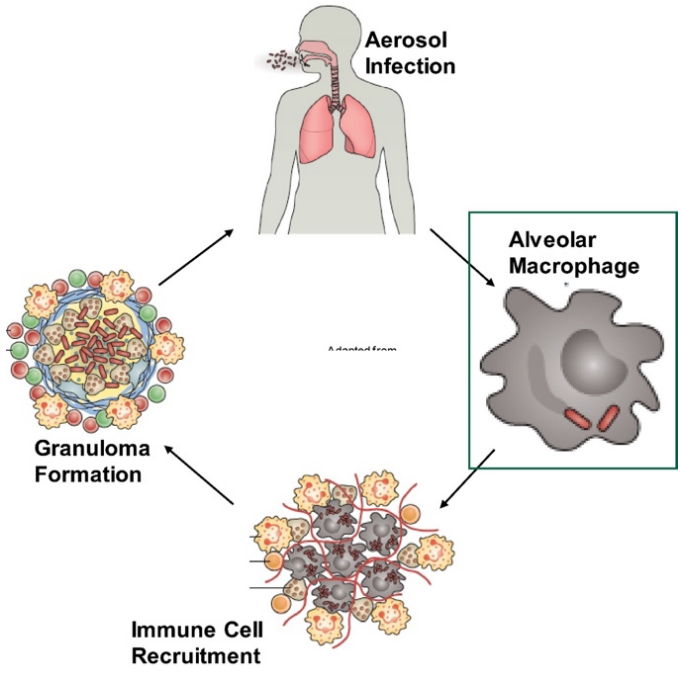

Figure 1.1 ***Mycobacterium tuberculosis*** **Infection Cycle.** Pictured here is the lifecycle of *Mycobacterium tuberculosis* which initiates with the inhalation of infected aerosols into the alveoli where alveolar macrophages will phagocytose the pathogen. Mtb proliferates within the phagosome leading to further recruitment of immune cells at the site of infection and formation of granulomas. Mtb then escapes from the granuloma and goes on to repeat this cycle. This figure was adapted from Nunes-Alves *et al.* (2014) [5].

1.2 Importance of iron for Mtb pathogenesis

1.2.1. Importance of iron & storage of iron in the human host

Iron is a crucial transition metal that is required for numerous essential biological processes for survival in most living organisms. Iron acts as a cofactor essential for the catalysis of various biochemical reactions by enzymes; hence its presence is crucial in many physiological mechanisms such as central carbon metabolism, DNA replication, and transcription. Furthermore, iron can be found in proteins involved in hydrogen fixation, nitrogen fixation, and dioxygen fixation through respiration. In addition, iron plays a role in generating energy by incorporating protoporphyrin containing ferrous iron (Fe^{2+}) into cytochromes [6, 7]. On average human beings consume ~4-5 g daily through diet. Since iron is abundant within the human host and is essential to pathogens, the host immune system has evolved various mechanisms by which they prevent pathogens from acquiring iron; this is known as nutritional immunity response. Approximately 25% of the iron in the human host is stored within tissue or liver as ferric iron within transferrin (Tf), lactoferrin (Lf) and ferritin (F) (Figure 1.2). The majority (>75%) of the iron in the human host is stored in the form of heme within red blood cells [8]. Heme is a small molecule consisting of a porphyrin ring with iron bound in the middle of the ring (Figure 1.2). Since free heme itself can be toxic, it is further complexed with proteins and stored in hemoproteins such as myoglobin and hemoglobin (Figure 1.2) [9]. Since, iron is very efficiently hidden in the human host, it is not surprising that bacterial pathogens like Mtb have evolved various mechanisms to acquire iron from the human host. Iron acquisition in bacteria can be categorized into two generalized mechanisms: **1)** iron acquisition from Lf, Tf or F by bacterial siderophores, or **2)** iron acquisition from heme [10].

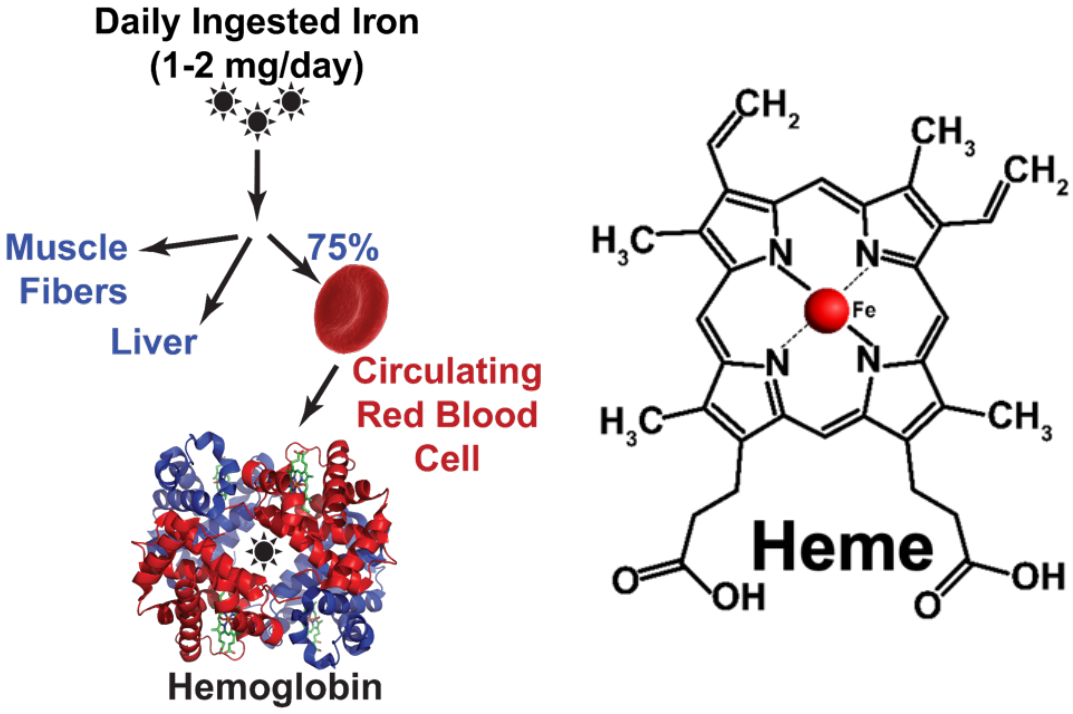

Figure 1.2. Storage of iron within the human host. Approximately 25% of iron is stored within lactoferrin, transferrin and ferritin within tissues. Greater than 75% of iron is stored in the form of heme.

1.2.2. Siderophore-dependent iron acquisition in bacteria and Mycobacterium tuberculosis

To acquire iron from Tf, Lf or F gram-negative and gram-positive bacteria secrete high affinity iron chelating molecules called siderophores [11]. Since, siderophores have high binding affinity for ferric iron, they essentially steal the ferric iron from Tf, Lf or F [12]. In gram-negative bacteria, iron-loaded siderophores are first bound by an outer channel membrane protein (OMP) located in the outer membrane (Figure 1.3). The transport of the Fe^{3+}-siderophore complex is energy-dependent because the complex is large and does not diffuse through the channel. At the inner membrane, a protein complex called the TonB-ExbBD system harnesses the energy generated by the proton motive force and transfers the energy to the OMP, which then transports the Fe^{3+}-siderophore complex into the periplasmic space [13]. In the periplasm substrate-binding proteins (SBPs) bind the complex and transfers it to an ATP-binding cassette (ABC) transporter located in the inner membrane, which then transports the Fe^{3+}-siderophore complex into the cytoplasm. Since, gram positive bacteria do not have an outer membrane, the Fe^{3+}-siderophore complex is first bound by a cell surface receptor(s) (Figure 1.3). After this the transport process through the periplasm is exactly the same as in gram-negative bacteria [14]. In the cytoplasm the iron is typically released from the Fe^{3+}-siderophore complex by reducing Fe^{3+} to Fe^{2+}, which releases the iron because siderophores do not bind Fe^{2+}. Sometimes, the iron can also be released by degrading the siderophore with enzymes [15].

