DISSERTATION

MINING THE *MYCOBACTERIUM TUBERCULOSIS* CELLULAR ENVELOPE FOR DIAGNOSTIC AND DRUG TARGETS

Submitted by

Lisa Marie Wolfe

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: Karen M. Dobos

John T. Belisle
Steven Dow
David Lewinsohn
Jessica Prenni

ABSTRACT

MINING THE *MYCOBACTERIUM TUBERCULOSIS* CELLULAR ENVELOPE FOR DIAGNOSTIC AND DRUG TARGETS

The cellular envelope of *Mycobacterium tuberculosis* is a highly complex structure containing many lipids, carbohydrates and proteins. Together these components maintain cellular homeostasis and play an active role in the establishment and maintenance of intracellular infection. The World Health Organization estimates that upon exposure, 90% of infected persons will not succumb to active disease and likely remain a potential source of *M. tuberculosis* transmission upon the emergence of clinical symptoms. Throughout this complex disease course, physiological and structural changes occur within the cell envelope of the bacillus and facilitate its survival within an infected host. These physiological changes also influence the immunological interplay between bacteria and host cell, further contributing to the success of the pathogen. To understand these changes we must have a comprehensive knowledge of the cell wall proteins that contribute to the overall makeup of the mycobacterial envelope, understand the unique antigens that reside in or that are secreted from the bacillus and monitor these proteins under different physiological disease profiles. In the last decade many large-scale descriptive studies focusing on gene transcript profiles and proteomic composition of subcellular fractions have pioneered efforts in the understanding of *M. tuberculosis* physiology, immunology and pathogenesis at the level of systems biology. These studies have shown that *M. tuberculsosis* is capable of retaining essential gene products for respiration, nutrient uptake and energy metabolism both *in vitro* and *in vivo* and that the proteins within the cell wall are highly immunogenic for *M.tuberculosis*-specific T lymphocytes. Our first objective established a comprehensive description of the cell wall proteome of *M. tuberculosis* using traditional two-

dimensional gel-based techniques and liquid-chromatogoraphy mass spectrometry (LC-MS). From this work, over 500 proteins were identified using a combination of differential detergent extraction and multi-dimensional-LC. A highly lipoprotein enriched fraction revealed that the majority of cell wall associated proteins were functionally annotated to mechanisms of intermediary metabolism (35%) and macromolecular synthesis and degradation (25%) building upon evidence that the *M. tuberculosis* cell wall is actively engaged in cellular homeostasis and remodeling events. Secondly, we investigated the role of the cell envelope proteins in the search for novel immunodiagnostic epitopes. It is well known that the cell wall of *M. tuberculosis* is highly immunogenic and contains both non-protein and protein antigens. Specifically, the proteins associated with the cell wall were shown to be uniquely responsible for the activation of human CD8 T cell clones generated from both actively and latently infected individuals. The immunological response to CD8 T cell antigens may be an effective means of distinguishing between latent TB infection (LTBI) and active disease. To broadly define the repertoire of CD8 T cell antigens, 56 proteins from the cell wall proteome study were included in the design of a synthetic peptide library. Exhaustive screening of the peptide library for novel antigens and epitopes that elicit an immunological response in TB patients, resulted in the identification of eight cell wall antigens that are currently being investigated for their clinical utility. In addition, the cell wall proteome was also mined in the identification of an HLA-E restricted CD8 T cell epitope. HLA-E has low polymorphism in the human population and seems to be enriched in *M. tuberculosis* - containing phagosomes, therefore identification of this antigen could be used as a novel diagnostic or vaccine candidate. Using a MS-based proteomics approach, we discovered the HLA-E antigen to be the post-translationally modified glycoprotein Mpt32 (45kDa/Apa). Glycosylated proteins and lipids within the mycobacterial cell envelope are dominant and the role of this modification in the host immune response can now be elucidated. Lastly, the

composition and integrity of the *M. tuberculosis* cell envelope facilitates its adaptation and survival within various microenvironments. These physiological functions are influenced by the presence or absence of functionally linked genes and proteins whose relative abundance may change over time or within altered metabolic states. Our final efforts used nucleotide analog probes, to specifically bind and enrich proteins with an ATP-binding function and measure their relative abundance between altered states of growth (i.e. between active disease and hypoxia-induced dormancy). With these efforts we classified 122 ATP-binding proteins in either metabolic state and demonstrated differential abundance patterns between actively growing and hypoxic cells within the functionally linked protein networks of energy metabolism, cell wall and lipid biosynthesis. These protein families represented in the *M. tuberculosis* ATPome are a subset of essential (60% of the *Mtb*-ATPome) gene products and may be relevant therapeutic targets for the future development of novel small molecule inhibitors against *M. tuberculosis*. The spectrum of studies undertaken to mine the cellular envelope for diagnostic and drug targets demonstrates a natural evolution of MS-based proteomics in the study of biologically relevant questions. From a purely descriptive characterization of the cell wall proteome, this data was utilized in a practical approach in the design of a high-throughput antigen/epitope-screening library and finally these studies culminate in a functionally relevant profile of the ATP-binding proteins of *M. tuberculosis*. Future work will continue to focus on developing hypothesis-driven proteomic studies for the identification of novel diagnostic antigens and drug targets.

ACKNOWLEDGEMENTS

I give my deepest gratitude to my advisor, Dr. Karen Dobos for her genuine support and heartfelt encouragement throughout my professional career and graduate education. Words cannot express the impact Dr. Dobos has had in my growth as a student, young scientist, wife and mother. I am grateful to not only call her my graduate education mentor, but one of my most valued friends and confidantes. Through her actions she has taught me to be an independent thinker, a competent writer and scientific collaborator. I acknowledge her sacrifice, dedication and perseverance, not only to be an impactful researcher, but also her efforts as a true counselor and teacher to all she advises.

I would also like to thank my graduate committee members, Dr. John Belisle, Dr. Steven Dow, Dr. David Lewinsohn and Dr. Jessica Prenni, for their enthusiastic support and valuable insight into my research. I especially thank Dr. Lewinsohn for his years as a trusted collaborator and Dr. John Belisle for his years of valued mentorship.

To my colleagues, I thank you all for your support and friendship. Without the comfort of laughter making it through this journey would not have been so easy.

I also give special thanks to my family - my mother Yolanda Wolfe for always providing me with guidance, love and faith. My father, Jeffery Wolfe whom I credit for my love of science and curiosity – may your spirit in Heaven always lead me forward. To my sister Dr. Amanda Wolfe for always letting me do the hard work while you "supervised". I am so proud of the woman you have become. Finally, I thank my husband and children for their understanding and patience in pursuing my professional goals.

DEDICATION

To my children:

Maia Danielle

Ezekiel Adam

and

Ana Marie

May you forever believe in yourself, make no excuses, and never give up.

Love you always,

Mom

TABLE OF CONTENTS

I. CHAPTER I: Tuberculosis, Physiology of the Bacillus and the Host Immune Response.

1.1 Tuberculosis and Current World Problems

Estimates from the World Health Organization (WHO) reported 8.5 to 9.2 million new cases of tuberculosis (TB) in 2010, with an overall global disease burden of over 2 billion people (1). *Mycobacterium tuberculosis (M. tuberculosis)* is transmitted into the lungs via close-contact aerosol and remains one of the leading causes of death worldwide (1.4 million deaths in 2010) despite extensive research, directly observed therapy short course (DOTS) initiatives and the widespread use of a vaccine (2). The majority of infected persons harbor the bacterium in a persistent state of metabolic dormancy, only 1 in 10 persons infected present with clinical symptoms. However, the frequency and scale of latent infections presents a major risk for the future re-emergence of active, transmittable disease. The burden of TB disease is monitored in over 204 countries and territories which encompasses ~ 99% of the world's population (3). The STOP TB initiative organized by the WHO was initiated in the early 1990's to control the resurgence of disease cases caused by *Mtb* infection. Their efforts included the rapid standardization and expansion of diagnostic methods and treatment regimens. The campaign objectives are currently stated as follows (www.stoptb.org):

- Achieve universal access to quality diagnosis and patient-centered treatment
- Reduce the human and socioeconomic burden associated with tuberculosis
- Protect vulnerable populations from TB, TB-HIV co-infection and drug-resistant TB
- Support development of new methods and enable their timely and effective use
- Protect and promote human rights in TB prevention, care and control

Because of these efforts, it is estimated that more than 36 million people were cured between 1995 and 2008, saving 6 million lives. However, worldwide the annual case rate of ~ 9

million people still remains and incidence of disease is falling at less than 1% peryear (2,3). Global initiatives for TB elimination have been set by the WHO and include a halt to the disease epidemic by 2015, with an overall goal of worldwide elimination of disease by 2050. While progress has been made, the emergence of multi-drug resistant (MDR) and extensively-drug resistant (XDR) tuberculosis presents a severe problem, as does the continuing epidemic of HIV in high burden countries (4).

TB is the leading cause of death among people living with HIV and is responsible for one-quarter of all HIV-related deaths, especially in sub-Saharan Africa where it estimated that 80% of the TB cases from African nations are among people co-infected with the virus (1). In addition to HIV-TB co-infection, diabetes mellitus (DM) is now recognized as a major impedance to the successful treatment of TB infection. Persons with DM have a significantly increased risk (2-3X higher) of active TB. In countries such as China and India the prevalence of DM has increased dramatically in recent years and the implementation of dual DM-TB screening efforts among local clinics has helped to identify cases of dual burden (5-7).

In order to fully realize the early and long-term goals set by the Stop TB global initiative the annual diagnoses of new TB cases must be dramatically improved from the estimated new case rate of 1.8 million per year. To achieve this goal there must be: (1) The development of rapid, accurate and inexpensive diagnostic tools to better identify cases of subclinical latent infection; (2) The design of new drugs and identification of new drug targets in order to circumvent problems of multi-drug resistance, and (3) Scientists and medical doctors need to understand the differences in virulence among clinically relevant strains and characterize the host factors that are responsible for determining the ultimate fate of disease outcome (i.e. progress into active or latent disease) (8).

1.1.2 Diagnosing Tuberculosis and the Need for Improved Diagnostic Tools

TB is most universally diagnosed by direct microscopy of patient sputa or through the positive culturing of *M. tuberculosis* in clinical settings (9). In high burden countries direct smear examinations of non-concentrated sputum samples remains the "gold standard" necessary for the diagnosis of tuberculosis in most patients. Smear examination is subjected to acid-fast staining procedures such as the Ziehl-Neelsen stain or the auramine O fluorescence acid-fast stain (10). Staining in such a manner is dependent on the mycobacteria (and related actinomycetes) to retain the red staining dye after treatment with a mild acid-alcohol solution. The retention of the dye is due to the unique physiochemical properties of the mycobacterial cell envelope (i.e. the hydrophobic nature of the mycolic acid core wall). Despite extensive effort by the WHO to improve the technical capabilities of clinical diagnosticians in endemic regions, the sensitivity of identification of *M. tuberculosis* in sputa is low (11). Less than 50% of cases deemed positive by culture are also positive by smear microscopy and detection rates in patients co-infected with HIV is down to approximately 30% (10). However, culturing the bacteria from sputum smears is more costly and requires specialized biosafety infrastructure and takes 6-8 weeks for positive colony growth to occur (12). Immunological-based diagnosis of TB begins with the tuberculin skin test (TST) and the measurement of delayed-type hypersensitivity reaction to soluble purified-protein derivative (PPD). More sophisticated is the measurement of T cell responses to select antigen panels of *Mtb,* using IFN-γ release assays (IGRA's) (13). IGRA's have an advantage over the TST, in that these assays can distinguish between those infected with *Mtb* from those vaccinated by *M. bovis*, BCG, however neither the TST or IGRA methods distinguish between active disease and latent TB infection.

Biomarker discovery studies for *M. tuberculosis* are needed in order to develop second-generation diagnostic assays in which the ability to distinguish between active and latent disease

is possible. Prognostically, *M. tuberculosis* biomarkers could help to determine if cases of latent infection would progress to active disease, and monitor the successes of trial drugs (12). In this context, efforts must be made at the systems level (i.e. transcriptome, proteome and metabolome) to understand these important distinctions among a variety of disease states and clinically relevant settings.

1.1.3 Drug Resistant Tuberculosis and the Search for Novel Drug Targets.

The emergence of drug-resistant tuberculosis was first observed shortly after the clinical use of streptomycin in the 1940's and 1950's (14,15). Ultimately the use of antibiotic monotherapy was replaced with a standard regimen of three to four drugs all working against distinct protein targets in the inhibition of cell wall biosynthesis (Isoniazid (INH), Ethambutol (EMB)) and protein translation (Rifampin (RIF), Pyrazinamide (PZA))(16). Currently the standard antibiotic regimen for active pulmonary TB is comprised of three drugs – INH, RIF and PZA. For susceptible *M. tuberculosis* strains, this regimen is prescribed for duration of 6 months. Multi-drug resistant TB (MDR-TB) accounts for approximately 4-6% of new TB cases worldwide (1). MDR-TB is defined as being resistant to both first-line drugs (i.e. INH/RIF). Treatment of resistant cases of infection requires the use of second line drugs for 18-24 months. This is a costly endeavor and is difficult to attain in resource-poor countries where MDR-TB may be commonplace (17). Drug resistant bacilli result from mutations of the genetic code. These changes result in the alteration of key amino acids within the primary sequences of target enzymes, negating the intracellular effectiveness of each drug. For INH resistance, mutations in two genes, *inhA* and *katG,* most dominantly define the drug-resistant phenotype (18-20). The primary target enzyme for isoniazid is the enoyl acyl reductase InhA (21). However INH in its pro-drug form must undergo activation by the catalase-peroxidase KatG prior to its binding

InhA. Mutations in either of these genes prevent the drug from inhibiting the synthesis of mycolic acid, a key component in the cellular envelope. For RIF, the cellular target is the RNA polymerase RpoB and mutations in the RpoB gene overwhelmingly represent RIF resistant clinical cases and are commonly associated with MDR-TB (22). In addition to the increasing prevalence of MDR/XDR cases caused by inheritable genetic resistance mechanisms, drug-tolerant bacilli confer resistance to drug treatment due to changes in their physiological and metabolic state (23) (Figure 1.1). The ability of *M. tuberculosis* to enter into a "phenotypically drug-resistant", non-replicating dormant lifestyle is clinically defined as Latent TB infection (LTBI) (24). Treatment of such infections with standard drug regimens is problematic and requires a prolonged treatment duration (23,25). Identification of new drugs, with novel modes of action that are active against *M. tuberculosis* during varying physiological stages of infection are needed to curtail the continuing emergence of drug resistant TB.

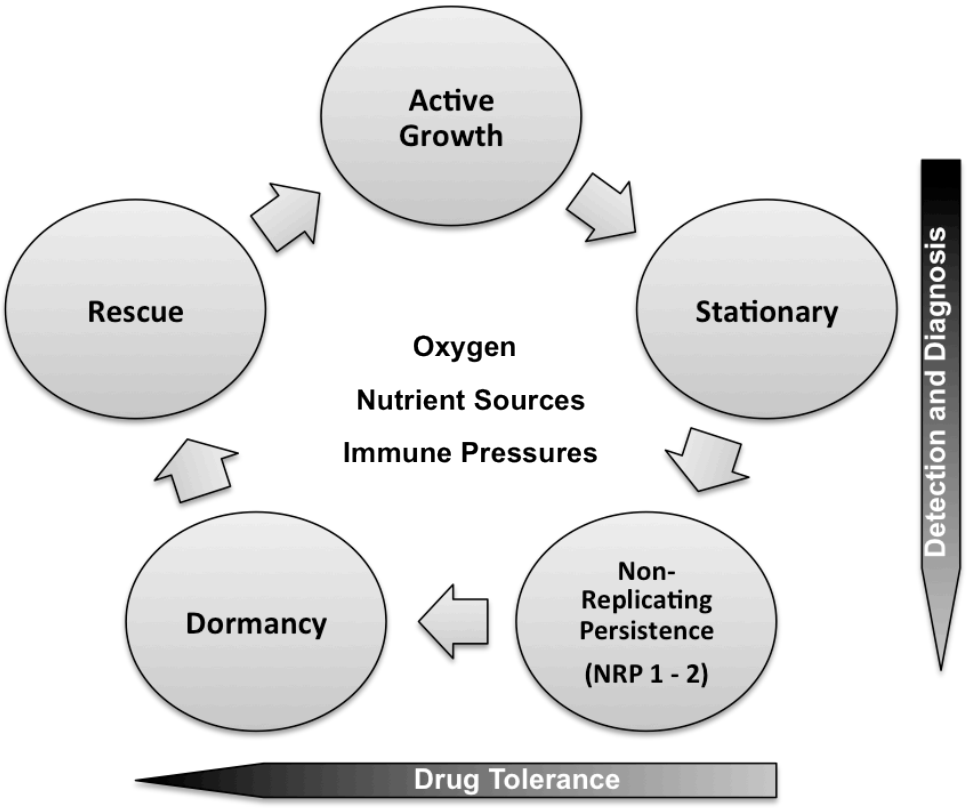

Figure 1.1 – Physiological Metabolic States of *Mycobacterium tuberculosis.*
In response to micro-environmental stressors, *M. tuberculosis* is capable of surviving in a multitude of physiological states. Survival throughout these stages depends on changes within the macromolecular architecture of the cellular envelope, which diminish our capacity to detect the pathogen and diagnose infection. Drug tolerance or resistance is characteristic of bacteria existing in a state of non-replicating persistence or dormancy and new drugs are needed to better target this form of clinically latent disease.

1.2 History of Tuberculosis

The etiologic agent of tuberculosis was first characterized in 1882 by Dr. Robert Koch however its history as a communicable disease of man and animal dates back thousands of years. The oldest proven case of mycobacterial disease was discovered in a 17,000-year-old bison in a North American cave (26) and biomarkers of *M. tuberculosis*-specific lipids remain detectable in this ancient specimen (27). In humans, evidence of tubercle lesions was first described in the 9,000-year old skeletal remains of a woman and child within present day Israel (28). Evidence of *Mtb* infection has also been found in Egyptian and Peruvian mummies (29,30).

Sociological and anthropological research has explored the impact Tuberculosis has had throughout history within social contexts. Analyses of ancient writings provide detailed descriptions of persons afflicted with pulmonary TB (31). The disease, believed to be caused by poverty, uncleanliness or immorality, had an overall negative social stigma. Evidence in ancient written records and death records seem to have downplayed the existence of TB disease (32). In the 17^{th}, 18^{th} and 19^{th} centuries, new understandings and major social reforms, such as improved sanitation and housing and better nutrition were undertaken to reduce the prevalence of disease (33). Late in the 1800's TB patients were isolated and treated in sanitariums in an active attempt to control disease outbreak (34).

It wasn't until the 1940's that scientists discovered that streptomycin was an effective treatment for *Mtb* infection (15) however treatment with streptomycin alone led to antimicrobial resistant bacilli. Isoniazid was introduced in the mid-1950's a miracle drug against *Mtb* and rifampin a few years later (35). By the next decade, it was recognized that a combination of drugs was the most effective treatment for TB (14,36) and cases began to substantially decrease until the 1980's and cases of multi-drug resistance began to emerge. Multi-drug resistant strains

are resistant to both rifampin and isoniazid and in 1993, the World Health Organization declared tuberculosis a global emergency (31).

1.3 Physiology of *M. tuberculosis* - The Cellular Envelope

M. tuberculosis is a rod shaped bacillus measuring 2 - 4 μm in length and has a doubling time of 18-24 hours. The cell envelope of *M. tuberculosis* is comprised of an inner plasma membrane, and a cell wall core built of peptidoglycan, arabinogalactan and mycolic acids. Surrounding this core is a capsule-like outer structure of non-covalently linked glycans, lipids and proteins (37) (Figure 1.2). The chemical structure of the cell wall core and its constituents has been extensively studied and reviewed (38-44). Professor David Minnikin first proposed a seminal model of the structural composition of the mycobacterial cell envelope in 1982. He described a lipid bilayer - type architecture with an inner leaflet of mycolic acids attached to arabinogalactan and an outer leaflet of extractable phospholipids, glycolipids, peptidolipids and mycosides (45,46). Full elucidation of the mycobacterial structural core demonstrated more sophisticated model of interconnected networks between covalently attached and non-covalently linked macromolecules. In this depiction, the extractable lipids are intercalated into both leaflets, along with the extending carbon chains of the mycolic acids (47). The mycolic acids (MA) are long chain (C_{60}-C_{90}) α-branched, β-hydroxylated fatty acids (45,46) and comprise a core structure covalently attached to an inner complex of arabinogalactan and peptidoglycan (mAGP). This material accounts for ~ 60% of the dry weight of the bacillus. The highly cross-linked peptidoglycan (PG) is comprised of repeating units of N-acetyl glucosamine (GlcNAc) and N-acetylated muramic acid (MurNAc) or N-acyl glycolic acid (MurNGlyc) (43,48).

