DISSERTATION

NOVEL EXTRACELLULAR PRODUCTS OF *MYCOBACTERIUM TUBERCULOSIS*: COMPOSITION, SYNTHESIS, AND RELEVANCE TO DISEASE

Submitted by

Andrés Obregón-Henao

Department of Microbiology, Immunology and Pathology

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Spring 2013

Doctoral Committee:

Advisor: John T. Belisle

Patrick J. Brennan
Norman P. Curthoys
Steven W. Dow

ABSTRACT

NOVEL EXTRACELLULAR PRODUCTS OF *MYCOBACTERIUM TUBERCULOSIS*: COMPOSITION, SYNTHESIS, AND RELEVANCE TO DISEASE

Mycobacterium tuberculosis (*Mtb*) is a bacterium causing great morbidity and mortality especially in developing countries. In order to identify possible areas of intervention to positively alter the history of the disease, a better identification and characterization of *Mtb* virulence determinants is required. Specifically, biosynthetic routes for these virulence determinants should be pursued. Furthermore, the interaction between the host and *Mtb* virulence determinants should be characterized at a molecular level. It is hoped that unraveling these pathogenesis mechanisms could lead to novel strategies to combat the infection.

In Chapter II, the identification of secreted *Mtb* molecules that induce macrophage apoptosis was performed. Apoptosis is a mechanism of host cell death and in the life cycle of *Mtb*, different modalities of host cell death have been suggested to tip the balance between bacterial eradication and multiplication. However, a systematic approach to identify and characterize secreted *Mtb* molecules that modulate host cell death, has not been performed. Surprisingly, extracellular *Mtb* RNA fragments were identified as a potent inducer of host cell apoptosis. This extracellular RNA was identified as predominantly rRNA and tRNA fragments that accumulated early during *in vitro* culture of *Mtb*. Mechanistic studies determined that the *Mtb* RNA induced macrophage apoptosis through a caspase-8-dependent, TNF-α-independent mechanism. Importantly, *Mtb* RNA abrogated the macrophage's ability to control an *Mtb* infection. In Chapter II, the first description of an extracellular *Mtb* RNA with potent biological

activity was performed. This opens an exciting field in research of host interactions with pathogen nucleic acids.

Chapters III and IV were devoted to identifying the biochemical pathway involved in α-L-polyGlutamine (α-L-polyGln) biosynthesis and determining its role in pathogenesis in the murine model of TB. α-L-polyGln is an *Mtb* and *Mycobacterium bovis* (*M. bovis*) specific product and its presence in virulent *Mycobacterium* spp., suggest that it could play an important role in pathogenesis. *Bacillus anthracis* (*B. anthracis*) synthesizes γ-D-polyGlutamate (γ-D-polyGlu), an amino acid polymer that is present in its capsule and is absolutely required for pathogenicity. As the pathway for *B. anthracis* γ-D-polyGlu biosynthesis has been well characterized, it was used as a model to start elucidating the *Mtb* α-L-polyGln biosynthetic pathway. Bioinformatics analysis suggested that Rv0574c and Rv2394 are the *Mtb* homologues for *B. anthracis* CapA and CapD, respectively. In Chapter III, a complete biochemical characterization of Rv2394 was performed. Similar to other γ-glutamyltranspeptidases (GGTs), Rv2394 had a conserved catalytic motif consisting of a Threonine (Thr) residue. Mutating this Thr residue to Alanine (Ala) abrogated the enzymatic activity of Rv2394, including its autocatalytic activation. In contrast to eukaryote GGT, Rv2394 was able to perform a GGT activity in the presence of physiological relevant acceptors such as di- or oligopeptides containing Glutamate (Glu) or Glutamine (Gln). In addition to its autocatalytic activation, Rv2394 was shown to be post-translationally modified with hexose residues. A putative phosphorylation and acylation modification also seemed to be present in Rv2394.

In Chapter IV, *Mtb* mutants for *rv0574c* and *rv2394* were engineered and characterized biochemically to determine if the amount of α-L-polyGln had been altered. Furthermore, the mutant's virulence was evaluated in the murine model of TB. Consistent with a putative role in

α-L-polyGln, both mutants had reduced amounts of Glu and ammonia in the cell wall. Furthermore, preliminary analysis suggested that the apolar lipid profiles were also altered by these mutations. In the murine model, *Mtb* mutants had a tendency to grow faster in the initial stages of disease. However, the difference between wild type (WT) and mutant strains was not statistically significant and normalized during the later stages of disease. Furthermore, mutant *Mtb* also seemed to induce more lung damage. In contrast to bacterial burden, this difference persisted throughout the course of the study. Altogether, these results suggest that Rv0574c and Rv2394 participate in the biosynthesis of α-L-polyGln. Remarkably, similar biochemical and phenotypic results were obtained for both mutants despite being encoded in different loci. These initial results provide the foundation for future studies characterizing the biochemical pathway involved in α-L-polyGln biosynthesis.

ACKNOWLEDGEMENTS

I would like to thank my advisor Dr. John T. Belisle for financing my education and giving me the opportunity to work in his lab. The experience at Dr. Belisle's lab was enlightening in both personal and academic aspects. Despite the difficulties, I'm satisfied I worked in these challenging projects as they helped me fulfill the goal of achieving a higher education. I would also like to thank my committee members Drs. Patrick J. Brennan, Norman T. Curthoys and Steven W. Dow for their time and suggestions along the way. I would like to thank my colleagues at Dr. Belisle's lab and the Mycobacteria Research Lab at Colorado State University, for their support and encouragement. Finally, I would like to acknowledge the contributions from the different collaborators that had a role in my thesis: Drs. Luis F. Garcia, Blanca Ortiz, Mauricio Rojas at Universidad de Antioquia (Chapter II), Drs. William Jacobs, Rainer Kalscheuer at Albert Einstein College of Medicine, Bronx, NY (Chapter IV), Drs. Angelo Izzo and Randall Basaraba at Colorado State University (Chapter IV).

DEDICATION

I would like to dedicate this work to my family, especially my wife Marcela and my two sons Emilio and Matias. Their love and happiness was a major driving force to complete this dissertation… to my parents and brothers who unconditionally motivated me during these years… to my uncle Lito for showing me the wonders of curiosity… to my family members that passed away during this time: Doña Chava, Don Feliz y Merce.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter I

Literature Review

1.1 Introduction

After the advent of antibiotics in the second half of the twentieth century, there was a feeling that infectious diseases were going to be eradicated. This feeling could be best exemplified by the following infamous phrase attributed to Dr. William H. Stewart, the US Surgeon General during the late 1960's: "It is time to close the book on infectious diseases, and declare the war against pestilence won" [1]. Twenty years later (1986) some infectious disease experts like Dr. Robert Petersdorf still believed that fellows in infectious diseases were soon to be jobless as he reflected on the number of students graduating each year: "the millennium where fellow in infectious disease will culture one another is almost here" [2]. Only seven years had passed after this comment when in 1993 the World Health Organization (WHO) declared tuberculosis (TB) a public health emergency. As described below, there was plenty of evidence that TB was once again a major health threat like it had been before the antibiotic era. In contrast to the optimistic quote from Dr. Stewart, the war against pestilence (specifically for TB) had basically re-started. WHO's declaration should be considered a major milestone in the field of TB as it brought the disease back to the spotlight. Enormous economical and intellectual efforts have started paying their dividends and some of the ominous trends in the field of TB have been regressing.

1.2 Epidemiology

When in 1993 WHO declared TB as a public health emergency, it was estimated that 6-8 million people per year were getting sick with *Mtb*. Furthermore, 1.3-1.6 million people died annually of TB. In the latest WHO report released for the year 2011 [3], the absolute numbers are similar as approximately 8.5-9.2 million new cases and 1.2-1.5 million deaths were estimated to occur. Taking into consideration the dramatic demographic explosion in the last decades, incidence and mortality rates per 100,000 people

have actually decreased significantly. It has been reported that the incidence has been decreasing at a rate of 1.3% annually since 2006 and more significantly, deaths have been reduced by one third since 1990. Indeed, if this trend continues the goal set by the Stop TB strategy of eliminating TB for the year 2050 (1 case/1 million population/year), might be met.

TB is a disease that predominantly affects men (64%) in their most productive time of their life (ages 15-59). Most cases of TB are diagnosed in 22 developing countries and approximately 65% of the cases occur in China and India [3].

The gold standard for diagnosing TB is culturing, however in most countries this is hardly performed. Instead, identification of acid fast bacilli by light microscopy is the predominant diagnosis technique. However, it has been estimated that only 57% of the diagnosed TB cases came from individuals with acid fast bacilli in their sputum. Thus, it can easily be appreciated that many TB patients are not being promptly diagnosed and treated. In fact, the current case detection rate is 65% and it is hoped that by 2050, all TB cases are diagnosed [3].

The current drug treatment against *Mtb* consists of 2 months of rifampicin, ethambutol, isoniazid (INH), and pyrazinamide, followed by 4 months of rifampicin and INH. When correctly applied, directly observed treatment short (DOTS) course, leads to a 90% success rate. Since the implementation of DOTS in 1995, more than 50 million people have been treated and approximately 7 million lives have been saved. However, treatment non-compliance is common and this seems to be leading to the appearance of antibiotic resistant strains. It is currently estimated that approximately 5% of all TB cases are attributed to multiple-drug-resistant (MDR) strains. Unfortunately, only 18% of MDR cases are reported. Ideally all *Mtb* isolates should be evaluated for drug susceptibility; however, this only occurs in 2% of the cases. In contrast to the 6 month treatment required for drug susceptible strains, treatment against MDR strains requires 20 months. Furthermore, treatment success rates against MDR strains are approximately 60-75% [3].

Fortunately, the international community has become aware of the magnitude of the TB epidemic and has doubled the funding in the last five years. Several new antibiotics are already in phase II and III clinical trials. In addition, nine vaccines are currently in some stage of clinical trials [3]. It is hoped that new successful interventions can significantly alter the trend so as to achieve Stop TB goals by 2050.

1.3 *Mtb* life cycle.

Mtb life cycle starts by the airborne transmission from an infected individual to a naïve one, reaching the alveolar sacs in the lower respiratory airways [4-6]. Resident macrophages engulf and internalize *Mtb* by phagocytosis and several outcomes can result from this initial interaction: bacteria can be eradicated, alternatively bacteria thrive intracellularly and replicate, or bacteria can remain in a metabolic quiescent state know as dormancy to microbiologists. As discussed above in the epidemiology section, statistical data suggests that this initial interaction is critical: 10% of the infected individuals progress to an acute form of the disease (bacteria thrived after the initial contact), whereas 90% of infected individuals harbor dormant bacilli and are referred as having latent disease [3]. A note of caution should be taken to avoid using the terms dormancy and latency interchangeably. Whereas the former term refers to a physiological and metabolic bacterial state characterized by minimal activity, the later terms refers to a clinical condition in which individuals are infected but there's no clinical evidence of ongoing disease. As discussed in Chapter V, failure to separate these terms created enough confusion in the TB field so as to affect some lines of experimentation. The third outcome of this initial interaction in which mycobacteria are eradicated after being deposited in the alveolar sacs, is certainly the ideal situation. Unfortunately, we don't know its frequency and if containment of low numbers of *Mtb* is sufficient for purified protein derivative (PPD) conversion, falsely incrementing the statistics for latent disease.

As *Mtb* starts replicating exponentially (either during acute infection or after re-emergence from dormancy), an initial inflammatory response begins accumulating at the site of infection. Animal models have been extremely useful to delineate the cell type and kinetics of this infiltrate. Consistently for both

the murine and guinea pig model, the initial inflammatory response occurring during the first 20 days of infection is characterized by monocytic mononuclear cells [7]. The guinea pig model has also revealed the presence of an initial wave of granulocytic cells termed heterophils, but this is not paralleled in the murine model [8]. As adaptive immunity (see below) takes hold during the third week of infection, greater influx of lymphocytic cells both $CD4^+$ and $CD8^+$ T cells as well as B cells is observed. In addition, activation of the adaptive immune response culminates the phase of exponential mycobacterial growth. Failure to do so such as in the interferon-γ knock-out (GKO) mice model, leads to the animal rapidly succumbing to disease [9]. In contrast, in the immunocompetent murine model, activation of the adaptive immune response results in a fairly constant mycobacterial burden lasting several months, whereas a gradual but persistent increase in mycobacterial numbers occurs in the guinea pig model.

The end result of accumulating inflammatory cells is the granuloma, a hallmark of TB and other infectious and non-infectious chronic diseases [10]. Histologically, the mycobacterial granuloma is characterized by a central area predominantly composed of monocytic mononuclear cells. Surrounding this central area is a rim of $CD4^+$ T cells, B cells and fibroblasts**.** Further outside, scuffs of $CD8^+$ T cells are observed in discrete patches. Similar histological structures are not only found in the lungs, but in other organs where *Mtb* has disseminated.

Recently, it has become evident that granulomas are more dynamic than previously recognized. Granulomas can necrotize and either liquefy or calcify. Alternatively, granulomas can regress if proper medication is given [11, 12]. Finally, granuloma structure can also disorganize in the presence of immunosuppression. Antibiotic-induced granuloma regression has been widely studied in animal models. In this case, the majority of cellular infiltrates are cleared as bacteria are removed and normal lung structure and function is left. Drug treatment does not remove necrotic granulomas [11, 13], which will be discussed below. In terms of disorganized granuloma architecture, the best examples have been obtained during acquired immunosuppression such as in AIDS patients and in humans or animals receiving anti-TNF-α treatment [14-16]. In these cases, disorganized granuloma structure correlates with increased

mycobacterial numbers clearly indicating that granuloma have an important role in mycobacteria containment. In regards to the pathogenesis and life cycle of *Mtb*, granuloma necrosis is a significant event. Two major outcomes of necrosis have been readily documented in both humans and animal models: necrotic granulomas can calcify and heal or liquefy. From a clinical perspective, the advantage of calcified granulomas was the possibility to identify them by simple X-ray techniques that became available in the late 1800s. Alternatively, necrotic granulomas that don't heal can rupture and the necrotic fluid containing multiple *Mtb* is discharged into nearby airways. If this process continues for some time, enough tissue disruption can originate cavity formation significantly enhancing *Mtb* transmission potential. This last process of cavity formation is rarely seen in animal models except in the rabbit [17-19]. Actually, several of the previous steps leading to cavity formation such as liquefaction and even granuloma necrosis also don't occur in animal models. Except for immunocompromised mice such as the GKO, recombinant activation gene (RAG) and nitric oxide synthase-2 (NOS-2) KO mice, granuloma necrosis doesn't occur [20]. Recently, it was appreciated that the immunocompetent C3HeJ/F strain also develops necrotic granulomas during *Mtb* infection [21]. Necrotic granulomas do occur in the guinea pig and rabbit model [19]. Lack of necrotic granulomas in mice, has been one of the major drawbacks for using the murine model to study some aspects of TB.

Necrotic granulomas have two additional properties worth considering in the life cycle of *Mtb*. A major characteristic of necrotic granulomas is their low oxygen tension. Since the beginning it was proposed that necrotic granulomas could be hypoxic. Recent experimental evidence has indeed confirmed this. Immunohistochemical staining with pimonidazole, a reagent binding hypoxic zones has identified the necrotic rim as the interphase where hypoxia begins [11]. Hypoxia would certainly extend into the necrotic area inside the rim but its low permeability hinders pimonidazole diffusion and labeling of necrotic contents. More importantly, subsequent data obtained by inserting oxygen-sensing electrodes into necrotic granulomas of living guinea pigs and rabbits, unequivocally confirmed hypoxia in these structures [22, 23]. In contrast, murine granulomas were determined to be micro-aerophilic using this

same technique [24]. From *Mtb's* perspective, hypoxia represents a condition hindering its metabolism and replication, possibly leading to mycobacterial dormancy. Life in the necrotic granuloma isn't completely deleterious to *Mtb* [25-27]. It has recently been determined that the necrotic milieu is rich in potential nutrients for *Mtb* such as fatty acids, cholesterol esters and diacylglycerol [28]. Metabolism of these nutrients is probably slowly occurring during dormancy and could certainly fulfill *Mtb* future metabolic demands once the necrotic granuloma ruptures, liquefies and is discharged into more aerophilic conditions. Additionally, even though granulomas have been considered a mechanism to wall off *Mtb*, necrotic granulomas also represent a niche where *Mtb* is protected from the immune response onslaught. From a clinical point of view, it's believed that these necrotic granulomas represent a major barrier limiting antibiotic levels during therapy.

Mtb life cycle is completed when bacteria are transmitted from the lung of a sick patient to the outside world. As *Mtb* resides in the lower respiratory tract, significant force through coughing is required for *Mtb* to become airborne. This explains why the young and elderly are usually not as contagious as young adults, and also why cavity formation has been associated with increase transmission risk (*Mtb's* adaptation to transmission will be discussed in Chapter V). In the following sections, several critical aspects of *Mtb* life cycle will be emphasized. Specifically, the initial encounter between *Mtb* and host phagocytes, the innate and adaptive immune response against *Mtb*, and transmission will be further discussed.

1.4 Initial encounter and innate immune response.

1.4.1 Interaction with the alveolar sac's acellular constituents.

Once inhaled into the lower respiratory tract, mycobacteria are housed in the alveolar sac, a special niche consisting of three principal components: surfactant, alveolar respiratory epithelium cells and resident macrophages. Surfactant is a fluid predominantly consisting of lipids and proteins [29], and is produced by specialized respiratory epithelial cells (type-II pneumocytes). Its role to reduce surface

tension between air (gas) and blood (liquid) is widely known due to the severe respiratory distress in premature infants lacking it. In contrast, it has only recently been appreciated that surfactant has a role in the lung's innate immunity. One of surfactant's major protein is collectin, a sugar-binding protein (lectin) that binds to sugars present on the surface of microorganisms. Collectin activates the complement pathway and efficiently opsonizes these invading microorganisms [30]. In turn, opszonization assists phagocytosis and microbial clearance by resident macrophages [31, 32]. The interaction of *Mtb* and surfactant has been evaluated *in vitro* [33]. *Mtb* is coated by surfactant and this interaction is predominantly dependent on the presence of lipoarabinomannan (LAM), a glycolipid present on the bacterium's surface [34]. Surfactant-coated *Mtb* is readily phagocytosed and its intra-cellular survival is reduced in comparison to untreated *Mtb* [33].

1.4.2 Interaction with the alveolar respiratory epithelium.

The most abundant cellular component present in alveolar sacs is the alveolar respiratory epithelium. The interaction of *Mtb* and respiratory epithelial cells has also been evaluated in *in vitro* models, but it has also been documented during post-mortem microscopical evaluation of human lung tissues [35]. In general, three major observations have been derived from these studies: alveolar epithelial cells can harbor intracellular *Mtb* and sustain its growth, *Mtb* can migrate to the basolateral side of alveolar epithelial cells and finally, alveolar epithelial cells can die by necrosis during the infectious process [36, 37]. Intracellular *Mtb* can gain access to the basolateral side by a process called trancytosis, a dynamic endosome shuttling that transports cargo (in this case *Mtb*) between the luminal and basolateral side [38]. Alternatively, *Mtb* also reaches the basolateral side by migrating through intercellular junctions between alveolar respiratory cells. The other outcome of *Mtb* and alveolar epithelial cell interaction is the demise of the alveolar cell by necrosis [39]. At a molecular level, necrosis of alveolar cells was shown to be dependent on the mycobacterial proteins encoded by the RD1 region, which collectively constitute a secretion apparatus [37] (esx-1, section 1.6.1.1). Even though the significance of alveolar cell necrosis

was not evaluated experimentally, necrosis of the alveolar epithelium would in essence generate an ulcer, allowing the bacterial's access to the basolateral side. Reminiscent of the intestinal epithelium necrosis induced by *Salmonella* and *Shigella*, accessing the basolateral space endows *Mtb* with the potential to disseminate via lymphatic or blood vessels.

The interaction between pneumocytes and *Mtb* has been characterized at the molecular level. Proteoglycans present on the surface of pneumocytes act as the ligand for at least two mycobacterial proteins: the mycobacterial protein heparin-binding haemagglutinin (HbHA) [40, 41] and the mycobacterial DNA-binding protein 1 [42]. In a cell free assay, enzymatic removal of the pneumocytes' proteoglycans abrogated HbHA's binding. Subsequently, it was shown that proteoglycan removal inhibited the interaction between *Mtb* and pneumocytes. As described below, this interaction between *Mtb* HbHA and pneumocytes was further shown to be critical for *Mtb in vivo* dissemination [40].

Altogether, these results suggest that alveolar epithelial cells do not represent a major barrier or defense against *Mtb*, beyond the production of surfactant by type II pneumocytes.

1.4.3 Interaction with the macrophage.

As macrophages are the host cell target of *Mtb*, understanding the outcome of this interaction has been a priority [43]. Contacts between *Mtb* and macrophages are mediated by their respective surface molecules. In addition, *Mtb* can be covered or opsonized by host molecules present in serum and surfactant [33], such as immunoglobulins (Ig), complement or lectins [44, 45]. In general, opsonization has several consequences benefitting the host such as enhancing phagocytosis, host cell activation and microbicidal activity. The macrophage is endowed with several receptors that can mediate phagocytosis and/or activation, and there's significant cross-talk between them. Regarding the process of phagocytosis, it's become evident that multiple receptors may participate during *Mtb* phagocytosis and there's really not a predominant receptor mediating this [46, 47].

For non-opsonized *Mtb*, phagocytosis can be mediated by lectin receptors such as the macrophage mannose receptor (MMR) [48], DC-SIGN (this applies for dendritic cells discussed below) and complement receptor 3 (CR3 or CD11b/CD18 [30, 49]. At the molecular level, the mycobacterial surface-exposed molecules mannosylated lipoarabinomannan (ManLAM), lipomannan (LM), phosphatidyl-inositol mannosides (PIMs) and glucan seem to be principally targeted by these receptors, respectively. For ManLAM, mannosylation of its non-reducing end is essential for binding to occur to MMR [50, 51]. Enzymatic removal of the ManLAM's non-reducing mannose residues readily diminishes uptake of LAM-coated beads. Furthermore, phagocytosis of *Mycobacterium smegmatis* (*M. smeg*), LAM-coated beads is severely hindered. This has been attributed to the fact that *M. smeg* LAM is not mannosylated at its non-reducing end, instead displaying alternative modifications such as inositol phosphates [52]. An additional line of evidence that phagocytosis of *Mtb* and ManLAM-coated beads occurs through the MMR, is the reduced rate at which this proceeds in the presence of antibodies blocking the MMR.

Mtb binding to macrophages has also been shown to be mediated by CR3 [30, 44, 53]. In addition to the C3 binding motifs, CR3 also has a well characterized lectin domain. At the molecular level, CR3 can bind to the mycobacterial surface exposed glucan, a poly-saccharide consisting exclusively of a polymer of glucose residues in α-1-4 linkage [54]. (See section 1.6.2.2 for glucan discussion)

In contrast to non-opsonic phagocytosis, opsonic phagocytosis has been associated with increased macrophage bactericidal activity. Several Fc receptors (FcR) are present on macrophages and for mycobacteria, receptors for IgA, IgG and IgM have been readily studied. IgA is the predominant Ig present in respiratory mucosa. An isolated study showed increased protection against *Mtb* when an anti-ManLAM vaccine was designed to enhance IgA levels against this mycobacterial glycolipid [55-57]. The majority of experiments evaluating immunoglobulin opsonization have been performed with either IgG or IgM. Two major cellular mechanisms have been found to increase the macrophage's mycobactericidal activity when macrophages engage and engulf antibody-opsonized mycobacteria through its FcR: increase phagosome-lysosome fusion and enhanced macrophage respiratory burst. During phagocytosis,

antibody-opsonized mycobacteria induce greater intracellular Ca^+ fluxes required for phagosomal maturation [58-60]. Ca^+ fluxes are targeted by non-opsonized mycobacteria through an active process as evidenced by the fact that opsonized but dead mycobacteria are incapable to do so. Despite the *in vitro* data suggesting that antibody-mediated immunity could be important to control mycobacterial infections, it should be pointed out that the *in vivo* role of antibodies in TB is highly controversial. Except for the previously discussed Man-LAM vaccines aimed at enhancing mucosal IgA levels, antibody generation is not a major objective of anti-mycobacterial vaccines.

Simultaneous to the mycobacteria engaging macrophage receptors directly involved in phagocytosis, additional interaction occurs with other macrophage surface receptors which participate in cell activation but have minor or no role in phagocytosis. Activation of these receptors leads to signal transduction across the plasma membrane, usually culminating in gene expression. Specifically to *Mtb*, probably the most important macrophage activation receptor is Toll-like receptor (TLR)-2. Several mycobacterial products have been shown to activate TLR-2. In addition to the mannosylated glycolipids mentioned above, mycobacterial lipoproteins have been consistently shown to activate macrophages through the TLR-2 pathway [61]. Mycobacterial lipoproteins belong to an abundant family of proteins which are characterized by the post-translational addition of lipid moieties in their respective N-terminus [62]. The presence of these mycobacterial lipids confers lipoproteins with TLR-2 triggering activity. In regards to the host-pathogen relationship, activation of TLR-2 has both beneficial and detrimental effects to the host and *Mtb*. From the host perspective, TLR-2 activation is important in downstream signaling events leading to macrophage activation. Indeed, in murine macrophages mycobacterial lipoprotein activation of TLR-2 was shown to generate mycobactericidal concentrations of nitric oxide [63]. Macrophage apoptosis leading to diminished *Mtb* survival was also observed when this interaction occurred (apoptosis will be discussed below). Furthermore, it was recently described that in human macrophages, TLR-2 activation induced a VitaminD-dependant mycobactericidal activity [64]. From the pathogen perspective, mycobacterial lipoproteins reduce major histocompatibility complex (MHC)-II

levels in macrophages [65, 66]. As MHC-II molecules participate in antigen presentation to T cells, reduction of MHC-II is an important *Mtb* virulence factor. Recent experimental evidence suggested that TLR-2 activation by *Mtb* induced arginase-I production in macrophages [67]. Reduction of arginine levels by this enzyme curtailed the efficiency of T cell response. It was hoped that the generation of TLR-2 KO mice could help clarify the role of this receptor. Unfortunately, contradictory reports were published regarding the effect of TLR-2's deficit on mycobacterial burden [68, 69] and infections with higher bacterial numbers were usually required to see an effect. In humans, polymorphisms in TLR-2 and other TLRs have been associated with a minor but statistical significant increase in susceptibility to TB [70, 71]. Altogether, these results highlight the complex downstream consequences emanating from host-pathogen interaction.

Another example of an important cell surface receptor that interacts with *Mtb* but doesn't participate during phagocytosis is Mincle. Mincle was identified as the receptor for trehalose-di-mycolate (TDM) or cord factor [72, 73]. TDM is a mycobacterial lipid with potent immunomodulatory and granulomatous-inducing activities and is widely accepted as being one of the *Mtb* virulence factors [74-77]. Consequently, the identification of the elusive host receptor for TDM had created great expectations to explain some of *Mtb* pathogenesis mechanisms. Similar to TLR-2, generation of Mincle KO mice did not show any overall effect in the pathogenesis of TB [78]. Besides being a warning for the importance of translational science in which *in vitro* experiments are evaluated in *in vivo* settings, one consideration is that the innate immune system is endowed with multiple and redundant mechanisms to ensure proper identification of foreign or dangerous signals.

Once particles or microorganisms are internalized into vesicles by phagocytosis, the default pathway is a complex cell biology event progressing to phagosome maturation and lysosomal fusion. This default pathway guarantees at least two events: microorganism inactivation and destruction in the more acidic lysosomal compartment, and antigen processing and presentation to activate the adaptive immune response. Early cell biology studies aimed at identifying the reason why *Mtb* thrived in macrophages,

identified failure of phagosome maturation as a major culprit interfering with *Mtb* eradication [79, 80]. These studies also suggested high ammonia concentrations being produced by *Mtb* as one of the mechanisms behind the failed maturation process [81]. This hypothesis has not been pursued to a great extent, but one conclusion that has endured years of experimentation is that arresting phagosome maturation is an active mycobacterial process requiring live bacilli. Furthermore, phagosomal arrest was exclusive to pathogenic bacteria such as *Mtb*, *M. bovis* and even the more attenuated *M. bovis* Bacillus Calmette-Guerin (BCG). In contrast, avirulent mycobacteria such as *M. smeg* lacked this property, thus phagosomal maturation arrest is still considered one of the most important *Mtb* virulence factor. Subsequent immunohistochemistry studies provided evidence that mycobacteria-containing phagosomes were stalled at the early endosomal stage [82, 83]. Mycobacterial phagosomes were identified to contain the transferrin receptor [82], a recycling receptor known to releases its cargo at the early endosomal stage before returning to the cell membrane. Later on, proteomic studies characterizing *Mtb*-containing phagosomes demonstrated these were devoid of H^+-ATPase responsible for lysosome acidification [84]. In conjunction with cell biology studies, the mycobacterial phagosome was also shown to retain the early endosomal marker rab5 but was lacking the late endosomal marker rab7 [85-88]. Finally, the lack of phosphatidyl-inositol-3-phosphate kinase (PI3K) on the phagosomal cytoplasmic surface was identified as one of the major culprits explaining the failed phagosomal maturation [87, 89, 90]. The lipid phosphatidylinositol-3-phosphate (PI3P) is the product of the enzymatic activity of PI3K. This lipid plays an important role in regulating protein-lipid interactions required for membrane fusion events [91]. More importantly, the similar structure of PI3P and some of *Mtb* phosphatidylmannosides could explain at the molecular level the reason for phagosomal maturation arrest. Indeed this was confirmed to be the case, and PIMs have been clearly identified as one of the principal molecules participating in this virulence mechanism [92, 93]. Other *Mtb* lipids like sulfolipids and TDMs have also been described to inhibit phagosome-lysosome fusion [94]. Thus, it seems *Mtb* is endowed with redundant mechanisms to guarantee this virulence mechanism is functional.

For *Mtb*, life in the macrophage's early endosome has several advantages. Intracellular *Mtb* are protected from extracellular innate immune components like serum Igs and complement. More importantly, it has become evident that its location in early endosomes will allow *Mtb* to access nutrients that the macrophage is internalizing through the endosomal pathway. In addition to the iron present in transferrin [95], *Mtb* also has access to lipoproteins en route to the lysosome [96]. Elemental analysis of the *Mtb* phagosome has confirmed the presence of other metals such as zinc [97], that could be used by *Mtb*. As described below, the cytokine IFN-γ potently activates macrophages to control *Mtb*. One mechanism by which IFN-γ activates macrophages is by circumventing *Mtb*-induced phagosomal arrest [98]. IFN-γ has been shown to induce the expression of additional GTPases which enhance phagosomal maturation as well as oxidative burst, autophagy and anti-microbial peptides [99]. As *Mtb* is efficiently shuttled to the lysosome, it becomes exposed to a low pH which is critical for activation of host proteases, lipases etc. that inflict damage to *Mtb [83]*. In addition, the *Mtb* would be denied of the nutrients available in the early endosomes. Once again, elemental analysis of *Mtb* phagosomes from IFN-γ activated macrophages has shown a significant reduction of the ions available in the early endosome [97]. In conclusion, *Mtb* has developed several mechanisms guaranteeing its presence in the comfortable niche of early endosomes but fortunately the immune system can circumvent this. The outcome of this "positional" battle is an important determinant of *Mtb* survival's or death.

As mentioned, IFN-γ is an important cytokine activating the macrophage to control *Mtb* [100]. Even though this cytokine is produced by innate immune cells such as NK and γδ T cells, the role of these cells in controlling *Mtb* is not completely understood. Instead, IFN- γ is the classical cytokine of a TH1 acquired immune response and is the major target sought by vaccination strategies. As this response is delayed by almost three weeks during *Mtb*, certainly other innate immune responses are in place to minimize mycobacterial growth while the development of the acquired adaptive immune response is achieved. In the last couple of years, several additional macrophage mechanisms have been described to participate in controlling *Mtb* growth. Autophagy, a process in which intracellular components such as

proteins or even organelles are targeted to the lysosome for destruction, has been recently described to play a significant role in innate immune responses to *Mtb* [101]. During autophagy, components destined for degradation are enclosed in a double membrane bi-layer as opposed to the single membrane bi-layer present in most organelles. Ultrastructural studies of *Mtb*-infected macrophages have confirmed that some *Mtb* phagosomes are actually enclosed with an additional membrane constituting an autophagic vesicle [101]. Similar to the classical endosomal pathway, this cellular process is also dependent on the production of PI3P and can be activated by GTPases. *Mtb* molecules blocking autophagy have not been described but PI3P production could also be targeted by mycobacterial lipids. Again, as occurs with the endosomal pathway, IFN-γ has also been shown to increase the formation of *Mtb*-loaded autophagic vesicles [102]. So in essence, the mechanism by which autophagy participates in controlling *Mtb* is by ensuring its delivery to the lysosome.

In contrast to autophagy, the mechanism by which macrophage apoptosis diminishes mycobacterial survival is not completely clear (apoptosis will be thoroughly discussed in Chapter II). Apoptosis is a cell death mechanism dependent on the activation of caspases which induce cell death by cleaving essential proteins required for normal cellular processes. This cell death mechanism is a tightly controlled process in contrast to necrosis which is a catastrophic mechanism of cell death. Initial studies suggested that apoptosis as opposed to necrosis led to mycobacterial death [103]. Mechanistic studies revealed that macrophage apoptosis during *Mtb* infection was dependent on several factors such as caspase activation, TNF-α production, mitochondrial depolarization and Ca+ fluxes [104-107]. Furthermore, downstream events during apoptosis such as membrane micro-disruptions and the translocation of phosphatidyserine (PS) to the extracellular leaflet of the plasma membrane were dependent on the balance of lipidic intermediates such as lipoxins and prostaglandins [108-110]. As expected, some mycobacterial products have been shown to inhibit apoptosis. ManLAM was shown to inhibit Ca^{+} fluxes as well as to induce the production of IL-10, a cytokine that counteracts some of the effects by TNF-α [104]. In addition, ManLAM was also shown to activate the survival pathway mediated

by Akt [111]. Downstream effects of Akt include the induction of gene transcription and phosphorylation of proteins such as BAD and BID, that control apoptosis through the mitochondrial pathway. As mentioned, despite the abundance of literature implicating macrophage apoptosis as a cellular process controlling mycobacterial growth, the mechanism(s) has not been fully elucidated. In some circumstances, macrophage apoptosis correlates with the induction of high concentrations of the mycobactericidal product nitric oxide [112]. It could be further argued that by committing suicide, the macrophage is altruistically eradicating the niche for mycobacterial growth.

1.5 Dissemination and acquired immune response.

1.5.1 Dissemination.

Even though *Mtb* is transmitted via the aerosol route, dissemination to other organs besides the lungs might seem like a dead end. *Mtb* affecting the majority of organs has been described, a testimony to *Mtb* potential to disseminate. *Mtb* dissemination causes several significant morbidity and/or mortality such as during meningeal, renal, pericardial and musculoskeletal TB amongst other forms of the disease. Fortunately, with the advent of antibiotic therapy and better diagnostic methods, these TB complications have been reduced. Furthermore, even though the available TB vaccine BCG doesn't seem to have a major impact in protecting against lung disease, it has a clear effect in reducing complicated TB in younger children.

In the murine model, it has been clearly determined that the thoracic lymph nodes are the first site of dissemination in pulmonary TB. Viable bacteria have been detected at approximately 10 days after an aerosol infection but the low numbers of mycobacteria precludes earlier detection [113]. This event is critical during the development of the acquired immune response as it has been shown that T cell priming occurs at the local lymph nodes rather than at the lungs (discussed in section 1.4.2) [113, 114]. It's presumed that dissemination to the thoracic lymph nodes is also an early event in humans as evidenced by the classical radiologic finding of a Ghon complex [10]. At the radiological level, the Ghon complex is

defined as the identification of a calcified focus in the lung parenchyma in addition to a calcified thoracic lymph node. Actually sometimes the Ghon complex is the only evidence for previous or latent TB. Besides the morbidity attributed to *Mtb* dissemination, it has been speculated that extrapulmonary *Mtb* might be responsible for reactivation or antibiotic failure. Recently, *Mtb* was discovered in the adipose tissue but interestingly these mycobacteria had a dormant phenotype [115]. As mycobacterial dormancy has been associated with enhanced antibiotic tolerance, adipose tissue-embedded mycobacteria could certainly predispose to reactivation.

One aspect of the dissemination process which is still debated is how *Mtb* disseminates. The classical story tells that dendritic cells with high migration capabilities are responsible for transporting peripherally encountered antigens or microbes to the lymph nodes. Dendritic cells have been shown to be infected with *Mtb* both *in vivo* and *in vitro* [116, 117]. It is widely accepted that *Mtb* has the potential to inhibit dendritic cell maturation and migration [118]. Alternatively, some studies have suggested that neutrophils shuttle mycobacteria to lymph nodes via afferent lymphatics. One characteristic of guinea pig pulmonary TB is the early lymphatic involvement. In this case, heterophils in the pulmonary lymphatic system (the guinea pig equivalents of neutrophils) were observed to be loaded with intracellular mycobacteria [119, 120]. In the murine model, this clear observation hasn't been reported yet. Nevertheless, neutrophils were also shown to be the cells transporting BCG from the site of subcutaneous inoculation to the lymph node [121]. In contrast to the "Trojan horse" mechanism in which *Mtb* is being transported intracellularly by host cells, mycobacteria could also be disseminating extracellularly. As mentioned above, mycobacteria have been reported to transverse respiratory epithelium to the basolateral side or alternatively, to de-epithelialize the respiratory mucosa by inducing cell death of the epithelium lining it. The *Mtb* mutant for HbHA was unable to bind to respiratory epithelial cells [40]. More importantly, this mycobacterial mutant was severely compromised in its ability to disseminate to the spleen after an aerosol infection. Meanwhile, the mutant *Mtb* was able to thrive in the spleens of animals infected by the IV route, ruling out any spleen-specific defect. Even though it was not directly evaluated

by the authors, it would be highly unlikely that host cells could not phagocytose the HbHA-mutant *Mtb* taking in consideration the redundant mechanisms involved in *Mtb* phagocytosis. Instead, failure to adhere to the respiratory epithelial cells and to disseminate could be a better indication that *Mtb* significantly disseminates extracellularly.

1.5.2 Acquired immune response.

Multiple lines of evidence have robustly confirmed that T cells are responsible for protection against *Mtb*. Adoptive transfer experiments initially pointed out the role of $CD4^+$ T cells as being the primary cells against *Mtb* infection [122-124]. In contrast, adoptive transfer of immune serum also ruled out a major role for the antibody response during TB [123]. As described above, the high susceptibility to TB in AIDS patients with low $CD4^+$ T cells, plainly confirmed their role in TB [3]. In murine models, depletion of $CD4^+$ T cells by injection of anti-$CD4^+$ antibodies has additionally substantiated this [125]. Furthermore, non-human primates infected with simian immune deficiency virus (SIV), are also highly susceptible to *Mtb* when a significant decrease in $CD4^+$ T cells has occurred [126, 127].

$CD4^+$ T cells are the archetype of "helper" cells. Not only do they "help" B cells during the generation of antibodies, but most importantly, they "help" macrophages control intracellular microorganisms like *Mtb* [100]. Activated $CD4^+$ T cells release cytokines such as TNF-α and IFN-γ which activate macrophages. Furthermore, activated T cells display co-stimulatory surface molecules or receptors which further enhance macrophage activation after interacting with their cognate receptors present on the macrophage's surface. It's well established that patients with mutations in the IFN-γ receptor pathway are highly susceptible to *Mtb* and a similar propensity has been reported for GKO mice [9, 128, 129]. A similar increased susceptibility to *Mtb* has also been observed in patients with mutations in the IL-12 pathway [129]. This cytokine is released by macrophages and dendritic cells, and activate the production of IFN-γ in T cells. Thus, the outcome to TB is critically dependent on the crosstalk between

the IL-12 produced by macrophages, leading to IFN-γ production by T cells, in turn reciprocally activating macrophages [100].

In TB, naïve $CD4^+$ T cells are primed by antigen presenting cells (APCs) displaying on its surface a mycobacterial peptide in the context of the Major Histocompatibility Complex (MHC)-II [100]. In the MHC-II pathway, proteolytic cleavage of proteins from endosomally-located microorganisms like *Mtb*, generate peptides which are loaded onto endosomally-residing MHC-II molecules. The complex of MHC-II-mycobacterial peptide is then shuttled to the macrophage's surface. If sufficient affinity is present between the MHC-II-mycobacterial peptide and a cognate T cell receptor (TCR) being displayed on the T cell's surface, T cell activation ensues. Besides the TCR, additional crosslinking between T cell molecules and cognate receptors on the APC's surface is important for efficient T cell activation. In this context, the interaction between the T cell's $CD4^+$ and the macrophage's MHC-II is critical, hence the importance of $CD4^+$ T cells in TB. As expected and as described above, *Mtb* residence in phagosomes as opposed to lysosomes, minimizes *Mtb* antigen processing and presentation. Furthermore, it has been described that *Mtb* lipoproteins reduce the expression of MHC-II on the APC's surface by reducing the class II transactivator [65, 66, 130]. T cell-derived IFN-γ counteracts and circumvents these immune evasion mechanisms by increasing phagosome-lysosome fusion and enhancing antigen presentation [98]. Recently, it was also described that T cell derived IFN-γ increased the macrophage's mycobactericidal activity by inducing the anti-microbial peptide cathelicidin [64, 131, 132]. More importantly, this pathway was dependent on the levels of VitaminD [133]. Lower levels of VitaminD are present in African-Americans and this correlates with an increase risk to TB. *In vitro* addition of VitaminD was able to correct the reduced mycobactericidal activity of macrophages obtained from African-Americans, thus it has been suggested that VitaminD should be used during anti-TB treatment.

Besides "helping" with macrophage activation via cytokines and cell surface receptors, a direct mycobactericidal effect has been described in $CD4^+$ T cells. In the skin of leprosy patients, $CD4^+$ T cells were shown to contain granulysin. Granulysin is an anti-microbial peptide that interacts with the

mycobacterial cell wall and by inducing permeability alterations, ultimately compromise the mycobacterial's integrity. Previously an anti-mycobacterial role had been proposed for granulysin from cytotoxic T cells (usually $CD8^+$ or $CD4^-CD8^-$ double negative cells) [134]. In contrast, granulysin was not only shown to be present in the granules of $CD4^+$ T cells in leprosy lesions, but most importantly its release from degranulating $CD4^+$ T cells had a mycobactericidal effect. The role of granulysin-containing $CD4^+$ T cells in TB has not been reported.

Detailed kinetic studies evaluating the acquired immune system in TB have concluded that priming of naïve $CD4^+$ T cells occur in the thoracic lymph nodes but not in the lung [114, 135]. Sufficient *Mtb* antigens are required for efficient T cell priming, but unfortunately this doesn't occur until 10-14 days after infection. The reason for this delay in T cell priming is unknown but could be attributed to the slower mycobacterial replication rate, in addition to the immune evasion mechanisms described above. Indeed, experiments in which animals have been infected with higher numbers of *Mtb* have detected earlier *Mtb*-specific T cell responses in the thoracic lymph nodes. This window of time between infection and T cell priming certainly benefits *Mtb* as unrestricted bacterial growth occurs for the first 20 days after infection. In turn, this could also limit the efficacy of anti-mycobacterial vaccines aimed at controlling infection.

Different subsets of $CD4^+$ T cells have been described to participate in the adaptive immune response to *Mtb* [136]. Besides the IFN-γ-producing TH1 $CD4^+$ T cells described above, TH17 $CD4^+$ T cells [137] and CD4+CD25+FoxP3+ T regulatory cells [138] have also been shown to have a role in TB. TH17 cells are characterized by the production of IL-17. This cytokine was shown to be important for efficient activation of the IL-12-IFN-γ axis. Mice with mutations in the IL-17 receptor were susceptible to *Mtb* and had increased bacterial burden [137]. As occurs with TH1 $CD4^+$ T cells, both effector and memory cells have been described for TH17 cells. In general terms, effector cells principally localize to the affected organ and could be described as the cells doing the "dirty work". These cells are activating the macrophages by both cell-contact and cytokine production [139]. In contrast, memory cells are

principally localized in the secondary lymphoid organs. As implied by their name, these cells have originated from a previous or ongoing infection and have the potential to generate effector cells when re-encountering their cognate antigen. Even though the distinct function and importance of effector and memory cells has been characterized in viral infections, this has not been the case in TB. Both cell subsets have been identified in acute and chronic infection [139]. Furthermore, BCG vaccination only leads to a temporal increase in memory $CD4^+$ T cells [139]. In fact, it has been hypothesized that BCG fails to protect against pulmonary TB because it's unable to induce sufficient memory T cells [140]. Further adoptive transfer studies should help understand the role of memory and effector cells in TB.

Several lines of evidence suggest that CD8 T cells also play an important role in the immune response against *Mtb*. Initially, it was observed that adoptive transfer of these cells could give some degree of protection to naïve animals infected with *Mtb* [141]. Conversely, it was also reported that antibody-mediated depletion of CD8 T cells enhanced susceptibility to *Mtb* [142]. This line of research got a major boost when it was reported that β2 microglobulin (β2M)-deficient mice were highly susceptible to *Mtb* infection [143]. β2M is a constituent of the MHCI that is required during CD8 T cell ontogeny [143]. Therefore, β2M-deficient mice do not express MCHI molecules and lack CD8 T cells [143]. However, it was later noted that β2M not only associated with MHCI molecules but also with CD1 (see below), transferrin receptor (see below) and other non-classical MHC molecules. Thus, other immunological defects besides the absence of CD8 T cells could explain the enhanced susceptibility of β2M-deficient mice. In order to accurately evaluate the role of CD8 T cells in the response against *Mtb*, mice with mutations in TAP-1, CD8α and perforin were subsequently evaluated [144, 145]. TAP-1 transports cytoplasm-derived peptides into the rough endoplasmic reticulum, where the peptides are loaded onto MHCI molecules [144]. Similar to the β2M, TAP-1 knockout mice are also deficient in CD8 T cells and are highly susceptible to *Mtb* infection [144, 145]. Perforin, on the other hand, is required for CD8 T cell cytolytic activity. Importantly, perforin deficient mice did not have any significant alterations in controlling *Mtb* infection [145]. This result suggested that the role of CD8 T cells in *Mtb* infection is

not dependent on its cytolytic function but could be attributed to other properties such as cytokine production [145]. One major caveat to these experiments was the fact that mice had been infected via the IV route with a high bacterial burden. Using the more physiological aerosol route, North reevaluated the role of CD8 T cells after infecting with a low dose of bacilli [146]. In this case, β2M-deficient mice were not as susceptible to *Mtb* infection; however, bacterial burden was 10 times higher than WT controls. Furthermore, depletion of CD8 T cells in animals already lacking CD4 T cells, worsen the outcome [146]. Altogether, these results suggest that CD8 T cells have a protective role against *Mtb*, albeit significantly less than CD4 T cells. In humans it has become clear that CD8 T cells are being primed against *Mtb* antigens [147, 148]. *In vitro* studies using CD8 T cell clones obtained from human patients have resulted in IFN-γ secretion when stimulated with their cognate *Mtb* antigens. Moreover, CD8 T cells have been shown to participate in the granulomatous response during TB and they localize in small clusters in the periphery of granulomas from infected animals [149]. Finally, it is noteworthy to mention that targeting CD8 T cells is being considered as way to increase the efficacy of anti-*Mtb* vaccines [150].

In addition to conventional CD4 and CD8 T cells, recent evidence has suggested that other subsets of non-classically activated T cells can also participate in the adaptive immune system against *Mtb* [151]. In humans, human leukocyte antigen (HLA)-E restricted $CD8^+$ T cells have been characterized [152]. In general, these T cells react with self-signal peptides derived from MHC molecules that are being presented in the context of HLA-E. HLA-E restricted T cells have been proposed to represent an immune surveillance mechanism to detect down regulation of MHC presentation occurring during some viral infections and cancer. Specifically in TB, several mycobacterial antigens have been confirmed to be presented via this specialized pathway [153]. The significance and mechanism of HLA-E restricted T cells is not completely known.

In addition to HLA-E-restricted T cells, CD1-restricted T cells have also been identified predominantly in the murine model. The CD1 pathway was discovered to represent a specialized antigen presentation pathway in which the presented antigens were lipids as opposed to proteins [154]. Taking in

consideration that *Mtb* is endowed with multiple immunomodulatory lipids, understanding the CD1 restricted pathway was a major endeavor [155]. In TB, initial experimental evidence for CD1 restricted T cells was suggested in KO models deficient for β2M [144]. This protein paired with CD1 for efficient antigen presentation and surface localization. Unfortunately β2M is also required for MHC-I restricted $CD8^+$ T cell response, in addition to the transferrin receptor pathway. In the later situation, deficiency in the transferrin receptor led to iron-overload and increased susceptibility to *Mtb*. Indeed, correction of the iron-overload present in β2M KO mice was able to almost normalize susceptibility to *Mtb* [156]. Final proof that CD1-restricted T cells were not critical in the immune response against *Mtb* was achieved when CD1 KO mice were finally obtained [144]. It should be pointed out though, that mice only have one isoform of CD1 (CD1d) whereas the situation in humans is more complex due to the presence of four isoforms. Therefore, in humans CD1-restricted T cells could still play an important part in the immune response against *Mtb*.

The severe susceptibility to *Mtb* in AIDS patients or in those having mutations in the IL-12 and IFN-γ axis has highlighted the importance of the adaptive immune response in TB [100]. Despite this, it has also become clear that the TB is a chronic inflammatory disease and the adaptive immune response could negatively be participating in the disease's pathogenesis. The adaptive immune system could be inducing severe inflammation and lung damage. Initial studies suggested in rabbits developing necrotic granulomas, reported that necrosis and tissue damage became evident once the adaptive immune response became operational [157]. Thus, the inflammation generated by the adaptive immune response could be affecting the lung's integrity and function. Supporting this concept that the adaptive immune system in TB causes lung damage, AIDS patients in which the $CD4^+$ -mediated adaptive immune system is non-functional, tend to have inconspicuous organ damage despite dramatic bacterial burdens [158]. Recently, it was provocatively hypothesized that the adaptive immune response could also be negatively impacting transmission [159]. After evaluating hundreds of *Mtb* T cell epitopes in humans, it became evident that these epitopes were more conserved amongst different *Mtb* isolates, than essential proteins. Intuitively,

the sequence of essential mycobacterial genes should be the most conserved in order to guarantee the protein's function. The majority of mutations affecting essential genes should therefore be synonymous in nature. That is, the nucleotide mutation should not lead to a change in the amino acid sequence. Even though this was corroborated for essential genes, the surprising finding was the T cell epitopes had fewer mutations than essential genes and furthermore, mutations in T cell epitopes were even more synonymous than in essential genes. A well-known immune evasion mechanism is antigen variation. The adaptive immune system cannot mount an effective response to antigen with a different sequence than the initial antigen that drove B or T cell priming. In contrast, these results suggest that *Mtb* has been under a selective pressure to conserve T cell epitopes. Inducing a strong T cell activation could lead to significant tissue inflammation and damage, facilitating transmission. The specific subset of $CD4^+$ T cells causing tissue damage has not been completely defined. One possibility is TH17 cells. Even though these cells have been shown to be important in the early protection against *Mtb*, they can also induce potent inflammation through IL-17 secretion and neutrophil recruitment. In fact, instead of increasing protection against *Mtb*, repetitive BCG vaccination severely compromised the lung of *Mtb*-infected mice by overdriving TH17 cells [160]. In essence, there's a delicate balance between what an adequate immune response to *Mtb* is and what causes tissue damage.

1.6 *Mtb* virulence determinants.

In the following section, specific *Mtb* virulence determinants will be covered in depth. For each virulence determinant, the identification, characterization and role in pathogenesis will be discussed. The rationale for selecting these virulence determinants is based on its relationship with the pathogenesis mechanisms evaluated in this dissertation. As a secreted RNA was identified in the *Mtb* CF, specialized secretion systems such as the Esx-1, SecA2 and Tat are reviewed. Taking in consideration α-L-polyGln's structure and putative biosynthetic pathway, the following products were addressed: glucan as an example

of an *Mtb* polymer, as well as PDIM and mycobactin as examples of *Mtb* products synthesized by the non-conventional non-ribosomal peptide and polyketide synthatase, respectively.

1.6.1 Protein secretion mechanisms.

1.6.1.1 Esx-1

Esx-1 is probably the most intensely studied, *Mtb* protein secretion system. Identification of this secretion machinery started with genomic studies aimed at determining the mechanism by which BCG became attenuated [161]. Subtractive genomic hybridization between BCG and *M. bovis* or *Mtb*, initially identified three regions of difference (RD) that were missing in BCG. The RD1 region was particularly interesting as it seemed to be missing in all BCG strains, suggesting that its deletion was a major attenuation mechanism. Subsequent genomic studies with DNA microarrays determined that BCG actually had 13 RD with *M. bovis*, however it was confirmed that deletion of RD1 was the first event in BCG attenuation as it was absent from all BCG strains [162]. *In silico* analysis suggested that the RD1 could be associated with virulence, as it was conserved in all virulent *Mycobacterium* including *M. leprae*, *M. bovis*, and *M. marinum* [163]. However, it should be noted that the RD1 was also detected in the avirulent *M. smeg*, therefore its presence is not sufficient for virulence. To confirm that the RD1 mutation led to attenuation this locus was complemented back into BCG and evaluated *in vivo* in the murine model of TB [164]. Remarkably, RD1-complemented BCG was more virulent as determined by increased number of granulomas and bacterial burden. In contrast, complementing with other RD did not have the same effect. Finally, it was shown that mutating RD1 in *Mtb* caused a significant *in vivo* attenuation that resembled the one observed for BCG [165]. Altogether, these results suggested that BCG attenuation was mostly attributed to the loss of the RD1.

Several lines of evidence suggested that the RD1 encoded for a specialized protein secretion system. Similar to the requirements for Sec-mediated protein transport, the RD1 encoded for some proteins with either multiple membrane spanning domains or potential ATPase activity [163].

Furthermore, proteomics analysis in BCG demonstrated that complementing the RD1 restored the secreted protein profile observed in *M. bovis* and *Mtb* [161]. ESAT-6 and CFP10, which are encoded in the RD1, were amongst the identified proteins accumulating in the CF after complementation of RD1 [164]. Subsequently, it was shown that restoring protein secretion required complementation with the entire region [166]. Conversely, mutating any individual RD1 gene would abrogate ESAT-6 and CFP10 secretion [167]. Extensive protein-protein interaction was reported for ESAT-6 and CFP10, as well as for other proteins encoded in the RD1 and this could explain why mutating one gene could have such a generalized effect during protein secretion. During analysis of the crystal structure for the ESAT-6/CFP10 heterodimer, it was noted that the C-terminal region in CFP10 did not participate in their interaction [168]. Instead, this region protruded from the protein and it was suggested that it could represent a signal peptide-like motif that would facilitate secretion. Indeed, it was observed that mutating this sequence impacted secretion of both proteins. Furthermore, this motif could mediate secretion of a fungal protein, clearly confirming its role as a signal peptide. Additional Esx-1-like systems have been identified in *Mtb* and other Gram positive bacteria such as *Listeria monocytogenes*, *Bacillus anthracis* (*B. anthracis*) and *Staphylococcus aureus [163]*. To keep in line with the nomenclature of other bacterial protein secretion systems, the Esx-like system has been named type VII [169].

In regards to the host-pathogen interaction and the critical role that Esx-1 plays in virulence, multiple functions have been reported for this secretion system. However, an underlying theme is that ESAT-6 has a significant ability to make pores in host membranes. Indeed, ESAT-6 but not CFP10 has been shown to interact with liposomes [170]. Furthermore, an Esx-1 dependent pore was observed by electron microscopy in host cell membranes after infection with *M. marinum* [171]. Pore formation has been associated primarily with three outcomes: host cell death, mycobacterial escape from the phagosome or secretion of bacterial effector molecules into the host cytoplasm. Esx-1 mediated host cell death has been reported both *in vitro* and *in vivo*. Initial studies reported that the Esx-1 system induced pneumocyte and macrophage cell death [172]. *Mtb* strains lacking Esx-1 could not colonize the lungs as efficiently as

WT strains due to a defect in the initial interaction with pneumocytes. *In vivo* studies have also suggested that the Esx-1 system is important for inducing the formation of necrotic granulomas [173]. In contrast to infections with WT *Mtb*, Esx-1 mutants fail to induce necrotic granulomas that normally develop in GKO mice. This could be explained with results coming out of the zebra fish model during *M. marinum* infection. It was observed that granuloma formation was Esx-1 dependent [26]. Smaller granulomas were observed in mutant *M. marinum* and this correlated with decreased bacterial burden [174]. Subsequent studies suggested that the Esx-1 secretion apparatus was inducing metalloproteinases by the host epithelium and this event led to macrophage recruitment [175]. In turn, freshly arrived macrophages could be a niche where *M. marinum* could thrive and cause macrophage death. Eventually, multiple rounds of this infectious cycle would lead to sufficient host cell death and possibly generate necrotic granulomas. Escaping from the phagosome is a second outcome induced by pore formation through an Esx-1 dependent mechanism. Indeed, initial studies with *M. marinum* observed that it would eventually localize in the host cytoplasm and spread to neighboring cells through actin polymerization [176]. Even though *Mtb* and *M. leprae* are generally not considered to reside in the host cytoplasm, some recent results suggest bacteria gain access to the cytoplasm after prolonged infection [177]. For these three *Mycobacterium* spp, translocation into the host cytoplasm required the Esx-1 secretion system. Finally, it has recently been reported that by creating pores in the host cell membrane, *Mtb* Esx-1 mediates secretion of effector molecules into the host cytoplasm [178]. Specifically, mycobacterial DNA has been suggested as one of the translocated molecules [179]. This is in line with the previous observation that *M. smeg* Esx-1 could participate in bacterial conjugation [180]. The presence of bacterial DNA inside the host cytoplasm has been shown to be a potent stimulus for the induction of type-I IFNs which seem to benefit bacterial survival [178].

In conclusion, Esx-1 is the prototype of a new protein (and maybe nucleic acid) secretion system that emerged from the *Mtb* field and opened many lines of research into bacterial pathogenesis.

1.6.1.2 SecA2 pathway.

One of the surprises encountered in 1998 when the *Mtb* genome was sequenced, was the presence of a second SecA protein with approximately 50% homology to the previously characterized and ubiquitous SecA1 [181]. At that time, it constituted the first example of a microorganism encoding for two copies of the SecA protein and the term SecA2 was coined to describe the second variant. Subsequently, it was determined that two copies of *secA* were also encoded in the genomes of most *Mycobacterium* spp, but most important, in several highly pathogenic bacteria such as *B. anthracis*, *M. leprae*, *Listeria monocytogenes*, *Staphylococcus aureus* and *Streptococcus pneumonia [182]*. The presence of SecA2 in these pathogenic bacteria suggested that it constituted a specialized secretion system involved in virulence.

The first attempt to characterize SecA2's function was performed by Braunstein *et al* in 2001 [183]. It was initially determined that SecA2 was not essential in *M. smeg* and *Mtb* whereas SecA1 was essential and could only be mutated in a merodiploid strain. Interestingly, despite the high amino acid sequence identity, SecA2 overexpression could not compensate for the loss of SecA1, suggesting that their function is non-redundant. The Δ*secA2* mutant was highly susceptible to sodium azide, routinely used to inhibit protein secretion. Thus, it was suggested that similar to SecA1, SecA2 could also participate in protein secretion. Using the alkaline phosphatase (AP) reporter system, it was shown that the Δ*secA2* mutant accumulated unprocessed AP, a good indication that protein secretion was inhibited. In order to identify specific SecA2-mediated secreted proteins, proteomics studies were subsequently performed to determine differences in the protein profile of CF obtained from WT *Mtb* or the Δ*secA2* mutant [184]. Despite an unremarkably similar protein profile between both strains, the superoxide dismutase A (SodA) was the most significant protein to be underrepresented in the CF of the Δ*secA2* mutant strain. Conversely, increased SodA levels were detected in the mutant's cell wall, suggesting that it was not being efficiently secreted. Interestingly, the SodA protein lacks a signal peptide in its N-terminus, a *sine qua non* for SecA mediated protein secretion. This was also the case for KatG and

Rv0390, also present in reduced amounts in the CF of the Δ*secA2* mutant strain. Interestingly, in *M. smeg* it was later shown that SecA2 could also mediate secretion of signal peptide-containing lipoproteins (Msmeg1712 and Msmeg1704) [185]. Secretion of these lipoproteins through the *M. smeg* SecA2 pathway seems to be the exception rather than the norm as mutating *secA2* did not affect other lipoproteins. In contrast to *M. smeg*, no defect in lipoprotein secretion has been identified in the *Mtb* Δ*secA2* mutant. However, a similar secretion defect was observed for Msmeg1704 after heterologous expression in the *Mtb* Δ*secA2* mutant. Thus, *Mtb* and *M. smeg* SecA2 seem to recognize similar features in proteins being secreted through this pathway. Finally, in regards to the protein's biochemical characterization, it has recently been suggested that SecA1 also participates in the secretion of SecA2 dependent proteins [186]. Site directed mutagenesis of the SecA2 ATPase motifs abrogated its activity and led to a dominant negative mutant that even seemed to affect the SecA1 pathway. Conversely, conditional mutants for SecA1 also affected secretion of the SecA2-dependent lipoprotein Msmeg1704. The current dogma for SecA2-mediated protein secretion is the following: proteins both with and without signal peptide are recognized by the SecA2 pathway. Once the protein is bound to SecA2, it's delivered to the SecYEG translocon and secreted with the help of SecA1.

In the context of host-pathogen interaction, it was shown that the Δ*secA2* mutant failed to thrive in BMMOs [187]. In the murine model of TB, reduced bacterial burden was obtained in the lungs and spleens of mice infected with the mutant strain [184, 187]. However, the mutant strain did persist in the chronic stage of the disease, albeit at lower levels. Furthermore, after an aerosol infection with the mutant strain, prolonged survival for both WT and SCID mice was observed [184, 187]. This result with SCID mice suggested that attenuation of the Δ*secA2* mutant strain was principally associated with a defect in counteracting the host innate immune system. Consistent with this hypothesis and in stark contrast to infecting BMMO with virulent *Mtb* strains, infection with the Δ*secA2* mutant led to higher nitric oxide and inflammatory cytokines production [187]. Similarly, the mutant strain failed to down regulate the expression of MHCII, a major immune evasion mechanism induced by virulent *Mtb* [187]. It was

assumed that the lack of SecA2-mediated SodA secretion could explain the mutant's phenotype. Surprisingly, the mutant strain was also compromised during *in vitro* infections of BMMOs obtained from *phox* mice lacking the ability to produce reactive oxygen species [187]. Thus, mutating *secA2* has a greater impact in *Mtb* physiology beyond SodA secretion. Recent studies have determined that the Δ*secA2* mutant is defective in preventing phagosome-lysosome fusion [188]. Indeed, pharmacologically blocking phagosome-lysosome fusion with bafilomycin restored the *secA2* mutant's ability to reside in BMMOs. It has been suggested that the Δ*secA2* mutant could have cell wall alterations that not only explain the activation of BMMO, but also its impotency in blocking phagosome-lysosome fusion. Interestingly, *Listeria* Δ*secA2* mutants have dramatic alterations in their cell wall as evidence by rough as opposed to the smooth colony morphology associated with virulence [189]. Specifically*, Listeria's* altered colony morphology has been attributed to SecA2-mediated secretion of autolysins that cleave peptidoglycan [190]. It remains to be determined if this is the case in *Mtb*. One interesting aspect that has emerged from studying *Mtb* SecA2 pathway is the possibility of using the Δ*secA2* mutants as a vaccine [191]. As mentioned, this mutant strain is attenuated both *in vitro* and *in vivo*, including in immunocompromised mice. Furthermore, the Δ*secA2* mutant potently stimulates the innate immune system, which is critical in activating the adaptive immune system. Significantly, when the Δ*secA2* mutant was used as a vaccine, an enhanced CD8 T cell response was observed. This was attributed to antigen cross-presentation associated with the Δ*secA2* mutant's ability to induce apoptosis. More importantly, vaccinating with the Δ*secA2* mutant significantly reduced bacterial burden and lung pathology in both infected mice and guinea pigs. In conclusion, the identification and characterization of a specialized *Mtb* protein secretion system, has led to a putative translational application such as a vaccine.

1.6.1.3 Twin-arginine transporter (Tat) pathway.

The field of horticulture discovered the Tat pathway in the mid 1990's when a SecA-independent protein secretion system was identified in the chloroplast's thylakoid membrane [192]. This alternate secretion system was not inhibited by sodium azide, known to block SecA-mediated protein transport. Instead, protein secretion by this new pathway could be abrogated with valinomycin or nigericin which dissipate the proton motif force (PMF) [193]. Taking in consideration that secretion through the SecA pathway requires signal peptides in the protein's N-terminus, it was considered that the two secretion pathways could mechanistically diverge because of significantly different signal peptides. However, signal peptides for both types of secreted proteins were remarkably similar, only differing in the conserved presence of an arginine-arginine (RR) motif in the signal peptide of proteins secreted by this new pathway. Thus it was suggested that the RR motif specifically targeted proteins being secreted by this alternate pathway. Indeed, mutagenesis experiments targeting the RR motif abrogated secretion through this new pathway [194]. Finally, while characterizing a maize mutant with defective protein transport into the chloroplast, a novel group of transporters was sequenced and shown to participate in this alternate secretion pathway [193]. Surprisingly, homology searches determined that similar proteins were encoded in the genomes of bacteria such as *E. coli*, *Mtb* and *M. leprae*. The term Tat (twin-arginine transporter) system was coined after confirming that mutating *E. coli's* homologues abrogated secretion of some periplasmic proteins [194].

While the Tat system was being identified in chloroplasts, sequencing of *Mtb* genome was being completed. Probably due to the incipient body of knowledge in the Tat system, during the initial genome annotation this system was not included as a member of the 'protein and peptide secretion' subcategory [181]. The driving force to characterize *Mtb* Tat system came from several observations linking this pathway to virulence mechanisms in bacteria such as *Legionella pneumophila*, *Pseudomonas aeruginosa* and some strains of *E. coli* [195]. Using the amino acid sequence for *E. coli's* Tat system, homologues for TatA, B and C were identified encoded in the *Mtb* genome. It is currently believed that the *Mtb* Tat

proteins have similar roles as their counterpart in *E.coli*: TatA oligomers form the translocation pore, whereas TatB and C seem to recognize and tether the target proteins. The genetic organization for *Mtb tat* genes was found to be conserved in *M. leprae*, *M. bovis*, *M. avium* and even *M. smeg* [195]. Furthermore, bioinformatics analysis suggested that at least 31 proteins could be secreted through the Tat system, including BlaC (a β-lactamase) and PlcA and B (phospholipases), previously shown to be important for *Mtb*, *Legionella* and *Pseudomonas* virulence [196]. However, this number of proteins seem to vary according to the algorithm and stringency applied during the analysis, and different studies have reported anywhere from 11 to almost 90 proteins putatively secreted through the Tat system [197-199]. Initial mutagenesis studies determined that in contrast to *Mtb*, the Tat system is not essential in *M. smeg* [195, 197, 198]. Thus, most studies have been performed in the latter microorganism. Consistent with the bioinformatics analysis suggesting that BlaC is a Tat-secreted protein, the Δ*tatC M. smeg* mutant was susceptible to β-lactams [195]. Complementation with WT BlaC was able to restore resistance to β-lactams, however mutating the RR motif or forcing BlaC secretion through the SecA pathway did not restore the WT phenotype. An additional study in *M. smeg* confirmed these results and furthermore, characterized a functional succinate transporter (DctP) that was secreted through the Tat pathway [197]. Interestingly, bioinformatics analysis performed in this and other studies, suggested that some *Mtb* lipoproteins also have Tat-compatible signal peptides containing the RR motif [198, 200]. Unfortunately, no experimental evidence has been obtained in *Mtb* or *M. smeg* confirming lipoprotein secretion through the Tat pathway.

Additional studies were performed to identify other *Mtb* proteins that could be secreted through the Tat pathway. Using stringent bioinformatics analysis 11 *Mtb*, putatively Tat-secreted proteins were identified [198]. Out of these 11 proteins, only Rv2525c is conserved in *M. leprae*. Taking in consideration the massive genomic decay present in *M. leprae*, conservation of an Rv2525c homologue is a good indication that it could have an important role in pathogenesis. Similar to the phenotype observed for the Δ*tatC M. smeg* mutant, mutating *rv2525c* also increased *Mtb* susceptibility to β-lactams and

slowed its replication rate in axenic cultures. Paradoxically, the Δ*rv2525c Mtb* mutant replicated faster in BMMO and was more virulent in the murine model of TB. The authors suggested that Rv2525c could have a role in cell wall metabolism, as a Δ*rv2525c Mtb* mutant had increased susceptibility to β-lactams. In agreement with this line of thought, structural homology with a transglycosidase was identified for Rv2525c using the Phyre secondary structure matching software [199]. In turn, the altered cell wall structure present in the Δ*rv2525c Mtb* mutant could increase the release of immunomodulatory molecules. Finally, a recent study reported that Ag85A and C are *bona fide* Tat-secreted proteins [201]. For these proteins, secretion through the Tat pathway depended on the oxidative potential in the cytoplasm. As occurs for AP, increased cytoplasmic oxidative conditions might assist in the folding of Ag85A and C. Interestingly, Ag85B does not seem to be secreted through the Tat pathway, instead being secreted through the conventional SecA pathway [202]. The reason why similar proteins would be secreted through different secretion systems is not known.