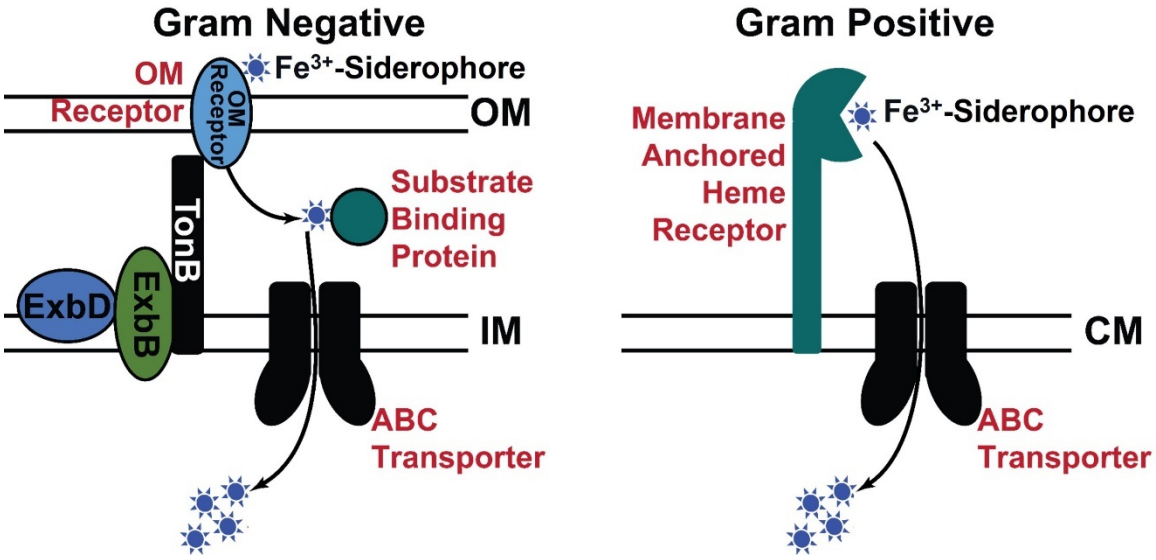

Figure 1.3. Models of siderophore-dependent iron acquisition in gram-negative and gram-negative bacteria. See section 1.2.2. for detailed description.

As in most bacterial pathogens, *Mycobacterium tuberculosis* acquires ferric iron from Lf, Tf or F by producing siderophores mycobactin (MBT) and carboxymycobactin (cMBT). Mycobactin is hydrophobic and remains bound to the cell surface of Mtb, whereas carboxymycobactin is hydrophilic and secreted into the extracellular space. Both siderophores are produced by Mtb under conditions when iron availability is low [16]. A model for siderophore dependent iron acquisition is shown in Figure 1.4. MBT and cMBT are synthesized in the cytoplasm and exported across the inner membrane by the MmpS4/S5 and MmpL4/L5 efflux pumps. It is not known how MBT and cMBT are exported across the Mtb outer membrane. It is

hypothesized that cMBT first binds ferric iron and can transfer the iron to the cell surface MBT. How Fe^{3+}-MBT or Fe^{3+}-cMBT is transported across the outer membrane is currently not known. After these complexes are transported into the cytoplasm, the inner membrane ABC transporters IrtA/IrtB transport them into the cytoplasm. In the cytoplasm, the Fe^{3+} is reduced to Fe^{2+} to release the iron and then siderophores are again recycled and secreted outside to bring in more iron [17, 18] [19].

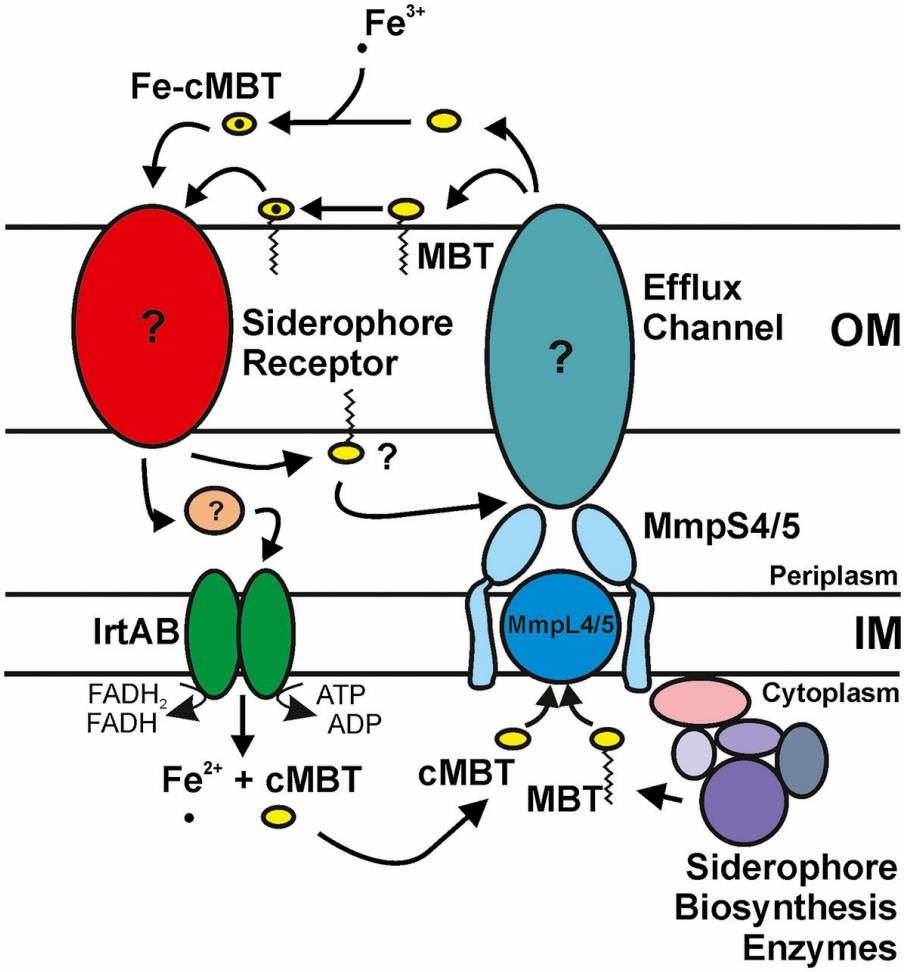

Figure 1.4. Model of siderophore-dependent iron acquisition in *Mycobacterium tuberculosis*. See section 1.2.2. for detailed description.

1.2.3. Heme-dependent iron acquisition in bacteria and Mycobacterium tuberculosis

The mechanism of heme acquisition is almost similar to that of siderophore-dependent iron acquisition. Briefly, heme is first bound by OMPs or cell surface receptors in gram-negative and gram-positive bacteria, respectively. In gram-negative bacteria heme transport across the OM requires energy, provided by TonB-ExbBD. Heme transport from the periplasm into the cytoplasm also requires a substrate binding protein and an ABC transporter. In the cytoplasm, iron is released from heme by degrading the porphyrin ring or heme is stored within proteins.