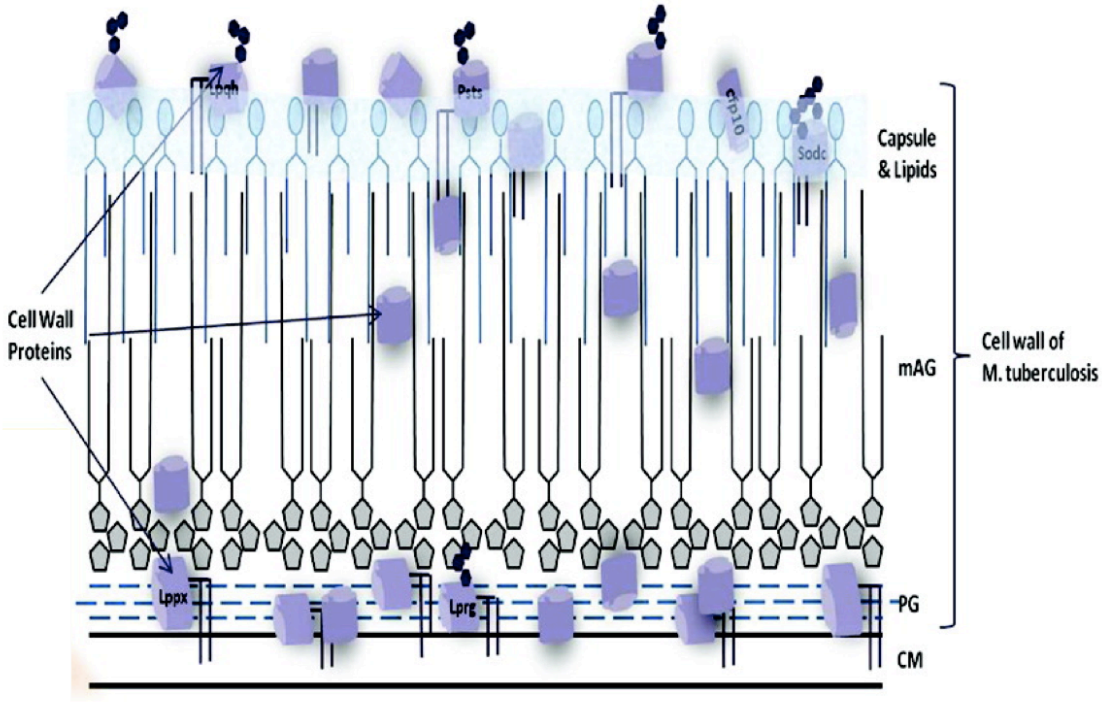

Figure 1.2 – Schematic Representation of the Major Components of the *Mycobacterium tuberculosis* cellular envelope. CM – Cellular (plasma) membrane; PG – Peptidoglycan; mAG – mycolyl-arabinogalactan. The outer capsule comprises non-covalently attached glycans, lipids and proteins.

The penta-peptide side chains of PG are comprised of L-alanyl-D-isoglutaminylmeso-diaminopimelyl-D-alanine (L-Ala-D-Glu-meso-DAP-D-Ala-D-Ala) (39). The degree of cross-linking in mycobacterial PG is estimated to be 2.5 times greater than in that seen in *E. coli* (70-80% versus 30-50-%, respectively) (38). Differential cross-linking patterns have been observed to occur between actively growing and dormant state bacteria potentially allowing the bacterium to survive through a variety of physiological and metabolic phenotypes (49). The L, D-traspeptidases, LdtA and LdtB catalyze this differential cross-linking. The PG layer is covalently attached to the arabinogalactan (AG) polysaccharide via the C_6 of some of the MurNAc or MurNGly residues (50). Structurally the AG is composed of elemental sugars of arabinose and galactose (51,52). The linker unit is characterized as a two-sugar phosphate bridge of L-Rha*(p)*

and D-GlcNAc (41). Like the structure of PG, the molecular features of AG are well defined (41). In addition to their utilization in the MA-core of the *M. tuberculosis* cell wall, additional mycolates are also key components of the virulence associated transport molecules trehalose monomycolate (TMM) and treholse dimycolate (TDM) (53,54). Additionally, several lipid and glycolipid components decorate the cell wall and contribute uniquely to the pathogenesis and virulence of the bacillus. These cell envelope associated virulence factors are summarized in Table 1.1.

Table 1.1 Cell Envelope Associated Virulence Factors

Molecule	Function	Virulence Factor
Mannose-capped lipoarabinomannan (manLAM)	Immunomodulation Outer leaflet of cell envelope	Inhibits phagosome maturation, apoptosis and IFN-γ signaling (55)
Lipomannan (LM)	Immunomodulation Outer leaflet of cell wall	Proinflammatory, TLR2 agonist Precursor to LAM (55)
Sulfolipid 1(SL-1)	Sulfated acyl trehalose (56) Immunomodulation	Modulates oxidative responses and cytokine secretion (57) Inhibits granuloma formation and TNF-α (58)
Phenolic glycolipid (PGL-Tb)	Methyl branched polyketide	Blocks release of TNF-α, IL-6 and IL-12 (59)
Threhalose dimycolate mannoside (TDM)	Neutral glycolipid Immunomodulation	Induces cytokines, granuloma formation (60) Toxic

1.4 The *M. tuberculosis* Cell Envelope in the Context of the Host

The unique composition of the mycobacterial cell wall allows the bacterium to adapt to its intracellular environment and promote its survival and pathogenicity within an infected host cell (42). The dense hydrophobic outer lipid mycolate layer and the capsule control the permeability of the bacterial cell to small hydrophilic molecules. The ability of such molecules to be shuttled both into and out of the cell remains poorly understood. Generally speaking however, pore-

forming proteins, hydrophobic diffusion and lipo-peptide/protein mediated transport are the likely mechanisms contributing to cell wall permeability (61). Like those of gram-negative bacteria, pore-forming proteins allow the import and export of nutrients and other macromolecules. The pore-forming outer membrane protein OmpA is considered a virulence factor that mediates survival within host phagosomes (62). In *Corynebacterium,* two-component channels within the cell envelope have been described (63). The lipoproteins of *M.tuberculosis* are highly enriched within the cell wall core (64), however their various cellular functions remain largely unknown. Several are strong immunogens (19kDa (lpqH) (65,66), 38kD (PstS1) (67) and are potent stimulators of IL-12 production in human macrophages. This induction is mediated by the pattern recognition receptor Toll-Like Receptor - 2 (TLR-2) (68). Specific functions for only a few proteins have been described. For example, the 38kDa lipoprotein is a phosphate-binding protein implicated in the active transport of nutrients (69), and LpqM has recently been described as a metalloproteinase necessary to mediate conjugal DNA transfer (70). Probably one of the most well characterized lipoproteins is LprG. In addition to its TLR2 agonist activity, LprG was shown to bind the triacylated glycolipids lipoarabinomannan (LAM), lipomannan (LM) and phosphatidylinositol mannoside (PIM). The binding of glycolipids by a cell envelope associated lipoprotein may be indicative of their functional roles in glycolipid assembly and transport (71). In addition, LAM, LM and PIM, as described in Table 1.1 play key roles in other immunoregulatory processes. For example, LAM from virulent *M. tuberculosis*, but not from the avirulent *M. smegmatis* is capable of blocking phagosome maturation, apoptosis and the secretion of inflammatory cytokines such as IL-12 and IFN-γ (55). This phenomenon is attributed to differences in the modifications of the molecule's terminal arabinose residues. LAM

isolated from *M. smegmatis* is pro-inflammatory due to the addition of a phospho-myo-inositol on the terminal arabinose sugar moiety (55).

1.5 *M. tuberculosis* Infection

M. tuberculosis is usually transmitted via close-contact aerosol (i.e. cough) originating from a person with active disease. Primary tuberculosis commonly occurs within twelve months of initial exposure, however the majority (90%) of infected persons effectively contain the bacteria without the onset of clinical symptoms (termed latent tuberculosis infection, LTBI). This immunological phenomenon is not well understood, however many factors are thought to contribute to the variance seen in the onset of clinical disease. Some of these factors are likely due to the virulence of the inhaled organism, the number of repeated exposures and the immunocompetence of the infected individual (72).

Once transmitted, *M. tuberculsosis* takes up residence within alveolar macrophages and dendritic cells. Inhalation of as little as 10 bacilli is enough to initiate infection (http://www.cdc.gov/tb/topic/laboratory/BiosafetyGuidance_xdrtb.htm). Macrophages and dendritic cells resident in the airway first ingest the bacterium and initiate the production of pro-inflammatory cytokines such as interleukin (IL)-12, IL-1β and TNF-α (73). IL-12 is critical for Th1 cell mediated immunity (74) and TNF activates macrophages to initiate killing and apoptosis and to secrete cytokines such as IFN-γ (72). Anti-inflammatory cytokines such as IL-10 and TGF-β, along with the immunosuppressant functions of T_{reg} cells can benefit the host by limiting inflammation and minimizing tissue damage in response to infection. However these functions also contribute to a less efficient bactericidal immune reaction (75). Infected host cells in the lungs and airway epithelium migrate to the thoracic lymph nodes to present antigen to CD4 and CD8 T cells (76). However, in the case of *M. tuberculosis* infected cells this priming

function is delayed for many days. Classic works by Wallgren and Poulsen clinically observed that the cellular immune response, as determined by TST skin test, could not be measured for at least 42 days in patients exposed and infected with *M. tuberculosis* (77,78). In mice, it has been shown that the migration of infected DCs and the appearance of antigen specific CD4 T cells does not occur before 11-14 days post-infection (79,80). It is unclear why the migration of infected cells to the draining lymph nodes is delayed for such a significant period of time, however it may have something to do with inhibition of migratory receptors such as the CC-chemokine receptor 7 as well as the apparent ability of *M. tuberculosis* to subvert apoptotic events initiated by neutrophils (81,82). Despite this delay in adaptive immune response, primed CD4 and CD8 T cells are able migrate back to the site of infection along with activated macrophages and B cells, initiating the formation of a granuloma (83). In addition, other cells such as monocytes, neutrophils and fibroblasts contribute to final architecture of the granulomatous lesion (84). Interesting studies using the zebrafish model of *M. marinum*, have demonstrated that early events in granuloma formation (i.e. the recruitment of uninfected cells) actually result in the subsequent infection of newly recruited cells (85). It is at this point the lifecycle of the bacterium and the immunological response to infection is thought to reach equilibrium (82). The actively replicating bacilli adopt alternative metabolic lifestyles (Figure 1.1) (23) and gain the ability to survive within the protected microenvironment of the granuloma.

1.5.1 Mycobacterial Antigens and the Host Immune Response

M. tuberculosis is an intracellular pathogen. A host cell has many defensive strategies in which invading bacteria (or viruses) must subvert in order to successfully propagate infection. *Mycobacteria spp*. have co-evolved with surrogate hosts for millennia and have developed exquisite evasion mechanisms designed to ensure immune modulation and survival. As with all

microbial species, the stimulation of the immune response is primarily initiated by the interaction between the cell envelope surface and host cell pattern recognition receptors. As mentioned in the previous section, the first demonstration of an antigen-induced immune response to mycobacterial species was the development of tuberculin by Robert Koch in 1890 (86). Tuberculin was a crude, non-specific preparation of mycobacterial products released into the culture medium upon heat inactivation of *M. tuberculosis*. Fatal reactions in "vaccinated" tuberculosis patients demonstrated the potent-immunoreactivity of the tuberculin preparation. Skin-reactivity, in the form of reddened induration was observed at the site of tuberculin injection in people previously known to have TB (87). The diagnostic potential of this preparation was further refined two decades later by Florence Seibert who determined the tuberculosis specific response (termed delayed-type hypersensitivity) could be assigned to a protein-enriched fraction of tuberculin (88). The proteomic composition of modern preparations of tuberculin (i.e. purified protein derivative, PPD) and relevant perspectives of immune-diagnostics of TB has been recently published (89,90).

Vaccination with BCG protects individuals throughout childhood and into adolescence. However, the WHO reports millions of newly infected persons each year. Childhood vaccination against *M. tuberculosis* is common practice among TB endemic regions worldwide. The vaccine is a live-attenuated strain of *Mycobacterium bovis*, BCG (1). It first came into use during the 1920's but didn't become routine practice until after World War II (91). BCG vaccination in children is critical for the protection against infection with milliary tuberculosis and tuberculosis meningitis (92). However, immunity to *M. tuberculosis* wanes over time, and seems to be variable among geographical regions, resulting in an overall measured efficacy of 80% (93). Additionally, as this vaccine is a live-attenuated strain of *M. bovis*, the characteristic delayed-

type hypersensitivity is induced when PPD is administered in diagnostic setting. This makes distinguishing BCG vaccination from latent infection difficult (94). Further, the immunological factors that distinguish active TB cases from those of "healthy" contacts (i.e. persons with close-contact to TB patients showing no signs of disease), or persons with LTBI are not fully understood (95-98). However it is known that antigen-specific T cells with the ability to secrete IFN-γ (and other inflammatory cytokines) are necessary to invoke a productive cellular immune response (73,82).

Characterization of both protein and non-protein agonists responsible for effective and long-lasting immune responses in human cases of *M. tuberculosis* infection will aid in the development of both novel diagnostic reagents and vaccine candidates. During the 1950's and 1960's, the immunological properties of the extracted polysaccharide fraction of the bacillus were first elucidated. In these studies, crude envelope preparations comprised of arabinomannan, arabinogalactan and mannan (then termed Wax D) were shown to precipitate with antibody (precipitin positive reaction), interact with complement and stimulate hemagglutination with rabbit antiserum (99-102). Processing and presentation of mycobacterial antigens was also studied in the context of whole, heat-killed bacterium (103) or by live infection of murine macrophages (104). These studies and others demonstrated the immunostimulatory properties of the cell wall and that these antigens could be presented to the host cell surface to activate the cell-mediated immune response (105).

Barnes and Hirschfield characterized the protein and peptide antigens of the cell envelope responsible for T cell stimulation and the induction of protective immunity (106,107). Cell wall associated antigens were tested in tuberculosis patients and healthy, tuberculin skin-test positive individuals. In the study of Barnes, et al., it was demonstrated that the protein-peptidoglycan

complex evoked similar immune responses to that induced by whole *M. tuberculosis* sonicate. Antigen-specific T cell lines were generated and identified several protein species ranging in size from 10 – 65 kDa, as having strong immunostimulatory activity. Three clones were also able to recognize protein antigens in the secreted culture filtrate isolated from *in vitro* grown *M. tuberculosis* (106). CD4 T cells respond to antigenic peptides presented on the cell surface of macrophages and dendritic cells in the context of major histocompatibility complex II (MHC II), while CD8 T cell antigens are presented on the host cell surface by MHC I molecules. The repertoire of CD4 and CD8 T cell antigens has been studied for some time and have associated the immune response to the antigenicity of many secreted proteins of the culture filtrate, including post-translationally modified proteins such as the 19kDA glyco-lipoprotein and the 45kDa glycoprotein Mpt32 (106,108-112). Antigen discovery efforts have expanded with the utilization of genomics and proteomics based strategies in which families of antigenic gene products could be derived from sequencing of the genome. From this information, using genome fragment libraries identified a number of novel CD4 T cell antigens (112). Lists of cell envelope associated CD4 and CD8 T cell antigens are provided in Tables 1.2 and 1.3, respectively.

Table 1.2 *M. tuberculosis* Antigens Recognized by Human CD4 T Cells

Name [1]	Gene Annotation	Function [2]
DnaK/Hsp70	Rv0350	Chaperone
GroEL2/Hsp65	Rv0440	Chaperone
GroES/Hsp10	Rv3418c	Chaperonin
HspX	Rv2031c	Latency Antigen
Ag85a	Rv3804c	Mycolyl-

		transferase/fibronectin binding
Ag85b	Rv1886c	Mycolyl-transferase/ fibronectin binding
Ag85c	Rv3803c	Mycolyl-transferase/fibronectin binding
Esat6	Rv3875	Virulence Factor
Cfp10	Rv3874	Virulence Factor
TB10.4/EsxH	Rv0288	Virulence Factor
TB10.3/EsxR	Rv3019c	Unknown
TB12.9/EsxQ	Rv3017c	Unknown
Mpt64	Rv1980c	Unknown
TB8.4	Rv1174c	Unknown
EsxN	Rv1793	Unknown
EsxV	Rv3619c	Unknown
EsxL	Rv1198	Unknown
Esxl	Rv1037c	Unknown
EsxO	Rv2346c	Unknown

PepA	Rv0125	Serine protease
PPE4	Rv0286	Unknown
PPE18	Rv1196	Unknown
PPE46	Rv3018c	Unknown
PPE47	Rv3021c	Unknown
PPE68	Rv3873	Unknown
PE35	Rv3872	Unknown
Hbha	Rv0475	Adhesin
LpqH	Rv3763	Lipoprotein/Unknown
PstS1	Rv0934	Phosphate Transport
Mpt32/45kDa	Rv1860	Glycoprotein/Unknown
Rv1733c	Rv1733c	Transmembrane Protein
Rv2029c	Rv2029c	Phosphofrucktokinase
Rv2627c	Rv2627c	Unknown
Rv2628	Rv2628	Unknown
Rv2653	Rv2653	Prophage
Rv2654	Rv2654	Prophage

Rv3407	Rv3407	Unknown

1. **Reference (113)**
2. **Tuberculist (http://tuberculist.epfl.ch/)**

Table 1.3 *M. tuberculosis* Antigens Recognized by Human CD8 T Cells

Name [1]	Gene Annotation	Restricting Allele[2]
LpqH	Rv3763	HLA-A2
Ag85A	Rv3804c	HLA-A2
Ag85B	Rv1886c	HLA-A2
Ag85 Complex	Rv3804c/Rv1886c/Rv3803c	HLA-B35
Ald	Rv2780	HLA-A2
CFP10	Rv3874	HLA-B[3]
ESAT-6	Rv3875	HLA-A6802/B53/A2
GlnS	Rv2220	HLA-A2
GroEL2/hsp65	Rv0440	HLA-A2
HspX	Rv2031c	HLA-A2
Mtb8.4	Rv1174c	HLA-B3514/1501
Mtb9.8	Rv0287	HLA-B0801/A0201
Mtb39	Rv1196	HLA-B44
Rv2093c	Rv2093c	HLA-B35
Rv0341	Rv0341	HLA-A2
SodA	Rv3846	HLA-A2

1. **Reference: (113)**
2. **Reference: (114)**

1.5.2 The CD4 T cell Immune Response

CD4 T cell immunity to mycobacterial infection is one of the essential defense mechanisms needed to minimize the progression of disease. Persons with acquired immunodeficiencies such as a positive HIV status, have an increased likelihood of developing TB disease from 5%-10% over a lifetime to 10% per year (115). Furthermore, persons with defects in genes that encode for the inflammatory response cytokines and/or their receptors (IFN-γ, IL-12 and IL-23) are unusually susceptible to severe mycobacterial infection (116). Some patients that develop autoantibodies against IFN-γ, for example also are prone to uncontrolled, disseminated infections. In cases of advanced stages of disease or disseminated infection within previously immunocompetent individuals, it is commonly observed that patients have diminished CD4 T cell responses to characteristic antigenic repertoires (113). In the laboratory, depletion of CD4 T cells in mice results in severe impairment of the immune response to *M. tuberculosis* (117). Secretion of IFN-γ by both CD4 and CD8 T cells is key to macrophage activation and killing of infected cells (83). And it has recently been demonstrated that CD4 T cells must effectively secrete this cytokine in order to obtain optimal stimulation and co-secretion from other T cell subsets during *M. tuberculosis* infection in the mouse (118). This study and others suggests that future vaccine strategies that focus on CD4 T cell-specific immunity are needed not only to enhance the primary response to acute infection, but are also necessary to improve the efficiency of disease clearance mechanisms employed by other T cell subtypes such as CD8 effector cells.

1.5.3 The CD8 T cell Immune Response

Immunity to infection with *M. tuberculosis* requires both CD4 and CD8 T cells for successful control of infection. In humans, both cellular phenotypes have redundant functions such as cytolysis and the release of IFN-γ and TNF-α (119,120). The targeting of infected host cells for apoptosis has also been demonstrated for both cell types through the release of granulysin and induction of the perforin-granzyme pathway or through the Fas-Fas ligand pathways (121,122). However several studies have addressed the unique role CD8 T cells can play in the host defense against mycobacterial infection. In the mouse model of TB, a protective role for CD8 T cells in the host response was demonstrated in studies of adoptive transfer or *in vivo* depletion of CD8 T cells (117,123). In humans, unlike for CD4 T cells and their depletion in cases of HIV+ co-infection, the role of CD8 T cells in protection against *M. tuberculosis* is not as clear, however several hypothesis have been put forth. For example, using CD8 T cell clones from latently and actively infected individuals, CD8 T cells were shown to preferentially recognize and kill heavily infected dendritic cells and macrophages (124). This phenomenon was not observed for populations of CD4 T cells. From this work it was hypothesized that CD8 T cells may work to recognize cells in which the intracellular mycobacterial containment within the phagosome has been subverted, in which case cytolytic processes could then be initiated. Furthermore it has been demonstrated that antigen specific CD8 T cells are capable of recognizing *M. tuberculosis* residing in MHC class II negative cells (i.e cells incapable of presenting antigen to CD4 T cells) (125). CD8 T cell antigens are presented in the context of MHC class I, a molecule that is present in all nucleated cells. This gives CD8 T lymphocytes an increased repertoire of cells in which to identify pathogen specific antigens. In individuals presenting with active TB disease, it is well established that *M. tuberculosis* resides within granulomatous lesions within the lungs, however much less is known about the nature of the

bacterial burden in humans during persistent infection (113). This increased capacity to detect a wider variety of infected cells potentially gives CD8 T lymphocytes an advantage over CD4 T lymphocytes during stages of clinical latency (124). The relative importance of CD8 T cell function during prolonged infection is further supported in small animal models, despite the inherent difficulty of modeling latency *in vivo*. Andersen, et al., demonstrated a pronounced affect over the control of latent infection and a higher incidence of reactivation in mice depleted of CD8 T cells compared to CD4-depleted mice (126). Also in the mouse model, mice incapable of presenting antigen via MHC class I ($\beta 2m^{-/-}$) had significantly increased bacterial burden within macrophages (127). Furthermore, *M. tuberculosis* specific CD8 T cells are more frequent in non-human primates exhibiting long-term infection (128) and just recently it has been published that in young children, the frequency and abundance of CD8 T cells are a robust indicator and diagnostic measure of primary *M. tuberculosis* infection (129). All of this information suggests that beyond their role in cytotoxicity and immune surveillance, CD8 T cells and elucidation of their cognate antigens may lead to improved diagnostics in both latently infected individuals and children, (9,130) as well as provide a target strategy for the development of novel vaccine candidates.