In conclusion, the *Mtb* Tat pathway is an essential, specialized protein secretion system that has a significant, yet not fully understood role in pathogenesis. Several Tat-secreted proteins like the phospholipases and Ag85A or C could have a direct immunomodulatory activity affecting the outcome of an infection. Alternatively, this dedicated protein secretion system could have an important role in cell wall remodeling, as suggested by the enzymatic activity present in Tat-secreted proteins such as Ag85A, Ag85C, Rv2525c and BlaC. The identification of additional Tat-mediated secreted proteins might enhance our understanding of this secretion system.

1.6.2 Specific *Mtb* products

1.6.2.1 PDIMs

Studies in the early 1900's identified the presence of waxy-like material that could be extracted from *Mtb* using organic solvents. Importantly, when inoculated intraperitoneally in rabbits, this material had biological activity consisting of granuloma formation with significant influx of langhan and

epitheliod cells [203]. Anderson developed a protocol to isolate four waxes (A-D) based on their differential solubility in different organic solvents [204]. It was determined that prolonged incubations with potassium hydroxide were required to saponify wax A into two optically active components [205]. The alcohol and acid components were named phthiocerol [205], and mycocerosic acid [206] respectively, in order to reflect their origin and composition. The structures of phthiocerol [207] and mycocerosic acid [208] were elucidated in the late 1950's using mass spectrometry and some interesting features were identified for both compounds. For instance, two alcohol groups were identified in the middle of the aliphatic chain of phthiocerol [207]. Mycocerosic acid was reported to contain 4 methyl groups [208], an unusual modification in fatty acids. Based on the structure of each individual component, it was suggested that both alcohol groups present in phthiocerol were esterified with a mycocerosic acid, hence the current name phthiocerol di-mycocerosate (PDIM) (Fig. 1.1).

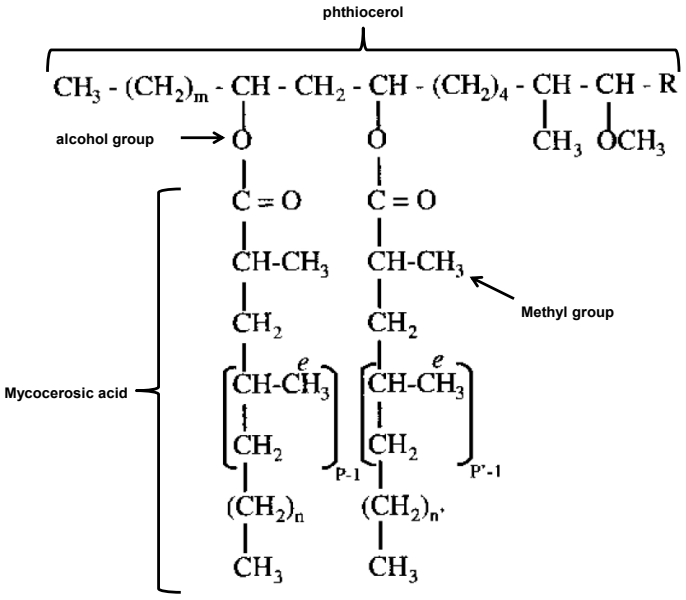


Figure 1.1 Structure of PDIM. PDIM consists of two mycocerosic acids covalently linked through an ester bond to the alcohol groups present in phthiocerol. Saponification of this ester bond yielded the

individual components present in PDIM. The presence of alcohol and methyl groups in this molecule suggested that biosynthesis was performed by a polyketide synthatase. Adapted from reference [213].

In the 1960's initial studies evaluating PDIM's biosynthesis suggested that both radiolabeled propionate and acetate could be incorporated into phthiocerol [204]. Similarly, in a cell free assay it was shown that methylmalonate could be incorporated into the mycocerosic acids [209]. Interestingly, enzymes synthesizing mycocerosic acid were identified in the cytoplasm [209] as opposed to the cell membrane where fatty acid synthesis had generally been detected. Furthermore, efficient incorporation of methylmalonate into mycocerosic acids required the presence of a fatty acid 'primer' with 18-20 carbon atoms [209]. The enzyme synthesizing mycocerosic acid was eventually purified [210], subjected to N-terminal sequencing and the resulting amino acid sequence was used to identify the *mas* (mycocerosic acid synthase) gene [211]. Additionally, it was determined that the MAS protein had a high homology to polyketide synthatases present in other organisms. Involvement of a polyketide synthatase in PDIM's biosynthetic pathway could explain the aforementioned observations regarding its metabolic labeling with radiolabeled acetate, propionate, and methylmalonate, the cytoplasmic localization for the enzymatic complex and the requirement for a fatty acid priming the reaction. Inactivation of the *mas* gene indeed confirmed that it encoded for a protein participating in both PDIM biosynthesis and phenolic glycolipid (PGL) [212]. Interestingly, during *Mtb* genome annotation, it was observed that the *mas* gene was localized in a region encoding other polyketide synthatases [181]. This genetic organization suggested that PDIM and PGL had common biosynthetic steps. Furthermore, the fact that >1% of the *Mtb* genome was devoted for PDIM/PGL biosynthesis suggested that these products had an important role in pathogenicity. Then in 1999, a transposon mutagenesis study identified three PDIM-deficient mutants in which the transposons had integrated into this genomic region [213]. From these studies, it was suggested that the *mmpL7* gene encoded for a PDIM transporter as this product accumulated in the *Mtb* cytoplasm but was not exported extracellularly into the CF. An interesting aspect about the genetic organization of

polyketide synthatases, was the presence of genes encoding for FadD proteins with high similarity to fatty acid synthases [181]. It was shown that these proteins activate fatty acids via an acyl-adenylate intermediate and are subsequently transferred to the polyketide synthatase [214]. The stage was ready to complete the description of PDIM biosynthesis. Briefly, mycocerosic acid is synthesized by the concerted action of FadD28 and Mas [215, 216]. The former protein activates C14-C16 fatty acids and transfers them to Mas. Mas catalyzes the incorporation of methylmalonate into the fatty acid [209, 215], generating this methyl-branched polyketide. Phthiocerol biosynthesis also starts with an equivalent fatty acid activation by FadD26 [216], which is then transferred to the Pps proteins. These proteins mediate the incorporation of malonate or methylmalonate [215]. In contrast to other fatty acid synthases, the modular organization of PpsA and PpsB is unique in that it lacks an enoylreductase and dehydratase domain, but conserves the ketoreductase domain. Thus, during the incorporation of malonate, the β-keto group is only reduced to a hydroxyl group and the subsequent generation of an alkyl chain is not completed [181, 215]. These phthiocerol hydroxyl groups are subsequently esterified with mycocerosic acid, a step catalyzed by PapA5 [217].

PDIM's critical role in virulence was identified early on by Goren [218], when they isolated an *Mtb* strain that despite producing large quantities of SL, was avirulent. Animals infected with this strain had reduced numbers of pulmonary granulomas and less bacterial burden. However, mutant bacteria were not cleared out from the tissues but persisted at lower numbers. Similar results have been reported for all the engineered PDIM mutants [213, 219]. An important observation made by Domenech was the spontaneous loss of PDIM during *in vitro* culturing of *Mtb* and this occurred both in laboratory strains and clinical isolates [220]. The authors recommended that before legitimately claiming that other *Mtb* products are important virulence determinants, the absence of PDIM should be ruled out. Despite multiple studies highlighting the importance of PDIM as a major virulence determinant, the mechanism is still unclear. Cell biology studies using purified PDIM have suggested that the product is rather inert in inducing cytokine production by both BMMOs and dendritic cells [219]. However, *in vitro* infections

with PDIM mutants lead to increased inflammatory cytokine production. Additionally, mutant strains fail to thrive in activated macrophages and this has been attributed to an enhanced susceptibility to RNis [219]. Recently, it was reported that PDIM could participate during the initial *Mtb*-macrophage interaction [221]. It was observed that PDIM could alter the macrophage's membrane fluidity, including the fluidity of cholesterol-rich lipid rafts which are important during phagocytosis. A PDIM-less mutant was phagocytosed through a CR3 and MR-independent pathway and was not able to efficiently block phagosome-lysosome fusion. Furthermore, as determined by the altered hydrophobicity, it has been suggested that PDIM could have an important role in the structure of the *Mtb* cell wall [222]. Further studies are required to completely elucidate PDIM's role in TB pathogenesis, however its importance has opened several lines of research into the pharmacological inhibition of *Mtb* polyketide synthatases.

1.6.2.2 Capsule and Glucan/glycogen.

The presence of an electron transparent zone (ETZ) surrounding mycobacteria had been reported since the 1950s [223-225] when electron microscopy was applied to study mycobacteria obtained from both axenic cultures or from infected tissues. Even though it was suggested that this ETZ could be attributed to a mycobacterial 'capsule', this hypothesis created significant controversy as the ETZ could simply represent an artifact during sample fixation for electron microscopy [225]. Furthermore, the ETZ could represent a host derived product as opposed to a mycobacterial product. Subsequently, several lines of evidence suggested that the ETZ was actually a microbial product as it could only be observed surrounding live bacteria but not dead bacteria [226, 227]. In addition, microscopic characterization of the 'capsule' identified that it was composed of fibrils not present in host tissues and mycobacterial waxes were obtained during preliminary biochemical characterization of the fibrils [226]. In the 1980s, Rastogi determined that the mycobacterial ETZ/capsule could easily be stained with ruthenium red and it was present in both virulent and avirulent *Mycobacterium* spp [228]. Interestingly, the 'capsules' width was variable and ranged from 5-100nm [227, 228]. The capsule's biochemical characterization was finally

performed by Daffé after purifying it from the surface of live *Mtb* and *M. bovis*, after gentle mechanical disruption [229]. The capsule was reported to account for approximately 2-3% of dried bacteria and be constituted predominantly of polysaccharides and protein (94-99%), as opposed to lipids (1-6%). Further characterization of the sugars present in the polysaccharides identified arabinose and mannan (from arabinomannan), but predominantly glucose residues forming a polymer which shared several of glycogen's chemical and biophysical properties. Similar to glycogen, mycobacterial capsules were shown to be repeating units of glucose residues in 1→4 linkage with short 1→6 branching occurring at the glucose residue in position 4 (Fig. 1.2) [54]. Glucan was not only present on the surface of mycobacteria, but it was also present in the CF suggesting that constant capsule remodeling was taking place [230]. Finally, a second glucose-containing polymer was found to be present inside mycobacteria and this polymer actually resembled glycogen more than the capsular glucan [231].

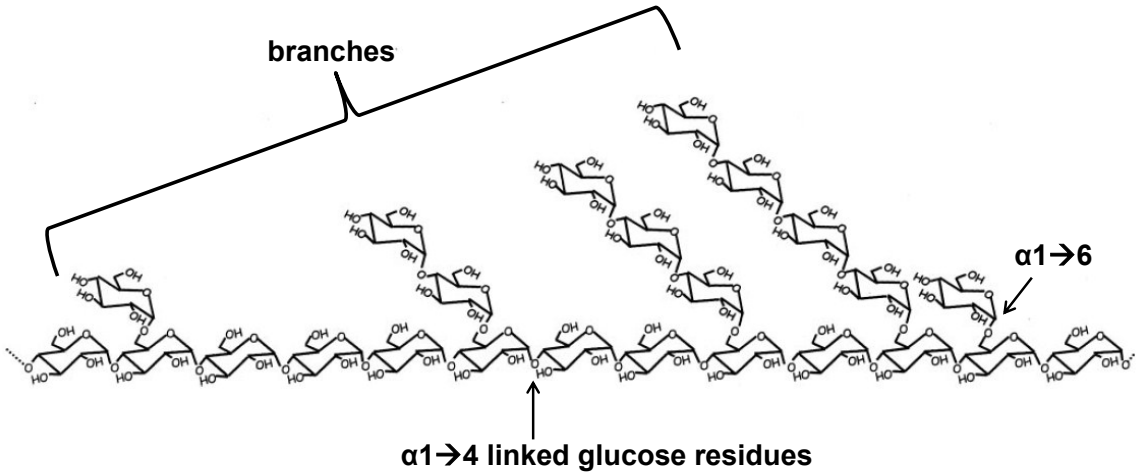


Figure 1.2 Structure of *M. bovis* capsular glucan. *Mycobacterium* glucan consists of a polymer of glucose residues in a α1→4 linkage. Some glucose residues are also derivatized at position 6 with short oligomers of α1→4 linked glucose residues, constituting the branches present in glucan. Adapted from reference [54].

Multiple experiments have been performed at the cell biology level to understand glucan's role in the host-pathogen interaction. Serendipitously, glucan was purified as the molecule responsible for BCG's anti-neoplastic activity [232]. It had been known for a while that BCG inoculation into different types of cancer induced tumor regression. In fact, BCG is still being used therapeutically against bladder cancer. As occasionally BCG has important side effects, it was decided that purification of the molecule(s) responsible for this activity could circumvent using a live bacteria. Glucan was purified from BCG extracts obtained after boiling and shown to mediate the anti-neoplastic effect. The mechanism involved in mediating this effect was not determined. However, subsequent studies with purified glucan have suggested that it has a detrimental effect in the activation of the immune response. Specifically, glucan has been shown to affect the differentiation of monocytes into dendritic cells [233]. In the presence of glucan, monocyte-derived dendritic cells have reduced levels of the co-stimulatory molecule CD80. Furthermore, these cells are devoid of CD1 and do not activate CD1 restricted T cells. Additional studies showed that glucan activated the dendritic cell receptor DC-SIGN which is known as being targeted by mycobacteria in order to affect antigen processing and presentation [234]. The immunosuppressive cytokine IL-10 was induced in dendritic cells stimulated with glucan. Alternatively, glucan's role in the host-pathogen interaction has been evaluated by mechanically removing the capsule from live mycobacteria and comparing it to encapsulated bacteria. Initial experiments evaluated this interaction in the context of CR3. When the non-phagocytic CHO cells were transfected in order to express CR3, encapsulated as opposed to non-encapsulated bacteria were uptaken [235]. The interaction between transfected CHO cells and encapsulated *Mtb* could be blocked with either purified glucan or commercial glycogen, further confirming the interaction's specificity. However, using macrophages as opposed to transfected non-phagocytic CHO cells suggested that the glucan capsule was actually inhibiting *Mtb* phagocytosis [236]. Capsule removal via sonication, increased the number of bacteria being uptaken by macrophages through a CR3-independent mechanism. It was suggested that this enhanced uptake was attributed to negative charges being exposed after glucan removal. The authors suggested that other

receptors besides CR3 could be mediating faster *Mtb* uptake, but the capsule was actually slowing down the process by targeting CR3 which is known not induce a respiratory burst.

In order to extrapolate the cell biology results to an *in vivo* animal model, *Mtb* mutants were engineered so as to disrupt genes putatively involved in glucan and/or glycogen biosynthesis [237]. As *Mtb* glucan and glycogen have a similar structure, mutations targeted *Mtb* genes with homology to *E.coli's glg* involved in glycogen biosynthesis. In *E.coli*, glycogen biosynthesis starts with the *glgC*-mediated glucose activation via an acyl phosphate. The main chain consisting of glucose in 1→4 linkage is performed by the polymerizing activity of *glgA* and the 1→6 linkage present in the branches by *glgB*. Homologues for these three genes were indeed identified in *Mtb* genome and the following observations were made in the mutagenesis experiment. At least two distinct biochemical pathways seem to be involved in glucan's biosynthesis as compensation for a *glgA* mutation could be compensated by *rv3032c*, an additional glycogen synthase encoded in *Mtb* genome. However, double mutations were not successful suggesting that glycogen/glucan production is essential. Interestingly, the branching enzyme *glgB* (*rv1326c*) was the only essential gene and it was hypothesized that a toxic polymer accumulated in its absence. As *glgC* involved in glucose activation was not shown to be essential, it was suggested than an additional redundant pathway could be present. Finally, and in contrast to the cell biology results highlighting glucan's importance in the host-pathogen interaction, glucan's absence did not affect the outcome of macrophage infections with mutant strains. *In vivo* studies, however, suggested that glucan could play a role in the persisting stage of chronic infections as reduced bacterial loads were observed for these mutants. As mentioned, the fact that *glgC* was not essential for glucan biosynthesis suggested that an additional pathway leading to glucose activation was present. Indeed, it was shown that a completely different pathway involved in glucan biosynthesis was present in mycobacteria and is dependent on the gene *glgE* (*rv1327c*) [238]. This enzyme was shown to transfer an activated glucose residue from maltose 1-phosphate (maltose is a glucose disaccharide consisting of 1→4 linkage). The gene *pep2* (*rv0127*) would phosphorylate and activate maltose's reducing glucose, thus explaining why *glgC* was not essential

for the activation step. As occurred with the *glcB* mutant, *glgE* mutation also essential. Both genes seem to be in an operon and it was postulated that the reason for their essentiality was the accumulation of a toxic product. Thus, inhibiting GlgE has become an exciting drug target.

Altogether, an initial observation of a capsule-like material surrounding mycobacteria, led to the characterization of two glucose polymers present in this bacterial genus: capsular glucan and intracellular glycogen. Redundant pathways leading to glucan/glycogen biosynthesis were identified and in combination, shown to be essential. Despite all the cell biology studies suggesting that glucan played an important role in the host-pathogen relationship, it is currently hypothesized that glucan/glycogen could have a more important role either as a glucose depot and/or a trehalose sink.

1.6.2.3 Mycobactin and Iron.

Iron is an essential micronutrient required by all microorganisms. However, ferric iron is highly insoluble and microbial acquisition requires dedicated molecules collectively called siderophores, which essentially 'solubilize' iron [239]. For pathogenic bacteria the situation is even more precarious as free iron inside the host is even scarcer, instead being associated with the host proteins lactoferrin, transferrin and ferritin. In addition, the acute phase response occurring in inflammation significantly lowers iron concentration further more [239]. Thus, it has been postulated that pathogen iron acquisition could be an alternative drug target. Conversely, it has been shown that iron overload in humans, increases the risk factor for *Mtb* morbidity and mortality [240].

In the early 1900's, a simple yet transforming observation initiated the field of study in *Mtb* iron procurement: *in vitro* growth of the until-then non-culturable *M. johnei*, was achieved by supplementing egg media with *Mtb* extracts (referenced in [241]). Using extracts from *M. phlei*, the molecule responsible for this effect was identified by Snow *et al* in 1953 and the name "mycobactin" was coined [241]. Mycobactin was obtained after subjecting *M. phlei* to organic extractions followed by adsorption to alumina. Initial biochemical analysis suggested that mycobactin's structure contained several rings and it

was shown to bind other metals such as copper. The structures of *M. phlei* [242] and *Mtb* mycobactin [243] were elucidated by Snow in 1965 and the presence of hydroxamic acids previously known to coordinate iron binding, suggested that mycobactin was an iron siderophore . During mycobactin's biochemical analysis it was observed that alkali treatment cleaved the molecule into two fragments: cobactin and mycobactic acid [243]. In turn, cobactin and mycobactic acid could be broken down to its basic constituents via acid hydrolysis. Mycobactic acid was shown to be composed of salicylic acid, a serine (Ser) residue, 2-amino-6-hydroxyaminohexanoic acid and fatty acids of different length. Instead, 2-amino-6-hydroxyaminohexanoic acid and butyrate were shown to be cobactin's constituents. Lysine residues were proposed to be modified to generate the 2-amino-6-hydroxyaminohexanoic acids containing the hydroxamic acids that mediate iron binding. From the N-terminus, mycobactin's unusual structure consists of salicylic acid forming an amide bond with a cyclized Ser or Thr residue (Fig. 1.3).

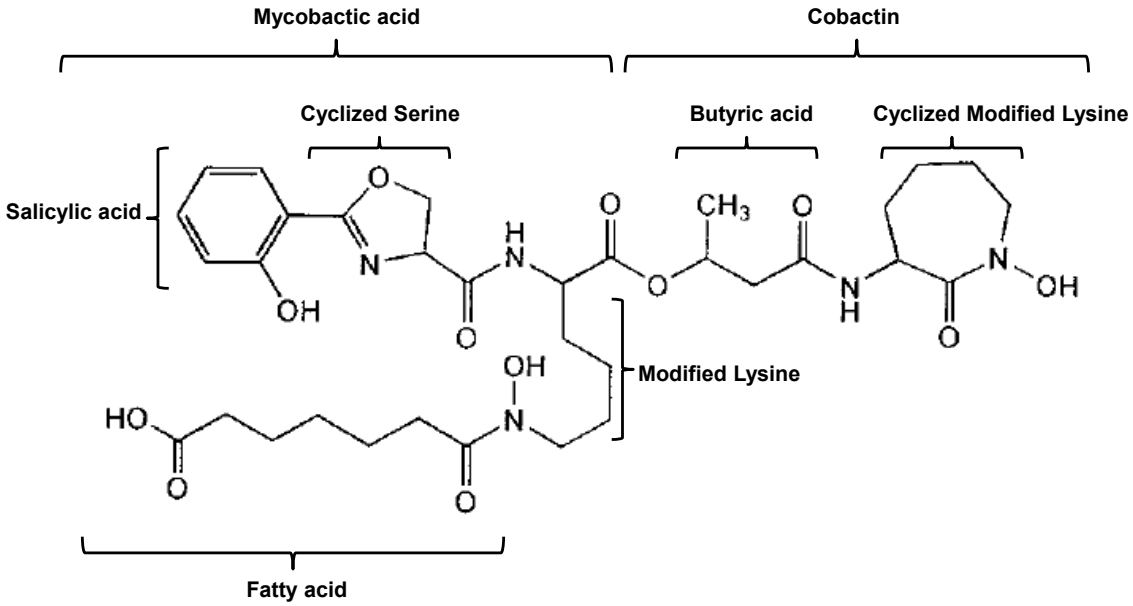


Figure 1.3 Structure of mycobactin. The structure of mycobactin was deduced from initial studies using alkali and acid hydrolysis. Alkali hydrolysis cleaved mycobactin at the ester bond between butyric acid and the modified lysine residue, yielding two fragments: mycobactic acid and cobactin. In turn, acid

hydrolysis cleaved mycobactic acid into its constituents: salicylic acid, a cyclized form of Ser, a modified lysine residue and a fatty acid that could vary in length. Cobactin was hydrolyzed in acidic conditions into butyric acid and a cyclized form of a modified lysine residue. The modified lysine residues provide iron-binding properties. Adapted from reference [238].

This oxazoline ring is also covalently attached via an amide bond to a modified lysine residue. At the ε-amino group, the lysine residue is substituted with both a hydroxyl group and a fatty acid, generating acylated 2-amino-6-hydroxyaminohexanoic acid. Both modifications generate one of the hydroxamic acids participating in iron binding. The modified lysine (acylated 2-amino-6-hydroxyaminohexanoic acid) is attached to keto-butyrate via an ester bond, which is the target for the aforementioned alkali-mediated mycobactin fragmentation into mycobactic acid and cobactin. Finally, an amide bond covalently links the keto-butyrate to the second modified lysine residue (2-amino-6-hydroxyaminohexanoic acid) which is also cyclized and provides the second hydroxamic acid. As mentioned, one of the initial findings during mycobactin characterization was the presence of fatty acid of varying lengths attached to the lysine residue [241]. It was later shown that long fatty acids are present in mycobactins, whereas shorter fatty acids ending with either a carboxy or a methyl ester are present in carboxy-mycobatins [244]. This structural difference accounts for the fact that mycobactins are only soluble in organic solvents whereas carboxy-mycobactins are soluble in both organics and aqueous conditions. As will be discussed below, this structural difference correlates with significantly different physiological roles during iron acquisition.

Once mycobactin and carboxy-mycobactin's structure had been determined, the field's focus was shifted to determine their biosynthesis and function. In agreement with the chemical groups characterized in mycobactin, initial metabolic labeling experiments determined that both radioactive lysine and acetate could be incorporated into the molecule at the 2-amino-6-hydroxyaminohexanoic acid and keto-butyrate, respectively [245]. This unusual molecule consisting of an aryl group, amino acids and fatty acids

suggested that mycobactin could be synthesized by a mixed poliketide/non-ribosomal synthetase. In fact, a mixed poliketide/non-ribosomal synthetase had been shown to participate in the synthesis of yersiniobactin, another siderophore highly resembling mycobactin (referenced in [246]). Thus, in 1998 Walsh *et al* performed bioinformatics analysis to characterize a putative mixed poliketide/non-ribosomal synthetase involved in mycobactin's synthesis [246]. The criteria applied to the bioinformatics analysis consisted of finding genes that contained several A domains involved in amino acid activation, a poliketide-like module synthesizing and ligating keto-butyrate, but most importantly, the first module had to be able to activate and carry or ligate an aryl group such as salicylate. Indeed, a 24 kb 10-gene cluster named *mbt*, was shown to encode proteins putatively having these enzymatic activities. The first biochemical characterization of proteins involved in mycobactin biosynthesis was also performed, and MbtA and MbtB were shown to activate and carry salicylate, respectively. Subsequently, mutagenesis studies targeting *mbtB* effectively abrogated mycobactin's production, further substantiating this locus' role in mycobactin synthesis [247]. However, it was noticed that the *mbt* locus did not encode proteins required for lysine hydroxylation and acylation [248]. Thus, transcriptional studies were performed to identify other genes induced in iron-limiting conditions, a stimulus known to induce genes in the *mbt* locus. This strategy successfully lead to the identification of a second locus named *mbtII*, which was biochemically characterized and confirmed to modify lysine's ε-amino group via hydroxylation and acylation. Finally, MbtE and MbtF were recently confirmed to participate in mycobactin synthesis by activating and inserting the modified lysine residues [249]. Despite some minor controversies, the pathway leading to mycobactin biosynthesis is pretty much defined.

As mentioned above, identification of hydroxamic acids in mycobactin's structure as well as its ability to chelate metals, suggested that mycobactin could be an iron siderophore. This was supported by the fact that iron deprivation led to increased mycobacterial mycobactin production [250]. Pioneering work by Ratledge *et al* in the 1970's and Horwitz in the 1990's, significantly contributed to understanding the mechanism whereby mycobactin obtains and transports iron into mycobacteria. As implied by its

solubility in organic solvents, Ratledge determined that mycobactin was exclusively present in the mycobacterial cell wall but not in the cytoplasm or secreted in the media [251]. Furthermore, detailed kinetic analysis showed that after pulsing mycobacteria with radiolabeled iron, it initially bound to and was extracted with mycobactin, however, at later time points this was not the case. Thus, it was hypothesized that mycobactin's role was to transport iron across the hydrophobic mycobacterial cell wall. A reductase dependent mechanism would then unload iron from mycobactin to finalize iron incorporation. Taking in consideration mycobactin's insolubility in aqueous buffers, it was unclear how insoluble iron or iron ligated to host proteins could gain access to mycobactin embedded in the mycobacterial cell wall. While evaluating this discrepancy, Ratledge identified a group of water soluble mycobacterial siderophores that could solubilize precipitated iron and transfer it to mycobactin [252]. These mycobacterial water soluble molecules were collectively named exochelins. However, subsequent work has determined that exochelins obtained from avirulent vs. virulent mycobacteria have substantially different structures [253] and it has been proposed that the name exochelin and carboxy-mycobactin be used in reference to siderophores from avirulent vs. virulent mycobacteria, respectively. Both exochelins and carboxy-mycobactins were shown to extract iron from ferritin [244, 253], the host iron storage protein. After chelating iron, these molecules would maintain its aqueous solubility and could diffuse so as to relay iron to mycobactin. Finally, it was shown that carboxy-mycobactins could also extract iron from transferrin and lactoferrin [254], host iron transport molecules. Altogether, the widely accepted model for mycobacterial iron acquisition starts with carboxy-mycobactins removing iron from host proteins [253, 254]. After diffusing to the mycobacterial cell wall, iron would be transferred to mycobactin and shuttled across the cell wall. An ABC-type transporter encoded in the *mbtII* locus has recently been reported to mediate mycobactin transport across the cell membrane [255]. Interestingly, conditional mutants in the mycobacterial esx-3 protein secretion apparatus [256] had a similar phenotype as mutants in the mycobactin ABC-type transporter. Not only are mutants compromised to grow in iron depleted media, but most important and in contrast to mutants in mycobactin synthesis, addition of

exogenous mycobactin is not able to reverse the defect. Thus, an uncharacterized link between esx-3 and the mycobactin ABC-type transporter has been suggested.

Several lines of evidence suggest that mycobactin is essential for *Mtb* physiology and pathogenesis. Initial transposon mutagenesis studies revealed that several genes of the *mbtI* locus are essential [257]. Even though alternative roles besides mycobactin import have not been defined for the esx-3 secretion apparatus, this secretion system is also essential [256]. In addition, mutation of non-essential genes such as *mbtB* results in *Mtb* failure to thrive inside macrophages [247]. Finally, mutation of the mycobactin ABC-transport system is not only deleterious for *Mtb in vitro* growth inside macrophages, but more importantly during *in vivo* growth in mice lungs [255]. Thus, pharmacological inhibition of mycobactin synthesis is actively being pursued as an alternative drug target [258].

1.7 Research Objectives.

After infection, the initial host-pathogen interaction is mediated by surface and secreted molecules. Characterization of the bacterial molecules participating in this initial interaction can lead to novel interventions aimed at reducing bacterial growth or minimizing tissue damage. Specifically for *Mtb*, the initial interaction with host pulmonary macrophages can lead to three principal outcomes: *Mtb* can replicate and induce host cell death, *Mtb* can be eradicated or alternatively, *Mtb* can persist in a dormant stage.

The first outcome of the host-*Mtb* interaction is intracellular replication leading to the host cell demise. Multiple rounds of *Mtb* replication and host cell death are presumed to lead to necrotic granulomas, a hallmark of human TB. Furthermore, dying macrophages provide a nutrient-replete milieu in which *Mtb* has been suggested to thrive. As macrophage cell death is a critical component of *Mtb* infection, understanding this event at a molecular level is warranted. In Chapter II, CF from virulent *Mtb* H37Rv was fractioned by chromatographic techniques in order to identify secreted molecules that

modulate macrophage cell death. Furthermore, the underlying mechanisms that participate in host cell death were also determined.

In Chapters III and IV, the biosynthesis and role in pathogenesis of α-L-polyGln was evaluated. α-L-polyGln is a cell wall associated product that has been reported to be exclusively present in *Mtb* and *M. bovis*. The presence of this molecule in virulent *Mycobacterium* spp, suggests that it could have an important role in pathogenesis. The identification of a pathway leading to α-L-polyGln synthesis could provide tools to evaluate its role in pathogenesis. In Chapter III, the biochemical characterization of an *Mtb* enzyme participating in α-L-polyGln biosynthesis was performed. In Chapter IV, *Mtb* mutants for α-L-polyGln biosynthesis were biochemically characterized and evaluated in the murine model of TB. The results suggest that α-L-polyGln could have a role in *Mtb* dormancy, the third outcome of the *Mtb*-host interaction.

Literature Cited

1. Spellberg, B., *Dr. William H. Stewart: mistaken or maligned?* Clin Infect Dis, 2008. **47**(2): p. 294.
2. Petersdorf, R.G., *Whither infectious diseases? Memories, manpower, and money.* J Infect Dis, 1986. **153**(2): p. 189-95.
3. World Health Organization., *Global tuberculosis control : WHO report 2011*. 2011, Geneva: World Health Organization. viii, 246 p.
4. Gengenbacher, M. and S.H. Kaufmann, *Mycobacterium tuberculosis: success through dormancy.* FEMS Microbiol Rev, 2012. **36**(3): p. 514-32.
5. Russell, D.G., *Mycobacterium tuberculosis: here today, and here tomorrow.* Nat Rev Mol Cell Biol, 2001. **2**(8): p. 569-77.
6. Russell, D.G., *Mycobacterium tuberculosis and the intimate discourse of a chronic infection.* Immunol Rev, 2011. **240**(1): p. 252-68.
7. Ordway, D., et al., *The cellular immune response to Mycobacterium tuberculosis infection in the guinea pig.* J Immunol, 2007. **179**(4): p. 2532-41.
8. Basaraba, R.J., *Experimental tuberculosis: the role of comparative pathology in the discovery of improved tuberculosis treatment strategies.* Tuberculosis (Edinb), 2008. **88 Suppl 1**: p. S35-47.
9. Cooper, A.M., et al., *Disseminated tuberculosis in interferon gamma gene-disrupted mice.* J Exp Med, 1993. **178**(6): p. 2243-7.
10. Kumar, V. and A. Maitra, *Robbins basic pathology*, in *Basic pathology*, V. Kumar, R.S. Cotran, and S.L. Robbins, Editors. 2003, Saunders: Philadelphia. p. 484-490.
11. Lenaerts, A.J., et al., *Location of persisting mycobacteria in a Guinea pig model of tuberculosis revealed by r207910.* Antimicrob Agents Chemother, 2007. **51**(9): p. 3338-45.
12. Ordway, D.J., et al., *Evaluation of standard chemotherapy in the guinea pig model of tuberculosis.* Antimicrob Agents Chemother, 2010. **54**(5): p. 1820-33.
13. Hoff, D.R., et al., *Location of intra- and extracellular M. tuberculosis populations in lungs of mice and guinea pigs during disease progression and after drug treatment.* PLoS One, 2011. **6**(3): p. e17550.
14. Algood, H.M., P.L. Lin, and J.L. Flynn, *Tumor necrosis factor and chemokine interactions in the formation and maintenance of granulomas in tuberculosis.* Clin Infect Dis, 2005. **41 Suppl 3**: p. S189-93.
15. Bigbee, C.L., et al., *Abatacept treatment does not exacerbate chronic Mycobacterium tuberculosis infection in mice.* Arthritis Rheum, 2007. **56**(8): p. 2557-65.
16. Bruns, H., et al., *Anti-TNF immunotherapy reduces CD8+ T cell-mediated antimicrobial activity against Mycobacterium tuberculosis in humans.* J Clin Invest, 2009. **119**(5): p. 1167-77.
17. Converse, P.J., et al., *Cavitary tuberculosis produced in rabbits by aerosolized virulent tubercle bacilli.* Infect Immun, 1996. **64**(11): p. 4776-87.
18. Dannenberg, A.M., Jr., *Liquefaction and cavity formation in pulmonary TB: a simple method in rabbit skin to test inhibitors.* Tuberculosis (Edinb), 2009. **89**(4): p. 243-7.
19. Dannenberg, A.M., Jr. and F.M. Collins, *Progressive pulmonary tuberculosis is not due to increasing numbers of viable bacilli in rabbits, mice and guinea pigs, but is due to a continuous host response to mycobacterial products.* Tuberculosis (Edinb), 2001. **81**(3): p. 229-42.
20. Scanga, C.A., et al., *The inducible nitric oxide synthase locus confers protection against aerogenic challenge of both clinical and laboratory strains of Mycobacterium tuberculosis in mice.* Infect Immun, 2001. **69**(12): p. 7711-7.

21. Driver, E.R., et al., *Evaluation of a mouse model of necrotic granuloma formation using C3HeB/FeJ mice for testing of drugs against Mycobacterium tuberculosis.* Antimicrob Agents Chemother, 2012. **56**(6): p. 3181-95.
22. Tsai, M.C., et al., *Characterization of the tuberculous granuloma in murine and human lungs: cellular composition and relative tissue oxygen tension.* Cell Microbiol, 2006. **8**(2): p. 218-32.
23. Via, L.E., et al., *Tuberculous granulomas are hypoxic in guinea pigs, rabbits, and nonhuman primates.* Infect Immun, 2008. **76**(6): p. 2333-40.
24. Aly, S., et al., *Oxygen status of lung granulomas in Mycobacterium tuberculosis-infected mice.* J Pathol, 2006. **210**(3): p. 298-305.
25. Bold, T.D. and J.D. Ernst, *Who benefits from granulomas, mycobacteria or host?* Cell, 2009. **136**(1): p. 17-9.
26. Davis, J.M. and L. Ramakrishnan, *The role of the granuloma in expansion and dissemination of early tuberculous infection.* Cell, 2009. **136**(1): p. 37-49.
27. Reece, S.T. and S.H. Kaufmann, *Floating between the poles of pathology and protection: can we pin down the granuloma in tuberculosis?* Curr Opin Microbiol, 2012. **15**(1): p. 63-70.
28. Kim, M.J., et al., *Caseation of human tuberculosis granulomas correlates with elevated host lipid metabolism.* EMBO Mol Med, 2010. **2**(7): p. 258-74.
29. Hawgood, S., *Pulmonary surfactant apoproteins: a review of protein and genomic structure.* Am J Physiol, 1989. **257**(2 Pt 1): p. L13-22.
30. Ferguson, J.S., et al., *Complement protein C3 binding to Mycobacterium tuberculosis is initiated by the classical pathway in human bronchoalveolar lavage fluid.* Infect Immun, 2004. **72**(5): p. 2564-73.
31. Gaynor, C.D., et al., *Pulmonary surfactant protein A mediates enhanced phagocytosis of Mycobacterium tuberculosis by a direct interaction with human macrophages.* J Immunol, 1995. **155**(11): p. 5343-51.
32. Ferguson, J.S., et al., *Surfactant protein D increases fusion of Mycobacterium tuberculosis-containing phagosomes with lysosomes in human macrophages.* Infect Immun, 2006. **74**(12): p. 7005-9.
33. Ferguson, J.S. and L.S. Schlesinger, *Pulmonary surfactant in innate immunity and the pathogenesis of tuberculosis.* Tuber Lung Dis, 2000. **80**(4-5): p. 173-84.
34. Ferguson, J.S., et al., *Surfactant protein D binds to Mycobacterium tuberculosis bacilli and lipoarabinomannan via carbohydrate-lectin interactions resulting in reduced phagocytosis of the bacteria by macrophages.* J Immunol, 1999. **163**(1): p. 312-21.
35. Hernandez-Pando, R., et al., *Persistence of DNA from Mycobacterium tuberculosis in superficially normal lung tissue during latent infection.* Lancet, 2000. **356**(9248): p. 2133-8.
36. Ashiru, O.T., M. Pillay, and A.W. Sturm, *Adhesion to and invasion of pulmonary epithelial cells by the F15/LAM4/KZN and Beijing strains of Mycobacterium tuberculosis.* J Med Microbiol, 2010. **59**(Pt 5): p. 528-33.
37. McDonough, K.A. and Y. Kress, *Cytotoxicity for lung epithelial cells is a virulence-associated phenotype of Mycobacterium tuberculosis.* Infect Immun, 1995. **63**(12): p. 4802-11.
38. Bermudez, L.E., et al., *The efficiency of the translocation of Mycobacterium tuberculosis across a bilayer of epithelial and endothelial cells as a model of the alveolar wall is a consequence of transport within mononuclear phagocytes and invasion of alveolar epithelial cells.* Infect Immun, 2002. **70**(1): p. 140-6.
39. Dobos, K.M., et al., *Necrosis of lung epithelial cells during infection with Mycobacterium tuberculosis is preceded by cell permeation.* Infect Immun, 2000. **68**(11): p. 6300-10.
40. Pethe, K., et al., *The heparin-binding haemagglutinin of M. tuberculosis is required for extrapulmonary dissemination.* Nature, 2001. **412**(6843): p. 190-4.
41. Pethe, K., et al., *Characterization of the heparin-binding site of the mycobacterial heparin-binding hemagglutinin adhesin.* J Biol Chem, 2000. **275**(19): p. 14273-80.

42. Aoki, K., et al., *Extracellular mycobacterial DNA-binding protein 1 participates in mycobacterium-lung epithelial cell interaction through hyaluronic acid.* J Biol Chem, 2004. **279**(38): p. 39798-806.
43. Rohde, K., et al., *Mycobacterium tuberculosis and the environment within the phagosome.* Immunol Rev, 2007. **219**: p. 37-54.
44. Schlesinger, L.S., *Mycobacterium tuberculosis and the complement system.* Trends Microbiol, 1998. **6**(2): p. 47-9; discussion 49-50.
45. Torrelles, J.B., et al., *Role of C-type lectins in mycobacterial infections.* Curr Drug Targets, 2008. **9**(2): p. 102-12.
46. Court, N., et al., *Partial redundancy of the pattern recognition receptors, scavenger receptors, and C-type lectins for the long-term control of Mycobacterium tuberculosis infection.* J Immunol, 2010. **184**(12): p. 7057-70.
47. Zimmerli, S., S. Edwards, and J.D. Ernst, *Selective receptor blockade during phagocytosis does not alter the survival and growth of Mycobacterium tuberculosis in human macrophages.* Am J Respir Cell Mol Biol, 1996. **15**(6): p. 760-70.
48. Kang, P.B., et al., *The human macrophage mannose receptor directs Mycobacterium tuberculosis lipoarabinomannan-mediated phagosome biogenesis.* J Exp Med, 2005. **202**(7): p. 987-99.
49. Schafer, G., et al., *Non-opsonic recognition of Mycobacterium tuberculosis by phagocytes.* J Innate Immun, 2009. **1**(3): p. 231-43.
50. Schlesinger, L.S., *Macrophage phagocytosis of virulent but not attenuated strains of Mycobacterium tuberculosis is mediated by mannose receptors in addition to complement receptors.* J Immunol, 1993. **150**(7): p. 2920-30.
51. Schlesinger, L.S., S.R. Hull, and T.M. Kaufman, *Binding of the terminal mannosyl units of lipoarabinomannan from a virulent strain of Mycobacterium tuberculosis to human macrophages.* J Immunol, 1994. **152**(8): p. 4070-9.
52. Khoo, K.H., et al., *Inositol phosphate capping of the nonreducing termini of lipoarabinomannan from rapidly growing strains of Mycobacterium.* J Biol Chem, 1995. **270**(21): p. 12380-9.
53. Schlesinger, L.S., et al., *Phagocytosis of Mycobacterium tuberculosis is mediated by human monocyte complement receptors and complement component C3.* J Immunol, 1990. **144**(7): p. 2771-80.
54. Dinadayala, P., et al., *Revisiting the structure of the anti-neoplastic glucans of Mycobacterium bovis Bacille Calmette-Guerin. Structural analysis of the extracellular and boiling water extract-derived glucans of the vaccine substrains.* J Biol Chem, 2004. **279**(13): p. 12369-78.
55. Hamasur, B., et al., *A mycobacterial lipoarabinomannan specific monoclonal antibody and its F(ab') fragment prolong survival of mice infected with Mycobacterium tuberculosis.* Clin Exp Immunol, 2004. **138**(1): p. 30-8.
56. Hamasur, B., et al., *Mycobacterium tuberculosis arabinomannan-protein conjugates protect against tuberculosis.* Vaccine, 2003. **21**(25-26): p. 4081-93.
57. Hamasur, B., G. Kallenius, and S.B. Svenson, *Synthesis and immunologic characterisation of Mycobacterium tuberculosis lipoarabinomannan specific oligosaccharide-protein conjugates.* Vaccine, 1999. **17**(22): p. 2853-61.
58. Malik, Z.A., G.M. Denning, and D.J. Kusner, *Inhibition of Ca(2+) signaling by Mycobacterium tuberculosis is associated with reduced phagosome-lysosome fusion and increased survival within human macrophages.* J Exp Med, 2000. **191**(2): p. 287-302.
59. Malik, Z.A., S.S. Iyer, and D.J. Kusner, *Mycobacterium tuberculosis phagosomes exhibit altered calmodulin-dependent signal transduction: contribution to inhibition of phagosome-lysosome fusion and intracellular survival in human macrophages.* J Immunol, 2001. **166**(5): p. 3392-401.
60. Malik, Z.A., et al., *Cutting edge: Mycobacterium tuberculosis blocks Ca2+ signaling and phagosome maturation in human macrophages via specific inhibition of sphingosine kinase.* J Immunol, 2003. **170**(6): p. 2811-5.

61. Brightbill, H.D., et al., *Host defense mechanisms triggered by microbial lipoproteins through toll-like receptors.* Science, 1999. **285**(5428): p. 732-6.
62. Sutcliffe, I.C. and D.J. Harrington, *Lipoproteins of Mycobacterium tuberculosis: an abundant and functionally diverse class of cell envelope components.* FEMS Microbiol Rev, 2004. **28**(5): p. 645-59.
63. Thoma-Uszynski, S., et al., *Induction of direct antimicrobial activity through mammalian toll-like receptors.* Science, 2001. **291**(5508): p. 1544-7.
64. Liu, P.T., et al., *Toll-like receptor triggering of a vitamin D-mediated human antimicrobial response.* Science, 2006. **311**(5768): p. 1770-3.
65. Gehring, A.J., et al., *Mycobacterium tuberculosis LprG (Rv1411c): a novel TLR-2 ligand that inhibits human macrophage class II MHC antigen processing.* J Immunol, 2004. **173**(4): p. 2660-8.
66. Noss, E.H., et al., *Toll-like receptor 2-dependent inhibition of macrophage class II MHC expression and antigen processing by 19-kDa lipoprotein of Mycobacterium tuberculosis.* J Immunol, 2001. **167**(2): p. 910-8.
67. El Kasmi, K.C., et al., *Toll-like receptor-induced arginase 1 in macrophages thwarts effective immunity against intracellular pathogens.* Nat Immunol, 2008. **9**(12): p. 1399-406.
68. Reiling, N., et al., *Cutting edge: Toll-like receptor (TLR)2- and TLR4-mediated pathogen recognition in resistance to airborne infection with Mycobacterium tuberculosis.* J Immunol, 2002. **169**(7): p. 3480-4.
69. Drennan, M.B., et al., *Toll-like receptor 2-deficient mice succumb to Mycobacterium tuberculosis infection.* Am J Pathol, 2004. **164**(1): p. 49-57.
70. Texereau, J., et al., *The importance of Toll-like receptor 2 polymorphisms in severe infections.* Clin Infect Dis, 2005. **41 Suppl 7**: p. S408-15.
71. Davila, S., et al., *Genetic association and expression studies indicate a role of toll-like receptor 8 in pulmonary tuberculosis.* PLoS Genet, 2008. **4**(10): p. e1000218.
72. Ishikawa, E., et al., *Direct recognition of the mycobacterial glycolipid, trehalose dimycolate, by C-type lectin Mincle.* J Exp Med, 2009. **206**(13): p. 2879-88.
73. Matsunaga, I. and D.B. Moody, *Mincle is a long sought receptor for mycobacterial cord factor.* J Exp Med, 2009. **206**(13): p. 2865-8.
74. Takimoto, H., et al., *Interferon-gamma independent formation of pulmonary granuloma in mice by injections with trehalose dimycolate (cord factor), lipoarabinomannan and phosphatidylinositol mannosides isolated from Mycobacterium tuberculosis.* Clin Exp Immunol, 2006. **144**(1): p. 134-41.
75. Fujita, Y., et al., *Molecular and supra-molecular structure related differences in toxicity and granulomatogenic activity of mycobacterial cord factor in mice.* Microb Pathog, 2007. **43**(1): p. 10-21.
76. Hunter, R.L., et al., *Multiple roles of cord factor in the pathogenesis of primary, secondary, and cavitary tuberculosis, including a revised description of the pathology of secondary disease.* Ann Clin Lab Sci, 2006. **36**(4): p. 371-86.
77. Rao, V., et al., *Trans-cyclopropanation of mycolic acids on trehalose dimycolate suppresses Mycobacterium tuberculosis -induced inflammation and virulence.* J Clin Invest, 2006. **116**(6): p. 1660-7.
78. Heitmann, L., et al., *Mincle is not essential for controlling Mycobacterium tuberculosis infection.* Immunobiology, 2012.
79. Clemens, D.L. and M.A. Horwitz, *Characterization of the Mycobacterium tuberculosis phagosome and evidence that phagosomal maturation is inhibited.* J Exp Med, 1995. **181**(1): p. 257-70.
80. Hart, P.D. and M.R. Young, *Interference with normal phagosome-lysosome fusion in macrophages, using ingested yeast cells and suramin.* Nature, 1975. **256**(5512): p. 47-9.

81. Gordon, A.H., P.D. Hart, and M.R. Young, *Ammonia inhibits phagosome-lysosome fusion in macrophages.* Nature, 1980. **286**(5768): p. 79-80.
82. Clemens, D.L. and M.A. Horwitz, *The Mycobacterium tuberculosis phagosome interacts with early endosomes and is accessible to exogenously administered transferrin.* J Exp Med, 1996. **184**(4): p. 1349-55.
83. Sturgill-Koszycki, S., U.E. Schaible, and D.G. Russell, *Mycobacterium-containing phagosomes are accessible to early endosomes and reflect a transitional state in normal phagosome biogenesis.* EMBO J, 1996. **15**(24): p. 6960-8.
84. Sturgill-Koszycki, S., et al., *Lack of acidification in Mycobacterium phagosomes produced by exclusion of the vesicular proton-ATPase.* Science, 1994. **263**(5147): p. 678-81.
85. Clemens, D.L., B.Y. Lee, and M.A. Horwitz, *Deviant expression of Rab5 on phagosomes containing the intracellular pathogens Mycobacterium tuberculosis and Legionella pneumophila is associated with altered phagosomal fate.* Infect Immun, 2000. **68**(5): p. 2671-84.
86. Clemens, D.L., B.Y. Lee, and M.A. Horwitz, *Mycobacterium tuberculosis and Legionella pneumophila phagosomes exhibit arrested maturation despite acquisition of Rab7.* Infect Immun, 2000. **68**(9): p. 5154-66.
87. Fratti, R.A., et al., *Role of phosphatidylinositol 3-kinase and Rab5 effectors in phagosomal biogenesis and mycobacterial phagosome maturation arrest.* J Cell Biol, 2001. **154**(3): p. 631-44.
88. Via, L.E., et al., *Arrest of mycobacterial phagosome maturation is caused by a block in vesicle fusion between stages controlled by rab5 and rab7.* J Biol Chem, 1997. **272**(20): p. 13326-31.
89. Vergne, I., J. Chua, and V. Deretic, *Mycobacterium tuberculosis phagosome maturation arrest: selective targeting of PI3P-dependent membrane trafficking.* Traffic, 2003. **4**(9): p. 600-6.
90. 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.
91. Birkeland, H.C. and H. Stenmark, *Protein targeting to endosomes and phagosomes via FYVE and PX domains.* Curr Top Microbiol Immunol, 2004. **282**: p. 89-115.
92. Vergne, I., et al., *Mycobacterium tuberculosis phagosome maturation arrest: mycobacterial phosphatidylinositol analog phosphatidylinositol mannoside stimulates early endosomal fusion.* Mol Biol Cell, 2004. **15**(2): p. 751-60.
93. Fratti, R.A., et al., *Mycobacterium tuberculosis glycosylated phosphatidylinositol causes phagosome maturation arrest.* Proc Natl Acad Sci U S A, 2003. **100**(9): p. 5437-42.
94. Goren, M.B., et al., *Prevention of phagosome-lysosome fusion in cultured macrophages by sulfatides of Mycobacterium tuberculosis.* Proc Natl Acad Sci U S A, 1976. **73**(7): p. 2510-4.
95. Olakanmi, O., et al., *Intraphagosomal Mycobacterium tuberculosis acquires iron from both extracellular transferrin and intracellular iron pools. Impact of interferon-gamma and hemochromatosis.* J Biol Chem, 2002. **277**(51): p. 49727-34.
96. Luo, M., E.A. Fadeev, and J.T. Groves, *Mycobactin-mediated iron acquisition within macrophages.* Nat Chem Biol, 2005. **1**(3): p. 149-53.
97. Wagner, D., et al., *Elemental analysis of Mycobacterium avium-, Mycobacterium tuberculosis-, and Mycobacterium smegmatis-containing phagosomes indicates pathogen-induced microenvironments within the host cell's endosomal system.* J Immunol, 2005. **174**(3): p. 1491-500.
98. Schaible, U.E., et al., *Cytokine activation leads to acidification and increases maturation of Mycobacterium avium-containing phagosomes in murine macrophages.* J Immunol, 1998. **160**(3): p. 1290-6.
99. Kim, B.H., et al., *A family of IFN-gamma-inducible 65-kD GTPases protects against bacterial infection.* Science, 2011. **332**(6030): p. 717-21.
100. Flynn, J.L. and J. Chan, *Immunology of tuberculosis.* Annu Rev Immunol, 2001. **19**: p. 93-129.
101. Gutierrez, M.G., et al., *Autophagy is a defense mechanism inhibiting BCG and Mycobacterium tuberculosis survival in infected macrophages.* Cell, 2004. **119**(6): p. 753-66.

102. Singh, S.B., et al., *Human IRGM induces autophagy to eliminate intracellular mycobacteria.* Science, 2006. **313**(5792): p. 1438-41.
103. Molloy, A., P. Laochumroonvorapong, and G. Kaplan, *Apoptosis, but not necrosis, of infected monocytes is coupled with killing of intracellular bacillus Calmette-Guerin.* J Exp Med, 1994. **180**(4): p. 1499-509.
104. Rojas, M., et al., *Mannosylated lipoarabinomannan antagonizes Mycobacterium tuberculosis-induced macrophage apoptosis by altering Ca+2-dependent cell signaling.* J Infect Dis, 2000. **182**(1): p. 240-51.
105. Rojas, M., et al., *TNF-alpha and IL-10 modulate the induction of apoptosis by virulent Mycobacterium tuberculosis in murine macrophages.* J Immunol, 1999. **162**(10): p. 6122-31.
106. Chen, M., H. Gan, and H.G. Remold, *A mechanism of virulence: virulent Mycobacterium tuberculosis strain H37Rv, but not attenuated H37Ra, causes significant mitochondrial inner membrane disruption in macrophages leading to necrosis.* J Immunol, 2006. **176**(6): p. 3707-16.
107. Duan, L., et al., *Critical role of mitochondrial damage in determining outcome of macrophage infection with Mycobacterium tuberculosis.* J Immunol, 2002. **169**(9): p. 5181-7.
108. Chen, M., et al., *Lipid mediators in innate immunity against tuberculosis: opposing roles of PGE2 and LXA4 in the induction of macrophage death.* J Exp Med, 2008. **205**(12): p. 2791-801.
109. Divangahi, M., et al., *Mycobacterium tuberculosis evades macrophage defenses by inhibiting plasma membrane repair.* Nat Immunol, 2009. **10**(8): p. 899-906.
110. Divangahi, M., et al., *Eicosanoid pathways regulate adaptive immunity to Mycobacterium tuberculosis.* Nat Immunol, 2010. **11**(8): p. 751-8.
111. Maiti, D., A. Bhattacharyya, and J. Basu, *Lipoarabinomannan from Mycobacterium tuberculosis promotes macrophage survival by phosphorylating Bad through a phosphatidylinositol 3-kinase/Akt pathway.* J Biol Chem, 2001. **276**(1): p. 329-33.
112. Rojas, M., et al., *Differential induction of apoptosis by virulent Mycobacterium tuberculosis in resistant and susceptible murine macrophages: role of nitric oxide and mycobacterial products.* J Immunol, 1997. **159**(3): p. 1352-61.
113. Chackerian, A.A., et al., *Dissemination of Mycobacterium tuberculosis is influenced by host factors and precedes the initiation of T-cell immunity.* Infect Immun, 2002. **70**(8): p. 4501-9.
114. Wolf, A.J., et al., *Initiation of the adaptive immune response to Mycobacterium tuberculosis depends on antigen production in the local lymph node, not the lungs.* J Exp Med, 2008. **205**(1): p. 105-15.
115. Neyrolles, O., et al., *Is adipose tissue a place for Mycobacterium tuberculosis persistence?* PLoS One, 2006. **1**: p. e43.
116. Gonzalez-Juarrero, M. and I.M. Orme, *Characterization of murine lung dendritic cells infected with Mycobacterium tuberculosis.* Infect Immun, 2001. **69**(2): p. 1127-33.
117. Marino, S., et al., *Dendritic cell trafficking and antigen presentation in the human immune response to Mycobacterium tuberculosis.* J Immunol, 2004. **173**(1): p. 494-506.
118. Wolf, A.J., et al., *Mycobacterium tuberculosis infects dendritic cells with high frequency and impairs their function in vivo.* J Immunol, 2007. **179**(4): p. 2509-19.
119. Basaraba, R.J., et al., *Lymphadenitis as a major element of disease in the guinea pig model of tuberculosis.* Tuberculosis (Edinb), 2006. **86**(5): p. 386-94.
120. Basaraba, R.J., et al., *Pulmonary lymphatics are primary sites of Mycobacterium tuberculosis infection in guinea pigs infected by aerosol.* Infect Immun, 2006. **74**(9): p. 5397-401.
121. Abadie, V., et al., *Neutrophils rapidly migrate via lymphatics after Mycobacterium bovis BCG intradermal vaccination and shuttle live bacilli to the draining lymph nodes.* Blood, 2005. **106**(5): p. 1843-50.
122. Orme, I.M. and F.M. Collins, *Protection against Mycobacterium tuberculosis infection by adoptive immunotherapy. Requirement for T cell-deficient recipients.* J Exp Med, 1983. **158**(1): p. 74-83.

123. Lefford, M.J., D.D. McGregor, and G.B. Mackaness, *Immune response to Mycobacterium tuberculosis in rats.* Infect Immun, 1973. **8**(2): p. 182-9.
124. Lefford, M.J., *Transfer of adoptive immunity to tuberculosis in mice.* Infect Immun, 1975. **11**(6): p. 1174-81.
125. Scanga, C.A., et al., *Depletion of CD4(+) T cells causes reactivation of murine persistent tuberculosis despite continued expression of interferon gamma and nitric oxide synthase 2.* J Exp Med, 2000. **192**(3): p. 347-58.
126. Diedrich, C.R., et al., *Reactivation of latent tuberculosis in cynomolgus macaques infected with SIV is associated with early peripheral T cell depletion and not virus load.* PLoS One, 2010. **5**(3): p. e9611.
127. Mattila, J.T., et al., *Simian immunodeficiency virus-induced changes in T cell cytokine responses in cynomolgus macaques with latent Mycobacterium tuberculosis infection are associated with timing of reactivation.* J Immunol, 2011. **186**(6): p. 3527-37.
128. Flynn, J.L., et al., *An essential role for interferon gamma in resistance to Mycobacterium tuberculosis infection.* J Exp Med, 1993. **178**(6): p. 2249-54.
129. Casanova, J.L. and L. Abel, *Genetic dissection of immunity to mycobacteria: the human model.* Annu Rev Immunol, 2002. **20**: p. 581-620.
130. Pai, R.K., et al., *Inhibition of IFN-gamma-induced class II transactivator expression by a 19-kDa lipoprotein from Mycobacterium tuberculosis: a potential mechanism for immune evasion.* J Immunol, 2003. **171**(1): p. 175-84.
131. Liu, P.T., et al., *Cutting edge: vitamin D-mediated human antimicrobial activity against Mycobacterium tuberculosis is dependent on the induction of cathelicidin.* J Immunol, 2007. **179**(4): p. 2060-3.
132. Fabri, M., et al., *Vitamin D is required for IFN-gamma-mediated antimicrobial activity of human macrophages.* Sci Transl Med, 2011. **3**(104): p. 104ra102.
133. Adams, J.S., et al., *Vitamin D in defense of the human immune response.* Ann N Y Acad Sci, 2007. **1117**: p. 94-105.
134. Gansert, J.L., et al., *Human NKT cells express granulysin and exhibit antimycobacterial activity.* J Immunol, 2003. **170**(6): p. 3154-61.
135. Reiley, W.W., et al., *ESAT-6-specific CD4 T cell responses to aerosol Mycobacterium tuberculosis infection are initiated in the mediastinal lymph nodes.* Proc Natl Acad Sci U S A, 2008. **105**(31): p. 10961-6.
136. Reiley, W.W., et al., *Distinct functions of antigen-specific CD4 T cells during murine Mycobacterium tuberculosis infection.* Proc Natl Acad Sci U S A, 2010. **107**(45): p. 19408-13.
137. Khader, S.A., et al., *IL-23 and IL-17 in the establishment of protective pulmonary CD4+ T cell responses after vaccination and during Mycobacterium tuberculosis challenge.* Nat Immunol, 2007. **8**(4): p. 369-77.
138. Scott-Browne, J.P., et al., *Expansion and function of Foxp3-expressing T regulatory cells during tuberculosis.* J Exp Med, 2007. **204**(9): p. 2159-69.
139. Henao-Tamayo, M.I., et al., *Phenotypic definition of effector and memory T-lymphocyte subsets in mice chronically infected with Mycobacterium tuberculosis.* Clin Vaccine Immunol, 2010. **17**(4): p. 618-25.
140. Orme, I.M., *The Achilles heel of BCG.* Tuberculosis (Edinb), 2010. **90**(6): p. 329-32.
141. Orme, I.M., *The kinetics of emergence and loss of mediator T lymphocytes acquired in response to infection with Mycobacterium tuberculosis.* J Immunol, 1987. **138**(1): p. 293-8.
142. Muller, I., et al., *Impaired resistance to Mycobacterium tuberculosis infection after selective in vivo depletion of L3T4+ and Lyt-2+ T cells.* Infect Immun, 1987. **55**(9): p. 2037-41.
143. Flynn, J.L., et al., *Major histocompatibility complex class I-restricted T cells are required for resistance to Mycobacterium tuberculosis infection.* Proc Natl Acad Sci U S A, 1992. **89**(24): p. 12013-7.

144. Behar, S.M., et al., *Susceptibility of mice deficient in CD1D or TAP1 to infection with Mycobacterium tuberculosis.* J Exp Med, 1999. **189**(12): p. 1973-80.
145. Sousa, A.O., et al., *Relative contributions of distinct MHC class I-dependent cell populations in protection to tuberculosis infection in mice.* Proc Natl Acad Sci U S A, 2000. **97**(8): p. 4204-8.
146. Mogues, T., et al., *The relative importance of T cell subsets in immunity and immunopathology of airborne Mycobacterium tuberculosis infection in mice.* J Exp Med, 2001. **193**(3): p. 271-80.
147. Grotzke, J.E. and D.M. Lewinsohn, *Role of CD8+ T lymphocytes in control of Mycobacterium tuberculosis infection.* Microbes Infect, 2005. **7**(4): p. 776-88.
148. Cho, S., et al., *Antimicrobial activity of MHC class I-restricted CD8+ T cells in human tuberculosis.* Proc Natl Acad Sci U S A, 2000. **97**(22): p. 12210-5.
149. Gonzalez-Juarrero, M., et al., *Temporal and spatial arrangement of lymphocytes within lung granulomas induced by aerosol infection with Mycobacterium tuberculosis.* Infect Immun, 2001. **69**(3): p. 1722-8.
150. Ottenhoff, T.H. and S.H. Kaufmann, *Vaccines against tuberculosis: where are we and where do we need to go?* PLoS Pathog, 2012. **8**(5): p. e1002607.
151. Gold, M.C., et al., *Human mucosal associated invariant T cells detect bacterially infected cells.* PLoS Biol, 2010. **8**(6): p. e1000407.
152. Rodgers, J.R. and R.G. Cook, *MHC class Ib molecules bridge innate and acquired immunity.* Nat Rev Immunol, 2005. **5**(6): p. 459-71.
153. Heinzel, A.S., et al., *HLA-E-dependent presentation of Mtb-derived antigen to human CD8+ T cells.* J Exp Med, 2002. **196**(11): p. 1473-81.
154. Sieling, P.A., et al., *CD1-restricted T cell recognition of microbial lipoglycan antigens.* Science, 1995. **269**(5221): p. 227-30.
155. Porcelli, S.A. and R.L. Modlin, *The CD1 system: antigen-presenting molecules for T cell recognition of lipids and glycolipids.* Annu Rev Immunol, 1999. **17**: p. 297-329.
156. Schaible, U.E., et al., *Correction of the iron overload defect in beta-2-microglobulin knockout mice by lactoferrin abolishes their increased susceptibility to tuberculosis.* J Exp Med, 2002. **196**(11): p. 1507-13.
157. Dannenberg, A.M., Jr., *Roles of cytotoxic delayed-type hypersensitivity and macrophage-activating cell-mediated immunity in the pathogenesis of tuberculosis.* Immunobiology, 1994. **191**(4-5): p. 461-73.
158. Mullerpattan, J.B. and Z.F. Udwadia, *Normal chest radiographs in sputum culture-positive pulmonary tuberculosis.* Int J Tuberc Lung Dis, 2009. **13**(1): p. 148, author reply 148-9.
159. Comas, I., et al., *Human T cell epitopes of Mycobacterium tuberculosis are evolutionarily hyperconserved.* Nat Genet, 2010. **42**(6): p. 498-503.
160. Cruz, A., et al., *Pathological role of interleukin 17 in mice subjected to repeated BCG vaccination after infection with Mycobacterium tuberculosis.* J Exp Med, 2010. **207**(8): p. 1609-16.
161. Mahairas, G.G., et al., *Molecular analysis of genetic differences between Mycobacterium bovis BCG and virulent M. bovis.* J Bacteriol, 1996. **178**(5): p. 1274-82.
162. Behr, M.A., et al., *Comparative genomics of BCG vaccines by whole-genome DNA microarray.* Science, 1999. **284**(5419): p. 1520-3.
163. Gey Van Pittius, N.C., et al., *The ESAT-6 gene cluster of Mycobacterium tuberculosis and other high G+C Gram-positive bacteria.* Genome Biol, 2001. **2**(10): p. RESEARCH0044.
164. Pym, A.S., et al., *Loss of RD1 contributed to the attenuation of the live tuberculosis vaccines Mycobacterium bovis BCG and Mycobacterium microti.* Mol Microbiol, 2002. **46**(3): p. 709-17.
165. Lewis, K.N., et al., *Deletion of RD1 from Mycobacterium tuberculosis mimics bacille Calmette-Guerin attenuation.* J Infect Dis, 2003. **187**(1): p. 117-23.
166. Pym, A.S., et al., *Recombinant BCG exporting ESAT-6 confers enhanced protection against tuberculosis.* Nat Med, 2003. **9**(5): p. 533-9.

167. Guinn, K.M., et al., *Individual RD1-region genes are required for export of ESAT-6/CFP-10 and for virulence of Mycobacterium tuberculosis.* Mol Microbiol, 2004. **51**(2): p. 359-70.
168. Champion, P.A., et al., *C-terminal signal sequence promotes virulence factor secretion in Mycobacterium tuberculosis.* Science, 2006. **313**(5793): p. 1632-6.
169. Abdallah, A.M., et al., *Type VII secretion--mycobacteria show the way.* Nat Rev Microbiol, 2007. **5**(11): p. 883-91.
170. de Jonge, M.I., et al., *ESAT-6 from Mycobacterium tuberculosis dissociates from its putative chaperone CFP-10 under acidic conditions and exhibits membrane-lysing activity.* J Bacteriol, 2007. **189**(16): p. 6028-34.
171. Smith, J., et al., *Evidence for pore formation in host cell membranes by ESX-1-secreted ESAT-6 and its role in Mycobacterium marinum escape from the vacuole.* Infect Immun, 2008. **76**(12): p. 5478-87.
172. 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.
173. Junqueira-Kipnis, A.P., et al., *Mycobacteria lacking the RD1 region do not induce necrosis in the lungs of mice lacking interferon-gamma.* Immunology, 2006. **119**(2): p. 224-31.
174. Volkman, H.E., et al., *Tuberculous granuloma formation is enhanced by a mycobacterium virulence determinant.* PLoS Biol, 2004. **2**(11): p. e367.
175. Volkman, H.E., et al., *Tuberculous granuloma induction via interaction of a bacterial secreted protein with host epithelium.* Science, 2010. **327**(5964): p. 466-9.
176. Stamm, L.M., et al., *Mycobacterium marinum escapes from phagosomes and is propelled by actin-based motility.* J Exp Med, 2003. **198**(9): p. 1361-8.
177. van der Wel, N., et al., *M. tuberculosis and M. leprae translocate from the phagolysosome to the cytosol in myeloid cells.* Cell, 2007. **129**(7): p. 1287-98.
178. Stanley, S.A., et al., *The Type I IFN response to infection with Mycobacterium tuberculosis requires ESX-1-mediated secretion and contributes to pathogenesis.* J Immunol, 2007. **178**(5): p. 3143-52.
179. Manzanillo, P.S., et al., *Mycobacterium tuberculosis activates the DNA-dependent cytosolic surveillance pathway within macrophages.* Cell Host Microbe, 2012. **11**(5): p. 469-80.
180. Flint, J.L., et al., *The RD1 virulence locus of Mycobacterium tuberculosis regulates DNA transfer in Mycobacterium smegmatis.* Proc Natl Acad Sci U S A, 2004. **101**(34): p. 12598-603.
181. Cole, S.T., et al., *Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence.* Nature, 1998. **393**(6685): p. 537-44.
182. Feltcher, M.E. and M. Braunstein, *Emerging themes in SecA2-mediated protein export.* Nat Rev Microbiol, 2012. **10**(11): p. 779-89.
183. Braunstein, M., et al., *Two nonredundant SecA homologues function in mycobacteria.* J Bacteriol, 2001. **183**(24): p. 6979-90.
184. Braunstein, M., et al., *SecA2 functions in the secretion of superoxide dismutase A and in the virulence of Mycobacterium tuberculosis.* Mol Microbiol, 2003. **48**(2): p. 453-64.
185. Gibbons, H.S., et al., *Identification of two Mycobacterium smegmatis lipoproteins exported by a SecA2-dependent pathway.* J Bacteriol, 2007. **189**(14): p. 5090-100.
186. Rigel, N.W., et al., *The Accessory SecA2 System of Mycobacteria Requires ATP Binding and the Canonical SecA1.* J Biol Chem, 2009. **284**(15): p. 9927-36.
187. Kurtz, S., et al., *The SecA2 secretion factor of Mycobacterium tuberculosis promotes growth in macrophages and inhibits the host immune response.* Infect Immun, 2006. **74**(12): p. 6855-64.
188. Sullivan, J.T., et al., *The Mycobacterium tuberculosis SecA2 system subverts phagosome maturation to promote growth in macrophages.* Infect Immun, 2012. **80**(3): p. 996-1006.
189. Lenz, L.L. and D.A. Portnoy, *Identification of a second Listeria secA gene associated with protein secretion and the rough phenotype.* Mol Microbiol, 2002. **45**(4): p. 1043-56.