Heme acquisition in most bacterial pathogens is very well studied since the 80s, but the mechanisms of heme acquisition in Mtb is poorly understood. The first study showing that Mtb can utilize heme as an iron source was shown by Jones *et al.* in 2011 [20]. In more recent studies

Mitra *et al.* has shown that Mtb uses novel outer membrane proteins of the PPE (proline-proline-glutamate) family, PPE36 and PPE62, to bind heme and transport it into the periplasm[21]. These proteins were also shown to be required for using hemoglobin as an iron source. These studies also identified that the inner membrane dipeptide ABC transporter called the Dpp transporter works with PPE36 and PPE62 to transport heme into the cytoplasm (Figure 1.5) [22]. The authors showed that deletion of either *ppe36* or *ppe62* or the *dpp* transporter significantly reduced growth of Mtb in growth medium where heme was the only iron source. But the authors also found that if serum albumin is added into the heme growth medium, then the *ppe36*, *ppe62* and *dpp* mutants would grow without any issues. This led to the discovery that Mtb has a completely separate heme acquisition pathway independent from PPE36-PPE62-Dpp and this second pathway was called the albumin-heme pathway.

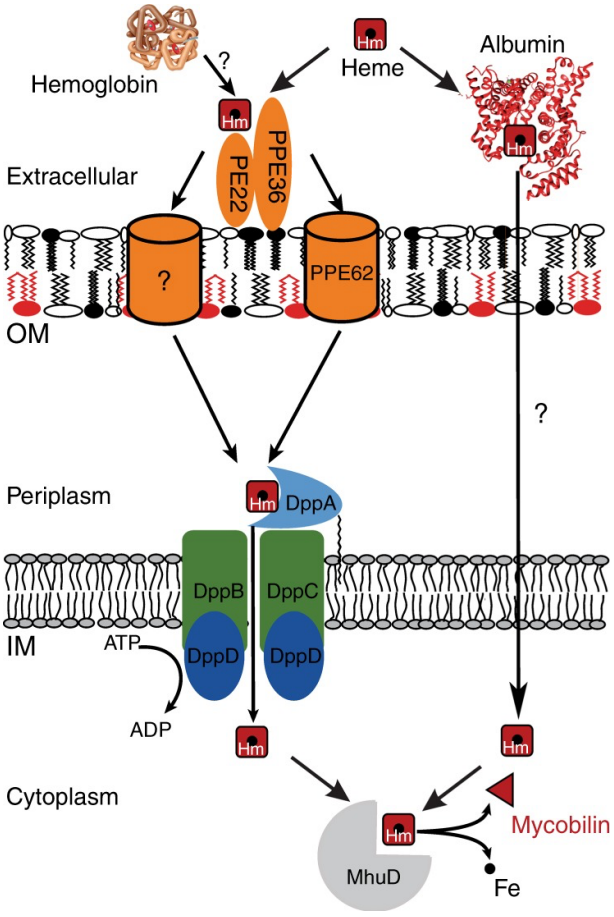

Figure 1.5. Model of heme iron acquisition in *Mycobacterium tuberculosis*. This Model of Mtb heme acquisition shows the known components involved in utilizing host heme iron. In one pathway, heme is transported by PPE36 and PPE62 across the outer membrane. In the periplasm, DppA binds heme and heme is transported by the Dpp ABC transporter into the cytoplasm. It is known from previous studies that in the cytoplasm heme is degraded by MhuD to release iron[23]. In a completely separate pathway, serum albumin helps in heme uptake into the Mtb cell. The components of this pathway are not known.

1.2.4. Goals of this study

The discovery of the albumin-heme pathway in Mtb led us to try to identify components of this pathway. Mitra *et al.* had performed transcriptomic analysis of Mtb to identify novel components of this albumin-heme pathway. Mtb was grown in growth medium containing heme and heme-albumin and the transcriptomic profile was analyzed by RNA sequencing (unpublished data). From this data, it was observed that expression of two genes *ppe64* and *rv0125* were increased in Mtb by approximately 8-fold in the presence of heme-albumin. The function of the genes *ppe64* and *rv0125* remain unknown, but the protein PPE64, which belongs to the PPE protein family is a predicted outer membrane protein while Rv0125 is a predicted protease which hydrolyzes proteins and peptides. As a result of the increased expression of the genes *ppe64* and *rv0125* the goal of this study was to clone these Mtb genes in an expression vector to purify recombinant proteins, so that we can characterize these proteins through biochemical methods.

1.2.5. Use of an inducible plasmid expression system

A *lac* inducible plasmid system was used in this study to express Mtb genes and purify recombinant proteins. The *lac* inducible system is widely used to control expression of genes and the knowledge is derived through studies of the *lac* operon of *E. coli*. In *E. coli*, the *lac* operon is composed of three genes *lacZ*, *lacY* and *lacA*, which are used to transport lactose into the cell [24]. In the absence of lactose, expression of these genes is blocked by LacI by binding to the promoter region (Figure 1.6). When *E. coli* cells are exposed to lactose, small amounts of lactose are transported into the cell which is converted to allolactose. Allolactose then binds to LacI and prevents it from binding to the promoter, which activates expression of the *lac* operon genes and results in increased transport of lactose into the cell. This basic biology knowledge has been used to design expression plasmids where a T7 promoter is put under inducible control of the *lac* system. In this system, an analog of allolactose called IPTG is used to activate expression of any target gene downstream of the T7 promoter. Both *ppe64* and *rv0125* were cloned downstream of the lac-controlled T7 promoter and gene expression was induced using IPTG and then recombinant proteins were purified using nickel chromatography.

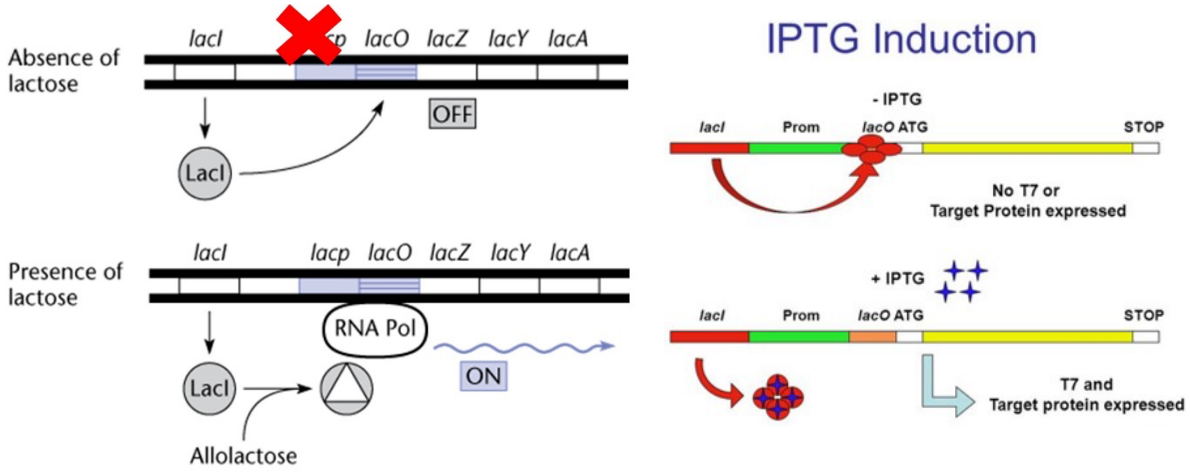

Figure 1.6. Mechanism of *lac* operon and inducible protein expression system. See section 1.2.5. for detailed description.