1.5.4 Antigen Processing & Presentation by CD4 and CD8 T Cells

Antigen processing pathways present in professional antigen presenting cells (APCs) (macrophages, monocytes and dendritic cells) transform protein antigens derived from the extracellular milieu or from the intracellular cytosol into peptides. Peptides are subsequently loaded onto MHC molecules (MHC class II for CD4 T cells and MHC I for CD8 T cells) for display on the cell surface to T lymphocytes. In general, antigen uptake occurs through endocytosis of soluble antigen or through the delivery of cytosolic proteins derived from

intracellular pathogens. Protein antigens are digested within acidic vesicular compartments of the APC generating MHC II – associated peptides or are processed through the proteosome into peptide fragments for MHC I association. The resulting peptides have structural characteristics required for binding within the peptide groove of either MHC class II molecules or I and act to stabilize the MHC-peptide complex for cell surface expression (131). In *M. tuberculosis*, the mechanisms involved in antigen processing and presentation on MHC II are more complex and are critical for the induction of an effective CD4 T cell immune response and subsequent control of infection (132). In addition to the classical endocytosis model of antigen uptake, the phagocytosis of whole bacilli by the APC also results in the introduction of antigen in to the endocytic-processing pathway (133). Resident within the phagosome, *M. tuberculosis* specific ligands such as glycolipids, lipoproteins and other soluble antigens (Table 1.2) are readily accessible to the processing machinery. These ligands are recognized by Toll-like receptors (TLRs) and nucleotide-binding oligomerization domain (NOD) proteins (134,135), which initializes the pro-inflammatory immune response and activation of naïve T cells. Several studies have noted one mechanism by which *M. tuberculosis*-infected cells may go undetected by the CD4 cell mediated immune response. Noss and colleagues reported that MHC II antigen processing and presentation in murine bone marrow macrophages infected with *M. tuberculosis* is diminished and that this decrease is to due to low levels of MHC II expression (136). Further studies, summarized in a review by Harding and Boom have demonstrated that the mechanism of immune evasion is mediated through the actions of TLR signaling initiated by the glycolipid and glycoprotein antigens. This results in an overall decreased expression of the MHC II processing and presentation pathway (132).

For CD8 T cells, activation occurs via the recognition of peptide epitopes associated with MHC I molecules. MHC I peptides are produced by the proteolytic digestion of cytosolic proteins derived from foreign antigens, viruses or as with tumor cells, an overabundance of mutated proteins, that may trigger processing and presentation by the MHC I pathway (131). Two pathways – the cytosolic and vacuolar models – represent the courses of antigen processing onto MHC I. In the cytosolic model, mycobacterial antigens internalized by phagosomes are able to escape the phagosomal compartments within the APC and gain access to the cytosolic processing machinery which consists of enzymes associated within the endoplasmic reticulum (ER) and the proteosome, a multiprotein enzyme complex responsible for degrading proteins into peptides(137). It has also been shown that several potent CD4 T cell antigens deriving from the secreted proteome of *M. tuberculosis* are able to escape the phagosome and be presented to CD8 T cells at the cell surface (Table 1.3) (138,139). In depth studies following the processing of known CD8 T cell antigens characterized the *M. tuberculosis* containing phagosome as an organelle that contains MHC I processing molecules (transporter associated with antigen processing – TAP and the protein disulfide isomerase – PDI) typically localized to the endoplasmic reticulum (140,141). Song and Harding first described the second pathway, termed the vacuolar model. In this system, antigen processing and loading onto MHC I occurs within the endosome and is TAP and proteosome independent (142). Further, presentation of exogenous antigens occurs with the binding of post-processed MHC I molecules (i.e. MHC I previously stabilized and released from the ER-Golgi complex) (143). This model is dependent on endosomal proteases and also seems to be more active in microenvironments of lower pH, such as those experienced by bacilli residing in phagosomes (144). It should also be noted that for CD8 T cells, presentation by MHC I molecules are described as being either classically or non-

classically restricted (145). MHC I complexes are described as classical or non-classical via their human leukocyte antigen (HLA) allele. Classically restricted HLA molecules (HLA-1a) HLA – A, B and C are the most genetically diverse and have over 1,000 known alleles (146). In contrast, non-polymorphic MHC class 1b molecules (HLA-EFG) present a limited array of peptides and represent a class of specialized receptors that can present leader sequences of endogenously synthesized class 1a molecules (147). In addition to peptide-presenting molecules, unconventional antigen presenting complexes have been described as being important for immunity against *M. tuberculosis* infection. For example, CD1 molecules are recognized by both CD8 and CD4 T cells and are designed to bind long acyl chains of lipid antigens. CD1 molecules are sub-typed as group I CD1 and include CD1a-c, or group II CD1, which is comprised of CD1d (148). Additionally, a new non-classical molecule termed MR1 has recently been characterized as an innate pathogen-reactive effector function (149) and may present alternative small-molecule metabolites derived from microbes (150).

1.6 Conclusion

Tuberculosis is an ancient disease. Its long history and co-evolution among a wide variety of animal species has made the successful eradication of this disease one of the most difficult challenges faced by modern society. Discovery of the tubercle bacilli and subsequent studies paved the way for detailed characterization of its physiology and biochemistry. The interaction of *M. tuberculosis* with the human host is almost too complex to properly summarize and new knowledge and advances in the study of host-pathogen interaction are evolving every year. The cellular envelope in particular is a key player to successful invasion and the establishment of successful intracellular growth. Among bacterial species, the chemical composition of its macromolecular structures has to be one of the most elegantly defined (Figure 1.2).

Immunologically, *M. tuberculosis* is a formidable foe. The organism is capable of not only surviving within the most harsh intracellular conditions of the phagosome, it is able to thrive within these environments. By inducing the formation of the granuloma, tubercle bacilli are able to persist within host tissues for decades in a constant battle between pathogen and host. This clinically quiescent state is vital to disease pathogenesis, as the bacterium can remain viable without overt damage to its host for many years before the onset of active disease and clinical symptoms. An active state of growth for the organism is in a way, its downfall. At this critical point, invading mycobacteria are most susceptible to the actions of first and second-line drugs and the damage to its host threatens its lifestyle. This is also when the bacteria are most virulent, most likely to increase its chances of transmission and propagation throughout the population. Tying the unique physiology of the bacillus, specifically the composition and integrity of its cellular envelope, to its disease pathogenesis is the bacterium's ability to transition into and out of dormancy (or non-replicating persistence) within its hosts. Further knowledge of the structural features and resident proteins intercalated within the cell wall and envelope complex is key to answering relevant biological questions regarding the unique antigenic capacity of the bacillus and how these unique features contribute to the immunopathogenesis of both acute and latent infection. Further, if we can increase our understanding of these host-pathogen interactions at the systems level, new and exciting developments in both the diagnosis of disease bio-signatures and the identification of functionally relevant protein-protein interactions responsible for persistent survival of the bacillus will become clinically useful for the treatment of tuberculosis.

1.7 LITERATURE CITED

1. WHO. (2011) WHO Report 2011: Global Tuberculosis Control. World Health Organization, Geneva
2. Brennan, M., Cobelens, F., Lanfranchi, B., Lienhardt, C., Sizemore, C., Waltzl, G., Werf, M. v. d., Young, D., and Zimmermann, N. (2011) *World Health Organization STOP TB Initiative*
3. Lonnroth, K., Castro, K. G., Chakaya, J. M., Chauhan, L. S., Floyd, K., Glaziou, P., and Raviglione, M. C. (2010) *Lancet* **375**, 1814-1829
4. CDC. (2012) *MMWR Morb Mortal Wkly Rep* **61**, 484-489
5. Harries, A. D., Lin, Y., Satyanarayana, S., Lonnroth, K., Li, L., Wilson, N., Chauhan, L. S., Zachariah, R., Baker, M. A., Jeon, C. Y., Murray, M. B., Maher, D., Bygbjerg, I. C., Enarson, D. A., Billo, N. E., and Kapur, A. (2011) *Int J Tuberc Lung Dis* **15**, 1436-1444, i
6. Lin, Y., Li, L., Mi, F., Du, J., Dong, Y., Li, Z., Qi, W., Zhao, X., Cui, Y., Hou, F., Zachariah, R., Kapur, A., Lonnroth, K., and Harries, A. D. (2012) *Trop Med Int Health*
7. Raviglione, M., Marais, B., Floyd, K., Lonnroth, K., Getahun, H., Migliori, G. B., Harries, A. D., Nunn, P., Lienhardt, C., Graham, S., Chakaya, J., Weyer, K., Cole, S., Kaufmann, S. H., and Zumla, A. (2012) *Lancet* **379**, 1902-1913
8. Castro, K. G., and LoBue, P. (2011) *Emerg Infect Dis* **17**, 337-342
9. Nyendak, M. R., Lewinsohn, D. A., and Lewinsohn, D. M. (2009) *Curr Opin Infect Dis* **22**, 174-182
10. Heifets, L., and Desmond, E. (2005) Clinical Mycobacteriology (Tuberculosis) Laboratory: Services and Methods. in *Tuberculosis and the Tubercle Bacillus* (Cole, S., Eisenach, K. D., McMurray, D. N., and Jacobs, W. R. eds.), ASM Press, Washington, DC. pp
11. Swai, H. F., Mugusi, F. M., and Mbwambo, J. K. (2012) *BMC Res Notes* **4**, 475
12. Maertzdorf J Fau - Weiner Iii, J., Weiner Iii J Fau - Kaufmann, S. H. E., and Kaufmann, S. H. (2012)
13. Pai, M., Dheda, K., Cunningham, J., Scano, F., and O'Brien, R. (2007) *Lancet Infect Dis* **7**, 428-438
14. Petit, A. (1955) *Acta Tuberc Belg* **46**, 459-504
15. Jellinek, K. (1949) *N Engl J Med* **240**, 680-683
16. Zumla, A., Abubakar, I., Raviglione, M., Hoelscher, M., Ditiu, L., McHugh, T. D., Squire, S. B., Cox, H., Ford, N., McNerney, R., Marais, B., Grobusch, M., Lawn, S. D., Migliori, G. B., Mwaba, P., O'Grady, J., Pletschette, M., Ramsay, A., Chakaya, J., Schito, M., Swaminathan, S., Memish, Z., Maeurer, M., and Atun, R. (2012) *J Infect Dis* **205 Suppl 2**, S228-240
17. Dalton, T., Cegielski, P., Akksilp, S., Asencios, L., Caoili, J. C., Cho, S. N., Erokhin, V. V., Ershova, J., Gler, M. T., Kazennyy, B. Y., Kim, H. J., Kliiman, K., Kurbatova, E., Kvasnovsky, C., Leimane, V., van der Walt, M., Via, L. E., Volchenkov, G. V., Yagui, M. A., and Kang, H. (2012) *Lancet*
18. Lavender, C., Globan, M., Sievers, A., Billman-Jacobe, H., and Fyfe, J. (2005) *Antimicrob Agents Chemother* **49**, 4068-4074
19. Hazbon, M. H., Bobadilla del Valle, M., Guerrero, M. I., Varma-Basil, M., Filliol, I., Cavatore, M., Colangeli, R., Safi, H., Billman-Jacobe, H., Lavender, C., Fyfe, J., Garcia-

Garcia, L., Davidow, A., Brimacombe, M., Leon, C. I., Porras, T., Bose, M., Chaves, F., Eisenach, K. D., Sifuentes-Osornio, J., Ponce de Leon, A., Cave, M. D., and Alland, D. (2005) *Antimicrob Agents Chemother* **49**, 3794-3802
20. Guo, H., Seet, Q., Denkin, S., Parsons, L., and Zhang, Y. (2006) *Journal of medical microbiology* **55**, 1527-1531
21. Vilcheze, C., Wang, F., Arai, M., Hazbon, M. H., Colangeli, R., Kremer, L., Weisbrod, T. R., Alland, D., Sacchettini, J. C., and Jacobs, W. R., Jr. (2006) *Nature medicine* **12**, 1027-1029
22. Telenti, A., Imboden, P., Marchesi, F., Lowrie, D., Cole, S., Colston, M. J., Matter, L., Schopfer, K., and Bodmer, T. (1993) *Lancet* **341**, 647-650
23. Zhang, Y. (2004) *Front Biosci* **9**, 1136-1156
24. Gomez, J. E., and McKinney, J. D. (2004) *Tuberculosis (Edinb)* **84**, 29-44
25. Karakousis, P. C., Williams, E. P., and Bishai, W. R. (2008) *J Antimicrob Chemother* **61**, 323-331
26. Rothschild, B. M., Martin, L. D., Lev, G., Bercovier, H., Bar-Gal, G. K., Greenblatt, C., Donoghue, H., Spigelman, M., and Brittain, D. (2001) *Clin Infect Dis* **33**, 305-311
27. Lee, O. Y., Wu, H. H., Donoghue, H. D., Spigelman, M., Greenblatt, C. L., Bull, I. D., Rothschild, B. M., Martin, L. D., Minnikin, D. E., and Besra, G. S. (2012) *PLoS One* **7**, e41923
28. Hershkovitz, I., Donoghue, H. D., Minnikin, D. E., Besra, G. S., Lee, O. Y., Gernaey, A. M., Galili, E., Eshed, V., Greenblatt, C. L., Lemma, E., Bar-Gal, G. K., and Spigelman, M. (2008) *PLoS One* **3**, e3426
29. Nerlich, A. G., Haas, C. J., Zink, A., Szeimies, U., and Hagedorn, H. G. (1997) *Lancet* **350**, 1404
30. Salo, W. L., Aufderheide, A. C., Buikstra, J., and Holcomb, T. A. (1994) *Proc Natl Acad Sci U S A* **91**, 2091-2094
31. Daniel, T. M. (2006) *Respir Med* **100**, 1862-1870
32. Dubos, R. J. (1952) *Bacteriol Rev* **16**, 145-151
33. Ayvazian, F. (1993) History of tuberculosis. in *Tuberculosis: A Comprehensive International Approach* (L.B., R., and Herschfield, E. S. eds.), Marcel Dekker, New York. pp 1-20
34. Shrestha-Kuwahara, R., Wilce, M., Joseph, H., Carey, J., Plank, R., and Sumartojo, E.
35. Daniel, T. M. (1969) *N Engl J Med* **280**, 615-616
36. Singh, B., and Mitchison, D. A. (1955) *J Gen Microbiol* **12**, 76-84
37. Crick, D. C., Mahapatra, S., and Brennan, P. J. (2001) *Glycobiology* **11**, 107R-118R
38. Matsuhashi, M. (1966) *Tanpakushitsu Kakusan Koso* **11**, 875-886
39. Lederer, E. (1971) *Pure Appl Chem* **25**, 135-165
40. McNeil, M., Wallner, S. J., Hunter, S. W., and Brennan, P. J. (1987) *Carbohydr Res* **166**, 299-308
41. Brennan, P. J., and Besra, G. S. (1997) *Biochem Soc Trans* **25**, 188-194
42. Draper, P. (1998) *Front Biosci* **3**, D1253-1261
43. Mahapatra, S., Yagi, T., Belisle, J. T., Espinosa, B. J., Hill, P. J., McNeil, M. R., Brennan, P. J., and Crick, D. C. (2005) *J Bacteriol* **187**, 2747-2757
44. Kaur, D., Guerin, M. E., Skovierova, H., Brennan, P. J., and Jackson, M. (2009) *Adv Appl Microbiol* **69**, 23-78

45. Minnikin, D. E., Minnikin, S. M., Goodfellow, M., and Stanford, J. L. (1982) *J Gen Microbiol* **128**, 817-822
46. Collins, M. D., Goodfellow, M., and Minnikin, D. E. (1982) *J Gen Microbiol* **128**, 129-149
47. Alderwick, L. J., Birch, H. L., Mishra, A. K., Eggeling, L., and Besra, G. S. (2007) *Biochem Soc Trans* **35**, 1325-1328
48. Mahapatra, S., Scherman, H., Brennan, P. J., and Crick, D. C. (2005) *J Bacteriol* **187**, 2341-2347
49. Lavollay, M., Arthur, M., Fourgeaud, M., Dubost, L., Marie, A., Veziris, N., Blanot, D., Gutmann, L., and Mainardi, J. L. (2008) *J Bacteriol* **190**, 4360-4366
50. Barry, C. E., Crick, D. C., and McNeil, M. R. (2007) *Infect Disord Drug Targets* **7**, 182-202
51. Vilkas, E., Amar, C., Markovits, J., Vliegenthart, J. F., and Kamerling, J. P. (1973) *Biochim Biophys Acta* **297**, 423-435
52. Daffe, M., Brennan, P. J., and McNeil, M. (1990) *J Biol Chem* **265**, 6734-6743
53. Minnikin, D. E., Kremer, L., Dover, L. G., and Besra, G. S. (2002) *Chem Biol* **9**, 545-553
54. Watanabe, M., Aoyagi, Y., Mitome, H., Fujita, T., Naoki, H., Ridell, M., and Minnikin, D. E. (2002) *Microbiology* **148**, 1881-1902
55. Briken, V., Porcelli, S. A., Besra, G. S., and Kremer, L. (2004) *Mol Microbiol* **53**, 391-403
56. Middlebrook, G., Coleman, C. M., and Schaefer, W. B. (1959) *Proc Natl Acad Sci U S A* **45**, 1801-1804
57. Zhang, L., Goren, M. B., Holzer, T. J., and Andersen, B. R. (1988) *Infect Immun* **56**, 2876-2883
58. Okamoto, Y., Fujita, Y., Naka, T., Hirai, M., Tomiyasu, I., and Yano, I. (2006) *Microbial pathogenesis* **40**, 245-253
59. Reed, M. B., Domenech, P., Manca, C., Su, H., Barczak, A. K., Kreiswirth, B. N., Kaplan, G., and Barry, C. E., 3rd. (2004) *Nature* **431**, 84-87
60. Lima, V. M., Bonato, V. L., Lima, K. M., Dos Santos, S. A., Dos Santos, R. R., Goncalves, E. D., Faccioli, L. H., Brandao, I. T., Rodrigues-Junior, J. M., and Silva, C. L. (2001) *Infect Immun* **69**, 5305-5312
61. Jarlier, V., and Nikaido, H. (1994) *FEMS Microbiol Lett* **123**, 11-18
62. Yang, Y., Auguin, D., Delbecq, S., Dumas, E., Molle, G., Molle, V., Roumestand, C., and Saint, N. (2010) *Proteins* **79**, 645-661
63. Barth, E., Barcelo, M. A., Klackta, C., and Benz, R. (2010) *J Bacteriol* **192**, 786-800
64. Wolfe, L. M., Mahaffey, S. B., Kruh, N. A., and Dobos, K. M. (2010) *J Proteome Res* **9**, 5816-5826
65. Abou-Zeid, C., Gares, M. P., Inwald, J., Janssen, R., Zhang, Y., Young, D. B., Hetzel, C., Lamb, J. R., Baldwin, S. L., Orme, I. M., Yeremeev, V., Nikonenko, B. V., and Apt, A. S. (1997) *Infect Immun* **65**, 1856-1862
66. Mohagheghpour, N., Gammon, D., Kawamura, L. M., van Vollenhoven, A., Benike, C. J., and Engleman, E. G. (1998) *J Immunol* **161**, 2400-2406
67. da Fonseca, D. P., Joosten, D., van der Zee, R., Jue, D. L., Singh, M., Vordermeier, H. M., Snippe, H., and Verheul, A. F. (1998) *Infect Immun* **66**, 3190-3197