190. Lenz, L.L., et al., *SecA2-dependent secretion of autolytic enzymes promotes Listeria monocytogenes pathogenesis.* Proc Natl Acad Sci U S A, 2003. **100**(21): p. 12432-7.
191. Hinchey, J., et al., *Enhanced priming of adaptive immunity by a proapoptotic mutant of Mycobacterium tuberculosis.* J Clin Invest, 2007. **117**(8): p. 2279-88.
192. Berks, B.C., *A common export pathway for proteins binding complex redox cofactors?* Mol Microbiol, 1996. **22**(3): p. 393-404.
193. Settles, A.M., et al., *Sec-independent protein translocation by the maize Hcf106 protein.* Science, 1997. **278**(5342): p. 1467-70.
194. Sargent, F., et al., *Overlapping functions of components of a bacterial Sec-independent protein export pathway.* EMBO J, 1998. **17**(13): p. 3640-50.
195. McDonough, J.A., et al., *The twin-arginine translocation pathway of Mycobacterium smegmatis is functional and required for the export of mycobacterial beta-lactamases.* J Bacteriol, 2005. **187**(22): p. 7667-79.
196. Raynaud, C., et al., *Phospholipases C are involved in the virulence of Mycobacterium tuberculosis.* Mol Microbiol, 2002. **45**(1): p. 203-17.
197. Posey, J.E., T.M. Shinnick, and F.D. Quinn, *Characterization of the twin-arginine translocase secretion system of Mycobacterium smegmatis.* J Bacteriol, 2006. **188**(4): p. 1332-40.
198. Saint-Joanis, B., et al., *Inactivation of Rv2525c, a substrate of the twin arginine translocation (Tat) system of Mycobacterium tuberculosis, increases beta-lactam susceptibility and virulence.* J Bacteriol, 2006. **188**(18): p. 6669-79.
199. DiGiuseppe Champion, P.A. and J.S. Cox, *Protein secretion systems in Mycobacteria.* Cell Microbiol, 2007. **9**(6): p. 1376-84.
200. McDonough, J.A., et al., *Identification of functional Tat signal sequences in Mycobacterium tuberculosis proteins.* J Bacteriol, 2008. **190**(19): p. 6428-38.
201. Marrichi, M., et al., *Genetic toggling of alkaline phosphatase folding reveals signal peptides for all major modes of transport across the inner membrane of bacteria.* J Biol Chem, 2008. **283**(50): p. 35223-35.
202. Braunstein, M., et al., *Identification of genes encoding exported Mycobacterium tuberculosis proteins using a Tn552'phoA in vitro transposition system.* J Bacteriol, 2000. **182**(10): p. 2732-40.
203. Sabin, F.R. and C.A. Doan, *The Biological Reactions in Rabbits to the Protein and Phosphatide Fractions from the Chemical Analysis of Human Tubercle Bacilli.* J Exp Med, 1927. **46**(4): p. 645-69.
204. Asselineau C and Asselineau J, *Waxes, Mycosides and Related Compounds*, in *The Mycobacteria a sourcebook*, G.P. Kubica and L.G. Wayne, Editors. 1984, Dekker: New York. p. 345-360.
205. Stodola Fh and Anderson RJ, *The Chemistry of the lipids of tubercle bacilli XLVI. Phthiocerol, a new alcohol from the wax of the human tubercle bacillus.* Journal of Biological Chemistry, 1936. **1936**: p. 467-472.
206. Ginger LG and Anderson RJ, *The Chemistry of the lipids of Tubercle Bacilli LXXII. Fatty acids occurring in the wax prepared from tuberculin residues, concerning mycocerosic acid.* Journal of Biological Chemistry, 1944. **157**: p. 203-211.
207. Demartreau-Ginsburg, H., et al., *Structure of phthiocerol.* Nature, 1959. **183**(4668): p. 117-9.
208. Asselineau J, Ryhage R, and Stenhage E, *Mass Spectrometry Studies of Long Chain Methyl Esters A Determination of the Molecular Weight and Structure of Mycocerosic Acid.* Acta Chemica Scandinavica, 1957. **11**(1): p. 196-198.
209. Rainwater, D.L. and P.E. Kolattukudy, *Synthesis of mycocerosic acids from methylmalonyl coenzyme A by cell-free extracts of Mycobacterium tuberculosis var. bovis BCG.* J Biol Chem, 1983. **258**(5): p. 2979-85.
210. Rainwater, D.L. and P.E. Kolattukudy, *Fatty acid biosynthesis in Mycobacterium tuberculosis var. bovis Bacillus Calmette-Guerin. Purification and characterization of a novel fatty acid*

synthase, mycocerosic acid synthase, which elongates n-fatty acyl-CoA with methylmalonyl-CoA. J Biol Chem, 1985. **260**(1): p. 616-23.
211. Mathur, M. and P.E. Kolattukudy, *Molecular cloning and sequencing of the gene for mycocerosic acid synthase, a novel fatty acid elongating multifunctional enzyme, from Mycobacterium tuberculosis var. bovis Bacillus Calmette-Guerin.* J Biol Chem, 1992. **267**(27): p. 19388-95.
212. Azad, A.K., et al., *Targeted replacement of the mycocerosic acid synthase gene in Mycobacterium bovis BCG produces a mutant that lacks mycosides.* Proc Natl Acad Sci U S A, 1996. **93**(10): p. 4787-92.
213. Cox, J.S., et al., *Complex lipid determines tissue-specific replication of Mycobacterium tuberculosis in mice.* Nature, 1999. **402**(6757): p. 79-83.
214. Trivedi, O.A., et al., *Enzymic activation and transfer of fatty acids as acyl-adenylates in mycobacteria.* Nature, 2004. **428**(6981): p. 441-5.
215. Trivedi, O.A., et al., *Dissecting the mechanism and assembly of a complex virulence mycobacterial lipid.* Mol Cell, 2005. **17**(5): p. 631-43.
216. Simeone, R., et al., *Delineation of the roles of FadD22, FadD26 and FadD29 in the biosynthesis of phthiocerol dimycocerosates and related compounds in Mycobacterium tuberculosis.* FEBS J, 2010. **277**(12): p. 2715-25.
217. Onwueme, K.C., et al., *Mycobacterial polyketide-associated proteins are acyltransferases: proof of principle with Mycobacterium tuberculosis PapA5.* Proc Natl Acad Sci U S A, 2004. **101**(13): p. 4608-13.
218. Goren, M.B., O. Brokl, and W.B. Schaefer, *Lipids of putative relevance to virulence in Mycobacterium tuberculosis: phthiocerol dimycocerosate and the attenuation indicator lipid.* Infect Immun, 1974. **9**(1): p. 150-8.
219. Rousseau, C., et al., *Production of phthiocerol dimycocerosates protects Mycobacterium tuberculosis from the cidal activity of reactive nitrogen intermediates produced by macrophages and modulates the early immune response to infection.* Cell Microbiol, 2004. **6**(3): p. 277-87.
220. Domenech, P. and M.B. Reed, *Rapid and spontaneous loss of phthiocerol dimycocerosate (PDIM) from Mycobacterium tuberculosis grown in vitro: implications for virulence studies.* Microbiology, 2009. **155**(Pt 11): p. 3532-43.
221. Astarie-Dequeker, C., et al., *Phthiocerol dimycocerosates of M. tuberculosis participate in macrophage invasion by inducing changes in the organization of plasma membrane lipids.* PLoS Pathog, 2009. **5**(2): p. e1000289.
222. Camacho, L.R., et al., *Analysis of the phthiocerol dimycocerosate locus of Mycobacterium tuberculosis. Evidence that this lipid is involved in the cell wall permeability barrier.* J Biol Chem, 2001. **276**(23): p. 19845-54.
223. Brieger, E.M. and A.M. Glauert, *Electron microscopy of the leprosy bacillus: a study of submicroscopical structure.* Tubercle, 1956. **37**(3): p. 195-206.
224. Yamamoto, T., et al., *Electron microscopy of ultra-thin sections of lepra cells and Mycobacterium leprae.* Int J Lepr, 1958. **26**(1): p. 1-8.
225. Chapman, G.B., J.H. Hanks, and J.H. Wallace, *An electron microscope study of the disposition and fine structure of Mycobacterium lepraemurium in mouse spleen.* J Bacteriol, 1959. **77**(2): p. 205-11.
226. Draper, P. and R.J. Rees, *Electron-transparent zone of mycobacteria may be a defence mechanism.* Nature, 1970. **228**(5274): p. 860-1.
227. Frehel, C., et al., *The electron-transparent zone in phagocytized Mycobacterium avium and other mycobacteria: formation, persistence and role in bacterial survival.* Ann Inst Pasteur Microbiol, 1986. **137B**(3): p. 239-57.
228. Rastogi, N., C. Frehel, and H.L. David, *Triple-layered structure of mycobacterial cell wall: Evidence for the existence of a polysaccharide-rich outer layer in 18 mycobacterial species.* Current Microbiology, 1986. **13**(5): p. 237-242.

229. Ortalo-Magne, A., et al., *Molecular composition of the outermost capsular material of the tubercle bacillus.* Microbiology, 1995. **141 (Pt 7)**: p. 1609-20.
230. Lemassu, A. and M. Daffe, *Structural features of the exocellular polysaccharides of Mycobacterium tuberculosis.* Biochem J, 1994. **297 (Pt 2)**: p. 351-7.
231. Dinadayala, P., et al., *Comparative structural analyses of the alpha-glucan and glycogen from Mycobacterium bovis.* Glycobiology, 2008. **18**(7): p. 502-8.
232. Wang, R., et al., *An anti-neoplastic glycan isolated from Mycobacterium bovis (BCG vaccine).* Biochem J, 1995. **311 (Pt 3)**: p. 867-72.
233. Gagliardi, M.C., et al., *Cell wall-associated alpha-glucan is instrumental for Mycobacterium tuberculosis to block CD1 molecule expression and disable the function of dendritic cell derived from infected monocyte.* Cell Microbiol, 2007. **9**(8): p. 2081-92.
234. Geurtsen, J., et al., *Identification of mycobacterial alpha-glucan as a novel ligand for DC-SIGN: involvement of mycobacterial capsular polysaccharides in host immune modulation.* J Immunol, 2009. **183**(8): p. 5221-31.
235. Cywes, C., et al., *Nonopsonic binding of Mycobacterium tuberculosis to complement receptor type 3 is mediated by capsular polysaccharides and is strain dependent.* Infect Immun, 1997. **65**(10): p. 4258-66.
236. Stokes, R.W., et al., *The glycan-rich outer layer of the cell wall of Mycobacterium tuberculosis acts as an antiphagocytic capsule limiting the association of the bacterium with macrophages.* Infect Immun, 2004. **72**(10): p. 5676-86.
237. Sambou, T., et al., *Capsular glucan and intracellular glycogen of Mycobacterium tuberculosis: biosynthesis and impact on the persistence in mice.* Mol Microbiol, 2008. **70**(3): p. 762-74.
238. Kalscheuer, R., et al., *Self-poisoning of Mycobacterium tuberculosis by targeting GlgE in an alpha-glucan pathway.* Nat Chem Biol, 2010. **6**(5): p. 376-84.
239. Quadri, L.E. and C. Ratledge, *Tuberculosis and the tubercle bacillus*, S.T. Cole, Editor. 2005, ASM Press: Washington, DC. p. 341-357.
240. Gangaidzo, I.T., et al., *Association of pulmonary tuberculosis with increased dietary iron.* J Infect Dis, 2001. **184**(7): p. 936-9.
241. Francis, J., et al., *Mycobactin, a growth factor for Mycobacterium johnei. I. Isolation from Mycobacterium phlei.* Biochem J, 1953. **55**(4): p. 596-607.
242. Snow, G.A., *The Structure of Mycobactin P, a Growth Factor for Mycobacterium Johnei, and the Significance of Its Iron Complex.* Biochem J, 1965. **94**: p. 160-5.
243. Snow, G.A., *Isolation and structure of mycobactin T, a growth factor from Mycobacterium tuberculosis.* Biochem J, 1965. **97**(1): p. 166-75.
244. Gobin, J., et al., *Iron acquisition by Mycobacterium tuberculosis: isolation and characterization of a family of iron-binding exochelins.* Proc Natl Acad Sci U S A, 1995. **92**(11): p. 5189-93.
245. Tateson, J.E., *Early steps in the biosynthesis of mycobactins P and S.* Biochem J, 1970. **118**(5): p. 747-53.
246. Quadri, L.E., et al., *Identification of a Mycobacterium tuberculosis gene cluster encoding the biosynthetic enzymes for assembly of the virulence-conferring siderophore mycobactin.* Chem Biol, 1998. **5**(11): p. 631-45.
247. De Voss, J.J., et al., *The salicylate-derived mycobactin siderophores of Mycobacterium tuberculosis are essential for growth in macrophages.* Proc Natl Acad Sci U S A, 2000. **97**(3): p. 1252-7.
248. Krithika, R., et al., *A genetic locus required for iron acquisition in Mycobacterium tuberculosis.* Proc Natl Acad Sci U S A, 2006. **103**(7): p. 2069-74.
249. McMahon, M.D., J.S. Rush, and M.G. Thomas, *Analyses of MbtB, MbtE, and MbtF suggest revisions to the mycobactin biosynthesis pathway in Mycobacterium tuberculosis.* J Bacteriol, 2012. **194**(11): p. 2809-18.

250. Antoine, A.D. and N.E. Morrison, *Effect of iron nutrition on the bound hydroxylamine content of Mycobacterium phlei.* J Bacteriol, 1968. **95**(1): p. 245-6.
251. Ratledge, C. and B.J. Marshall, *Iron transport in Mycobacterium smegmatis: the role of mycobactin.* Biochim Biophys Acta, 1972. **279**(1): p. 58-74.
252. Macham, L.P. and C. Ratledge, *A new group of water-soluble iron-binding compounds from Mycobacteria: the exochelins.* J Gen Microbiol, 1975. **89**(2): p. 379-82.
253. Macham, L.P., C. Ratledge, and J.C. Nocton, *Extracellular iron acquisition by mycobacteria: role of the exochelins and evidence against the participation of mycobactin.* Infect Immun, 1975. **12**(6): p. 1242-51.
254. Gobin, J. and M.A. Horwitz, *Exochelins of Mycobacterium tuberculosis remove iron from human iron-binding proteins and donate iron to mycobactins in the M. tuberculosis cell wall.* J Exp Med, 1996. **183**(4): p. 1527-32.
255. Rodriguez, G.M. and I. Smith, *Identification of an ABC transporter required for iron acquisition and virulence in Mycobacterium tuberculosis.* J Bacteriol, 2006. **188**(2): p. 424-30.
256. Siegrist, M.S., et al., *Mycobacterial Esx-3 is required for mycobactin-mediated iron acquisition.* Proc Natl Acad Sci U S A, 2009. **106**(44): p. 18792-7.
257. Sassetti, C.M., D.H. Boyd, and E.J. Rubin, *Genes required for mycobacterial growth defined by high density mutagenesis.* Mol Microbiol, 2003. **48**(1): p. 77-84.
258. Juarez-Hernandez, R.E., S.G. Franzblau, and M.J. Miller, *Syntheses of mycobactin analogs as potent and selective inhibitors of Mycobacterium tuberculosis.* Org Biomol Chem, 2012. **10**(37): p. 7584-93.

Chapter II

Stable Extracellular RNA Fragments of *Mycobacterium tuberculosis* Induce Human Monocyte Apoptosis via a Caspase-8 Dependent Mechanism[1]

2.1 Introduction

Apoptosis mainly occurs following activation of the caspase cascade by the mitochondrial or apical pathways [1, 2]. Mitochondrial release of organelle constituents such as cytochrome c and other macromolecules [3], activates the caspase cascade via pro-caspase-9. Alternatively the apical pathway is mediated by activation of pro-caspase-8 after cross linking of cell surface receptors belonging to the Tumor Necrosis Factor Receptor (TNFR) super family [4]. Both pathways converge to activate the executioner caspase-3 [1, 3]. Other alternative pathways designated as "type II" that do not rely on caspase 3 activity were described recently [5]. Cell death ensues after cleavage of the cell's cytoskeleton and intracellular enzymes involved in cellular homeostasis, leading to apoptosis hallmarks such as exposure of PS in the cell membrane, DNA fragmentation, and cell shrinkage [1, 2].

Apoptosis is a process associated with multiple infectious diseases including tuberculosis [6]. Several lines of evidence suggest that apoptosis is an important process occurring during mycobacterial infection [7]. Recently, caspase activation and TUNEL staining demonstrated that apoptosis is a major event occurring in the caseous necrosis of granulomas from *Mtb*-infected humans [8]. More importantly, *sst-1* and *slca11* (formerly *nramp-1*), the best characterized murine genetic loci determining susceptibility/resistance to mycobacterial infections, were associated with the degree of macrophage apoptosis [9-11]. It also is known that macrophage infection with live *Mtb* or stimulation with PPD activate both the apical and mitochondrial apoptotic pathways by inducing TNF-α production, caspase

[1] (presented in Obregón-Henao A *et al.* Stable extracellular RNA fragments of *Mycobacterium tuberculosis* induce early apoptosis in human monocytes via a caspase-8 dependent mechanism. PLoS One. 2012; 7(1):e299702012)

activation and calcium influx, [11-15]. In contrast dead *Mtb* and ManLAM inhibit these pathways by inducing IL-10, activating Akt and blocking calcium influx [14-18].

Several mycobacterial molecules of diverse chemical nature were reported to induce apoptosis in various types of host cells. In particular, the 19 kilodalton (kDa) lipoprotein was shown to trigger murine and human macrophages apoptosis through TLR-2 via nitric oxide dependent and independent pathways [19-22]. Glycolipids such as TDM and LM from *Mtb* also activate apoptosis [23-25]. TDM induces apoptosis of murine natural killer and naïve T cells [23, 24], while LM targets human macrophages [25]. Contrary to the consistent finding that macrophage apoptosis requires infection with live, metabolically active mycobacteria [11], the aforementioned mycobacterial molecules are cell wall associated products purified from inactivated cells. However, a systematic purification of secreted products released by actively growing *Mtb* that promote apoptosis has not been performed.

To further address this issue we purified the major apoptosis-inducing product released by virulent *Mtb* strain H37Rv, and unexpectedly discovered this primary apoptotic factor was stable mycobacterial RNA fragments and not a protein or lipid. Additional experimentation revealed the nature of these stable RNA molecules and that their induction of apoptosis was via a caspase-8 dependent, and TNF-α and caspase-1 independent mechanism that affected the ability of human monocytes to control mycobacterial infection.

2.2 Materials and Methods

2.2.1 Growth of *Mtb*.

To prepare bacilli for monocyte infections, *Mtb* H37Rv (kindly provided by Laboratorio de Micobacterias, Instituto Nacional de Salud, Bogota, Colombia) was grown as a pellicle in Middlebrook 7H9 medium (Becton Dickinson, Sparks, MD) supplemented with glycerol (Promega, Madison, WI), oleic acid-albumin-dextrose-catalase (OADC) (Becton Dickinson), and Tween 80 (Becton Dickinson). Cultures were collected and *Mtb* cells washed twice with phosphate buffered saline (PBS). To disrupt

bacterial clumps, *Mtb* was suspended in RPMI-1640 with 20% glycerol and probe sonicated (Model CV33 Sonics Vibra Cell, Newtown, CT) five times at 2.5 W output for 1 min at 4º C. The cell suspensions were stored at -70^{o} C until use. CFUs were determined prior to monocyte infections by plating 100 µL of serial dilutions of an *Mtb* stock on Middlebrook 7H10 agar (Becton Dickinson) supplemented with glycerol and OADC. Colonies were counted after two weeks of incubation at 37º C.

2.2.2 Isolation of Extracellular RNA.

For isolation of biologically active fractions, *Mtb* strain H37Rv was grown in glycerol alanine salts (GAS) medium under constant rotation in roller bottles for two weeks at 37°C and culture filtrate (CF) was obtained as previously described [26]. Similarly, CF was obtained from *Mtb* HN878, H37Rv ΔRD-1 (obtained from Dr. Jacobs) in order to evaluate if RNA release also occurred in different strains including clinical isolates. Lyophilized CF was suspended overnight at 4°C at 1 mg/ml protein concentration in ConcanavalinA (ConA) binding buffer [27], and passed over a ConA-Sephadex column (Sigma-Aldrich, St. Louis, MO) to remove mannosylated glycoproteins and lipoglycans (LM and LAM). The flow-through fraction was collected, concentrated, and washed three times with 20 mM Tris (pH 8.1). The CF depleted of mannosylated glycoconjugates (25 mg) was applied to a 1 x 10 cm column packed with 5 ml of DEAE-Sepharose (Amersham Biosciences Corporation, Piscataway, NJ) equilibrated with Tris 20 mM (pH 8.1). The column was washed with 20 mM Tris 20 (pH 8.1) and eluted with stepwise increases (25 to 1000 mM) of NaCl in 20 mM Tris (pH 8.1). Each fraction was concentrated and dialyzed against 10 mM ammonium bicarbonate, filtered sterilized, lyophilized and stored at -20°C until used.

Initial biochemical analysis of mycobacterial RNA was accomplished by digesting fraction 7 for increasing periods of time at 37°C with either RNaseA, RNaseT or RNaseVI (Ambion, Austin, TX).

To prepare highly purified RNA, the material eluted from DEAE-sepharose with 500-1000 mM NaCl was suspended in 0.1 ml of 10 mM ammonium bicarbonate, 1 mM $MgCl_2$ to a final concentration of 0.4 mg/ml and digested overnight at 37°C with RNase-free DNaseI (Sigma-Aldrich), followed by RNase-

free Proteinase K (Sigma-Aldrich) for 8 h at 37°C at a final concentrations of 0.04 mg/ml. The digested material was partitioned with an equal vol of phenol:chloroform:isoamyl alcohol (25:24:1) followed by an equal vol of chloroform:isoamyl alcohol (24:1). RNA was precipitated from the final aqueous phase with 0.1 vol of 3 M sodium acetate (pH 5.3) and 2.5 vol of ethanol, and collected by centrifugation (13,000 x g, 30 min). Final purification was performed by preparative gel electrophoresis using a denaturing urea poly-acrylamide gel [28]. RNA preparations were suspended in PBS, and filter sterilized. The gel purified RNA was incubated overnight at 37° C with or without 0.1 U/µl RNaseV1 (Ambion) and lyophilized. Endotoxin contamination was determined using the Lymulus Amebocyte Lysate assay (Bio-Whittaker, Walkersville, Maryland). RNA concentration was determined by densitometry using Quantity One software (Version 4.1.1 Bio-Rad, Hercules, CA) after resolving the *Mtb* RNA and standards by denaturing urea poly-acrylamide gel electrophoresis and staining with ethidium bromide. To assess for the presence of LAM, samples (1 µg) were subjected to Western blot analysis using the LAM specific monoclonal antibody CS-35 as the probe [29]. As an alternative and faster method to obtain mycobacterial RNA, lyophilized CF was directly extracted with organic solvents and precipitated with ethanol as described above.

To evaluate the kinetics of mycobacterial RNA release into the CF, low passage *Mtb* H37Rv was centrifuged at 3,000 rpm and washed three times with medium. Bacterial pellets were suspended in GAS medium and used to inoculate 400 ml cultures. CF was collected at different time points of incubation and processed. Alternatively, CF was precipitated overnight at -20°C with three volumes of ice cold acetone. The precipitated material was collected by centrifugation and analyzed by SDS-PAGE [30].

2.2.3 RNA cloning and sequencing.

Gel purified RNA was cloned as described by Elbashir and Lau for small interfering (si)RNA [28, 31]. Briefly, the RNA was dephosphorylated and ligated to the 3'adaptor [5' phosphorylated uuu AAC CGC ATC CTT CTC iT-3'; lower case letters ribonucleotides, capitalized letters deoxyribonucleotides,

(Dharmacon, Lafayette, CO)]. The ligation products were gel purified, phosphorylated at the 5' end with T4 polynucleotide kinase (New England Biolabs, Beverly MA) and ligated to the 5' adaptor sequence [5' TAC TAA TAC GAC TCA CT aaa 3' (Dharmacon)]. Synthesis of cDNA from adaptor-ligated RNA was accomplished using the Thermoscript kit (Invitrogen) and the reverse primer (5' GAC TAG CTG GAA TTC AAG GAT GCG GTT AAA-3'). Double stranded DNA was generated by PCR amplification with the reverse primer, and the forward primer (5'CAG CCA ACG GAA TTC ATA CGA CTC ACT AAA-3'). An aliquot of the PCR products was digested with *EcoR*I (New England Biolabs) and ligated to form concatemers that were resolved by agarose gel electrophoresis and extracted. Overhangs of poly-A were added with Taq Platinum polymerase in the presence of dNTPs. The DNA fragments were ligated with pCR 2.1 TOPO vector (Invitrogen), transformed into TOP10 *E. coli* (Invitrogen) and individual clones grown for plasmid isolation. The cloned DNA fragments were sequenced through Macromolecular Resources Facility (Colorado State University). Plasmid sequences were analyzed with the Vector NTI software (BioExchange, San Francisco, CA) to identify the 5' and 3' adaptor sequences, and the intervening regions were searched against the *Mtb* genome, via BLAST analysis.

2.2.4 Monocyte infection and stimulation.

Venous blood (120 ml) was obtained from healthy volunteers after signing an informed consent form describing the protocol for human subjects, which was approved by the Ethics Committee of the Facultad de Medicina, Universidad de Antioquia. The blood was defibrinated by continuous agitation with glass beads, centrifuged, and the buffy coat suspended in 3 vol of PBS. Mononuclear cells were separated on Histopaque, density 1.077 g/ml (Sigma-Aldrich) by centrifugation at 400 x *g,* and cell viability determined by trypan blue exclusion. Cells were suspended at $1x10^6$ cells/ml in RPMI-1640 medium (Gibco BRL) without antibiotics and supplemented with 0.5 % heat-inactivated autologous serum (AS). The ratio of monocytes was determined by flow cytometry with anti–CD14-PE staining (clone M5E2, Becton Dickinson-Pharmingen). CD14+ cells ($2x10^5$/well) were plated in 48-well flat-

bottomed culture plates (Corning, Corning, NY) for 4 h at 37º C, and non-adherent cells removed by repeatedly washing with prewarmed PBS with 0.5% AS. RPMI-1640 supplemented with 10% AS (1 ml) was added to each well. The number of adherent cells was determined by mechanical removal and counting in a haemocytometer.

Adhered monocytes were infected with viable *Mtb* at a multiplicity of infection of 5:1 for 4 h. Extracellular bacilli were removed by repeatedly washing with prewarmed PBS. RPMI-1640 supplemented with 10% AS (1 ml) was added to each well. The CF and RNA preparations were added at 1 μg/ml protein or nucleic acid concentration, respectively. Cultures were incubated at 37º C, 5% CO_2 for 96 h, and the number of viable bacilli determined by plate counting after lysis of host cells with water containing 0.1% saponin [12].

For analysis of apoptosis induction or cytokine production the CF, individual CF fractions, and purified mycobacterial RNA or rabbit mRNA (Sigma) were added at 1 μg/ml protein or nucleic acid concentration, respectively. PPD was used as a control at 10 μg/ml. Induction of human monocytes with anti-CD95 was performed by incubating 10^6 cells with 2 μg/ml of anti-human CD95 (BD Pharmingen).

2.2.5 Determination of intracellular TNFα and IL-10 by flow cytometry.

Evaluation of intracellular cytokines was based on previous methods [14]. Briefly, cells were cultured for 18 h with mycobacterial products at 37°C and 5% CO_2. Brefeldin A (1 μg/ml) was subsequently added and the cultures were incubated for 6 h. Cells were washed twice with PBS, fixed with 2% paraformaldehyde in 0.1 M NaH_2PO_4 for 20 min at room temperature, and harvested. Intracellular cytokine labeling was achieved by washing the cells once with permeabilization buffer [PBS (pH7.4) containing 1% PHS, 1% BSA, 0.1% sodium azide and 0.1% saponin], and incubating in 200 μl of permeabilization buffer containing 2.5 μg mAb anti-TNF-α-FITC and 2.5 μg anti-IL-10-PE, or IgG2a-FITC and IgG2b-PE as isotype controls for 30 min at 4ºC. Cells were washed three times with PBS containing 1% PHS, 1% BSA, and 0.1% sodium azide, pH 7.4. As control for the intracellular signal,

non-permeabilized cells were stained with the TNF-α and IL-10 specific antibodies. Thereafter cells were analyzed by flow cytometry with a FACS EPICS XL (Coulter, Hialeah, FL). The percentage of positive cells and the mean fluorescence intensity were determined using Windows Multiple Document Interface 2,8 software (WinMDI, Scrips Research Institute, La Jolla, CA).

2.2.6 Determination of cell death by flow cytometry using annexinV and propidium iodide.

Cells were cultured for 48 h in the presence of the mycobacterial products and incubated for 30 min at RT in darkness with 5 μl of FITC-labeled annexinV (Molecular Probes, Invitrogen), 10 μl of propidium iodide at 1 μg/ml (ICN Biomedicals, Costa Mesa, CA), and 100 μl of PBS. Cells were washed with PBS and analyzed by FACS [11].

To assess whether apoptosis was dependent on TNF-α signalling, monocytes were incubated for 1 h with purified anti–TNF-α (MAb11, Becton Dickinson-Pharmingen), followed by the addition of the different stimuli for 36 h. Thereafter, annexinV staining was evaluated.

2.2.7 Caspase activation.

Caspase activation was measured at 48 h post addition of *Mtb* products to the adhered monocytes. The activation of caspase-1, 3, and 8 was determined by flow cytometry as previously described [12]. To evaluate whether caspase activation was required for apoptosis, monocytes were incubated with 10 nM of caspase-1 (YVAD.fmk), caspase-3 (DEVD.fmk), or caspase-8 (IEMD.fmk) specific inhibitors (Sigma-Aldrich) for 1 h prior to the addition of mycobacterial products. Dimethylsulphoxide and trichloroacetic acid were also assessed as buffer controls. Exposure of PS was then evaluated by staining with annexinV.

2.2.8 Statistical analysis.

Experiments were performed a minimum of three independent times. Data were analyzed with GraphPad Prism, version 4 (GraphPad, San Diego, CA). Comparisons between treatments were

performed by one-way ANOVA. Statistical significance was tested at $p < 0.05$ as the critical value. Data are presented as the mean ± SEM.

2.3 Results

2.3.1 Purification and identification of apoptosis inducer.

Infection of human derived monocytes for 48 h with live, virulent *Mtb* (MOI of 5:1) induces early apoptotic events that are defined by annexinV positive and propidium iodide negative staining [12]. However, apoptosis does not occur with formaldehyde-fixed or heat-killed *Mtb.* Thus, it was hypothesized that induction of apoptosis in human monocytes was due to a secreted product of *Mtb* [11]. Towards purifying the responsible mycobacterial product(s), the CF from *in vitro* cultures of *Mtb* H37Rv was shown to induce monocyte cell membrane damage similar to the live infection [12]. As a first step in the purification scheme the ManLAM, previously shown to inhibit macrophage apoptosis [15, 17, 18], was removed by ConA-affinity chromatography to give Man-LAM depleted CF (CF-Man). This material was further resolved on DEAE-Sephadex with a stepwise gradient of sodium chloride from 0-1000 mM, resulting in seven fractions (Fig. 2.1).

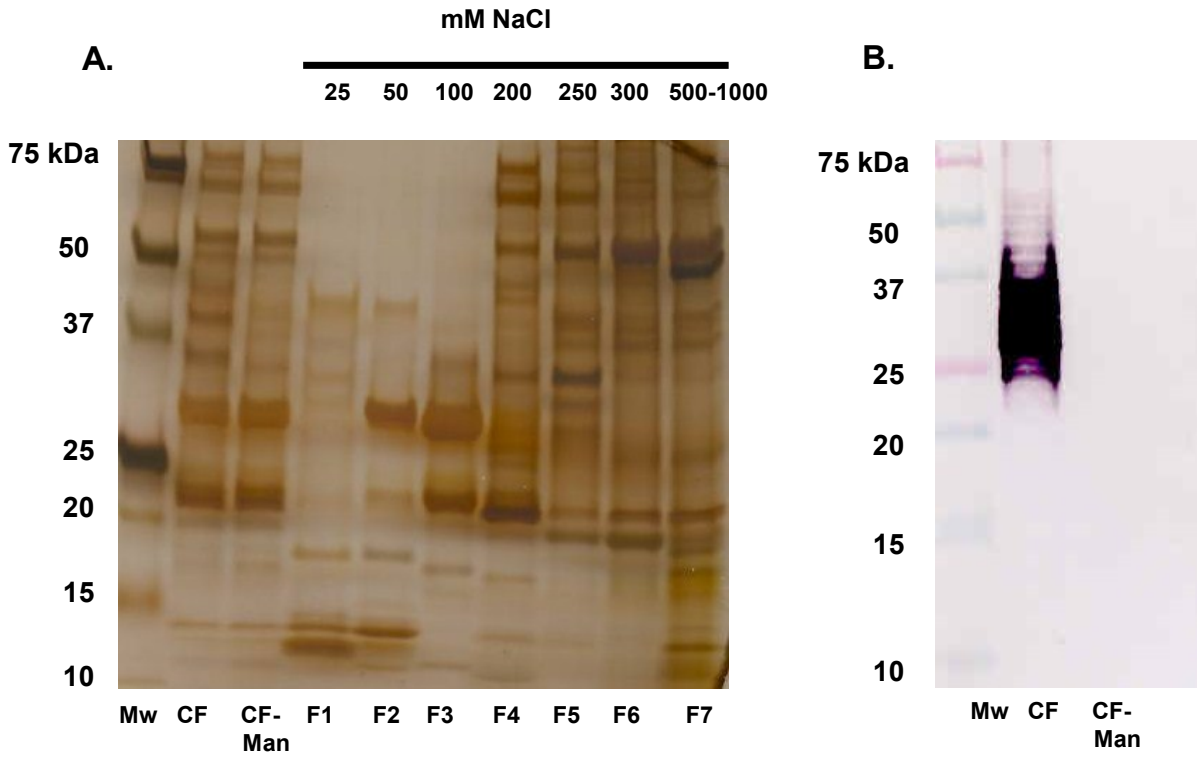


Figure 2.1 Fractionation of CF by chromatography. A. SDS-PAGE and silver staining of fractions of CF after ConA affinity (CF-Man) and DEAE-Sepharose chromatography (F1 to F7). B. Western blot with CS35 anti-LAM showing that ConA chromatography efficiently removed ManLAM from the CF. Mw denotes the molecular mass standards.

Evaluation of each fraction for induction of apoptosis via annexinV staining, demonstrated that fraction 7 induced the strongest response with nearly 70% of human monocytes displaying early cell membrane damage (Fig. 2.2) (a small percentage of cells was dually labeled with annexinV and propidium iodide, data not shown). Considerable activity was also associated with fraction 5 and 6. However, further purification and characterization was targeted to fraction 7 as it was the most potent.

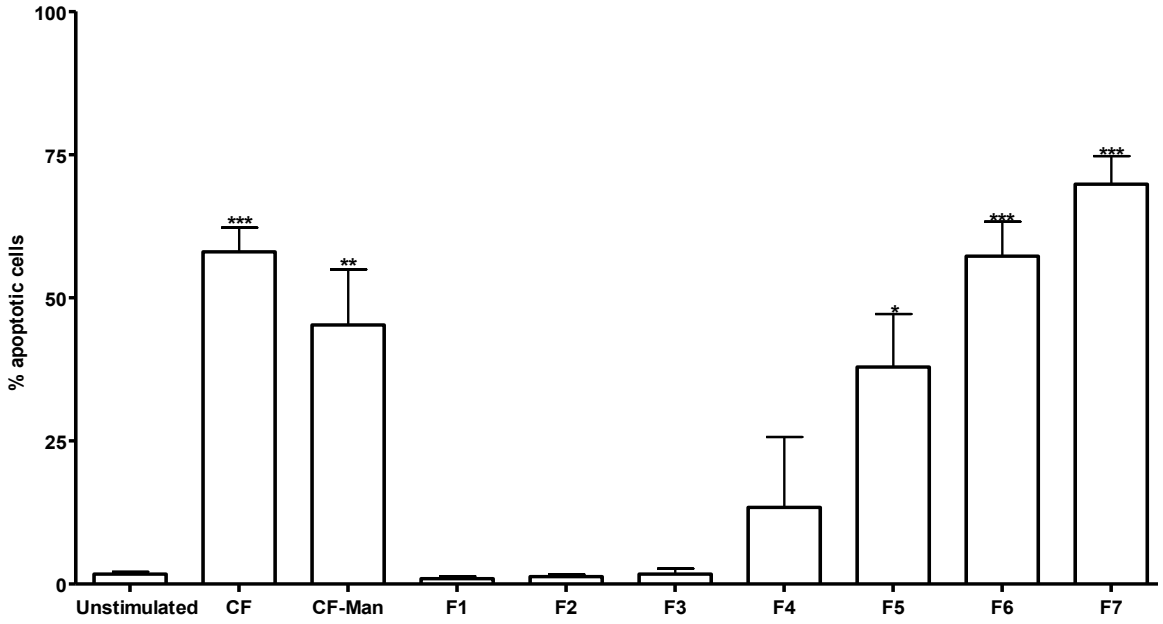


Figure 2.2 Apoptosis-inducing activity of CF fractions. Separation of CF by ConA affinity chromatography and DEAE-Sepharose chromatography yielded a fraction with potent apoptotic activity in human monocytes. Apoptosis induced by DEAE-Sepharose fractions 1 to 7 (F1 to F7) and unfractionated CF was measured by flow cytometry and presented as percentage of human monocytes that stained annexinV positive *** significance $p<0.001$, ** significance $p<0.01$, * significance $p<0.05$ as compared to unstimulated control.

The chemical nature of the apoptotic-inducing mycobacterial molecule(s) present in fraction 7 was initially determined by assessing the loss of stimulation after differential enzymatic digestion with Proteinase K, DNaseI, or RNaseV1. Unexpectedly, the apoptosis-inducing activity of fraction 7 was completely abrogated by treatment with RNaseV1, as there was a 95% reduction in the number of annexinV positive cells. Digestion with DNaseI had no impact, whereas Proteinase K treatment increased the number of monocytes in early apoptosis (80% of monocytes), but was not statistically significant (Fig. 2.3). Treatment of monocytes with enzymes alone had a negligible effect on apoptosis (data not shown).

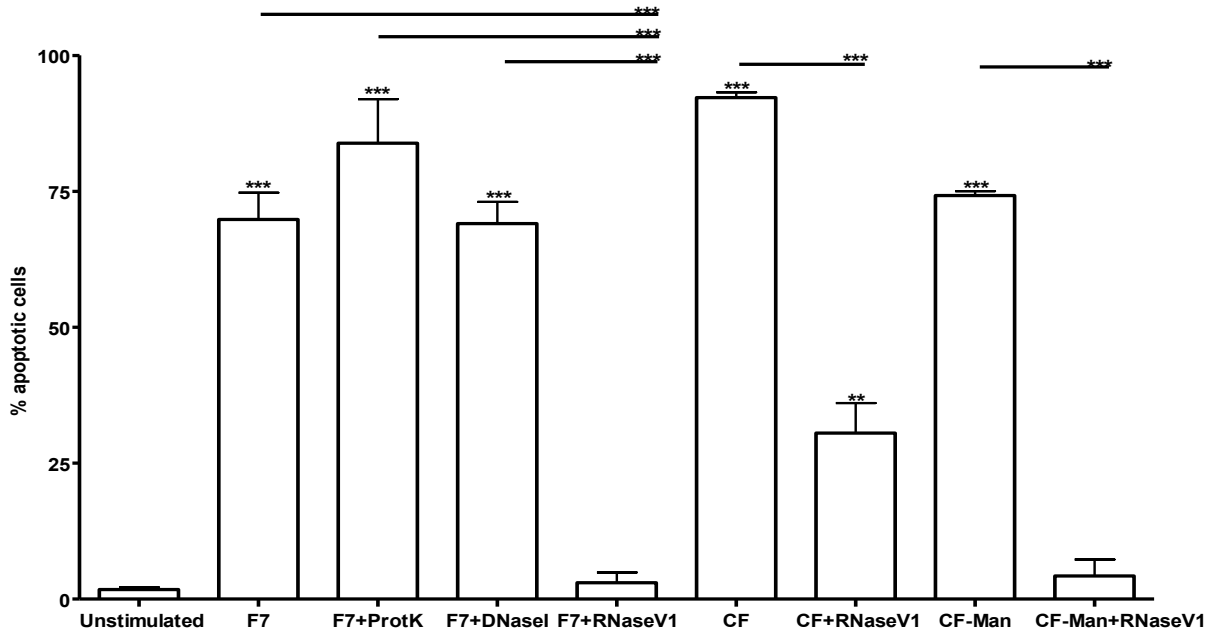


Figure 2.3 RNA in DEAE-Sepharose Fraction 7 induces apoptosis in human monocytes. Apoptosis induced by DEAE-Sepharose Fraction 7 (F7), F7 treated with proteinase K (F7+ProtK), F7 treated with DNase1 (F7 + DNase1), F7 treated with RNaseV1 (F7 + RNaseV1), CF, CF treated with RNaseV1 (CF+RNaseV1), CF-Man, CF-Man treated with RNaseV1 (CF-Man+RNaseV1). Apoptosis was measured by flow cytometry and presented as percentage of human monocytes that stained annexinV positive. *** significance $p<0.001$, ** significance $p<0.01$ as compared to unstimulated control or between treated samples.

Evaluation of fraction 7 by SDS-PAGE and ethidium bromide or silver staining pre and post enzymatic digestions confirmed the presence of nucleic acids in fraction 7. These nucleic acids migrated at < 20 kDa by SDS-PAGE and were only eliminated by digestion with RNaseV1 (Fig. 2.4A and B). It was also noted that Proteinase K, but not RNaseV1 or DNaseI treatment resulted in the elimination of proteins present in fraction 7.

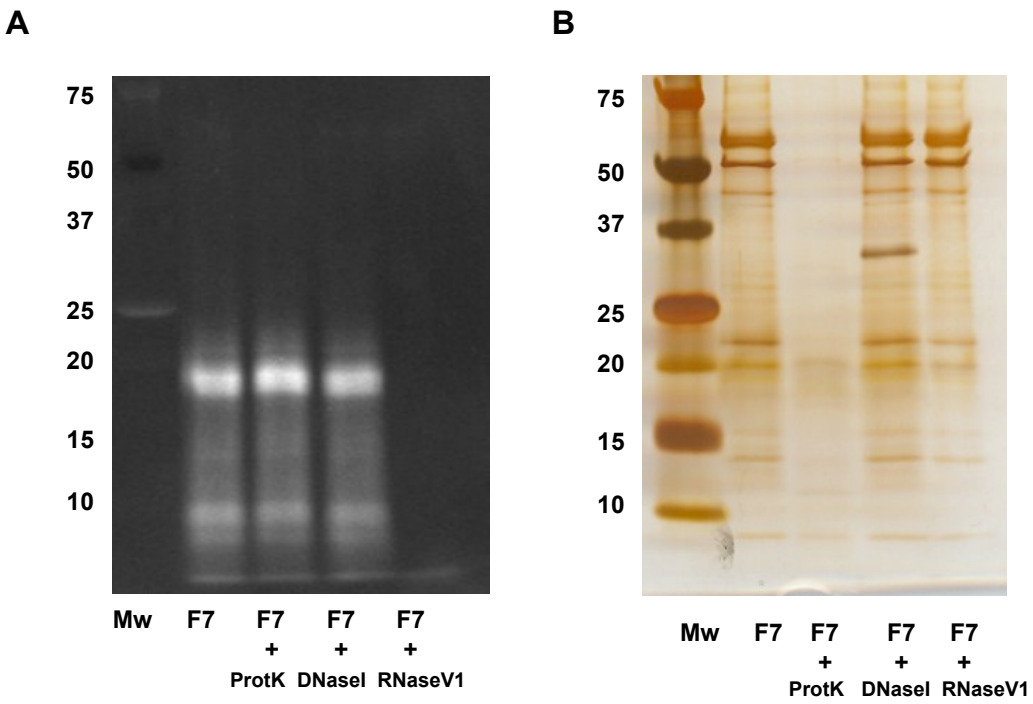


Figure 2.4 Fraction 7 is enriched with RNA. A. SDS-PAGE with ethidium bromide staining and B. silver staining of F7 and F7 treated with proteinase K, DNase1, or RNaseV1. By silver staining, it appeared like proteinase K treatment digested the majority of molecules present in fraction 7. Instead, ethidium bromide staining confirmed that RNA was present in significant concentrations in fraction 7, remaining after incubating with proteinase K or DNaseI but not with RNaseV1.

To obtain structural information about the *Mtb* RNA, it was subjected to enzymatic digestion with RNases targeting single or double stranded RNA. As observed in Figure 2.5, the *Mtb* RNA was extremely susceptible to RNaseV1, which cleaves double stranded bases. In contrast, the *Mtb* RNA displayed intermediate to high resistance when enzymatically digested with RNaseA and RNaseT1 respectively, which cleave single stranded bases. Altogether, these results suggested that the *Mtb* RNA was predominantly double stranded.

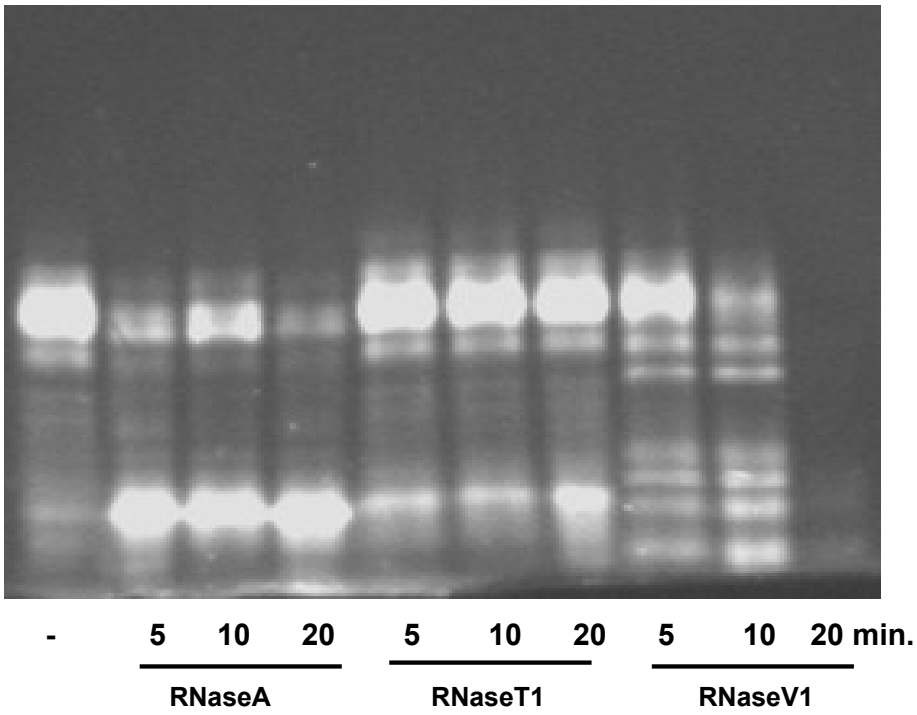


Figure 2.5 Treatment with RNaseV1 efficiently digested RNA present in fraction 7. Ethidium bromide staining of a SDS-PAGE, showed that enzymatic digestion with RNaseV1, led to rapid and complete digestion of the RNA present in fraction 7.

Furthermore, RNaseV1 treatment was additionally performed on CF and CF-Man to determine the RNA's contribution to monocyte cell membrane damage induced by these complex fractions. As determined by evaluating the samples pre- and post-RNaseV1 digestion, the RNA was responsible for almost 70 and 90 % of the biological activity from CF and CF-Man, respectively (Fig. 2.3).

2.3.2 Definition of Biologically Active RNA.

To confirm that RNA was responsible for the biological activity, fraction 7 was digested with DNaseI and Proteinase K, extracted with phenol/chloroform/isoamyl alcohol (25:24:1) to remove contaminating lipids and undigested proteins and the resulting RNA was gel purified from a denaturing urea-polyacrylamide gel (Fig. 2.6A).

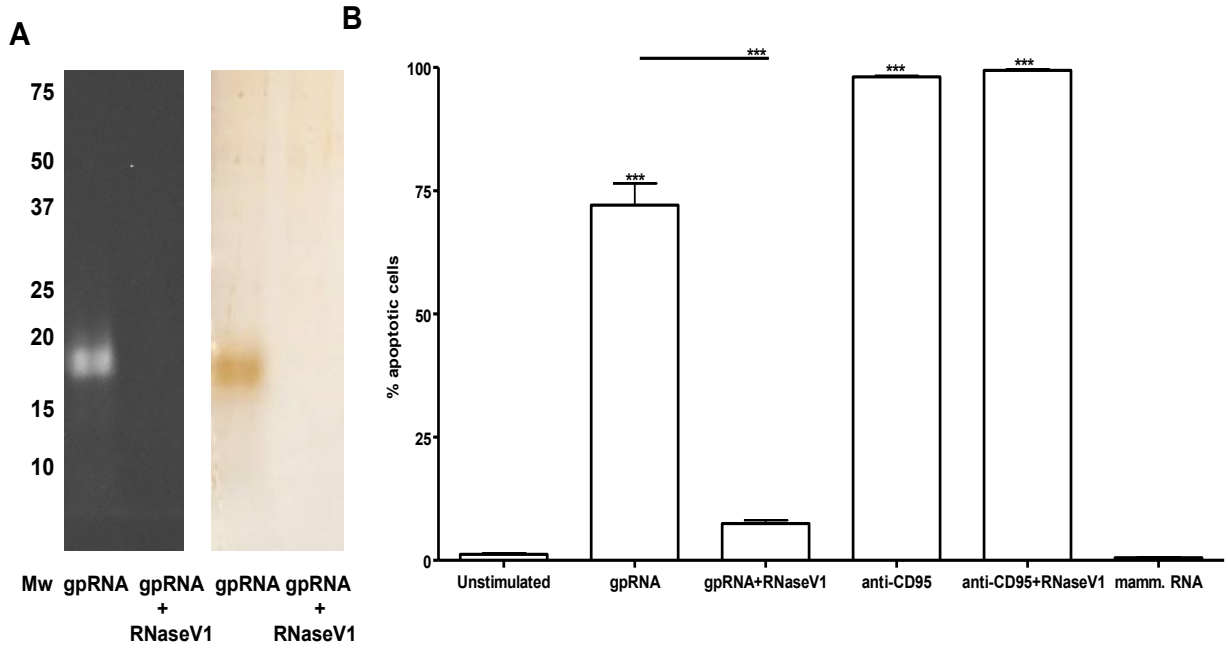


Figure 2.6 Human monocyte apoptosis is specifically induced by gel purified mycobacterial RNA. A. SDS-PAGE with ethidium bromide (left) and silver stained (right) RNA gel purified from F7 (gpRNA) and treated with RNaseV1 (gpRNA+RNaseV1). B. Monocyte apoptosis induced by gpRNA untreated or treated with RNaseV1, anti-CD95 untreated or treated with RNaseV1 (anti-CD95+RNaseV1) and rabbit mRNA. Significance as described above.

Gel purified RNA was subjected to RNaseV1 treatment and compared to untreated RNA in the biological assay with human monocytes. The apoptotic activity of the gel purified RNA was retained after purification (72% apoptotic monocytes), and this activity was abrogated by RNaseV1 treatment (7.5% apoptotic monocytes) (Fig. 2.6A and B). As a control to ensure RNaseV1 was not modulating monocyte apoptosis via a different mechanism, we evaluated cell membrane damage induced by the well characterized anti-CD95 pathway in the presence or absence of RNaseV1. No difference in AnnexinV positive cells was observed when monocytes were co-incubated with anti-CD95 and RNaseV1 versus anti-CD95 alone (Fig 2.6B). Finally, to evaluate whether monocyte cell membrane damage could be

induced by other sources of exogenous RNA, monocytes were incubated with rabbit mRNA; no induction of monocyte apoptosis was observed (Fig. 2.6B). These results confirmed that fraction 7 obtained from *Mtb* H37Rv CF is enriched in RNA, which was the sole inducer of human monocytes cell membrane damage.

To further define the RNA present in the CF the same strategy utilized to clone siRNA was adopted [28, 31]. From a library of cDNA clones plasmids were purified from a total of 33 clones. Sequence analyses of the plasmid inserts revealed that the *Mtb* RNA present in the CF predominantly consisted of tRNA and rRNA with lengths between 30 to 70 bases (Table 2.1).

Table 2.1 Sequence analysis of cloned extracellular mycobacterial RNA fragments present in *Mtb* CF

Cloned RNA	Number of Clones[2]
$tRNA^{Asp}$ [1]	10
$tRNA^{Asn}$	4
$tRNA^{Lys}$	1
$tRNA^{Thr}$	2
23S rRNA[1]	9
16S rRNA[1]	6
mRNA	1

[1]The specific regions of the RNA sequences represented by each clone are depicted in Figure 2.7
[2]The number of clones represents the total number of sequences observed for each RNA species

Of the 17 tRNA sequences, a majority corresponded to $tRNA^{Asp}$ fragments and a total of 15 clones represented 16S or 23S rRNA fragments, indicating selectivity in the RNA species released into the CF. Sequence alignments for both the 16S and 23S rRNA showed that the majority of fragments originated from different regions of the mature rRNA and only occasionally did two fragments correspond to the same region (Fig. 2.7A and B). Alignment of the cloned tRNA fragments determined that they were truncated at both the 5' and 3' ends (Fig. 2.7C).

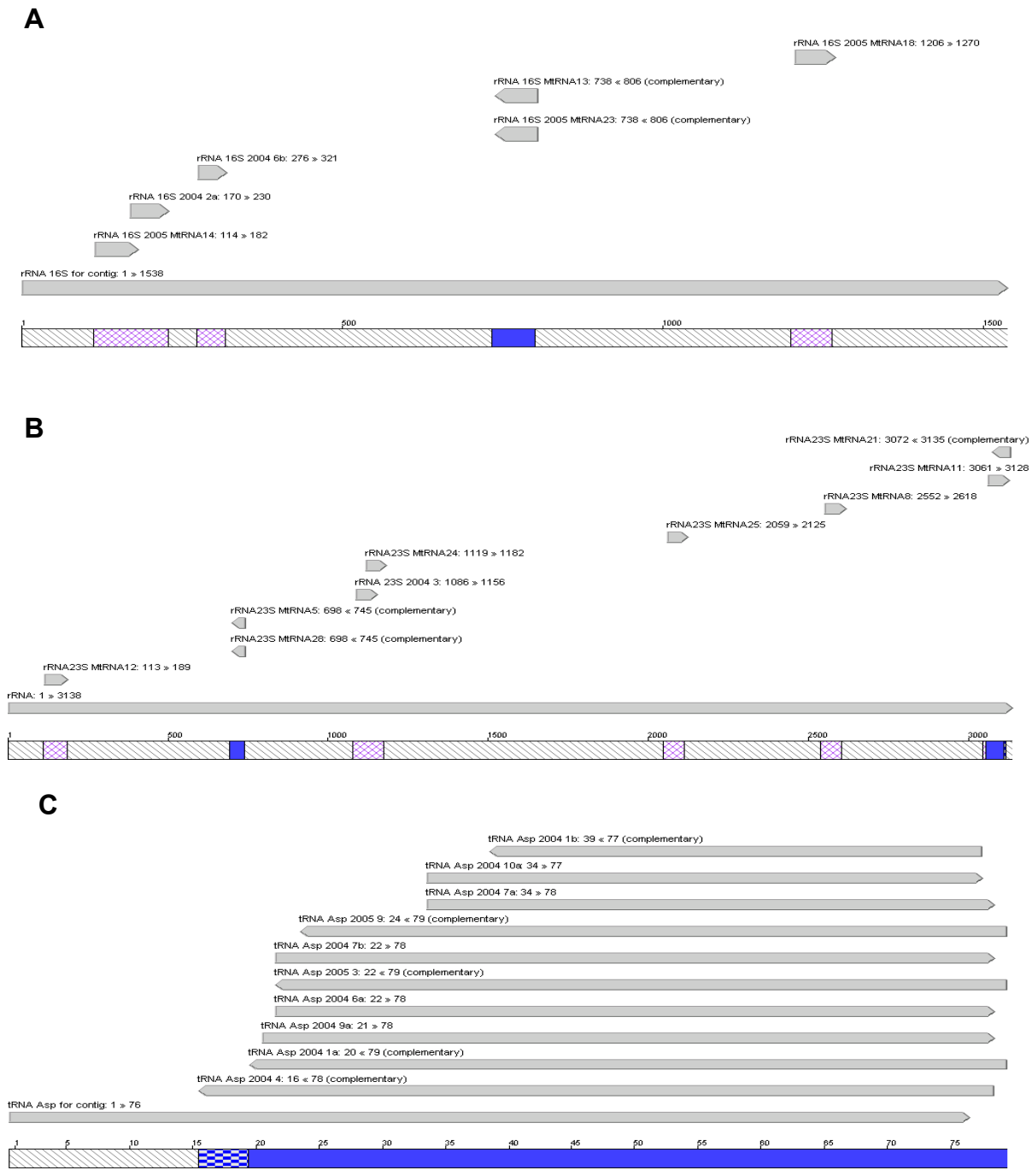


Figure 2.7 Contig alignments between sequenced RNA fragments and 16S rRNA, 23S rRNA or tRNA^{Asp}. Contigs were generated using Vector NTI software to align sequences from the cloned RNA fragments with the respective H37Rv sequences obtained from Tuberculist. Depicted below are the

contigs obtained for 16S rRNA (A), 23S rRNA (B) or $tRNA^{Asp}$ (C). For each sequenced fragment, numbers represent the nucleotide position in the respective gene.

2.3.3 Kinetics of RNA Release.

As tRNA and rRNA are highly abundant intracellular constituents, kinetic studies were performed to evaluate when the *Mtb* RNA started to accumulate in the CF and if its presence was attributable to the CF being harvested from late logarithmic (14 day), presumably autolytic, cultures . After extensively washing cells with GAS medium, CF was collected at different time points (0 to 28 days) and analyzed by SDS-PAGE and ethidium bromide staining. Interestingly, the presence of *Mtb* RNA paralleled *Mtb* protein secretion into the CF. Specifically, the RNA began to accumulate between two and three days of culture (early log phase), becoming more abundant as the culture progressed (Fig. 2.8).

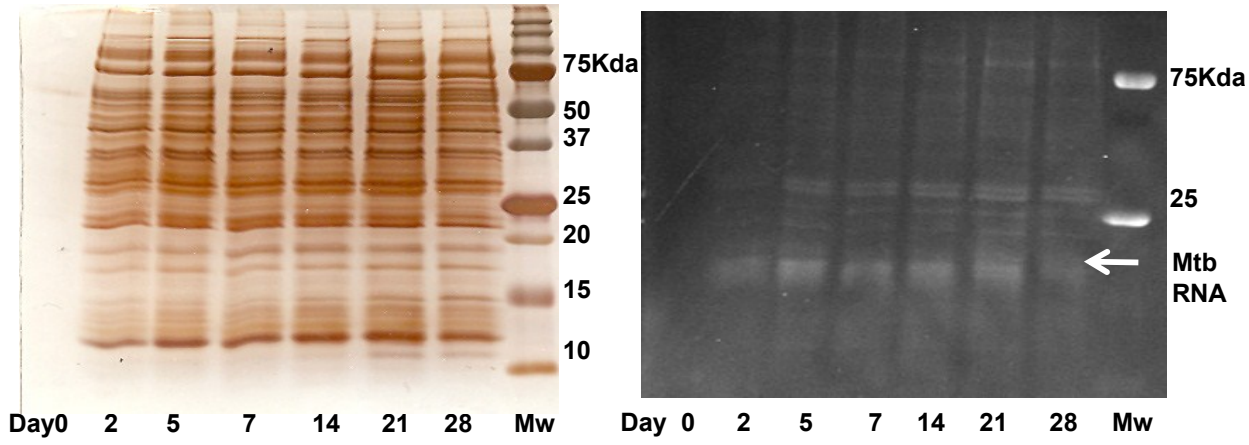


Figure 2.8 Extracellular mycobacterial RNA accumulate in the CF with similar kinetics as the rest of the mycobacterial secretome. Low passage *Mtb* H37Rv was thoroughly washed three times in GAS media and cultured in roller bottles. At different days (indicated below Figure) CF was harvested, filter

sterilized and concentrated. CF was analyzed by SDS-PAGE plus silver (left gel) or ethidium bromide (right gel) staining.

Similar results were obtained when the CF was immediately precipitated with acetone after harvesting, to rule out the possibility that higher molecular weight RNA species were being degraded to 30 to 70 bases during the protracted processing of CF (not shown). These results suggest that in addition to secreted proteins, *Mtb* H37Rv actively releases rRNA and tRNAs into the CF and these accumulate as stable products starting in early log phase.

2.3.4 Different *Mtb* strains also release RNA into the CF.

To rule out that RNA release by *Mtb* was not an isolated event occurring only in a laboratory strain, CF was generated from the *Mtb* clinical isolates HN878 and CDC1551. Additionally, H37Rv and H37RvΔRD1strains obtained from Albert Einstein Institute, Bronx, NY (kindly provided by Dr. Jacobs) were also evaluated to determine if RNA release was a characteristic of CSU-adapted strains. As can be observed in Figure 2.9, after 14 days of culture, CF obtained from all the strains contained RNA migrating at approximately 20 kDa (CDC1551 not shown). Thus it was concluded that RNA release into the CF is an inherent characteristic of *Mtb*.

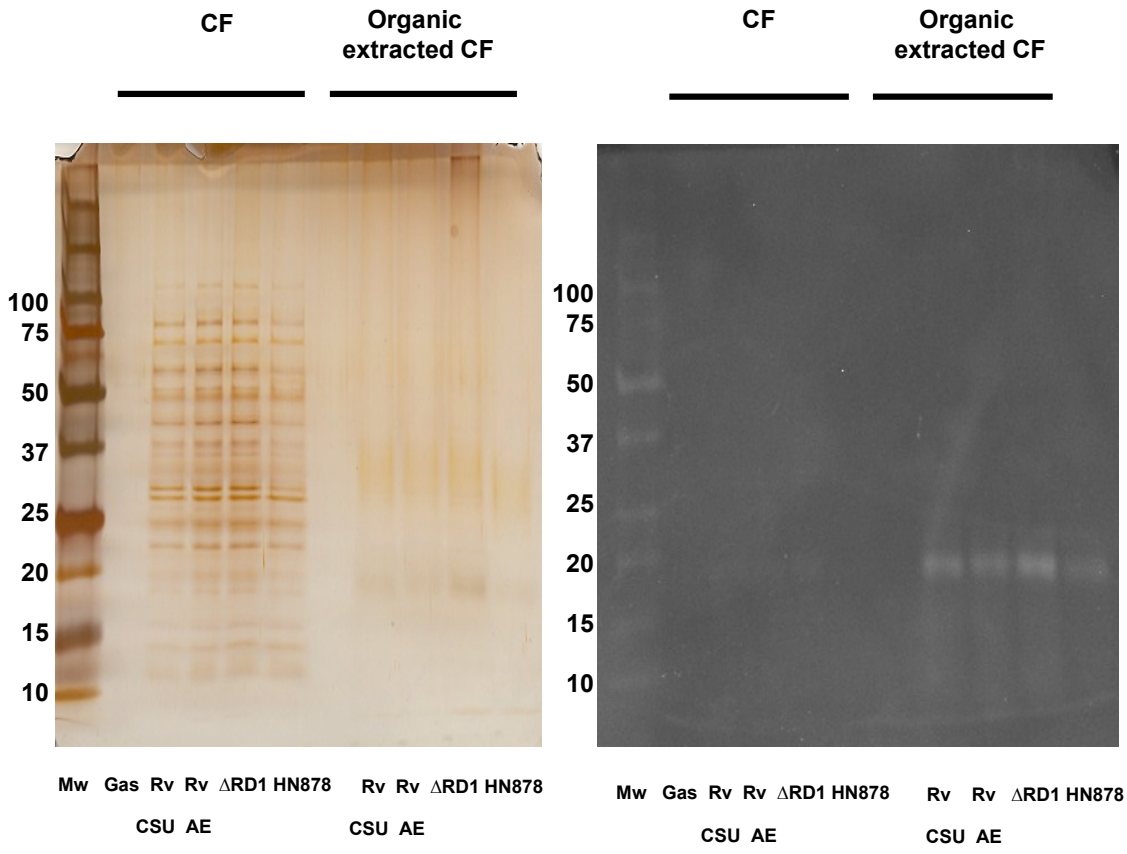


Figure 2.9 RNA is also released by other strains of *Mtb*. CF was generated from a clinical isolate (HN878), an H37Rv mutant (ΔRD1) and the lab strain H37Rv obtained from two different sources (CSU and Albert Einstein Institute, AE). Lyophilized CF was extracted or not with organic solvents and precipitated with ethanol. After resolving by SDS-PAGE and visualization by silver or ethidium bromide staining, RNA migrating <20 kDa was detected in the CF of all the strains

2.3.5 Mechanism of RNA Induced Apoptosis.

In order to assess the mechanism by which RNA induced apoptosis, TNF-α and IL-10 production was evaluated since the balance between these two cytokines is a factor in modulating cell death [12-14, 16]. *Mtb* RNA induced the intracellular production of these two cytokines to a level similar to that

induced by the positive control, PPD (Fig. 2.10A). However, in contrast to PPD, the cell membrane alterations induced by RNA were not blocked with anti-TNF-α (Fig. 2.10B).

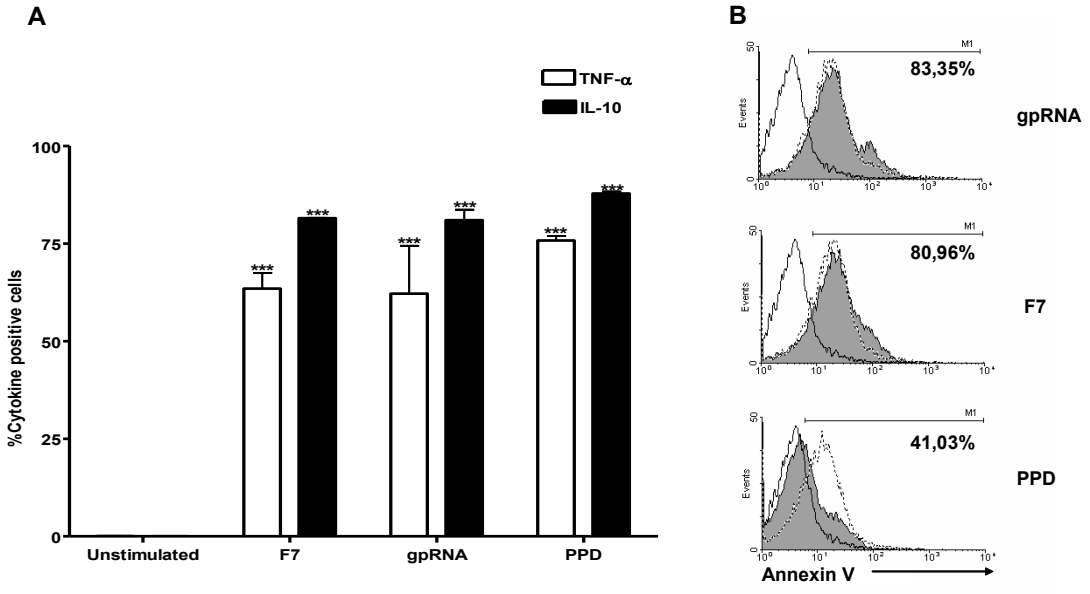


Figure 2.10 Human monocyte apoptosis induced by extracellular *Mtb* H37Rv RNA is TNF-α independent. A. TNF-α (open bars) and IL-10 (closed bars) induced in human monocytes treated with DEAE-Sepharose Fractions 7 (F7), gel purified *Mtb* H37Rv (gpRNA) and purified protein derivative (PPD) from *Mtb* were measured by flow cytometry. B. Treatment of monocytes with anti-TNF-α for 1 h prior to stimulation with F7, gpRNA, or PPD demonstrated no alteration in annexinV staining for monocytes incubated with F7 or gpRNA, but did result in decreased annexinV staining for monocytes incubated with PPD (closed histogram). The open histograms with a solid line correspond to unstimulated monocytes.

Given the importance of caspase activity in TNF-α dependent and independent apoptotic pathways [1, 4], and recent reports that bacterial RNA activates caspase-1 [32], experiments were

performed to determine whether *Mtb* H37Rv RNA activates caspases. *Mtb* RNA strongly activated caspase-8 (LETD-FMK), intermediately activated caspase-3 (DEVD-FMK) and had negligible effect on caspase-1 (YVAD-FMK), as determined by flow cytometry (Fig. 2.11A). In contrast, PPD activated caspase-1, 3 and 8 to similar extents. Consistent with these data, only caspase-8 inhibitor abrogated the RNA induced cell membrane alterations, whereas all the caspase inhibitors significantly abrogated PPD's activity (Fig. 2.11B). These findings provide evidence that *Mtb* RNA activates apoptosis via a caspase-8 signaling mechanism and that this route of activation differs from those observed for other mycobacterial products [12].