2. METHODS

2.1 Polymerase chain reaction

Polymerase chain reaction (PCR) was used to amplify the open reading frames (ORF) of the *ppe64* and *rv0125* gene DNA sequences. Forward and reverse primers used to amplify genes were designed using VectorNTI software. For both genes, the forward and reverse primers contained NdeI and XhoI restriction sites, respectively (Table 1). The reverse primer replaced the stop codon of the respective genes to allow addition of the C-terminal 6His-tag from pET21a+ plasmid (Figure 2). Lyophilized primers obtained from IDT were solubilized with millipore water (mH_2O) at 100 µM concentration for long-term storage. Working primer stocks were prepared by diluting freezer stocks in mH_2O to 10 µM. PCR mix setup for *ppe64* and *rv0125* was performed in a 50 µl volume containing mH_2O, 25 µl of 2X KOD polymerase premix, 1.25 µl of forward primer (gene-F/ NdeI), 1.25 µl of reverse primer (gene-R/XhoI), 2 µl of DMSO and 40 ng of Mtb H37Rv DNA which was obtained from ATCC®. PCR reactions consist of denaturing, annealing, and extension cycled through and repeated 40 times. The denaturing step consists of heating the reaction at a high temperature (94°C) to separate the DNA strand so that the primers can bind to the DNA stands in the following steps. After denaturing, the reaction is cooled during the annealing step to allow the binding of the primers to the single-stranded template DNA. The temperature for the annealing step is calculated based on the melting temperature of the forward and reverse primer pairs (57ºC for *ppe64*, 63ºC for *rv0125*). Lastly, during the extension step, the reaction temperature is raised to 72°C allowing KOD polymerase to extend the primers leading to the synthesis of new DNA strands. The resulting PCR samples were mixed with 3 µl of loading buffer and then separated by electrophoresis in a 1% agarose gel containing ethidium bromide. Electrophoresis was performed at 140 V for 50 minutes, and PCR products were then visualized using a BioRad ChemiDoc MP imaging system.

Table 1. Primers used in this study.

ppe64-21a-F/NdeI	atatCATATGGCTCATTTTTCGGTGTT
ppe64-21a-R/XhoI	atatCTCGAGCCCCAACAGCTGGCGCAGGT
rv0125-21a-F/NdeI	atatCATATGgccccgccggccttgtcgca
rv0125-21a-R/XhoI	atatCTCGAGggccgggggtccctcggcca

2.2 Restriction digestion of plasmid and PCR products for cloning

Our goal was to clone the gene PCR products into the pET21a+ vector (Figure 2). To perform the cloning process, pET21a+ was first digested with NdeI and HindIII restriction enzymes, which were purchased from NEB. Plasmid digestion mix was prepared in a 50 µl volume containing mH_2O, 5.0 µl Cutsmart buffer, 3 µg of pET21a+ DNA, and 0.3 µL of each restriction enzyme, and the mix was incubated at 37ºC for 2 h. Digested pET21a+ was then purified using the Qiagen PCR purification protocol. The 50 µl digestion reaction was first mixed with 250 µl PB buffer, transferred to a silica column, and then bound to the column by centrifugation at 17,900 g for 30 s. Then the column was washed with 750 µl PE buffer by centrifugation at 17,900 g for 30 s and centrifuged again at full speed for 1.5 min to remove any residual PE buffer. Finally, 40 µl of mH_2O was added to the column, incubated at room temperature for 2 min, and the DNA was eluted

into a fresh 1.5 ml tube by centrifugation at full speed for 2 min. The concentration and purity of the DNA sample were determined using the Take3 plate in a BioTek plate reader. Additionally, the 5'end of the digested pET21a+ was dephosphorylated to increase the efficiency of PCR product ligation. The pET21a+ dephosphorylation reaction was set up in a 20 µl volume containing mH_2O, 2 µl Antarctic Phosphatase buffer, 1 µl Antarctic phosphatase enzyme, and 1 µg of digested pET21a+. The dephosphorylation reaction was incubated at 37°C for 30 minutes and then immediately heat-inactivated at 80°C for 2 min. Dephosphorylated pET21a+ was then purified using the PCR purification protocol as described above. Gene PCR products of *ppe64* and *rv0125* were also similarly digested with NdeI and HindIII and then purified as described above. Digested PCR products were not dephosphorylated.

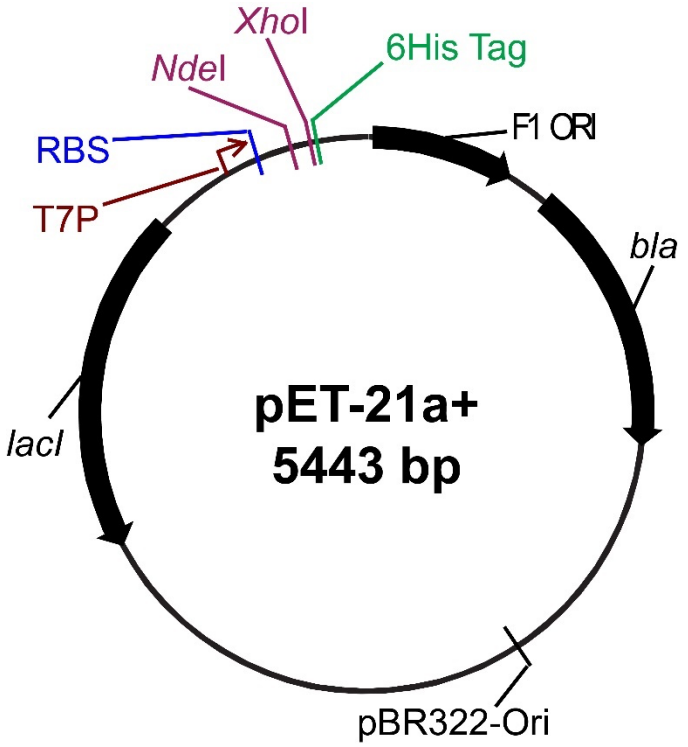

Figure 2. pET21a+ Expression Vector. Both *ppe64* and *rv0125* were cloned into pET21a+ at the NdeI and XhoI restriction sites. The native stop codon of both genes was replaced to allow read through and usage of the 6-His and stop codon sequence from pET21a+.

2.3 Ligation, transformation, and validation of cloning

Digested and dephosphorylated pET21a+ was ligated with similarly digested *ppe64* and *rv0125* using T4 DNA ligase. Ligation reactions were prepared in a 20 µL volume containing mH_2O, 2 µL of T4 DNA ligase buffer, 50 ng of pET21a+, 150 ng of PCR product, 1 µl of T4 DNA ligase and the mix was incubated overnight at 16ºC. The ligation reaction was then transformed into competent *E. coli* DH5α cells. The transformation mix was prepared in a 1.5 ml tube containing 20 µl of the ligation mix and 80 µl of DH5α. In the following sequence, the mix was incubated on ice for 30 min, heat shocked at 42ºC for 90 sec, incubated on ice for 5 min and then resuspended with 1 ml of LB. The resuscitated cells were incubated at 37ºC with shaking at 200 rpm for 1 h. The transformation mix was then plated on LB agar plates containing 100 µg/ml ampicillin and

incubated overnight at 37ºC. Putative transformed colonies were then grown in LB-Amp and plasmid was isolated using the Qiagen Miniprep kit. Putative cloned plasmids were then validated by restriction digestion with NdeI and XhoI.