68. Brightbill, H. D., Libraty, D. H., Krutzik, S. R., Yang, R. B., Belisle, J. T., Bleharski, J. R., Maitland, M., Norgard, M. V., Plevy, S. E., Smale, S. T., Brennan, P. J., Bloom, B. R., Godowski, P. J., and Modlin, R. L. (1999) *Science* **285**, 732-736
69. Chang, Z., Choudhary, A., Lathigra, R., and Quiocho, F. A. (1994) *J Biol Chem* **269**, 1956-1958
70. Nguyen, K. T., Piastro, K., and Derbyshire, K. M. (2009) *J Bacteriol* **191**, 2721-2727
71. Drage, M. G., Tsai, H. C., Pecora, N. D., Cheng, T. Y., Arida, A. R., Shukla, S., Rojas, R. E., Seshadri, C., Moody, D. B., Boom, W. H., Sacchettini, J. C., and Harding, C. V. (2010) *Nat Struct Mol Biol* **17**, 1088-1095
72. Flynn, J. L., Chan, J., and Lin, P. L. (2011) *Mucosal Immunol* **4**, 271-278
73. van Crevel, R., Ottenhoff, T. H., and van der Meer, J. W. (2002) *Clin Microbiol Rev* **15**, 294-309
74. Cooper, A. M., Solache, A., and Khader, S. A. (2007) *Curr Opin Immunol* **19**, 441-447
75. Shafiani, S., Tucker-Heard, G., Kariyone, A., Takatsu, K., and Urdahl, K. B. (2010) *J Exp Med* **207**, 1409-1420
76. Wolf, A. J., Linas, B., Trevejo-Nunez, G. J., Kincaid, E., Tamura, T., Takatsu, K., and Ernst, J. D. (2007) *J Immunol* **179**, 2509-2519
77. Wallgren, A. (1948) *Tubercle* **29**, 245-251
78. Poulsen, A. (1950) *Acta tuberculosea Scandinavica* **24**, 311-346
79. Wolf, A. J., Desvignes, L., Linas, B., Banaiee, N., Tamura, T., Takatsu, K., and Ernst, J. D. (2008) *J Exp Med* **205**, 105-115
80. Chackerian, A. A., Alt, J. M., Perera, T. V., Dascher, C. C., and Behar, S. M. (2002) *Infect Immun* **70**, 4501-4509
81. Blomgran, R., and Ernst, J. D. (2011) *J Immunol* **186**, 7110-7119
82. Ernst, J. D. (2012) *Nat Rev Immunol* **12**, 581-591
83. Flynn, J. L. (2004) *Tuberculosis (Edinb)* **84**, 93-101
84. Davis, J. M., and Ramakrishnan, L. (2009) *Cell* **136**, 37-49
85. Volkman, H. E., Pozos, T. C., Zheng, J., Davis, J. M., Rawls, J. F., and Ramakrishnan, L. (2009) *Science* **327**, 466-469
86. Koch, R. (1891) *Br Med J* **1**, 125-127
87. Von Pirquet, C. (1909) *JAMA: The Journal of the American Medical Association* **LII**, 675-678
88. Seibert, F. B. (1926) *Science* **63**, 619c-620c
89. Cho, Y. S., Dobos, K. M., Prenni, J., Yang, H., Hess, A., Rosenkrands, I., Andersen, P., Ryoo, S. W., Bai, G. H., Brennan, M. J., Izzo, A., Bielefeldt-Ohmann, H., and Belisle, J. T. (2012) *Proteomics* **12**, 979-991
90. Yang, H., Kruh-Garcia, N. A., and Dobos, K. M. (2012) *FEMS Immunol Med Microbiol*
91. Rosenthal, S. R. (1957) *Am J Med Technol* **23**, 354-360
92. Trunz, B. B., Fine, P., and Dye, C. (2006) *Lancet* **367**, 1173-1180
93. Fine, P. E. (1995) *Lancet* **346**, 1339-1345
94. Colditz, G. A., Brewer, T. F., Berkey, C. S., Wilson, M. E., Burdick, E., Fineberg, H. V., and Mosteller, F. (1994) *JAMA* **271**, 698-702
95. Havlir, D. V., Wallis, R. S., Boom, W. H., Daniel, T. M., Chervenak, K., and Ellner, J. J. (1991) *Infect Immun* **59**, 665-670
96. Lewinsohn, D. M., Alderson, M. R., Briden, A. L., Riddell, S. R., Reed, S. G., and Grabstein, K. H. (1998) *J Exp Med* **187**, 1633-1640

97. Barry, C. E., 3rd, Boshoff, H. I., Dartois, V., Dick, T., Ehrt, S., Flynn, J., Schnappinger, D., Wilkinson, R. J., and Young, D. (2009) *Nat Rev Microbiol* **7**, 845-855
98. Deenadayalan, A., Heaslip, D., Rajendiran, A. A., Velayudham, B. V., Frederick, S., Yang, H. L., Dobos, K., Belisle, J. T., and Raja, A. (2010) *Mol Cell Proteomics* **9**, 538-549
99. Azuma, I., Kimura, H., Niinaka, T., Aoki, T., and Yamamura, Y. (1968) *J Bacteriol* **95**, 263-271
100. Azuma, I., Kimura, H., and Yamamura, Y. (1968) *J Bacteriol* **96**, 567-568
101. Yamamura, Y., Onoue, K., and Azuma, I. (1968) *Ann N Y Acad Sci* **154**, 88-97
102. Kotani, S., Kato, T., Matsuda, T., Kato, K., and Misaki, A. (1971) *Biken J* **14**, 379-387
103. Bhardwaj, V., and Colston, M. J. (1988) *Eur J Immunol* **18**, 691-696
104. Mshana, R. N., Hastings, R. C., and Krahenbuhl, J. L. (1988) *Immunobiology* **177**, 40-54
105. Dannenberg, A. M., Jr. (1989) *Rev Infect Dis* **11 Suppl 2**, S369-378
106. Barnes, P. F., Mehra, V., Hirschfield, G. R., Fong, S. J., Abou-Zeid, C., Rook, G. A., Hunter, S. W., Brennan, P. J., and Modlin, R. L. (1989) *J Immunol* **143**, 2656-2662
107. Hirschfield, G. R., McNeil, M., and Brennan, P. J. (1990) *J Bacteriol* **172**, 1005-1013
108. Rook, G. A., Steele, J., Barnass, S., Mace, J., and Stanford, J. L. (1986) *Clin Exp Immunol* **63**, 105-110
109. Andersen, P., Askgaard, D., Ljungqvist, L., Bentzon, M. W., and Heron, I. (1991) *Infect Immun* **59**, 1558-1563
110. Young, D., Garbe, T., Lathigra, R., Abou-Zeid, C., and Zhang, Y. (1991) *Bull Int Union Tuberc Lung Dis* **66**, 47-51
111. Dhiman, N., Verma, I., and Khuller, G. K. (1997) *Microbiol Immunol* **41**, 495-502
112. Alderson, M. R., Bement, T., Day, C. H., Zhu, L., Molesh, D., Skeiky, Y. A., Coler, R., Lewinsohn, D. M., Reed, S. G., and Dillon, D. C. (2000) *J Exp Med* **191**, 551-560
113. Ottenhoff, T. H., Lewinsohn, D. A., and Lewinsohn, D. M. (2008) Human CD4 and CD8 T Cell Responses to Mycobacterium tuberculosis: Antigen Specificity, Function, Implications and Applications. in *Handbook of Tuberculosis: Immunology and Cell Biology* (Kaufmann, S. H., and Britton, W. J. eds.), Wiley-VCH, Germany. pp 119-143
114. Lewinsohn, D. A., Winata, E., Swarbrick, G. M., Tanner, K. E., Cook, M. S., Null, M. D., Cansler, M. E., Sette, A., Sidney, J., and Lewinsohn, D. M. (2007) *PLoS Pathog* **3**, 1240-1249
115. Diedrich, C. R., and Flynn, J. L. *Infect Immun* **79**, 1407-1417
116. Ottenhoff, T. H., Verreck, F. A., Lichtenauer-Kaligis, E. G., Hoeve, M. A., Sanal, O., and van Dissel, J. T. (2002) *Nature genetics* **32**, 97-105
117. Flory, C. M., Hubbard, R. D., and Collins, F. M. (1992) *J Leukoc Biol* **51**, 225-229
118. Bold, T. D., and Ernst, J. D. (2012) *J Immunol* **189**, 2530-2536
119. Mutis, T., Cornelisse, Y. E., and Ottenhoff, T. H. (1993) *Eur J Immunol* **23**, 2189-2195
120. Ochoa, M. T., Stenger, S., Sieling, P. A., Thoma-Uszynski, S., Sabet, S., Cho, S., Krensky, A. M., Rollinghoff, M., Nunes Sarno, E., Burdick, A. E., Rea, T. H., and Modlin, R. L. (2001) *Nature medicine* **7**, 174-179
121. Canaday, D. H., Wilkinson, R. J., Li, Q., Harding, C. V., Silver, R. F., and Boom, W. H. (2001) *J Immunol* **167**, 2734-2742
122. Lewinsohn, D. M., Bement, T. T., Xu, J., Lynch, D. H., Grabstein, K. H., Reed, S. G., and Alderson, M. R. (1998) *J Immunol* **160**, 2374-2379

123. Leveton, C., Barnass, S., Champion, B., Lucas, S., De Souza, B., Nicol, M., Banerjee, D., and Rook, G. (1989) *Infect Immun* **57**, 390-395
124. Lewinsohn, D. A., Heinzel, A. S., Gardner, J. M., Zhu, L., Alderson, M. R., and Lewinsohn, D. M. (2003) *Am J Respir Crit Care Med* **168**, 1346-1352
125. Hernandez-Pando, R., Jeyanathan, M., Mengistu, G., Aguilar, D., Orozco, H., Harboe, M., Rook, G. A., and Bjune, G. (2000) *Lancet* **356**, 2133-2138
126. van Pinxteren, L. A., Cassidy, J. P., Smedegaard, B. H., Agger, E. M., and Andersen, P. (2000) *Eur J Immunol* **30**, 3689-3698
127. Flynn, J. L., Goldstein, M. M., Triebold, K. J., Koller, B., and Bloom, B. R. (1992) *Proc Natl Acad Sci U S A* **89**, 12013-12017
128. Lewinsohn, D. M., Tydeman, I. S., Frieder, M., Grotzke, J. E., Lines, R. A., Ahmed, S., Prongay, K. D., Primack, S. L., Colgin, L. M., Lewis, A. D., and Lewinsohn, D. A. (2006) *Microbes Infect* **8**, 2587-2598
129. Lancioni, C., Nyendak, M., Kiguli, S., Zalwango, S., Mori, T., Mayanja-Kizza, H., Balyejusa, S., Null, M., Baseke, J., Mulindwa, D., Byrd, L., Swarbrick, G., Scott, C., Johnson, D. F., Malone, L., Mudido-Musoke, P., Boom, W. H., Lewinsohn, D. M., and Lewinsohn, D. A. (2012) *Am J Respir Crit Care Med* **185**, 206-212
130. Lewinsohn, D. A., Gennaro, M. L., Scholvinck, L., and Lewinsohn, D. M. (2004) *Int J Tuberc Lung Dis* **8**, 658-674
131. Abbas, A. K., Lichtman, A. H., and Pillai, S. (eds). (2010) *Cellular and Molecular Immunology*, Vol. 6, Saunders Elsevier, Philadelphia, PA
132. Harding, C. V., and Boom, W. H. (2010) *Nat Rev Microbiol* **8**, 296-307
133. Ramachandra, L., Noss, E., Boom, W. H., and Harding, C. V. (2001) *J Exp Med* **194**, 1421-1432
134. Krutzik, S. R., and Modlin, R. L. (2004) *Seminars in immunology* **16**, 35-41
135. Ferwerda, G., Girardin, S. E., Kullberg, B. J., Le Bourhis, L., de Jong, D. J., Langenberg, D. M., van Crevel, R., Adema, G. J., Ottenhoff, T. H., Van der Meer, J. W., and Netea, M. G. (2005) *PLoS Pathog* **1**, 279-285
136. Noss, E. H., Harding, C. V., and Boom, W. H. (2000) *Cellular immunology* **201**, 63-74
137. Mazzaccaro, R. J., Gedde, M., Jensen, E. R., van Santen, H. M., Ploegh, H. L., Rock, K. L., and Bloom, B. R. (1996) *Proc Natl Acad Sci U S A* **93**, 11786-11791
138. Lewinsohn, D. M., Grotzke, J. E., Heinzel, A. S., Zhu, L., Ovendale, P. J., Johnson, M., and Alderson, M. R. (2006) *J Immunol* **177**, 437-442
139. Grotzke, J. E., Siler, A. C., Lewinsohn, D. A., and Lewinsohn, D. M. (2010) *J Immunol* **185**, 4336-4343
140. Grotzke, J. E., Harriff, M. J., Siler, A. C., Nolt, D., Delepine, J., Lewinsohn, D. A., and Lewinsohn, D. M. (2009) *PLoS Pathog* **5**, e1000374
141. Ackerman, A. L., Giodini, A., and Cresswell, P. (2006) *Immunity* **25**, 607-617
142. Song, R., and Harding, C. V. (1996) *J Immunol* **156**, 4182-4190
143. Chefalo, P. J., and Harding, C. V. (2001) *J Immunol* **167**, 1274-1282
144. Gromme, M., and Neefjes, J. (2002) *Mol Immunol* **39**, 181-202
145. Stroynowski, I., and Lindahl, K. F. (1994) *Curr Opin Immunol* **6**, 38-44
146. de Villartay, J. P., Rouger, P., Muller, J. Y., and Salmon, C. (1985) *Tissue antigens* **26**, 12-19
147. Joyce, S., Tabaczewski, P., Angeletti, R. H., Nathenson, S. G., and Stroynowski, I. (1994) *J Exp Med* **179**, 579-588

148. Moody, D. B., Reinhold, B. B., Guy, M. R., Beckman, E. M., Frederique, D. E., Furlong, S. T., Ye, S., Reinhold, V. N., Sieling, P. A., Modlin, R. L., Besra, G. S., and Porcelli, S. A. (1997) *Science* **278**, 283-286
149. Gold, M. C., Eid, T., Smyk-Pearson, S., Eberling, Y., Swarbrick, G. M., Langley, S. M., Streeter, P. R., Lewinsohn, D. A., and Lewinsohn, D. M. (2012) *Mucosal Immunol*
150. Kjer-Nielsen, L., Patel, O., Corbett, A. J., Le Nours, J., Meehan, B., Liu, L., Bhati, M., Chen, Z., Kostenko, L., Reantragoon, R., Williamson, N. A., Purcell, A. W., Dudek, N. L., McConville, M. J., O'Hair, R. A., Khairallah, G. N., Godfrey, D. I., Fairlie, D. P., Rossjohn, J., and McCluskey, J. (2012) *Nature* **491**, 717-723

II. CHAPTER II: Mass Spectrometry Approaches in the Study of the *M. tuberculosis* Proteome for Diagnostic and Drug Targets

2.1 Introduction

Proteomics is often defined as the large-scale study of proteins within a system (1). For several generations the techniques and technologies utilized in this multidisciplinary field have rapidly evolved. Fundamentally, however three central themes in proteomics have remained the same: (1) The large-scale identification of proteins, (2) the identification and characterization of post-translational modifications and (3) the analysis of protein-protein interactions (2,3). Developments in the next decade in the field of mass spectrometry-based proteomics will see new applications of proteome analysis (4), including the labeling and characterization of functional sub-proteomes (5), advances in the accuracy and sensitivity of quantitative measurements (6) and improvements in search algorithms and data analysis tools (7). With these advances proteomic studies are moving beyond descriptive analyses, answering hypothesis driven biological questions and further developing into unique assay platforms whereby hundreds of proteins can be quantitatively monitored in a given system (8). For tuberculosis research, proteomics and mass spectrometry have become powerful tools to answer biologically relevant questions. These studies have contributed useful insights into pathogen biology, including a more comprehensive understanding of strain diversity, cellular physiology, immunogenicity and interactions between the host and pathogen.

2.2 The *M. Tuberculosis* Proteome

The genome sequence of *M. tuberculosis* was published by Cole et al., in 1998 and which facilitated the rapid advances in the fields of genomics/transcriptomics and proteomics. Comprehensive studies of the *M. tuberculosis* proteome focused on the characterization of

subcellular fractions such as cytosol, cell wall and membrane, including the proteins secreted from the bacillus during *in vitro* growth. This information was often utilized to study changes is protein profles under various controlled conditions or to monitor the pathogen-specific proteome in the context of host cells. Proteomic studies were also frequently used to gather information on immunodominate proteins, and identify novel B and T cell antigens for the development TB vaccines and diagnostic tools.

2.2.1 Discovery & Descriptive Analyses

Gel-based techniques for the study of *M. tuberculosis* proteins provided the first two dimensional gel electrophoresis (2DGE) reference maps of soluble protein fractions. Prior to the publication of the genome sequence of *M. tuberculosis*, H_{37}Rv, Sonnenberg et al., successfully resolved the secreted immunodominant culture filtrate proteins of the bacillus, visualizing over 200 protein spots. Using a combination of western blot analysis, N-terminal Edman sequencing and LC-MS of digested protein spots, approximately 40 proteins were characterized. This study also described the first observation of a then novel member of the Ag85 complex, Mpt51(9). Subsequent studies using 2DGE improved the resolving power of the gels by expanding the range in the first dimension separation, using immobilized pH gradients. The combination of the complete genome sequence (10) and the successful resolution of both secreted and cell associated protein fractions by 2DGE, facilitated the identification of hundreds of protein species by tandem mass spectrometry (LC-MS/MS). These included the first proteomic description of differences between the lab strain of H_{37}Rv and vaccine strain *M. bovis*, BCG (11,12), and the development of the first online 2DGE database of *M. tuberculosis* proteins (http://www.mpiib-berlin.mpg.de/2D-PAGE) (13). Despite the success of these initial studies, there were known limitations of 2DGE technology, especially in the identification of low-abundance proteins and proteins within highly insoluble fractions such as membrane and cell wall. Turning to liquid

chromatography-tandem mass spectrometry (LC-MS/MS), facilitated the identification of many more proteins using non-gel based approaches... Using LC-MS/MS in combination with 2DGE resulted in more comprehensive descriptions of the secreted proteins of the culture filtrate (14), cytosolic and cell wall proteins (15,16), whole cell lysate (17) and membrane fractions (18-21). Studies such as these represent a more than two decades of work in the field of proteomics. Computationally, each study pushed the limits of available bioinformatics tools that were built upon genomic sequence information and from this work, the reliability and accuracy of predictive algorithms (i.e. secretion signal prediction) could be verified (22). For other computational tools, such as transmembrane domain prediction, validating the presence or absence of membrane proteins was another challenge taken on in the study of the membrane proteome. The routine use of LC-MS/MS facilitated the identification and characterization of many predicted transmembrane proteins. Increased instrument sensitivity also allowed for the identification of other membrane-associated proteins with no physiological transmembrane features, thus new protein complexes were potentially defined (23,24).

2.2.2 Proteomic Analysis *in vivo* & Conditions Mimicking Host Environment

Proteomic studies of *M. tuberculosis* grown in liquid culture are very valuable to the research community as baseline descriptions of the presence or absence of proteins within a fraction of interest. However, the ability of *M. tuberculosis* to invade the body and occupy unique niches within a host is a defining characteristic of the bacillus and recognizing proteins that are present during natural states of infection will aid the development of second-generation diagnostic reagents and identify relevant proteomic targets in the design of novel chemotherapeutics. Recently, the *M. tuberculosis* proteome was described during *in vivo* infection of guinea pigs. Proteomic profiles of *M. tuberculosis* isolated from granulomas were

characterized at day 30 post-aerosol challenge and compared to the proteomic profile of isolated bacilli at day 90. In this study, over 500 proteins were identified from *M. tuberculosis* isolated from the granulomas of infected guinea pigs (25). Proteins involved in nitrogen assimilation and cation transport were shown to be the least variable throughout the course of infection. When compared to representative *in vitro* models of gene transcription, the most correlative model was that of non-replicating persistence (NRP) in which the gene expression profile of the microaerophilic (NRP-1) was compared to the anaerobic, dormant state (NRP-2)(26,27). However, it was noted there was very little overlap between the *in vivo*-generated dataset and prior *in vitro* models simulating the host environment (26,28-31). Interestingly, the proteomic profile of bacilli isolated from the infected lung tissue of the guinea pig demonstrated the least correlation with models of nutrient depletion, leading to the hypothesis that the bacilli within an infected host are not as deprived of nutrients as once thought, as many proteins involved in lipid degradation were identified in this work (25). A second notable contrast between *in vivo* proteomic data and expression profiles generated during conditions mimicking the host environment was the absence of known secreted, immunodominant antigens such as HspX and other typically dominant proteins such as GroEl and DnaK (32). Kruh et al., hypothesized this was due to the rapid secretion of these proteins into the host milieu where their trafficking to the draining lymph node may occur or alternatively these proteins could be rapidly degraded. To date, this has been the only proteome-based identification of proteins in the context of a mammalian host.

As mycobacteria infect the cell, the bacilli are engulfed within phagosomal compartments in which they encounter several hostile host factors. Thus, several proteomic profiles unique to intraphagosomally grown *M. tuberculosis* have identified proteins unique to this phase of

virulence. Lee and colleagues first observed changes at the protein level when metabolically labeled proteins were monitored during intracellular growth in a human macrophage cell line. Labeled proteins were separated by 2DGE and unique changes in protein profiles were observed under conditions of heat-shock, low pH and in the presence of peroxides (33). Following this work, Monahan and Betts et al., studied the intraphagosomal protein profile of the vaccine strain *M. bovis*, BCG. They determined six proteins to have increased abundance inside macrophages, most of which are involved in the stress response or transition into dormancy (34). These included chaperones GroEL1/2 and the heat shock protein HspX as well as the hypothetical protein Rv2623 – a protein known to be regulated by DosR. In 2006, Mattow and colleagues identified eleven proteins exclusive to phagasomal infection, including the malate dehydrogenase (Mdh) and the peptidyl-proline-cis-trans isomerase (PpiA) (35).

Studies of the *in vivo* (or *in situ*) characterization of mycobacterial proteomes are relatively scarce. Most likely this is due to the lack of suitable methods for overcoming the extremely high abundance of host proteins in comparison to bacterial proteins and their accurate identification and quantification in high background. However, the methods employed within the *in vivo* proteome dataset demonstrate improvements in technical and computing resources that will facilitate similar, more targeted studies in the future.

2.2.3 Comparison of Virulent & Avirulent Strains

As mentioned above, the first proteomic analysis of both the virulent $H_{37}Rv$ and the avirulent *M. bovis*, BCG vaccine strain of mycobacteria was undertaken over a decade ago in an attempt to describe novel virulence determinants (11,13). The induction of protective immunity via BCG vaccination is attributed to the highly abundant secreted proteins which are presented to the immune system more efficiently with a live vaccine (36). Proteomic comparisons by 2DGE identified several proteins unique to each species. The *M. tuberculosis* specific proteins included

the 40kDa antigen Ald, the immunostimulatory antigen Mpt64, and members of the ESAT-6-like protein cluster, including Cfp10. ESAT-6 and Cfp10 are now used as *M. tuberculosis*-specific diagnostic reagents in the quantiferon-Gold immunoassay (37,38). Other antigens unique to *M. tuberculosis* include several hypothetical proteins (Rv0020c, Rv1684, Rv1893, Rv3046c, Rv3881c, Rv1198 and Rv1793), the chaperone GroES and elongation factor EF-Tu (39). These proteins could potentially be used in the design of novel vaccines and/or diagnostic methods. Comparing the strains of *M. tuberculosis* H_{37}Rv and *M. tuberculosis* H_{37}Ra has also resulted in the identification of proteins unique to virulent strains. One study observed the upregulation of eleven proteins in H_{37}Rv and the down regulation of three (40). Most recently, Målen and colleagues compared the proteomes of these two strains using label-free quantitative proteomics (described below) where more than 1700 proteins were identified from both strains. They identified 29 membrane-associated proteins. 19 proteins were higher in abundance for H_{37}Rv, while 10 proteins were higher in H_{37}Ra. Among the more abundant proteins identified in H_{37}Rv were the ABC-transporters (Rv0933, Rv1273c and Rv1819c), as well as the protein-export membrane protein SecF (21). These studies and others demonstrate that the virulence factors at the proteome level are quite complex and differing pathological phenotypes cannot be attributed to the presence or absence of one factor or another.