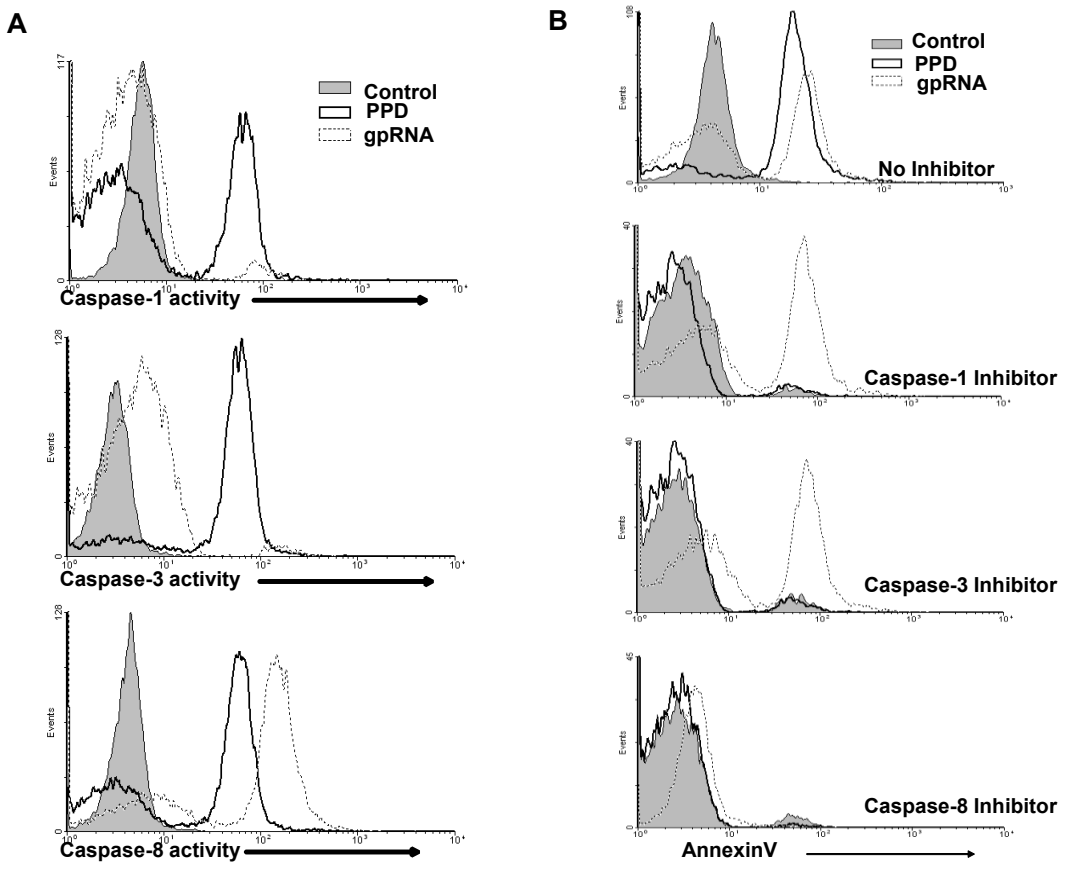


Figure 2.11 The *Mtb* RNA's activity is caspase-8 dependent. A. Flow cytometry demonstrated *Mtb* H37Rv RNA (gpRNA) activated caspase-8. Human monocytes were stimulated with 1 μg/ml of purified

Mtb H37Rv RNA gpRNA (open histogram with dotted line) or PPD (open histograms with solid line) for 48 h and caspase activity was determined by flow cytometry after staining with FLICA-YVAD-FMK (caspase-1), FLICA-DEVD-FMK (caspase-3), or FLICA-LETD-FMK (caspase-8). Tissue culture medium without RNA or PPD was used as the control (closed histogram). B. Stimulation of apoptosis based on annexinV staining was measured for monocytes stimulated with gpRNA or PPD (top histogram) and when the monocytes were pretreated for 1h with 10 nM caspase-1 (YVAD-FMK), caspase-3 (DEVD-FMK), or caspase-8 (LETD-FMK) inhibitors (lower histograms).

Lastly, the ability of RNA to alter monocyte's control of mycobacterial growth was examined by treating *Mtb* infected human monocytes with purified RNA. Determination of colony forming units (CFUs) after four days of incubation revealed a greater than two fold increase in the number of bacilli associated with RNA treated vs. untreated monocyte cultures (Fig. 2.12). Furthermore, digestion with RNaseV1 abrogated the RNA's deleterious effect on monocyte control.

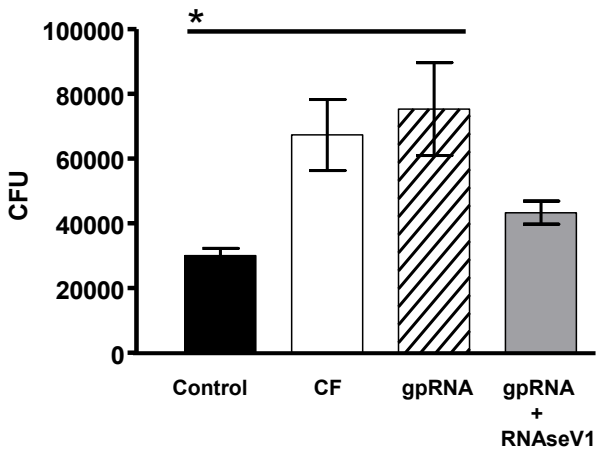


Figure 2.12 *Mtb* H37Rv RNA altered human monocyte's ability to control *Mtb* infection. CFUs were determined for human monocytes infected with *Mtb* H37Rv and incubated for four days in the presence of

1 µg/ml of *Mtb* H37Rv CF, purified RNA (gpRNA), or purified RNA digested with RNaseV1 (gpRNA + RNaseV1). The presence of CF or gpRNA resulted in a significant increase in CFUs as compared to the untreated infected monocytes (Control). Data represent the mean ± SEM of 3 replicates of the same experiment (*$p<0.03$).

2.4 Discussion.

Understanding apoptosis at a molecular level is critical to elucidating the pathogenesis of tuberculosis. Recent *in vivo* evidence suggests that apoptosis is an active mechanism in necrotic granulomas [8], a pathological hallmark of tuberculosis that is widely accepted as being key to the bacterium's persistence and transmission [33]. Our search for a *Mtb* product that replicated the apoptotic activity observed with a live *Mtb* infection [11], led to the surprising discoveries firstly that low molecular weight RNA fragments were secreted into the CF, long known as a source of major *Mtb* protein antigens, and secondly that these RNA fragments were major apoptotic factors. The role of mycobacterial RNA in the pathogenesis of tuberculosis was initially proposed by Youmans and colleagues in the 1960s and 70s [34, 35]. In their studies, ribosomal RNA obtained from whole cells increased bacterial burden when introduced into infected animals. Our experiments demonstrated that apoptosis induced by RNA significantly altered the ability of the monocytes to control *Mtb* growth. Although these studies did not assess *in vivo* growth of *Mtb* in a manner similar to Youmans and colleagues, they do provide support and a potential explanation for the early observations.

The activation of specific caspases is central to the induction of apoptosis and classically differentiates between the mitochondrial and apical pathways [1, 5]. The strong induction of caspase-8 activity with the *Mtb* RNA fragments and the prevention of apoptosis with a caspase-8 inhibitor provided convincing evidence for the induction of apoptosis via the apical pathway. The *Mtb* RNA fragments also stimulated the production of TNF-α and although TNF-α is primarily recognized as a mediator of inflammation, it can also induce apoptosis via caspase-8 when it binds to the TNF-R1 [4, 5]. However,

anti-TNF-α antibodies did not inhibit the induction of apoptosis by *Mtb* RNA fragments. Our previous work as well as that of others also indicated a potential role for caspase-1 in *Mtb* mediated apoptosis [14, 15]; however, the current studies demonstrated that *Mtb* RNA induction of apoptosis was caspase-1 independent. This lack of caspase-1 involvement also suggests that the *Mtb* RNA is not inducing an inflammasome response [36]. We did not, however, investigate whether the RNA fragments inhibited caspase-1 activity as has been recently reported to occur with *Mtb* infections [37]. In the apical pathway of apoptosis, caspase-3 is defined as the effector caspase and is downstream of caspase-8 [1, 5]. Our studies unexpectedly revealed that treatment of monocytes with *Mtb* RNA fragments only resulted in a modest increase in caspase-3 activity, as compared to that induced by PPD, and inhibition of caspase-3 failed to prevent *Mtb* RNA induced apoptosis. Thus, while the caspase-8 data suggest that the excreted *Mtb* RNA fragments stimulate monocyte apoptosis via the apical pathway, the non-essentiality of TNF-α or caspase-3 indicated an alternative mechanism of induction. One potential mechanism is that described for the "type II" apoptotic cells where capase-8 activation directly induces a mitochondrial pathway of apoptosis and is associated with FAS or TNF-related apoptosis-inducing ligand (TRAIL) mediated signaling [5]. Future work will assess the specific apoptotic pathway and signaling events stimulated by excreted *Mtb* RNA.

As evidenced by susceptibility to RNaseV1, the RNA fragments characterized in this study possess numerous dsRNA hairpin structures. Studies with other viral and bacterial dsRNAs demonstrate that interactions with TLR-8 or cryopyrin led to caspase activation [32, 38, 39]. Interactions between dsRNA and TLR-8 result in caspase-8 activity [38, 39] and TLR-8 polymorphisms are associated with increased susceptibility to tuberculosis [40]. Thus, activation of host monocytes via TLR8 would fit the data presented for the *Mtb* RNA fragments. In contrast, bacterial RNA interactions with cryopyrin activate caspase-1 [32], an event that did not occur with the *Mtb* RNA fragments. It must also be considered that tRNA and rRNA undergo multiple species specific modifications [41, 42] that could alter their ability to activate different pattern recognition receptors (PRRs) and signaling pathways. In contrast,

mammalian mRNA which is known to possess different modifications [42], did not activate monocyte apoptosis as the mycobacterial RNA fragments. Therefore, other PRRs for the *Mtb* RNA fragments need to be considered in future studies such as Protein Kinase R (PKR) that is activated in macrophages infected with *M. bovis* BCG [43], the endosomal TLR-3 and TLR-7 [44, 45], and cytoplasmic receptors Rig-1, MDA-5, DAI and Nod [46].

Multiple classes of pathogen derived macromolecules are defined as modulators of innate immunity, including RNA and DNA [47]. However bacterial nucleic acids typically are not described as being extracellular products. Thus, the isolation of the small RNA fragments from the CF of *Mtb* as the primary apoptotic factor was unexpected. The possibility that release of RNA fragments is due to lysis of *Mtb* was investigated by following the kinetics of RNA release. Such analyses revealed that accumulation of the secreted RNA fragments paralleled the kinetics of *Mtb* protein secretion. RNA was observed in the CF as early as 2 days, the same point in time used to distinguish truly secreted proteins from those resulting from autolysis [48]. Additionally, there did not appear to be a massive accumulation of extracellular RNA during late log-phase that would imply cell lysis as the underlying mechanism. Although a specific mechanism of active transport of the RNA fragments was not identified, there are suggestions by others that the ESX (Type VII) secretion apparatus of *Mycobacterium* spp. is involved in DNA transfer or conjugation [49]. Accumulation of mycobacterial RNA was also observed in the CF of H37RvΔRD-1, ruling out the specific ESX apparatus encoded by this region. Nevertheless, there are multiple ESX like secretion modules encoded on the *Mtb* genome [50] and investigation into whether one or more of these is involved in secretion of the RNA fragments is on-going. An additional, novel secretion mechanism that could be responsible for RNA release was recently unveiled in mycobacterial species [51]. It was demonstrated that mycobacteria release vesicles with potent immunomodulatory activity both *in vivo* and *in vitro*. Even though most of this activity was attributed to lipoprotein enrichment in these vesicles, it was also shown that cytoplasmic components such as DnaK (and possibly RNA) were also present in the vesicle lumen. Our current data does not completely rule out that early

bacterial lysis or cell division contributed to RNA release. It is well known that other intracellular molecules such as heat shock proteins, strongly activate the host immune system during *in vivo* mycobacterial infections [52, 53]. It is possible that early bacterial lysis allows for the release of significant quantities of mycobacterial proteins and RNA in host granulomas where it could induce host cell apoptosis leading to necrotic granulomas. It was originally expected that bacterial lysis only occurred during the chronic phase of *in vivo* infection. However, recent studies using plasmid loss as a "replication clock", revealed that bacterial lysis is an ongoing event during the log phase of bacterial replication [54]. Thus mycobacterial RNA could also be accumulating and modulating the environment as granulomas are developing, in order to favor bacterial survival. In this context, it was interesting to observe that mycobacterial RNA fragments affected the monocyte's ability to control infection despite stimulating TNF-α production. TNF-α production is required for both granuloma formation and bacterial control but alone is not sufficient [33]. The anti-inflammatory cytokine IL-10 which was also induced by the RNA fragments could be counteracting TNF-α. Alternatively, it was recently observed that intranasal delivery of the synthetic RNA poly-IC, increased bacterial counts in tuberculosis infected animals by stimulating type-I IFNs [55]. Even though we did not evaluate this family of cytokines, we obtained similar enhanced monocyte "permissiveness" to *Mtb* infection when stimulating with mycobacterial instead of a synthetic RNA. The stability of the *Mtb* RNA fragments must also be considered with respect to accumulation within the host. The complex secondary structure of the tRNA and rRNA might protect against host as well as bacterial RNase activity, and allow for its stable accumulation in the granuloma. Experiments with TLR-9 KO mice have suggested that mycobacterial DNA is being sensed by the host immune response [56], thus it is not unconceivable to expect a similar situation occurring with mycobacterial RNA.

A hallmark of *Mtb* is the bacterium's ability to interact with the host immune response in a manner that allows for its own survival and at the same causes active disease in only a relatively small percentage of the individual it infects. With this perspective it is not surprising that this pathogen has evolved a means to take advantage of the eukaryotic mechanism of programmed cell death. Likewise, the

involvement in this process of a bacterial product that is released or secreted only during active growth is not surprising. Thus, the continued elucidation of the exact signaling mechanisms exploited by *Mtb* to induce apoptosis as well as the process the bacterium utilizes to release RNA fragments will allow further definition of tuberculosis pathogenesis and the biology of this highly successful pathogen.

Literature Cited

1. Salvesen, G.S. and S.J. Riedl, *Caspase mechanisms.* Adv Exp Med Biol, 2008. **615**: p. 13-23.
2. Galluzzi, L., et al., *Cell death modalities: classification and pathophysiological implications.* Cell Death Differ, 2007. **14**(7): p. 1237-1243.
3. Vaculova, A. and B. Zhivotovsky, *Caspases: determination of their activities in apoptotic cells.* Methods Enzymol, 2008. **442**: p. 157-81.
4. Gupta, S., *Molecular steps of tumor necrosis factor receptor-mediated apoptosis.* Curr Mol Med, 2001. **1**(3): p. 317-24.
5. Guicciardi, M.E. and G.J. Gores, *Life and death by death receptors.* FASEB J, 2009. **23**(6): p. 1625-37.
6. Weinrauch, Y. and A. Zychlinsky, *THE INDUCTION OF APOPTOSIS BY BACTERIAL PATHOGENS.* Annual Review of Microbiology, 1999. **53**(1): p. 155-187.
7. Lee, J., M. Hartman, and H. Kornfeld, *Macrophage apoptosis in tuberculosis.* Yonsei Med J, 2009. **50**(1): p. 1-11.
8. Leong, A.S., P. Wannakrairot, and T.Y. Leong, *Apoptosis is a major cause of so-called "caseous necrosis" in mycobacterial granulomas in HIV-infected patients.* J Clin Pathol, 2008. **61**(3): p. 366-72.
9. Pan, H., et al., *Ipr1 gene mediates innate immunity to tuberculosis.* Nature, 2005. **434**(7034): p. 767-72.
10. Kramnik, I., *Genetic dissection of host resistance to Mycobacterium tuberculosis: the sst1 locus and the Ipr1 gene.* Curr Top Microbiol Immunol, 2008. **321**: p. 123-48.
11. Rojas, M., et al., *Differential induction of apoptosis by virulent Mycobacterium tuberculosis in resistant and susceptible murine macrophages: role of nitric oxide and mycobacterial products.* J Immunol, 1997. **159**(3): p. 1352-61.
12. Arcila, M.L., et al., *Activation of apoptosis, but not necrosis, during Mycobacterium tuberculosis infection correlated with decreased bacterial growth: role of TNF-alpha, IL-10, caspases and phospholipase A2.* Cell Immunol, 2007. **249**(2): p. 80-93.
13. Spira, A., et al., *Apoptosis genes in human alveolar macrophages infected with virulent or attenuated Mycobacterium tuberculosis: a pivotal role for tumor necrosis factor.* Am J Respir Cell Mol Biol, 2003. **29**(5): p. 545-51.
14. Rojas, M., et al., *TNF-alpha and IL-10 modulate the induction of apoptosis by virulent Mycobacterium tuberculosis in murine macrophages.* J Immunol, 1999. **162**(10): p. 6122-31.
15. Rojas, M., et al., *Mannosylated lipoarabinomannan antagonizes Mycobacterium tuberculosis-induced macrophage apoptosis by altering Ca+2-dependent cell signaling.* J Infect Dis, 2000. **182**(1): p. 240-51.
16. Balcewicz-Sablinska, M.K., H. Gan, and H.G. Remold, *Interleukin 10 produced by macrophages inoculated with Mycobacterium avium attenuates mycobacteria-induced apoptosis by reduction of TNF-alpha activity.* J Infect Dis, 1999. **180**(4): p. 1230-7.
17. Maiti, D., A. Bhattacharyya, and J. Basu, *Lipoarabinomannan from Mycobacterium tuberculosis promotes macrophage survival by phosphorylating Bad through a phosphatidylinositol 3-kinase/Akt pathway.* J Biol Chem, 2001. **276**(1): p. 329-33.
18. Nigou, J., et al., *Mycobacterial lipoarabinomannans: modulators of dendritic cell function and the apoptotic response.* Microbes Infect, 2002. **4**(9): p. 945-53.
19. Lopez, M., et al., *The 19-kDa Mycobacterium tuberculosis protein induces macrophage apoptosis through Toll-like receptor-2.* J Immunol, 2003. **170**(5): p. 2409-16.

20. Ciaramella, A., et al., *Mycobacterial 19-kDa lipoprotein mediates Mycobacterium tuberculosis-induced apoptosis in monocytes/macrophages at early stages of infection.* Cell Death Differ, 2000. **7**(12): p. 1270-2.
21. Thoma-Uszynski, S., et al., *Induction of direct antimicrobial activity through mammalian toll-like receptors.* Science, 2001. **291**(5508): p. 1544-7.
22. Brightbill, H.D., et al., *Host defense mechanisms triggered by microbial lipoproteins through toll-like receptors.* Science, 1999. **285**(5428): p. 732-6.
23. Ozeki, Y., et al., *In vivo induction of apoptosis in the thymus by administration of mycobacterial cord factor (trehalose 6,6'-dimycolate).* Infect Immun, 1997. **65**(5): p. 1793-9.
24. Nuzzo, I., et al., *Apoptosis modulation by mycolic acid, tuberculostearic acid and trehalose 6,6'-dimycolate.* J Infect, 2002. **44**(4): p. 229-35.
25. Dao, D.N., et al., *Mycobacterium tuberculosis lipomannan induces apoptosis and interleukin-12 production in macrophages.* Infect Immun, 2004. **72**(4): p. 2067-74.
26. Sonnenberg, M.G. and J.T. Belisle, *Definition of Mycobacterium tuberculosis culture filtrate proteins by two-dimensional polyacrylamide gel electrophoresis, N-terminal amino acid sequencing, and electrospray mass spectrometry.* Infect Immun, 1997. **65**(11): p. 4515-24.
27. Dobos, K.M., et al., *Evidence for glycosylation sites on the 45-kilodalton glycoprotein of Mycobacterium tuberculosis.* Infect Immun, 1995. **63**(8): p. 2846-53.
28. Elbashir, S.M., W. Lendeckel, and T. Tuschl, *RNA interference is mediated by 21- and 22-nucleotide RNAs.* Genes Dev, 2001. **15**(2): p. 188-200.
29. Kaur, D., et al., *Characterization of the epitope of anti-lipoarabinomannan antibodies as the terminal hexaarabinofuranosyl motif of mycobacterial arabinans.* Microbiology, 2002. **148**(Pt 10): p. 3049-57.
30. Laemmli, U.K., *Cleavage of structural proteins during the assembly of the head of bacteriophage T4.* Nature, 1970. **227**(5259): p. 680-5.
31. Lau, N.C., et al., *An abundant class of tiny RNAs with probable regulatory roles in Caenorhabditis elegans.* Science, 2001. **294**(5543): p. 858-62.
32. Kanneganti, T.D., et al., *Bacterial RNA and small antiviral compounds activate caspase-1 through cryopyrin/Nalp3.* Nature, 2006. **440**(7081): p. 233-6.
33. Saunders, B.M. and W.J. Britton, *Life and death in the granuloma: immunopathology of tuberculosis.* Immunol Cell Biol, 2007. **85**(2): p. 103-11.
34. Youmans, A.S., *Biological Activities of Mycobacterial Ribosomal and RNA Vaccines.*, in *Tuberculosis*, G.P. Youmans, Editor. 1979, W.B. Saunders Company: Philadelphia. p. 236-279.
35. Youmans, G.P., *The Nature of the Immunizing Activity of Mycobacterial and other Ribosomal and RNA Vaccines*, in *Tuberculosis*, G.P. Youmans, Editor. 1979, W.B. Saunders Company: Philadelphia. p. 268-277.
36. Martinon, F., A. Mayor, and J. Tschopp, *The inflammasomes: guardians of the body.* Annu Rev Immunol, 2009. **27**: p. 229-65.
37. Master, S.S., et al., *Mycobacterium tuberculosis prevents inflammasome activation.* Cell Host Microbe, 2008. **3**(4): p. 224-32.
38. Takahashi, K., et al., *Roles of caspase-8 and caspase-10 in innate immune responses to double-stranded RNA.* J Immunol, 2006. **176**(8): p. 4520-4.
39. Beisner, D.R., et al., *Cutting edge: innate immunity conferred by B cells is regulated by caspase-8.* J Immunol, 2005. **175**(6): p. 3469-73.
40. Davila, S., et al., *Genetic association and expression studies indicate a role of toll-like receptor 8 in pulmonary tuberculosis.* PLoS Genet, 2008. **4**(10): p. e1000218.
41. McCloskey, J.A. and J. Rozenski, *The Small Subunit rRNA Modification Database.* Nucleic Acids Res, 2005. **33**(Database issue): p. D135-8.
42. Kariko, K., et al., *Suppression of RNA recognition by Toll-like receptors: the impact of nucleoside modification and the evolutionary origin of RNA.* Immunity, 2005. **23**(2): p. 165-75.

43. Cheung, B.K., et al., *A role for double-stranded RNA-activated protein kinase PKR in Mycobacterium-induced cytokine expression.* J Immunol, 2005. **175**(11): p. 7218-25.
44. Heil, F., et al., *Species-specific recognition of single-stranded RNA via toll-like receptor 7 and 8.* Science, 2004. **303**(5663): p. 1526-9.
45. Alexopoulou, L., et al., *Recognition of double-stranded RNA and activation of NF-kappaB by Toll-like receptor 3.* Nature, 2001. **413**(6857): p. 732-8.
46. Meylan, E. and J. Tschopp, *Toll-like receptors and RNA helicases: two parallel ways to trigger antiviral responses.* Mol Cell, 2006. **22**(5): p. 561-9.
47. Kumagai, Y., O. Takeuchi, and S. Akira, *Pathogen recognition by innate receptors.* J Infect Chemother, 2008. **14**(2): p. 86-92.
48. Wiker, H.G., M. Harboe, and S. Nagai, *A localization index for distinction between extracellular and intracellular antigens of Mycobacterium tuberculosis.* J Gen Microbiol, 1991. **137**(4): p. 875-84.
49. Flint, J.L., et al., *The RD1 virulence locus of Mycobacterium tuberculosis regulates DNA transfer in Mycobacterium smegmatis.* Proc Natl Acad Sci U S A, 2004. **101**(34): p. 12598-603.
50. Simeone, R., D. Bottai, and R. Brosch, *ESX/type VII secretion systems and their role in host-pathogen interaction.* Curr Opin Microbiol, 2009. **12**(1): p. 4-10.
51. Prados-Rosales, R., et al., *Mycobacteria release active membrane vesicles that modulate immune responses in a TLR2-dependent manner in mice.* J Clin Invest. **121**(4): p. 1471-83.
52. Silva, C.L., *The potential use of heat-shock proteins to vaccinate against mycobacterial infections.* Microbes Infect, 1999. **1**(6): p. 429-35.
53. Walker, K.B., J. Keeble, and C. Colaco, *Mycobacterial heat shock proteins as vaccines - a model of facilitated antigen presentation.* Curr Mol Med, 2007. **7**(4): p. 339-50.
54. Gill, W.P., et al., *A replication clock for Mycobacterium tuberculosis.* Nat Med, 2009. **15**(2): p. 211-4.
55. Antonelli, L.R., et al., *Intranasal Poly-IC treatment exacerbates tuberculosis in mice through the pulmonary recruitment of a pathogen-permissive monocyte/macrophage population.* J Clin Invest. **120**(5): p. 1674-82.
56. Bafica, A., et al., *TLR9 regulates Th1 responses and cooperates with TLR2 in mediating optimal resistance to Mycobacterium tuberculosis.* J Exp Med, 2005. **202**(12): p. 1715-24.

Chapter III

Characterization of Rv2394, a homologue of the γ-glutamyltranspeptidase CapD

3.1 Introduction.

In higher eukaryotes, γ-glutamyltranspeptidases (GGTs) were initially described to participate in the glutathione (GSH) cycle, a term coined by Meister [1-3]. GSH is a tri-peptide consisting of a Glu, cysteine (Cys) and glycine (Gly) residues. In contrast to the conventional peptide link between the alpha-carboxyl and amino groups of two adjacent amino acids, the Glu and Cys residues of GSH are covalently linked via a γ-linkage between a Glu γ-carboxyl group and a Cys amino group. In turn, the Cys is linked to the Gly residue by a conventional peptide bond. GSH is critical for maintaining intracellular redox homeostasis due to the reducing power of Cys thiol group [3]. GSH is also released extracellularly where it participates in detoxification of endogenous products such as leukotrienes or foreign products such as xenobiotics. Detoxification begins with the enzyme glutathione-S-transferase which covalently links GSH's free thiol group to a wide array of compounds. After GGT removes the γ-linked Glu residue, dipeptidases cleave the C-terminus Gly and finally the free N-terminus of the thiol-linked Cys is acetylated [3]. Mercapturic acids are the end products of this detoxification reaction and are readily excreted by the body. Any remaining extracellular GSH that is not used for detoxification is recovered through a process that also starts by the GGT-mediated cleavage of the γ-linked Glu residue. This is of upmost importance specially for recovering Cys which is in limited intracellular concentration [1].

Several different functions have been attributed to conserved, extracellular bacterial and lower eukaryote GGTs. In *E. coli* it was shown by a simple yet elegant study that several amino acid auxotrophs could survive in medium supplemented with a di-peptide of Glu linked via its γ-carboxyl group to the essential amino acid [4]. In contrast, auxotrophs in which the extracellular GGT had been mutated did not thrive under these same conditions suggesting that uptake of the essential amino acid was dependant on the GGT-mediated di-peptide hydrolysis [4]. In *H. pylori*, an extracellular GGT was shown to participate

in Glu uptake when presented with substrates such as GSH and glutamine (Gln) [5]. *H. pylori* does not seem to encode for a Gln transporter and instead has a functional Na^+-dependant Glu transporter. Thus, when using Gln as the Glu source, the GGT performs a hydrolysis reaction generating Glu and ammonia [5]. In contrast, *Saccharomyces cerevisiae* and *Francisella tularensis* (*F. tularensis*) GGT participate in the GSH cycle required for sulfur and Cys uptake, thus resembling the role of GGT in higher eukaryotes [6, 7]. Finally, two novel roles were recently attributed to extracellular GGTs in *H. pylori* and *H. capsulatum*. *H. pylori* GGT was identified as a protein inducing gastric epithelial cell apoptosis [8, 9]. This was confirmed by showing that recombinant *H. pylori* GGT expressed in *E. coli* recapitulated the native GGT's apoptosis-inducing activity. Furthermore, gastric epithelial cell apoptosis was attributed to the GGT-mediated catabolism of GSH into Glu and Cys-Gly [9]. This latter product is highly reactive and led to increased oxidative stress of gastric epithelial cells. Similarly, the increased reactivity of Cys-Gly generated by *H. capsulatum* GGT was harnessed to reduce insoluble ferric ion to the more soluble and acquirable ferrous ion [10]. Therefore, it can be concluded that despite sharing significant homology between different species, GGTs are rather promiscuous in terms of substrate specificity and function. In essence, GGT can accomplish multiple functions by targeting the γ-linkage between Glu and other functional groups as diverse as ammonia (in Gln), amino acids (as in the auxotrophs), peptides (GSH) or any GSH-modified compound (lipids and xenobiotics) targeted for excretion as mercapturic acids.

A completely different role was identified for CapD, a GGT present in *Bacillus anthracis* (*B. anthracis*). In addition to the hydrolysis reactions described above, CapD was shown to mediate a significantly more important transpeptidase reaction participating in *B. anthracis* capsule biosynthesis [11]. *B. anthracis* capsule consists of poly-γ-D-Glu, a polymer of D-Glu residues linked through a γ-linkage [11-14]. CapD was shown to link poly-γ-D-Glu to diamino pimelic acid residues (DAP) present in *B. anthracis* cell wall [11]. Furthermore, it was suggested that CapD could participate in the elongation of poly-γ-D-Glu by crosslinking γ-D-Glu oligopeptides. Interestingly, *Mtb* has been shown to synthesize a similar cell wall product consisting of a polymer of L-Gln residues in alpha linkages (see Chapter IV for

further discussion of *Mtb* α-L-poly-Gln) [15]. In line with CapD's role in *B. anthracis* poly-γ-D-Glu synthesis, it was reasoned that *Mtb* could have a homologous GGT participating in α-L-poly-Gln's biosynthesis. Indeed, bio-informatic analysis confirmed that two *capD* homologues are present in *Mtb* genome (*rv2394* and *rv0773c*). The experimentation presented in this Chapter provides several lines of evidence to demonstrate the function of Rv2394 as a CapD homologue. This represents the first biochemical characterization of Rv2394, as a *Mtb* GGT with enzymatic activity that could participate in the biosynthesis of α-L-poly-Gln by crosslinking α-L-Gln oligopeptides through a γ-linkage.

3.2 Materials and Methods.

3.2.1 Bioinformatic analyses.

The amino acid sequence of Rv2394 was obtained from Tuberculist webpage and was analyzed using the following bioinformatic analyses tools: BLAST® from NCBI (http://blast.ncbi.nlm.nih.gov/), SignalP (http://www.cbs.dtu.dk/services/SignalP/), LipoP (http://www.cbs.dtu.dk/services/LipoP/) and NetOGlyc (http://www.cbs.dtu.dk/services/NetOGlyc/) available at the Expasy Proteomics Server website. Multiple sequence alignments were performed using Invitrogen VectorNTI (Invitrogen, Carlsbad, CA). Genome organization for the different genes was performed using the genome region comparison tool at JCVI comprehensive microbial resource webpage.

3.2.2 PCR and cloning.

rv2394 was amplified by PCR using Platinum *Pfx* DNA Polymerase (Invitrogen). Briefly, 100 ng of *Mtb* H37Rv genomic DNA was incubated with 1X Amplification Buffer, 0.3 mM dNTPs, 1 mM $MgSO_4$, 0.3 µM forward and reverse primers (see Table 3.1), and 1 U of *Pfx* DNA Polymerase in a final volume of 50 µl. The thermocycler was programmed to amplify the DNA by a two-step cycling as follows: initial DNA denaturation for 5 minutes at 95°C, followed by 30 cycles of 1 min at 95°C and an elongation for 2 min at 68°C. The amplified reaction was run in a 1% agarose gel in 1X TAE at 100 volts

for 1.5 h. The gel was stained with SYBR green 1/10,000 in 1X TAE for 30 min, visualized under UV light using GelDoc System (BioRad, Hercules, CA), and bands of expected size were excised and extracted using Qiagen Gel Extraction kit (Qiagen, Valencia, CA) in a final volume of 30 µl of EB buffer.

Table 3.1 PCR Primers for Rv2394 constructs

Primer	Sequence
For Rv2394[1]	CAT ATG AGT GTT TGG TTG CGA GCG GG
Rev Rv2394[2]	AAG CTT GGC ATC GTC GCC CAT GAC CGC
For Rv2394 TA[3]	CCG CCA GTG CCT GAG CAT GGC GCC AGC CAC CTC AGC GTC GTC GAT TCG
Rev Rv2394 TA	CGA ATC GAC GAC GCT GAG GTG GCT GGC GCC ATG CTC AGG CAC TGG CGG
For Rv2394 TS[4]	CCG CCA GTG CCT GAG CAT GGC AGC AGC CAC CTC AGC GTC GTC GAT TCG
Rev Rv2394 TS	CGA ATC GAC GAC GCT GAG GTG GCT GCT GCC ATG CTC AGG CAC TGG CGG
For Rv2394 SP-I[5]	GGT GGC TGC AGT GAT GCT GTC GGC GAA CGC AGA TCC CGG CTT CCA CGC GGG TGC GCC
Rev Rv2394 SP-I	GGC GCA CCC GCG TGG AAG CCG GGA TCT GCG TTC GCC GAC AGC ATC ACT GCA GCC ACC
For Rv2394 -20SP[1,6]	CAT ATG TGT GGC GGC TTC CAC GCG GG

[1] *NdeI* site is underlined.
[2] *HindIII* site is underlined.
[3] Thr to Ala substitution (TA). Mutated base is underlined.
[4] Thr to Ser substitution (TS). Mutated base is underlined.
[5] Enforced signal peptidase-I mediated secretion (SP-I). Mutated bases are underlined.
[6] Without the initial twenty amino acids encoding for the signal peptide (-20SP).

A' overhangs were incorporated into the gel purified *rv2394* PCR product with Taq polymerase (GE Health Systems, Waukesha, WI) for 60 min at 72°C and subsequently introduced into the pGEM-T Easy cloning system (Promega, Madison, WI) following the manufacture's protocol. This cloning system was used to clone all PCR products. After incubating overnight at 4°C, 1 µl of the cloning reaction was used to transform chemically competent TOP10 *E. coli* cells (Invitrogen) following the manufacture's recommendations. Bacteria were plated on LB agar containing 100 µg/ml of ampicillin and incubated overnight at 37°C.

Single colonies were picked up and grown overnight at 37°C in LB medium containing 100 µg/ml of ampicillin with constant agitation at 200 rpm. Plasmids were then purified from *E. coli* cells

using the Qiagen Mini Prep Isolation kit. Plasmids were subjected to restriction enzyme digestion in order to confirm the inserted PCR product. Briefly, 4 µl of each plasmid were incubated at 37°C for 3 h with 0.5 µl of 10X buffer and 0.25 µl of *NdeI* and *HindIII* (New England Biolabs, Ipswich, MA). Digested plasmids were analyzed by agarose gel electrophoresis and SYBR green staining as described above. Plasmids containing inserts of expected molecular weight were submitted for sequencing at Macromolecular Lab Resources, Colorado State University. Sequences were analyzed using VectorNTI (Invitrogen). Table 3.2 describes the plasmids that were engineered for this Chapter.

Table 3.2 Plasmid names and description

Plasmids	Description
pVV16	**Shuttle plasmid for protein expression in *mycobacterium* spps. Contains a His-Tag at the carboxyl terminus.**
pVV16.2394	**Expression of WT Rv2394.**
pVV16.2394TS	**Expression of a mutant Rv2394 in which the Thr residue at position 446 is replaced by a Ser.**
pVV16.2394TA	**Expression of a mutant Rv2394 in which the Thr residue at position 446 is replaced by an Ala.**
pVV16.2394-20SP	**Expression of a mutant Rv2394 lacking the signal peptide corresponding to the first 20 amino acids.**
pVV16.2394-20SP TA	**Expression of a double mutant Rv2394 lacking the signal peptide and the Thr residue at position 446 is replaced by an Ala.**
pVV16.2394 SP-I	**Expression of a mutant Rv2394 in which the acylated cysteine is replaced in order to inhibit lipoprotein processing. Engineered so that the signal peptide of Rv2394 is removed by SP-I as opposed to SP-II.**
pVV16.2394 SP-I TA	**Similar to pVV16.2394 SP-I with an additional mutation consisting of replacing Thr residue at position 446 with Ala.**

For protein expression, the shuttle vector pVV16 was obtained from the NIH, NIAID N01-AI-40091 program at Colorado State University. To insert *rv2394* or the engineered mutants into pVV16, 40 µl of both pVV16 and the pGEM-T Easy plasmid containing the construct were simultaneously digested for 4 h at 37°C with 5 µl of 10X buffer and 2.5 µl of *NdeI* and *HindIII* (New England Biolabs). Digested

pVV16 and cloned gene were purified using a Gel Extraction Kit (Qiagen) and ligated O/N at 14ºC with 400 U of T4 DNA ligase (New England Biolabs). Thereafter, ligated expression vectors were transformed into chemically competent *E. coli* TOP10 cells and transformed bacteria were plated on LB agar plates with 50 and 100 µg/ml of kanamycin and hygromycin, respectively. Plates were incubated O/N at 37ºC, colonies and plasmids were analyzed as described above. Once again, sequencing was performed to corroborate that no mutations had been introduced during the cloning process.

3.2.3 Site Directed Mutagenesis.

Site directed mutagenesis of *rv2394* was accomplished following the SiteDirected Mutagenesis Kit (Stratagene, La Jolla, CA) protocol with minor variations. Briefly, 100 ng of the plasmid to be mutated was incubated with 1X Reaction Buffer, 125 ng of each mutated primer (see Table 3.1), 1 µl of dNTPs and 2.5 U of *Pfu* Turbo DNA polymerase in a final volume of 50 µl. The desired mutation was introduced by performing the following PCR: an initial DNA denaturing step of 30 sec at 95°C, followed by 18 cycles of 30 sec at 95°C and elongation for 8 min at 68°C. Thereafter, 10 U of *DpnI* were added to the PCR product and incubated for 1h at 37°C to digest the initial unmutated plasmid. Finally, the mutated plasmid was transformed into XL-1 Blue Supercompetent *E. coli* as described in the protocol and selected on LB agar plates containing 100 µg/ml of ampicillin. See Table 3.2 for a description of the Rv2394 mutants that were engineered for this Chapter.

3.2.4 Generation and Transformation of Electrocompetent *M. smeg* mc^2 155.

M. smeg mc^2 155 was grown at 37°C in 7H9 medium (BD Difco) supplemented with 0.05% Tween 80 until reaching an OD between 0.4-0.6 at an absorbance of 600 nm. Bacteria were harvested by centrifugation at 3000 x g for 10 min at 4°C. After decanting the supernatant, bacteria were washed twice with ice cold water by inverting the cells and centrifuging as before. Finally, *M. smeg* was washed with ice cold 10% glycerol in water by centrifuging as described above. After decanting, bacteria were

resuspended in 5 ml of 10% glycerol in water and 500 µl aliquots were stored at -80°C [16]. Electrocompetent *M. smeg* was thawed in ice and 60µl aliquots were incubated with 100 ng of plasmid for 5-10 min in ice. Cells were then added to 340 µl of water that had been dispensed into 0.2 mm electroporation cuvettes (BioRad). Electroporation was performed at 2000 Volts. Thereafter, cells were recovered from the electroporation cuvettes, added to 600 µl of 7H9 medium and incubated for 4 h at 37°C. Bacteria were then recovered by centrifugation at 3000 x g for 5 min, and after decanting, were plated on 7H11 agar plates containing 50 and 100 µg/ml kanamycin and hygromycin, respectively. Plates were incubated at 37°C for 5-7 days and individual colonies were used to inoculate 5 ml of 7H9 liquid medium with the same antibiotics as the selection plates. After growing bacteria for 5 days at 37°C on an orbital shaker at 100 rpm, 1 ml of bacteria was harvested to determine recombinant protein production as indicated below (Western blotting section). Colonies producing a His-Tag reactive band corresponding to the protein's predicted molecular weight, were harvested by centrifugation as described above and inoculated in 100 ml of 7H9 plus 50 µg/ml Kanamycin. After 5 days at 37°C on an orbital shaker, cells were harvested and used to inoculate Fernbacks containing 900 ml of 7H9 medium plus 50 µg/ml Kanamycin. Finally, bacteria were harvested by centrifugation and pellets were stored at -80°C until performing protein purification.

3.2.5 Protein Purification.

Bacterial pellets of recombinant *M. smeg* were thawed in ice and washed once with binding buffer (500 mM sodium chloride, 10 mM Tris-Hydrochloride, 5 mM Imidazole, pH 7.9). Binding buffer containing 1% TritonX-100 [17] and 1 tablet of Complete EDTA-free protease inhibitors (Roche Applied Sciences, Mannheim, Germany) in 50 ml of buffer, was added to bacterial pellets and cells were sonicated in ice for 5 cycles of 1 min intermittent pulsing in a Vibra Cell VCX750 Sonicator (Sonicator and Materials, Inc, Newton, CT). Bacterial slurries were then passed 7 times through a French Press at 1800 psi, in order to fully break the bacteria. After removing unbroken bacteria by centrifuging at 3,000 rpm

for 5 min, supernatants were clarified by centrifuging at 4°C for 30 min at 27,000 x g. Thereafter, clarified supernatants were subjected to nickel affinity chromatography by binding to 1 ml of charged HisBind resin (EMD Chemicals, Gibbstown, NJ), at 4°C using gravity. The flow through fraction was collected and after washing with 10 column volumes of binding buffer containing 1% Triton X-100 [17], elution was performed with binding buffer containing a stepwise increase of 10, 15, 20, 30, 40, 60 and 1000 mM imidazole. Fractions were then analyzed by SDS-PAGE and silver staining (see below) to determine which fractions contained protein of the expected molecular weight, with the least amount of contaminating bands. These fractions were pooled and concentrated by centrifuging at 3500 x g, in an Amicon with 10 kDa molecular weight cutoff (Millipore, Billerica, MA). Finally, the concentrated fraction was dialyzed in 10 mM ammonium bicarbonate and stored at 4°C. The concentration of purified proteins was determined by the BCA protein Assay (Pierce, Rockford, IL) using the manufacturer's procedure. After incubating at 37°C for 30 min, optical density was determined at a wavelength of 500 nm using a Bio-Rad plate reader. Protein concentration was determined by extrapolating the optical density of samples with the optical density of a standard curve containing known concentrations of bovine serum albumin (BSA) (Pierce).

3.2.6 Western Blotting.

One µg of purified protein was resolved by SDS-PAGE in 4-12% Tris-Glycine gels (Invitrogen) by running in 1X MES buffer (Invitrogen) at 200 Volts for 45 min. Thereafter, gels were either silver stained as described before [18] or transferred to a nitrocellulose membrane (Bio-Rad) at 50 Volts for 1 h.

For Western blotting, membranes to be probed with antibodies were blocked at RT for 1 h with 1% BSA in 1X TBS, whereas membranes to be probed with ConA were blocked with 1 % BSA diluted in PBS [19]. BSA-blocked membranes were incubated with either 1/10,000 mouse anti-His antibody (Qiagen) diluted in 1X TBS or biotinylated-ConcavalinA (Sigma, St. Louis, MI) at 1 µg/ml in PBS. Membranes were incubated O/N at 4°C with the primary antibody and then washed thrice for 5 min with

0.5 % Tween-20 in TBS or PBS. After washing, Western blots were developed with NBT/BCIP (Sigma) as recommended by the manufacturer. For lectin blotting, washed membranes were incubated with 1 µg/ml peroxidase-labeled streptavidin for 1 h. Finally, lectins blots were developed with chloronapthol (Sigma) as recommended by the manufacturer.

3.2.7 N-Terminal Sequencing.

For N-terminal sequencing, 10 µg of protein were resolved by SDS-PAGE using 4-12% Tricine gels (Invitrogen). Proteins were electroblotted onto PVDF nylon membranes at 50 volts for 1 h. Thereafter, blotted membranes were stained with Coomassie blue for 30 min and destained in methanol [20]. Membranes were submitted to the Proteomics Facility at UC-Davis for N-terminal sequencing.

3.2.8 Mass Spectrometry analysis.

Protein Identification and Detection of glycopeptides. Five µg of protein was resolved by SDS-PAGE using 4-12% Tris-Glycine gels and in-gel digestion was performed as described before [21].

Peptides were separated by applying to a C18 chip column with 5µm particle size, a 40 nL trap, and dimensions of 75µm X 43 mm (Agilent, Santa Clara, CA). Peptides were eluted with an increasing linear gradient (15 to 80%) of acetonitrile in 0.1% formic acid over 17 min, using an Agilent 1260 capillary HPLC system at a flow rate of 4 µl/min. Eluted peptides were introduced in a qTOF Agilent C520 (Agilent).

Peptide identification was determined using Mascot (Matrix Science Ltd, www.matrixscience.com) software. Peptide identities were accepted if a 95% or higher probability was obtained using the Peptide Prophet algorithm. To identify glycopeptides, the MS/MS data was searched for a neutral loss of m/z 162 corresponding to a hexose. Analysis was performed using the XCalibur software version 2.0 SR2.

3.2.9 GGT Enzymatic Assay.

The GGT assay was performed as previously described [2]. Briefly, enzymes were incubated with 1 mM of the donor L-γ-glutamyl-ρ-nitroanilide (L-Glu-γ-ρNA, Sigma) and 20 mM of the acceptor glycylglycine (Gly-Gly, Sigma) in a 0.1 M Tris-HCl, pH8 buffer. Alternatively, Gly-Gly was replaced with a dipeptide of Glu residues (Glu-Glu, Sigma), oligopeptide of Glu residues (polyGlu, Sigma), a dipeptide of Gln residues (diGln, Pi Proteomics, Huntsville, AL), or a pentapeptide of Gln residues (pentaGln,Pi Proteomics) to evaluate other acceptors. The plate was kept on ice until the reaction was started and then incubated in a spectrophotometer plate (Bio-Rad) at 37°C for 15-30 min. OD readings at an absorbance of 405 nm were recorded every 3-5 min. GGT from equine kidney (Sigma) was used as a positive control.

3.3 Results

3.3.1 Bioinformatics analyses. The *B. anthracis capA-capE* operon located in the pX02 plasmid has been shown to encode for proteins involved in the biosynthesis of poly-γ-D-Glu, a capsule product essential for its virulence [12, 13]. Interestingly, virulent *Mycobacterium* such as *Mtb* uniquely synthesize α-L-poly-Gln, a similar peptidic polymer which could also have a role in *Mtb* virulence [11]. Taking in consideration that CapD is probably the best characterized protein encoded in the *capA-capE* operon, as a first approach to studying the biosynthesis of *Mtb* α-L-poly-Gln we decided to determine if *capD* homologue(s) are encoded by the *Mtb* H37Rv genome, The CapD sequence was obtained from the NCBI protein database and used for homology searches by BLAST analysis in the Tuberculist webpage (http://genolist.pasteur.fr/TubercuList/). As can be observed in Figure 3.1, Rv2394 was identified as the protein with highest homology to CapD. Specifically, both proteins were 26% identical, had 41% conserved amino acids and overall e-value of 5^{-36}. This protein is annotated as ggtB in the Tuberculist webpage.

Rv0773c (annotated as ggtA) was identified as a second mycobacterial protein with significant homology to both CapD and Rv2394 (not shown). Analysis was limited to Rv2394 as it has higher homology to CapD than Rv0773c. Furthermore, additional bio-informatic analyses suggested that similar to CapD, Rv2394 would also be localized to the extracellular space. In contrast, Rv0773c seems to be localized in the cytoplasm as it lacks a conventional or a lipoprotein signal peptide (not shown). Finally, it was determined that virulent *Mtb* as opposed to avirulent *Mycobacterium* have a conserved genomic organization in the *rv2394* locus (see below). This could suggest that *rv2394* participates in a unique pathogenesis mechanism not present in avirulent *Mycobacterium*.

```
>M. tuberculosis H37Rv|Rv2394|ggtB PROBABLE
          GAMMA-GLUTAMYLTRANSPEPTIDASE PRECURSOR GGTB
          (GAMMA-GLUTAMYLTRANSFERASE) (GLUTAMYL TRANSPEPTIDASE)
         Length = 643

 Score =  145 bits (367), Expect = 5e-36
 Identities = 124/472 (26%), Positives = 196/472 (41%), Gaps = 81/472 (17%)

Query: 49  YGVSASHPLAVEEGMKVLKNGGSAVDAAIVVSYVLGVVELHASXXXXXXXMLIISKDKET 108
           Y  + ++PLA +   +VL++GG+A DA +    VLG+VE  +S       ++       +
Sbjct: 80  YAAATANPLATQVACRVLRDGGTAADAVVAAQAVLGLVEPQSSGIGGGGYLVYFDARTGS 139

Query: 109 FIDY--RETTPYFTGNQ-------------KPH-------IGVPGFVAGMEYIHDNYGSL 146
              Y  RE  P                    +P+       IGVPG +  +E +H+ +G
Sbjct: 140 VQAYDGREVAPAAATENYLRWVSDVDRSAPRPNARASGRSIGVPGILRMLEMVHNEHGRT 199

Query: 147 PMGELLQPAINYAEKGFKVDDSLTMRLDLAKPRIYSDKLSIFY---PNGEPIETGETLIQ 203
           P  +L  PA+  A+ GF +   +   +  A P++  D  +  Y   P+G P   G  L
Sbjct: 200 PWRDLFGPAVTLADGGFDISARMGAAISDAAPQLRDDPEARKYFLNPDGSPKPAGTRLTN 259

Query: 204 TDLARTLKKIQKEGAKGFYEGGVARAISKTAK----------ISLEDIKGYKVEVRKPVK 253
              ++TL  I   GA  FY G +A  I   A           +++ED+ GY  + R+P+
Sbjct: 260 PAYSKTLSAIASAGANAFYSGDIAHDIVAAASDTSNGRTPGLLTIEDLAGYLAKRRQPLC 319

Query: 254 GNYMGYDVYTAPPPFSGVTLLQMLKLAE---------KKEVYKDVDHTATYMSKMEEISR 304
             Y G ++    P   GV +   L + E          K        T   + + E  R
Sbjct: 320 TTYRGREI-CGMPSSGGVAVAATLGILEHFPMSDYAPSKVDLNGGRPTVMGVHLIAEAER 378

Query: 305 IAYQDRKKNLGDPNYVNM---------DPNKMVSDKYISTMKNENGDA-----------L 344
           +AY DR + + D ++V +          DP  + +   + + ++  G A
Sbjct: 379 LAYADRDQYIADVDFVRLPGGSLTTLVDPGYLAARAALISPQHSMGSARPGDFGAPTAVA 438

Query: 345 SEAEHESTTHFVIIDRDGTVVSSTNTLSNFFGTGKYTAGFFLNNQLQNF--------GSE 396
                T+H  ++D  G   + T T+ + FG+     GF LNNQL +F        GS
Sbjct: 439 PPVPEHGTSHLSVVDSYGNAATLTTTVESSFGSYHLVDGFILNNQLSDFSAEPHATDGSP 498

Query: 397 GFNSYEPGKRSRTFMAPTVLKKDGETIG-------IGSPGGNRIPQILTPIL 441
             N  EPGKR R+ MAPT L  D  + G       +GSPGG+ I Q +   L
Sbjct: 499 VANRVEPGKRPRSSMAPT-LVFDHSSAGRGALYAVLGSPGGSMIIQFVVKTL 549
```

Figure 3.1 Amino acid alignment between *B. anthracis* CapD and *Mtb* Rv2394. The amino acid sequence for CapD was obtained from *B. anthracis* str. Ames ancestor and compared to *Mtb* proteins using the BLAST tool in the Tuberculist webpage http://genolist.pasteur.fr/TubercuList/. As can be observed, Rv2394 was identified as a CapD homologue with a high percent of identical and conserved amino acids.

Initial bio-informatic analysis was focused on predicting whether Rv2394 could have GGT activity similar to CapD [11]. The amino acid sequence for Rv2394 was obtained from Tuberculist (http://genolist.pasteur.fr/TubercuList/) and analyzed using BLAST® tool from NCBI (http://blast.ncbi.nlm.nih.gov/Blast.cgi). As can be observed in Figure 3.2, BLAST® analysis determined that Rv2394 has a putative GGT domain (pfam0109) with and e-value of 7.98^{-168}.

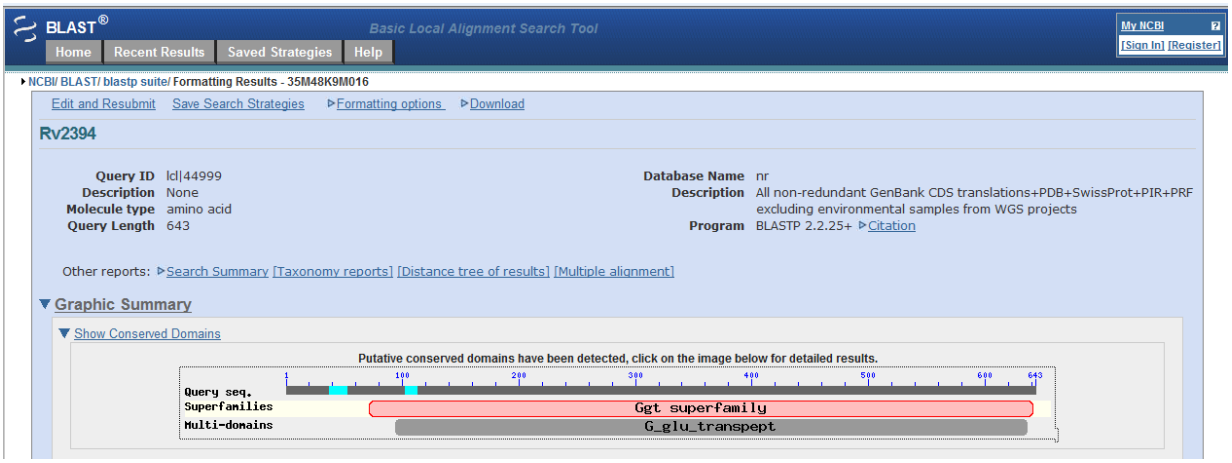


Figure 3.2 Rv2394 has a putative GGT domain. The amino acid sequence for Rv2394 was obtained from Tuberculist data base and analyzed with the BLAST® tool from NCBI. As can be observed, a putative GGT domain was predicted to be present in Rv2394.

The catalytic motif amongst GGTs is highly conserved and consists of a Thr residue located at the N-terminus of the enzyme's catalytic domain [22, 23]. Interestingly, this Thr residue is not the mature protein's N-terminal amino acid. An autocatalytic intra-chain cleavage splits the pre-protein into two subunits that remain non-covalently associated: the large (L) and small (S) subunits originate from the protein's N and C-terminal regions, respectively [24]. The catalytic Thr residue becomes the N-terminus of the cleaved S subunit. Usually the L subunit remains anchored to the extracellular leaflet of the cell membrane. To determine if Rv2394 conserves the amino acids present in the catalytic motif of

biochemically characterized GGTs, multiple sequence analysis was performed using VectorNTI (Invitrogen). In Figure 3.3, it can be observed that similar to other GGTs, Rv2394 possess a conserved Thr residue (position 446) present in the catalytic motif. Additional amino acids surrounding this Thr residue are also conserved. As suggested by the bioinformatics analysis, Rv0773c also has the conserved Thr residue.

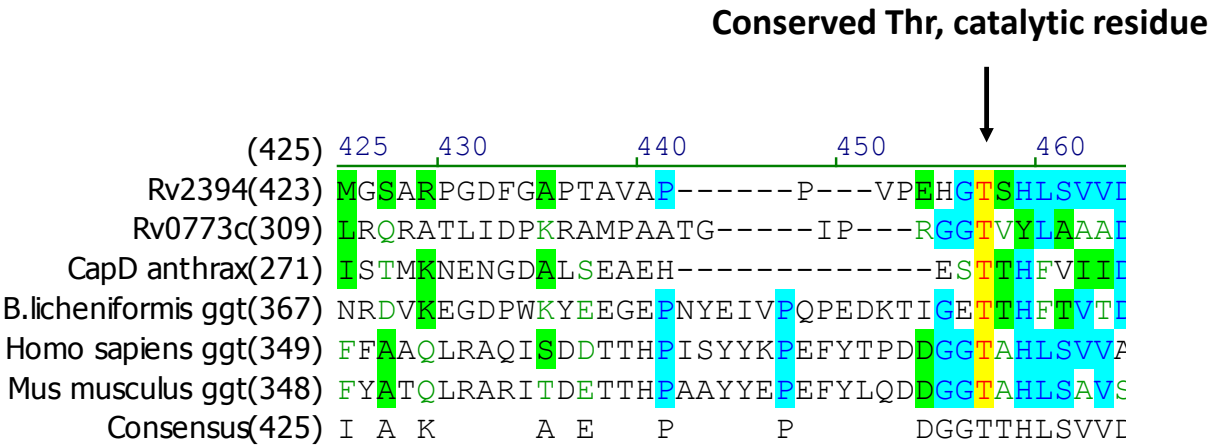


Figure 3.3 Rv2394 has conserved amino acid residues present in the catalytic motif of GGTs. Using VectorNTI's multiple sequence alignment tool, the sequence of Rv2394 was aligned with other biochemically characterized GGT sequences. Similar to other GGTs, a Thr residue that is critical for enzymatic activity and autocatalytic cleavage (marked with an arrow) is also present in Rv2394. As depicted in the consensus sequence, additional amino acids surrounding the catalytic Thr are also conserved in Rv2394.

B. anthracis CapD has been localized to the extracellular compartment, consistent with the subcellular localization of its substrate poly-γ-D-Glu [11]. Even though it hasn't been determined experimentally, secretion of CapD to the extracellular space is probably determined by a putative N-terminal signal peptide identified in our own bio-informatic analysis (not shown). Similar to *B. anthracis*

poly-γ-D-Glu, sub-cellular fractionation has also suggested that *Mtb* α-L-poly-Gln is located outside the mycobacterial cytoplasmic membrane [11]. Thus, it was predicted that Rv2394 could have a signal peptide similar to CapD and this was evaluated by subjecting the amino acid sequence of Rv2394 to bio-informatic analysis using SignalP 4.0 Server tool ([25], http://www.cbs.dtu.dk/services/SignalP/), and applying the algorithm for Gram positive organisms and both neural networks and hidden Markov models. This analysis indicated the presence of a possible signal peptide that would be cleaved between amino acids 28 and 29 by signal peptidase I (SP-I) (Fig. 3.4). Thus the Pro residue located at position 29, would become the mature protein's N-terminus after being translocated outside the cytoplasmic membrane. However, it is noted that the score for the SP-I cleavage site at amino acid 28-29 is relatively weak: the cutoff for signal peptides is set at 0.45 and Rv2394 obtained a value of 0.588.

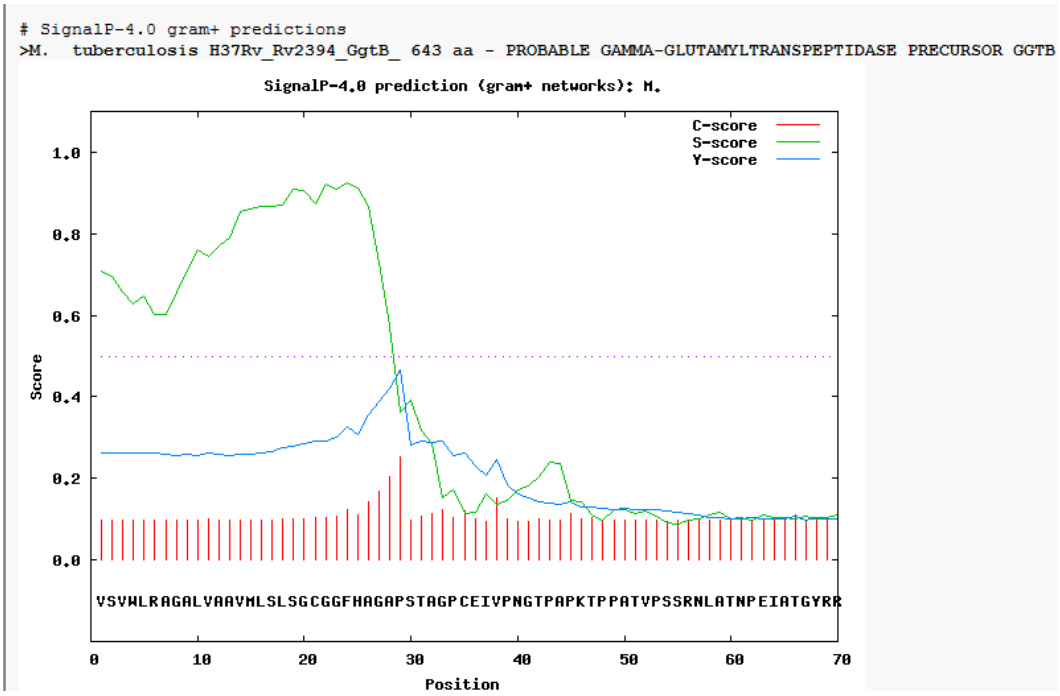


Figure 3.4 Signal peptide prediction in Rv2394 using the tool SignalP 4.0. The protein sequence for Rv2394 was obtained from Tuberculist data base and analyzed with SignalP 4.0. For analysis, the

algorithm for Gram-positive organisms and Markov method were used. A putative cleavage site was predicted to occur between alanine (Ala) at position 28 and proline (Pro) at position 29.

Alternatively, some bacterial proteins containing signal peptides are post-translationally modified with acyl groups at an N-terminal Cys residue exposed after cleavage by SP-II [26, 27]. Once again, *B. anthracis* CapD hasn't been experimentally shown to be anchored to the cell membrane through acyl groups but our own bio-informatic analysis highly suggests this (not shown). In order to evaluate if Rv2394 could be acylated as suggested by the annotation found in the Tuberculist webpage, its sequence was analyzed using the LipoP tool ([26], http://www.cbs.dtu.dk/services/LipoP/) available in Expasy (Fig. 3.5).

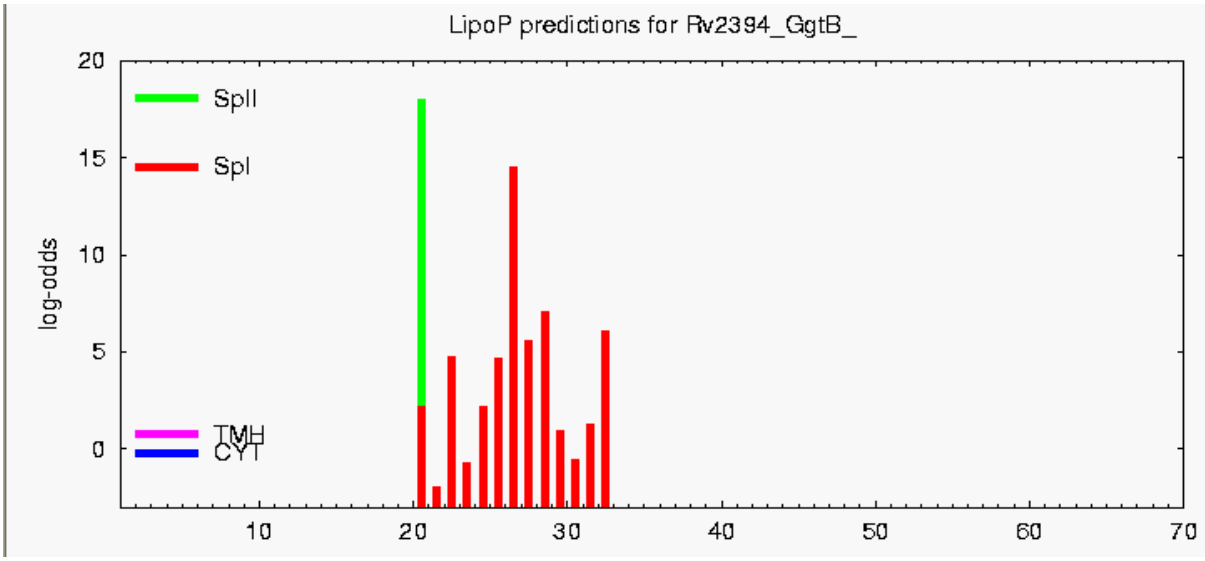


Figure 3.5 Cleavage of Rv2394's signal peptide is putatively mediated by signal peptidase II (SP-II). The amino acid sequence for Rv2394 was obtained from the Tuberculist data base and analyzed with LipoP tool. In addition to the presence of a putative signal peptide already predicted by SignalP, LipoP predicted that cleavage of Rv2394's signal peptide is probably mediated by SP-II between amino acid 20 and 21. Thus amino acid 21 (Cys) could represent Rv2394's N-terminus which could further be acylated.

Instead of the putative SP-I-mediated cleavage site occurring between amino acids 28-29 (score: 14.5), LipoP predicted a greater probability that Rv2394's signal peptide could be processed by the SP-II pathway (score: 18). Compatible with the well characterized mycobacterial acylation pathway [27], SP-II-mediated cleavage between amino acids 20 and 21 would leave Cys 21 as the new N-terminus. In turn this N-terminal Cys could become acylated.

As described in Chapter IV, a CapA homologue (Rv0574c) was identified in *M. tuberculosis* but no homologues to the other *cap* genes seem to be present. In bacteria, genes involved in a metabolic pathway tend to cluster and form operons. Therefore, genome organization analysis of the *rv2394* locus was performed to determine if neighboring genes could be involved in the biosynthesis of α-L-poly-Gln. Interestingly, in *Mtb* and *M. bovis*, the sulfite metabolism and Cys transport loci are split by the insertion of *rv2394* and 2 additional genes [28] (Fig. 3.6).

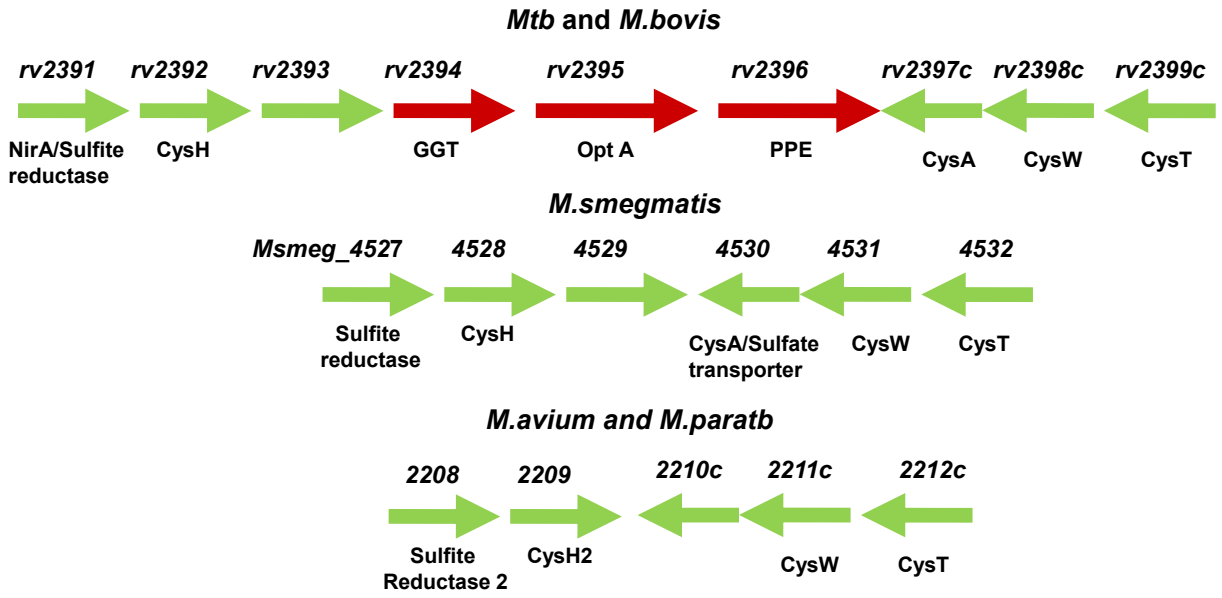


Figure 3.6 Comparison of the genome organization for the *rv2394* locus in different *Mycobacterium* spp. In *Mtb* and *M. bovis*, *rv2394* is located between the sulfite reductase and Cys transport loci. In

contrast, *M. smeg*, *M. avium* and *M. paratb* these two loci are contiguous without additional genes inserted in between.

In contrast, these loci are contiguous and face each other in the saprophytic *M. smeg* and the opportunistic *M. avium* and *M. paratuberculosis*. Thus, this unique genome organization only present in virulent *Mycobacterium* could suggest that these genes are involved in a unique pathogenesis mechanism that isn't present in avirulent bacteria. Rv2395 was identified as a putative oligopeptide transporter from the Opt family, whereas Rv2396 was determined to be a member of the PPE family. Even though it could be envisioned that Rv2395 participates in the transport of poly-Gln oligopeptides, a specific role in the biosynthesis of α-L-poly-Gln for these two proteins would require additional experimentation.

Altogether, results obtained from bio-informatic analyses suggested that Rv2394 is a putative lipoprotein with GGT activity dependent on a conserved Thr residue. This lipoprotein would be located in the extracellular compartment where *M. tuberculosis* α-L-poly-Gln has been identified by subcellular localization [15]. These results are in agreement with those obtained with *B. anthracis* poly-γ-D-Glu and CapD. Thus, it was hypothesized that Rv2394 could be involved in a GGT reaction putatively occurring during α-L-poly-Gln biosynthesis (see Chapter IV). CapD's enzymatic activity, protein motifs and post-translational modifications are well characterized [12]. Further evaluation of Rv2394 was therefore performed using recombinant protein constructs produced in *M. smeg.,* to evaluate the likelihood that Rv2394 could be CapD's ortholog.

3.3.2 Recombinant expression of *rv2394* constructs and evaluation of the resulting products.

To further explore Rv2394's enzymatic activity as well as post-translational modifications, Rv2394 was PCR-amplified from H37Rv genomic DNA (gDNA) and cloned into the mycobacterial expression vector pVV16 (see Table 3.1 for PCR primers and restriction sites).