2.4 Testing expression of recombinant Mtb proteins

Validated cloned plasmids were transformed into competent *E. coli* BL21 cells and transformants were selected on LB-Amp agar plates. A 5 ml overnight culture in LB-Amp was started from a single colony, which was then inoculated into 200 ml LB-Amp at an initial OD_{600} of 0.05. This culture was grown to OD_{600} 0.5, and 50 ml of culture was removed for uninduced timepoint. To the 150 ml culture, IPTG was added to a final concentration of 1 mM and 50 ml of culture was removed after 1 h, 2 h and 4 h. Cells for all time points were pelleted by centrifugation at 5,000 g for 5 min and then resuspended in 10 ml of base buffer (BB: 50 mM NaPi, 150 mM NaCl, pH 7.4) and then lysed by sonication. Lysates were centrifuged at 1,500 g for 10 min to remove cell debris and supernatant (whole cell lysate, WCL) was transferred to a fresh tube and protein concentration for all samples was determined by BCA assay. The resulting collected fractions were mixed with 3 µl of blue protein loading dye and approximately 10 µg of protein was analyzed by SDS-PAGE in a 10% acrylamide gel to determine protein expression levels. Electrophoresis was performed at 160 V for 50 min. The protein gel was then fixed with Colloidal fixing solution until gel had visibly shrunk, washed three times with acidified water for 5 min, dyed with Colloidal Coomassie dye solution, and then de-stained with mH_2O. The gel was then visualized using the Epson Perfection V850 Pro Photo Scanner.

2.5 Purification of recombinant PPE64 from *E. coli* inclusion bodies

Recombinant PPE64 (rPPE64) cell whole cell lysates were prepared as described above and then Triton X-100 was added to WCL to final concentration of 1% (v/v) and incubated on ice for 30 min to solubilize membrane proteins. The lysate was then centrifuged at 10,000 g for 15 min at 4ºC and the supernatant was removed to a separate tube. The inclusion body pellet was then washed twice with 40 ml of BB and pelleted each time by centrifugation at 10,000 g for 15 min at 4ºC. Finally, the inclusion body pellet was resuspended and solubilized in BB containing 8 M Urea (BBU) overnight at room temperature. Denatured rPPE64 was then isolated using nickel chromatography. His Pur Ni-NTA resin was first equilibrated with BBU and then incubated with solubilized rPPE64 inclusion bodies to bind $rPPE64_{6His}$ at RT for 2 h. After binding the flow through was collected and then the resin was allowed to settle for 10 min, after which the column was drained slowly into a Falcon 50 ml conical tube to collect the flow through. In sequence, the resin was then washed with BBU+10 mM imidazole base buffer, BBU+20 mM imidazole and BBU+30 mM imidazole base buffer. $rPPE64_{6His}$ was then eluted with BBU+250 mM imidazole. All fractions were then analyzed by SDS-PAGE.

2.6 Refolding and characterizing rPPE64 by size exclusion chromatography

PPE64 purified through nickel chromatography was refolded in the presence of two detergents OPOE and DDM. To a 3 ml sample of denatured $rPPE64_{6His}$, OPOE and DDM was added at 0.5% and 0.02% final concentration, respectively. Urea and imidazole were removed, and the protein was refolded by step wise dialysis using a 3kDa dialysis cassette for a total of 48 h at 4ºC. After 24 h, the dialysis cassette was placed in fresh dialysis buffer for 24 h at 4 ºC. The 3 ml dialyzed protein sample was then concentrated to a final volume of 150 µl using 3kDa ultracentrifugation. Protein concentration was determined by BCA assay and then 100 µl were loaded onto a BioRad

Enrich 650 column and analyzed by size exclusion chromatography. The protein was separated using base buffer containing the appropriate detergent.

3. RESULTS

3.1 PCR amplification of Mtb genes

The expected size for *ppe64* and *rv0125* PCR products were 1673 bp and 989 bp, respectively. Using the primers described in Table 1, we successfully amplified both *ppe64* and *rv0125* from the Mtb H37Rv genome. As shown in Figure 3.1, the observed DNA products from both PCR reactions match the expected size.

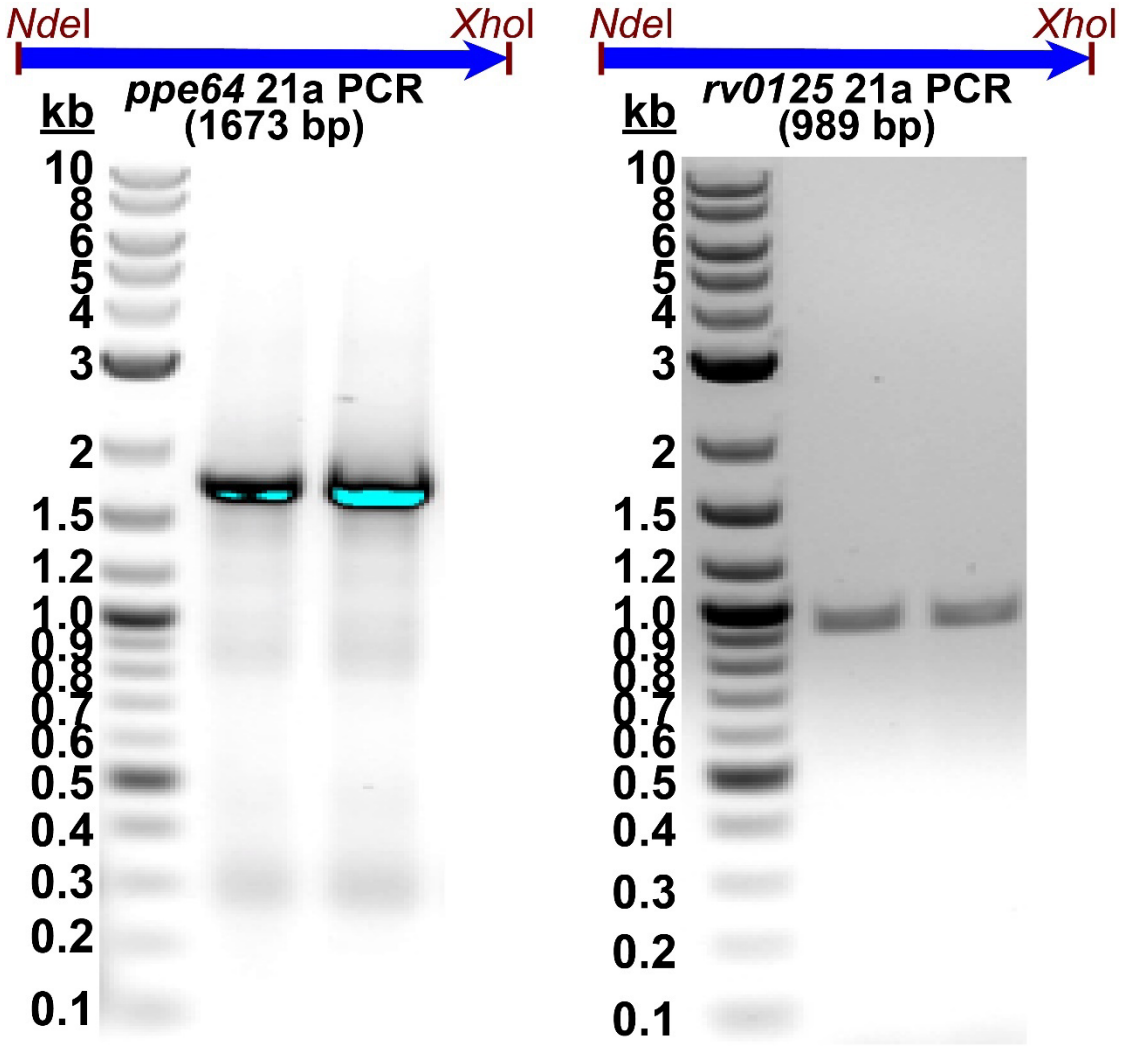

Figure 3.1. Analysis of *ppe64* and *rv0125* PCR products by DNA electrophoresis.