2.3 Proteomic Technologies

The typical workflow in a proteomics experiment involves the enzymatic digestion of a protein sample, usually with site-specific enzymes such as trypsin. Pre-fractionation of protein/peptide samples via liquid chromatography, one-dimensional or two-dimensional gel electrophoresis is employed to minimize the saturation levels of high abundance proteins and overall, yields higher proteome coverage (41). For several decades, two-dimensional gel

electrophoresis was the method of choice for gel-based pre-fractionation of protein samples, prior to enzymatic treatment. Gel-based fractionation techniques, were the technique of choice prior to the large-scale use of mass spectrometry (MS) for protein identification. Advances in throughput and MS technologies facilitated liquid-chromatography based fractionation techniques for the generation of peptide fractions. Once peptides are introduced into a mass spectrometer, fragment spectra (MS/MS spectra) are generated and recorded. Search algorithms such a SEQUEST (42) or MASCOT (43) are able to align sequences in the database to the acquired data. Detection and identification of peptides unique to specific proteins confirm their presence within a given sample, where typically two unique peptides are enough to identify a protein (6). The following sections discuss the use of two-dimensional gel electrophoresis and multi-dimensional liquid chromatography techniques, with a brief review of MS technologies.

2.3.1 Two Dimensional Gel Electrophoresis

Early resolution of multi-protein sample sets utilized two-dimensional gel electrophoresis (2DGE) to generate protein maps present within cells or tissues. 2DGE technology separated proteins first via iso-electric point (iso-electric focusing), usually throughout relatively neutral pH ranges (pH 4-8) and then by molecular mass via sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) (44). Whole protein maps could be monitored over time, and changes to the protein profile observed (2,45). 2DGE immunoblots could be probed for reactivity against immune sera or protein antibodies to identify immunoreactive polypeptides and select protein targets (46). As a precursor to the identification of peptides via mass spectrometry, protein spots were excised out of the gel and processed for amino acid analysis via Edman degradation (47), or subjected to trypsin digestion for sequence analysis (48,49). Fragment peptide sequences were searched against available databases, and proteins could be identified if sequences, molecular masses and predicted pI values could be cross-matched.

In addition to protein identification, post-translational modifications (PTM), such as phosphorylation and glycosylation can be visualized via 2DGE as these properties affect both the iso-electric point and molecular weight of proteins (50). While a powerful technique, limitations to the broad applicability of 2DGE involve relatively low dynamic range, whereby the detection of low abundance proteins is often variable. Also, solubility issues especially with complex mixtures of hydrophobic proteins (i.e. membrane associated) are problematic (41) and increase the incidence of gel-to-gel variability. These factors, among others, limit the utility of traditional 2DGE to studies of relative abundance versus absolute quantification (50). New gel-based technologies involving differential fluorescent labeling techniques such as those described for two-dimensional difference in gel electrophoresis (2D DIGE), expand the capabilities of traditional 2DGE and allow quantification of all protein spots relative to an internal standard (51). However, despite these quantitative advances 2DGE is often complemented with non-gel based proteomic approaches.

2.3.2 Liquid Chromatography

The successful resolution of complex protein samples dominantly relies on at least two independent physiochemical properties such as those described for 2DGE. Similar to the separation of proteins by iso-electric point and molecular weight, complex protein samples may be resolved in the post-digestion process at the peptide level. Pioneering work in this field was accomplished over a decade ago in the Yates laboratory (52), and has since become the leading sample processing technique for peptide identification via liquid chromatography mass spectrometry (LC-MS) (53,54). Modern trends in peptide LC are focused on improving peak capacity through ultra-high performance liquid chromatography systems as well improving the capacity of multi-dimensional chromatographic techniques to streamline the processing of highly crude fractions prior to their introduction into a mass spectrometer (8).

2.3.4 Multi-Dimensional Liquid Chromatography

In a 2DLC-based separation scheme, a protein mixture is digested by a protease (i.e. trypsin) and peptides are resolved in the first dimension over a strong-cation exchange (SCX) column, using increasing concentrations of salt in the mobile phase. These "salt-bumps" are then subsequently resolved over a C18 reverse phase (RP) column and directly introduced into the mass spectrometer for peptide identification (54). The two-dimensional approach can be tackled "offline" or "in-line" with the MS. In the offline method, the SCX and RP columns are setup on separate high-performance liquid chromatography (HPLC) systems. Fractions are collected manually and then undergo analysis in the MS. The LC-MS approach is advantageous because its peptide profiles are relatively unbiased, unlike gel-based systems in which membrane or other hydrophobic proteins are poorly represented. Also it is adaptable to multiple types of site-specific (or non-specific) enzymes, increasing the probability that the majority of proteins can be represented in the mass spectrometer by at least one peptide (6).

2.3.5 Affinity Chromatography

Affinity chromatography separates proteins on the basis of a reversible interaction between a protein (or group of proteins) and a specific ligand coupled to a chromatography matrix. For LC-MS based proteomics, affinity capture can be a powerful intermediate, purification tool prior to more precise methods such as RP-HPLC. Target proteins (or peptides) are selected out of complex mixtures using appropriate ligands and elution conditions. One example of affinity purification involves the purification of histidine-tagged recombinant proteins using immobilized metal affinity chromatography (55). Other examples are the use of streptavidin-conjugated resin to enrich for biotinylated molecules (56) or the enrichment of glycoproteins and polysaccharides with the hexose binding lectin, concavalin A (ConA) (57). Substances bound to either streptavidin or ConA resins are eluted off in the presence of competing ligand or by alterations in

pH. Eluted proteins are then subjected to protease digestion and peptides introduced into a mass spectrometer. They glycoproteome of M. tuberculosis was elucidated by first enriching mannosylated proteins through a ConA column purification prior to digestion with trypsin (58). As another example, the phosphoproteome of *M. tuberculosis* was characterized under a variety of environmental conditions and investigators were very successful at elucidating sites of phosphorylation for over 300 proteins. This study also employed the use of phospho-peptide affinity chromatography (59).

2.4 Mass Spectrometry

Mass spectrometry measures the mass-to-charge ratios (*m/z*) of ionized molecules and in MS-based proteomics, is used to determine the amino acid composition of peptides. A mass spectrometer is built from three compartments: an ion source, which converts sample molecules into ions, a mass analyzer that measures the *m/z* of the charged particles and a detector which measures the intensity of ions generated for each *m/z* value (3). The 2002 Nobel Prize in Chemistry was co-awarded for the development of two soft-ionization techniques -electrospray ionization (ESI) (60) and laser-desorption ionization (61), or matrix-assisted laser desorption ionization (MALDI). ESI is capable of inducing ions from complex mixtures of molecules in solution and generates multiply charged species of parent ions. The ability to ionize particles in solution makes ESI instrumentation readily suitable for liquid-samples such as those generated from LC separation methods or enzymatic digests. MALDI techniques do not introduce multiple charge states of ions and are capable of analyzing peptide mixtures that have not undergone prior separation over an LC (3). For mass analysis, the ion trap, time-of-flight (TOF), quadropole and Fourier transform ion cyclotron (FTCIR-MS) mass analyzers are the most common in use for proteomics research. Ion traps, first invented in the 1950's and 60's can be considered the workhorse for routine protein identification platforms. Ions are captured in a column of

quadropoles and subjected to further fragmentation to generate MS/MS spectra. Traditional linear ion traps have relatively low mass accuracy, however are very versatile, robust and can be used in combination with other mass analyzers (62). FTICR-MS analyzers are also trapping instruments, with extremely high mass accuracy, resolvation power and dynamic range. However their routine operation in proteomics research is limited due to the high cost of instrumentation and the need for a large magnet (63). For TOF analyzers, the time it takes for ions with similar charge to reach a detector is measured, and this value is proportional to their *m/z*. Thus large ions will move more slowly than small ions. TOFs are usually coupled with MALDI ionization methods whose spectra of parent (single charged) ions are simpler to visualize and have been termed "peptide mass fingerprints" (64). MS/MS fragmentation of MALDI-TOF spectra can be accomplished by separating two TOF sections with a collision cell ("TOF-TOF) (65) where amino acid sequences o peptides can be derived. TOF analyzers have high sensitivity, resolution and mass accuracy and can be coupled to an ESI ionization source (Q-TOF instruments). Using either ESI or MALDI-based instruments, raw spectra outputs are acquired and search algorithms interrogate comprehensive protein sequence databases, which are wholly dependent on the quality of available genome sequence information for a given organism. Whole-genome sequencing makes the provision of complete DNA/RNA sequence libraries possible and thus the quality of protein databases will only improve. The two most widely used search algorithms in the proteomics field are SEQUEST™ (66) and MASCOT™ (43). New architectures of powerful MS instrumentation based on triple-quadropole (QQQ) and Orbitrap technologies have drastically improved sensitivity and data acquisition speed and have evolved the utility of MS-based proteomics from characterization of protein identifications to robust, quantitative assay platforms (8).

2.4.1 Quantitative Mass Spectrometry in the Study of the *M. tuberculosis* Proteome

Techniques in quantitative mass spectrometry are either label-based or label-free. And facilitate the relative or absolute quantification of protein abundance changes among biological conditions. Several excellent reviews summarize the details of each labeling platform and their applicability in quantitative MS (6,8,67). Labeling techniques using radioactive isotopes were first used early in the days of 2DGE (45) and since then the use of stable isotopes to label cells is common practice (6), however for organisms with slow metabolic rates, such as *M. tuberculosis* these metabolic labeling approaches are not very efficient (68). Fortunately, chemical labeling techniques such as isotope-coded affinity tags (ICAT) and isobaric tags for relative and absolute quantitation (iTRAQ), as well as label-free spectral counting methods have proven to be very informative in comparing protein abundance differences between bacteria grown under various metabolic or environmental conditions, or characterizing protein factors that may be responsible for the hyper- or hypovirulent phenotypes observed for clinical isolates.

2.4.2 Chemical Labeling Strategies

ICAT technology first described in 1999 is a chemical labeling technique used to study the relative abundance of peptides (i.e. proteins) under two different experimental conditions (6). It is not based on gel separation of proteins and thus is a much more versatile tool when comparing whole proteomes. The ICAT reagents (one heavy with 8 deuteriums, and one light) modify cysteine residues of peptides produced in each biological sample. A drawback to chemical labeling with ICAT reagents is the reliance on cysteine-containing peptides. The genome of *M. tuberculosis* does not encode for very many cysteine residues (10) and thus limits the utility of this chemical labeling approach. Technologies such as iTRAQ or Tandem Mass Tag (TMT) labels are specific for free amines generated after the digestion of protein samples (8) and are thus more universally applicable. These assays also enable the simultaneous study of multiple

experimental conditions – 8-plex systems for iTRAQ and 6-plex systems for TMT. As can be expected the data analysis and statistical validation that is required for the most complex of experimental designs is very intense, often requiring the use of mixed modeling approaches to find statistically significant protein abundance changes (69).

One of the first uses of this labeling technology in the study of mycobacterial proteomics utilized isotope coded affinity tags (ICAT) to measure the differential abundance of large numbers (> 500) of *M. tuberculosis* proteins. Changes in the protein profile of bacteria in a state of NRP were observed using this approach. Characterization of proteins involved in small molecule metabolism, degradation and energy metabolism were observed to contribute to the unique protein profiles of *M. tuberculosis* in different metabolic states (27). The ICAT method relies on the labeling of cysteine residues of proteins. If the amino acid sequence of the gene product is devoid of cysteine residues, the labeling technique is not effective. Therefore, a less-biased approach would be needed to make comparison between proteomes. Mehaffy et al., utilized iTRAQ to characterize the protein variability between closely related clinical isolates of *M. tuberculosis*. Using this approach, over 100 proteins in either the secreted culture filtrate or cytosolic protein fraction demonstrated quantitative differences in protein abundance among the three isolates tested. Basic metabolic proteins such as those involved in the TCA cycle, nutrient acquisition and proteins related to cell wall and cellular processes were found to be the most significantly different under the conditions tested (69).

2.4.3 Label-free MS Quantitation

Label-free methods, as the name implies, do not utilize the incorporation of stable isotopes into protein samples, nor do they rely on chemical modification of peptides. This method of measurement is especially useful when comparing large datasets. Proteins can be directly measured using intensity-based methods or indirectly measured using peptide spectral

counts (8). The spectral count technique has high dynamic range and is sufficiently reproducible if care is taken in the experimental setup (70).

For *M. tuberculosis*, differences between hypo-virulent and hyper-virulent clinical isolates were quantified using a label-free spectral counting method. Similar to the iTRAQ study of Mehaffy, this study focused on identifying differences in virulence determinants between strains. Overall, 48 proteins were over-represented in the hypovirulent isolate, whereas 53 were over-represented in the hypervirulent. Functional classification of these proteins also identified proteins involved in cell wall organization and homeostasis and DNA transcription regulatory proteins. Also of note were the changes observed in the reduced repertoire of highly immunogenic proteins (Esx-like proteins and Mpt51) in the highly virulent strain (68).

Advanced, but highly specific label-free techniques such as those used for absolute quantification of proteins within complex samples (single-reaction monitoring or multiple-reaction monitoring) are becoming more widely utilized, especially in the identification of disease biomarkers. Applications of this technique afford the precise measurement of 100's to theoretically 1000's of specific proteins (71). This approach can also be used for comprehensive quantitative analysis measuring changes within primary metabolic pathways in cells grown or isolated from different metabolic states (72). Recently, a comprehensive set of proteins from the human pathogen *Leptospirosa interrogans* were monitored in 25 different disease states. This study demonstrated that in a single LC-MS/MS experiment around 5000 peptides, covering 1680 *L. interrogans* proteins can be consistently detected and quantified during dynamic changes in the proteome (73). The utilization of MRM technology for *M. tuberculosis* can be modeled based on the studies described above. For example, differences among clinically relevant strains in regards to their immunogenic proteome can now be absolutely quantified. For example, a study

initiated in our laboratory has utilized targeted MRM strategies to characterize the differential abundance of the highly immunogenic Ag85 complex among six clinical isolates of *M. tuberculosis* that represent three different genomic clades (74). This is important to have characterized, as the major component for many second-generation subunit vaccines include proteins from this repertoire of antigens. As was observed previously, phenotypically hyper or hypo-virulent strains may alter their antigenic repertoire, making the immune response to infection (and by extension vaccination) less effective.

2.5 Perspective on Chemical Proteomics & Drug Discovery for *M. tuberculosis*

Beyond descriptive analyses, mass spectrometry (MS) and the field of chemical proteomics are being applied to answer more direct questions of protein function and regulation of enzyme activity. In addition, chemical proteomics in the drug-discovery pipeline enables the identification of direct or indirect profiling of drug-target interactions and lends itself to large-scale selectivity profiling of hundreds of protein targets assayed against multiple drug chemotypes (75). Within the last decade, multiple chemoproteomic techniques and activity based protein probes have been developed for profiling inhibitors of serine hydrolases (76), kinases (77) and histone de-acetylases (78) in the human proteome. By utilizing classical drug affinity chromatography techniques coupled to mass spectrometry and bioinformatic analysis (5), investigation of drug profiles in the physiologically relevant proteome of the mycobacterial cell can be accomplished. This is especially useful for the characterization of new drugs with novel modes of action, such as has been demonstrated in the high-throughput screening studies of *M. tuberculosis*-specific kinase inhibitors (79-83). For ATP-small molecule inhibitors specifically (i.e. kinase inhibitors), chemoproteomic platforms utilize active-site specific probes that are designed to bind target proteins in competetion with either broad range or specific inhibitors

(84). These probes then facilitate the enrichment of ATP-binding proteins for profiling via MS . Inhibitor profiling has been successful in a wide number of disease applications (77,85,86), however none have investigated the utility of this approach for antimicrobial drug discovery. *M. tuberculosis* is one model organism for which these technologies can be adapted and where application of these techniques can have substantial impact on the identification of drug-target interactions and facilitate the development of novel small-molecule inhibitors for treatment of tuberculosis.

2.6 Rationale and Objectives

The advances described in the last several decades in the relatively new field of proteomics demonstrate the vast scope and potential this science has to offer, especially in the new era of systems biology-driven research (87,88). In contrast to genomic or transcriptomic studies, proteomics directly focuses on the characterization and quantification of specific gene products (6). For *M. tuberculosis*, the application of proteomic platforms to answer disease relevant questions related to the multi-stage infection profile of the bacillus is of great importance not only in the characterization of the physiology of the organism, but also in the detection of true *M. tuberculosis* infection in the absence of clinical symptoms (latent TB). Furthermore, this complicated disease lifecycle requires the identification and development of relevant progress-markers, measured during the course of antimicrobial therapy. Within the next several years, advances in TB diagnostic research will see the development of LC-MS based methods capable of challenging the accuracy and throughput of traditional immunoassays in the unambiguous identification and quantification of relevant target proteins isolated from biofluids, such as saliva or urine (8,89). Similarly, in the area of drug discovery, quantitative chemoproteomics platforms that isolate and characterize critical subproteomes can be utilized to

identify novel drug-target interactions of antimicrobial inhibitors in the treatment of tuberculosis. The work described in this dissertation applies a range of proteomic technologies and methods in the study of M. tuberculosis physiology and immunogenicity. We base these works in the hypothesis that the identification of new diagnostic reagents and drug targets can be gleaned from a global, whole proteome perspective of the mycobacterial cell envelope. The results of this study should contribute to the characterization of the mechanisms that are utilized to maintain bacterial survival within multiple disease states as well as the elucidation of novel epitopes unique to proteins of the cell wall.

2.7 LITERATURE CITED

1. Wilkins, M. R., Sanchez, J. C., Gooley, A. A., Appel, R. D., Humphery-Smith, I., Hochstrasser, D. F., and Williams, K. L. (1996) *Biotechnol Genet Eng Rev* **13**, 19-50
2. Shevchenko, A., Jensen, O. N., Podtelejnikov, A. V., Sagliocco, F., Wilm, M., Vorm, O., Mortensen, P., Boucherie, H., and Mann, M. (1996) *Proc Natl Acad Sci U S A* **93**, 14440-14445
3. Aebersold, R., and Mann, M. (2003) *Nature* **422**, 198-207
4. Schirle, M., Bantscheff, M., and Kuster, B. (2012) *Chem Biol* **19**, 72-84
5. Rix, U., and Superti-Furga, G. (2009) *Nat Chem Biol* **5**, 616-624
6. Ong, S. E., and Mann, M. (2005) *Nat Chem Biol* **1**, 252-262
7. Eng, J. K., Searle, B. C., Clauser, K. R., and Tabb, D. L. (2011) *Mol Cell Proteomics* **10**, R111 009522
8. Bantscheff, M., Lemeer, S., Savitski, M. M., and Kuster, B. (2012) *Anal Bioanal Chem* **404**, 939-965
9. Sonnenberg, M. G., and Belisle, J. T. (1997) *Infect Immun* **65**, 4515-4524
10. Cole, S. T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S. V., Eiglmeier, K., Gas, S., Barry, C. E., 3rd, Tekaia, F., Badcock, K., Basham, D., Brown, D., Chillingworth, T., Connor, R., Davies, R., Devlin, K., Feltwell, T., Gentles, S., Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., Krogh, A., McLean, J., Moule, S., Murphy, L., Oliver, K., Osborne, J., Quail, M. A., Rajandream, M. A., Rogers, J., Rutter, S., Seeger, K., Skelton, J., Squares, R., Squares, S., Sulston, J. E., Taylor, K., Whitehead, S., and Barrell, B. G. (1998) *Nature* **393**, 537-544
11. Urquhart, B. L., Atsalos, T. E., Roach, D., Basseal, D. J., Bjellqvist, B., Britton, W. L., and Humphery-Smith, I. (1997) *Electrophoresis* **18**, 1384-1392
12. Urquhart, B. L., Cordwell, S. J., and Humphery-Smith, I. (1998) *Biochem Biophys Res Commun* **253**, 70-79
13. Jungblut, P. R., Schaible, U. E., Mollenkopf, H. J., Zimny-Arndt, U., Raupach, B., Mattow, J., Halada, P., Lamer, S., Hagens, K., and Kaufmann, S. H. (1999) *Mol Microbiol* **33**, 1103-1117
14. Malen, H., Berven, F. S., Fladmark, K. E., and Wiker, H. G. (2007) *Proteomics* **7**, 1702-1718
15. Mawuenyega, K. G., Forst, C. V., Dobos, K. M., Belisle, J. T., Chen, J., Bradbury, E. M., Bradbury, A. R., and Chen, X. (2005) *Mol Biol Cell* **16**, 396-404
16. Wolfe, L. M., Mahaffey, S. B., Kruh, N. A., and Dobos, K. M. (2010) *J Proteome Res* **9**, 5816-5826
17. Schmidt, F., Donahoe, S., Hagens, K., Mattow, J., Schaible, U. E., Kaufmann, S. H., Aebersold, R., and Jungblut, P. R. (2004) *Mol Cell Proteomics* **3**, 24-42
18. Gu, S., Chen, J., Dobos, K. M., Bradbury, E. M., Belisle, J. T., and Chen, X. (2003) *Mol Cell Proteomics* **2**, 1284-1296
19. Sinha, S., Kosalai, K., Arora, S., Namane, A., Sharma, P., Gaikwad, A. N., Brodin, P., and Cole, S. T. (2005) *Microbiology* **151**, 2411-2419
20. Xiong, Y., Chalmers, M. J., Gao, F. P., Cross, T. A., and Marshall, A. G. (2005) *J Proteome Res* **4**, 855-861
21. Malen, H., De Souza, G. A., Pathak, S., Softeland, T., and Wiker, H. G. (2011) *BMC Microbiol* **11**, 18