The pVV16 expression vector encodes for a C-terminus hexa-Histidine tag (His-Tag), enabling purification by nickel affinity chromatography. Correct subcellular localization would be mediated by an N-terminus signal peptide. Additional Rv2394 constructs were engineered to evaluate the protein's enzymatic activity after: a) mutating the conserved catalytic Thr residue (Fig. 3.7A), b) forcing secretion through the SP-I pathway to abrogate N-terminus acylation (Fig. 3.7B) and c) mistargeting Rv2394 to the cytoplasmic compartment (Fig. 3.7C).

By site-directed mutagenesis, the catalytic Thr residue was modified to either a conserved amino acid like Ser (TS mutation), or a non-conserved amino acid like Ala (TA mutation) (Fig. 3.7A).

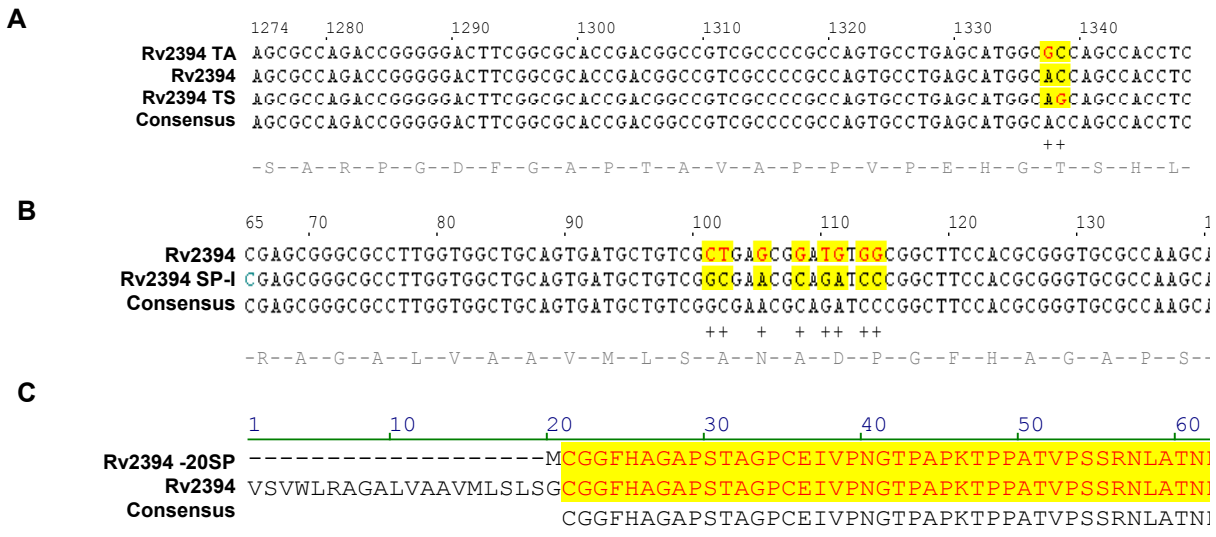


Figure 3.7 Site directed mutagenesis of additional Rv2394 constructs. Additional Rv2394 constructs were engineered to evaluate: **A**) Enzymatic activity after site-directed mutagenesis of the catalytic Thr residue (position 446, codon ACC), to either Ser (TS, codon AGC) or Ala (TA, codon GCC). **B**) Forced secretion through the SP-I pathway. By site-directed mutagenesis, sequence 15-MLS**LSGCG**G-23 was converted to 15-MLS**ANADP**G-23. This new sequence affecting the 3 most C-terminus amino acids of the signal peptide and the 2 most N-terminus amino acids of the mature protein, would abrogate Cys acylation and force SP-I mediated secretion instead of the default SP-II pathway. **C**) Protein

mislocalization by removing the signal peptide. The first 20 amino acids encoding for Rv2394's signal peptide were removed by PCR. Cys 21, corresponding to the putatively acylated amino acid during SP-II mediated secretion, would instead represent the second, non-acylated amino acid of a mistargeted, cytosolic protein. New constructs were verified by DNA sequencing as shown in Figure 3.7.

Site-directed mutagenesis was also used to force Rv2394's secretion through the SP-I pathway as opposed to the expected SP-II pathway (construct Rv2394 SP-I, Fig 3.7B). For this purpose, the sequence LSG**C**G was replaced with the sequence ANA**D**P. This mutation removes the putatively acylated Cys (bold and underscored), corresponding to the first amino acid of the mature Rv2394. Furthermore, the new sequence ANADP (obtained from MPT-32 or 45-kDa glycoprotein), conforms to the rule A-X-A-**D**-P which has been used in algorithms to predict mycobacterial protein secretion [29]. Signal peptide cleavage at the N-terminus of the aspartic acid residue (D) would leave the sequence DP as the mature protein's 2 most N-terminal amino acids. A third construct was engineered to mistarget a putative extracellular protein like Rv2394 to the cytoplasmic compartment (Fig. 3.7C). For this, the DNA sequence encoding for the signal sequence was removed by PCR to generate the construct Rv2394-20SP. Finally, both Rv2394 SP-I and Rv2394-20SP were subjected to a second mutation affecting the catalytic Thr residue to generate constructs RV2394 SP-I/TA and Rv2394-20SP/TA, respectively. After confirming by DNA sequencing that these modifications had been achieved (Fig. 3.7), plasmids were electroporated into *M. smeg* and proteins expressed as previously described.

Anti-His Western blots targeting the C-terminal histidine-tag were performed in order to evaluate expression of Rv2394 constructs. Rv2934's expected molecular weight is 66.5 kDa, but autocatalytic cleavage should result in a C-terminus histidine-tagged S domain with an expected molecular weight of 23 kDa. Consistent with this, a 23 kDa anti-His reactive band was detected by Western blot A when expressing both wild type Rv2394 and Rv2394 SP-I constructs (Fig. 3.8). A similar result was observed with the Rv2394 TS construct suggesting that mutating Thr to a conserved amino acid like Ser had no

apparent effect on the enzyme's autocatalytic activity. In contrast, a band migrating between 73-75 kDa was detected in all the constructs containing the TA mutation (Rv2394 TA, Rv2394 SP-I/TA and Rv2394-20SP/TA), suggesting that this mutation abrogated the enzyme's autocatalytic activity regardless of the protein's sub-cellular localization.

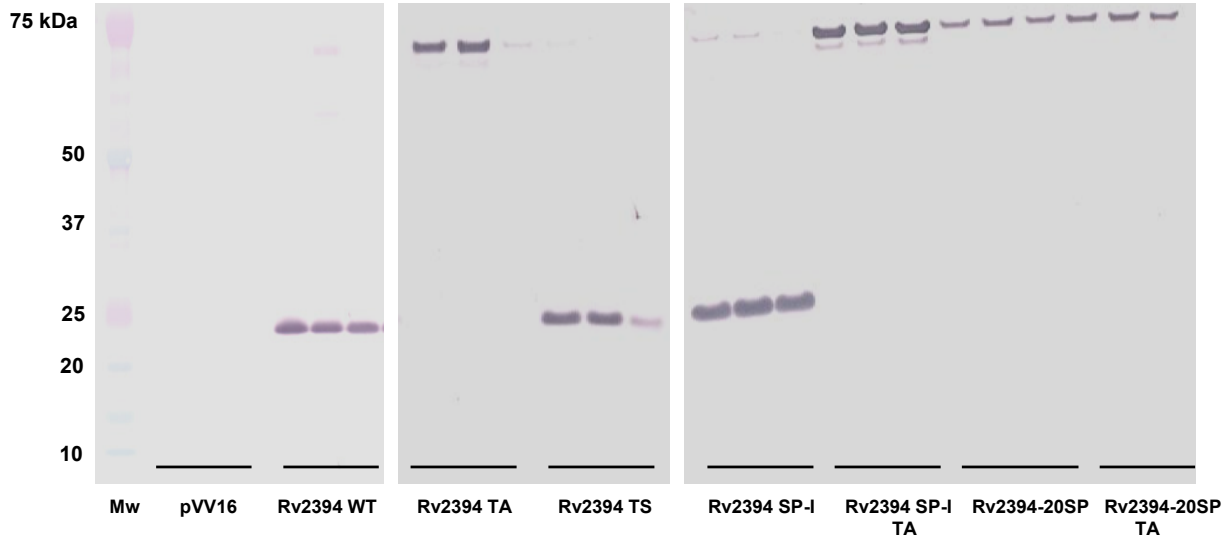


Figure 3.8 Expression of Rv2394 constructs in *M. smeg*. *M. smeg* was electroporated with the different Rv2394 constructs. Protein expression was evaluated for three different colonies each, by anti-His Western blot as all the constructs have a C-terminus His-Tag. A anti-His reactive 23 kDa band was observed for constructs Rv2394 WT, Rv2394 TS, Rv2394 SP-I. In contrast, a 70 kDa anti-His reactive band was observed for all the constructs having the TA mutation. Unexpectedly, Rv2394 – 20SP also migrated at 70 kDa despite having a WT catalytic site. By protein expression, it seemed that both constructs lacking a signal peptide were expressing less efficiently, therefore there were not pursued further. No anti-His reactive bands were observed for cells electroporated with the vector control (pVV16).

Notably, despite encoding for the wild type enzymatic motif, expression of the construct Rv2394-20SP had a similar effect to the TA mutants and no autocatalytic processing was detected. Thus, at least for this construct, sub-cellular localization affected the protein's enzymatic activity. As low expression levels were detected by Western blot for both cytoplasmic-localized constructs, they were not pursued further as protein yield could be limiting. In contrast, removal of the acylation site and forcing secretion through the SP-I pathway, had the opposite effect and resulted in increased band intensity. Thus, further biochemical analyses were mainly performed on Rv2394 SP-I and the TA version with abrogated autocatalytic activity (some experiments were also performed with the putatively acylated WT or TA and TS mutants).

3.3.3 Purification of Rv2394 SP-I.

Ni^{+} affinity chromatography was used to purify Rv2394 SP-I and Rv2394 SP-I/TA via interaction with the recombinant proteins' C-terminus His-Tag. For Rv2394 SP-I, analysis by SDS-PAGE and silver staining revealed the expected presence of two predominant bands migrating at approximately 23 and 50 (Fig. 3.9).

As previously observed during protein expression checks, Western blots confirmed that the band migrating at 23 kDa was responsible for the anti-His reactivity present in the S subunit. No anti-His reactivity was observed for the L subunit migrating at 50 kDa. In contrast and in accordance with Rv2394's SP-I/TA predicted size after abrogation of the autocatalytic activity, a band migrating at approximately 73 kDa was observed both on the anti-His Western blot and silver stained gel (Fig. 3.9). The protein's identity was confirmed by mass spectrometry analysis of in-gel digested bands. In addition, N-terminus sequencing of both bands further confirmed the protein's identity (Table 3.3).

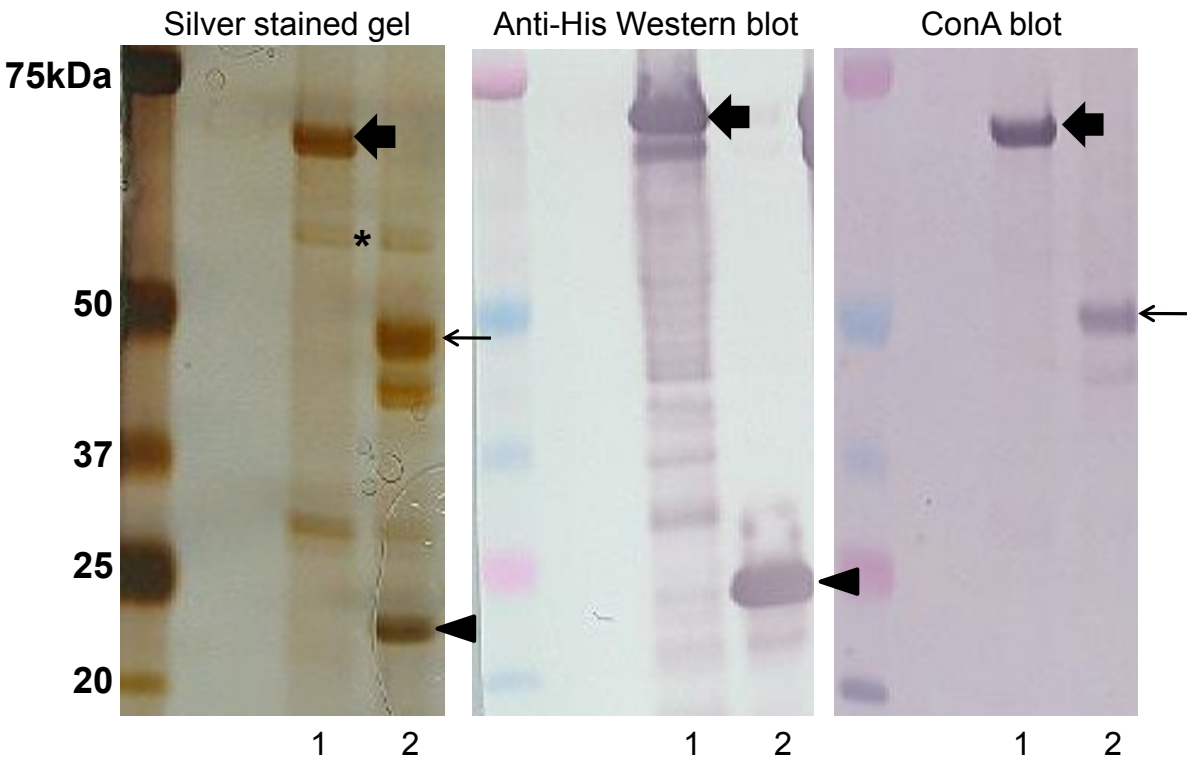


Figure 3.9 Rv2394 consists of two non-covalently associated subunits. Silver stained gel (left panel) revealed that Rv2394 SP-I (lane 2) consists of two subunits migrating at 50 and 23 kDa approximately (thin arrow and arrow head, respectively). In contrast, Rv2394 SP-I/TA (lane 1) lacking autocatalytic activity, migrates at approximately 70-73 kDa (thick arrow). For Rv2394 SP-I, anti-His reactivity (middle panel) is restricted to the 23 kDa band (arrow head). In contrast, reactivity using the mannose-binding lectin ConA (right panel), is restricted to the 50 kDa band (thin arrow). As expected, anti-His and ConA blots react with the 70-73 kDa Rv2394 SP-I/TA (thick arrow) resulting from both subunits remaining covalently attached due to the abrogated autocatalytic activity. GroEL1 is denoted with an *.

Table 3.2 N-terminal sequencing of Rv2394 constructs

Protein	Amino acid sequence
Rv2394 SP-I S-subunit	**Thr-Ser-His-Leu**
Rv2394 TA	**Blocked**
Rv2394 SP-I TA	**Gly-Ala-Pro-Ser-Thr**
Rv2394 SP-I L-subunit	**Gly-Ala-Pro-Ser-Thr**

As predicted by bio-informatic analyses comparing other GGTs (Fig. 3.4), Thr 446 was unequivocally identified at the 23 kDa band's N-terminus. Data obtained from N-terminus sequencing of the 50 and 73 kDa bands of Rv2394 SP-I and its TA version respectively, resulted in several signals suggesting that signal peptide cleavage of Rv2394 SP-I was heterogeneous. Alternatively, additional trimming of N-terminal amino acids by other proteases could have occurred after SP-I cleavage. Nevertheless, the predominant signal obtained from N-terminal sequencing of these bands corresponded to Gly-Ala-Pro-Ser-Thr, which is located four amino acids away from the carboxy terminus of the engineered cleavage site.

Despite multiple attempts to obtain pure protein, this was not completely successful for several reasons. Non-denaturing conditions were used during purifications to conserve the protein's enzymatic activity (see below). The protein did not bind strongly to the Ni^{+} resin, thus it started eluting during the wash steps. In turn, this resulted in low protein yields. Furthermore, GroEL1 a 60 kDa mycobacterial heat shock protein containing a histidine-rich motif, was routinely found as a contaminating band (Fig. 3.9, asterisk in left panel). Enzymatic assays were therefore performed with enriched Rv2394 and mass spectrometry analysis was done from gel extracted, SDS-PAGE-separated proteins.

3.3.4 Identification of post-translational modifications present in Rv2394 SP-I.

As mentioned above, bio-informatic analyses suggested that during secretion through the SP-II pathway, Rv2394 could be post-translationally modified with acyl groups on the N-terminus Cys residue. Two indirect approaches were taken to evaluate if the WT Rv2394 could be acylated as predicted by LipoP (Fig. 3.6, see above). Initially, N-terminal sequencing of the purified protein was attempted (Table 3.3, see above). In contrast to the constructs that were engineered to be secreted through the SP-I pathway, no sequencing was possible because the N-terminus was blocked possibly due to the acyl group. As N-terminal sequencing is abrogated by multiple amino acid modifications besides acylation, an additional technique of detergent phase partitioning with Tx-114 was performed [30]. After inducing partitioning at 37°C, lipoproteins are enriched in the detergent phase due to the presence of its acyl group. As can be observed in Figure 3.10, anti-His reactivity targeting the protein's C-terminus of the S domain was detected in the detergent phase but not in the aqueous phase.

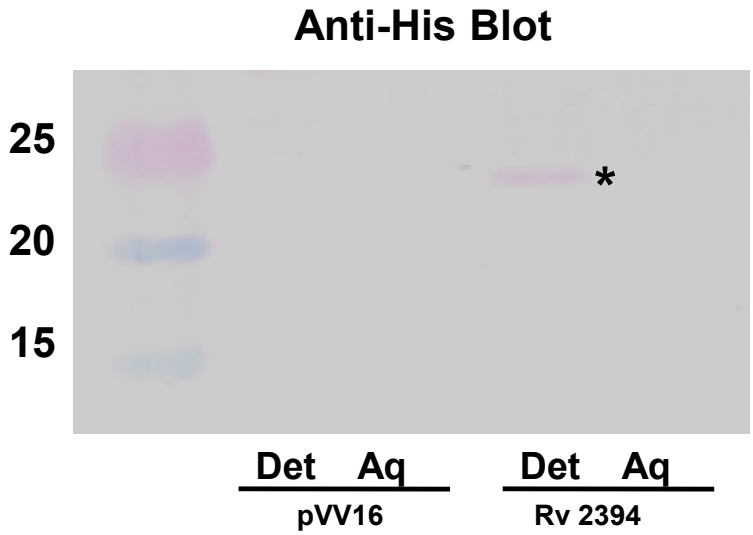


Figure 3.10 Detergent phase partitioning in Tx-114 suggest Rv2394 could be acylated. Whole cell lysate of *M. smeg* overexpressing WT Rv2394 (right lanes) or the vector control (pVV16, left lanes) were subjected to detergent phase partitioning in Tx-114. This is a common technique highly suggestive of

protein acylation. After acetone precipitation, proteins were resolved in SDS-PAGE and analyzed by anti-His Western blot. As expected, no anti-His reactivity at 23 kDa was observed in lanes containing vector control. In contrast, anti-His reactivity was observed in the detergent phase of whole cell lysates of *M. smeg* overexpressing Rv2394 (asterisk). As acylation occurs in the N-terminal domain but the His-Tag is present in the C-terminal domain, this result suggests non-covalent interactions between both subunits are stronger than detergent hydrophobic forces separating them.

As acylation occurs in the N-terminus domain, this result suggests that the non-covalent interaction between both domains was stronger than the hydrophobic forces during Tx-114 partitioning. More conclusive data suggesting that Rv2394 is acylated could be obtained if Tx-114 partitioning was attempted in Rv2394 TA. This construct should be acylated at the N-terminal Cys and as it has an abrogated autocatalytic activity, the full length protein (73 kDa) should partition in the detergent phase. Definitive evidence proving acylation of WT Rv2394 could be obtained by mass spectrometry techniques such as MALDI or lipid analysis by GC-MS.

An additional post-translational modification commonly observed in secreted mycobacterial lipoproteins is glycosylation. Bio-informatic analyses using NetOGlyc tool (http://www.cbs.dtu.dk/services/NetOGlyc/) in the Expasy website [31], suggested some of the Thr residues of Rv2394 could be glycosylated (Fig. 3.11).

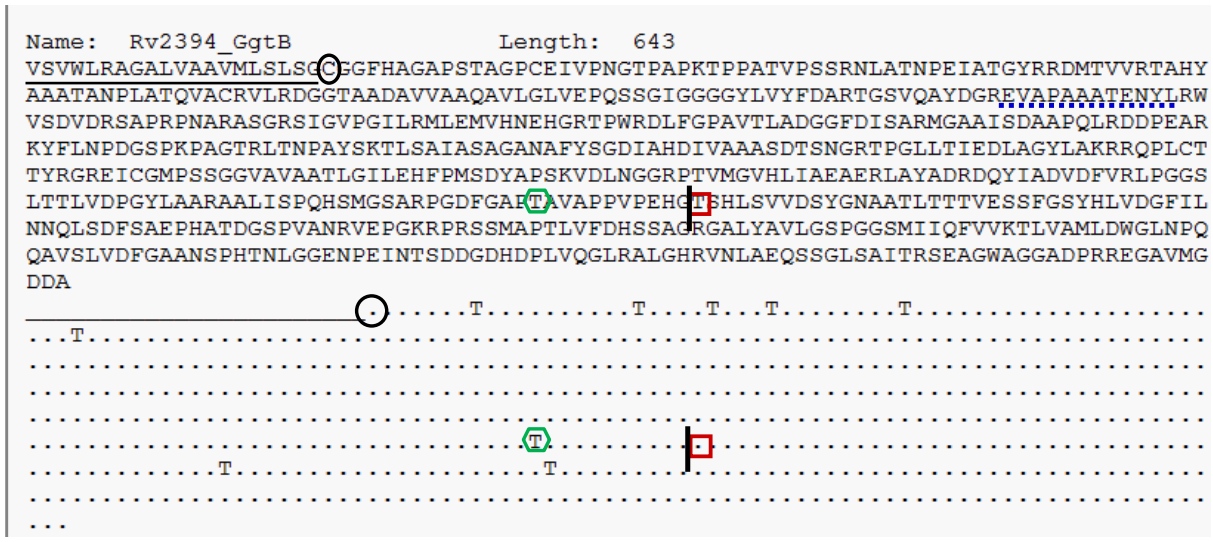


Figure 3.11 Rv2394's glycosylation prediction by NetOGlyc Rv2394's sequence was obtained from Tuberculist and analyzed by NetOGlyc tool. The protein's amino acid sequence is shown above, the putatively glycosylated Thr residues (T) are depicted below. In addition to detecting the previously recognized signal peptide (underlined), this tool also predicted that 9 T residues could be glycosylated. The majority of these T residues are clustered at the protein's N-terminus in proximity to the acylated Cys residue (black circle). Interestingly, NetOGlyc did not predict glycosylation of T 446 (red box), the putative catalytic residue which cleaves itself from the preceding Gly residue (vertical black line). The green diamond shows a T residue predicted to be glycosylated and confirmed by mass spectrometry. The blue dotted line corresponds to a putatively glycosylated and phosphorylated peptide.

Consistent with its putative catalytic function Thr 446 (Fig. 3.11, red box), was not predicted to be glycosylated. The majority of the predicted glycosylated Thr residues were clustered in the L subunit in proximity to the N-terminal acylated Cys (Fig. 3.11, circle). Initial evaluation to determine if Rv2394 SP-I could be glycosylated was performed by Western blot with ConA as the probe. As observed in Figure 3.9 (right panel), ConA interacted with the 50 kDa L subunit but not with the anti-His reactive 23 kDa S subunit. This is consistent with the bioinformatic analysis predicting the majority of glycosylated

residues are present in the N-terminus L subunit. Furthermore, ConA also bound to the 73 kDa Rv2394 SP-I/TA containing both the N-terminus (glycosylated L domain) and the C-terminus (His-tagged S domain). Definite confirmation that Rv2394 SP-I is glycosylated was performed by mass spectrometry analysis of in-gel trypsin-digested Rv2394 SP-I. Mass spectrometry data was analyzed for the neutral loss of 162.05 amu, corresponding to the mass of one hexose moiety. Identification of the de-glycosylated peptide was performed by MS/MS and compared to a theoretical MS/MS sequencing obtained from Protein Prospector (http://prospector.ucsf.edu/prospector/mshome.htm). As can be observed in Figure 3.12, one of the identified peptides corresponds to the L subunit's C-terminus generated after autocatalytic cleavage.

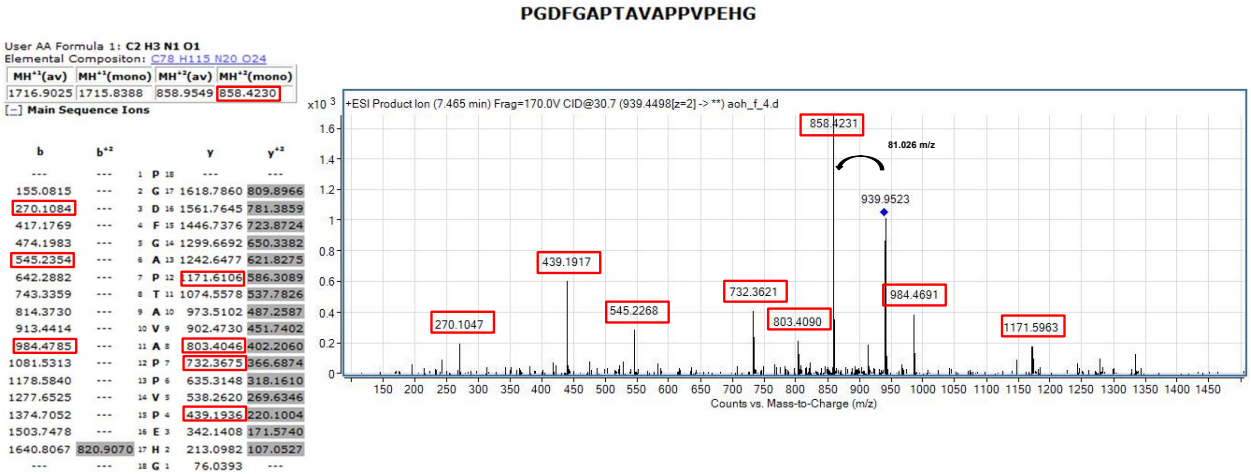


Figure 3.12 Rv2394 SP-I is post-translationally modified with hexose residues. After in-gel trypsin digestion of Rv2394 SP-I, peptides were analyzed by mass spectrometry for neutral loss of 162.05Da, corresponding to one hexose residue. As can be observed in the right panel, a doubly charged peptide with an m/z of 939.9523 was converted to an 858.423 m/z peptide. The 81.52Da difference corresponds to a hexose on a doubly charged peptide. MS/MS of the 858.423 m/z peptide resulted in sequencing of the peptide PGDFGAPTAVAPPVPEHG as can be observed in the left panel.

As predicted by bio-informatic analysis with NetOGlyc (Fig. 3.11, green hexagon), the only Thr residue present in this peptide could be the modified amino acid. Longer versions of this peptide with additional amino acids at the N-terminus were also observed by mass spectrometry to have a neutral loss of 1 hexose residue (not shown).

Serendipitously, two peptides with m/z of 1025 and 1054 were observed to undergo the loss of a hexose residue in addition to a neutral loss of 79.98 amu (Fig. 3.13). This neutral loss is characteristic of phosphate residues [32].

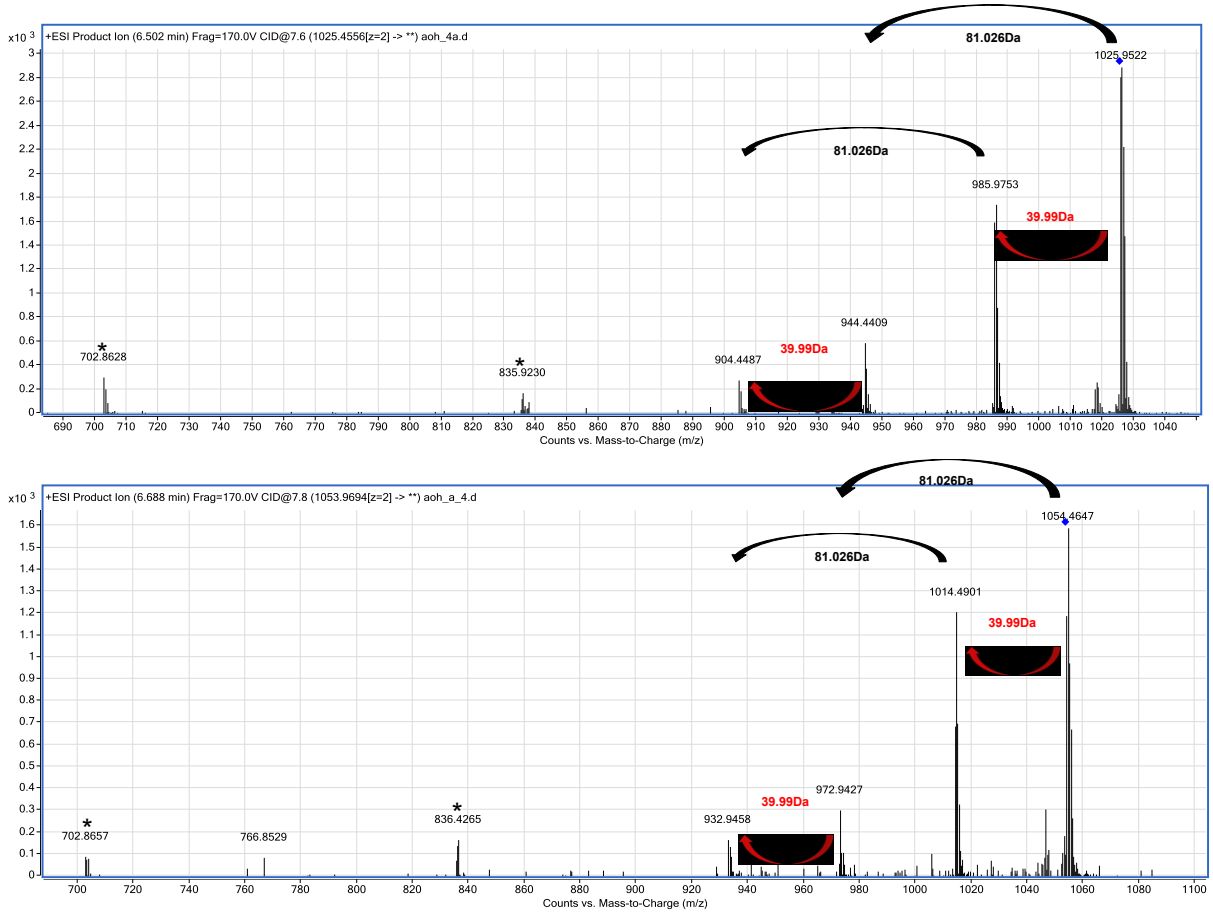


Figure 3.13 Mass spectrometry evidence that Rv2394 is phosphorylated. Two peptides with m/z of 1025.95 and 1054.46, had a neutral loss corresponding to a hexose residue (81.026) and surprisingly, an

additional neutral loss corresponding to a phosphate group with m/z of 39.99. The neutral loss of the hexose or phosphate group was consecutive but not simultaneous, suggesting that the phosphate group was not present on the sugar. The 1054.46 m/z peptide seems to represent a longer version of the 1025.95 m/z peptide as suggested by the presence of similar ions (702.86 and 835.92 m/z, *) during ms/ms.

These neutral losses were sequential but not simultaneous, indicating that phosphorylation occurred elsewhere besides the hexose residue. The peptide with m/z of 1054.46 seems to represent a longer version of the 1025.95 m/z peptide as similar ions with m/z of 702.86 and 835.92, were observed during MS/MS. Characterization of this glyco-phospho peptide has not been fully achieved probably attributed to additional modifications present in the surrounding amino acids. Nevertheless, the ion with m/z of 702.86 was unequivocally identified as EVAPAAAATENYLR from the parent MS where this ion was observed along with those for the modified peptide (Fig. 3.14). This peptide is located in the L subunit of Rv2394 but was not predicted by NetOGlyc to be glycosylated (Fig. 3.11).

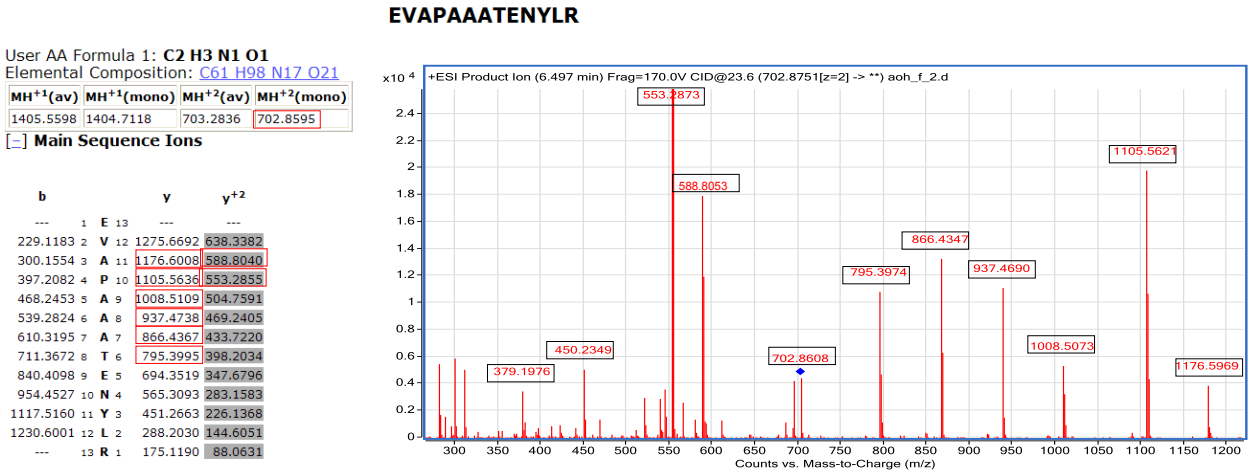


Figure 3.14 Identification of the 702.85 m/z ion. Identification of the peptide previously observed in the scans from the 1054.46 and 1025.95 peptides, was achieved by comparing an experimental vs theoretical

mass spectrometry digest obtained from ProteinProspector. The peptide's sequence is displayed on the left.

Altogether, in addition to the observed auto-catalytic activity, the obtained experimental data is compatible with the bio-informatic analyses predicting that Rv2394 could undergo two additional post-translational modifications associated with secretion: acylation of the N-terminus Cys residue and glycosylation of Thr residues. Furthermore, Rv2394 was also observed to putatively undergo phosphorylation. A summary of all the engineered Rv2394 constructs and possible post-translational modifications is depicted in Figure 3.15).

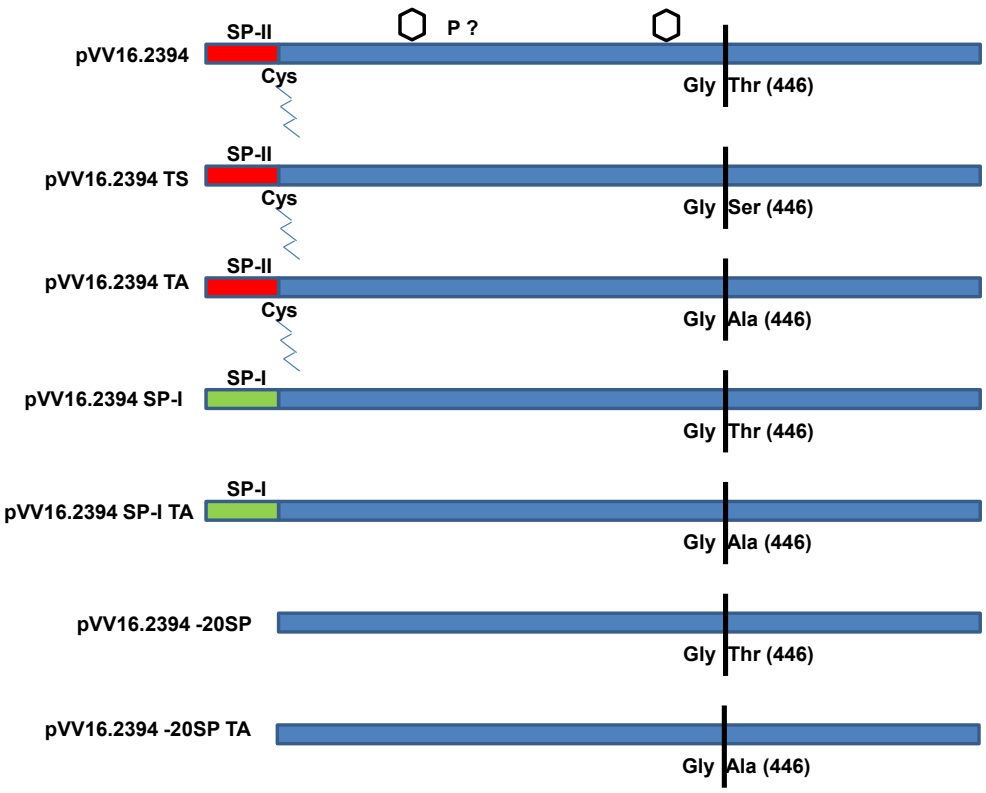


Figure 3.15 Summary of the engineered Rv2394 constructs and putative post-translational modifications. Diagram representing the Rv2394 constructs that were evaluated in this study. WT

Rv2394 (pVV16.2394) has a SP-II motif (red box) at the N-terminus with a properly located Cys residue that might be acylated (zig-zag line). Autocatalytic activation results in cleavage between the Gly and Thr residues located at position 445 and 446 (vertical black line). This event is catalyzed by the Thr residue at position 446. Site directed mutagenesis of this amino acid was performed to change the catalytic Thr residue to a conserved Ser or non-conserved Ala residues. Furthermore, the SP-II motif was mutated so as to convert it to a SP-I motif (green box) which does not undergo acylation (pVV16.2394 SP-I). A third mutant was made in which the SP motif was completely removed to affect protein secretion (pVV16.2394 -20SP). Double mutants were generated for pVV16.2394 SP-I and pVV16.2394-20SP, in which the catalytic Thr residue was mutated to an Ala residue. In addition to acylation Rv2394 also undergoes glycosylation (diamond) and possibly, phosphorylation (P ?).

3.3.5 Enzymatic activity of WT and mutant Rv2394.

Finally, Rv2394's GGT activity was evaluated by performing enzyme kinetic assays as described elsewhere [33]. In this reaction, the donor's γ-linked Glu moiety is transferred to either water (hydrolysis) or the N-terminus of an acceptor such as amino acids or peptides (transpeptidase). For this assay, L-Glu-γ-pNA was used as the donor and L-amino acids or peptides were used as acceptors [34]. GGT-mediated release of pNA was followed over time by measuring the absorbance at 405 nm, but unfortunately this method does not allow differentiating between a hydrolysis or transpeptidase reaction. Initial enzyme assays were performed to determine the effect of mutating the catalytic Thr residue to a conserved or non-conserved amino acid such as Ser and Ala, respectively. Similar to a eukaryote GGT that was used as a positive control, WT Rv2394 was able to cleave L-Glu-γ-ρNA in the presence of Gly-Gly (Fig. 3.16), the classical acceptor used in GGT assays.

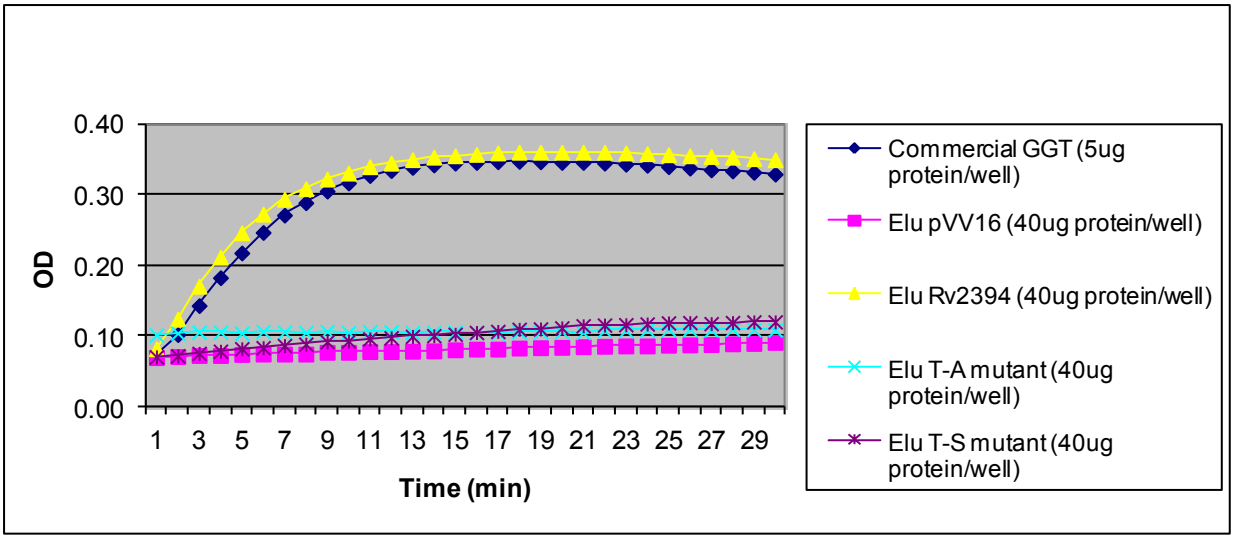


Figure 3.16 WT Rv2394 has GGT activity as opposed to the TS and TA mutants. GGT activity was evaluated using the acceptor Gly-Gly and following the release of ρNA from the donor L-Glu-γ-ρNA. The WT Rv2394 had considerable enzymatic activity and similar kinetics as a eukaryote GGT used as a positive control. In contrast, both TA and TS mutants lacked any enzymatic activity. Similarly, no activity was observed in control (pVV16) *M. smeg* which was processed identically as the recombinant Rv2394 clones.

Surprisingly, despite having auto-catalytic activity (Fig. 3.8), Rv2394 TS lacked GGT activity and no increase in absorbance could be observed during the 30 min reaction. As expected from other reports evaluating TA mutants devoid of auto-catalytic activity, no detectable GGT activity was observed for Rv2394 TA.

Furthermore, acceptors mimicking the structure and amino acid composition of α-L-poly-Gln were additionally evaluated. Initially, the Glu oligomers α-L-Glu-Glu (Di-Glu) and α-L-poly-Glu (poly-Glu) were used as they were readily available. Importantly, WT Rv2394 could also use Di-Glu and poly-Glu as acceptors to a similar extent as Gly-Gly (Fig 3.17). In contrast, the commercial eukaryote GGT could not use Di-Glu or polyGlu as acceptors.

As Gln oligomers resemble α-L-poly-Gln better than Glu oligomers, the former were subsequently evaluated. Rv2394 was able to use di or penta-α-L-Gln as acceptors for the GGT activity, in a similar fashion as the commonly used Gly-Gly (Fig. 3.18A). Once again and in comparison to Gly-Gly, the commercial GGT displayed minimal activity when using di or penta-α-L-Gln as acceptors (Fig. 3.18B). Finally, the effect that bicarbonate might have on Rv2394's enzymatic activity was evaluated. The rationale for including bicarbonate in the reaction buffer was the description that it enhances *B. anthracis* capsule formation [35]. Even though bicarbonate's effect on capsule synthesis has been principally associated with increased expression of the *cap* genes [36], its effect on the enzymatic activities of the Cap proteins has not been considered. Interestingly, it was noted that bicarbonate containing buffers significantly enhanced Rv2394's GGT activity towards all the acceptors. Furthermore, the presence or absence of bicarbonate in the reaction's buffer didn't significantly enhance the commercial GGT's activity (Fig. 3.18B). For eukaryote GGTs, an allosteric effect had been previously reported for the non-physiological maleate [37]. Whether bicarbonate's effect on Rv2394 activity is physiologically relevant, remains to be explored.

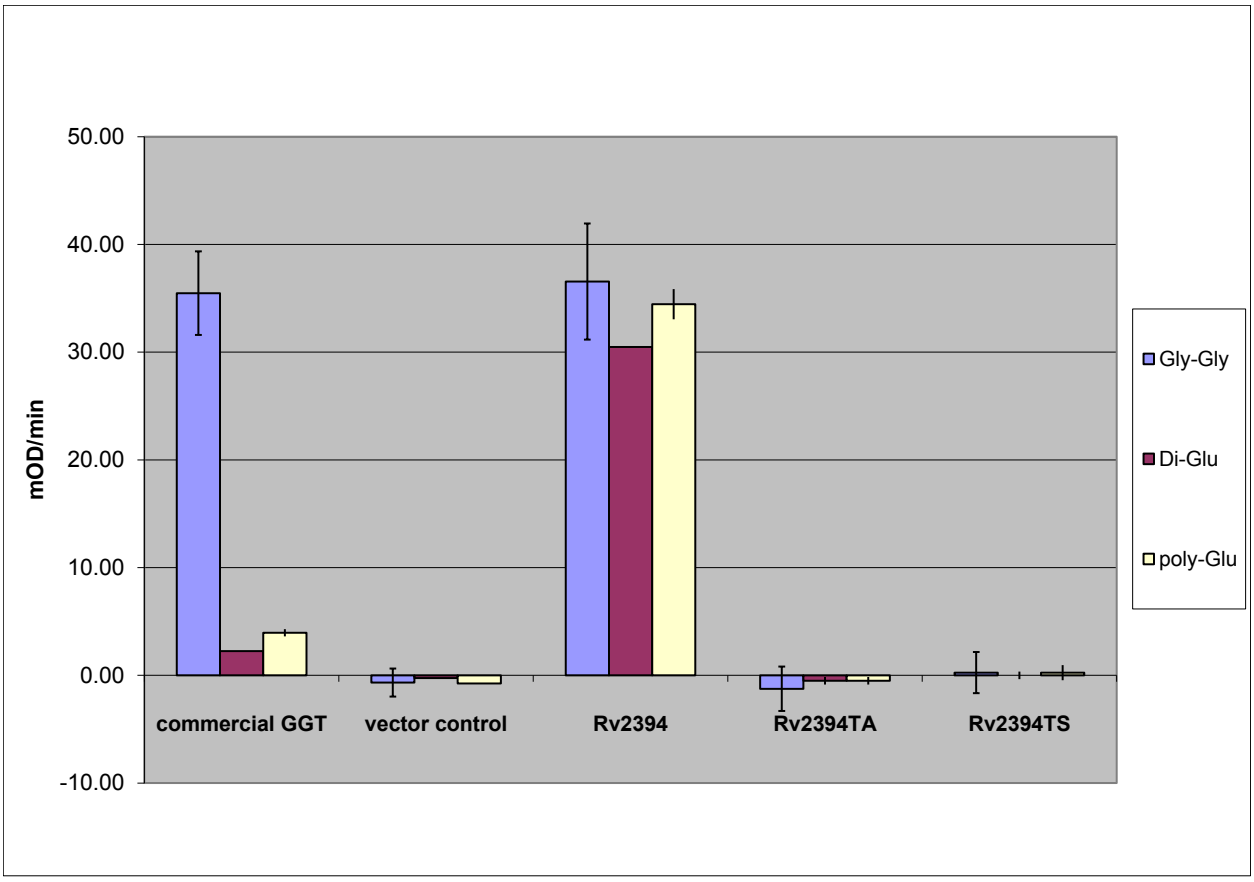
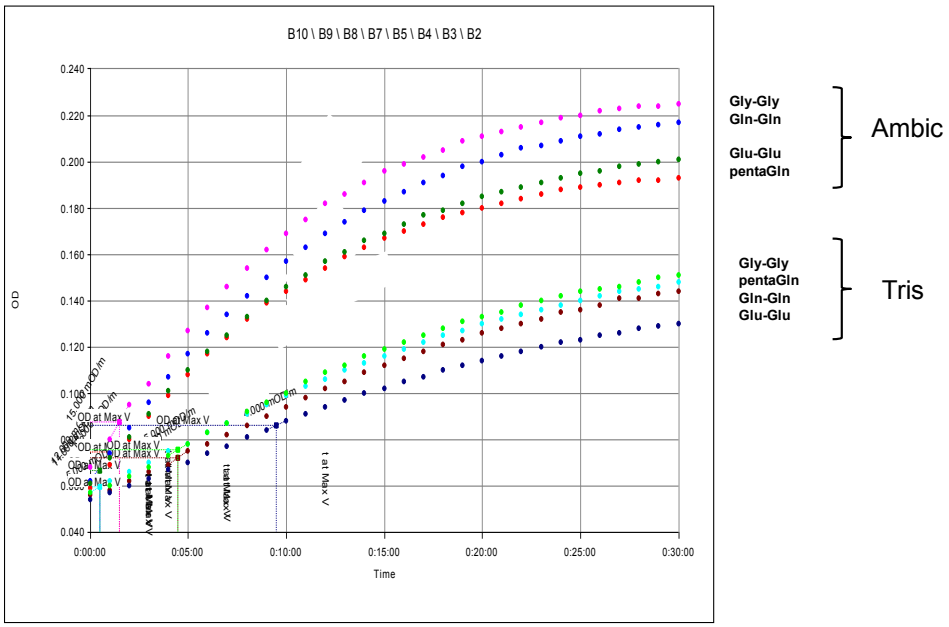
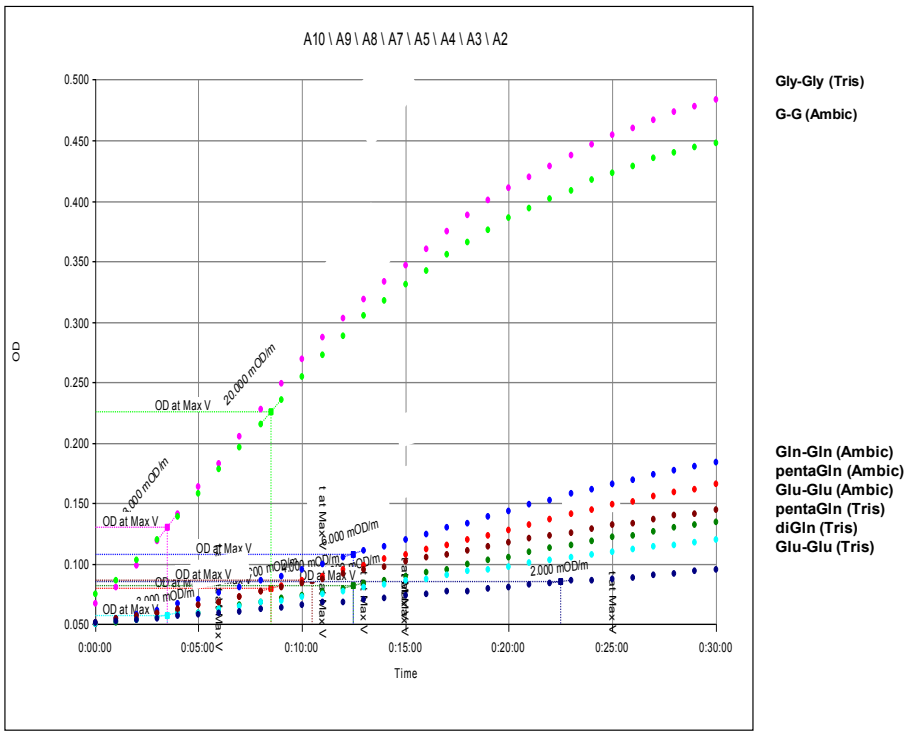


Figure 3.17 Glu di/oligopeptides can be used as acceptors by WT Rv2394. GGT activity was evaluated using different acceptors as depicted in the legend. WT Rv2394 could use short L-Glu-containing peptides as acceptors. In contrast, eukaryote GGT and both Rv2394 mutants lacked enzymatic activity when including these acceptors. As shown in Figure 3.16, both Rv2394 mutants also lacked activity for the dipeptide Gly-Gly.



A. Rv2394



B. Commercial GGT

Figure 3.18 Gln di or pentapeptides can be used as acceptors by Rv2394 and its activity is enhanced by bicarbonate containing buffers. A) As α-L-polyGln is composed of Gln residues, Gln di or

pentapetide was used as an acceptor to evaluate the GGT activity of Rv2394. In Tris containing buffers, these Glnd-based acceptors were used by Rv2394 as efficiently as Gly-Gly. Interestingly, the presence of bicarbonate in the reaction of buffer (as in ammonium bicarbonate, Ambic) as opposed to Tris, significantly enhanced Rv2394's activity using all the acceptors. B) In contrast to its activity in the presence of the acceptor Gly-Gly, and as observed for Glu-based acceptors, the commercial GGT didn't efficiently used Gln-based acceptors. In addition, the presence of bicarbonate (Ambic) didn't have a major effect on the enzyme's activity towards any of the evaluated acceptors.

In conclusion, GGT activity was confirmed for Rv2394. In contrast to a commercial eukaryote GGT used as control, Rv2394 was capable of using Glu or Gln based di or oligo peptides as acceptors, which resemble the structure of α-L-poly-Gln.

3.4 Discussion.

Since the discovery that transpeptidases are targeted by the widely used β-lactam antibiotics, these enzymes have become a major focus of study in the field of bacterial physiology and pathogenesis [38]. The classic transpeptidase/transglycosase penicillin-binding proteins were shown to mediate cross-linking between the third and fourth amino acids of neighboring peptidoglycan pentapeptides (4→3 linkage). Subsequently, three additional groups of bacterial transpeptidases have been characterized and interestingly, several of these enzymes have been involved in remodeling or modifying peptidoglycan through different mechanisms. L,D-transpeptidases (Ldts) were initially identified to catalyze a β-lactam-resistant, non-classical 3→3 linkage present in Gram positive bacteria [39]. It has become evident that Ldts are not restricted to Gram positive bacteria as they were recently shown to participate in cross-linking Braun lipoprotein to *E. coli* peptidoglycan [40]. Furthermore, Ldts were also shown to mediate 3→3 linkages present in the peptidoglycan of stationary *Mtb* [41]. Inactivation of this enzyme was shown to compromise *Mtb* survival in the chronic phase of tuberculosis [42]. Additionally, it rendered the

bacillus susceptible to a combination of β-lactam antibiotics and β-lactamase inhibitors. Sortases are an additional group of transpeptidases shown to covalently link proteins to peptidoglycan in Gram positive bacteria [43]. These enzymes mediate a transpeptidase reaction between the third amino acid of the peptidoglycan pentapeptide and a specific sequence in the protein to be cross-linked. Recently, CapD a GGT present in *B. anthracis* was included as an additional member of peptidoglycan-modifying transpeptidases [11, 14, 44]. CapD was shown to covalently link polymers of γ-D-Glu to meso-DAP and most importantly, genetic or pharmacological inhibition of *B. anthracis* CapD reduced lethality in the mouse model [14]. In our search for CapD homologs that could be involved in the bio-synthesis of *Mtb* α-L-poly-Gln, Rv2394 was identified as the best candidate due to its high homology to CapD, putative extracellular localization and its presence in a highly conserved genomic locus only encountered amongst virulent *Mycobacterium*. Significantly and as described in Chapter IV, *Mtb* mutants for Rv2394 had reduced concentrations of Glu and ammonia in their cell wall, further substantiating its role in α-L-poly-Gln biosynthesis.

The major reason to characterize *Mtb* Rv2394 was its homology to CapD, the *B. anthracis* enzyme participating in poly-γ-D-Glu synthesis. The exact role of CapD in poly-γ-D-Glu biosynthesis had been controversial until recently. Initially, the enzyme was characterized as catalyzing a hydrolysis reaction leading to *B. anthracis* capsule depolymerization. Hence the protein was initially named Dep instead of CapD [45, 46]. CapD's most straightforward evaluable activity still consists of incubating the enzyme with high molecular weight *B. anthracis* capsule [11, 47, 48]. By agarose gel electrophoresis and methylene blue staining, it can be observed that this incubation leads to progressive loss of the high molecular weight capsule with concomitant generation of smaller fragments that can even run out of the gel. Furthermore, incubation of *B. anthracis* with CapD or alternatively CapD overexpression, led to the accumulation of small γ-D-Glu capsule fragments in culture supernatants [47, 48]. Altogether, these results led to the hypothesis that CapD was involved in capsule degradation or turnover. In contrast to the held dogma, it was recently shown that mutation of the *capD* gene in *B. anthracis* led to a similar

accumulation of capsule γ-D-Glu oligopeptides in the culture supernatants [46]. Significantly, mutants were also devoid of capsule. This phenotype was also observed when incubating bacteria with capsidin, a CapD inhibitor [14, 44]. Thus, the hypothesis was modified and currently it is accepted that CapD participates in capsule anchoring to the cell wall. The controversy regarding CapD's function was finally settled with the demonstration that in cell free assays, the enzyme could mediate a transpeptidase reaction cross-linking oligopeptides of γ-D-Glu to peptidoglycan's meso-DAP [14]. In order for CapD to perform the GGT reaction, both the specific donor and acceptor had to be present. In the absence of acceptor (DAP-containing pentapeptides or muropeptides), CapD principally catalyzed a hydrolysis reaction. Therefore, in cell-free assays the aforementioned controversy involving CapD-mediated capsule hydrolysis was due to enzymatic reactions lacking CapD's physiological acceptors [14]. A similar hydrolysis reaction has also been reported for a member of the SortC family of transpeptidases when the physiologically relevant acceptors are omitted during the enzymatic reaction.

Using CapD as a model for a GGT involved in the biosynthesis of an amino acid polymer such as a *B. anthracis* poly-γ-D-Glu, a similar function was envisioned for Rv2394 during *Mtb* biosynthesis of a peptidic polymer such as α-L-poly-Gln. As discussed above, not only was CapD shown to cross-link poly-γ-D-Glu and peptidoglycan, but it also mediated an auto-transpeptidation reaction resulting in the elongation of higher molecular weight γ-D-Glu oligomer [14]. The fact that in contrast to eukaryote GGT, recombinant Rv2394 was able to catalyze a GGT reaction in the presence of acceptors mimicking α-L-poly-Gln such as di-Glu, di-Gln, penta-Gln and oligo-Glu, supports this activity. Furthermore, reduced Gln and ammonia content in cell-walls of *Mtb* mutants for Rv2394, strongly substantiates the hypothesis that this enzyme participatates in α-L-poly-Gln biosynthesis (see Chapter IV). It should be kept in mind that *Mtb* α-L-poly-Gln was described as exclusively consisting of α-linkages [15]. *In vitro* modeling of poly-CAG extending diseases (the CAG codon encodes for Gln), has shown that Gln oligopeptides consisting of more than 40 α-linked residues are highly insoluble due to extensive β-sheet formation [49]. If kinks are introduced in the polymer's structure, its insolubility is reduced [50]. Similar to the protein-

aggregating phenotype explaining the pathology of poly-GAG extending diseases, generation of such insoluble products could possibly present problems to *Mtb*. As an alternative during α-L-poly-Gln biosynthesis, smaller but soluble α-L-Gln oligopepeptides could be synthesized intracellularly. After being exported, the α-L-Gln soluble oligopeptides could be cross-linked by Rv2394 through a γ-linkage as depicted in Figure 3.19.

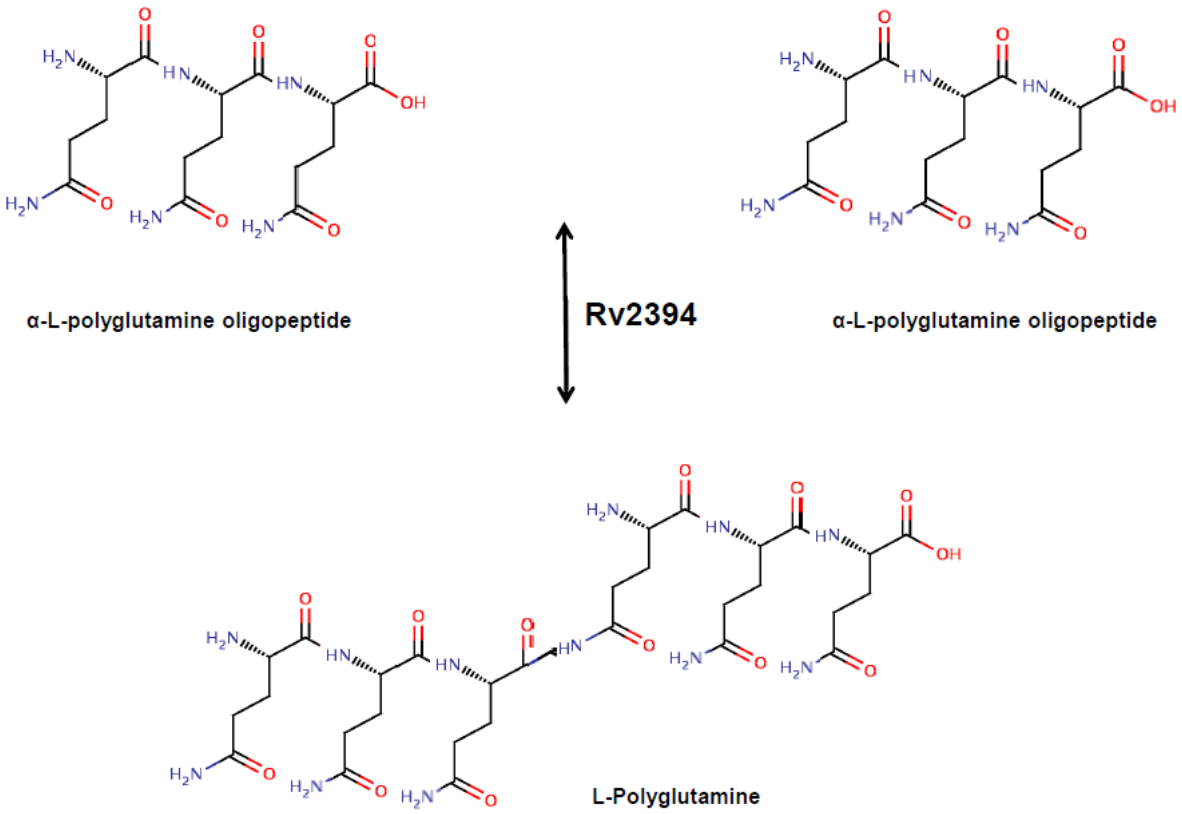


Figure 3.19 Role of Rv2394 in α-L-polyGln biosynthesis. α-L-polyGln oligopeptides would be synthesized in the cytoplasm and transported extracellularly by the Rv0547c locus (see Chapter IV). In the extracellular compartment, Rv2394 would crosslink α-L-polyGln oligopeptides through a γ-linkage. Consistent with the currently proposed structure, most Gln residues would be in an α-linkage and only a minority of residues would be in a γ-linkage. Similar to CapD, Rv2394 could have a polymerizing or depolymerizing activity depending on the physiological conditions.

Detecting a minor population of γ-linkages in α-L-poly-Gln's structure could thus be challenging, resulting in the erroneous assumption that this polymer exclusively consists of α-linkages [15].

In *B. anthracis*, two distinct "populations" of poly-γ-D-Glu have been identified. The first "population" corresponds to the aforementioned poly-γ-D-Glu which is covalently linked to the cell wall [14, 46]. Furthermore, a significant amount of free poly-γ-D-Glu is present in the culture filtrate representing a second "population" [46]. In contrast, only cell wall-associated but non-covalently bound α-L-poly-Gln has been described in *Mtb*. The conclusion that α-L-poly-Gln is not covalently attached to the *Mtb* cell wall, was drawn by the fact that both products could simply be separated by a physical method such as density gradient ultracentrifugation [15]. It should be stressed that the cell wall purification method for *B. anthracis* poly-γ-D-Glu analysis is significantly more stringent than the equivalent for α-L-poly-Gln analysis in *Mtb* cell walls. *Mtb* cell wall purification basically consists of three 2% SDS extractions at 60°C for 2 h [15]. In contrast, *B. anthracis* cell wall purification includes boiling in 1% SDS, protease digestion for 12 h, extraction with 8M $LiCl_2$, and finally, hydrolysis with hydrofluoric acid to remove polysaccharides [11]. Some of these procedures such as the extraction with 8M $LiCl_2$ and hydrofluoric acid treatment could be included during the extraction of *Mtb* cell wall in order to unequivocally conclude that α-L-poly-Gln is definitely non-covalently attached to the *Mtb* cell wall. Alternatively and as described for poly-γ-D-Glu, two "populations" of α-L-poly-Gln could be associated to the *Mtb* cell wall by both covalent and non-covalent bonds. Free α-L-poly-Gln would remain non-covalently associated to the *Mtb* cell wall, as its insolubility would preclude diffusion into the culture filtrate. Meanwhile and paralleling CapD's activity, Rv2394 could be covalently cross-linking a second "population" of α-L-poly-Gln to the *Mtb* cell wall. If this is the case, then meso-DAP present at the third position of mycobacterial pentapeptides would be the most probable candidate. For all peptidoglycan-modifying transpeptidases (PBPs, Ldts, sortases and GGTs), the amino acid located at the third position seems to be the targeted "hot spot" [14, 38-41]. As with CapD, the likelihood of detecting Rv2394's putative transpeptidase reaction in a cell free assay would be increased by including acceptors

that better resemble *Mtb* peptidoglycan (such as DAP containing pentapeptides), as well as donors mimicking α-L-poly-Gln (such as di or penta-Gln).

One major characteristic that was experimentally shown to be important in Rv2394 and is shared amongst GGTs is autocatalytic activation [24, 51]. In contrast to the other transpeptidases such as the Penicillin-Binding Proteins, sortases and Ldts, GGTs belong to the family of N-terminal nucleophile (Ntns) proteins [52, 53]. Ntns include proteins with diverse functions such as the proteasome, asparaginase, penicillin acylase, aspartylglucosaminidase, and Gln phosphoribosyl-pyrophosphate amidotransferase [52]. More importantly, Ntns share a three-dimensional structure characterized by the pre-protein's autocatalytic loop lying in close proximity to the enzymes' catalytic site, which facilitate the activation process [52-54]. Specifically for GGTs, the crystal structure for *B. anthracis* CapD, *H. pylori* and *E. coli* GGT crystal structures, has confirmed this [51, 55, 56]. In all GGTs described so far, autocatalytic activation is dependent on a conserved Thr residue. Ser and Cys are alternative nucleophiles in some members of the Ntn family [52]. Consistent with the characterized GGTs, mutating Rv2394's conserved Thr residue to Ala completely abrogated its autocatalytic and transpeptidase activity. In contrast, mutation of the Thr residue to a conserved amino acid such as Ser did not affect Rv2394's autocatalytic activation. Interestingly, transpeptidase activity over a 30 min time frame was not observed for the Rv2394 TS construct. Similarly, despite the occurrence of autocatalytic activation for an *E. coli* GGT TS mutant, an altered Km and V_{max} was reported for this mutant's transpeptidase reaction [24]. *E. coli's* TS mutant required days instead of minutes to perform the transpeptidase reaction [24]. Furthermore, some Ntns such as the proteasome are highly susceptible to a TS mutation [57]. Thus, as observed for Rv2394 TS, a conserved mutation affecting the catalytic Thr residue uncoupled the autocatalytic and transpeptidase reactions. Uncoupling both reactions in GGTs has additionally been achieved pharmacologically. Whereas 6-Diazo-5-OxoNorleucine (DON) readily inhibited transpeptidase activity, GGT's autocatalytic activity was preserved [24]. Finally, a rather unexpected finding was observed when expressing the Rv2394–20SP construct. As this construct had a wild type catalytic site, it

was expected to undergo autocatalytic activity. In contrast, an un-cleaved pre-protein form was observed during expression checks. Except for *Neisseria meningitides* GGT [58] and maybe *Mtb* Rv0773c which seems to lack a signal peptide, the remaining GGTs have been characterized as extracellular proteins. Thus, auto-catalytic activation might have evolved to occur only after the GGT has correctly folded once in the extracellular compartment. In *Mycobacterium*, one major post-translational modification occurring in the extracellular compartment is glycosylation [59]. Indeed, Rv2394 was shown to be glycosylated and in agreement with human GGT [60], the lack of this post-translational modification might affect its folding and autocatalytic activation. Alternatively, Rv2394 is unique amongst bacterial GGTs in having Cys residues in the extracellular domain. In eukaryote GGTs, it has been well established that these Cys residues participate in disulphide bond formation which are critical for proper folding and autocatalytic activation [61]. By remaining in the cytoplasm, Rv2394–20SP might lack disulphide bonds, thus the protein might not fold properly to support autocatalytic activation.