3.2 Validation of cloning of PCR products into pET21a+

To validated cloning of the gene PCR products into pET21a+, putative plasmids were digested with NdeI and XhoI restriction enzymes, which were the same enzymes used to perform cloning. If the plasmid contained the appropriate PCR product, then we expected to see two DNA bands in the gel. We would expect one band at ~5400 bp showing the plasmid backbone and one band at either ~1600 bp (*ppe64*) or ~1000 bp (*rv0125*). Figure 3.2 shows restriction digestion validation for three plasmid clones for each gene. As expected, for all clones we observed the plasmid backbone at ~5400 bp and the expected DNA band for the respective genes. These results demonstrated that our ligation and plasmid cloning were completed successfully.

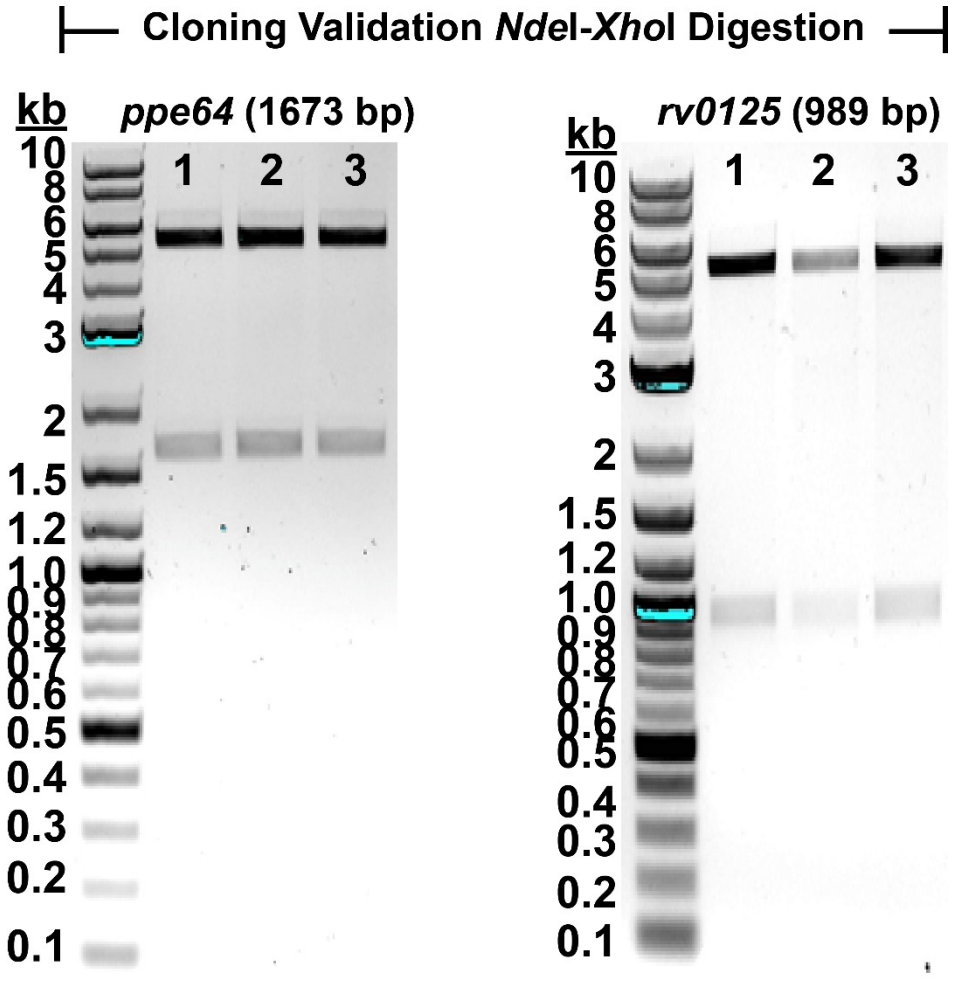

Figure 3.2. Valiation of plasmid clones by restriction digestion. Putative plasmid clones were digested with NdeI and XhoI and then analyzed by DNA electrophoresis.

3.3 Expression of recombinant Rv0125 in *E. coli* BL21

To check for expression recombinant Rv0125 (rRv0125), whole cell lysates from uninduced and IPTG induced cells were analyzed by SDS-PAGE. The expected molecular weight of rRv0125 is ~37 kDa. Figure 3.3 shows protein samples from all timepoints and only a faint ~37 kDa band is observed at the 4 h timepoint. We attempted to purify rRv0125 using the whole cell lysate from these time points. These whole cell lysates were obtained by low-speed centrifugation (1,500 g) which only removed cell debris and prevented the protein from going into the pellet. Our attempts to purify rRv0125 were unsuccessful because we were not able to obtain any protein sample. These results demonstrated that Rv0125 expression is either extremely low or not expressed in *E. coli* BL21.

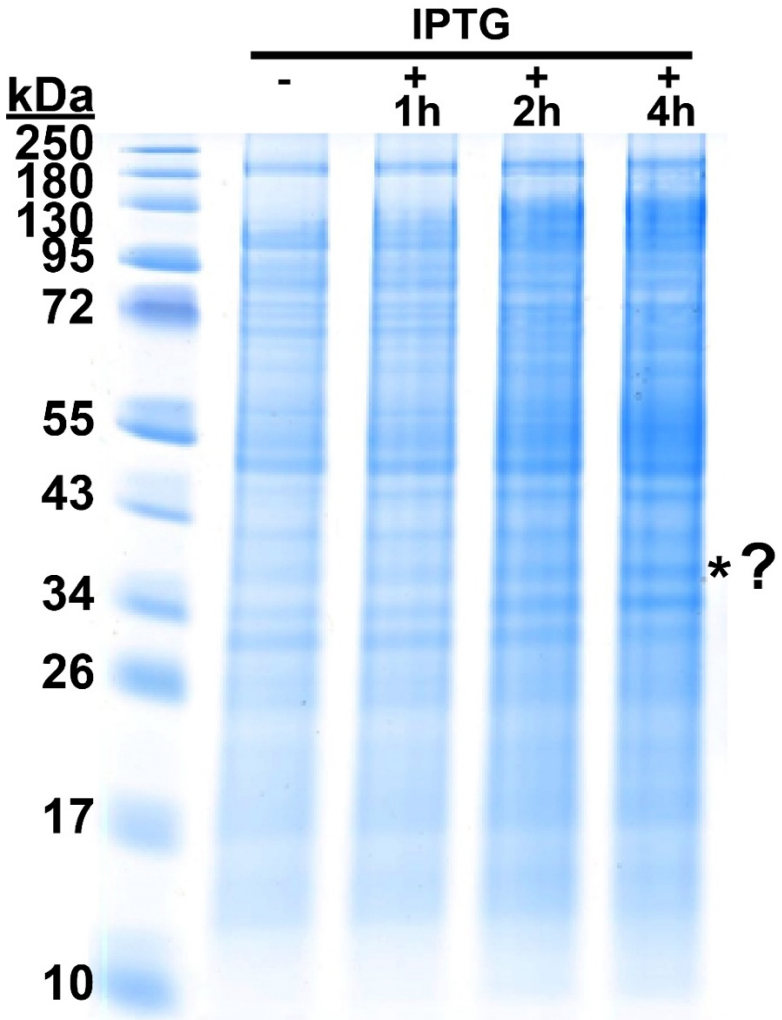

Figure 3.3. SDS-PAGE analysis of Rv0125 expression. Whole cell lysates from *E. coli* BL21 expressing Mtb Rv0125 was analyzed at different timepoints. The expected size of the protein is approximately 37 kDa. The asterisk shows possible induction of the protein.