22. Leversen, N. A., de Souza, G. A., Malen, H., Prasad, S., Jonassen, I., and Wiker, H. G. (2009) *Microbiology* **155**, 2375-2383
23. Malen, H., Berven, F. S., Softeland, T., Arntzen, M. O., D'Santos, C. S., De Souza, G. A., and Wiker, H. G. (2008) *Proteomics* **8**, 1859-1870
24. Zheng, J., Wei, C., Leng, W., Dong, J., Li, R., Li, W., Wang, J., Zhang, Z., and Jin, Q. (2007) *Proteomics* **7**, 3919-3931
25. Kruh, N. A., Troudt, J., Izzo, A., Prenni, J., and Dobos, K. M. (2010) *PLoS One* **5**, e13938
26. Muttucumaru, D. G., Roberts, G., Hinds, J., Stabler, R. A., and Parish, T. (2004) *Tuberculosis (Edinb)* **84**, 239-246
27. Cho, S. H., Goodlett, D., and Franzblau, S. (2006) *Tuberculosis (Edinb)* **86**, 445-460
28. Sherman, D. R., Voskuil, M., Schnappinger, D., Liao, R., Harrell, M. I., and Schoolnik, G. K. (2001) *Proc Natl Acad Sci U S A* **98**, 7534-7539
29. Betts, J. C., Lukey, P. T., Robb, L. C., McAdam, R. A., and Duncan, K. (2002) *Mol Microbiol* **43**, 717-731
30. Fisher, M. A., Plikaytis, B. B., and Shinnick, T. M. (2002) *J Bacteriol* **184**, 4025-4032
31. Voskuil, M. I., Visconti, K. C., and Schoolnik, G. K. (2004) *Tuberculosis (Edinb)* **84**, 218-227
32. Nagai, S., Wiker, H. G., Harboe, M., and Kinomoto, M. (1991) *Infect Immun* **59**, 372-382
33. Lee, B. Y., and Horwitz, M. A. (1995) *J Clin Invest* **96**, 245-249
34. Monahan, I. M., Betts, J., Banerjee, D. K., and Butcher, P. D. (2001) *Microbiology* **147**, 459-471
35. Mattow, J., Siejak, F., Hagens, K., Becher, D., Albrecht, D., Krah, A., Schmidt, F., Jungblut, P. R., Kaufmann, S. H., and Schaible, U. E. (2006) *Proteomics* **6**, 2485-2494
36. Harboe, M., Nagai, S., Patarroyo, M. E., Torres, M. L., Ramirez, C., and Cruz, N. (1986) *Infect Immun* **52**, 293-302
37. Efthimiou, P., and Sood, S. (2007) *Ann Rheum Dis* **66**, 276
38. Mazurek, G. H., Jereb, J., Lobue, P., Iademarco, M. F., Metchock, B., and Vernon, A. (2005) *MMWR Recomm Rep* **54**, 49-55
39. Mattow, J., Schaible, U. E., Schmidt, F., Hagens, K., Siejak, F., Brestrich, G., Haeselbarth, G., Muller, E. C., Jungblut, P. R., and Kaufmann, S. H. (2003) *Electrophoresis* **24**, 3405-3420
40. He, X. Y., Zhuang, Y. H., Zhang, X. G., and Li, G. L. (2003) *Microbes Infect* **5**, 851-856
41. Gygi, S. P., Corthals, G. L., Zhang, Y., Rochon, Y., and Aebersold, R. (2000) *Proc Natl Acad Sci U S A* **97**, 9390-9395
42. Yates, J. R., 3rd, Eng, J. K., McCormack, A. L., and Schieltz, D. (1995) *Anal Chem* **67**, 1426-1436
43. Perkins, D. N., Pappin, D. J., Creasy, D. M., and Cottrell, J. S. (1999) *Electrophoresis* **20**, 3551-3567
44. O'Farrell, P. H. (1975) *J Biol Chem* **250**, 4007-4021
45. Snyder, E. L., Dunn, B. E., Giometti, C. S., Napychank, P. A., Tandon, N. N., Ferri, P. M., and Hofmann, J. P. (1987) *Transfusion* **27**, 335-341
46. Dunn, B. E., Blaser, M. J., and Snyder, E. L. (1987) *Infect Immun* **55**, 1564-1572
47. Gooley, A. A., Ou, K., Russell, J., Wilkins, M. R., Sanchez, J. C., Hochstrasser, D. F., and Williams, K. L. (1997) *Electrophoresis* **18**, 1068-1072

48. Aebersold, R. H., Leavitt, J., Saavedra, R. A., Hood, L. E., and Kent, S. B. (1987) *Proc Natl Acad Sci U S A* **84**, 6970-6974
49. Rosenfeld, J., Capdevielle, J., Guillemot, J. C., and Ferrara, P. (1992) *Anal Biochem* **203**, 173-179
50. Rabilloud, T. (2002) *Proteomics* **2**, 3-10
51. Marouga, R., David, S., and Hawkins, E. (2005) *Anal Bioanal Chem* **382**, 669-678
52. Link, A. J., Eng, J., Schieltz, D. M., Carmack, E., Mize, G. J., Morris, D. R., Garvik, B. M., and Yates, J. R., 3rd. (1999) *Nat Biotechnol* **17**, 676-682
53. Washburn, M. P., Wolters, D., and Yates, J. R., 3rd. (2001) *Nat Biotechnol* **19**, 242-247
54. Wolters, D. A., Washburn, M. P., and Yates, J. R., 3rd. (2001) *Anal Chem* **73**, 5683-5690
55. Porath, J. (1992) *Protein Expr Purif* **3**, 263-281
56. Gretch, D. R., Suter, M., and Stinski, M. F. (1987) *Anal Biochem* **163**, 270-277
57. Hunt, L. A. (1982) *Biochem J* **205**, 623-630
58. Gonzalez-Zamorano, M., Mendoza-Hernandez, G., Xolalpa, W., Parada, C., Vallecillo, A. J., Bigi, F., and Espitia, C. (2009) *J Proteome Res* **8**, 721-733
59. Prisic, S., Dankwa, S., Schwartz, D., Chou, M. F., Locasale, J. W., Kang, C. M., Bemis, G., Church, G. M., Steen, H., and Husson, R. N. (2010) *Proc Natl Acad Sci U S A* **107**, 7521-7526
60. Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F., and Whitehouse, C. M. (1989) *Science* **246**, 64-71
61. Karas, M., and Hillenkamp, F. (1988) *Anal Chem* **60**, 2299-2301
62. Douglas, D. J., Frank, A. J., and Mao, D. (2005) *Mass Spectrom Rev* **24**, 1-29
63. Marshall, A. G., Hendrickson, C. L., and Jackson, G. S. (1998) *Mass Spectrom Rev* **17**, 1-35
64. Thiede, B., Hohenwarter, W., Krah, A., Mattow, J., Schmid, M., Schmidt, F., and Jungblut, P. R. (2005) *Methods* **35**, 237-247
65. Shevchenko, A., Loboda, A., Ens, W., and Standing, K. G. (2000) *Anal Chem* **72**, 2132-2141
66. Eng, J., McCormack, A., and Yates, J. (1994) *Journal of the American Society for Mass Spectrometry* **5**, 976-989
67. Kline, K. G., Finney, G. L., and Wu, C. C. (2009) *Brief Funct Genomic Proteomic* **8**, 114-125
68. de Souza, G. A., Fortuin, S., Aguilar, D., Pando, R. H., McEvoy, C. R., van Helden, P. D., Koehler, C. J., Thiede, B., Warren, R. M., and Wiker, H. G. (2010) *Mol Cell Proteomics* **9**, 2414-2423
69. Mehaffy, C., Hess, A., Prenni, J. E., Mathema, B., Kreiswirth, B., and Dobos, K. M. (2010) *Proteomics*
70. Liu, H., Sadygov, R. G., and Yates, J. R., 3rd. (2004) *Anal Chem* **76**, 4193-4201
71. Malmstrom, J., Beck, M., Schmidt, A., Lange, V., Deutsch, E. W., and Aebersold, R. (2009) *Nature* **460**, 762-765
72. Costenoble, R., Picotti, P., Reiter, L., Stallmach, R., Heinemann, M., Sauer, U., and Aebersold, R. (2011) *Mol Syst Biol* **7**, 464
73. Schmidt, A., Beck, M., Malmstrom, J., Lam, H., Claassen, M., Campbell, D., and Aebersold, R. (2011) *Mol Syst Biol* **7**, 510
74. Kruh-Garcia, N. A. (in preparation)
75. Schirle, M., Bantscheff, M., and Kuster, B. *Chem Biol* **19**, 72-84

76. Patricelli, M. P., Giang, D. K., Stamp, L. M., and Burbaum, J. J. (2001) *Proteomics* **1**, 1067-1071
77. Bantscheff, M., Eberhard, D., Abraham, Y., Bastuck, S., Boesche, M., Hobson, S., Mathieson, T., Perrin, J., Raida, M., Rau, C., Reader, V., Sweetman, G., Bauer, A., Bouwmeester, T., Hopf, C., Kruse, U., Neubauer, G., Ramsden, N., Rick, J., Kuster, B., and Drewes, G. (2007) *Nat Biotechnol* **25**, 1035-1044
78. Salisbury, C. M., and Cravatt, B. F. (2007) *Proc Natl Acad Sci U S A* **104**, 1171-1176
79. Lougheed, K. E., Osborne, S. A., Saxty, B., Whalley, D., Chapman, T., Bouloc, N., Chugh, J., Nott, T. J., Patel, D., Spivey, V. L., Kettleborough, C. A., Bryans, J. S., Taylor, D. L., Smerdon, S. J., and Buxton, R. S. (2011) *Tuberculosis (Edinb)* **91**, 277-286
80. Magnet, S., Hartkoorn, R. C., Szekely, R., Pato, J., Triccas, J. A., Schneider, P., Szantai-Kis, C., Orfi, L., Chambon, M., Banfi, D., Bueno, M., Turcatti, G., Keri, G., and Cole, S. T. (2010) *Tuberculosis (Edinb)* **90**, 354-360
81. Mak, P. A., Rao, S. P., Ping Tan, M., Lin, X., Chyba, J., Tay, J., Ng, S. H., Tan, B. H., Cherian, J., Duraiswamy, J., Bifani, P., Lim, V., Lee, B. H., Ling Ma, N., Beer, D., Thayalan, P., Kuhen, K., Chatterjee, A., Supek, F., Glynne, R., Zheng, J., Boshoff, H. I., Barry, C. E., 3rd, Dick, T., Pethe, K., and Camacho, L. R. (2012) *ACS Chem Biol* **7**, 1190-1197
82. Reynolds, R. C., Ananthan, S., Faaleolea, E., Hobrath, J. V., Kwong, C. D., Maddox, C., Rasmussen, L., Sosa, M. I., Thammasuvimol, E., White, E. L., Zhang, W., and Secrist, J. A., 3rd. (2011) *Tuberculosis (Edinb)* **92**, 72-83
83. Triola, G., Wetzel, S., Ellinger, B., Koch, M. A., Hubel, K., Rauh, D., and Waldmann, H. (2009) *Bioorg Med Chem* **17**, 1079-1087
84. Drewes, G. (2012) *Methods Mol Biol* **803**, 15-21
85. Fischer, J. J., Michaelis, S., Schrey, A. K., Graebner, O. G., Glinski, M., Dreger, M., Kroll, F., and Koester, H. (2010) *Toxicol Sci* **113**, 243-253
86. Li, J., Rix, U., Fang, B., Bai, Y., Edwards, A., Colinge, J., Bennett, K. L., Gao, J., Song, L., Eschrich, S., Superti-Furga, G., Koomen, J., and Haura, E. B. (2010) *Nat Chem Biol* **6**, 291-299
87. Nagaraj, N., Kulak, N. A., Cox, J., Neuhauser, N., Mayr, K., Hoerning, O., Vorm, O., and Mann, M. *Mol Cell Proteomics* **11**, M111 013722
88. Nagaraj, N., Wisniewski, J. R., Geiger, T., Cox, J., Kircher, M., Kelso, J., Paabo, S., and Mann, M. *Mol Syst Biol* **7**, 548
89. Whiteaker, J. R., Zhao, L., Lin, C., Yan, P., Wang, P., and Paulovich, A. G. (2011) *Mol Cell Proteomics* **11**, M111 015347

III. CHAPTER III - Proteomic Definition of the Cell Wall of *Mycobacterium tuberculosis*

Partially presented in Lisa M. Wolfe, Spencer B. Mehaffy, Nicole A. Kruh, and Karen M. Dobos. 2010. Proteomic Definition of the Cell Wall of *Mycobacterium tuberculosis*. Journal of Proteome Research. Nov 5;9(11):5816-26. Epub 2010 Sep 29.

3.1 Introduction

The cell wall of *M. tuberculosis* is a complex and diverse reservoir of proteins intermingled in a matrix of peptidoglycan, mycolic acids, lipids and carbohydrates. Collectively, the localization and secretion of these molecules into the extracellular milieu both stimulates (1-4) and represses (5-7) the host immune system, ultimately facilitating the survival of the bacillus throughout active infection and during stages of non-replicating persistance (8). Immunologically, resident and secreted proteins of the cell wall are known to mediate specific T cell responses, (9-12). The molecules and proteins of the cell wall also provide active means of antimicrobial resistance (13-15), and are generally associated with the critical processes of bacterial survival and immune modulation. The need for modern tools in the fight against tuberculosis necessitates the study and exploitation of the physiological interplay between the pathogen and host. In this study, we hypothesized that the cell wall proteome of *M. tuberculosis* was highly complex and that protein species associated with the highly insoluble subcellular fraction could lead to functional characterization of a subset of proteins that may play a role in host-pathogen interactions, facilitate T cell mediated immunity, and can be exploited in the identification of novel drug targets.

3.1.1 *M. tuberculosis* Cell Wall

For decades the macromolecular features of the mycobacterial cell wall, including the mycolic acid and arabinogalactan core, have been studied in detail (16). Structurally, the cell wall of *M. tuberculosis* is comprised of a distinct inner core of mycolic acid-arabinogalactan-peptidoglycan (mAGP). In addition to the covalently attached lipids and carbohydrates, it is well known that the free lipids, lipoglycans and phosphotidyl inositols that reside in the outer core of the cell wall play key roles in modulation of the host immune response (5). Specifically, the molecules, lipoarabinomannan (LAM), lipomannan (LM) and phosphatidyl inositol mannoside (PIM) are known to aid the process of host immune evasion (17). In addition, virulence lipids such as trehalose dimycolate/monomycolate (TDM/TMM), phthiocerol dimycocerosate (PDIM) and sulfolipids (SL) and the protein machinery, such as MmpLs, required for their export are intercalated in the cell wall (18). Numerous cell wall associated proteins, including many lipo and lipoglycoproteins, have also been described (19-21). For example, the TLR2 agonists, lpqH (19kDa), phoS1 (38kDa) and lprG (Rv14llc) are all found in the cell wall (20-23) (24), where they function to regulate the action of macrophages and dendritic cells (25). PhoS1 also plays a role in bacterial escape from the host macrophage through apoptosis (26). Many other lipoproteins of unknown function are identified in the secreted proteome of *M. tuberculosis* culture filtrate (27-29).

3.1.2 *M. tuberculosis* Proteomics

The study of the secreted proteome of *M. tuberculosis* was driven by the search for novel immunodominant antigens, drug targets and biomarkers for disease (30-33). To build upon this work, a number of studies employed advances in proteomic technologies such as two-dimensional gel electrophoresis (2DGE) and liquid chromatography mass spectrometry (LC-MS/MS) to further mine additional subcellular compartments derived from the bacillus, such as

the intracellular cytosol, membrane and cell wall (34-39). More recently, the proteomic definition of the bacillus in the context of the granuloma revealed the majority of the secreted proteome to be strikingly absent from the tubercle lesions (40). This finding could indicate either a rapid and efficient trafficking of the antigens in to the periphery or a rapid and efficient degradation of secreted protein in the host milieu. Bacterial proteins that were most dominant *in vivo* were those predicted to reside in the *M. tuberculosis* cellular envelope, indicating their overall importance in the establishment and maintenance of infection. Historically, the proteins within the cell wall have been difficult to resolve and identify by traditional 2DGE methods (32,41). Mawuenyega et al., employing two-dimensional liquid chromatography coupled with mass spectrometry (2DLC/MS) were the most successful at defining the cell wall proteome with the identification of 306 proteins (37). In this study, we set out to comprehensively describe the *M. tuberculosis* cell wall proteome in an effort to exploit additional proteins that may play a role in host-pathogen interactions and define new potential drug targets via discovery of unique biosynthetic or metabolic processes.

3.2 Results

3.2.1 Identification of Cell Wall Proteins

To maximize the resolution and protein identification of cell wall proteins, the delipidated cell wall of *M. tuberculosis* was subjected to differential detergent extraction to maximize the solubility of a wide range of proteins. Guanidine hydrochloride (GuHCl) acts as a denaturant, sodium dodecyl sulfate (SDS) is a cationic detergent and Triton X-114 (TX-114) is a nonionic detergent. Proteins were resolved via two-dimensional gel electrophoresis (2DGE) using the experimental outline as depicted in Figure 3.1.

In comparison with the SDS and GuHCl extracts, the 2DGE separation was insufficient in resolving a significant number of proteins in the TX-114 detergent extract (Figure 3.2).

This was most likely due to its increased hydrophobic content (42,43). To better resolve the proteins in this fraction, a multi-dimensional chromatography approach was employed to enhance the resolution of any unique or uncharacterized protein families within the TX-114 subset. From the 2DGE analysis, 290 proteins were identified with 122, 131 and 37 proteins identified in the GuHCl, SDS and TX-114 fractions respectively. For all fractions, over half were unique to each subset. Chromatographic separation via strong cation exchange (SCX), combined with reverse phase chromatography of a trypsin digest of this CWP preparation, allowed the resolution of 364 proteins, including an additional 294 proteins not found in any of the 2DGE

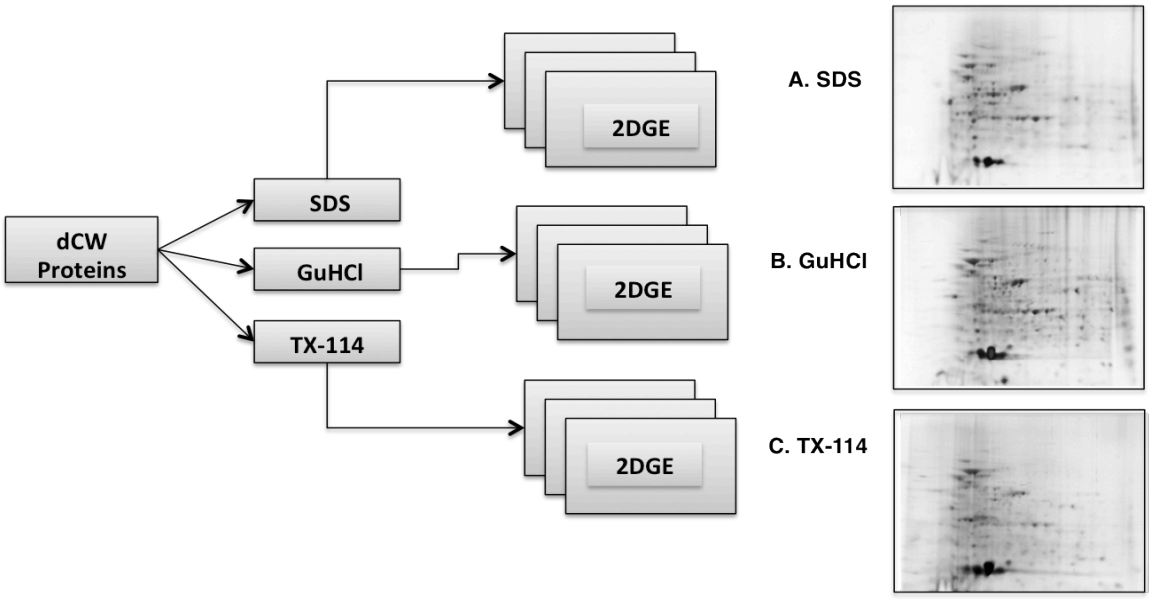

Figure 3.1 - Schematic Diagram of Cell Wall Proteome Workflow. The cell wall fraction was subjected to extraction with organic solvents to remove non-covalently bound lipids and lipoglycans and remaining proteins were either denatured in guanidine hydrochloride or solubilized with cationic (Sodium Dodecyl Sulfate, SDS) or nonionic (Triton X-114) surfactants. These extracts were resolved by 2DGE and spots analyzed by mass spectrometry for protein identification.

preparations (Figure 3.3). Combining the data sets of both gel and non-gel based protein IDs led to the detection of 528 proteins, which are listed in Table form in Appendix I.