Besides autocatalytic activation, three additional post-translational modifications including glycosylation, acylation and phosphorylation were putatively identified in recombinant Rv2394. Both glycosylation and acylation of Rv2394 were expected based on the accurate and high predictive value obtained when subjecting mycobacterial proteins to bio-informatics analyses such as NetOGlyc, SignalP and LipoP, respectively. Our own bio-informatic analyses suggest such modifications might also be present in CapD. The presence of hexose residues attached to Rv2394 was confirmed by ConA lectin blot and mass spectrometry analysis showing the characteristic neutral loss of 162.0528 m/z. Glycosylated peptides identified by mass spectrometry and ConA reactivity was limited to the L-subunit, whereas Rv2394's S-subunit was not found to be glycosylated using either technique. As seen for other mycobacterial glyco-lipoproteins [62], Rv2394's putative glycosylated Thrs/Ser residues clustered at the N-terminus (L subunit). In agreement with our result, ConA reactivity of Rv2394's L-subunit had been reported in a study identifying glycosylated *M. tuberculosis* proteins obtained from culture filtrates[63]. As our study complemented ConA reactivity with mass spectrometry analysis to confirm glycosylation of

Rv2394, to the best of our knowledge this is the first biochemical demonstration for a glycosylated bacterial GGT. The fact that most studies evaluating bacterial GGT activity and crystal structure have been performed with recombinant proteins expressed intracellularly in *E. coli*, could explain why glycosylation had not been previously reported for bacterial GGTs. In contrast to bacterial GGTs, there is ample evidence that higher eukaryote and fungal GGTs are heavily glycosylated [17, 60]. For human GGT, approximately one third of the protein's molecular weight has been attributed to N-linked carbohydrates [60]. The S-subunit has also been shown to be glycosylated in eukaryote GGTs. Interestingly, a recent report with human GGT identified glycosylation as an essential step required for its auto-catalytic activity [60]. On the other hand, enzymatic de-glycosylation of the processed GGT did not affect its transpeptidase activity, further confirming that both enzymatic activities can be uncoupled [60]. Regarding Rv2394 acylation, this post-translational modification was evaluated by Tx-114 partitioning [30]. As reported for acylated proteins, Rv2394 partitioned to the detergent phase as determined by anti-His reactivity against Rv2394's S-subunit. It should be kept in mind that after Rv2394's autocatalytic activation, the acyl groups responsible for detergent partitioning are present in the L-subunit. In contrast, the His-Tag is present in the S-subunit. Thus both subunits seemed to remain associated during partitioning despite only being non-covalently associated. The crystal structures of *H. pylori*, *E. coli* and *B. anthracis* GGTs have identified considerable interaction between both subunits and this could explain why they remained associated during detergent partitioning [51, 55, 56]. A comprehensive study evaluating putative mycobacterial lipoproteins by Tx-114 partitioning also detected Rv2394 in the detergent phase [64]. Furthermore, this study reported the detection by mass spectrometry of peptides from the S-subunit of Rv2394. Detection of peptides for the S-subunit of Rv2394 could be attributed to the presence of a minor population of immature Rv2394 at the pre-protein stage when both subunits are still covalently linked. Alternatively, the S-subunit of Rv2394 could have long hydrophobic domains that could be forcing its partitioning into the detergent phase. Using the Kyte/Doolittle algorithm at ProtScale ([65], http://web.expasy.org/cgi-bin/protscale/protscale.pl), two hydrophobic segments of approximately

30 amino acids each were present in the S-subunit (not shown). As this subunit consists of 198 amino acids, the presence of 60 hydrophobic amino acids accounting for almost one third of the subunit could have a substantial impact during detergent partitioning. More conclusive biochemical evidence is required to definitely confirm the presence of acyl groups in Rv2394's L-subunit.

Unexpectedly, phosphorylation seemed to be an additional post-translationally modification present in enzymatically-active Rv2394. A neutral loss of 79.98 m/z corresponding to a phosphate residue [32], was clearly and consistently seen in several scans during mass spectrometry analysis of Rv2394. This serendipitous finding was even more remarkable as glycosylation was also present in the phosphorylated peptide. Consecutive neutral losses of 162.0528 and 79.98 m/z corresponding to a hexose and phosphate residue respectively, were detected for these peptides suggesting that the hexose residue isn't the phosphorylated group. In *Mycobacterium*, glycosylation has been shown to occur in the external cell membrane leaflet during SecA-mediated export of secreted proteins and lipoproteins [59]. Thus, the presence of a sugar residue is a very good indicator that this glyco-phospho peptide has undergone the mycobacterial secretion pathway en route to its extracellular location. If this is the case, then Rv2394 would constitute the first extracellular mycobacterial protein proven to be phosphorylated. Unfortunately it was not possible to conclusively identify the glyco-phospho peptide, probably attributed to amino acid modifications occurring during protein purification and/or while perfoming trypsin in-gel digestion for mass spectrometry. Nevertheless, by analyzing the fragmentation of the glyco-phosphopeptide, it was possible to circumscribe it to longer versions of the peptide EVAPAAATENYLR. This peptide is localized in the L-subunit and was not predicted by NetOGlyc to be glycosylated. Longer versions of this peptide would have at least two Thr residues where both glycosylation and phosphorylation could co-exist. The significance and extent of this putative phosphorylation is not known, but some of the mentioned enzymatic assays were performed with the same batch of Rv2394 that was subsequently shown to be phosphorylated.

The current dogma is that phosphorylation occurs intracellularly where ATP is synthesized, but the presence and concentration of extracellular ATP, as well as the function of extracellular kinases is currently unknown. It is envisioned that Rv2394 phosphorylation could occur either intracellularly before secretion through the Sec translocon or extracellularly after protein folding. In contrast to eukaryotes, bacterial secretion through the Sec translocon does not necessarily occur simultaneously with translation (co-translational translocation) [66]. Considerable folding has been known to occur before translocation. Furthermore, *Mycobacterium* do not have a SecB protein which mediates ribosome stalling until docking to a translocon [67]. Therefore, it is possible that Rv2394 could be phosphorylated during its brief intracellular stage, specially taking in consideration that the putative phosphorylated residue lays close to the protein's N-terminus. Other alternative secretion mechanisms described in *Mycobacterium* spp, could be participating during Rv2394 secretion. The SecA2 pathway has been shown to mediate secretion of folded proteins lacking a signal peptide like superoxide dismutase and catalase [68]. Recently, this pathway was also shown to participate during secretion of specific lipoproteins containing a signal peptide and a lipobox [69]. Therefore it would be important to evaluate the SecA2-mediated secretion of a lipoprotein like Rv2394, which by folding intracellularly could become phosphorylated prior to secretion. An alternative secretion system that has been shown to mediate translocation of partially folded proteins is the twin-arginine transport (Tat) system. Interestingly, some mycobacterial β-lactamases involved in cell wall structure were recently shown to be exported by the Tat system [70], however Rv2394 does not have a Tat signal sequence. Finally, a previous report had documented PknB-mediated phosphorylation of Rv0016c, an extracellular penicillin-binding protein. Compatible with the requirements for tranlocation by the Tat system, our own bio-informatic analysis identified two arginines in the signal peptide of Rv0016c [71]. Interestingly, the phosphorylated residues are located extracellularly. This report did not address the mechanism of how extracellular residues become phosphorylaed.

Besides a role in α-L-poly-Gln biosynthesis, Rv2394 could be participating in other biochemical pathways, specifically in sulfur metabolism. In contrast to other *Mycobacterium* spp, *rv2394* lies in

between the sulfate reductase and Cys uptake loci of *Mtb* and *M. bovis [28]*. Sulfate acquisition by this intracellular bacteria could potentially be more complex that in other *Mycobacterium* spp., requiring the presence of these additional genes. As occurs during the GSH cycle, one possibility could be that Rv2394 participates in sulfur uptake after cleavage of host GSH. A similar role was described for *F. tularensis* CapD homologue, whereas no role in capsule biosynthesis was identified for this protein [6]. In this regard, it should be mentioned that in *Mycobacterium*, a different operon encoded by *rv1280c-rv1283c* has already been characterized as a GSH or Cys-Gly uptake system [72]. Paradoxically, these two peptides are toxic to both *Mtb* and *M. bovis* and mutations to this operon abrogated their GSH susceptibility [72, 73]. Thus, it is currently thought that both *Mtb* and *M. bovis* only have one functional GSH uptake system and this could rule out the possibility that Rv2394 has this role. Nevertheless, an uncharacterized or alternative function in sulfate metabolism should be further pursued for Rv2394 and not disregarded. For instance, recombinant Rv2394 could be incubated with GSH and the appearance of Glu or Cys-Gly monitored over time.

In conclusion, we report the first biochemical characterization of Rv2394, a *Mtb* GGT with significant homology to CapD. We believe the presence of transpeptidases in clinically relevant Gram positive and negative bacteria, and their involvement in critical aspects of bacterial pathogenesis, should drive the characterization of additional members of this group of enzymes.

Literature Cited

1. Meister, A., *Glutathione metabolism.* Methods Enzymol, 1995. **251**: p. 3-7.
2. Meister, A., S.S. Tate, and O.W. Griffith, *Gamma-glutamyl transpeptidase.* Methods Enzymol, 1981. **77**: p. 237-53.
3. Zhang, H., H.J. Forman, and J. Choi, *Gamma-glutamyl transpeptidase in glutathione biosynthesis.* Methods Enzymol, 2005. **401**: p. 468-83.
4. Suzuki, H., W. Hashimoto, and H. Kumagai, *Escherichia coli K-12 can utilize an exogenous gamma-glutamyl peptide as an amino acid source, for which gamma-glutamyltranspeptidase is essential.* J Bacteriol, 1993. **175**(18): p. 6038-40.
5. Shibayama, K., et al., *Metabolism of glutamine and glutathione via gamma-glutamyltranspeptidase and glutamate transport in Helicobacter pylori: possible significance in the pathophysiology of the organism.* Mol Microbiol, 2007. **64**(2): p. 396-406.
6. Alkhuder, K., et al., *Glutathione provides a source of cysteine essential for intracellular multiplication of Francisella tularensis.* PLoS Pathog, 2009. **5**(1): p. e1000284.
7. Elskens, M.T., C.J. Jaspers, and M.J. Penninckx, *Glutathione as an endogenous sulphur source in the yeast Saccharomyces cerevisiae.* J Gen Microbiol, 1991. **137**(3): p. 637-44.
8. Flahou, B., et al., *Gastric epithelial cell death caused by Helicobacter suis and Helicobacter pylori gamma-glutamyl transpeptidase is mainly glutathione degradation-dependent.* Cell Microbiol, 2011. **13**(12): p. 1933-55.
9. Shibayama, K., et al., *A novel apoptosis-inducing protein from Helicobacter pylori.* Mol Microbiol, 2003. **47**(2): p. 443-51.
10. Zarnowski, R., et al., *Histoplasma capsulatum secreted gamma-glutamyltransferase reduces iron by generating an efficient ferric reductant.* Mol Microbiol, 2008. **70**(2): p. 352-68.
11. Candela, T. and A. Fouet, *Bacillus anthracis CapD, belonging to the gamma-glutamyltranspeptidase family, is required for the covalent anchoring of capsule to peptidoglycan.* Mol Microbiol, 2005. **57**(3): p. 717-26.
12. Candela, T. and A. Fouet, *Poly-gamma-glutamate in bacteria.* Mol Microbiol, 2006. **60**(5): p. 1091-8.
13. Fouet, A., *The surface of Bacillus anthracis.* Mol Aspects Med, 2009. **30**(6): p. 374-85.
14. Richter, S., et al., *Capsule anchoring in Bacillus anthracis occurs by a transpeptidation reaction that is inhibited by capsidin.* Mol Microbiol, 2009. **71**(2): p. 404-20.
15. Hirschfield, G.R., M. McNeil, and P.J. Brennan, *Peptidoglycan-associated polypeptides of Mycobacterium tuberculosis.* J Bacteriol, 1990. **172**(2): p. 1005-13.
16. Braunstein, M., S.S. Bardarov, and W.R. Jacobs, Jr., *Genetic methods for deciphering virulence determinants of Mycobacterium tuberculosis.* Methods Enzymol, 2002. **358**: p. 67-99.
17. Hughey, R.P. and N. Curthoys, *Comparison of the size and physical properties of gamma-glutamyltranspeptidase purified from rat kidney following solubilization with papain or with Triton X-100.* J Biol Chem, 1976. **251**(24): p. 8763-70.
18. Morrissey, J.H., *Silver stain for proteins in polyacrylamide gels: a modified procedure with enhanced uniform sensitivity.* Anal Biochem, 1981. **117**(2): p. 307-10.
19. Dobos, K.M., et al., *Evidence for glycosylation sites on the 45-kilodalton glycoprotein of Mycobacterium tuberculosis.* Infect Immun, 1995. **63**(8): p. 2846-53.
20. Speicher, K.D., N. Gorman, and D.W. Speicher, *N-Terminal Sequence Analysis of Proteins and Peptides*, in *Current Protocols in Protein Science*. 2001, John Wiley & Sons, Inc.
21. Hellman, U., et al., *Improvement of an "In-Gel" digestion procedure for the micropreparation of internal protein fragments for amino acid sequencing.* Anal Biochem, 1995. **224**(1): p. 451-5.

22. Inoue, M., et al., *Identification of catalytic nucleophile of Escherichia coli gamma-glutamyltranspeptidase by gamma-monofluorophosphono derivative of glutamic acid: N-terminal thr-391 in small subunit is the nucleophile.* Biochemistry, 2000. **39**(26): p. 7764-71.
23. Castonguay, R., et al., *Kinetic characterization and identification of the acylation and glycosylation sites of recombinant human gamma-glutamyltranspeptidase.* Biochemistry, 2007. **46**(43): p. 12253-62.
24. Suzuki, H. and H. Kumagai, *Autocatalytic processing of gamma-glutamyltranspeptidase.* J Biol Chem, 2002. **277**(45): p. 43536-43.
25. Petersen, T.N., et al., *SignalP 4.0: discriminating signal peptides from transmembrane regions.* Nat Methods, 2011. **8**(10): p. 785-6.
26. Sutcliffe, I.C. and D.J. Harrington, *Pattern searches for the identification of putative lipoprotein genes in Gram-positive bacterial genomes.* Microbiology, 2002. **148**(Pt 7): p. 2065-77.
27. Sutcliffe, I.C. and D.J. Harrington, *Lipoproteins of Mycobacterium tuberculosis: an abundant and functionally diverse class of cell envelope components.* FEMS Microbiol Rev, 2004. **28**(5): p. 645-59.
28. Pinto, R., et al., *Sulfite reduction in mycobacteria.* J Bacteriol, 2007. **189**(18): p. 6714-22.
29. Wiker, H.G., M.A. Wilson, and G.K. Schoolnik, *Extracytoplasmic proteins of Mycobacterium tuberculosis - mature secreted proteins often start with aspartic acid and proline.* Microbiology, 2000. **146 (Pt 7)**: p. 1525-33.
30. Radolf, J.D., et al., *Identification and localization of integral membrane proteins of virulent Treponema pallidum subsp. pallidum by phase partitioning with the nonionic detergent triton X-114.* Infect Immun, 1988. **56**(2): p. 490-8.
31. Julenius, K., et al., *Prediction, conservation analysis, and structural characterization of mammalian mucin-type O-glycosylation sites.* Glycobiology, 2005. **15**(2): p. 153-64.
32. Mann, M., et al., *Analysis of protein phosphorylation using mass spectrometry: deciphering the phosphoproteome.* Trends Biotechnol, 2002. **20**(6): p. 261-8.
33. Tate, S.S. and A. Meister, *gamma-Glutamyl transpeptidase from kidney.* Methods Enzymol, 1985. **113**: p. 400-19.
34. Orlowski, M. and A. Meister, *GAMMA-GLUTAMYL-P-NITROANILIDE: A NEW CONVENIENT SUBSTRATE FOR DETERMINATION AND STUDY OF L- AND D-GAMMA-GLUTAMYLTRANSPEPTIDASE ACTIVITIES.* Biochim Biophys Acta, 1963. **73**: p. 679-81.
35. Thorne, C.B., C.G. Gomez, and R.D. Housewright, *Synthesis of glutamic acid and glutamyl polypeptide by Bacillus anthracis. II. The effect of carbon dioxide on peptide production on solid media.* J Bacteriol, 1952. **63**(3): p. 363-8.
36. Makino, S., et al., *Cloning and CO2-dependent expression of the genetic region for encapsulation from Bacillus anthracis.* Mol Microbiol, 1988. **2**(3): p. 371-6.
37. Tate, S.S. and A. Meister, *Stimulation of the hydrolytic activity and decrease of the transpeptidase activity of gamma-glutamyl transpeptidase by maleate; identity of a rat kidney maleate-stimulated glutaminase and gamma-glutamyl transpeptidase.* Proc Natl Acad Sci U S A, 1974. **71**(9): p. 3329-33.
38. Tipper, D.J. and J.L. Strominger, *Mechanism of action of penicillins: a proposal based on their structural similarity to acyl-D-alanyl-D-alanine.* Proc Natl Acad Sci U S A, 1965. **54**(4): p. 1133-41.
39. Mainardi, J.L., et al., *A novel peptidoglycan cross-linking enzyme for a beta-lactam-resistant transpeptidation pathway.* J Biol Chem, 2005. **280**(46): p. 38146-52.
40. Magnet, S., et al., *Identification of the L,D-transpeptidases responsible for attachment of the Braun lipoprotein to Escherichia coli peptidoglycan.* J Bacteriol, 2007. **189**(10): p. 3927-31.
41. Lavollay, M., et al., *The peptidoglycan of stationary-phase Mycobacterium tuberculosis predominantly contains cross-links generated by L,D-transpeptidation.* J Bacteriol, 2008. **190**(12): p. 4360-6.

42. Gupta, R., et al., *The Mycobacterium tuberculosis protein LdtMt2 is a nonclassical transpeptidase required for virulence and resistance to amoxicillin.* Nat Med, 2010. **16**(4): p. 466-9.
43. Dramsi, S., et al., *Covalent attachment of proteins to peptidoglycan.* FEMS Microbiol Rev, 2008. **32**(2): p. 307-20.
44. Wu, R., et al., *Crystal structure of Bacillus anthracis transpeptidase enzyme CapD.* J Biol Chem, 2009. **284**(36): p. 24406-14.
45. Uchida, I., et al., *Identification of a novel gene, dep, associated with depolymerization of the capsular polymer in Bacillus anthracis.* Mol Microbiol, 1993. **9**(3): p. 487-96.
46. Makino, S., et al., *Effect of the lower molecular capsule released from the cell surface of Bacillus anthracis on the pathogenesis of anthrax.* J Infect Dis, 2002. **186**(2): p. 227-33.
47. Scorpio, A., et al., *Capsule depolymerase overexpression reduces Bacillus anthracis virulence.* Microbiology, 2010. **156**(Pt 5): p. 1459-67.
48. Scorpio, A., et al., *Poly-gamma-glutamate capsule-degrading enzyme treatment enhances phagocytosis and killing of encapsulated Bacillus anthracis.* Antimicrob Agents Chemother, 2007. **51**(1): p. 215-22.
49. Perutz, M.F., *Glutamine repeats and inherited neurodegenerative diseases: molecular aspects.* Curr Opin Struct Biol, 1996. **6**(6): p. 848-58.
50. Sharma, D., et al., *Peptide models for inherited neurodegenerative disorders: conformation and aggregation properties of long polyglutamine peptides with and without interruptions.* FEBS Lett, 1999. **456**(1): p. 181-5.
51. Boanca, G., et al., *Autoprocessing of Helicobacter pylori gamma-glutamyltranspeptidase leads to the formation of a threonine-threonine catalytic dyad.* J Biol Chem, 2007. **282**(1): p. 534-41.
52. Oinonen, C. and J. Rouvinen, *Structural comparison of Ntn-hydrolases.* Protein Sci, 2000. **9**(12): p. 2329-37.
53. Brannigan, J.A., et al., *A protein catalytic framework with an N-terminal nucleophile is capable of self-activation.* Nature, 1995. **378**(6555): p. 416-9.
54. Seemuller, E., A. Lupas, and W. Baumeister, *Autocatalytic processing of the 20S proteasome.* Nature, 1996. **382**(6590): p. 468-71.
55. Okada, T., et al., *Crystal structures of gamma-glutamyltranspeptidase from Escherichia coli, a key enzyme in glutathione metabolism, and its reaction intermediate.* Proc Natl Acad Sci U S A, 2006. **103**(17): p. 6471-6.
56. Okada, T., et al., *Crystal structure of the gamma-glutamyltranspeptidase precursor protein from Escherichia coli. Structural changes upon autocatalytic processing and implications for the maturation mechanism.* J Biol Chem, 2007. **282**(4): p. 2433-9.
57. Kisselev, A.F., Z. Songyang, and A.L. Goldberg, *Why does threonine, and not serine, function as the active site nucleophile in proteasomes?* J Biol Chem, 2000. **275**(20): p. 14831-7.
58. Takahashi, H. and H. Watanabe, *Post-translational processing of Neisseria meningitidis gamma-glutamyl aminopeptidase and its association with inner membrane facing to the cytoplasmic space.* FEMS Microbiol Lett, 2004. **234**(1): p. 27-35.
59. VanderVen, B.C., et al., *Export-mediated assembly of mycobacterial glycoproteins parallels eukaryotic pathways.* Science, 2005. **309**(5736): p. 941-3.
60. West, M.B., et al., *Autocatalytic cleavage of human gamma-glutamyl transpeptidase is highly dependent on N-glycosylation at asparagine 95.* J Biol Chem, 2011. **286**(33): p. 28876-88.
61. Kinlough, C.L., et al., *Gamma-glutamyltranspeptidase: disulfide bridges, propeptide cleavage, and activation in the endoplasmic reticulum.* Methods Enzymol, 2005. **401**: p. 426-49.
62. Sartain, M.J. and J.T. Belisle, *N-Terminal clustering of the O-glycosylation sites in the Mycobacterium tuberculosis lipoprotein SodC.* Glycobiology, 2009. **19**(1): p. 38-51.
63. Gonzalez-Zamorano, M., et al., *Mycobacterium tuberculosis glycoproteomics based on ConA-lectin affinity capture of mannosylated proteins.* J Proteome Res, 2009. **8**(2): p. 721-33.

64. Malen, H., et al., *Definition of novel cell envelope associated proteins in Triton X-114 extracts of Mycobacterium tuberculosis H37Rv.* BMC Microbiol, 2010. **10**: p. 132.
65. •Gasteiger E., et al., *The proteomics protocols handbook*, in *Springer protocols*, J.M. Walker, Editor. 2005, Humana Press: Totowa, NJ. p. 571-607.
66. Feltcher, M.E., J.T. Sullivan, and M. Braunstein, *Protein export systems of Mycobacterium tuberculosis: novel targets for drug development?* Future Microbiol, 2010. **5**(10): p. 1581-97.
67. DiGiuseppe Champion, P.A. and J.S. Cox, *Protein secretion systems in Mycobacteria.* Cell Microbiol, 2007. **9**(6): p. 1376-84.
68. Braunstein, M., et al., *SecA2 functions in the secretion of superoxide dismutase A and in the virulence of Mycobacterium tuberculosis.* Mol Microbiol, 2003. **48**(2): p. 453-64.
69. Gibbons, H.S., et al., *Identification of two Mycobacterium smegmatis lipoproteins exported by a SecA2-dependent pathway.* J Bacteriol, 2007. **189**(14): p. 5090-100.
70. McDonough, J.A., et al., *The twin-arginine translocation pathway of Mycobacterium smegmatis is functional and required for the export of mycobacterial beta-lactamases.* J Bacteriol, 2005. **187**(22): p. 7667-79.
71. Dasgupta, A., et al., *The serine/threonine kinase PknB of Mycobacterium tuberculosis phosphorylates PBPA, a penicillin-binding protein required for cell division.* Microbiology, 2006. **152**(Pt 2): p. 493-504.
72. Green, R.M., A. Seth, and N.D. Connell, *A peptide permease mutant of Mycobacterium bovis BCG resistant to the toxic peptides glutathione and S-nitrosoglutathione.* Infect Immun, 2000. **68**(2): p. 429-36.
73. Venketaraman, V., et al., *Role of glutathione in macrophage control of mycobacteria.* Infect Immun, 2003. **71**(4): p. 1864-71.

Chapter IV

Mtb α-L-Polyglutamine: Biosynthesis and Role in Pathogenesis

4.1 Introduction.

Mtb is the causative agent of human TB, a disease annually affecting approximately 7-8 million people and causing almost 2 million deaths [1]. Even though much has been learned about this pathogen since its genome was sequenced, big gaps in knowledge are still present. During genome annotation [2], it was observed that approximately 16% of the open reading frames were hypothetical conserved proteins with unknown function. Furthermore, a precise biochemical function could only be attributed to 40% of the open reading frames and for the remaining 44% only a putative function could be suggested [2]. It is hoped that further biochemical characterization of *Mtb*'s proteome could lead to better understanding its virulence factors and how they contribute to pathogenesis. For instance, considerable proteome differences have been observed for several related *Mtb* clinical isolates causing outbreaks in New Jersey [3]. How and why proteome differences exist in a genetically conserved microorganism, is currently unknown.

One clearly identified *Mtb* virulence factor is its cell wall [4]. From a clinical point of view, the mycobacterial cell wall is a big challenge as its hydrophobicity represents a major barrier against antibiotic diffusion [5]. Furthermore, the cell wall is responsible for mycobacterial acid fastness which is conveniently used during routine microbiological staining. The mycobacterial cell wall is composed of extractable lipids, the mycolyl-arabinogalactan-peptidoglycan (mAGP) complex and a cell membrane [6]. Altogether they constitute a remarkably different structure than Gram positive and negative cell wall. The major extractable lipids are PDIM, the mannosylated glycolipids PIMs, LM, ManLAM, and the trehalose based lipids TDM and SL, amongst others [7]. Starting from the outside, mAGP consists of mycolic acids covalently linked to the non-reducing end of the arabinogalactan chain [6]. Mycolic acids are long α-alkyl-β-hydroxyl fatty acids with more than 60 carbon units and are considered the major hydrophobic barrier [8]. Arabinogalactan is an unusual polymer of furanose D-arabinan and D-galactose residues

which forms a covalent bridge between mycolic acids and peptidoglycan [6]. GlcNAc and MurNAc residues constitute the repeating disaccharide present in peptidoglycan and the MurNac residues are derivatized with a peptidic chain composed of L-Ala, D-Glu, mesoDAP and D-Ala [6]. This peptidic chain is crosslinked by PBPs forming the classical 3→4 linkage (DAP→D-Ala) [9] and/or Ldts making a 3→3 linkage (DAP→DAP) [9-11].

In comparison to other microorganisms, several chemical modifications have been documented in *Mtb*'s peptidoglycan and seem to be important for the bacterium's physiology. For instance, it has been shown that the D-Glu residue could be amidated or modified with a Gly residue [6]. Addtionally, the MurNAc residue is glycolated [12] as opposed to the usual acetylation present in most bacteria. It has been well documented that the mycobacterial cell wall confers *Mtb* with some of the properties making it such a resilient microorganism. At the cell biology level, *Mtb* cell wall components have been shown to participate in blocking phagosome-lysosome fusion [13]. Furthermore, it has been suggested that the cell wall protects *Mtb* from the macrophage's oxidative burst [14], and, *in vivo* studies have shown that cell wall components participate in granuloma formation [15] and immune response modulation [16]. Taking in consideration the importance of *Mtb*'s cell wall, its biosynthesis is currently a major antibiotic target. Specifically, INH and ethambutol are included in the first line of anti-TB drugs as they interfere with mycolic acid and arabinogalactan biosynthesis, respectively [5]. It is hoped that more targets can be identified by understanding the *Mtb*'s cell wall biosynthesis. This is exemplified by the recent observation that Ldt activity mediates a different peptidoglycan crosslinking pattern required for *Mtb*'s persistence in chronic infections [10]. Thus, Ldt inhibition has been considered a new antibiotic target [17].

Several decades ago it was noticed that the cell wall obtained from virulent vs. avirulent *Mycobacterium* spp, differed in some immunomodulatory properties. For example, the cell wall of *M. bovis* BCG was shown to have adjuvant properties not present in *M. smeg* and this was particularly important as BCG but not *M. smeg* had therapeutic benefits in several types of cancer [18, 19]. Thus, it was considered that unique molecules in virulent *Mycobacterium* could account for these

immunomodulatory properties and characterization of mycobacterial cell walls indeed confirmed biochemical differences [16]. SL and PDIM were identified in the pool of extractable lipids obtained exclusively from virulent *Mycobacterium* spp. TDM was also initially characterized from virulent *Mtb* and determined to confer its characteristic serpentine colony morphology, therefore it is also known as 'cord factor'. However, it was later shown that TDM was present in other avirulent *Mycobacterium*. In this case, biochemical differences between the TDM obtained from virulent vs avirulent strains is attributed to subtle chemical modifications such as in its mycolate chains. In addition to lipid modifications, it was reported that cell walls of virulent *Mycobacterium* also contained two polymers consisting of hydrosoluble glucan (a polymer of glucose) [20] and insoluble poly-Glu/Gln that was present in peptidoglycan preparations [21]. As a similar polymer consisting exclusively of Glu residues had been identified in *B. anthracis* capsule and shown to be important for its virulence [22-24], further studies were carried out to characterize *Mtb*'s poly-Glu/poly-Gln.

The initial observation that led to the eventual characterization of *Mtb* poly-Glu/poly-Gln was done by Migliore *et al* in 1966 [21], when his group identified high Glu concentrations in the cell wall of *Mtb* clinical isolates. The high Glu concentrations lead to a Glu/DAP ratio of more than 30:1, in contrast to the expected ratio of 1:1 (one Glu and one DAP residue per peptidoglycan chain). The high Glu/DAP ratio was present after delipidation, urea extraction and trypsin/chymotrypsin digestion of cell walls, suggesting the excess Glu did not originate from conventional cell wall proteins and was not present as free Glu [21]. Furthermore, cell walls obtained from *Mtb* H37Rv, H37Ra and *M. bovis* BCG had significantly less Glu but still the ratio of Glu/DAP was higher than the expected ratio of 1:1 [21]. Taking in consideration that during amino acid analysis peptide bonds are cleaved during acid hydrolysis, it was suggested that the excess Glu was covalently linked and constituted a polymer similar *B. anthracis* poly-Glu [21]. Indeed it was subsequently shown that a Glu-containing polymer could be extracted from both *Mtb* and *M. bovis* cell walls after mild acid hydrolysis [25]. In contrast to *B. anthracis* poly-Glu, enzymatic analysis with L-Glu decarboxylase showed the Glu to be in L-configuration and covalently

joined α-linkages [25-27]. Milder purification techniques of *Mtb*'s polymer resulted in the identification that some Glu residues could be amidated, corresponding to Gln residues instead [27]. Finally, purification of *Mtb*'s polymer by sucrose gradient ultracentrifugation without acid hydrolysis confirmed the polymer was composed entirely of L-Gln residues covalently joined by α-linkages [28], hence the name α-L-polyGln which will be used hereafter.

The capsule of *B. anthracis* is entirely composed of poly-γ-D-Glu and it's essential for virulence [22, 29]. The biosynthetic machinery of *B. anthracis* capsule is well understood and is known to be encoded in genes of the pX02 plasmid [23, 24]. *B. anthracis* that loose this plasmid are not encapsulated and are avirulent [23, 24]. Conversely, introducing the pX02 plasmid into non-encapsulated *Bacillus* spp, led to bacterial encapsulation [23]. Capsule biosynthesis activity has been specifically mapped to the *capA-capE* operon in the pX02 plasmid [29]. Biosynthesis of poly-D-Glu starts in the cell's cytoplasm where CapB activates the substrate in an ATP-dependent fashion [22, 30]. In concert with CapC, these two enzymes catalyze the covalent linkage of D-Glu residues via γ-linkages [30, 31]. CapA and CapE transport the elongating polymer to the extracellular space [30, 32], where CapD covalently links the capsule to DAP residues of *B. anthracis* peptidoglycan [33-35]. In contrast, biosynthesis of *Mtb*'s α-L-polyGln has not been elucidated. It has been reported that reduced quantities of α-L-polyGln were obtained in *Mtb*'s cell wall after inhibition of Rv2220 [36], an abundantly secreted Gln synthetase present in the CF [37]. Despite the presence of three additional Gln synthetase encoded in *Mtb*'s genome, Rv2220 accounts for the majority of the bacillus Gln synthetase activity [38]. Even though Rv2220 could have a direct role in α-L-polyGln biosynthesis, this would be highly unlikely taking in consideration that this enzyme is present and highly conserved in other *Mycobacterium* spp lacking α-L-polyGln [39]. Instead, inhibition of Rv2220 could have an indirect effect on α-L-polyGln as limiting intracellular Glu/Gln concentrations could be available for this polymer's biosynthesis.

The main objectives of this Chapter were to identify genes participating in the biosynthesis of *Mtb*'s α-L-polyGln and determine if this polymer had a role in animal models of TB. Taking in

consideration the biochemical similarities between *B. anthracis* poly-γ-D-Glu and *Mtb* α-L-polyGln, homologs of *B. anthracis capA-capE* operon were sought for in *Mtb*'s genome. Rv0574c and Rv2394 (discussed in Chapter III), were identified as proteins with significant homology to *B. anthracis* CapA and CapD, respectively. Reduced concentrations of ammonia and Glu residues were present in the cell walls of *Mtb* mutants for both genes, a good indication Rv0574c and Rv2394 could participate in α-L-polyGln biosynthesis. Amino acid analysis of WT and complemented strains was unsuccessful as the WT and possible the complemented strains stopped accumulating α-L-polyGln in the cell wall. Nevertheless, preliminary studies in the murine model of TB suggested that mutants for α-L-polyGln grew and disseminated faster in the initial stages of disease, but this difference with the WT strain normalized thereafter. Despite the subsequent normalization of lung and spleen bacterial burden, it appears that mutants for α-L-polyGln biosynthesis induce more lung inflammation and consolidation throughout the course of the infection. Additional studies with mutant and complemented strains should be pursued in order to confirm these results.

4.2 Materials and Methods.

4.2.1 Bioinformatic analyses.

The amino acid sequences for CapA, B, C, D and E were obtained from the annotated *B. anthracis* strain, 'Ames Ancestor' pXO2 plasmid (Accession number NC_007323). These sequences were individually analyzed by BLAST against the *Mtb* H37Rv proteome using Tuberculist. Rv0574c and Rv2394 were identified as CapA and CapD homologues, respectively. (Characterization of Rv2394 was discussed in Chapter III). The amino acid sequence for Rv0574c was obtained from Tuberculist and analyzed using the following bioinformatic analyses tools: BLAST® from NCBI and PROSITE scan, and ScanSite pI/Mw available at the Expasy Proteomics Server website. Genomic organization for the *rv0574c* locus was performed using the genome region comparison tool at JCVI comprehensive microbial resource webpage.

4.2.2 Genetic inactivation and complementation of *rv0574c* and *rv2394*.

Inactivation of *rv0574c* and *rv2394* was accomplished in *Mtb* H37Rv by a phage transduction system [40]. This system accomplishes an insertion/deletion event in which a substantial fragment of the open reading frame is removed and replaced by inserting a hygromycin resistance cassette (performed by Dr. Rainer Kalscheuer at Dr. William Jacobs' lab, Albert Einstein College of Medicine, Bronx, NY).

For complementation of the *rv0574c* mutant, two amplicons were obtained by PCR. The first amplicon consisted of *rv0574c* plus 500 bp upstream of the gene's start codon in order to include the putative endogenous promoter. The second amplicon encompassed both *rv0574c* and *rv0573c*, plus 500 bp upstream. This amplicon was engineered to circumvent the possibility that mutation of *rv0574c* also had a polar effect on the expression of the downstream gene *rv0573c*. A similar strategy was used for complementation of the *rv2394* mutant. Both amplicons contained *rv2394* plus 500 bp upstream of the gene's start codon. The second amplicon contained *rv2395* as well in case a polar effect had been introduced. Plasmids engineered in this Chapter are described in Table 4.1.

Table 4.1 Plasmid names and description

Plasmids	Description
pGEM-T+rv2462c	**Plasmid used to clone *rv2462c* used to synthesize Southern blot probe**
pMV306+rv0574c	**Integrative plasmid to complement the *rv0574c* mutant. Includes 500bp upstream of *rv0574c* start codon in order to use endogenous promoter. ***
pMV306+rv0574c_rv0573c	**Integrative plasmid to complement the *rv0574c* mutant. Also encodes for the downstream gene, *rv0573c*, in order to complement polar effects.**
pMV306+rv2394	**Integrative plasmid to complement the *rv2394* mutant.**
pMV306+rv2394_rv2395	**Integrative plasmid to complement the *rv2394* mutant. Also encodes for the downstream gene, *rv2395*, in order to complement polar effects.**

*** For all complementation plasmids, the constructs started 500bp upstream of the mutated gene's start codon, in order to include their putative promoter, ensuing gene expression in its physiological context.**

PCR primers for complementation studies are described in Table 4.2. PCR and cloning was performed as described in Chapter III. The PCR products were inserted in the *XbaI* and *ClaI* restriction sites present in pMV306, an integrative plasmid containing a kanamycin resistance cassette.

Table 4.2 PCR and qRT-PCR primers

A) PCR primers

[a] forRv2462c	CAT ATG AAG AGC ACC GTC GAG CAG TTG AGC CC
[b] revRv2462c	AAG CTT CGT TGT CGC TTC GTC GGA CG
[c] forRv0574c	ACT AGT GGC CGG GTG CAT CTC GCT GCT CGG
[d] revRv0574c	ATC GAT TTA CTC CTT GCT CGT TAG GTT GGC AGC
[d] revRv0573c	ATC GAT GTT CTA GGG TCG TTT GGC CTT CGC CCC G
[c] forRv2394	ACT AGT GCG TTG GTC ACC GGG TAC GCC AGG CCA GTG T
[d] revRv2394	ATC GAT TCA GGC ATC GTC GCC CAT GAC CGC GC
[d] revRv2395	ATC GAT TTA CCC AGG CAA ATC GGG CAG TGG CT

B) qRT-PCR primers

[e] rv0574c.14	CTG ATG TGG TGA CGG TGC TG
rv0574c.129	GGT CGC ATC CCG CAT ATA CC
rv0574c.366	GCT GGC CAA CAA CCA CAT TC
rv0574c.511	CAT GGC CAA CCG TGA CTA GC
rv0574c.1011	CTC CCA GAC CGA CAC CGA AT
rv0574c.1130	GTT AGG TTG GCA GCG GGA AC
rv0573c.75	CGA AAG AAT GTC GGG CAC AG
rv0573c.223	CAC GCA GGT AAC GCA GAT CC
rv1330c.732	CAC GCT GCT GGT GGA TAC CT
rv1330c.840	TAC CCC AAG CTC ACC GGA AT
rv2395.1093	GGT TAC ATG GCC GGG TTG AT
rv2395.1198	CCG GAC CAT ACG CAG TCT TG
rv2394.529	GGA CGG TCG ATC GGA GTA CC
rv2394.663	CCT GGC GCT GAT GTC AAA AC
rv2394.1382	TGA CGA CGA CGG TGG AAT CT
rv2394.1498	CCA CCG GTG ATC CGT CAG TA
rv0773c.1126	TTC ACT GTG GTG CCG AGA CA
rv0773c.1242	CAT CAC CCC GAA GCT CAT CA

[a] Underlined *NdeI* site
[b] Underlined *HindIII* site
[c] Underlined *SpeI* site
[d] Underlined *ClaI* site
[e] qRT-PCR numbers next to primer's name refers to the location of the primer's first base in the gene's sequence

4.2.3 Generating Electrocompetent *Mtb*.

With minor modifications, electrocompetent *Mtb* was generated as described before [41]. Both *rv0574c* and *rv2394* mutants were grown at 37°C in 7H9 (BD Difco) supplemented with OADC (BD Difco), 0.05% Tween 80 and 100 µg/ml of hygromycin (Invitrogen). When cultures reached an OD 600nm between 0.4-0.6, bacteria were harvested by centrifugation at 3,000 x g for 10 min at 4°C. After decanting the supernatant, bacteria were washed twice with ice cold water by inverting the cells and centrifuging as before. Finally, mutant H37Rv was washed with ice cold 10% glycerol in water followed by centrifugation. After decanting, bacteria were resuspended in 5 ml of 10% glycerol in water and 500 µl aliquots were stored at -80°C.

4.2.4 Transformation of mutant *Mtb* H37Rv.

Except for minor modifications, transformation of electrocompetent mutant *Mtb* H37Rv was accomplished as described for *M.smeg* (Chapter III) [41]. After plasmid electroporation, transformed *Mtb* was allowed to recover for 24 h at 37°C in 7H9 media supplemented with OADC and 100 µg/ml of hygromycin. Bacteria were then plated in 7H11 agar plates supplemented with OADC and containing 50 and 100 µg/ml kanamycin (Fisher) and hygromycin, respectively. Plates were incubated at 37°C for 3-4 weeks until single colonies were obtained. Thereafter, single colonies for each mutant were grown in 10 ml of 7H9 media supplemented with OADC and kanamycin plus hygromycin, as described above. After approximately 2 weeks at 37°C with gentle agitation, 5 ml of culture were harvested by centrifugation at 3,000 x g for 10 min and after decanting supernatants, bacterial pellets were frozen at -80°C until genomic DNA was extracted for Southern Blots. The remaining 5 ml cultures were upscaled to 100 ml in order to obtain cells for biochemical analysis and animal infections.

4.2.5 g DNA extraction.

Mtb H37Rv g DNA was extracted as described before with minor modifications [42]. After thawing frozen bacterial pellets, they were incubated O/N at 37°C in a solution consisting of 50% sucrose, 1 M Tris pH 8.0, 0.5 M EDTA pH 8.0 and 600 µg/ml of lysozyme. Thereafter, bacteria were incubated with 200 µg/ml of proteinase K (Sigma), 1% SDS for 4 h at 55°C and then extracted with an equal volume of 25:24:1 phenol:chloroform:isoamylalcohol (Sigma) for 30min at RT. After centrifugation at 13,000 x g for 10 min, the aqueous layer was transferred to a new tube and the DNA was precipitated O/N at -20°C after the addition of 1/10 the volume of 3 M sodium acetate pH 5.2 and 2 volumes of 100% ice cold ethanol. Precipitated DNA was recovered by centrifuging at 13,000 x g for 30 min, supernatants were discarded and the pellet was washed with 1 ml of 70% ethanol. Centrifugation was carried out as before, the supernatant was removed and the DNA was allowed to dry by evaporation. DNA was resuspended in 100 µl of TE buffer pH 8 and quantified with a NanoDrop (ThermoScientific).

4.2.6 Southern Blot.

To generate the probe for Southern blot, *rv2462c* was PCR amplified using the primers in Table 4.2. As described in Chapter III, the PCR amplicon was inserted in pGEM-T Easy (Promega) and cloned into TOP10 *E.coli* (Invitrogen). Plasmids obtained from resulting colonies were confirmed by sequencing at PMF in CSU. A plasmid containing the correct insert was digested with *NdeI* and *HindIII* and the dropout consisting of *rv2462c* was gel-purified. Gel-purified *rv2462c* was used to generate a biotinylated probe, according to the manufacturer's protocol (New England Biolabs). Briefly, 100 ng of gel purified *rv2462c* was denatured by boiling for 5 min and incubated with 1X labeling mix (containing biotinylated random primers), dNTPs (containing biotinylated ATP) and 5 U Klenow fragment (3'-5' exo). The PCR reaction was allowed to proceed for 3 h at 37°C and after quenching with EDTA, DNA was precipitated with ethanol. Precipitated probe was pelleted by centrifugation at 13,000 x g for 30 min, supernatant was decanted and the probe was allowed to dry.

Five µg of *Mtb* H37Rv g DNA was digested O/N at 37°C with 20 U of *AvrII* (New England Biolabs). Digested DNA was resolved in a 0.8% agarose gel and depurination was peformed with 0.25 M HCl for 10 min. This reaction was stopped with 3 washes of 0.4 M NaOH for 15 min. The DNA was subsequently blotted O/N to an Amersham Hibond-N nylon membrane (GE Healthcare Life Sciences) by capillarity using 20x SSC buffer. Using an UV Cross linker (Fisher), blotted DNA was cross-linked to the membrane with UV light at 1200 x 100 µJ/cm^2. Thereafter, the membrane was prehybridized in a hybridization oven (VWR) at 68°C for 1h with 5 X Denhardt's reagent, 0.5% SDS, 6X saline-sodium-phosphate-EDTA (SSPE) and 100 µg/ml of Salmon Sperm DNA (Sigma). Biotinylated probe was denatured by boiling for 5 min and after keeping in ice for 2 min, was added to the prehybridization solution. Hybridization was performed O/N at 68°C. The reaction was detected using the Phototope®-Star Detection Kit (New England Biolabs). Briefly, the membrane was blocked for 5 min with a solution containing 5% SDS, 25 mM sodium phosphate, pH 7.2. Thereafter, streptavidin to a final concentration of 1 µg/ml was added to the blocking solution and incubated for 5 min. After washing two times, the membrane was incubated for 5 min with 0.4 µg/ml biotinylated alkaline phosphatase in blocking solution. Once again, the membrane was washed twice and incubated for 5 min with 0.25 mM CDP-Star Reagent in 1X CDP-Star Assay Buffer. Finally, the membrane was washed and exposed for 5 min to a fragment of clear blue X-ray film. The film was developed using an X-OMat Processor (Eastman Kodak, Rochester, NY).

4.2.7 qRT-PCR.

With some modifications mRNA was obtained as described before [43]. Briefly bacterial pellets were transferred to 1 ml of Trizol (Invitrogen) containing silica beads and stored at -80°C until analyzed. After thawing samples in ice, bacteria were broken by 3 cycles of bead beating at 4800 rpm for 3 min. Trizol was transferred to a tube with an equal volume of chloroform and after centrifuging at 13,000 x g for 10 min, the aqueous layer was obtained and subjected to a second chloroform extraction. The aqueous

phase was precipitated with an equal volume of isopropanol and centrifuged at 13,000 x g for 10 min to pellet the RNA. Thereafter, the RNA pellet was washed with 80% ethanol and DNA was digested with DNaseI (Fermentas) for 20 min at 37°C. DNaseI was removed by extracting with Phenol and centrifugation as described above. Finally, the aqueous layer was precipitated at -80°C with ammonium acetate and ethanol as described for the DNA. RNA pellets were obtained by centrifugation, resuspended in RNase free water and quantified by spectrophotometry with a NanoDrop.

RNA was converted to cDNA using the Superscript III First-Strand Synthesis for RT-PCR kit (Invitrogen) [43]. Briefly, RNA was incubated with 50 ng of random hexamers, 200 U of Superscript III, 40 U of RNaseOut, 20 mM DTT, 6 mM $MgCl_2$ and 1X RT buffer. cDNA synthesis mix was then incubated for 10 min at 25°C, and the reverse transcriptase step was performed for 50 min at 50°C. Finally, after cDNA synthesis, RNA was removed with RNaseH for 20 min at 37°C. cDNA was then normalized to 50 ng/ul for the qRT step.The mix was prepared with 50 nM forward and reverse primers (see qRT primers in Table I), 50 ng of cDNA and 1X SYBR green. The thermocycler was programmed to perform the following: step 1, 55°C for 2 min; step 2, 95°C for 2 min; step 3, 45 cycles of 95°C for 15 sec, 60°C for 30 sec and 72°C for 45 sec; step 4, 72°C for 5 min. After the reaction was completed, specificity was evaluated by performing a melting curve between 30°-94°C. Data was analyzed using the iQ™5 Optical System Software Version 2.1 (Bio-Rad). To obtain mRNA concentration for each gene, CT values were extrapolated to curves containing 1, 10 and 100 ng of g DNA. Data was normalized to *sigA* or 16SrRNA (not shown).

4.2.8 Bacterial culture for biochemical analysis.

Mutant and complemented *Mtb* H37Rv were grown at 37°C with constant agitation in 100 ml of GAS medium containing 50 µg/ml of Kanamycin and 100 µg/ml of Hygromcyin. Thereafter, cultures were upscaled by transferring bacterial pellets to FB containing 900 ml of the same medium and grown for 15 days with constant agitation. Finally, bacterial pellets were used to inoculate FB containing GAS

plus antibiotics but cultures were left at 37°C without agitation (static cultures) to induce pellicle formation. Similar to Migliore *et a*l [21], the starting material for the current cell wall analysis was bacteria harvested from pellicles. Approximately 30 days after inoculating the FB, a thick pellicle had grown on the surface of the media at the interphase between air and liquid. Bacterial pellicles were gently harvested with a cell scraper avoiding bacteria that had sedimented to the bottom of the FB instead of growing on the media's surface. Bacteria were pelleted by centrifugation at 3,500 x g for 15 min, the supernatant was discarded and after freezing at -80°C, bacteria were inactivated by γ-irradiation.

4.2.9 SDS-extracted cell wall.

Cell wall extraction was performed as described by Hirschfield *et al* [28]. γ-irradiated *Mtb* was resuspended as a 2:1 weight: volume ratio in PBS containing 1 μg/ml of both DNaseI (Sigma) and RNaseA (Sigma). Probe sonication was performed with 3 cycles of 90 sec intermittent pulsing at 50% duty cycle in a Cole Palmer (Vernon Hills, IL) sonicator. The slurry was then passed seven times through a French Pressure Cell press (Spectronic Instrument, Rochester, NY) at 1200 psi to break cells. Unbroken cells were removed by centrifugation at 3000 x g for 5 min and the supernatant was centrifuged at 27,000 x g for 1 h in order to obtain cell walls. Thereafter, the cell wall was extracted for 2 h at 60°C with constant stirring in PBS containing 2% SDS. After SDS extraction, cell wall was obtained by centrifugation at 27,000 x g for 30 min, supernatants were discarded and the SDS extraction procedure was repeated two more times.

To remove SDS, cell walls were washed three times with water for 30 min with constant stirring and collected by centrifugation as described above. Finally, cell walls were extracted with 80% acetone for 30 min in constant stirring. After pelleting the cell wall by centrifugation, it was dried under a stream of nitrogen and kept at -80°C until further analysis.

4.2.10 Cell wall amino acid analysis.

Samples were analyzed as recommended by the manufacturer of the AccQ Tag amino acid analysis kit (Waters Corporation, Milford, MA). Briefly, cell wall was resuspended at 25 mg/ml in 6 N HCl and hydrolyzed under constant reflux for 20 h at 112°C in 13x100 glass tubes. After centrifuging tubes for 15 min at 3,000 x g to remove insoluble material, the supernatant was harvested. An aliquot of the hydrolyzates was diluted 60 fold in water so as to reduce the HCl concentration to 0.1 N. Samples were then filtered through at 0.2 µM filter to further remove any insoluble material. As internal standard (IS), α-aminobutyric acid (MP Biomedicals, Santa Ana, CA) was added to the samples. The sample's pH was then neutralized by mixing the same volume of sample and sodium borate. Neutralized samples (20 µl) were derivatized with 5 µl of the AccuTag reagent at 55°C for 10 min. Thereafter, 20 µl of derivatized sample was resolved on a 3.9 x 150 mm AccQ Tag Amino Acid Analysis Column using the Waters HPLC 2695 Separation Module System. The column was kept at 37°C and run at a constant rate of 1 ml/min. Using this derivatization protocol, 500 pmols of IS would be present in the 20 µl of analyzed sample. Amino acid standards (Pierce) were analyzed the same way after addition of an equivalent concentration of DAP (Sigma) and IS. It should be noted that commercial DAP was obtained as meso-DAP and in the amino acid analysis two peaks with similar intensity (250 pmol) were observed.

Amounts of ammonia, Glu and Ala were determined as follows =

$$\frac{500 \text{ pmol} \times (\text{area of the amino acid in the sample/area of the IS in the sample})}{(\text{area of the amino acid in the control/area of the IS in the control})}$$

Amount of DAP was determined as follows=

$$\frac{250 \text{ pmol} \times (\text{area of the amino acid in the sample/area of the IS in the sample})}{(\text{area of the amino acid in the control/area of the IS in the control})}$$

Data was normalized to the amount of DAP and expressed as a ratio of the respective amino acid (or ammonia)/DAP.

4.2.11 Cell wall lipid analysis.

Whole lipids were extracted from 100 mg of γ-irradited cells, by an O/N incubation under gentle stirring with 5 ml of 2:1 chloroform:methanol [44]. Bacteria were pelleted by centrifugation at 3,000 x g for 10 min and supernatants were saved. Bacterial pellets were extracted two additional times following the same procedure. Supernatants from the three extractions were pooled and dried under a nitrogen stream. As described in the results section, lipids were analyzed by TLC in solvents with different polarities and developed by cupric sulfate spraying.

Mycolic acids were extracted from 10mg of SDS-extracted cell wall core. Mycolic acids were hydrolyzed from mAGP by incubating with 15% tetrabutyl ammonium hydroxide for 2 h at 110°C [44]. Released mycolic acids were then methylated with dichloromethane and iodomethane by agitation at RT for 1hr. After centrifuging at 3,000 x g for 10 min, the organic layer was transferred to a new tube, resuspended in 1ml diethyl ether and the soluble material was transferred to another tube. After drying under a nitrogen stream, mycolic acid methyl esters (MAMEs) were resuspended in dichloromethane. Finally, MAMEs were analyzed by TLC using 95:5 petroleum ether:ether and visualized after spraying with cupric sulfate.

4.2.12 Animal studies.

Six to nine week female C57BL/6 mice were purchased from the Jackson Laboratory (Bar Harbor, ME). Mice were infected with the Glas-col (Terre Haute, IN) aerosol chamber and the target dose was 100 CFUs per animal [45]. At day 1 post-infection, three mice were euthanized by CO_2 asfixiation, their lungs homogenized and plated on 7H11 plus OADC plates to enumerate bacteria. CFUs were counted 3-4 weeks after incubating plates at 37°C.

At days 30, 90, and 180 post-infection, 5 mice per bacterial strain were euthanized. One lobe was perfused with 2% formalin and stored in 2% formalin until processed for histological analysis. The remaining lung lobes and spleen were processed for CFUs.

4.2.13 Histopathological analysis.

A stereology based analysis referred to as the Area Fraction Fractionator, was performed to determine the percent lung consolidation [46]. Representative slides were used to determine the lung and lesion/inflammation area. Multiple fields were randomly selected by the computer in order to minimize bias.

4.3 Results.

4.3.1 Bioinformatics analyses.

Taking in consideration the similarities between *Mtb* and *B. anthracis* α-L-polyGln and poly-γ-D-Glu respectively, it was reasoned that a common biosynthetic pathway could be present in both microorganisms. Thus, in the Tuberculist webpage, BLAST analysis was performed to identify *Mtb* homologues of the *B. anthracis* Cap proteins, which synthesize poly-γ-D-Glu. Whereas no homologues to CapB, C or E were detected to be encoded in the *Mtb* genome, Rv0574c (Fig. 4.1) and Rv2394 (see Chapter III) were identified as CapA and CapD homologues, respectively. Specifically Rv0574c and *B. anthracis* CapA shared 25% identical amino acids with an e score of 7^{-17}. Homology between these two proteins is mainly present at the protein's N-terminus, but not at the C-terminus. As no additional domains were found in Rv0574c, it is currently not possible to attribute a function to Rv0574c's C-terminus.

```
>M. tuberculosis H37Rv|Rv0574c|Rv0574c CONSERVED HYPOTHETICAL
          PROTEIN
         Length = 380

 Score = 82.0 bits (201), Expect = 7e-17
 Identities = 74/287 (25%), Positives = 124/287 (43%), Gaps = 45/287 (15%)

Query: 60  LTMTMVGDIMMGRHVKEIVNRYGTDYV----FRHVSPYLKNSDYVSGNFEHPV------- 108
           +T+ + GD+M+GR V +I+   G   +     R  + Y++ ++ V+G    PV
Sbjct: 8   VTVLLGGDVMLGRGVDQILPHPGKPQLRERYMRDATGYVRLAERVNGRIPLPVDWRWPWG 67

Query: 109 ----LLEDKK--------------NYQKAD-KNIHLSAKEETVKAVKEAGFTVLNLANNH 149
               +LE+                + + AD K +      + V A+      V  LANNH
Sbjct: 68  EALAVLENTATDVCLINLETTITADGEFADRKPVCYRMHPDNVPALTALRPHVCALANNH 127

Query: 150 MTDYGAKGTKDTIKAFKEADLDYVGAGENFKDVKNIVYQNVNGVRVATLGFTDAFVAGA- 208
           + D+G +G  DT+ A   A +  VGAG +    +      V   R   +G   A  +G
Sbjct: 128 ILDFGYQGLTDTVAALAGAGIQSVGAGADLLAARRSALVTVGHERRVIVGSVAAESSGVP 187

Query: 209 ---IATKEQPGS-LSMNP------DVLLKQISKAKDPKKGNADLVVVNTHWGEEYDNKPS 258
              A +++PG  L  +P      D +  Q+   K P     D+ +V+ HWG  +    +
Sbjct: 188 ESWAARRDRPGVWLIRDPAQRDVADDVAAQVLADKRP----GDIAIVSMHWGSNWGYATA 243

Query: 259 PRQEALAKAMVDAGADIIVGHHPHVLQSFDVYKQGIIFYSLGNFVFD 305
           P   A A  ++DAG D++ GH  H  +  ++Y+   I Y  G+ V D
Sbjct: 244 PGDVAFAHRLIDAGIDMVHGHSSHHPRPIEIYRGKPILYGCGDVVDD 290
```

Figure 4.1 *Mtb* Rv0574c is a homologue of *B. anthracis* CapA. The Cap proteins present in *B. anthracis* pXO2 plasmid have been identified to synthesize poly-γ-D-Glu. Thus, the CapA-E proteins were analyzed by the BLAST tool in Tuberculist web page to identify *Mtb* homologues that could be involved in α-L-polyGln biosynthesis. In addition to a CapD homologue (Rv2394, see Chapter III), Rv0574c was found to have high homology to CapA, specifically at the N-terminus. Several identical and conserved amino acids are depicted.

To identify conserved protein domains and additional Rv0574c homologues, the protein sequence of Rv0574c was analyzed by BLAST analysis using the non-redundant NCBI protein database. As can be observed in Figure 4.2, a CapA or PgsA domain (pfam09587) was identified in Rv0574c. The difference in nomenclature is in reference to capsular or secreted poly-Glu, respectively.

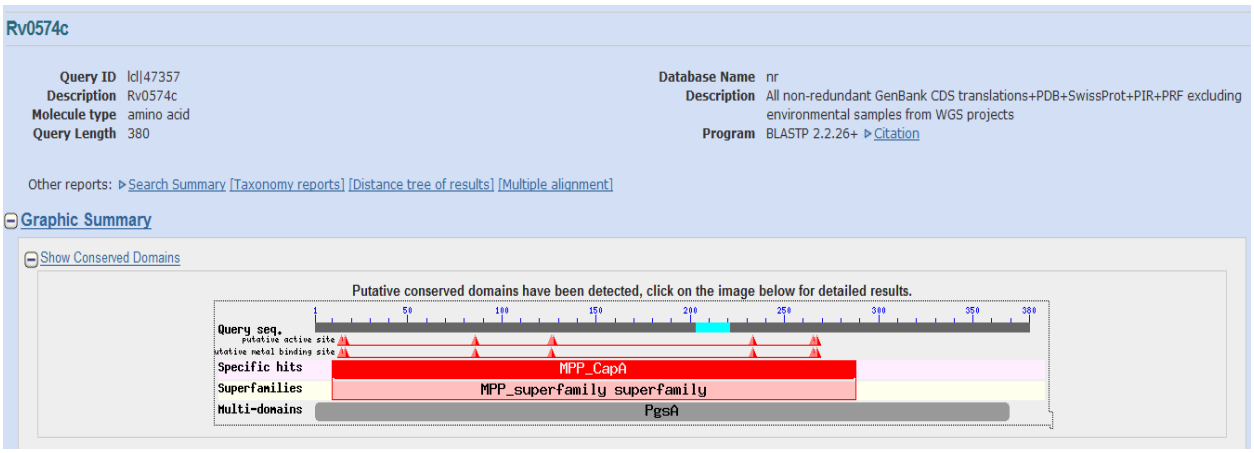


Figure 4.2 Rv0574c has a conserved CapA domain present in proteins involved in poly-Glu biosynthesis. The amino acid sequence for Rv0574c was obtained from Tuberculist and analyzed against non-redundant proteins, using the BLAST tool from NCBI. Domains present in CapA or PgsA proteins involved in poly-Glu biosynthesis were also identified in Rv0574c, suggesting that this protein could be CapA's homologue.

An Rv0574c homologue seems to be present in other *Mycobacterium* spp. As shown in Table 4.3, the genome for some members of the *Mtb* complex encodes for an identical protein with 99% identity (the only difference was in the start codon). An Rv0574c homologue with 91% identity is also present in *M. cannetti*, which also belongs to the *Mtb* complex. In Table 4.3 it can also be seen that some atypical *Mycobacterium* have an Rv0574c homologue with varying degrees of identity.

Table 4.3 Identity between Mtb H37Rv Rv0574c and other *Mycobacterium* spp

***Mycobacterium* spp.**	**Identity**
Mtb* H37Ra** ***Mtb* CDC1551** ***M. bovis ***M. bovis* BCG** ***M. africanum***	**99%**
M. cannetti	**91%**
M. kansasii	**76%**
M. chubuense	**70%**
M. xenopi	**58%**
M. thermoresistibile	**60%**

In *B. anthracis* the CapA protein is encoded in an operon encompassing *capA-capE* [29]. As described above, it was observed that *Mtb rv0574c* and *rv2394*, the homologues of *capA* and *capD* respectively, are not clustered in an operon. However, it was important to explore the genetic context around *rv0574c* to evaluate its conservation amongst other *Mycobacterium* spp encoding an Rv0574c homologue. Also, bioinformatic analyses might help determine if genes encoded in the *rv0574c* locus could participate in a putative pathway involved in α-L-polyGln biosynthesis. The genetic context of the *Mtb* H37Rv *rv0574c* locus is conserved amongst other *Mtb* strains, including clinical isolates, as well as in *M. bovis* and BCG which have Rv0574c homologues with the highest identity (Fig. 4.3). This locus is also present in *M. cannetti* and *africanum*, additional members of the *Mtb* complex. Interestingly, the locus is also conserved in *M. kansassii* and to a lesser degree in *M. chubuense* (not included in Fig. 4.3 because the genomes for these later *Mycobacterium* spp, are not available for comparison in the Comprehensive Microbial Resource webpage).

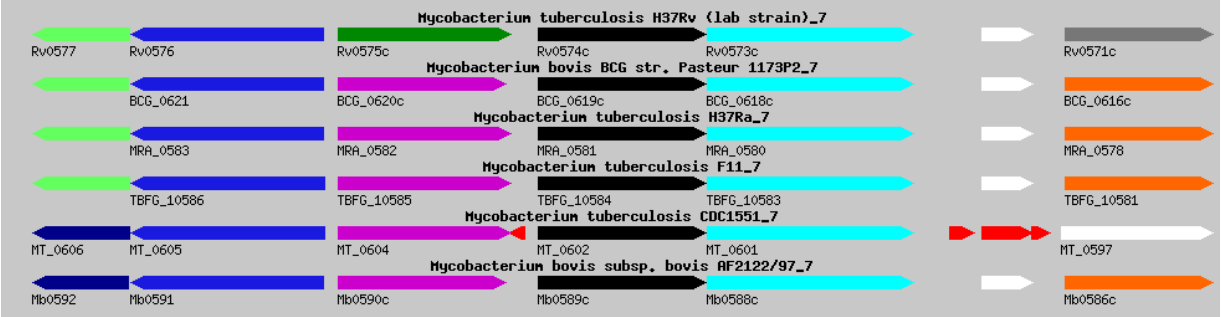


Figure 4.3 Genomic organization of the *rv0574c* locus is conserved in *Mtb* and *M. bovis*. Using the comparative genomic organization in the Comprehensive Microbial Resource page, the locus for *Mtb* H37Rv *rv0574c* was compared to other *Mycobacterium* spp. This locus was present in *Mtb* and *M. bovis* strains, including clinical isolates such as CDC1551. The same genomic organization was also observed in the attenuated *Mtb* H37Ra and *M. bovis* BCG. The region encompassing *rv0575c-rv0571c* was not present in *M. smeg*, *M. avium*, *M. paratuberculosis*, *M. leprae* or *M. abscessus*. However, it was also detected in *M. cannetti*, *M. kansassi* and with some variations, in *M. chubuense* (not shown).

For the other proteins encoded in the *rv0574c* locus, it was not possible to definitely allocate a role in a putative pathway involved in α-L-polyGln biosynthesis. Bioinformatic analysis suggested that Rv0575c has a putative oxido-reductase domain (TIGR03219) (Fig. 4.4).

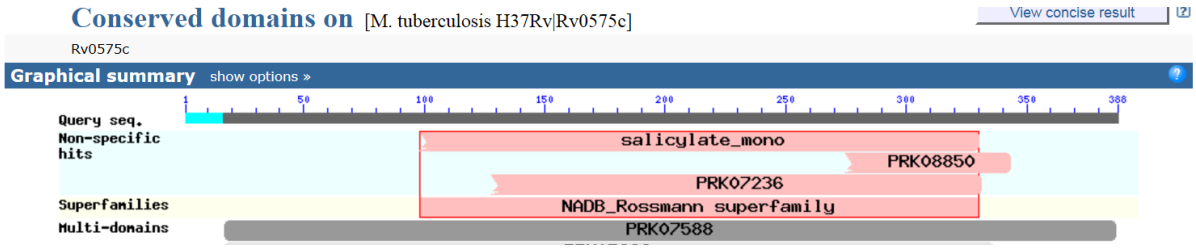


Figure 4.4 Rv0575c has an oxygenase domain. Rv0575c's amino acid sequence was analyzed by BLAST analysis against the non-redundant database in the NCBI webpage. An oxygenase domain present in several oxidoreductases was detected.

It has been previously reported that Rv0573c has a putative nicotinic phosphoribosyltransferase, similar to Rv1330c [47] (see discussion). Rv0572c is a hypothetical protein and during bioinformatics analysis, conserved protein domains were not detected (not shown). Finally, Rv0571c seems to be composed of two completely different domains: a phosphoribosyltransferase (COG1926) and a hydrolase domain (pfam12695) (Fig. 4.5, arrow and arrowhead, respectively).

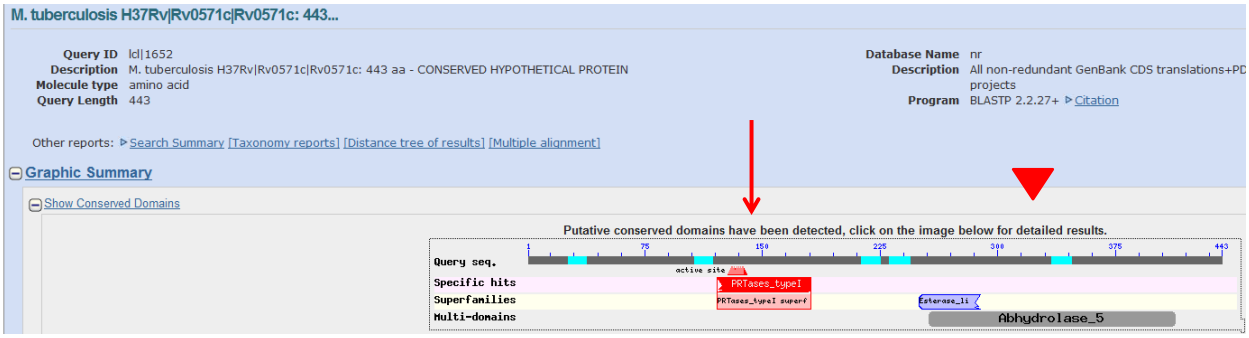


Figure 4.5 Rv0571c has two domains: a hydrolyase and a phosphoribosyl transferase domain. Rv0571c's amino acid sequence was analyzed by BLAST analysis against the non-redundant database in the NCBI webpage. A phosphoribosyl transferase domain is present towards the protein's N-terminus

(denoted with arrow), whereas a hydrolase domain is present in the C-terminus (denoted with an arrowhead).

Thus, if these proteins are in fact involved in α-L-polyGln biosynthesis, the biochemical pathway would be novel and completely different to the one involved in poly-γ-D-Glu biosynthesis. A plausible hypothesis is presented in the discussion (see below).

4.3.2 Mutation of *rv0574c* and *rv2394* in *Mtb* H37Rv.

The assumption that Rv0574c and Rv2394 could be involved in α-L-polyGln biosynthesis was based on homology to the Cap proteins and conserved genomic organization in virulent strains. In order to obtain biochemical data supporting this hypothesis, mutants in *rv0574c* and *rv2394* were engineered in *Mtb* H37Rv. These genes were mutated by an insertion/deletion event using a phage transduction system in which part of the open reading frame is removed and replaced with a hygromycin resistance cassette [40] (Mutants were constructed by Dr. Rainer Kalscheuer at Dr. Jacobs lab). Mutants were evaluated by Southern blot hybridization at the Jacobs lab. In Figure 4.6, it can be observed that gene mutation correlated with a shift in the hybridization pattern (no additional information regarding the Southern blot methodology was provided).

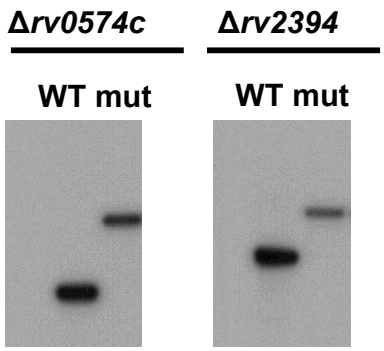


Figure 4.6 Southern blot analysis of Δ*rv0574c* and Δ*rv2394*. Genomic DNA from the mutant strains was subjected to restriction enzyme digestion and analyzed by Southern blot to demonstrate gene

inactivation. Mutants and WT strains were hybridized with probes against the respective genes (performed by Dr. Kalscheuer at Dr. Jacobs lab, Albert Einstein College of Medicine, Bronx, NY).

4.3.3 Biochemical analysis of mutant strains.

Even though by transposon mutagenesis analysis it has been reported that Rv0574c and Rv2394 are non-essential genes [48], it was important to determine if mutants had *in vitro* growth defects in the absence of selection pressures. If so, it would be difficult to ascertain if *in vivo* growth defects have significant physiological relevance. As observed in Figure 4.7, no difference was observed between the mutant and WT strain when grown *in vitro* in 7H9 media supplemented with OADC.