3.4 Expression and purification of recombinant PPE64 from *E. coli* BL21

To check for expression of recombinant PPE64 (rPPE64), whole cell lysates from uninduced and IPTG induced cells were analyzed by SDS-PAGE. PPE64 is a predicted Mtb membrane protein and previous studies have shown that when Mtb membrane proteins are expressed in *E. coli*, they are produced in inclusion bodies. Proteins in inclusion bodies are always present in the insoluble pellet fractions when the whole cell lysate is separated into insoluble and soluble fractions by centrifugation. Figure 3.4 shows that in the uninduced samples no protein band is present at the expected ~62 kDa (rPPE64) molecular weight. However, when cells are induced with IPTG, a specific protein band is present at ~62 kDa only in the insoluble pellet fraction demonstrating that rPPE64 is produced in inclusion bodies. Based on this observation, we purified rPPE64$_{6His}$ using nickel chromatography under denaturing conditions from the inclusion bodies (Figure 3.5).

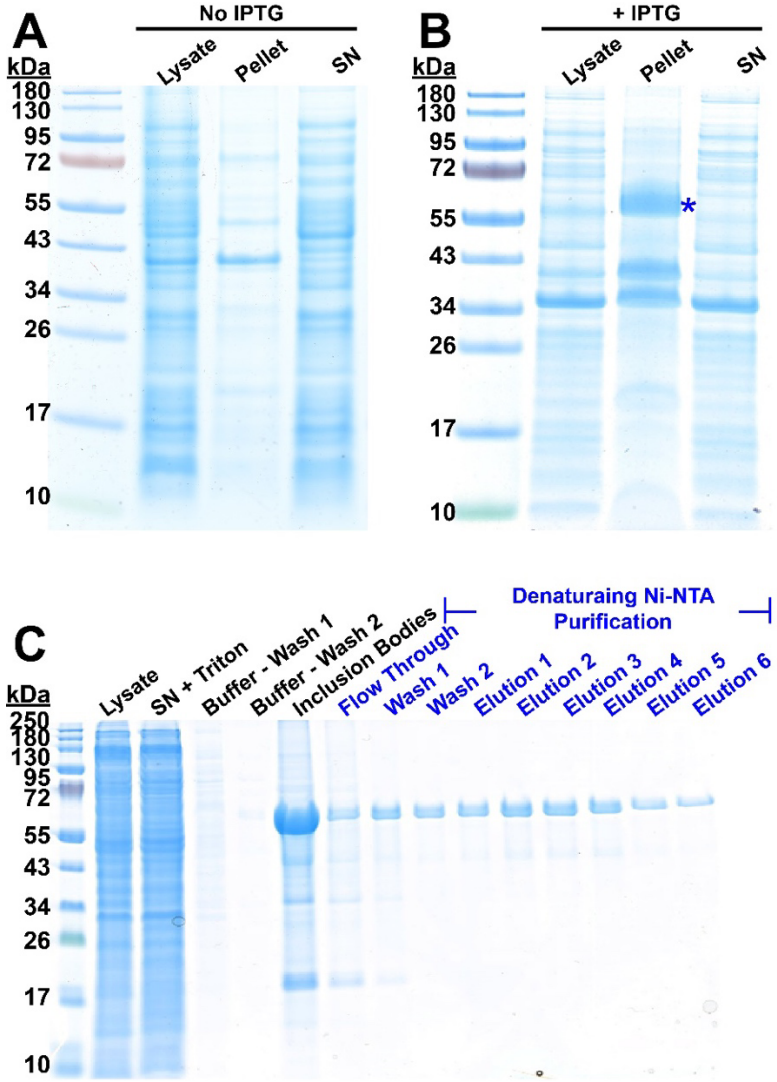

Figure 3.4. Expression and purification of recombinant PPE64 expression. Lysate, insoluble and soluble fractions of *E. coli* BL21 expressing Mtb PPE64 was analyzed from uninduced **(A)** and induced **(B)** samples. **C.** rPPE64$_{6His}$ was purified using nickel chromatography under denaturing conditions.

3.5 Expression and purification of recombinant PPE64 from *E. coli* BL21

Our attempts to remove urea from the purified rPPE64$_{6His}$ fractions (Figure 3.4C) in the absence of detergents always resulted in the protein precipitation. This is consistent with previous data showing that PPE64 is a predicted membrane protein [25], which require detergents for proper folding and stability. As such, we tested the refolding of rPPE64 in the presence of two detergents, OPOE and DDM. We were successfully able to remove urea from Ni-NTA purified rPPE64 fractions without any precipitation (Figure 3.5). These results confirmed that PPE64 is indeed a membrane protein. Refolded rPPE64 was then further characterized through size exclusion chromatography to determine efficiency of refolding (Figure 3.6). In the presence of OPOE, PPE64 refolds into numerous different species. However, PPE64 refolds into a more consistence state in the presence of DDM.

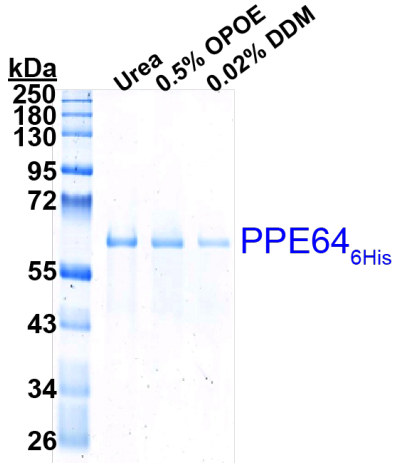

Figure 3.5. Refolding and stability of rPPE64 in OPOE and DDM.

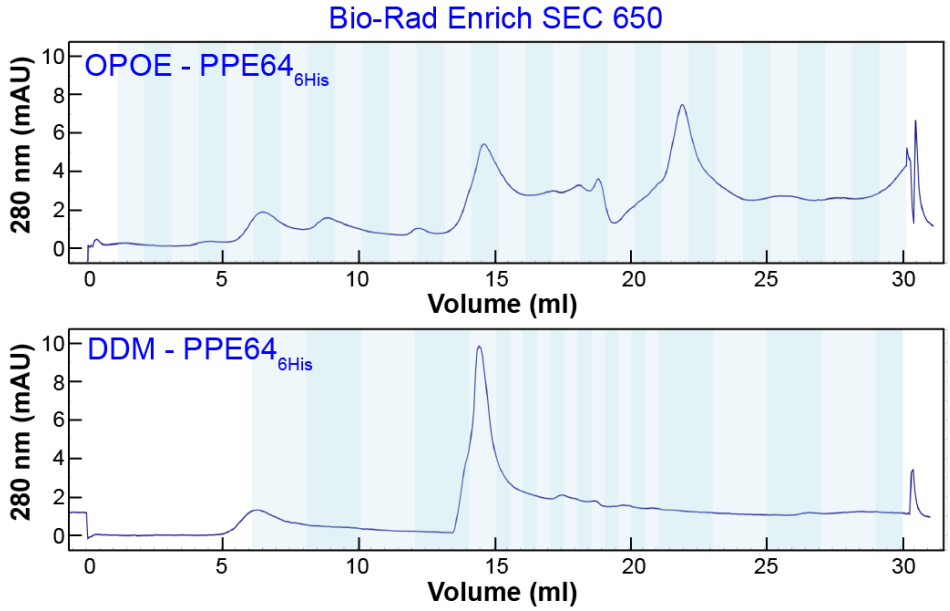

Figure 3.6. Characterization of refolding PPE64 by size exclusion chromatography. OPOE and DDM refolded PPE64 was analyzed by size exclusion chromatography using the Bio Rad Enrich SEC 650 on a BioRad NGC Chromatography System.