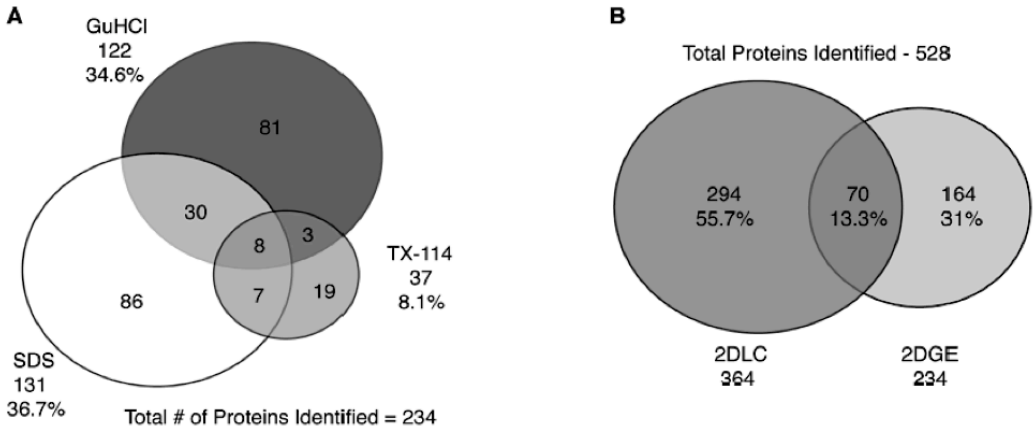

Figure 3.2 - Identification of Cell Wall Proteins by Two-Dimensional Gel Electrophoresis and Two-Dimensional Liquid Chromatography. A. Detergent extraction and separation of cell wall protein subsets via 2DGE demonstrated the majority of proteins to be resolved using SDS and GuHCl. The lipoprotein –rich TX-114 fraction was less amenable to gel separation. See text for details. B. Cumulatively, separation of cell wall proteins by 2DGE resulted in the identification of 234 proteins. Separating the TX-114 protein subset by multidimensional chromatography (2DLC-MS/MS) resolved an additional 294 proteins.

One hundred and five of the 528 cell wall proteins identified were not reported in previous *M. tuberculosis* proteome publications at the time of manuscript submission (29,32,35,37,41), including two comprehensive proteomic databases http://www.ssi.dk/sw14644.asp (41) and http://web.mpiib-berlin.mpg.de/cgi-bin/pdbs/2d-page/extern/index.cgi (44) (Table 3.1).

Table 3.1 – Cell Wall Proteins Unique to this Study

Rv no.	protein name	protein function	class ID TUBERCULIST	class ID SANGER	CWP fraction[a]
Rv0007	Rv0007	POSSIBLE CONSERVED MEMBRANE PROTEIN	3	V	D
Rv0014c	pknB	TRANSMEMBRANE SERINE/THREONINE-PROTEIN KINASE B PKNB (PROTEIN KINASE B)	9	I.J.3	D
Rv0018c	pstP	POSSIBLE SERINE/THREONINE PHOSPHATASE PSTP	9	I.J.3	D
Rv0092	ctpA	PROBABLE CATION TRANSPORTER P-TYPE ATPASE A CTPA	3	III.A.2	B
Rv0142	Rv0142	CONSERVED HYPOTHETICAL PROTEIN	10	VI	A, B
Rv0327c	cyp135A1	POSSIBLE CYTOCHROME P450 135A1 CYP135A1	7	IV.F	B
Rv0361	Rv0361	PROBABLE CONSERVED MEMBRANE PROTEIN	3	II.C.5	D
Rv0431	Rv0431	PUTATIVE TUBERCULIN RELATED PEPTIDE	3	II.C.2	D
Rv0497	Rv0497	PROBABLE CONSERVED TRANSMEMBRANE PROTEIN	3	II.C.5	D
Rv0512	hemB	PROBABLE DELTA-AMINOLEVULINIC ACID DEHYDRATASE HEMB	7	I.G.12	C
Rv0614	Rv0614	CONSERVED HYPOTHETICAL PROTEIN	10	V	A
Rv0622	Rv0622	POSSIBLE MEMBRANE PROTEIN	3	II.C.5	B
Rv0638	secE1	PROBABLE PREPROTEIN TRANSLOCASE SECE1	3	III.D	D
Rv0712	Rv0712	CONSERVED HYPOTHETICAL PROTEIN	10	V	B
Rv0771	Rv0771	POSSIBLE 4-CARBOXYMUCONOLACTONE DECARBOXYLASE (CMD)	7	II.B.6	D
Rv0858c	Rv0858c	PROBABLE AMINOTRANSFERASE	7	IV.H	A
Rv0892	Rv0892	PROBABLE MONOOXYGENASE	7	I.B.7	B
Rv0899	ompA	OUTER MEMBRANE PROTEIN A OMPA	3	II.C.2	D
Rv0902c	prrB	TWO COMPONENT SENSOR HISTIDINE KINASE PRRB	9	I.J.2	D
Rv0926c	Rv0926c	CONSERVED HYPOTHETICAL PROTEIN	10	V	A
Rv0954	Rv0954	PROBABLE CONSERVED TRANSMEMBRANE PROTEIN	3	II.C.2	D
Rv1016c	lpqT	PROBABLE CONSERVED LIPOPROTEIN LPQT	3	II.C.1	D
Rv1029	kdpA	PROBABLE POTASSIUM-TRANSPORTING ATPASE A CHAIN KDPA	3	III.A.2	A
Rv1096	Rv1096	POSSIBLE GLYCOSYL HYDROLASE	7	I.A.1	D
Rv1100	Rv1100	CONSERVED HYPOTHETICAL PROTEIN	10	V	D
Rv1128c	Rv1128c	CONSERVED HYPOTHETICAL PROTEIN	5	IV.B.2	B
Rv1151c	Rv1151c	PROBABLE TRANSCRIPTIONAL REGULATORY PROTEIN	9	I.J.1	A
Rv1184c	Rv1184c	POSSIBLE EXPORTED PROTEIN	3	V	D
Rv1196	PPE18	PPE FAMILY PROTEIN	6	IV.C.2	D
Rv1209	Rv1209	CONSERVED HYPOTHETICAL PROTEIN	10	V	D
Rv1230c	Rv1230c	POSSIBLE MEMBRANE PROTEIN	3	II.C.5	B
Rv1256c	cyp130	PROBA BLE CYTOCHROME P450 130 CYP130	7	IV.F	B
Rv1266c	pknH	PROBABLE TRANSMEMBRANE SERINE/THREONINE-PROTEIN KINASE H PKNH (STPK H)	9	I.J.3	B,D
Rv1273c	Rv1273c	PROBABLE DRUGS-TRANSPORT TRANSMEMBRANE ATP-BINDING PROTEIN ABC TRANSPORTER	3	II.C.5	B
Rv1330c	Rv1330c	CONSERVED HYPOTHETICAL PROTEIN	10	V	B
Rv1361c	PPE19	PPE FAMILY PROTEIN	6	IV.C.2	A
Rv1379	pyrR	PROBABLE PYRIMIDINE OPERON REGULATORY PROTEIN PYRR	9	I.J.1	D
Rv1393c	Rv1393c	PROBABLE MONOOXYGENASE	7	I.B.7	C
Rv1442	bisC	PROBABLE BIOTIN SULFOLIDE REDUCTASE BISC (BDS reductase) (BSO reductase)	7	I.G.1	A,D
Rv1450c	PE_PGRS27	PE-PGRS FAMILY PROTEIN	6	IV.C.1.b	A
Rv1451	ctaB	PROBABLE CYTOCHROME C OXIDASE ASSEMBLY FACTOR CTAB	7	I.B.6.a	B
Rv1466	Rv1466	CONSERVED HYPOTHETICAL PROTEIN	10	V	D
Rv1586c	Rv1586c	PROBABLE PHIRV1 INTEGRASE	5	IV.B.3	D
Rv1594	nadA	PROBABLE QUINOLINATE SYNTHETASE NADA	7	I.G.7	B
Rv1599	hisD	PROBABLE HISTIDINOL DEHYDROGENASE HISD (HDH)	7	I.D.5	D
Rv1602	hisH	PROBABLE AMIDOTRANSFERASE HISH	7	I.D.5	B
Rv1623c	cydA	PROBABLE INTEGRAL MEMBRANE CYTOCHROME D UBIQUINOL 0XIDASE (SUBUNIT I) CYDA	7	I.B.6.c	D
Rv1659	argH	PROBABLE ARGININOSUCCINATE LYASE ARGH	7	I.D.1	D
Rv1666c	cyp139	PROBABLE CYTOCHROME P450 139 CYP139	7	IV.F	A
Rv1748	Rv1748	HYPOTHETICAL PROTEIN	16	VI	D
Rv1777	cyp144	PROBABLE CYTOCHROME P450 144 CYP144	7	IV.F	A
Rv1801	PPE29	PPE FAMILY PROTEIN	6	IV.C.2	A
Rv1836c	Rv1836c	CONSERVED HYPOTHETICAL PROTEIN	10	V	D
Rv1957	Rv1957	HYPOTHETICAL PROTEIN	16	VI	D
Rv1969	mce3D	MCE-FAMILY PROTEIN MCE3D	0	IV.A	A
Rv2047c	Rv2047c	CONSERVED HYPOTHETICAL PROTEIN	10	V	D
Rv2063c	Rv2063c	HYPOTHETICAL PROTEIN	0	VI	B
Rv2095c	Rv2095c	CONSERVED HYPOTHETICAL PROTEIN	10	V	D
Rv2122c	hisE	PROBABLE PHOSPHORIBOSYL-AMP PYROPHOSPHATASE HISE	7	I.D.5	D
Rv2171	lppM	PROBABLE CONSERVED LIPOPROTEIN LPPM	3	II.C.1	D
Rv2198c	mmpS3	PROBABLE CONSERVED MEMBRANE PROTEIN MMPS3	3	II.C.4	D
Rv2205c	Rv2205c	CONSERVED HYPOTHETICAL PROTEIN	10	V	B
Rv2211c	gcvT	PROBABLE AMINOMETHYLTRANSFERASE GCVT (GLYCINE CLEAVAGE SYSTEM T PROTEIN)	7	I.C.1	D
Rv2223c	Rv2223c	PROBABLE EXPORTED PROTEASE	3	II.C.2	B
Rv2247	accD6	ACETYL/PROPIONYL-CoA CARBOXLASE (BETA SUBUNIT) ACCD6	1	I.H.1	A,D

Table 1. Continued

Rv no.	protein name	protein function	class ID TUBERCULIST	class ID SANGER	CWP fraction[a]
Rv2260	Rv2260	CONSERVED HYPOTHETICAL PROTEIN	10	V	D
Rv2307c	Rv2307c	CONSERVED HYPOTHETICAL PROTEIN	10	V	B
Rv2466c	Rv2466c	CONSERVED HYPOTHETICAL PROTEIN	10	V	D
Rv2518c	lppS	PROBABLE CONSERVED LIPOPROTEIN LPPS	3	II.C.1	D
Rv2556c	Rv2556c	CONSERVED HYPOTHETICAL PROTEIN	10	V	D
Rv2712c	Rv2712c	HYPOTHETICAL PROTEIN	16	VI	A
Rv2720	leXA	REPRESSOR LEXA	9	I.J.1	D
Rv2789c	fadE21	PROBABLE ACYL-CoA DEHYDROGENASE FADE21	1	I.A.3	A
Rv2864c	Rv2864c	POSSIBLE PENICILLIN-BINDING LIPOPROTEIN	3	II.C.3	B
Rv2880c	Rv2880c	CONSERVED HYPOTHETICAL PROTEIN	10	V	A
Rv2881c	cdsA	PROBABLE INTEGRAL MEMBRANE PHOSPHATIDATE CYTIDYLYLTRANSFERASE	1	I.H.3	C
Rv2931	ppsA	PHENOLPTHIOCEROL SYNTHESIS TYPE-I POLYKETIDE SYNTHASE PPSA	1	I.I	A
Rv2936	drrA	PROBABLE DAUNORUBICIN-DIM-TRANSPORT ATP-BINDING PROTEIN ABC TRANSPORTER DRRA	3	III.A.6	D
Rv2938	drrC	PROBABLE DAUNORUBICIN-DIM-TRANSPORT ABC TRANSPORTER DRRC	3	III.A.6	A,B
Rv3012c	gatC	PROBABLE GLUTAMYL-TRNA(GLN) AMIDOTRANSFERASE (GLU-ADT SUBUNIT C)	2	II.A.3	D
Rv3083	Rv3083	PROBABLE MONOOXYGENASE (HYDROYLASE)	7	I.B.7	C
Rv3086	adhD	PROBABLE ZINC-TYPE ALCOHOL DEHYDROGENASE ADHD (ALDEHYDE REDUCTASE)	7	I.B.7	B
Rv3107c	agpS	POSSIBLE ALKYLDIHYDROXYACETONEPHOSPHATE SYNTHASE AGPS (ALKYL-DHAP SYNTHASE)	1	I.B.7	B
Rv3136	PPE51	PPE FAMILY PROTEIN	6	IV.C.2	D
Rv3206c	moeB1	PROBABLE MOLYBDENUM COFACTOR BIOSYNTHESIS PROTEIN(MPT-SYNTHASE SULFURYLASE)	7	I.G.4	D
Rv3211	rhlE	PROBABLE ATP-DEPENDENT RNA HELICASE RHLE	2	II.A.7	D
Rv3243c	Rv3243c	HYPOTHETICAL PROTEIN	16	VI	A
Rv3281	Rv3281	CONSERVED HYPOTHETICAL PROTEIN	10	V	D
Rv3282	Rv3282	CONSERVED HYPOTHETICAL PROTEIN	10	V	A
Rv3373	echA18	PROBABLE ENOYL-CoA HYDRATASE(ENOYL HYDRASE) (UNSATURATED ACYL-CoA HYDRATASE)	1	I.A.3	A
Rv3416	whiB3	TRANSCRIPTIONAL REGULATORY PROTEIN WHIB-LIKE WHIB3	9	I.J.1	A
Rv3421c	Rv3421c	CONSERVED HYPOTHETICAL PROTEIN	10	V	C
Rv3448	Rv3448	PROBABLE CONSERVED INTEGRAL MEMBRANE PROTEIN	3	II.C.5	D
Rv3496c	mce4D	MCE-FAMILY PROTEIN MCE4D	0	IV.A	C
Rv3515c	fadD19	PROBABLE FATTY-ACID-CoA LIGASE FADD19 (FATTY-ACID-CoA SYNTHETASE)	1	I.A.3	B
Rv3534c	Rv3534c	PROBABLE 4-HYDROXY-2-OXOVALERATE ALDOLASE (HOA)	7	II.B.6	C
Rv3565	aspB	POSSIBLE ASPARTATE AMINOTRANSFERASE ASPB (TRANSAMINASE A) (ASPAT)	7	I.D.2	B
Rv3615c	Rv3615c	CONSERVED HYPOTHETICAL PROTEIN	10	V	D
Rv3651	Rv3651	CONSERVED HYPOTHETICAL PROTEIN	10	VI	D
Rv3666c	dppA	PROBABLE PERIPLASMIC DIPEPTIDE-BINDING LIPOPROTEIN DPPA	3	III.A.1	C
Rv3688c	Rv3688c	CONSERVED HYPOTHETICAL PROTEIN	10	V	D
Rv3690	Rv3690	PROBABLE CONSERVED MEMBRANE PROTEIN	3	VI	D
Rv3852	Hns	POSSIBLE HISTONE-LIKE PROTEIN HNS	2	II.A.4	D
Rv3877	Rv3877	PROBABLE CONSERVED TRANSMEMBRANE PROTEIN	3	V	B
Rv3903c	Rv3903c	HYPOTHETICAL ALANINE AND PROLINE RICH PROTEIN	16	VI	B

[a] Cell Wall Protein fractions: A, 2% SDS; B, 6M GuHCl; C, 4% TX-114; D, 2DLC TX-114.

Tuberculist Class IDs :

0 - Virulence, detoxification, adaptation
1 – Lipid metabolism
2 – Information pathways
3 – Cell wall and processes
4 – Stable RNAs
5 – Insertion sequences
6 – PE/PPE
7 – Intermediary Metabolism
8 – Unknown
9 – Regulatory Proteins
10 – Conserved Hypotheticals
16 – Hypothetical Protein, not in *M, bovis BCG*. (category removed in 2011)

3.2.2 Functional Annotation of Cell Wall Proteins

All identified proteins were grouped by functional category as defined by Institute Pasteur (45), demonstrating *M. tuberculosis* protein families present in each extract. Proteins in categories 3 (cell wall and cell wall processes) and 7 (intermediary metabolism) were consistently overrepresented among the CWP preparations (Figure 3.3). These were further classified into functional groups as defined by the Sanger Institute (Figure 3a-c). The majority of proteins were classified in either Category I – Small Molecule Metabolism (35%) or Category II – Macromolecule synthesis and degradation (25%) (Figure 3.4). Subclasses of category I showed roughly 5% of CW proteins sorting into small molecule degradation (I.A), 11% sorting into small molecule metabolism for energy (I.B), and the remaining ~19% classified as other (I.X), which included the classes: Central intermediary metabolism (I.C) and Amino acid biosynthesis (I.D). Proteins sorted into Category II, Macromolecule synthesis and degradation demonstrated

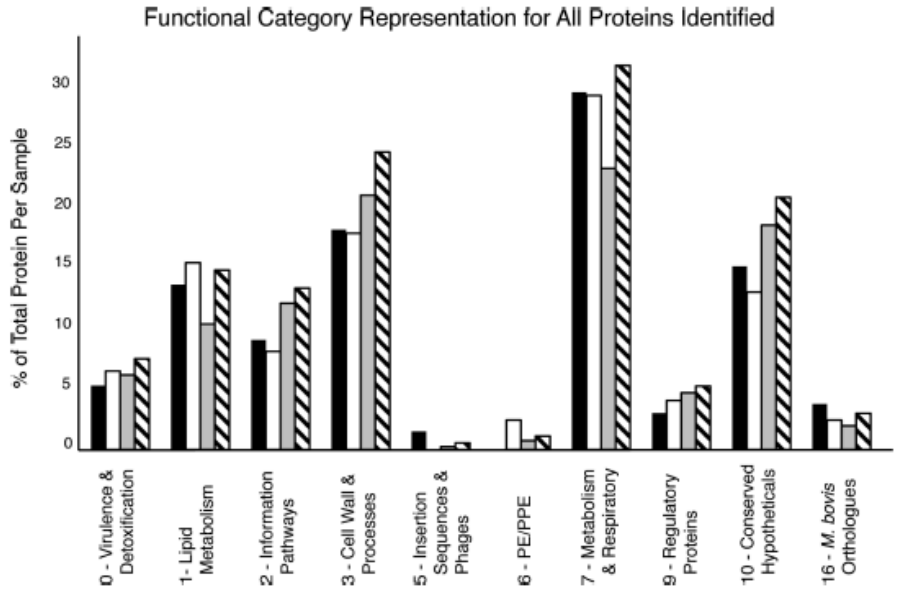

Figure 3.3 - The Functional Category Distribution by Biological Sample. Proteins identified for each extraction method (GuHCl, SDS, TX114) were sorted by functional category as a percentage of total proteins per sample. Representation of all protein groups were identified with the exception of groups 4 and 8 (data not shown). Seventy percent of the proteins identified in the cell wall are associated with lipid metabolism (group 1), cell wall processes (group 3) and intermediary metabolism (group 7). In addition, 21% are conserved hypotheticals. Solid Black, SDS; White, GuHCl; Gray, TX114; hash, total.

an even distribution of proteins with subclasses pertaining to the synthesis, modification, and/or degradation of macromolecules (II.A, II.B – 12.5%) and cell envelope proteins (II.C – 12.1%).

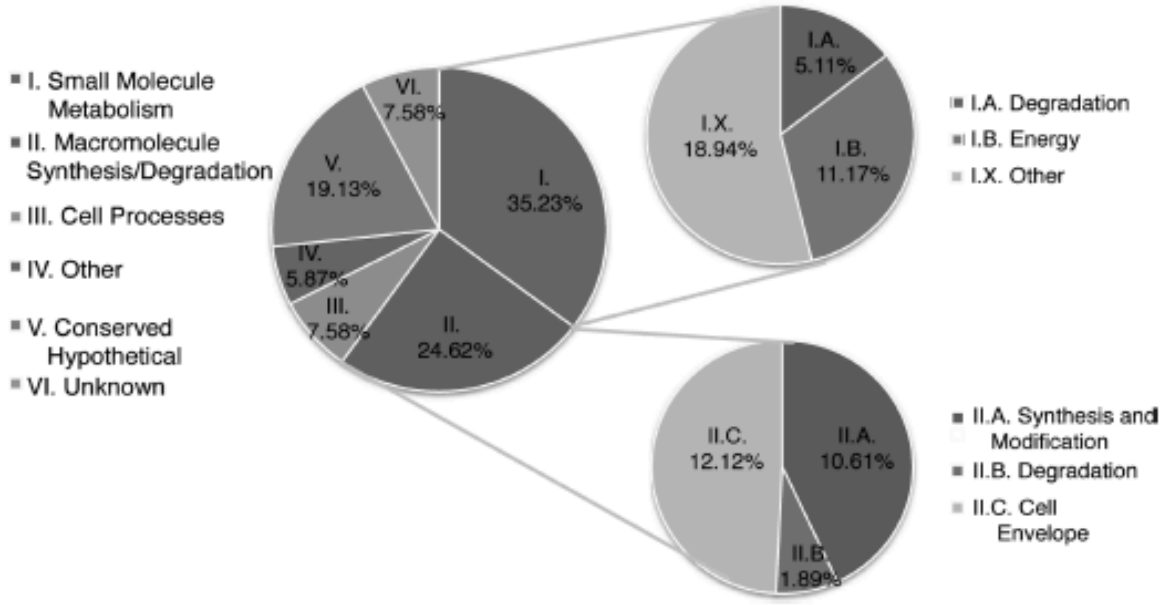

Figure 3.4 - The Functional Category Distribution of the Cell Wall Proteome. Assignments were made based on the Sanger Institute gene database. The distributions are among the major functional groups and the subgroups within functional groups I and II. The percentage for each subgroup indicates the percentage of the total number of identified proteins in its major functional group. I.X is defined by proteins involved in central intermediary metabolism, amino acid biosynthesis, polyamine synthesis, biosynthesis of cofactors, prosthetic groups, and carriers, lipid biosynthesis, polyketide and non-ribosomal peptide synthesis and broad regulatory functions.