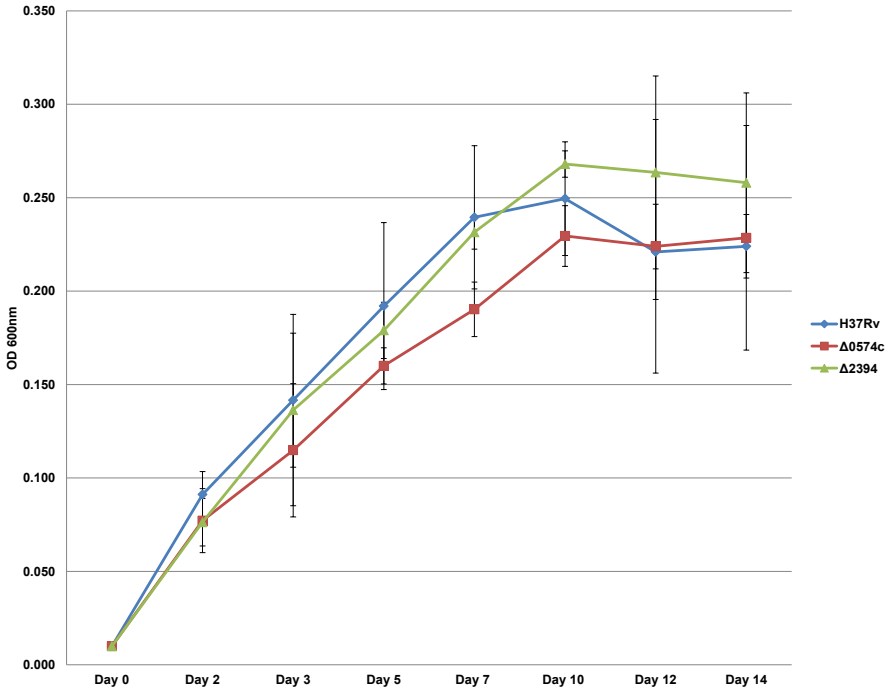


Figure 4.7 Growth curves of H37Rv WT, Δ*rv0574c*, Δ*rv2394*. H37Rv WT and mutant strains Δ*rv0574c and* Δ*rv2394* were grown in 7H9 supplemented with OADC and 0.05% Tween-80. Bacterial growth was determined by following the OD 600nm over the indicated days. No significant difference in growth rates and patterns were observed.

Biochemical analyses were then performed on the cell wall core of mutant strains to determine amino acid composition. To obtain cell wall core, cell walls were extracted thrice with 2% SDS at 60°C as this protocol had been previously shown to remove mycobacterial proteins without affecting α-L-polyGln structure and possibly its concentration [28]. SDS-extracted cell walls were subjected to acid hydrolysis in order to obtain free amino acids amenable to amino acid analysis. As observed in the representative chromatogram obtained during amino acid analysis of cell wall cores, only Ala, Glu, DAP and ammonia were detected (Fig. 4.8, ammonia is also measurable with the AccuTag® amino acid analysis reagent). Thus, only the peptidoglycan-associated amino acids remained after cell wall extraction. Excess Glu and ammonia could be attributed to α-L-polyGln reported to remain associated with peptidoglycan during this extraction procedure. Interestingly, for both mutants significantly smaller Glu and ammonia peaks could be observed in the chromatograms.

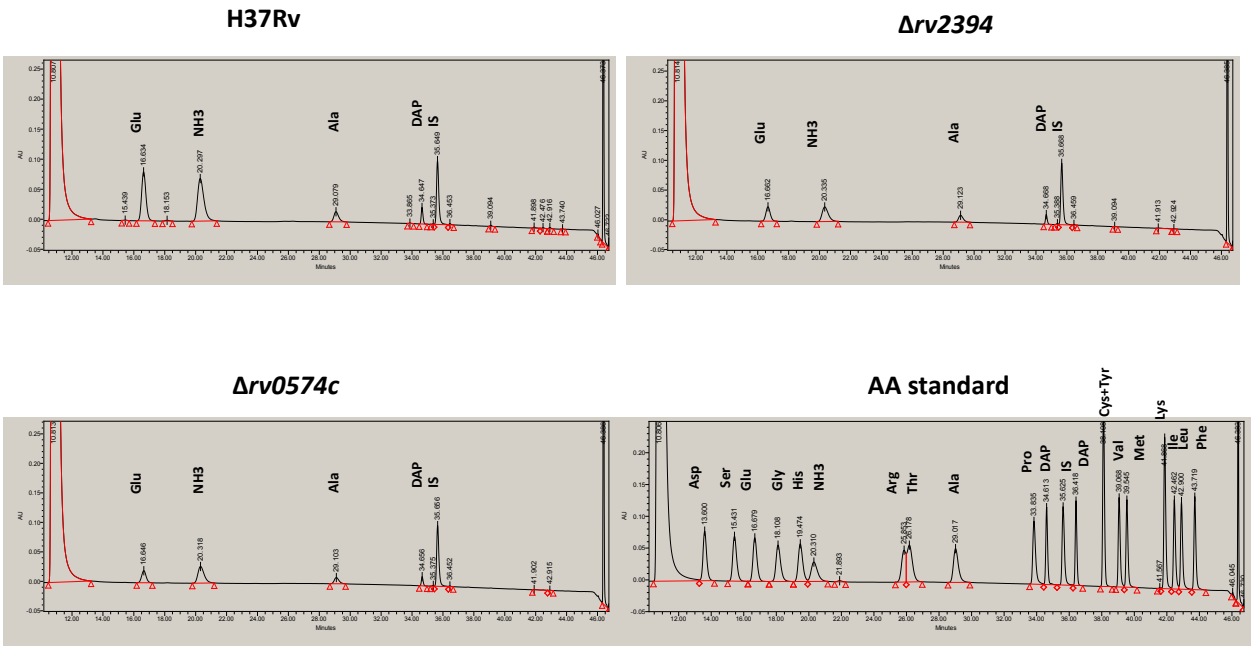


Figure 4.8 Representative chromatograms obtained after amino acid analysis show reduced Glu and ammonia amounts. SDS-extracted cell wall core was subjected to acid hydrolysis with 6N HCl and

hydrolysates were analyzed with the AccuTag® amino acid analysis kit (Waters). An amino acid standard (Pierce) was used to determine retention times and the control area for quantitative analysis. In comparison to WT H37Rv, both Δ*rv0574c* and Δrv2394 had smaller peaks for both Glu and ammonia. The other peaks observed during amino acid analysis of SDS-extracted cell wall core correspond to Ala and DAP. α-aminobutyric acid is used the IS.

Alternatively, data was normalized as the ratio between the amount of each amino acid or ammonia, to the amount of DAP (Fig. 4.9).

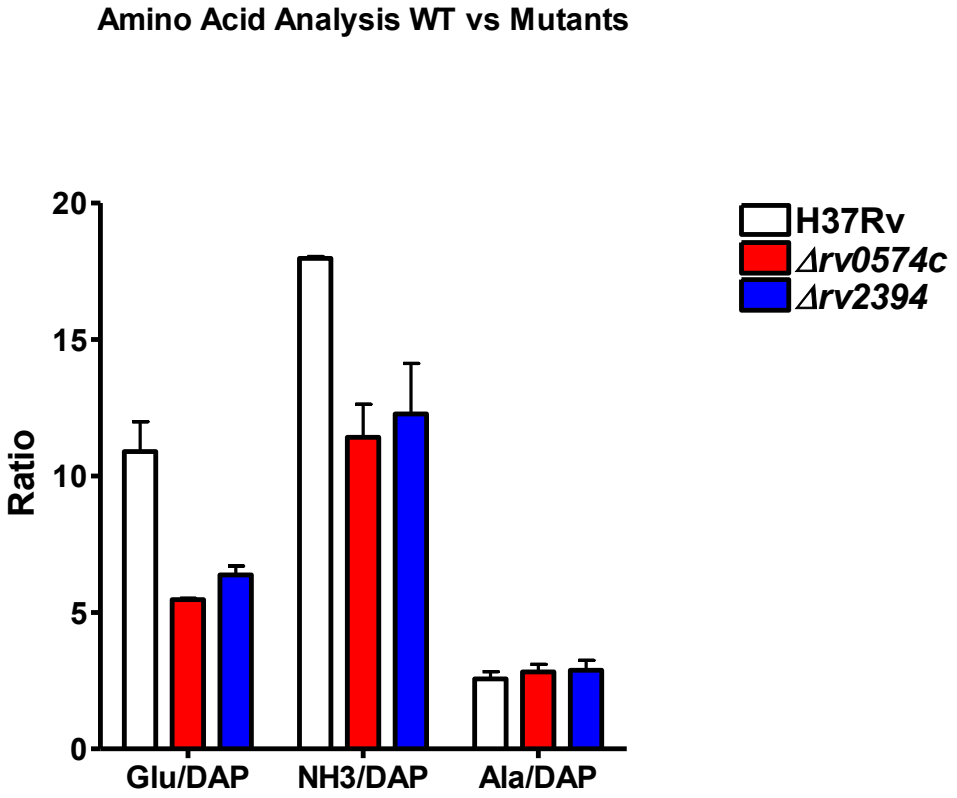


Figure 4.9 Reduced ratios of Glu/DAP and ammonia/DAP were detected for both mutants. Analysis of acid hydrolyzates of SDS-extracted cell wall core were performed after normalizing for the amount of DAP. As can be observed, Ala/DAP ratios were similar in mutants and WT H37Rv with an approximate

ratio of 2.2:1 as expected based on the mycobacterial cell wall structure. Ratios of Glu/DAP and ammonia (NH3)/DAP were reduced in both mutants in comparison to WT H37Rv. Interestingly, even in the mutants, the ratio of Glu/DAP was much higher than expected taking in consideration the presence of one Glu residue per peptidoglycan pentapeptide.

As reported previously [21] and experimentally determined here, the Glu/DAP ratio in WT H37Rv is higher than the expected ratio of 1:1. Even though in both mutants the Glu/DAP ratio was still greater than 1:1, an almost 0.5-fold reduction was observed in comparison to WT cell wall. The reduction of ammonia/DAP ratio was proportional to the reduction of Glu/DAP, suggesting that this ammonia was originating from hydrolyzed Gln. In contrast, the ratio of Ala/DAP remained constant suggesting no peptidoglycan changes occurred for both the *rv0574c* and *rv2394* mutants.

Analyses of the total extractable lipids and the covalently linked mycolic acids were also performed on WT and mutants strains. No difference was observed for the mycolic acids, again suggesting that the mAGP complex had not been affected in these mutants (Fig. 4.10).

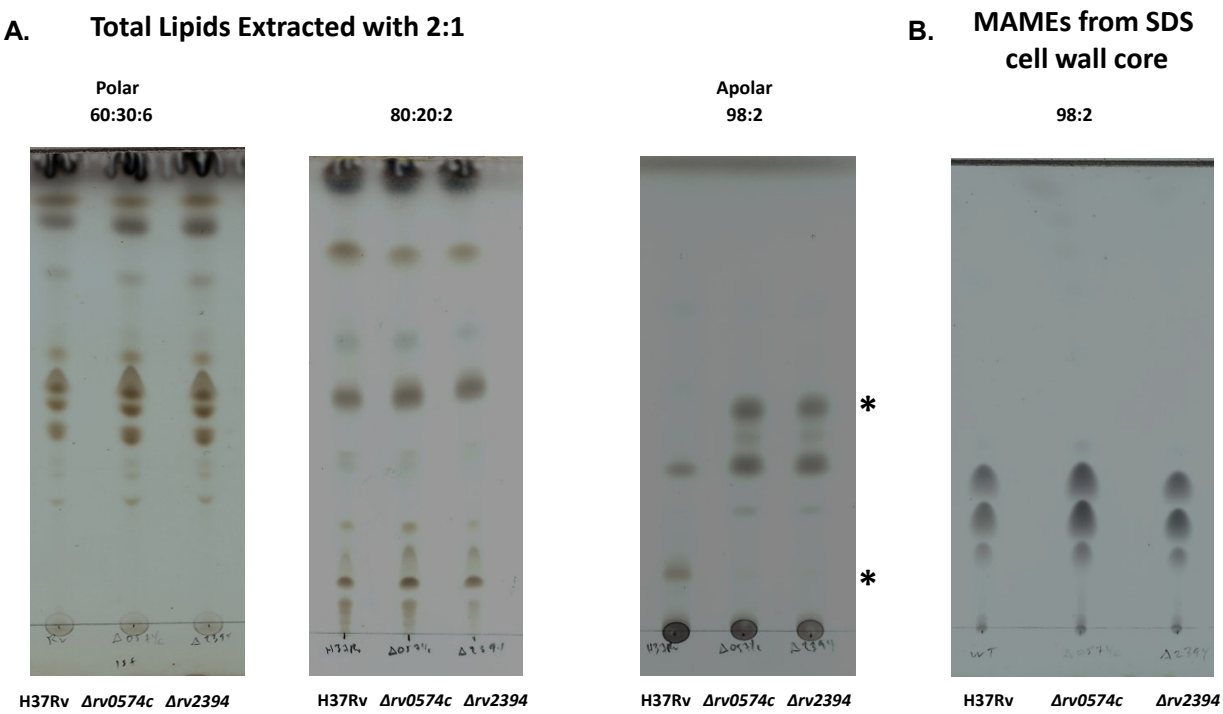


Figure 4.10 Lipid analysis of H37Rv, Δ*rv0574c* and Δ*rv2394* revealed differences in apolar lipids.

A. Total lipids were extracted from whole cells with 2:1 chloroform:methanol and analyzed by TLC on silica plates using different mobile solvent systems, as depicted above. As noted by *, differences between WT and mutant *Mtb*, were only observed when resolving total lipids with an apolar solvent system (98:2, chloroform:methanol). Interestingly, a similar pattern was observed for both the *rv0574c* and *rv2394* mutant.

B. Mycolic acids methyl esters (MAMEs) were obtained from SDS-extracted cell wall core and analyzed after resolving in a silica TLC. No differences were observed for mycolic acids. TLCs were developed with cupric sulfate.

The only noticeable difference was during TLC analysis of total lipids in an apolar solvent system, consisting of 98:2 chloroform:methanol (Fig. 4.10, third TLC, denoted with *). Unfortunately, this was not explored further and the identity of these lipids was not confirmed. However, others have

published that during apolar lipid analysis, the upper band corresponds to PDIMs whereas the lower band corresponds to triacylglycerols [49, 50] (see discussion).

4.3.4 Evaluation of the mutant *Mtb* strains in the murine model of TB.

C57BL/6 mice were infected with a low dose aerosol of WT H37Rv, Δ*rv0574c* and Δ*rv2394.* At the indicated time points, mice were euthanized and both lungs and spleens were harvested and plated in nutrient 7H11 agar. At day 30 and 180, mice infected with either mutant had higher bacterial burdens in the lungs (Fig. 4.11A). Additionally, higher bacterial burdens were also observed at day 30 in the spleens of mice infected with either mutant strain (Fig. 4.11B). At early time points, higher extra pulmonary bacterial count is a good indicator of increased dissemination. However, as disease progressed, tendencies reverted to the one observed for infections with the WT strain. In contrast, histopathological analysis suggested that lung lesions in mice infected with the mutant strains were more severe than in mice infected with the WT strain, and furthermore, this difference persisted throughout the infection (Fig. 4.11C). It should be mentioned that differences were not statistically significant by ANOVA.

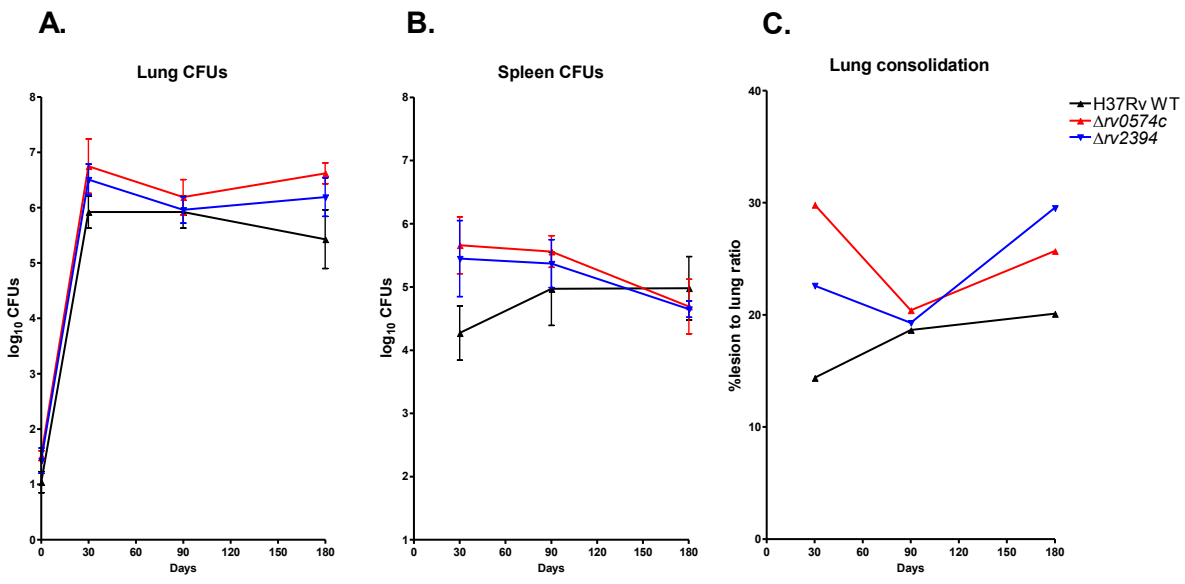


Figure 4.11 Increased lung and spleen bacterial burden, as well as lung lesion severity in mice infected with Δ*rv0574c* and Δ*rv2394*. C57BL/6 mice were infected with a low dose aerosol of WT H37Rv, Δ*rv0574c* and Δ*rv2394*. At the indicated time points, mice were euthanized and both lungs and spleens were harvested and plated in nutrient 7H11 agar. A. Mice infected with either mutant had higher bacterial burdens in the lungs at day 30 and day 180. B. Faster dissemination to the spleen was observed when mice were infected with either mutant. C. The severity of the lung lesions was increased in mice infected with mutant strains. Differences were not statistically significant.

4.3.5 Genetic complementation of mutant strains.

The integrative plasmid pMV306 (Fig. 4.12, left panel) was used to complement the mutant strains by cloning the mutated genes in the plasmid's *XbaI* and *ClaI* site (Fig. 4.13B and C, only the vectors complementing *rv0574c* are shown). All the constructs included 500 bp upstream of the mutated gene's start codon, in order to obtain expression of the complemented protein in a physiological context. As explained in the material and methods section (see above), two constructs were engineered for each mutant. Besides complementing the mutated gene *rv0574c* (Fig. 4.13B), the second construct also

included the downstream gene *rv0573c* (Fig. 4.13C). As *rv0574c* and *rv0573c* are only 9 bp apart, the second construct was used to circumvent the possibility that a polar effect had been introduced when the mutant was engineered.

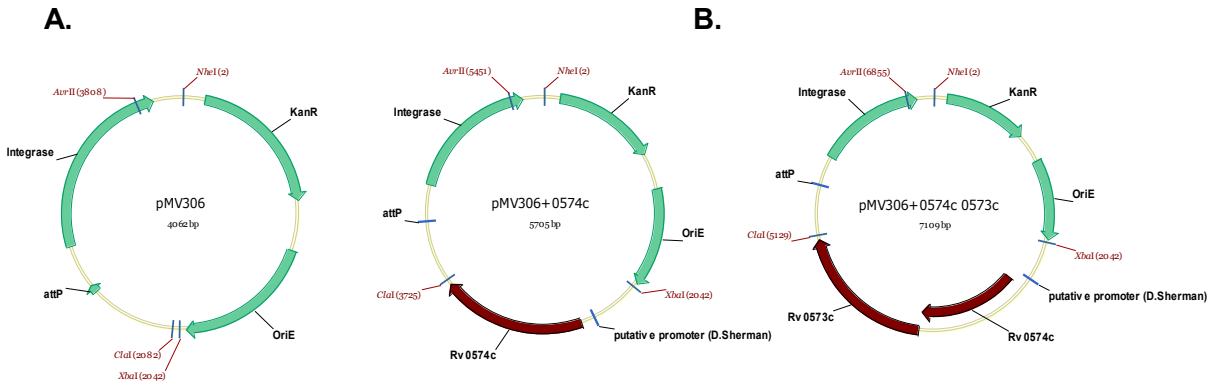


Figure 4.12 Maps of the integrative vector pMV306 and both constructs engineered to complement the *rv0574c* mutation. A. The pMV306 plasmid is an integrative vector that encodes for an integrase mediating a recombination event between the plasmids *attP* site and the *attB* site present in *Mtb*'s genome (Figure 4.13). Furthermore, it encodes for a Kanamycin resistance cassette which can be used as a selection marker. Two vectors containing the endogenous putative promoter (as described by D. Sherman), were engineered to complement the *rv0574c* mutation. B. The first construct only included the mutated gene *rv0574c.* C. The second construct included the downstream gene *rv0573c* in case a polar effect had been introduced when mutating *rv0574c*. Both constructs were inserted in pMV306's *ClaI* and *XbaI* site. The *AvrII* site was used during Southern blot analysis of the complemented bacteria (see Figure 4.14).

pMV306 encodes for an integrase which performs a site-specific recombination event between the plasmid's (Fig. 4.12A) and *Mtb*'s genome (Fig. 4.13) *attP* and *attB* site, respectively. Furthermore, it encodes for a kanamycin resistance cassette that allows selection of transformed bacteria. As can be

observed in Figure 4.14, the *Mtb*'s *attb* site (denoted with an arrow head), is located within the gene encoding for $tRNA^{Gly}$ and in close proximity to *rv2462c*.

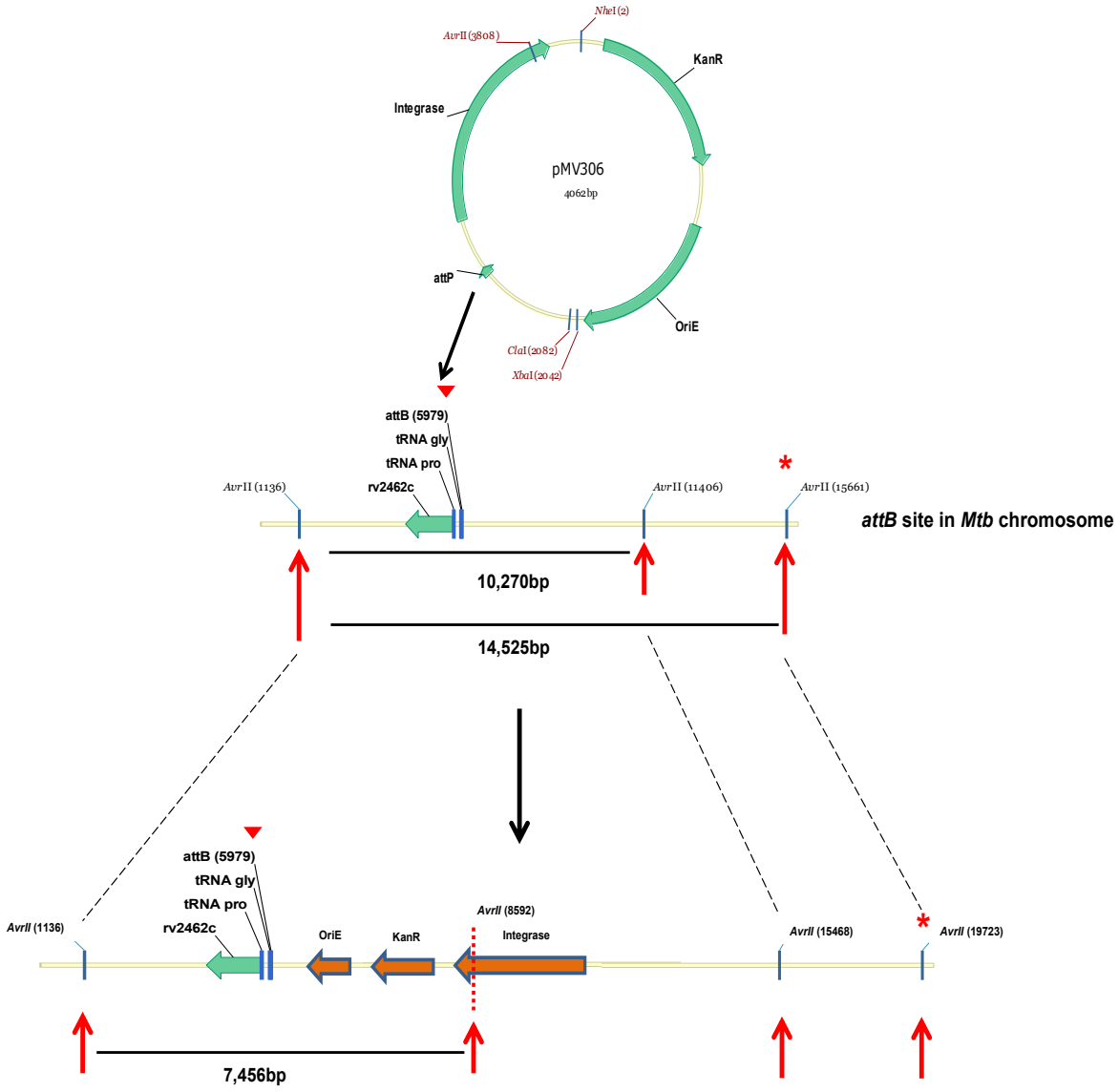


Figure 4.13 Schematic representation of a recombination event between *Mtb*'s chromosome and pMV306. The integrase encoded in pMV306 (top) mediates a recombination event between the *attP* and *attB* site present in the plasmid and *Mtb* chromosome (middle), respectively. In *Mtb* genome, the *attB* site lies in the $tRNA^{gly}$ gene. This genomic region is flanked by two *AvrII* sites (red arrows) that are separated

by 10,270 bp. This fragment contains the gene *rv2462c* which was used to generate the Southern blot probe. A second *AvrII* site is present downstream and is separated from the first one by 14,525 bp (red arrow and *). Integration of pMV306 inserts an *AvrII* site (bottom, dotted line), and *rv2462c* becomes embedded in a 7,456 bp fragment instead. During Southern blot analysis of WT and mutant strains using a probe against *rv2462c*, it was observed that *AvrII* digestion resulted in a fragment of approx. 14 kb (see Fig. 4.15). This probably reflects that *AvrII* digestion occurred in the last site (denoted by red arrow and *). After complementation with pMV306-based vectors (see fig. 4.13), Southern blot analysis confirmed the *AvrII* insertion and the *rv2462c*-reactive bands were proportional to the size of the different engineered constructs.

It was decided that a general way to screen for a recombination event between all of the engineered pMV306-based complementation plasmids and *Mtb*'s genome, was to perform Southern blots targeting *rv2462c*. Digestion of WT genomic DNA with *AvrII* should result in a 10,270 bp fragment hybridizing to a probe targeting *rv2462c* (Fig. 4.13 arrows point to the closest *AvrII* restriction sites flanking *rv2462c*).

In contrast, a band migrating above 10 kb was observed (Fig. 4.14), suggesting that digestion occurred at the next downstream *AvrII* site, resulting in a 14,525 bp fragment instead (Fig. 4.13, denoted by *).

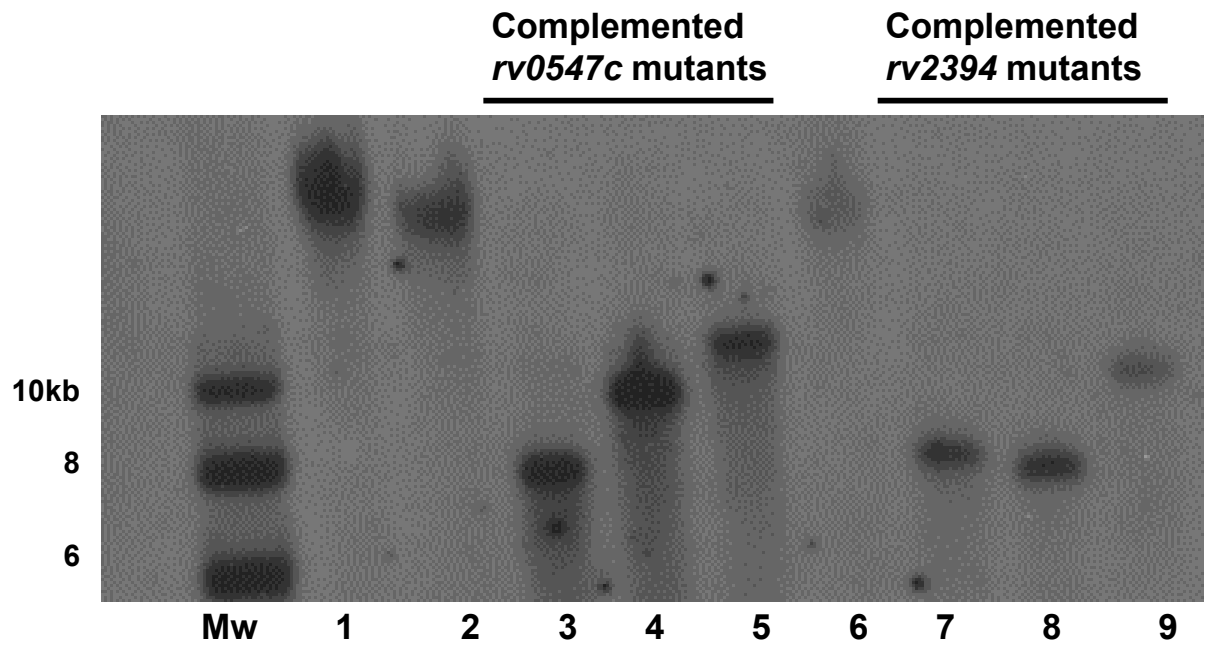


Figure 4.14 Southern blot analysis of complemented Δ*rv0574c* and Δ*rv2394*. Mutant complementation was achieved with pMV306-derived vectors. The *attB* site where pMV306 recombines is in close proximity to *rv2462c* which was used to synthesize the biotinylated probe used in this Southern blot. A higher than expected band was obtained for WT, Δ*rv0574c* and Δ*rv2394*. Bands for complemented strains were consistent with expected size (see Fig. 4.13). Lane 1, WT H37Rv, lane 2 Δ*rv0574*c, lane 3 Δ*rv0574c* complemented with pMV306 (vector control), lane 4 Δ*rv0574c* complemented with pMV306+rv0574c, lane 5 Δ*rv0574c* complemented with pMV306+0574_0573c. As *rv0574c* and *rv0573c* overlap, complementation of Δ*rv0574c* was also performed using a construct containing both genes (lane 5). Lane 6, Δ*r2394*, lane 7 Δ*r2394* complemented with pMV306, note the same hybridization pattern as lane 3 corresponding to Δ*rv0574c* complemented with pMV306. Lane 8, Δ*r2394* complemented with pMV306+2394. Note that complementation with *rv2394* results in a smaller fragment than with vector alone, attributed to an additional *AvrII* site present in the gene's promoter coding sequence. Lane 9, Δ*r2394* complemented with pMV306+2394_2395.

Digestion with *AvrII* consistently led to a bigger fragment than expected and currently the reason for this is unknown. One possible is point mutations which have been described to occur even within a same lab [51]. Despite this, several lines of evidence suggested that it was prudent to continue. First, the plasmid encoding *rv2462c* was sequenced. Second, the biotinylated probe was obtained from a gel-purified *rv2462c* dropout obtained after restriction enzyme digestion. Third, during Southern blot analysis of the complemented strains, the *rv2462c* probe hybridized to a unique and very sharp band (Fig. 4.14). This is consistent with the fact that the *Mtb* genome does not encode for other homologues of *rv2462c*. Finally and most important, analysis of the bands resulting after complementation corresponded to the expected size. In Figure 4.14, it can be observed that recombination between *Mtb*'s DNA and pMV306 would insert an *AvrII* site, resulting in a 7456 bp band during Southern blot analysis. The size of this band is in accordance to the one detected experimentally (Fig. 4.14, lane 3). Complementation with pMV306+0574c, or pMV306+0574c_0573c should generate bands corresponding to 9,099 and 10,503 bp, respectively similar to the bands detected in the Southern blot (Fig. 4.15, lanes 4 and 5). For the Δ*2394* mutant, complementation with pMV306+2394 should result in a smaller band than complementation with empty vector, as *rv2394's* promoter encodes for an additional *AvrII* site (Fig. 4.14 lane 7 vs lane 8). When the Δ*2394* mutant is complemented with pMV306+2394_2395, a band of approximately 10 kb should be observed taking in consideration that *rv2395* is almost 2000 bp (Fig. 4.14, lane 9).

Finally, complementation of mutant strains was evaluated by qRT-PCR to determine that complemented genes were being expressed to similar levels as WT control. Several sets of primers were used to analyze each of the mutated genes (see Table 4.2 for qRT-PCR primers). Furthermore, qRT-PCR was also performed on the gene downstream of the mutation in order to evaluate polar effects. As *rv2394* (see Chapter III) and *rv0773c* both seem to encode a GGT, expression of *rv0773c* was evaluated in the *rv2394* mutant to determine if it could be compensating the mutation. Similarly, since *rv0573c* and

rv1330c have been suggested to encode for redundant phosphoribosyl transferases [47], expression of the later gene was evaluated to determine compensatory mechanisms in case mutating *rv0574c* had a polar effect on *rv0573c*'s expression. As observed in Figure 4.15A, mutating *rv0574c* effectively abolished its expression as determined by the three sets of primers being evaluated. Complementation with both constructs restored the expression of *rv0574c* and in addition, this mutation didn't seem to have a significant polar effect on the expression of *rv0573c*. No difference in *rv1330c* expression was observed for the *rv0574c* mutant.

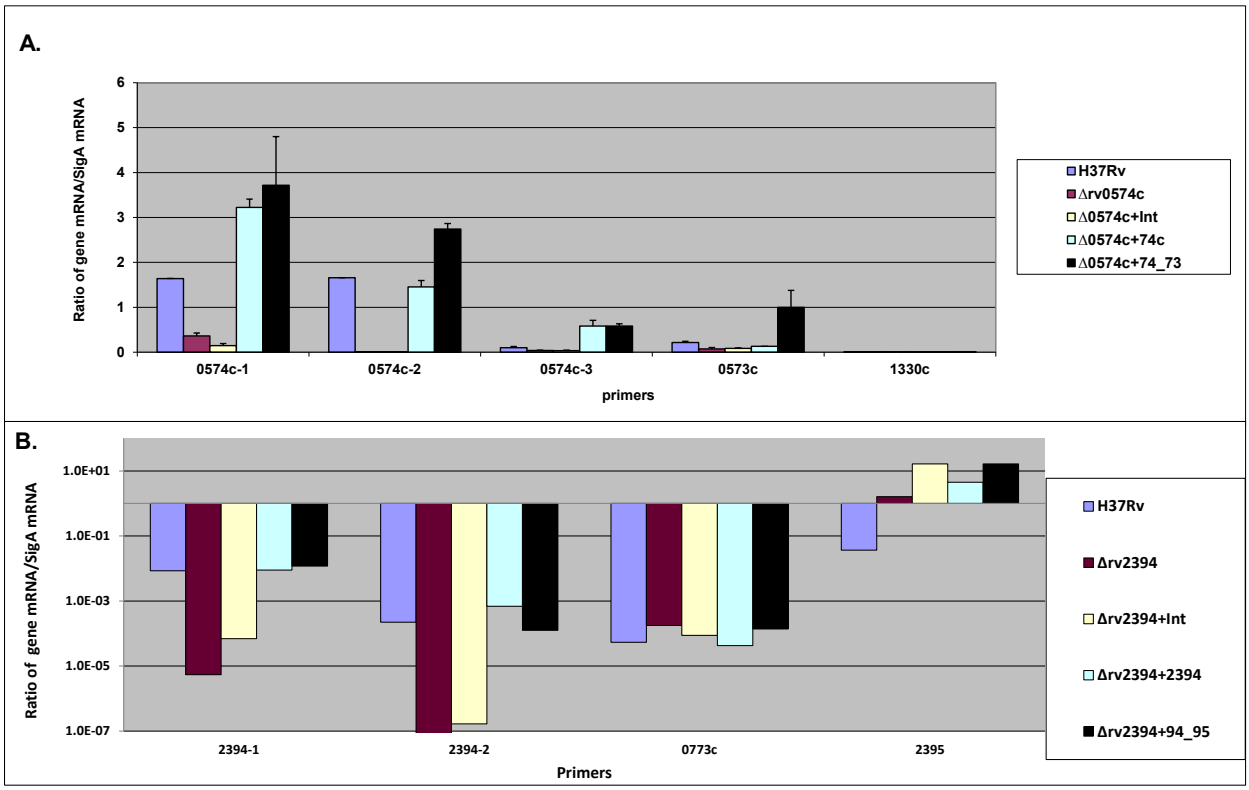


Figure 4.15 qRT-PCR analysis confirming the lack of Δ*rv0574c* and Δ*rv2394* expression in mutants and efficient complementation. mRNA was obtained for *rv0574c* and *rv2394* mutants and complemented strains. qRT-PCR analysis was performed to evaluate the mRNA expression levels for the different genes. Three and two sets of qRT primers were used to evaluate *rv0574c* and *rv2394* mutants

and complemented strains, respectively. Data is expressed as the ratio of mRNA for each respective gene/mRNA for SigA A. Primers 0574c-1 and 0574c-2 did not amplify any mRNA for Δ*rv0574c* and Δ*rv0574c* complemented with an empty vector (+Int). In contrast, complemented strains including gene *rv0574c* or *rv0574c_rv0573c* restored mRNA expression levels. Levels of mRNA for *rv0573c* were enhanced when co-expressing both *rv0574c_rv0573c* in the integrative plasmid. No differences were observed for *rv1330c*, an *rv0573c* homologue suggesting that this gene was not being up-regulated to compensate for the lack of *rv0573c*. B. Expression levels for *rv2394* were less than for *rv0574c* even in the WT H37Rv. Nevertheless it can be observed that for the Δ*rv2394* mutant or the mutant complemented with an empty vector (Int), both primers amplified between 1000 to 10,000-fold less mRNA. Expression levels for *rv2394* were restored to WT levels when complementing with either *rv2394* or *rv2394_rv2395*. Interestingly, mutants and complemented strains had increased mRNA expression levels for *rv2395*, suggesting that mutation in *rv2394* had a polar but positive effect on the expression of *rv2395*. Expression of *rv0773c*, an additional ggt was not affected by *rv2394*'s mutation.

Similar to what was observed for the *rv0574c* mutant, mutating *rv2394* abrogated its expression (Fig. 4.15B). Furthermore, no compensation was observed for *rv0773c*, the gene encoding a putative GGT homologue. Complementing the mutated gene restored *rv2394*'s expression to WT levels. Interestingly, mutating *rv2394* had a polar but positive effect on the expression of *rv2395*, the downstream gene. If this later gene has a role in α-L-polyGln biosynthesis, then the uncontrolled expression of *rv2395* could affect the polymer's concentration. This remains to be investigated.

4.3.6 Amino acid analysis of complemented strains.

After confirming by both Southern blot and qRT-PCR that complementation of mutant strains had been achieved, amino acid analysis of SDS-extracted cell wall cores was performed (Fig. 4.16). Unfortunately the trends obtained during the initial analysis of the mutant strains, were not reproduced.

Instead of obtaining a similar ratio of Glu/DAP of approx. 12 (see Fig. 4.9), the new ratio was approx. 6, very similar to the ratios previously obtained for the mutants. These new results precluded further experimentation and require re-evaluating the phenotype of the mutant and WT strains. It is known that the *Mtb* genome accrues considerable number of mutations during *in vitro* culturing, which tend to abrogate its virulence [51, 52]. It has been suggested that in order to maintain *Mtb* virulence, bacterial strains should frequently be subjected to the selection pressure of *in vivo* conditions [51]. Future experiments should address if the strains used in this study, lost their phenotype during *in vitro* growth.

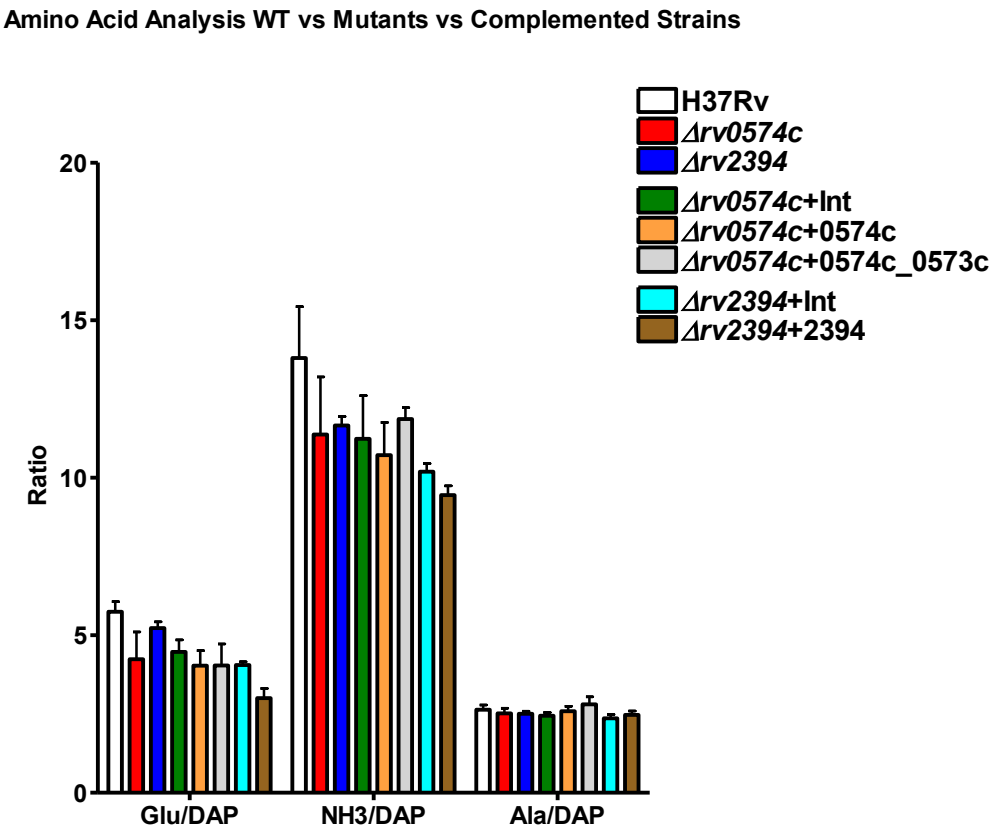


Figure 4.16 Amino acid analysis of SDS-extracted cell wall core in WT, mutant and complemented strains. Amino acid analysis was performed on SDS-extracted cell wall cores from WT, mutant and

complemented strains. Unfortunately, overall trends were different than observed in Fig. 4.9. The WT, mutant and complemented strains had similar ratios of Glu, NH3 and Ala to DAP.

4.4 Discussion.

The presence of pathogen-specific products is a unique opportunity to understand virulence mechanisms. In the TB field, comparison between avirulent BCG and virulent *M. bovis* [53], led to the discovery that the RD1 encodes for a specialized protein secretion system. Other examples of virulence determinants present in *Mtb* and *M. bovis* as opposed to environmental *Mycobacterium* include the identification of PDIM and SL in the former strains. In contrast, no specific role in pathogenesis has been identified in some *Mtb*-specific products such as SL [54]. However, this could be attributed to problems in experimental designs or models. As genome reduction is a common theme in the evolution of pathogens [55], the fact that a metabolic pathway is preserved in a virulent microorganism should mandate its evaluation. Similar to *B. anthracis*, *Mtb* synthesizes an amino acid polymer that is present in the bacterium's cell wall. Since *B. anthracis* requires poly-γ-D-Glu for its virulence [22], it was hypothesized that α-L-polyGln, an *Mtb* specific product, could also have a role in TB. In this Chapter, the identification of genes involved in the biosynthesis of the *Mtb*-specific α-L-polyGln was sought using the *B. anthracis* model of poly-γ-D-Glu biosynthesis. Rv0574c and Rv2394 were determined to be the most likely candidates involved in α-L-polyGln biosynthesis, as they share high homology to *B. anthracis* CapA and D, respectively. Mutants for both genes were generated and evaluated biochemically, as well as in animal models of TB.

As mentioned, the initial strategy to identify *Mtb* genes that could be involved in α-L-polyGln biosynthesis was based on the polymer's similar structure and biochemical composition to *B. anthracis* poly-γ-D-Glu. Rv0574c and Rv2394 were identified as *Mtb* proteins with high homology to *B. anthracis* CapA and D respectively, which are involved in poly-γ-D-Glu biosynthesis. In the *Bacillus* model, CapA participates in transporting poly-γ-D-Glu to the extracellular space [30, 32] and CapD in crosslinking it to

the peptidoglycan [33-35]. However, no *Mtb* homologues were identified for CapB, the actual protein involved in polymerizing D-Glu residues [22, 30]. Furthermore, initial bioinformatics analysis suggested that the proteins encoded in either the *rv0574c* or *rv2394* loci could not perform an amino acid polymerization reaction. Other alternatives were therefore sought based on the following premises. The covalent polymerization of Gln residues to make α-L-polyGln requires the formation of an energy-dependent peptide bond. Three major mechanisms have evolved in nature to catalyze covalent linkages between amino acids. For transpeptidases (see Chapter III), energy released from a cleaved bond is harnessed to make the new bond. More commonly, energy is consumed in the form of ATP in order to activate the donor or acceptor molecule via amino acyl adenylates or acyl phosphate intermediates. As can be observed in Figure 4.17, in acyl phosphates the substrate is directly phosphorylated and ADP is released. In contrast, in acyl adenylates the substrate is directly linked to AMP through a phosphodiester bond and pyrophosphate is released.

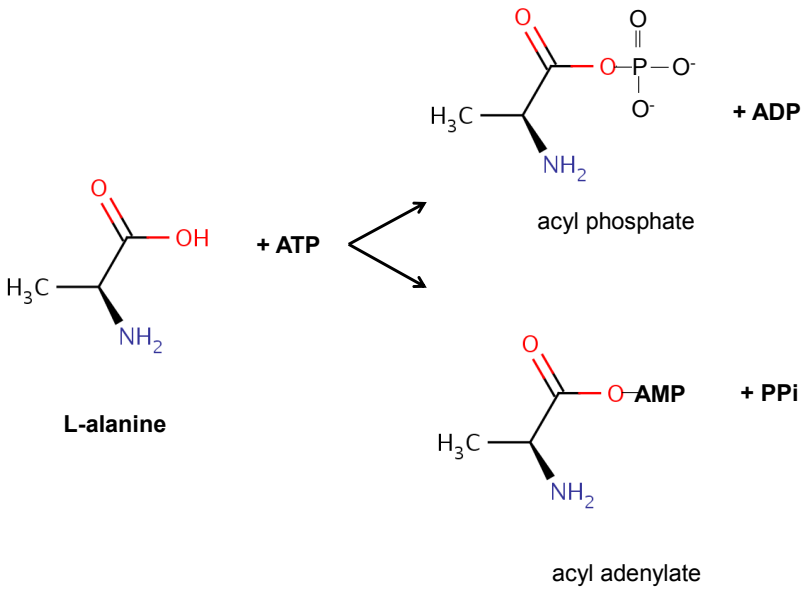


Figure 4.17 Activation of substrates as acyl phosphates or acyl adenylates. Activation of L-alanine is depicted to occur either through the generation of an acyl phosphate (upper reaction) or acyl adenylate

(lower reaction). In both cases, ATP is consumed. During the generation of acyl phosphate, a phosphate residue is directly conjugated to the substrate and ADP is released. In contrast, during acyl adenylate generation, the substrate is conjugated to AMP and pyrophosphate (PPi) is released. In both cases, high energy bonds are generated and will provide the energy for subsequent reactions.

Formation of amino acyl adenylates is the more frequent way to activate amino acids and this occurs in both ribosomal and non-ribosomal peptide synthesis [56, 57]. In conventional ribosomal protein synthesis, tRNA synthetase generate amino acyl adenylates which are rapidly transferred to the 3'OH group of the tRNA [58]. Homology searches for proteins with a tRNA synthetase-like activity that could be involved in α-L-polyGln biosynthesis, did not yield additional proteins besides tRNA synthetases (not shown). Non-ribosomal peptide synthetase was also considered as another alternative for the biosynthesis of α-L-polyGln. *Mtb* only has two non-ribosomal peptide synthetase: the Mbt proteins involved in mycobactin synthesis [59] (see Chapter I) and Rv0101. Rv0101 is composed of two adenylation domains that have homology to the ones present in Mbt. However, Rv0101 was considered a second tier candidate for α-L-polyGln biosynthesis, as *M. smeg* encodes for a similar protein (*Msmeg*_4511) despite not producing α-L-polyGln. Recently, several lines of evidence have suggested that Rv0101 encodes for a non-ribosomal peptide synthetase involved in the biosynthesis of a hereto uncharacterized lipopeptide [60, 61]. In regards to acyl phosphates involved in the formation of peptide bonds, the classical example is the Mur proteins participating in peptidoglycan biosynthesis [62]. Interestingly, in this case the accepting peptidoglycan stem is the activated molecule instead of the incoming amino acid (donor). The folC protein (Rv2447c) that catalyzes the biosynthesis of folyl-polyglutamates was the only protein identified during bioinformatics analysis searching for Mur-like proteins that could be involved in α-L-polyGln biosynthesis (not shown). This strong resemblance between the Mur proteins and folC had been reported several years ago and a special family of amide ligases was proposed [62]. Interestingly, CapB also belongs to this family of amide ligases including Mur and folC [62]. In line with the enzymatic

activity of the latter proteins, experimental evidence suggests that during poly-γ-D-Glu biosynthesis, CapB activates via an acyl phosphate the elongating polymer as opposed to the incoming D-Glu [30]. However, folC is also conserved in many microorganisms [63] and was not considered to participate in α-L-polyGln biosynthesis. Altogether, after this extensive but unsuccessful analysis searching for other genes involved in *Mtb's* α-L-polyGln biosynthesis, it was decided that the *rv0574c* and *rv2394* loci would be the initial focus of this study and future work would be dictated by the biochemical and phenotypic analysis of the mutant strains. Indeed, the major finding in this study consisting of reduced amount of Glu and ammonia in the cell wall of both *rv0574c* and *rv2394* mutants tends to support the hypothesis that these loci are involved in α-L-polyGln biosynthesis. Noticeably, both mutants behaved similarly during biochemical analysis and animal studies, further substantiating that Rv0574c and Rv2394 could be involved in the same pathway despite not being encoded in the same locus.

Capsular poly-γ-D-Glu is completely essential for *B. anthracis* virulence [22]. The proteins involved in this biosynthetic pathway are encoded in the pX02 plasmid [23, 24], one of the two *B. anthracis* plasmids. The other plasmid, pX01, is also essential for its virulence as it encodes for the potent protective toxin [64]. The role of poly-γ-D-Glu in *B. anthracis* virulence is intimately related to its subcellular localization and structure. This product is present in the bacillus capsule which is exposed to the immune system. In contrast to non-encapsulated *B. anthracis,* the poly-γ-D-Glu capsule inhibits phagocytosis [29]. Thus, the bacterium remains extracellular where it can readily multiply. Furthermore, as the capsule is composed of D-Glu residues forming γ-linkages [65], host proteases are less able to digest this product. In turn, lack of antigen processing and presenting makes poly-γ-D-Glu poorly immunogenic [65]. In terms of vaccine development, poor immunogenicity for bacterial surface products had been encountered with the polysaccharides present in *Neisseria meningitides*, *Haemophilus influenza* and *Streptococcus pneumonia* [66]. In order to circumvent this lack of immunogenicity, successful vaccines were engineered by covalently linking the polysaccharides to a carrier protein [66]. Interestingly, a similar strategy against *B. anthracis* poly-γ-D-Glu has also been successful. Initially, immune serum

developed against poly-γ-D-Glu conjugated to a carrier protein was shown to enhance *in vitro* opsonophagocytosis [67]. More importantly, recent *in vivo* studies using the same vaccine approach partially prevented the development of anthrax in guinea pigs [68]. *S. epidermidis* has also been shown to synthesize capsular poly-Glu, but in contrast to *B.anthracis*, the polymer consists of both D and L amino acids [69]. However, *S. epidermidis* polymer also protects against neutrophil phagocytosis and anti-microbial peptides. *S. epidermidis* mainly causes nosocomial infections, specifically those involving catheters. Despite successful attachment and biofilm formation *ex vivo* in catheters, *S. epidermidis* lacking poly-Glu did not thrive *in vivo* after the contaminated catheter was implanted in an animal [69]. Altogether, results with these bacteria have highlighted a role for poly-Glu in predominantly inhibiting the innate immune system.

In contrast to *B. anthracis* poly-γ-D-Glu, no role in pathogenesis has been attributed to *Mtb*'s α-L-polyGln. Mutants for the *rv0574c* and *rv2394* genes putatively involved in α-L-polyGln biosynthesis were engineered in order to test in animal models the hypothesis that the polymer's presence in virulent *Mycobacterium* is associated to pathogenicity. Interestingly, at early time points increased bacterial burden was observed in the lungs and spleens of mice infected with these mutant strains. As disease progressed, the difference waned and mutant and WT strains behaved similarly in terms of bacterial burden. In contrast, histological analysis revealed that lungs from mice infected with mutant strains had a tendency to be more compromised and this persisted throughout the acute and chronic stages of infection. This could rule out a CFU-dependent effect on pathological severity. Instead, phenotypic differences between mutant and WT strains could more likely account for this observation. If Rv0574c and Rv2394 participate in α-L-polyGln biosynthesis, than the more pronounced effect early after infection with these mutants, might suggest that this polymer could also affect the innate immune system as described for *B. anthracis* poly-γ-D-Glu. However, despite its localization in *Mtb*'s cell wall, α-L-polyGln is not currently considered to be surface exposed [70]. If this is true, it probably does not have a direct role during the early stages of the innate immune system including opsonization and phagocytosis.

Alternative roles for *Mtb's* α-L-polyGln in TB pathogenesis could be sought in poly-Gln expanding diseases. This group of diseases is characterized by mutations that increase the numbers of Gln repeats normally present in human proteins [71]. Pathogenesis in poly-Gln expanding diseases is attributed to a loss of function of the mutated protein and more importantly, to the insolubility and aggregating tendency of longer Gln stretches [72], a biophysical effect that could also explain α-L-polyGln insolubility. A major finding in expanding poly-Gln diseases is proteasome malfunction attributed to its inability to degrade the aggregated proteins [73, 74]. A similar situation might occur with WT *Mtb* and its α-L-polyGln could constitute an evasion mechanism against antigen presentation. Loss of α-L-polyGln in both mutants could lead to more efficient proteasome-mediated antigen processing and presentation. In turn, enhanced inflammation could ensue due to increased Tcell priming and activation. Furthermore, diminished NF-κB function has been observed in expanding poly-Gln diseases as a consequence of proteasome malfunction [75, 76]. In the absence of α-L-polyGln, enhanced NF-κB-mediated signaling could increase the inflammatory response against mutant *Mtb*. Cell biology studies with purified α-L-polyGln could help elucidate a direct role for this polymer. After all, it seems that during the course of *Mtb* infection the immune system encounters this product, as anti-α-L-polyGln antibodies were detected in the rabbit model of TB [70].

Abrogating *Mtb* α-L-polyGln biosynthesis could have an indirect as opposed to a direct role in TB pathogenesis. For instance, cell wall modifications could probably induce a compensatory effect [4]. In fact, mutations abrogating the synthesis of diacyl and polyacyltrehalose led to alterations in *Mtb's* capsule and this affected the bacillus-host cell interaction [77]. Specifically for α-L-polyGln deficient mutants *rv0574c* and *rv2394*, it seemed that increased concentration of PDIM might be present in the pool of extractable cell wall lipids. Thus, *in vivo* results with α-L-polyGln mutants could instead be attributed to the pro-inflammatory effect mediated by PDIM, a major *Mtb* virulence factor [78, 79]. Compensatory effects involving PDIM have also been reported for mutants in SL biosynthesis [80]. In this case, excess methylmalonyl-CoA is present as mutant *Mtb* does not use it for SL biosynthesis. Instead, this extra pool

of methylmalonyl-CoA is shunted into PDIM biosynthesis resulting in SL mutants containing higher abundance of longer PDIM molecules [80]. In turn, increased PDIM concentration could explain why no *in vivo* growth defect for SL mutants was observed in the guinea pig model [54]. A similar metabolic effect leading to enhanced PDIM biosynthesis could also be envisioned for both α-L-polyGln mutants *rv0574c* and *rv2394*.

It is well known that Gln is an excellent anaplerotic molecule that enters the TCA cycle as ketoglutarate after consecutively losing its amido and amino group [81]. If the *rv0574c* and *rv2394* mutants do not incorporate Gln into α-L-polyGln, then excess Gln in the form of ketoglutarate could be entering the TCA cycle. In the TCA cycle, excess ketoglutarate would be converted into succinyl-CoA, a precursor of methylmalonyl-CoA [80]. Thus, mutations in both SL and α-L-polyGln could have similar metabolic consequences consisting of excess methylmalonyl-CoA, which could be shunted into PDIM biosynthesis. As a corollary to this hypothesis, PDIM length could also be increased in α-L-polyGln mutants. From a metabolic perspective, an additional interesting observation with both mutants was the absence of TAGs during apolar lipid analysis. Accumulation of TAGs is considered an important marker of dormant *Mtb* [82, 83]. In fact, *rv1330c*, the gene encoding the most active triacylglycerol synthase, is part of the dormancy regulon [84, 85]. Interestingly, the *rv0574c* locus described herein to participate in α-L-polyGln, is also part of the dormancy regulon [86]. One consideration is that α-L-polyGln is a Gln depot system (discussed in Chapter V). Thus, the dormancy regulon could be shifting *Mtb* into "hibernation" mode in which carbon and nitrogen sources are stored as TAG and α-L-polyGln, respectively. In contrast, reduced carbon or nitrogen depots seem to be present in the *rv0574c* and *rv2394* mutants and these strains would be less likely to enter a dormant phase. A constantly higher metabolic activity in mutant strains could also lead to the increased inflammation observed in infected mice, as this could be associated with enhanced protein secretion known to contain multiple enzymatic activities [87].

Extrapolating results from animal models to human TB can be challenging taking in consideration the different pathogenesis mechanisms and in addition, the contrived infectious

methodology in animal models. However, results reported herein are especially compelling in the context of some recent epidemiological observations done in cattle and humans. A strain of *M. bovis* was detected to have quickly spread in cattle in West Africa [88]. Molecular epidemiology techniques confirmed that this strain of *M. bovis* had a deletion affecting *Mb0586c* to *Mb0590c*, equivalent to *Mtb's rv0575c* to *rv0571c*. Furthermore, mutations encompassing *rv0571c-rv0572c* have been identified in human clinical isolates [89]. In both studies it was suggested that mutant strains had a selective advantage over WT and this advantage had only recently become important as these genes are conserved in most other clinical isolates. Even though these studies were exclusively epidemiological and were not accompanied with histopathological analysis, results obtained in the murine model with the *rv0574c* and *rv2394* mutant could give some insight into the epidemiological findings. As occurred in the murine model, mutations did not seem to affect microbial fitness in either humans or cattle. Selective advantage for the mutants could instead be attributed to increased transmission, explaining its fast dissemination in West Africa [88]. For both α-L-polyGln deficient mutants, faster replication kinetics was observed in the initial stages of infection and this correlated with enhanced inflammation. It is well known that *Mtb* transmission requires substantial inflammation and tissue damage. Thus, one possibility is that these clinical strains are being transmitted quicker because of their initial faster growth and lung damage leading to cavitation [90]. Development of transmission models could help evaluate this hypothesis.

If the role of Rv0574c and Rv2394 is to transport α-L-polyGln and crosslink it in or to the cell wall respectively, which *Mtb* gene(s) encode for proteins involved in amino acid activation and polymerization? After thoroughly analyzing both loci, the following α-L-polyGln biosynthetic pathway is proposed on theoretical grounds but experimental evidence is required (*rv2394* locus is discussed in Chapter III). The proposed pathway starts with Rv0573c, encoded only 9 bp downstream of the *rv0574c* gene mutated in this study. This genetic proximity is a common finding when consecutive genes are in an operon, constitute a transcriptional unit and participate in the same biochemical pathway. However, Rv0573c is annotated as a putative nicotinic phosphoribosyl transferase, an activity redundant to the one

described for Rv1330c [47]. Interestingly, recombinant Rv1330c was indeed shown to have nicotinic phosphoribosyl transferase activity but this was not the case for recombinant Rv0573c. Instead, Rv0573c's nicotinic phosphoribosyl transferase activity was inferred from mutagenesis experiments. Whereas single mutations of either *rv1330c* or *rv0573c* did not affect the nicotinic acid pool, this was accomplished by double mutations. However and consistent with the lack of nicotinic phosphoribosyl transferase activity in recombinant Rv0573c, it is possible that in the context of an Rv1330c mutation, a 'moonlighting' nicotinic phosphoribosyl transferase could be present in Rv0573c. Hence, a 'pseudo-redundancy' might be concluded. It is clear that Rv1330c's substrates are phosphoribosyl pyrophosphate (PRPP) and nicotinic acid and the enzyme catalyzes a nicotinic phosphoribosyl transferase activity (Fig. 4.18, left reaction). Recombinant Rv0573c did not utilize nicotinic acid, but taking in consideration its high homology to Rv1330c, its respective cognate substrate should be similar to nicotinic acid. Picolinic acid (Fig. 4.18 center, top row), is a catabolyte from the tryptophan pathway [91] and is an isomer of nicotinic acid only differing in the position of the carboxylate in reference to the pyridine ring. Remarkably, picolinic acid's structure is very similar to pyrroline carboxylic acid (Fig. 4.18, right, top row), a cyclized form derived from Glu-γ-semialdehyde [92], an intermediate in Glu, Pro and Ornithine metabolism. Taking in consideration the similar structures of these compounds, a possibility is that Rv0573c catalyzes a reaction in which PRPP and pyrroline carboxylic acid are linked together by a glycosidic bond and form a novel compound, pyrroline carboxylate-D-ribonucleotide (Fig. 4.18, reaction on the right).

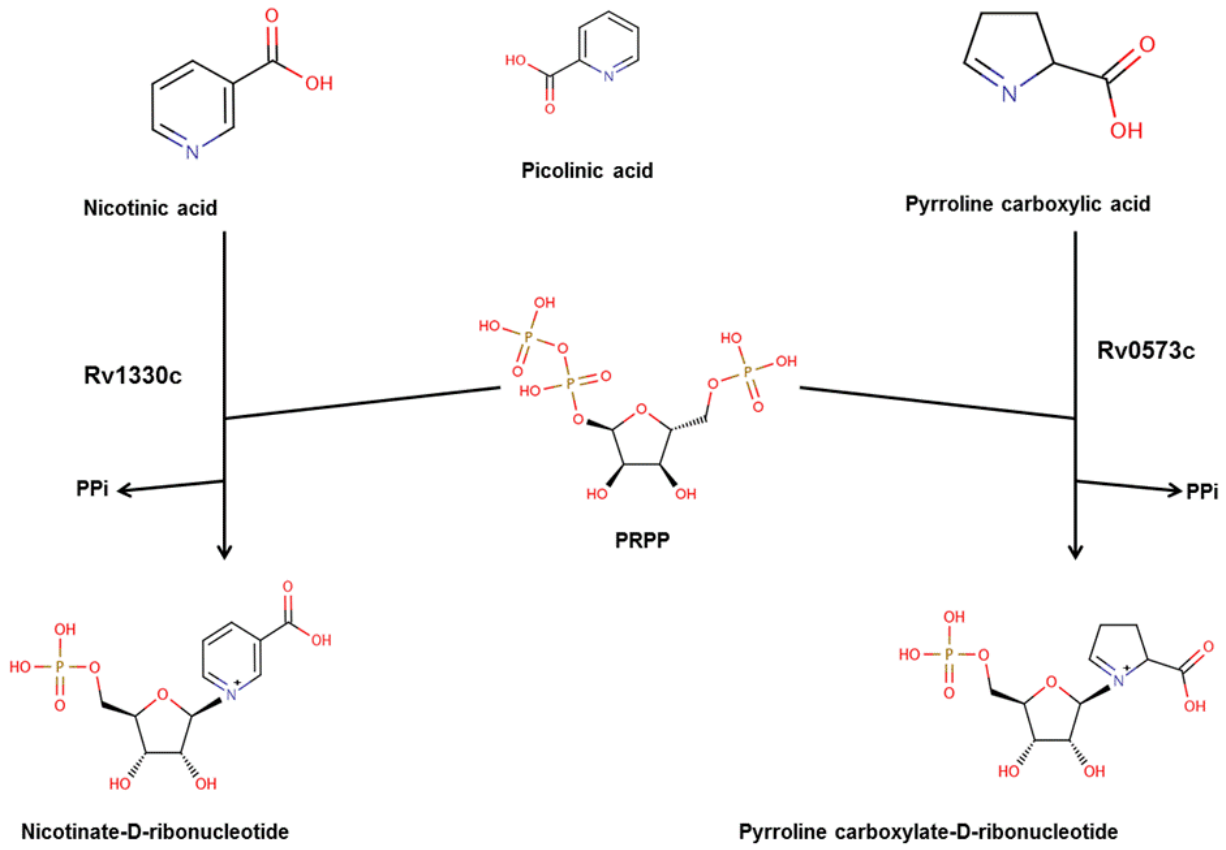


Figure 4.18 Rv0573c's substrate could be pyrroline carboxylic acid which resembles nicotinic and picolinic acid. Picolinic acid (top row, center) is an isomer of nicotinic acid (top row, left). The carboxylic group is located in the second and third position of the pyridine ring, respectively. Pyrroline carboxylic acid only has a five atom ring (top row, right). However, for both pyrroline carboxylic acid and picolinic acid, notice the similar position of the carboxylic acid in reference to the nitrogen group. Rv1330c's activity is depicted on the left, resulting in the generation of nicotinate-D-ribonucleotide. Rv0573c's hypothetical activity is depicted on the right. Its activity produces a hypothetical compound: pyrroline carboxylate-D-ribonucleotide. PRPP, phosphoribosyl pyrophosphate. PPi, pyrophosphate.

It is possible that this reaction could contribute to the previously reported increased hepatic production of PRPP and ribose-phosphate in mice receiving an IV injection of pyrroline carboxylate [93].

Pyrroline carboxylate-D-ribonucleotide could then be the substrate for Rv0575c. As observed in Figure 4.4, an oxygenase domain was detected in Rv0575c. Interestingly, Rv0575c's oxygenase domain is homologous to the one present in salicylate monooxygenase.

As implied by its name, this latter enzyme adds a hydroxyl group to salicylic acid, a carboxylated ring-based molecule sharing some features to the pyrroline carboxylate moiety putatively present in pyrroline carboxylate D-ribonucleotide. Thus, Rv0575c could introduce one oxygen atom to the ring structure present in pyrroline carboxylate D-ribonucleotide (Fig. 4.19). Double bonds would spontaneously reaccommodate in this hydroxylated molecule, to form pyroglutamate-D-ribonucleotide. Pyroglutamate (pyroGlu) is a cyclized form of Glu/Gln.

The last steps of this pathway would be catalyzed by Rv0571c, hypothesized to be a bi-functional protein with two independent domains (see Fig. 4.5). The C-terminus domain of Rv0571c has homology to hydrolases, in particular to dienelactone hydrolases capable of hydrolyzing ring structures [94]. Thus, this domain would hydrolyze the pyroGlu ring present in pyroGlu-D-ribonucleotide and convert it to Glu-D-ribonucleotide (Fig 4.19).

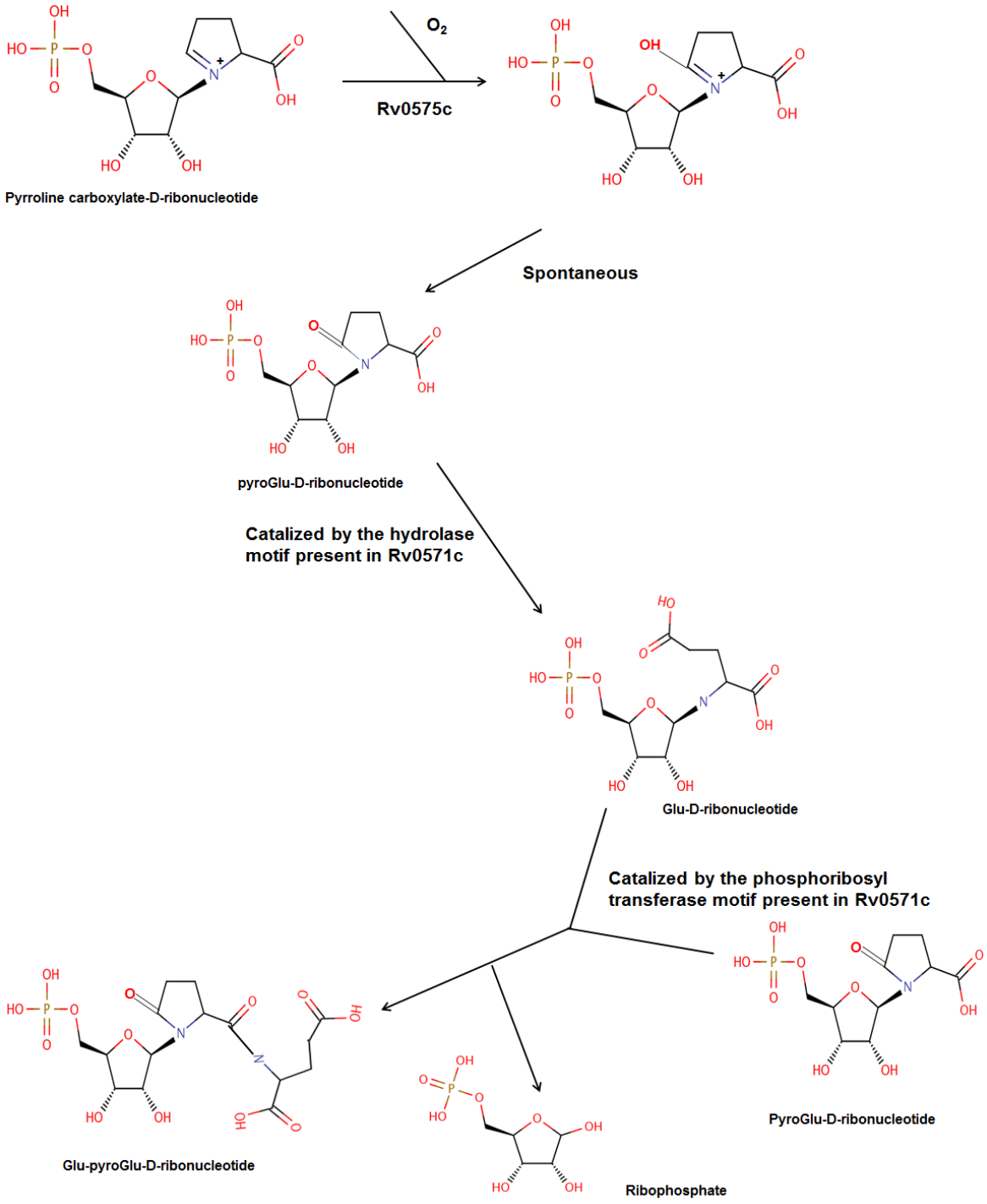


Figure 4.19 Role of the genes encoded in the *rv0574c* locus in a putative biochemical pathway involved in the α-L-polyGln biosynthesis. The pathway begins with Rv0573c which would transfer

pyrroline carboxylate to PRPP in an analogous way as nicotinic phosphoribosyl transferase (Fig. 4.17). Rv0575c would then insert an oxygen atom to pyrroline carboxylate-D-ribonucleotide, which becomes pyroGlu-D-ribonucleotide after spontaneous reaccommodation of the double bonds. Rv0571c has two domains: a hydrolase domain and a phosphoribosyl transferase domain. The hydrolase domain would open the pyroGlu ring and generate Glu-D-ribonucleotide. Finally, the phosphoribosyl transferase domain would release Glu and catalyze a covalent bond between this Glu residue and the free carboxylate moiety from another pyroGlu-D-ribonucleotide. Ribophosphate would be a by-product of this reaction.