4. CONCLUSION

The results of our study have allowed us to identify some of the characteristics of the proteins encoded by the two heme-induced genes, *rv0125,* and *ppe64*, of Mtb. One of the goals of this study was to clone these genes and attempt to purify the encoded recombinant proteins from *E. coli*. This first goal was achieved for both genes of interest (Figure 3.2). The validations by restriction digestion after transformation into *E. coli* DH5α confirmed that our expression vector contained a DNA fragment near the known size of our targets of interest. The *rv0125* gene validation showed two fragments, one between 5.0 kb and 6.0 kb, approximately the same size as our expression vector, and one band at 1.0 kb, which correlates with the size of the *rv0125* gene. In the case of *ppe64,* there was a fragment between 6.0 kb and 5.0 kb since the same expression vector was used for cloning both genes. Additionally, the *ppe64* gel had a fragment present right above the 1.5 kb correlating with the known size of *ppe64.* Furthermore, the genes amplified during PCR of the genes of interest were the same size as the genes in the cloned expression vector.

The second goal of this study was to characterize these proteins through various biochemical experiments. This included analyzing expression of these and then purifying the encoded protein by nickel chromatography. Induction with IPTG did not lead to successful production of Rv0125 in *E. coli*. Our results show that expression of *rv0125* is either very low or non-existent because we were unable to purify any protein. It is possible that transcribed *rv0125* mRNA is not stable in *E. coli* or that the recombinant Rv0125 is unable to fold properly in *E. coli* which results in protein degradation. Additionally, Rv0125 is a predicted protease and overproduction might lead to protein degradation which could be the cause of low protein levels. An alternative approach would be to test Rv0125 protein production level by using a soluble protein fusion tag like Maltose Binding Protein (MBP) to prevent proteolytic degradation. In the case of *ppe64,* we were able to induce gene expression at high enough levels which resulted in PPE64 being produced in inclusion bodies, which is likely due to expression in a non-native environment (Figure 3.4). We successfully purified PPE64 from the inclusion bodies under denaturing conditions. Our results that PPE64 could only be stabilized in the presence of detergents also confirmed previous observations that PPE64 is a membrane protein. We can also conclude that relative to OPOE, DDM is a better detergent for refolding PPE64 because our gel filtration experiments show that PPE64 elutes as a more monodisperse peak in the presence of DDM.

There is no doubt that further studies need to be performed to optimize expression conditions to be able to purify Rv0125. It may be necessary to purify Rv0125 from mycobacteria under native conditions. Future studies for PPE64 will include determining if PPE64 is a channel forming outer membrane protein and if it is a heme binding protein. We will also further characterize the role of PPE64 in Mtb heme utilization by deleting *ppe64* and observing growth of the mutant in the presence of heme.

REFERENCES

1. Khan, A., et al., *Macrophage heterogeneity and plasticity in tuberculosis.* J Leukoc Biol, 2019. **106**(2): p. 275-282.
2. Houben, R.M. and P.J. Dodd, *The Global Burden of Latent Tuberculosis Infection: A Re-estimation Using Mathematical Modelling.* PLoS Med, 2016. **13**(10): p. e1002152.
3. Rabahi, M.F., et al., *Tuberculosis treatment.* J Bras Pneumol, 2017. **43**(6): p. 472-486.
4. Organization, W.H., *Global Tuberculosis Report.* 2020.
5. Nunes-Alves, C., et al., *In search of a new paradigm for protective immunity to TB.* Nat Rev Microbiol, 2014. **12**(4): p. 289-99.
6. Hood, M.I. and E.P. Skaar, *Nutritional immunity: transition metals at the pathogen-host interface.* Nat Rev Microbiol, 2012. **10**(8): p. 525-37.
7. Skaar, E.P., *The battle for iron between bacterial pathogens and their vertebrate hosts.* PLoS Pathog, 2010. **6**(8): p. e1000949.
8. Runyen-Janecky, L.J., *Role and regulation of heme iron acquisition in gram-negative pathogens.* Front Cell Infect Microbiol, 2013. **3**: p. 55.
9. Ogun, A.S., N.V. Joy, and M. Valentine, *Biochemistry, Heme Synthesis*, in *StatPearls*. 2021, StatPearls Publishing Copyright © 2021, StatPearls Publishing LLC.: Treasure Island (FL).
10. Doherty, C.P., *Host-pathogen interactions: the role of iron.* J Nutr, 2007. **137**(5): p. 1341-4.
11. Chao, A., et al., *Iron Acquisition in Mycobacterium tuberculosis.* Chem Rev, 2019. **119**(2): p. 1193-1220.
12. Cassat, J.E. and E.P. Skaar, *Iron in infection and immunity.* Cell Host Microbe, 2013. **13**(5): p. 509-519.
13. Higgs, P.I., R.A. Larsen, and K. Postle, *Quantification of known components of the Escherichia coli TonB energy transduction system: TonB, ExbB, ExbD and FepA.* Mol Microbiol, 2002. **44**(1): p. 271-81.
14. Berntsson, R.P., et al., *A structural classification of substrate-binding proteins.* FEBS Lett, 2010. **584**(12): p. 2606-17.
15. Lau, C.K., K.D. Krewulak, and H.J. Vogel, *Bacterial ferrous iron transport: the Feo system.* FEMS Microbiol Rev, 2016. **40**(2): p. 273-98.
16. Arnold, F.M., et al., *The ABC exporter IrtAB imports and reduces mycobacterial siderophores.* Nature, 2020. **580**(7803): p. 413-417.
17. Zhang, L., et al., *Comprehensive analysis of iron utilization by Mycobacterium tuberculosis.* PLoS Pathog, 2020. **16**(2): p. e1008337.
18. Sandhu, P. and Y. Akhter, *Siderophore transport by MmpL5-MmpS5 protein complex in Mycobacterium tuberculosis.* J Inorg Biochem, 2017. **170**: p. 75-84.
19. Jones, C.M., et al., *Self-poisoning of Mycobacterium tuberculosis by interrupting siderophore recycling.* Proc Natl Acad Sci U S A, 2014. **111**(5): p. 1945-50.
20. Jones, C.M. and M. Niederweis, *Mycobacterium tuberculosis can utilize heme as an iron source.* J Bacteriol, 2011. **193**(7): p. 1767-70.
21. Mitra, A., et al., *PPE Surface Proteins Are Required for Heme Utilization by Mycobacterium tuberculosis.* mBio, 2017. **8**(1).
22. Mitra, A., et al., *Heme and hemoglobin utilization by Mycobacterium tuberculosis.* Nat Commun, 2019. **10**(1): p. 4260.

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. 10101-10109.
24. Osbourn, A.E. and B. Field, *Operons.* Cell Mol Life Sci, 2009. **66**(23): p. 3755-75.
25. Song, H., et al., *Identification of outer membrane proteins of Mycobacterium tuberculosis.* Tuberculosis (Edinb), 2008. **88**(6): p. 526-44.