Glu-D-ribonucleotide is analogous to an amino acyl-tRNA involved in conventional ribosomal protein synthesis [58]. In amino acyl-tRNA, both molecules are linked by an ester bond between the amino acid's α-carboxyl group and the ribose's 2' hydroxyl group. Instead, in Glu-D-ribonucleotide the linkage would be a glycosidic bond between Glu's α-amino group and ribose's anomeric carbon. During Rv0571c's phosphoribosyl transferase activity, the amino terminus of the Glu residue would be released from the ribose ring and be covalently linked to the carboxyl group of and incoming pyroGlu D-ribonucleotide. Similar to transpeptidases which crosslink amino acids by harnessing the energy released from a cleaved bond, the energy to join the Glu and pyroGlu D-ribonucleotide could come from the cleaved bonds catalyzed by Rv0571c: pyroGlu hydrolysis and/or phosphoribosyl transferase. An iterative reaction consisting of crosslinking Glu and pyroGlu-D-ribonucleotide would occur until the growing poly-Glu chain has achieved a critical size.

Multiple unknowns remain in this hypothetical pathway. For instance, identifying a role for Rv0572c is hampered by its lack of characterized functional domains. In addition, enzymes and subcellular localization of Glu amidation converting poly-Glu to α-L-polyGln is not known. Finally, besides postulating that Rv0574c could participate during extracellular transport of α-L-polyGln, the mechanism and source of energy is still unknown. The proposed pathway could be evaluated using *in vitro* cell free assays or *in vivo* using recombinant technology to express the operon(s) in *M. smeg* [40,

95]. In the first case, individual proteins from the *rv0574c* locus would be heterologous expressed in *E. coli* and recombinant proteins would be purified. As described in [47], an initial phosphoribosyl reaction would be performed using PRPP and pyrroline carboxylic acid. In this case, pyrroline carboxylic acid would be the radiolabeled substrate as it is the precursor for all subsequent steps. In a step by step sequence described in Figure 4.19, each product would be subsequently exposed to the respective enzyme. Intermediate products could be confirmed using TLC and/or mass spectrometry analysis. The final product should constitute of small oligopeptides of Glu residues in an α-linkage. Alternatively, the pathway could be evaluated in *M. smeg* after introducing a cosmid with the *rv0574c* locus. As previously described in the discussion, several genes in the *rv0574c* locus are regulated by the dormancy regulon [86]. Thus, expression of the genes included in this cosmid would require inducing dormancy in *M. smeg*. Fortunately, it has been described that *M. smeg* conserves the two-component regulatory system that regulates dormancy in the presence of low oxygen concentrations [96]. In this case, after inducing dormancy in recombinant *M. smeg* containing the cosmid, oligopeptides of Glu could be obtained in the cytoplasm. Glu oligopeptides could also be recovered in the supernatant instead of the cell wall, as this recombinant *M. smeg* would lack the *rv2394* locus that encodes for the GGT that crosslinks the oligopeptides (see Chapter III).

In conclusion, in this Chapter candidate genes for α-L-polyGln biosynthesis were identified and mutated in order to obtain preliminary results into a biosynthetic pathway and a putative role in pathogenesis. Biochemical analysis identified reduced amounts of Glu and ammonia in the cell wall of *rv0574c* and *rv2394* mutant *Mtb* strains. As Gln is hydrolyzed into Glu and ammonia, lower amounts of these compounds is a good indication that biosynthesis of α-L-polyGln had been affected in the mutant strains. Finally, mice infected with *Mtb* mutants for *rv0574c* or *rv2394* had higher bacterial burden in the early stages of disease and this correlated with faster extra-pulmonary dissemination. In these mice, more severe lung pathology was present during the earlier stage of disease and persisted throughout the course of infection, suggesting that the presence of α-L-polyGln in *Mtb's* cell wall could have an

immunomodulatory function. Unfortunately, the initial phenotype consisting of high Glu/DAP ratios for the WT strain, was not observed while evaluating complementation. Instead, similar Glu/DAP ratios were seen for WT, mutant and complemented strains. Thus, it was not possible to document if complementation could restore the normal ratios in the mutant strains. Further experiments are warranted to obtain more robust data confirming these results.

Literature Cited

1. World Health Organization., *Global tuberculosis control : WHO report 2011*. 2011, Geneva: World Health Organization. viii, 246 p.
2. Cole, S.T., et al., *Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence.* Nature, 1998. **393**(6685): p. 537-44.
3. Mehaffy, C., et al., *Descriptive proteomic analysis shows protein variability between closely related clinical isolates of Mycobacterium tuberculosis.* Proteomics, 2010. **10**(10): p. 1966-84.
4. Barry, C.E., 3rd, *Interpreting cell wall 'virulence factors' of Mycobacterium tuberculosis.* Trends Microbiol, 2001. **9**(5): p. 237-41.
5. Fitzgerald, D.W., T.R. Sterling, and H. D.W., *Mycobacterium tuberculosis*, in *Mandell, Douglas and Bennett's principles and practice of infectious diseases*. 2010, Churchill Livingstone/Elsevier: Philadelphia, PA. p. 3129-3163.
6. Crick, D.C., S. Mahapatra, and P.J. Brennan, *Biosynthesis of the arabinogalactan-peptidoglycan complex of Mycobacterium tuberculosis.* Glycobiology, 2001. **11**(9): p. 107R-118R.
7. Brennan, P.J. and H. Nikaido, *The envelope of mycobacteria.* Annu Rev Biochem, 1995. **64**: p. 29-63.
8. Barry, C.E., 3rd, et al., *Mycolic acids: structure, biosynthesis and physiological functions.* Prog Lipid Res, 1998. **37**(2-3): p. 143-79.
9. Wietzerbin, J., et al., *Occurrence of D-alanyl-(D)-meso-diaminopimelic acid and meso-diaminopimelyl-meso-diaminopimelic acid interpeptide linkages in the peptidoglycan of Mycobacteria.* Biochemistry, 1974. **13**(17): p. 3471-6.
10. Gupta, R., et al., *The Mycobacterium tuberculosis protein LdtMt2 is a nonclassical transpeptidase required for virulence and resistance to amoxicillin.* Nat Med, 2010. **16**(4): p. 466-9.
11. Lavollay, M., et al., *The peptidoglycan of stationary-phase Mycobacterium tuberculosis predominantly contains cross-links generated by L,D-transpeptidation.* J Bacteriol, 2008. **190**(12): p. 4360-6.
12. Raymond, J.B., et al., *Identification of the namH gene, encoding the hydroxylase responsible for the N-glycolylation of the mycobacterial peptidoglycan.* J Biol Chem, 2005. **280**(1): p. 326-33.
13. Fratti, R.A., et al., *Mycobacterium tuberculosis glycosylated phosphatidylinositol causes phagosome maturation arrest.* Proc Natl Acad Sci U S A, 2003. **100**(9): p. 5437-42.
14. Chan, J., et al., *Microbial glycolipids: possible virulence factors that scavenge oxygen radicals.* Proc Natl Acad Sci U S A, 1989. **86**(7): p. 2453-7.
15. Hamasaki, N., et al., *In vivo administration of mycobacterial cord factor (Trehalose 6, 6'-dimycolate) can induce lung and liver granulomas and thymic atrophy in rabbits.* Infect Immun, 2000. **68**(6): p. 3704-9.
16. Ryll, R., Y. Kumazawa, and I. Yano, *Immunological properties of trehalose dimycolate (cord factor) and other mycolic acid-containing glycolipids--a review.* Microbiol Immunol, 2001. **45**(12): p. 801-11.
17. Dubee, V., et al., *Inactivation of Mycobacterium tuberculosis l,d-transpeptidase LdtMt(1) by carbapenems and cephalosporins.* Antimicrob Agents Chemother, 2012. **56**(8): p. 4189-95.
18. Yamamura, Y., et al., *Immunotherapy of cancer with cell wall skeleton of Myocabacterium bovis-Bacillus Calmette-Guerin: experimental and clinical results.* Ann N Y Acad Sci, 1976. **277**(00): p. 209-27.
19. Wang, R., et al., *An anti-neoplastic glycan isolated from Mycobacterium bovis (BCG vaccine).* Biochem J, 1995. **311 (Pt 3)**: p. 867-72.

20. Ortalo-Magne, A., et al., *Molecular composition of the outermost capsular material of the tubercle bacillus.* Microbiology, 1995. **141 (Pt 7)**: p. 1609-20.
21. Migliore, D., N.P. Acharya, and P. Jolles, *[Characterization of important quantities of glutamic acid in the walls of human virulent strains of mycobacteria].* C R Acad Sci Hebd Seances Acad Sci D, 1966. **263**(11): p. 846-8.
22. Candela, T. and A. Fouet, *Poly-gamma-glutamate in bacteria.* Mol Microbiol, 2006. **60**(5): p. 1091-8.
23. Green, B.D., et al., *Demonstration of a capsule plasmid in Bacillus anthracis.* Infect Immun, 1985. **49**(2): p. 291-7.
24. Uchida, I., et al., *Association of the encapsulation of Bacillus anthracis with a 60 megadalton plasmid.* J Gen Microbiol, 1985. **131**(2): p. 363-7.
25. Wietzerbin-Falszpan, J., et al., *The amino acids of the cell wall of Mycobacterium tuberculosis var. bovis, strain BCG. Presence of a poly(L-glutamic acid).* Eur J Biochem, 1973. **32**(3): p. 525-32.
26. Phiet, P.H., et al., *Analysis of the cell wall of five strains of Myocbacterium tuberculosis BCG and of an attenuated human strain, W 115.* Infect Immun, 1976. **13**(3): p. 677-81.
27. Wietzerbin, J., F. Lederer, and J.F. Petit, *Structural study of the poly-l-Glutamic acid of the cell wall of Mycobacterium tuberculosis var hominis, strain Brevannes.* Biochem Biophys Res Commun, 1975. **62**(2): p. 246-52.
28. Hirschfield, G.R., M. McNeil, and P.J. Brennan, *Peptidoglycan-associated polypeptides of Mycobacterium tuberculosis.* J Bacteriol, 1990. **172**(2): p. 1005-13.
29. Makino, S., et al., *Molecular characterization and protein analysis of the cap region, which is essential for encapsulation in Bacillus anthracis.* J Bacteriol, 1989. **171**(2): p. 722-30.
30. Ashiuchi, M., K. Soda, and H. Misono, *A poly-gamma-glutamate synthetic system of Bacillus subtilis IFO 3336: gene cloning and biochemical analysis of poly-gamma-glutamate produced by Escherichia coli clone cells.* Biochem Biophys Res Commun, 1999. **263**(1): p. 6-12.
31. Urushibata, Y., S. Tokuyama, and Y. Tahara, *Characterization of the Bacillus subtilis ywsC gene, involved in gamma-polyglutamic acid production.* J Bacteriol, 2002. **184**(2): p. 337-43.
32. Candela, T., M. Mock, and A. Fouet, *CapE, a 47-amino-acid peptide, is necessary for Bacillus anthracis polyglutamate capsule synthesis.* J Bacteriol, 2005. **187**(22): p. 7765-72.
33. Candela, T. and A. Fouet, *Bacillus anthracis CapD, belonging to the gamma-glutamyltranspeptidase family, is required for the covalent anchoring of capsule to peptidoglycan.* Mol Microbiol, 2005. **57**(3): p. 717-26.
34. Wu, R., et al., *Crystal structure of Bacillus anthracis transpeptidase enzyme CapD.* J Biol Chem, 2009. **284**(36): p. 24406-14.
35. Richter, S., et al., *Capsule anchoring in Bacillus anthracis occurs by a transpeptidation reaction that is inhibited by capsidin.* Mol Microbiol, 2009. **71**(2): p. 404-20.
36. Harth, G., et al., *Treatment of Mycobacterium tuberculosis with antisense oligonucleotides to glutamine synthetase mRNA inhibits glutamine synthetase activity, formation of the poly-L-glutamate/glutamine cell wall structure, and bacterial replication.* Proc Natl Acad Sci U S A, 2000. **97**(1): p. 418-23.
37. Harth, G., D.L. Clemens, and M.A. Horwitz, *Glutamine synthetase of Mycobacterium tuberculosis: extracellular release and characterization of its enzymatic activity.* Proc Natl Acad Sci U S A, 1994. **91**(20): p. 9342-6.
38. Harth, G., et al., *All four Mycobacterium tuberculosis glnA genes encode glutamine synthetase activities but only GlnA1 is abundantly expressed and essential for bacterial homeostasis.* Mol Microbiol, 2005. **58**(4): p. 1157-72.
39. Tullius, M.V., G. Harth, and M.A. Horwitz, *High extracellular levels of Mycobacterium tuberculosis glutamine synthetase and superoxide dismutase in actively growing cultures are due*

to high expression and extracellular stability rather than to a protein-specific export mechanism. Infect Immun, 2001. **69**(10): p. 6348-63.
40. Braunstein, M., S.S. Bardarov, and W.R. Jacobs, Jr., *Genetic methods for deciphering virulence determinants of Mycobacterium tuberculosis.* Methods Enzymol, 2002. **358**: p. 67-99.
41. Jacobs, W.R., Jr., et al., *Genetic systems for mycobacteria.* Methods Enzymol, 1991. **204**: p. 537-55.
42. van Helden PD, et al., *Mycobacterium tuberculosis protocols*, in *Methods in molecular medicine*, T. Parish and N.G. Stoker, Editors. 2001, Humana Press: Totowa, NJ. p. 19-30.
43. Slayden, R.A. and J.T. Belisle, *Morphological features and signature gene response elicited by inactivation of FtsI in Mycobacterium tuberculosis.* J Antimicrob Chemother, 2009. **63**(3): p. 451-7.
44. R.A., S. and B. C.E, *Mycobacterium tuberculosis protocols*, in *Methods in molecular medicine*, T. Parish and N.G. Stoker, Editors. 2001, Humana Press: Totowa, NJ. p. 229-245.
45. Ordway, D.J. and I.M. Orme, *Animal models of mycobacteria infection.* Curr Protoc Immunol, 2011. **Chapter 19**: p. Unit19 5.
46. Sharpe, S.A., et al., *Determination of lesion volume by MRI and stereology in a macaque model of tuberculosis.* Tuberculosis (Edinb), 2009. **89**(6): p. 405-16.
47. Boshoff, H.I., et al., *Biosynthesis and recycling of nicotinamide cofactors in mycobacterium tuberculosis. An essential role for NAD in nonreplicating bacilli.* J Biol Chem, 2008. **283**(28): p. 19329-41.
48. Sassetti, C.M., D.H. Boyd, and E.J. Rubin, *Genes required for mycobacterial growth defined by high density mutagenesis.* Mol Microbiol, 2003. **48**(1): p. 77-84.
49. Huet, G., et al., *A lipid profile typifies the Beijing strains of Mycobacterium tuberculosis: identification of a mutation responsible for a modification of the structures of phthiocerol dimycocerosates and phenolic glycolipids.* J Biol Chem, 2009. **284**(40): p. 27101-13.
50. Giovannini, D., et al., *A new Mycobacterium tuberculosis smooth colony reduces growth inside human macrophages and represses PDIM Operon gene expression. Does an heterogeneous population exist in intracellular mycobacteria?* Microb Pathog, 2012. **53**(3-4): p. 135-46.
51. Ioerger, T.R., et al., *Variation among genome sequences of H37Rv strains of Mycobacterium tuberculosis from multiple laboratories.* J Bacteriol, 2010. **192**(14): p. 3645-53.
52. Domenech, P. and M.B. Reed, *Rapid and spontaneous loss of phthiocerol dimycocerosate (PDIM) from Mycobacterium tuberculosis grown in vitro: implications for virulence studies.* Microbiology, 2009. **155**(Pt 11): p. 3532-43.
53. Behr, M.A., et al., *Comparative genomics of BCG vaccines by whole-genome DNA microarray.* Science, 1999. **284**(5419): p. 1520-3.
54. Rousseau, C., et al., *Sulfolipid deficiency does not affect the virulence of Mycobacterium tuberculosis H37Rv in mice and guinea pigs.* Infect Immun, 2003. **71**(8): p. 4684-90.
55. Moran, N.A., *Microbial minimalism: genome reduction in bacterial pathogens.* Cell, 2002. **108**(5): p. 583-6.
56. Lautru, S. and G.L. Challis, *Substrate recognition by nonribosomal peptide synthetase multi-enzymes.* Microbiology, 2004. **150**(Pt 6): p. 1629-36.
57. Finking, R. and M.A. Marahiel, *Biosynthesis of nonribosomal peptides1.* Annu Rev Microbiol, 2004. **58**: p. 453-88.
58. Ibba, M. and D. Soll, *Aminoacyl-tRNA synthesis.* Annu Rev Biochem, 2000. **69**: p. 617-50.
59. De Voss, J.J., et al., *The salicylate-derived mycobactin siderophores of Mycobacterium tuberculosis are essential for growth in macrophages.* Proc Natl Acad Sci U S A, 2000. **97**(3): p. 1252-7.
60. Wang, F., et al., *Identification of a type III thioesterase reveals the function of an operon crucial for Mtb virulence.* Chem Biol, 2007. **14**(5): p. 543-51.

61. Chhabra, A., et al., *Nonprocessive [2 + 2]e- off-loading reductase domains from mycobacterial nonribosomal peptide synthetases.* Proc Natl Acad Sci U S A, 2012. **109**(15): p. 5681-6.
62. Eveland, S.S., D.L. Pompliano, and M.S. Anderson, *Conditionally lethal Escherichia coli murein mutants contain point defects that map to regions conserved among murein and folyl poly-gamma-glutamate ligases: identification of a ligase superfamily.* Biochemistry, 1997. **36**(20): p. 6223-9.
63. Bermingham, A. and J.P. Derrick, *The folic acid biosynthesis pathway in bacteria: evaluation of potential for antibacterial drug discovery.* Bioessays, 2002. **24**(7): p. 637-48.
64. Mock, M. and A. Fouet, *Anthrax.* Annu Rev Microbiol, 2001. **55**: p. 647-71.
65. Zwartouw, H.T. and H. Smith, *Polyglutamic acid from Bacillus anthracis grown in vivo; structure and aggressin activity.* Biochem J, 1956. **63**(3): p. 437-42.
66. Weintraub, A., *Immunology of bacterial polysaccharide antigens.* Carbohydr Res, 2003. **338**(23): p. 2539-47.
67. Schneerson, R., et al., *Poly(gamma-D-glutamic acid) protein conjugates induce IgG antibodies in mice to the capsule of Bacillus anthracis: a potential addition to the anthrax vaccine.* Proc Natl Acad Sci U S A, 2003. **100**(15): p. 8945-50.
68. Lee, D.Y., et al., *Poly-gamma-d-glutamic acid and protective antigen conjugate vaccines induce functional antibodies against the protective antigen and capsule of Bacillus anthracis in guinea-pigs and rabbits.* FEMS Immunol Med Microbiol, 2009. **57**(2): p. 165-72.
69. Kocianova, S., et al., *Key role of poly-gamma-DL-glutamic acid in immune evasion and virulence of Staphylococcus epidermidis.* J Clin Invest, 2005. **115**(3): p. 688-94.
70. Hirschfield, G.R., *Cell wall polypeptides of Mycobacterium tuberculosis physiological and immunological significance*, 1990.
71. Ross, C.A., et al., *Polyglutamine pathogenesis.* Philos Trans R Soc Lond B Biol Sci, 1999. **354**(1386): p. 1005-11.
72. Perutz, M.F., *Glutamine repeats and inherited neurodegenerative diseases: molecular aspects.* Curr Opin Struct Biol, 1996. **6**(6): p. 848-58.
73. Takahashi, T., S. Katada, and O. Onodera, *Polyglutamine diseases: where does toxicity come from? what is toxicity? where are we going?* J Mol Cell Biol, 2010. **2**(4): p. 180-91.
74. Li, X., H. Li, and X.J. Li, *Intracellular degradation of misfolded proteins in polyglutamine neurodegenerative diseases.* Brain Res Rev, 2008. **59**(1): p. 245-52.
75. Goswami, A., et al., *Expression of expanded polyglutamine proteins suppresses the activation of transcription factor NFkappaB.* J Biol Chem, 2006. **281**(48): p. 37017-24.
76. Reijonen, S., et al., *Downregulation of NF-kappaB signaling by mutant huntingtin proteins induces oxidative stress and cell death.* Cell Mol Life Sci, 2010. **67**(11): p. 1929-41.
77. Rousseau, C., et al., *Deficiency in mycolipenate- and mycosanoate-derived acyltrehaloses enhances early interactions of Mycobacterium tuberculosis with host cells.* Cell Microbiol, 2003. **5**(6): p. 405-15.
78. Goren, M.B., O. Brokl, and W.B. Schaefer, *Lipids of putative relevance to virulence in Mycobacterium tuberculosis: phthiocerol dimycocerosate and the attenuation indicator lipid.* Infect Immun, 1974. **9**(1): p. 150-8.
79. Cox, J.S., et al., *Complex lipid determines tissue-specific replication of Mycobacterium tuberculosis in mice.* Nature, 1999. **402**(6757): p. 79-83.
80. Jain, M., et al., *Lipidomics reveals control of Mycobacterium tuberculosis virulence lipids via metabolic coupling.* Proc Natl Acad Sci U S A, 2007. **104**(12): p. 5133-8.
81. Brunengraber, H. and C.R. Roe, *Anaplerotic molecules: current and future.* J Inherit Metab Dis, 2006. **29**(2-3): p. 327-31.
82. Garton, N.J., et al., *Intracellular lipophilic inclusions of mycobacteria in vitro and in sputum.* Microbiology, 2002. **148**(Pt 10): p. 2951-8.

83. Garton, N.J., et al., *Cytological and transcript analyses reveal fat and lazy persister-like bacilli in tuberculous sputum.* PLoS Med, 2008. **5**(4): p. e75.
84. Daniel, J., et al., *Induction of a novel class of diacylglycerol acyltransferases and triacylglycerol accumulation in Mycobacterium tuberculosis as it goes into a dormancy-like state in culture.* J Bacteriol, 2004. **186**(15): p. 5017-30.
85. Sirakova, T.D., et al., *Identification of a diacylglycerol acyltransferase gene involved in accumulation of triacylglycerol in Mycobacterium tuberculosis under stress.* Microbiology, 2006. **152**(Pt 9): p. 2717-25.
86. Park, H.D., et al., *Rv3133c/dosR is a transcription factor that mediates the hypoxic response of Mycobacterium tuberculosis.* Mol Microbiol, 2003. **48**(3): p. 833-43.
87. Raynaud, C., et al., *Extracellular enzyme activities potentially involved in the pathogenicity of Mycobacterium tuberculosis.* Microbiology, 1998. **144 (Pt 2)**: p. 577-87.
88. Muller, B., et al., *African 1, an epidemiologically important clonal complex of Mycobacterium bovis dominant in Mali, Nigeria, Cameroon, and Chad.* J Bacteriol, 2009. **191**(6): p. 1951-60.
89. Tsolaki, A.G., et al., *Functional and evolutionary genomics of Mycobacterium tuberculosis: insights from genomic deletions in 100 strains.* Proc Natl Acad Sci U S A, 2004. **101**(14): p. 4865-70.
90. Rodrigo, T., et al., *Characteristics of tuberculosis patients who generate secondary cases.* Int J Tuberc Lung Dis, 1997. **1**(4): p. 352-7.
91. Bosco, M.C., et al., *Macrophage activating properties of the tryptophan catabolite picolinic acid.* Adv Exp Med Biol, 2003. **527**: p. 55-65.
92. Hu, C.A., et al., *Human Delta1-pyrroline-5-carboxylate synthase: function and regulation.* Amino Acids, 2008. **35**(4): p. 665-72.
93. Boer, P. and O. Sperling, *Stimulation of ribose-5-phosphate and 5-phosphoribosyl-1-pyrophosphate generation by pyrroline-5-carboxylate in mouse liver in vivo: evidence for a regulatory role of ribose-5-phosphate availability in nucleotide synthesis.* Biochem Med Metab Biol, 1991. **46**(1): p. 28-32.
94. Holmquist, M., *Alpha/Beta-hydrolase fold enzymes: structures, functions and mechanisms.* Curr Protein Pept Sci, 2000. **1**(2): p. 209-35.
95. Bange, F.C., F.M. Collins, and W.R. Jacobs, Jr., *Survival of mice infected with Mycobacterium smegmatis containing large DNA fragments from Mycobacterium tuberculosis.* Tuber Lung Dis, 1999. **79**(3): p. 171-80.
96. Mayuri, et al., *Molecular analysis of the dormancy response in Mycobacterium smegmatis: expression analysis of genes encoding the DevR-DevS two-component system, Rv3134c and chaperone alpha-crystallin homologues.* FEMS Microbiol Lett, 2002. **211**(2): p. 231-7.

Chapter V

Final Discussion

"For we cannot stop where we are, nor at once accept as a solution to all our difficulties the theory which has its avowed basis in experiment and not in clinical observation, and which is as yet rather the fruit of the laboratory than of the hospital" James E. Pollock, Croonian Lectures, 1883. [1]

Epidemiological studies evaluating skin test positivity have estimated that one third of the world's population is infected with *Mtb* [2]. Approximately 90% of the infected individuals will never develop TB and are therefore considered to have latent TB infection (LTBI) [3]. In contrast, 10% of these infected individuals will develop active disease during their life time and maintain the infectious cycle of *Mtb*. Thus, 3% (10% of 30%) of the world's population will develop TB at some point during their life and this staggering number of sick patients has overwhelmed health systems, leading to the current global health problem. In developed countries where TB isn't a significant health problem, guidelines have been clearly established to identify and treat individuals with LTBI having underlying risk factors for developing active disease [4]. Even though 9 months of prophylactic treatment with INH effectively accomplishes this [5], treatment of individuals with LTBI in developing countries is simply not feasible. Alternative methods to control developing active disease from LTBI are urgently needed and probably will come from better understanding the bacteria's and immune system's physiology during this stage. Unfortunately, one major drawback is that LTBI is mainly observed in human beings the natural host of *Mtb* and this limits the amount of interventions that can be applied to collect significant information [6]. Even though several animals are susceptible to *Mtb* and have been fundamental in understanding many aspects of TB, our current animal models don't recapitulate LTBI. Besides some experimental data obtained with non-human primates [7], LTBI is not seen in the more commonly used guinea pig, murine and rabbit models. Instead of LTBI and in stark contrast to only 10% of infected humans, a chronic but progressive disease is observed in 100% of experimentally infected animals [8]. It should be stated that this holds true even when infecting with doses as low as 10 bacilli [9], which mimic the estimated *Mtb* infectiousness in humans. A more promising approach to understand LTBI has been to naturally infect

guinea pigs with air obtained from hospital wards housing TB patients [10-12]. By skin testing, it was shown the majority of exposed guinea pigs got infected and developed an acquired delayed-type hypersensitivity (DTH) to *Mtb* antigens [12]. Remarkably and in agreement with human infection, only 12% of infected animals developed active TB. Thus, the remaining infected guinea pigs with a positive skin test behaved as humans with LTBI as they didn't develop active disease [12]. In this final Chapter and in the context of *Mtb* physiology and metabolism, I attempt to give an alternate insight into the role that *Mtb* dormancy could play during transmission from infected to naïve indiviudals. This could have implications towards understanding human LTBI and could be applied to develop animal models of LTBI for drug and vaccine development.

LTBI is a clinical term that refers to individuals with a positive skin test to *Mtb* antigens in the absence of clinical symptoms [3, 13]. This DTH is mediated by Tcells and is attributed to a recall response to a previous exposure to *Mtb*. It is well known that false positive results during skin testing can be attributed to BCG vaccination [13] and exposure to other non-tuberculosis mycobacteria, but this will not be further discussed in this Chapter. As patients with LTBI are asymptomatic, skin testing is currently the principal technique to detect them. Ideally, further evaluation of these patients is performed by obtaining a chest X-ray [3]. This allows identification of patients with asymptomatic active disease that require immediate treatment, as well as leading to antibiotic treatment of exposed high risk individuals that might develop active disease in the future [4, 7]. Unfortunately, the majority of the patients with LTBI don't have any radiological evidence of prior infection, instead having inconspicuously normal chest X-rays [3]. In these patients, a DTH response is the only evidence that a previous infection occurred. In contrast and as reported in the Profit study in England, approximately 10% of patients who converted to a positive skin test against *Mtb* antigens, had X-ray findings compatible with a primary infection [14]. These findings principally consist of what has been referred to as the Ghon complex [15] in which usually a small lesion is observed in the lung parenchyma as well as in the draining lymph nodes

(the primary lesion). If the primary lesion heals and calcifies, the Ghon complex is referred to as the Renke complex [15] and in the clinical setting, this is associated with good prognosis.

From a pathogenesis perspective, the Ghon complex is a reflection that *Mtb* was able to colonize and multiply, recruiting host immune cells which ultimately coalesce to form the primary lesion. In fact, it has been estimated that the primary lesion can harbor 10^2 to 10^4 bacilli [3, 16]. The obvious question is why the majority of individuals with LTBI fail to develop a primary lesion (as evidenced by a lack of radiological findings) but have a DTH response to *Mtb* antigens? Besides the aforementioned possibility that this DTH response could be due to cross reaction to BCG vaccination or exposure to non-tuberculosis mycobacteria, several alternative hypotheses could be envisioned. A simple scenario is that *Mtb* was efficiently eradicated by the host immune system and was not able to multiply. This is consistent with the results obtained when guinea pigs were exposed to air from TB hospital wards [12]. Alternatively, the amount of bacterial multiplication was not sufficient to recruit enough inflammatory cells. Furthermore, the individual's immune system could be compromised such that it is not able to mount a significant inflammatory response. This has been observed in some AIDS patient with normal X-rays despite being diagnosed with TB by sputum culture [17]. One last consideration is that *Mtb* was able to colonize but remains in a dormant stage with minimum multiplication (bacterial dormancy will be discussed below). Indeed, several studies have confirmed the presence of *Mtb* in completely normal lung tissues. Several decades ago, it was shown that transferring macroscopically normal human lung tissue to guinea pigs would result in the animal developing TB [18]. Additionally, conventional acid fast staining and molecular techniques such as PCR have identified bacilli and *Mtb* gDNA respectively, in microscopically normal human lung tissue obtained from individuals dying of other cause besides TB [16]. In contrast in the guinea pig model, imaging techniques such as magnetic resonance imaging (MRI) have observed primary granulomas in equivalent numbers to the desired infection dose [19]. Thus, it seems that in contrast to humans, in our animal models every bacillus replicates, induces significant inflammatory cell influx and animals develop primary lesions.

As mentioned, LTBI could be attributed to the presence of *Mtb* in a dormant state. Dormancy refers to a bacterial physiology state consisting of low replication and metabolism. *In vitro* studies characterizing *Mtb* dormancy were pioneered by Wayne *et al* [20]. In their studies, the stimulus to induce dormancy of *Mtb* was gradual oxygen depletion [21]. As determined by densitometry measurements and methylene blue reduction, *Mtb* would replicate until microaerophilic conditions occurred. At this point, small increases in OD were observed and attributed to an increase in cell size rather than increased bacterial numbers. When complete hypoxia occurred, OD of *Mtb* cultures completely stopped increasing. Dormant bacilli could resume growth by simply transferring cultures to normal oxygen conditions. The fact that *Mtb* stopped growing during hypoxia, resumed growth during normoxia, and more importantly to separate dormancy (bacterial physiology) from latency (clinical term), led to some investigators coining the term "non-replicative persistence" instead [20]. It is widely known that *Mtb* is an aerophilic microorganism and this characteristic is used as a biochemical parameter to identify *Mtb* in the clinical setting [3]. Even though oxygen deprivation still remains as the "gold standard" to induce *Mtb* dormancy, recent studies have shown this also occurs when cultures are subjected to nutrient deprivation [22], high nitric oxide [23], and carbon monoxide concentrations [24], low pH etc. Indeed, it has been suggested that all these conditions could coexist *in vivo* and *Mtb* dormancy was observed *in vitro* in a model simultaneously replicating all of these [25]. Thus, it seems that bacterial dormancy is a general response when unfavorable conditions incompatible with bacterial multiplication are present. From a biochemical perspective, *Mtb* dormancy was associated with increased nitrate reduction as a surrogate terminal electron acceptor when oxygen is lacking [26]. Increased isocitrate lyase (icl) activity was also found in dormant mycobacteria and suggested to metabolize glyoxylate resulting from Gly deamination [27].

A major boost in the field of *Mtb* dormancy occurred with the identification of defined biochemical pathways leading to it. Dormancy was shown to be dependent on the activation of a two-component regulatory system involving the kinases DosS and DosT and the response regulator DosR, which acts as a transcription factor [28-30]. Approximately 50 genes were shown to be upregulated by

DosS/DosR and the name dormancy regulon was coined to refer to them [29]. The function of several proteins belonging to the dormancy regulon has been characterized and include proteins such as Rv1915 (icl1, isocitrate lyase) [31] , Rv3130c (tgs1, a triglyceride synthase) [32], Rv3097c (lipY, a lipase) [33], Rv2031c (hspx, a small chaperonin) [28] etc. It should be mentioned that more important functions have been recently attributed to icl beyond Wayne's suggestion of its role in Gly metabolism [31]. Mutating the *icl1* gene abrogated the bacillus ability to survive in *in vitro* cultures containing lipids as principal carbon source. Significantly, mutated bacteria did not thrive during the chronic phase of the murine TB model [34]. The icl shunt or bypass pathway is critical in organisms using lipids as only carbon source as it allows conservation of carbon units for gluconeogenesis [35]. Thus, it was hypothesized that during the chronic phase of infection *Mtb* switches to using lipids as carbon source. However, it was later shown by mutating both *icl1* and *icl2* genes that *Mtb* probably uses lipids as carbon source during its entire host life cycle [34]. The proposed carbon source switch is probably a reflection of the fact that *Mtb* is principally grown *in vitro* in media containing glycerol or glucose as major carbon sources (discussed below). Paradoxically, it was reported that *icl* expression predominantly in non-necrotic granulomas where hypoxia is not present [36]. Further studies are required to explain why this gene is highly upregulated during *in vitro* oxygen deprivation but not so *in vivo* in necrotic granulomas where hypoxia and dormancy have been well documented. In addition to specific protein upregulation and function, several phenotypic characteristics have been determined to be present in dormant *Mtb* and became important biomarkers for this physiological state. For instance, Rv3130c's triglyceride synthase activity leads to the bacillus accumulating intracellular lipid bodies [25, 32, 37], which are readily detected by Nile red staining and this staining technique is routinely used in dormancy studies [38]. However, one aspect that has not been considered during dormancy is the counterintuitive upregulation of enzymes simultaneous synthesizing (Rv3130c) and catabolizing lipids (Rv3097c), possibly ensuing a futile cycle. An additional critical phenotypic characteristic is that dormant bacteria are resistant to antibiotics [22, 39]. Antibiotic targets are minimally expressed in dormant bacteria, hence the phenotypic resistance. It should be remembered that

antibiotics against *Mtb* have only been around for approximately 60-70 years. In contrast, the dormancy regulon is conserved in several members of the *Mtb* complex that underwent different selection pressures after diverging thousands of years ago [40]. Even though dormant *Mtb* is a cause for the protracted antibiotic treatment of TB, antibiotics have not been the driving cause for the conservation of the dormancy regulon. Instead, the conservation of this regulon in different mycobacterial species is a clear testimony that it has a key but still unknown role in pathogenesis.

Considering the critical role of DosS/DosR in dormancy and the possibility that it could answer some questions about TB pathogenesis, it was highly anticipated that mutants in this pathway could be enlightening. However, this was not the case and at least in mice it seems this pathway doesn't have a significant role [41]. In other animal models, conflicting results have emerged with DosS/DosR mutants [42-44] but it seems like many in the TB field have moved away from it. Instead, focus has been shifted to a set of genes known as the enduring hypoxic response (EHR) [41]. This name highlights the fact that in contrast to genes in the dormancy regulon, expression of genes in the EHR remains for a prolong period of time after *Mtb* cultures are moved into hypoxic conditions. Unfortunately, so far no *in vivo* data has been reported for mutants in the EHR. It was assumed that *Mtb* mutants in the dormancy regulon would not thrive in the chronic phase of TB. This seems like an erroneous assumption for several reasons. In our current animal models, curves enumerating bacterial burden show exponential growth until approximately day 30 post-infection. This is followed by a plateau characterized by a constant number of bacilli. It had therefore been considered that bacterial multiplication did not occur in the chronic phase of TB [45]. However and in contrast to dormant bacteria, recent mechanistic data following the loss of a plasmid as a "replicative clock" has contradicted this belief and it seems clear that *Mtb* replication continues during the chronic phase [46]. Furthermore and again in contrast to dormant bacteria, most *Mtb* present in the chronic murine model (>30 days) of TB are susceptible to antibiotics [47, 48]. This is also observed in the chronic guinea pig model as bacterial burden rapidly declines when conventional antibiotic treatment is initiated even 60 days post *Mtb* infection [49]. In conclusion, most bacteria

surviving or persisting into chronic stages of TB do not display at least two major characteristics present in dormant *Mtb*: lack of replication and phenotypic resistance to antibiotics. Thus bacterial persistence in chronic stages of the disease should not be equated to dormant bacteria. As only a small fraction of bacilli would be dormant in chronic stages of TB, dormancy mutants wouldn't significantly impact bacterial burden. Lack of distinction between these two situations could have led the field to prematurely move away from the dormancy regulon.

What then is the role of the dormancy regulon in the pathogenesis of TB and LTBI? Two independent observations could hold the key to answer these questions. Abundant lipid-loaded *Mtb*, an indicator of dormant bacteria (see above), were recently reported to be present in sputum of TB patients [38, 50]. In addition to staining with Nile Red, transcriptional analysis confirmed the expression of dormancy genes quite similar to those reported from *in vitro* dormant cultures [50]. By immunohistochemistry techniques, multiple phenotypes of *Mtb* have been reported in lung lesions [51, 52] and the presence of bacteria with a dormant phenotype in sputum could simply be a casual finding. But what if dormant bacteria are actually important during *Mtb* transmission, hence their presence in sputum? The second observation was reported decades ago and actually represents one of the initial suggestions that *in vivo Mtb* preferentially uses lipids as carbon source [53]. Importantly, *Mtb* isolated from infected lungs could only be stimulated to respire in the presence of infected but not naïve lung extracts. *In vitro* subculturing of *Mtb* isolated from tissues would abrogate the bacteria's respiration reliance on extracts from infected tissues. This result could be attributed to essential metabolites conducive to bacterial respiration and multiplication, being present only in infected but not naïve lungs. Metabolic differences between naïve and infected tissue should be expected taking in consideration the drastic changes occurring during inflammation. For instance, the enhanced vascular permeability occurring during inflammation could increase metabolite tissue concentrations. Furthermore, as inflammatory cells become activated, their metabolism is radically changed. Instead of using lipids or glucose via oxidative phosphorylation, activated cells have been shown to perform aerobic glycolysis (the

Warburg effect) [54]. Despite the presence of oxygen, activated cells metabolize glucose into lactate and perform increased glutaminolysis in order to provide building blocks for macromolecular biosynthesis. Indeed, metabolite differences compatible with the Warburg effect were observed in nuclear magnetic resonance (NMR) analysis of whole lungs obtained from naïve vs. *Mtb* infected animals [55]. Further insight into the metabolite composition of *Mtb* infected lungs was achieved after laser assisted dissection of human necrotic granulomas. By thin layer chromatography (TLC) and mass spectrometry analysis, abundant triacylglycerols and cholesterylesters were characterized in these lesions and could represent the carbon source being accumulated in lipid bodies of extracellular *Mtb* [56]. In fact, *Mtb* accumulation of lipid bodies could start during its intracellular life cycle inside macrophages. A pathological observation in TB is the presences of 'foamy' macrophages harboring intracellular *Mtb* [57]. 'Foamy' macrophages also accumulate intracellular lipid bodies giving the cell its 'foamy' appearance. It has been reported that *Mtb* can obtain its carbon source from the lipid bodies present inside 'foamy' macrophages [58, 59].

Besides carbon, a source of nitrogen is also required for macromolecular biosynthesis. Specifically for *Mtb*, it has been shown that Gln could be an important nitrogen source. The NMR study previously mentioned, also reported that *Mtb* infected lungs contain an abundant concentration of Glu [55], the precursor for Gln synthesis. *Mtb* genome encodes for four Gln synthatases, with Rv2220 (glnA1) being the most physiologically relevant one [60]. *In vitro* studies have shown that *rv2220* mutants failed to thrive in macrophages [61]. In addition, pharmacological inhibition of *Mtb* Gln synthatase was shown to reduce *in vivo* bacterial burden in the guinea pig model [62], and therefore has been considered as an alternative antibiotic target. Interestingly and as discussed in Chapter IV, the dormancy regulon includes *rv0574c-0571c* [29], the putative α-L-polyGln biosynthetic machinery. As implied by its name, α-L-polyGln is a polymer of L-Gln residues [63]. Even though the function of this polymer is currently unknown, one possibility is that analogous to the role of glycogen as a glucose storage system, α-L-polyGln could represent an L-Gln depot. In a concerted action between Rv2220 and the biosynthetic machinery for α-L-polyGln, *Mtb* could take advantage of the high Glu concentrations present in infected

lungs and replenish its L-Gln stores for future needs. Indeed, inhibition of *Mtb* Gln synthatase with antisense nucleotides led to reduced α-L-polyGln [64]. Similar to the role played by *B. anthracis* CapD in regulating polymerization or release of poly-γ-D-Glu [65], Rv2394 (Chapter III) could have a dual role in the physiological role played by *Mtb* α-L-polyGln. While *Mtb* is storing Gln as α-L-polyGln, Rv2394 could be crosslinking oligopeptides of α-L-Gln into an insoluble polymer that remains associated with the cell wall. In contrast, when *Mtb* requires Gln, Rv2394 could be releasing Gln oligopeptides that *Mtb* could import and use. The dual function in Rv2394 as a hydrolase and a transpeptidase, empowers the enzyme to regulate α-L-polyGln concentration. Indeed, experimental evidence was obtained suggesting that in the presence of physiologically relevant acceptors, Rv2394 could mediate a transpeptidase reaction. If acceptors were not present, Rv2394 could perform a hydrolase reaction and cleave the donor. Finally, in regards to nutrient storage, it was recently reported that inflamed lungs contained high intra and extracellular deposits of iron [66], an additional micronutrient required by *Mtb*. In conclusion, in contrast to naïve lungs, inflamed lungs constitute a nutrient-replete environment where *Mtb* can obtain and accumulate carbon and nitrogen sources, as well as iron. These metabolite alterations are not circumscribed to the inflamed tissue but are also present systemically. Specifically in chronic inflammatory conditions such as TB and cancer, adipose tissue and muscle catabolism resulting in cachexia are probably fueling tissular accumulation of metabolites.

In this order of ideas, I propose that during transmission from a sick to a naïve individual, dormant as opposed to actively growing *Mtb* gives the bacillus the best probability for efficient colonization of the new host. This is in contrast to the current dogma that actively growing *Mtb* is responsible for transmission. As *Mtb* is transmitted from infected to naïve lungs, the bacilli will suddenly be confronted with a nutrient-deficient environment. Similar to what occurs during an abrupt oxygen withdrawal [21], this sudden loss of nutrients could be catastrophic to actively growing bacilli. This is consistent with the abrupt respiration shutdown in an *ex vivo* model of *Mtb* transmission [53]. In contrast, this change wouldn't affect dormant bacilli as much. It is known that dormant bacilli have gradually

minimized respiration and can actually survive with significantly reduced intracellular ATP concentrations [39]. Furthermore, dormant *Mtb* would have an abundant supply of its required carbon and nitrogen source stored in its lipid bodies and α-L-polyGln, respectively as well as iron stored in bacterioferritin. In fact, triacylglycerols present in mycobacterial lipid bodies have been shown to be consumed when bacilli are transferred from hypoxic to normoxic conditions [67]. Thus, acquisition of nutrient depots in the former host empowers dormant bacilli with the ability to colonize after transmission into the new host. These dormant bacilli would slowly adapt and start obtaining host nutrients as its depots are limited, but sufficient for the colonization process. In essence, colonizing dormant bacilli could silently persist without inducing lung damage and be responsible for LTBI seen in the majority of *Mtb* infected individuals. In contrast, after transmission and in order to acquire nutrients, actively growing *Mtb* might rely on the induction of inflammation of the recently colonized lung. A similar scenario is envisioned for *Mtb* with mutations in genes participating in the dormancy regulon. As these mutants would lack a source of endogenous nutrients, inducing inflammation to facilitate nutrient acquisition is a priority. Indeed, faster and more significant inflammation was observed for mice infected with *Mtb* mutants lacking the ability to synthesize α-L-polyGln (Chapter IV). A similar phenotype has been observed for *Mtb* mutants lacking either the two-component system regulating the dormancy operon [42] and for hspX, which is highly expressed in dormancy [68]. Finally, in the context of inducing inflammation to acquire nutrients, an important product that could mediate this event is the extracellular *Mtb* RNA described in Chapter II. The mechanism allowing RNA release is not clear, however, it was observed that a progressive increase in its extracellular concentration paralleled bacterial mass accumulation and the quantity of secreted proteins. A possibility is that RNA is simply being released during *Mtb* replication. It has become evident that an innate immune response can be elicited against *Shigella* and *Legionella* RNA and this has also been correlated to bacterial replication [69, 70]. Thus, RNA released during *Mtb* replication could be an early event modulating the host immune response. It was observed that *Mtb* RNA enhanced bacterial replication inside macrophages. For *Mtb* RNA, a positive

feedback loop can be envisioned in which as the bacilli starts to replicate, the released RNA deactivates the macrophage, in turn leading to more *Mtb* replication. The fact that *Mtb* RNA also induced host cell apoptosis can start the cell death process that eventually leads to necrotic granulomas and further down the line, liquefaction and cavitation.

As a corollary for this discourse and if wanting to mimic the natural history of human TB in animals models, the methodology for *in vitro* culturing infectious stocks of *Mtb* should be revisited. In general, the current protocols involve culturing in media containing glucose or glycerol as opposed to the abundant body of literature pointing out to lipids as the primary carbon source. Even though oleic acid or polysorbate 80 (Tween-80) is routinely supplemented in media, the concentration of glycerol and/or glucose still exceeds that of lipids. As *in vitro Mtb* has been reported to simultaneously metabolize sugars and lipids [21], the inclusion of both carbon sources could have biochemical and physiological consequences in a bacterium that apparently only uses lipids when transferred into the host [34, 71]. Furthermore, cultures are harvested during logarithmic growth when most bacilli are actively growing. To the best of my knowledge, there has only been one attempt to infect animals with dormant bacteria and this was accomplished after intranasal inoculation [72]. Unfortunately and in order to follow Wayne's protocol for inducing *Mtb* dormancy, *in vitro* culturing was performed in Dubos media which has abundant glycerol and glucose concentrations. In addition, even though it has been noticed that *in vitro Mtb* grows better in a carbon dioxide (CO^2)-containing environment which resemble *in vivo* conditions [3], inclusion of CO^2 is seldom, if at all, done when preparing infectious stocks. Finally and to make matters worse, infectious stocks are frozen at -80˚C for prolonged periods of time. How this affects the bacterial's physiology is unknown. To better mimic LTBI in animal models and in descending order of ideal methods for experimental infections, the following could be considered: a) as performed with guinea pigs breathing air from tuberculosis wards, animals could be infected with airborne bacilli being expectorated by patients, b) animals could be infected with bacilli obtained from sputum, c) animals could be infected with bacilli isolated from infected tissues. For obvious reasons this wouldn't be possible in

most cases. Instead *in vitro* culturing is required in most circumstances, so a fourth option could be to infect animals with bacilli grown inside a CO^2 incubator in medium composed of extracts from infected lungs.

In conclusion, I propose that transmission of dormant bacteria could be responsible for LTBI. By having reduced metabolic requirements and carrying its own source of nutrients in the form of depots, dormant as opposed to actively growing *Mtb* would be better equipped to deal with the colonization process of the nutrient-deprived environment encountered in naïve lungs. Actively growing *Mtb* procures its source of nutrients by inducing inflammation but this becomes a two-edged sword as the bacillus could be eradicated by the influx of inflammatory cells. In contrast, dormant *Mtb* could persist for long periods of time without inducing tissue damage (latency), until the presence of more favorable conditions minimizing the risk of being eradicated. The role of dormancy in the physiology and pathogenesis of *Mtb* should therefore be sought during transmission and not chronicity, the first vs. the last step in *Mtb* life cycle, respectively. In agreement with Pollock's quoted statement, clinical observation should be the driving force to understand TB. In this regard, understanding human LTBI and the role played by bacterial dormancy is highly warranted.

Literature Cited

1. Pollock, J.E., *Croonian Lectures on Modern Theories and Treatment of Pthisis.* Lancet, 1883. **121**(3110): p. 583-585.
2. World Health Organization., *Global tuberculosis control : WHO report 2011*. 2011, Geneva: World Health Organization. viii, 246 p.
3. Fitzgerald, D.W., T.R. Sterling, and H. D.W., *Mycobacterium tuberculosis*, in *Mandell, Douglas and Bennett's principles and practice of infectious diseases*. 2010, Churchill Livingstone/Elsevier: Philadelphia, PA. p. 3129-3163.
4. Horsburgh, C.R., Jr. and E.J. Rubin, *Clinical practice. Latent tuberculosis infection in the United States.* N Engl J Med, 2011. **364**(15): p. 1441-8.
5. Comstock, G.W., *How much isoniazid is needed for prevention of tuberculosis among immunocompetent adults?* Int J Tuberc Lung Dis, 1999. **3**(10): p. 847-50.
6. Lin, P.L. and J.L. Flynn, *Understanding latent tuberculosis: a moving target.* J Immunol, 2010. **185**(1): p. 15-22.
7. Capuano, S.V., 3rd, et al., *Experimental Mycobacterium tuberculosis infection of cynomolgus macaques closely resembles the various manifestations of human M. tuberculosis infection.* Infect Immun, 2003. **71**(10): p. 5831-44.
8. McMurray, D.N., *Guinea pig model of tuberculosis*, in *Tuberculosis: pathogenesis, protection, and control*, B.R. Bloom, Editor. 1994, ASM Press: Washington, D.C. p. 135-147.
9. Ordway, D.J. and I.M. Orme, *Animal models of mycobacteria infection.* Curr Protoc Immunol, 2011. **Chapter 19**: p. Unit19 5.
10. Mills, C.C., F. O'Grady, and R.L. Riley, *Tuberculin conversion in the "naturally infected" guinea pig.* Bull Johns Hopkins Hosp, 1960. **106**: p. 36-45.
11. Riley, R.L., et al., *Aerial dissemination of pulmonary tuberculosis. A two-year study of contagion in a tuberculosis ward. 1959.* Am J Epidemiol, 1995. **142**(1): p. 3-14.
12. Dharmadhikari, A.S., et al., *Natural infection of guinea pigs exposed to patients with highly drug-resistant tuberculosis.* Tuberculosis (Edinb), 2011. **91**(4): p. 329-38.
13. Barry, C.E., 3rd, et al., *The spectrum of latent tuberculosis: rethinking the biology and intervention strategies.* Nat Rev Microbiol, 2009. **7**(12): p. 845-55.
14. Burnet, F.M., *The natural history of tuberculosis.* Med J Aust, 1948. **1**(3): p. 57-63.
15. Kumar, V. and A. Maitra, *Robbins basic pathology*, in *Basic pathology*, V. Kumar, R.S. Cotran, and S.L. Robbins, Editors. 2003, Saunders: Philadelphia. p. 484-490.
16. Hernandez-Pando, R., et al., *Persistence of DNA from Mycobacterium tuberculosis in superficially normal lung tissue during latent infection.* Lancet, 2000. **356**(9248): p. 2133-8.
17. Mullerpattan, J.B. and Z.F. Udwadia, *Normal chest radiographs in sputum culture-positive pulmonary tuberculosis.* Int J Tuberc Lung Dis, 2009. **13**(1): p. 148, author reply 148-9.
18. opie el and aronson jd, *Tubercle bacillin in latent tuberculous lesions and in lung tissue without tuberculous lesions.* Archives of Pathology and Laboratory Medicine, 1927. **4**(1): p. 1-21.
19. Kraft, S.L., et al., *Magnetic resonance imaging of pulmonary lesions in guinea pigs infected with Mycobacterium tuberculosis.* Infect Immun, 2004. **72**(10): p. 5963-71.
20. Wayne, L.G. and C.D. Sohaskey, *Nonreplicating persistence of mycobacterium tuberculosis.* Annu Rev Microbiol, 2001. **55**: p. 139-63.
21. Wayne, L.G. and L.G. Hayes, *An in vitro model for sequential study of shiftdown of Mycobacterium tuberculosis through two stages of nonreplicating persistence.* Infect Immun, 1996. **64**(6): p. 2062-9.

22. Gengenbacher, M., et al., *Nutrient-starved, non-replicating Mycobacterium tuberculosis requires respiration, ATP synthase and isocitrate lyase for maintenance of ATP homeostasis and viability.* Microbiology, 2010. **156**(Pt 1): p. 81-7.
23. Voskuil, M.I., et al., *Inhibition of respiration by nitric oxide induces a Mycobacterium tuberculosis dormancy program.* J Exp Med, 2003. **198**(5): p. 705-13.
24. Kumar, A., et al., *Heme oxygenase-1-derived carbon monoxide induces the Mycobacterium tuberculosis dormancy regulon.* J Biol Chem, 2008. **283**(26): p. 18032-9.
25. Deb, C., et al., *A novel in vitro multiple-stress dormancy model for Mycobacterium tuberculosis generates a lipid-loaded, drug-tolerant, dormant pathogen.* PLoS One, 2009. **4**(6): p. e6077.
26. Wayne, L.G. and L.G. Hayes, *Nitrate reduction as a marker for hypoxic shiftdown of Mycobacterium tuberculosis.* Tuber Lung Dis, 1998. **79**(2): p. 127-32.
27. Wayne, L.G. and K.Y. Lin, *Glyoxylate metabolism and adaptation of Mycobacterium tuberculosis to survival under anaerobic conditions.* Infect Immun, 1982. **37**(3): p. 1042-9.
28. 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.
29. Park, H.D., et al., *Rv3133c/dosR is a transcription factor that mediates the hypoxic response of Mycobacterium tuberculosis.* Mol Microbiol, 2003. **48**(3): p. 833-43.
30. Roberts, D.M., et al., *Two sensor kinases contribute to the hypoxic response of Mycobacterium tuberculosis.* J Biol Chem, 2004. **279**(22): p. 23082-7.
31. McKinney, J.D., et al., *Persistence of Mycobacterium tuberculosis in macrophages and mice requires the glyoxylate shunt enzyme isocitrate lyase.* Nature, 2000. **406**(6797): p. 735-8.
32. Sirakova, T.D., et al., *Identification of a diacylglycerol acyltransferase gene involved in accumulation of triacylglycerol in Mycobacterium tuberculosis under stress.* Microbiology, 2006. **152**(Pt 9): p. 2717-25.
33. Deb, C., et al., *A novel lipase belonging to the hormone-sensitive lipase family induced under starvation to utilize stored triacylglycerol in Mycobacterium tuberculosis.* J Biol Chem, 2006. **281**(7): p. 3866-75.
34. Munoz-Elias, E.J. and J.D. McKinney, *Mycobacterium tuberculosis isocitrate lyases 1 and 2 are jointly required for in vivo growth and virulence.* Nat Med, 2005. **11**(6): p. 638-44.
35. Kornberg, H.L. and H. Beevers, *A mechanism of conversion of fat to carbohydrate in castor beans.* Nature, 1957. **180**(4575): p. 35-6.
36. Fenhalls, G., et al., *In situ detection of Mycobacterium tuberculosis transcripts in human lung granulomas reveals differential gene expression in necrotic lesions.* Infect Immun, 2002. **70**(11): p. 6330-8.
37. Daniel, J., et al., *Induction of a novel class of diacylglycerol acyltransferases and triacylglycerol accumulation in Mycobacterium tuberculosis as it goes into a dormancy-like state in culture.* J Bacteriol, 2004. **186**(15): p. 5017-30.
38. Garton, N.J., et al., *Intracellular lipophilic inclusions of mycobacteria in vitro and in sputum.* Microbiology, 2002. **148**(Pt 10): p. 2951-8.
39. Rao, S.P., et al., *The protonmotive force is required for maintaining ATP homeostasis and viability of hypoxic, nonreplicating Mycobacterium tuberculosis.* Proc Natl Acad Sci U S A, 2008. **105**(33): p. 11945-50.
40. Garnier, T., et al., *The complete genome sequence of Mycobacterium bovis.* Proc Natl Acad Sci U S A, 2003. **100**(13): p. 7877-82.
41. Rustad, T.R., et al., *The enduring hypoxic response of Mycobacterium tuberculosis.* PLoS One, 2008. **3**(1): p. e1502.
42. Parish, T., et al., *Deletion of two-component regulatory systems increases the virulence of Mycobacterium tuberculosis.* Infect Immun, 2003. **71**(3): p. 1134-40.
43. Malhotra, V., et al., *Disruption of response regulator gene, devR, leads to attenuation in virulence of Mycobacterium tuberculosis.* FEMS Microbiol Lett, 2004. **231**(2): p. 237-45.

44. Converse, P.J., et al., *Role of the dosR-dosS two-component regulatory system in Mycobacterium tuberculosis virulence in three animal models.* Infect Immun, 2009. **77**(3): p. 1230-7.
45. Munoz-Elias, E.J., et al., *Replication dynamics of Mycobacterium tuberculosis in chronically infected mice.* Infect Immun, 2005. **73**(1): p. 546-51.
46. Gill, W.P., et al., *A replication clock for Mycobacterium tuberculosis.* Nat Med, 2009. **15**(2): p. 211-4.
47. Scanga, C.A., et al., *Reactivation of latent tuberculosis: variations on the Cornell murine model.* Infect Immun, 1999. **67**(9): p. 4531-8.
48. Baek, S.H., A.H. Li, and C.M. Sassetti, *Metabolic regulation of mycobacterial growth and antibiotic sensitivity.* PLoS Biol, 2011. **9**(5): p. e1001065.
49. Dutta, N.K., et al., *Rifapentine is not more active than rifampin against chronic tuberculosis in guinea pigs.* Antimicrob Agents Chemother, 2012. **56**(7): p. 3726-31.
50. Garton, N.J., et al., *Cytological and transcript analyses reveal fat and lazy persister-like bacilli in tuberculous sputum.* PLoS Med, 2008. **5**(4): p. e75.
51. Ryan, G.J., et al., *Multiple M. tuberculosis phenotypes in mouse and guinea pig lung tissue revealed by a dual-staining approach.* PLoS One, 2010. **5**(6): p. e11108.
52. Seiler, P., et al., *Cell-wall alterations as an attribute of Mycobacterium tuberculosis in latent infection.* J Infect Dis, 2003. **188**(9): p. 1326-31.
53. Bloch, H. and W. Segal, *Biochemical differentiation of Mycobacterium tuberculosis grown in vivo and in vitro.* J Bacteriol, 1956. **72**(2): p. 132-41.
54. Vander Heiden, M.G., L.C. Cantley, and C.B. Thompson, *Understanding the Warburg effect: the metabolic requirements of cell proliferation.* Science, 2009. **324**(5930): p. 1029-33.
55. Somashekar, B.S., et al., *Metabolic profiling of lung granuloma in Mycobacterium tuberculosis infected guinea pigs: ex vivo 1H magic angle spinning NMR studies.* J Proteome Res, 2011. **10**(9): p. 4186-95.
56. Kim, M.J., et al., *Caseation of human tuberculosis granulomas correlates with elevated host lipid metabolism.* EMBO Mol Med, 2010. **2**(7): p. 258-74.
57. Russell, D.G., et al., *Foamy macrophages and the progression of the human tuberculosis granuloma.* Nat Immunol, 2009. **10**(9): p. 943-8.
58. Peyron, P., et al., *Foamy macrophages from tuberculous patients' granulomas constitute a nutrient-rich reservoir for M. tuberculosis persistence.* PLoS Pathog, 2008. **4**(11): p. e1000204.
59. Daniel, J., et al., *Mycobacterium tuberculosis uses host triacylglycerol to accumulate lipid droplets and acquires a dormancy-like phenotype in lipid-loaded macrophages.* PLoS Pathog, 2011. **7**(6): p. e1002093.
60. Harth, G., et al., *All four Mycobacterium tuberculosis glnA genes encode glutamine synthetase activities but only GlnA1 is abundantly expressed and essential for bacterial homeostasis.* Mol Microbiol, 2005. **58**(4): p. 1157-72.
61. Tullius, M.V., G. Harth, and M.A. Horwitz, *Glutamine synthetase GlnA1 is essential for growth of Mycobacterium tuberculosis in human THP-1 macrophages and guinea pigs.* Infect Immun, 2003. **71**(7): p. 3927-36.
62. Harth, G. and M.A. Horwitz, *Inhibition of Mycobacterium tuberculosis glutamine synthetase as a novel antibiotic strategy against tuberculosis: demonstration of efficacy in vivo.* Infect Immun, 2003. **71**(1): p. 456-64.
63. Hirschfield, G.R., M. McNeil, and P.J. Brennan, *Peptidoglycan-associated polypeptides of Mycobacterium tuberculosis.* J Bacteriol, 1990. **172**(2): p. 1005-13.
64. Harth, G., et al., *Treatment of Mycobacterium tuberculosis with antisense oligonucleotides to glutamine synthetase mRNA inhibits glutamine synthetase activity, formation of the poly-L-glutamate/glutamine cell wall structure, and bacterial replication.* Proc Natl Acad Sci U S A, 2000. **97**(1): p. 418-23.

65. Candela, T. and A. Fouet, *Bacillus anthracis CapD, belonging to the gamma-glutamyltranspeptidase family, is required for the covalent anchoring of capsule to peptidoglycan.* Mol Microbiol, 2005. **57**(3): p. 717-26.
66. Basaraba, R.J., et al., *Increased expression of host iron-binding proteins precedes iron accumulation and calcification of primary lung lesions in experimental tuberculosis in the guinea pig.* Tuberculosis (Edinb), 2008. **88**(1): p. 69-79.
67. Low, K.L., et al., *Triacylglycerol utilization is required for regrowth of in vitro hypoxic nonreplicating Mycobacterium bovis bacillus Calmette-Guerin.* J Bacteriol, 2009. **191**(16): p. 5037-43.
68. Hu, Y., et al., *Deletion of the Mycobacterium tuberculosis alpha-crystallin-like hspX gene causes increased bacterial growth in vivo.* Infect Immun, 2006. **74**(2): p. 861-8.
69. Jehl, S.P., et al., *IFNgamma inhibits the cytosolic replication of Shigella flexneri via the cytoplasmic RNA sensor RIG-I.* PLoS Pathog, 2012. **8**(8): p. e1002809.
70. Monroe, K.M., S.M. McWhirter, and R.E. Vance, *Identification of host cytosolic sensors and bacterial factors regulating the type I interferon response to Legionella pneumophila.* PLoS Pathog, 2009. **5**(11): p. e1000665.
71. Pethe, K., et al., *A chemical genetic screen in Mycobacterium tuberculosis identifies carbon-source-dependent growth inhibitors devoid of in vivo efficacy.* Nat Commun, 2010. **1**: p. 57.
72. Woolhiser, L., et al., *In vivo adaptation of the Wayne model of latent tuberculosis.* Infect Immun, 2007. **75**(5): p. 2621-5.

List of Abbreviations

ABC	ATP Binding cassette
Ala	Alanine
ATP	Adenosine-5'-triphosphate
β2M	β-2 microglobulin
BCG	Bacillus Calmette-Guerin
B. anthraci s	*Bacillus anthracis*
BCA	Bicinchoninic acid
BLAST	Basic Local Alignment Search Tool
BSA	Bovine serum albumin
Bp	Base pair
CF	Culture filtrate
CF-Man	Culture filtrate depleted of ManLAM
CFU	Colony forming unit
ConA	ConcanavalinA
CR	Complement receptor
Cys	Cysteine
DAP	Diaminopimelic acid
diGln	dipeptide of glutamine residues
DON	6-Diazo-5-OxoNorleucine
DOTS	Directly observed treatment short
DNA	Deoxyribonucleic acid
dsRNA	double stranded RNA
DTH	Delayed Type Hypersensitivity
DTT	Dithiothreitol
ETZ	Electron Transparent Zone
FcR	Fc receptor
F. tularensis	*Francisella tularensis*
GAS	Glycerol alanine salts
g DNA	genomic DNA
GGT	γ-glutamyl transpeptidase
GKO	γ-Interferon knock out mice
GlcNAC	N-acetyl glucosamine
Gln	Glutamine
Glu	Glutamate
Glu-Glu	dipeptide of glutamate residues
gpRNA	gel purified RNA
GSH	Glutathione
Gly	Glycine
GlyGly	GlycylGlycine
H	hour
HbHA	Heparin-binding Heamagluttinin
His	Histidine
HLA	Human Leukocyte Antigen
IFN	Interferon
IFN-γ	Interferon gamma
Icl	Isocitrate lyase

Ig	Immunoglobulin
IL	Interleukin
INH	Isoniazid
Kb	kilobase
kDa	kilodalton
kg	kilogram
KO	knock out
LAM	Lipoarabinomannan
L-Glu-γ-ρNA	L-γ-glutamyl-ρ-nitroanilide
LM	Lipomannan
LB	Luria-Bertani
Ldts	L,D-transpeptidase
LC	Liquid chromatography
LTBI	Latent tuberculosis infection
mAGP	mycolyl-arabinogalactan-peptidoglycan
ManLAM	Mannosylated lipoarabinomannan
MAMEs	Mycolic acid methyl esthers
MDR	Multi drug resistance
mg	Miligram
MHC	Major Histocompatibility Complex
min	minute
ml	milliliter
mm	millimeter
mM	Millimolar
MMR	Mannose macrophage receptor
MS	Mass spectrometry
M. bovis	*Mycobacterium bovis*
M. leprae	*Mycobacterium leprae*
M. smeg	*Mycobacterium smegmatis*
MRI	Magnetic resonance imaging
Mtb	*Mycobacterium tuberculosis*
MurNAc	N-acetyl muramic acid
NCBI	National Center for Biotechnology Information
NLR	Nod-like receptors
nm	Nanometer
nM	Nanomol
NMR	Nuclear magnetic resonance
NOS	Nitric oxice synthase
Ntn	N-terminal nucleophile
OADC	Oleic acid, albumin, dextrose, catalase
OD	Optical density
ORF	Open reading frame
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffers saline
PCR	Polymerase chain reaction
PDIM	Phthiocerol dimycocerosate
pentaGln	pentapeptide of glutamine residues
PI3K	Phosphatidyl-inositol 3 kinase
PI3P	Phosphatidyl-inositol 3 phosphate
PIMs	Phosphatidyl-inositol mannosides

polyGlu	oligopeptides of glutamate residues
polyGln	oligopeptides of glutamine residues
PPi	Pyrophosphate
PPD	Purified Protein Derivative
PRPP	Phosphoribosyl pyrophosphate
PRRs	Pattern recognition receptors
PS	Phosphatidylserine
Pyroglu	Pyroglutamic acid
RAG	Recombinase activation gene
RD1	Region of Difference 1
Ref	Reference
ROS	Reactive oxygen species
RNA	Ribonucleic acid
RNIs	Reactive oxygen intermediates
RR	arginine arginine
SDS	Sodium dodecyl sulphate
Sec	Secretion system
Ser	Serine
SL	Sulfolipid
Sp I	Signal peptidase I
Sp II	Signal peptidase II
Spp	Species
Tat	Twin-arginine translocation
TB	Tuberculosis
TBS	Tris buffer saline
T cell	T lymphocyte
TCR	T cell receptor
TDM	Trehalose dimycolate
Th	T helper cell
Thr	Threonine
TLR	Toll-like receptor
TNF-α	Tumor necrosis factor alpha
TNFR	Tumor necrosis factor receptor
Tx	Triton X
µg	Microgram
µl	Microliter
µM	Micromolar
WHO	World Health Organization