3.2.3 Putative Secreted Proteins within the Cell Wall

The secreted proteins of *M. tuberculosis* have traditionally been characterized as important antigens and immune-modulators (4,11,12,46). Prior to being exported, many of these proteins are resident within the cell wall where their function remains largely unknown. To find the putative secreted proteins all identified proteins were subject to interrogation against Neural Network (NN) and Hidden Markov model (HMM) algorithms (SignalP - http://www.cbs.dtu.dk/services/SignalP/) (47). Which revealed 18%, 19%, 27% putative

secreted proteins in GuHCl, SDS, TX-114 CWP samples respectively, and 13% in the 2DLC resolved TX-114 CWP fraction. Cumulatively, of the 528 proteins identified in this study 87 proteins were predicted to contain secretion signals and included the identification of 23 proteins uniquely found in this study (Table 3.2). The majority (60%) of the CWP associated with secretion are within the category of the cell wall and its processes (Figure 3.5). Further, many of these CWP are also membrane associated – either by description (35,37,38) or functional annotation. This finding is inline with the inclusion criteria for category 3 as defined by Cole and Camus (48). Of the 23 secreted CWP unique to this study, eleven have been described to be involved with small molecule and peptide binding. Additionally, 23 of the 87 secreted proteins are classified as hypothetical or unknown (Categories V and VI of Sanger Institute) illustrating that to a large extent, the functions of these exported proteins are poorly understood.

Table 3.2 – Putative Secreted Proteins and Lipoproteins

		CWP fractions				secreted			
Rv no.	protein name	GuHCl	SDS	TX-114	2DLC	experimental[b]	predicted[c]	NetOglyc motif	reference
Rv0125	pepA	x				yes	yes		42, 16
Rv0129c	fbpC	x	x		x	yes	yes		42, 19, 16
Rv0169	mce1A				x				24
Rv0172	mce1D				x		yes		16
Rv0173	**lprK**				x		yes		16
Rv0174	mce1F				x				16, 19
Rv0194	Rv0194		x						24
Rv0202c	mmpL11	x	x	x					24
Rv0227c	Rv0227c				x				24, 46
Rv0237	**lpqI**				x		yes	yes	16, 46
Rv0291	mycP3				x		yes		16
Rv0292	Rv0292				x				46
Rv0309	Rv0309				x		yes		16
Rv0338c	Rv0338c	x							24, 46
Rv0402c	mmpL1	x					yes		16
Rv0431	Rv0431				x				
Rv0432	**sodC**			x	x			yes	45, 46, 19
Rv0480c	Rv0480c	x							24
Rv0507	mmpL2		x						24
Rv0512	hemB			x					
Rv0559c	Rv0559c				x		yes		42, 16
Rv0583c	**lpqN**				x		yes	yes	16, 19
Rv0622	Rv0622		x						
Rv0830	Rv0830	x							46
Rv0858c	Rv0858c	x							
Rv0892	Rv0892		x						
Rv0902c	prrB				x				
Rv0928	**pstS3**				x		yes	yes	16, 45
Rv0934	**pstS1**		x	x	x	yes	yes	yes	42, 19, 16, 46
Rv1006	**Rv1006**		x		x				24, 46
Rv1016c	**lpqT**				x			yes	46
Rv1029	kdpA	x							
Rv1075c	Rv1075c		x				yes		16, 19
Rv1096	Rv1096				x				
Rv1130	Rv1130	x							24
Rv1184c	Rv1184c				x				
Rv1209	Rv1209				x				
Rv1230c	Rv1230c		x						
Rv1244	**lpqZ**				x				46, 19
Rv1269c	Rv1269c				x		yes		42, 16
Rv1270c	**lprA**			x	x		yes	yes	16, 46, 19
Rv1273c	Rv1273c		x						
Rv1275	**lprC**				x				24, 46
Rv1307	atpH				x				24, 46
Rv1368	**lprF**			x	x			yes	46, 24, 22
Rv1411c	**lprG**		x		x			yes	46, 24, 22
Rv1450c	PE_PGRS27	x							
Rv1488	Rv1488				x		yes		16, 24, 22
Rv1522c	mmpL12		x						22
Rv1815	Rv1815		x				yes		16
Rv1886c	fbpB	x	x			yes	yes		42, 16, 22
Rv1911c	**lppC**	x					yes	yes	22, 46
Rv1926c	mpt63				x	yes	yes		42, 16
Rv1969	mce3D	x							
Rv1980c	mpt64				x	yes	yes		42, 16
Rv1984c	cfp21	x					yes		16
Rv2095c	Rv2095c				x				
Rv2171	**lppM**				x				
Rv2216	Rv2216				x				22
Rv2223c	Rv2223c		x						
Rv2307c	Rv2307c		x						
Rv2345	Rv2345				x				46, 22
Rv2376c	cfp2	x				yes	yes		19, 16
Rv2518c	**lppS**				x				
Rv2531c	Rv2531c		x						24
Rv2536	Rv2536				x				24
Rv2576c	Rv2576c	x							16
Rv2625c	Rv2625c		x						22
Rv2631	Rv2631		x						16

		CWP fractions				secreted			
Rv no.	protein name	GuHCl	SDS	TX-114	2DLC	experimental[b]	predicted[c]	NetOglyc motif	reference
Rv2712c	Rv2712c	x							
Rv2721c	Rv2721c	x			x				16
Rv2756c	hsdM		x						24
Rv2864c	**Rv2864c**		x						
Rv2945c	**lppX**			x	x			yes	16, 46, 24, 22
Rv3006	**lppZ**				x			yes	16, 46, 22
Rv3106	fprA				x				22
Rv3193c	Rv3193c				x				16, 22
Rv3224	Rv3224	x			x	yes			42, 19, 24, 22
Rv3243c	Rv3243c	x							
Rv3244c	**lpqB**		x					yes	16, 46
Rv3584	**lpqE**				x				16, 22
Rv3623	**lpqG**			x				yes	24, 22
Rv3666c	**dppA**			x					
Rv3682	ponA2	x							16, 24, 22
Rv3725	Rv3725		x						16
Rv3763	**lpqH**			x	x	yes		yes	19, 46, 24, 22
Rv3804c	fbpA		x		x	yes	yes		42, 19, 16, 24, 22

[a] Proteins in bold indicate putative lipoprotein. [b] Experimental evidence reported for secretion. See reference. [c] Predicted Signal Peptidase I cleavage (SignalP, http://www.cbs.dtu.dk/services/SignalP/).

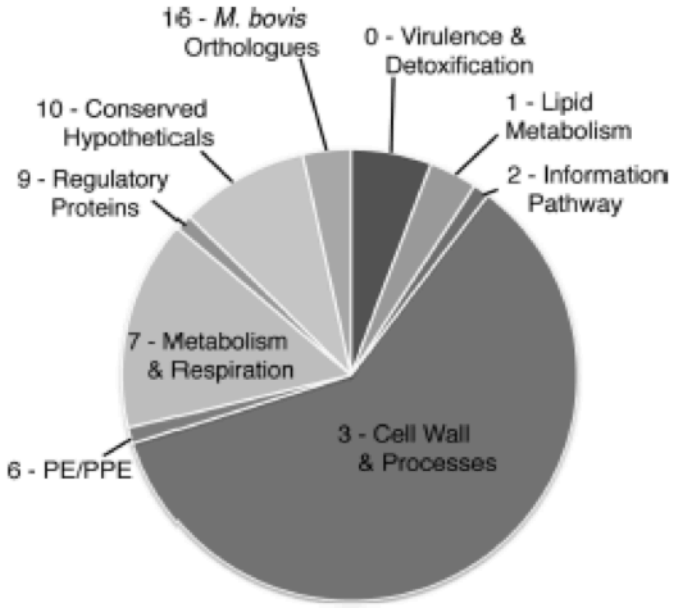

Figure 3.5 - Functional Classification for Putative Secreted Proteins and Lipoproteins.
Eighty-seven cell wall proteins were predicted to contain secretion signals and putative lipoprotein motifs. Sixty percent of these proteins are classified within the cell wall and cell wall processes functional category.

3.2.4 Putative Lipoproteins

Next, we interrogated the CWP to identify those with putative lipoprotein motifs. Here, a total of 16 proteins were predicted by LipoP (LipoP - http://www.cbs.dtu.dk/services/LipoP/) (49) to contain a signal peptidase II cleavage motif, and an additional 8 proteins were identified when compared to the 99 putative lipoproteins (2.5% of the genome) that reportedly exist within the *M. tuberculosis* proteome (20). The majority of the lipoproteins contain no additional functional classification, however 4 proteins (Rv0928, Rv0934, Rv2864c, Rv3666c) are involved with substrate binding and transport within the periplasm. One of two superoxide dismutases in *M. tuberculosis*, SodC (Rv0432), was also identified within the lipoprotein-enriched TX-114 CWP fraction. This protein has recently been defined as a highly glycosylated protein and is a defined B-cell antigen (50,51). Notably, no evidence of lipidation of this protein was characterized in this study suggesting that annotation of mycobacterial lipoproteins remains incomplete. Other putative lipoproteins identified in this study contain O-glycosylation motifs as predicted by NetOglyc and were found within the *M. tuberculosis* glycoproteome (Table 3.2) (52). The function of these dual-modified lipo-glycoproteins, as well as the nature of their modification remain largely undefined – however recent characterization of lprG (Rv1411c) has revealed its function as a lipoglycan carrier of triacylated glycolipids (53).

3.3 Discussion

In this study, we focused on the gel-based two-dimensional separation of the cell wall and complemented that separation by liquid chromatography of digested TX-114 CWP. Identified protein families included known antigens – Fbp A, B, and C (Rv 3804c, Rv1886c, and Rv0129c), CFP10 (Rv3874) and Esat-6 like proteins EsxJ and EsxL (Rv1038c, Rv1198) along with antigens lpqH (Rv3763), lprG (Rv1411c) and lprA (Rv1270c). Others included members of the MmpL and Mce protein family of lipid export machinery. In concordance with previous

work (39,54), a number of ribosomal proteins are reported here, which may indicate a high level of protein synthesis occurring at the cytosol – cell wall interface, most likely to facilitate the entry of proteins into the cellular envelope. For over a decade, the advancement of techniques in 2DGE, mass spectrometry and increased access to bioinformatic tools greatly enhanced proteomic studies of *Mycobacterium tuberculosis*. Largely descriptive, these studies were undertaken to identify novel virulence factors and drug targets (31,32,41). Subsequent studies set out to understand functional relationships between proteins and discover antigens responsible for adaptive T-cell responses (36,37,39,46,55).

The quantitative techniques of mass spectrometric profiling - isotope coded affinity tag (iCAT) and isobaric tagging for relative and absolute quantification (iTRAQ) - have afforded the accurate monitoring of up and down regulation of proteins on a global scale (36,56). Previous work mining subcellular fractions of *M. tuberculosis* focused on the cytosol and culture filtrate fractions, resulted in thorough characterization and the creation of 2D gel databases (4,32,44,57,58). Efforts have been made more recently to resolve the insoluble cell wall fraction. As mentioned previously Mawuenyega et al was able to identify 300 proteins within the cell wall using 2DLC as a separation method. Their study demonstrated functional relationships among key protein families within the fatty-acid synthesis pathway.

Using these studies as experimental platforms, we were able to focus on the *M. tuberculosis* cell wall and were successful in reporting over 100 additional proteins that had not been identified previously. The elucidation of the cell wall proteome can facilitate subsequent studies of a more targeted nature where specific proteins can be monitored in response to various environmental, nutritional or drug pressures. One of the most interesting yet poorly understood protein families residing within the cell wall of *M. tuberculosis* are the triacylated lipoproteins.

As stated above, Sutcliffe et al predicted 2.5% of the genome to encode for lipoproteins. Defining these proteins beyond their antigenic and immune modulatory potential remains quite elusive with only a few studies predicting functional roles for *M. tuberculosis* lipoproteins (15,26,59,60). In this study, the vast majority of proteins found were involved in small molecule and macromolecule metabolism. Specifically, many of the CWP were annotated as small molecule and nutrient binding. This builds upon the evidence that clearly defines two *M. tuberculosis* CW lipoproteins, LppX and LprG, as aiding and/or facilitating the transport of lipids and small molecules (15,61). The presence of these lipoproteins and perhaps other secreted, CW resident proteins provides a conduit between various nutrient transportation pathways that are required for cellular survival and growth while allowing the mycobacterial cell wall to maintain its rigid hydrophobic integrity.

3.4 Materials and methods

3.4.1 Preparation of the *M. tuberculosis* Cell Wall

M. tuberculosis strain $H_{37}Rv$ was cultured in 2 L of glycerol alanine salts (GAS) medium (14) in roller bottles for 14 d at 37 °C with gentle agitation. Cells were harvested (37), washed with phosphate-buffered saline (PBS) pH 7.4, and inactivated by gamma-irradiation. Cells were disrupted in a French press, free lipids removed and the cell wall was obtained as previously described (3). Briefly, 1 gram of lyophilized cell wall was subjected to two extractions of 2 h each followed by one 18 h extraction with chloroform:methanol (2:1 v/v) at a ratio of 30 mL per gram of cell wall. Extractions were performed at 22^{o}C with agitation. Centrifugation at 27,000-x g for 30 min was performed to collect cell wall material. 2:1 extracted cell wall was dried under N_2 and further extracted twice for 2 h and one 18 h extraction with chloroform:methanol:water (10:10:3 v/v/v) each at 22^{o}C. The fully delipidated cell wall was dried under N_2 and resuspended

in PBS pH 7.4. Cell wall protein was quantified by bicinchoninic acid (BCA) assay (Thermo Pierce).

3.4.2 Detergent Extractions of Cell Wall Proteins.

Cell wall protein (CWP) was solubilized by one of three methods using (i) Guanidine HCl (GuHCl), (ii) Sodium Dodecyl Sulfate (SDS) or (iii) TritionX-114 (TX-114). (i) 75mg of CWP was incubated at 22^{o}C for 4 h with agitation. The sample was centrifuged at 27,000 x g for 30 min. Solubilized proteins were exchanged into 0.01 M NH_4HCO_3, and the protein amount determined by BCA assay. (ii) SDS soluble proteins were generated by extraction of 100 mg of CWP with 2% SDS in PBS (w/v) as described previously (3). Briefly, the sample was bath sonicated for 3 h at 90^{o}C and centrifuged at 27,000-x g at 22^{o}C for 30 min, collected and centrifuged again. The fully cleared supernatant was then subjected to paired-ion extraction for removal of SDS (62). Proteins were concentrated by centrifugation as above and pellets washed with acetone at -20^{o}C for 4 h. The final pellet was resuspended in 0.01 M NH_4HCO_3 and protein amount determined by BCA. (iii) A stock solution of 32% TX-114 in PBS was added to 300 mg of CWP to a final concentration of 4% detergent. Primary extraction occurred at 4^{o}C for 16 h. The extract was allowed to bi-phase at 37^{o}C for 30 min and was fully separated by centrifugation at 27,000-x g for 30 min. Each phase was back extracted two additional times for 2 h each. The TX-114 detergent phases were pooled, and proteins collected by cold acetone precipitation (63). Final TX-114 proteins were resuspended as above and protein amount determined by BCA assay.

3.4.3 Two Dimensional Gel Electrophoresis (2DGE).

Samples of 200 μg of each CWP preparation were solubilized for 9 to 14 h in rehydration buffer, 7M urea/2M thiourea, 1% Amidosulfobetaine-14 (ASB-14), 1% 3-[(3-Cholamidopropyl)

dimethylammonio]-1propanesulfonate (CHAPS), 1 mM Dithiothreitol (DTT), 0.25% NP-40, 0.625% ZOOM™ carrier ampholytes 3-10, 1.9% ZOOM carrier ampholytes pH4-7 (Invitrogen). Solubilized CWP was applied on ZOOM pre-cast immobilized pH 4-7 linear gradient strips (7.0 cm; Invitrogen) according to manufacturer instructions. Focusing was achieved using a stepwise voltage gradient of 200, 450, 700 and 1000 V, for 10 min each followed by focusing at 2000 V for 2 h. SDSPAGE of the iso-electric focusing (IEF) strips was performed using ZOOM 4-12% Bis-Tris SDS-PAGE gradient gels (Invitrogen). The gels were stained by Coomassie blue R-250 (Bio-Rad).

3.4.4 Two-Dimensional Liquid Chromatography.

5.0 mg of TX-114 CWP was digested with modified trypsin (Roche Diagnostics) at a ratio of 1:20 (E:S), in 0.1 M Urea, 20 mM methylamine. The digest was desalted using Sep-Pak Light C18 cartridge (Waters Inc.) and concentrated under vacuum. The resultant peptides were separated by strong cation-exchange (SCX) chromatography using a polysulfylethyl A column (460 um X 200 mm, 300 A; Poly LC Inc.) connected to a Waters Alliance analytical HPLC with a 2487 UV detector. The digest was applied in buffer A (5 mM K2PO4, 20% acetonitrile (ACN) pH 3.0) using a gradient of 0-80% Buffer B (A with 0.5 M KCl) over 75 min with a flow rate of 1 mL/min. The elution of peptides was monitored at 214 nm, and fractions (3 mL each) were pooled based on UV absorbance. Each fraction (6 total) was concentrated under vacuum and resuspended into reverse phase Buffer A (0.1% TFA in H2O). Individual SCX pools were subjected to further separation by reversed-phase HPLC (RP-HPLC) using a monomeric C18 column (4.6 mm x 150 mm, Vydac). Peptides were eluted using a gradient of 0-50% reverse-phase buffer B (90% ACN in A) over 40 min at a flow rate of 1 mL/min. Peptides was manually

collected based on UV absorbance at 214 nm. A total of 139 RP-HPLC peptide fractions were collected and concentrated under vacuum.

3.4.5 Mass Spectrometry

Coomassie blue stained spots were excised from 2DGE gels and subjected to in-gel digestion (64,65) with modified trypsin (Roche Diagnostics). Digests and 2DLC peptide fractions were resolved by liquid chromatography-mass spectroscopy (LC-MS) using either an LCQ (2DGE) or LTQ (2DLC) as described previously (66). Tandem mass spectra were extracted, charge state deconvoluted and deisotoped by BioWorks version 3.2 (Thermo Finnigan, San Jose, CA). All MS/MS samples were analyzed using Sequest (Thermo Finnigan; version 27, rev. 12) Sequest was set up to search the TB genome database (version 2.0, Genbank accession #AL123456, 3912 entries) assuming the digestion enzyme trypsin, a fragment ion mass tolerance of 1.0 Da, a parent ion tolerance of 2.5 Da, and 3 allowable missed cleavages. Oxidation of methionine and acrylamide adduct of cysteine (2DGE spots) were specified in Sequest as variable modifications.

3.4.6 Criteria for Peptide and Protein Identification

Peptide identifications were accepted if they exceeded specific database search engine thresholds. Sequest identifications required at least deltaCn scores of greater than 0.3 and XCorr scores of greater than 1.5, 2.2, 2.5 and 2.5 for singly, doubly, triply and quadruply charged peptides (provided in Appendix II. Scaffold (version Scaffold-01_05_21, Proteome Software Inc., Portland, OR) was used to validate MS/MS based peptide and protein identifications. Peptide identifications were accepted if they could be established at greater than 90% probability as specified by the Peptide Prophet algorithm (67). Protein identifications were accepted if they could be established at greater than 90% probability as assigned by the Protein Prophet algorithm

(67) and contained at least two unique peptides. Protein identifications at the lower threshold of acceptance were manually inspected for spectra quality. The protein probability false discovery rate for the CWP identifications was 3.5% (68). Proteins that contained similar peptides and could not be differentiated based on MS/MS analysis alone were grouped to satisfy the principles of parsimony.

3.4.7 SignalP and LipoP Interrogation

FASTA sequence files for each annotated *M. tuberculosis* protein were collected and organized into small groups of 50. Each group was then submitted to the SignalP server (http://www.cbs.dtu.dk/services/SignalP/) (47,69,70) using the gram-positive option. The output was saved in HTML format and a Perl script written to combine the results from all HTML files into one file and then converted to tab delimited format. This generated a complete list of results that were sorted in a spreadsheet to create the subsets of Neural Network positive, Hidden Markov Model positive, and the union of the NN and HMM sets. These lists were input into a java application along with the list of proteins isolated from the cell wall fractions. This application simply compared the predicted list to those isolated and identified the proteins found in the cell wall that were also predicted in each subset as having a signal peptide. To predict lipoproteins, the *M. tuberculosis* peptide FASTA file from above was broken down into 1000 sequences per file. These were run through the downloadable version of LipoP (http://www.cbs.dtu.dk/services/LipoP/). The output files were combined in a spreadsheet, and from the sorted data, a list was extracted with every signal peptidase II cleavage site generating a predicted lipoprotein list. This list was input into a Java application that would compare the cell wall proteins, identifying the predicted lipoproteins from each cell wall fraction. The positive results verified that both Signal P models had predicted a signal peptide.

3.5 Conclusion & Current Perspectives

This work was undertaken in order to 1. Develop new techniques for solubilization and extraction of proteins from a highly insoluble matrix of cell wall lipids, carbohydrates and mycolic acids, 2. Optimize separation of these proteins via two-dimensional gel electrophoresis and multi-dimensional liquid chromatography and 3. Utilize techniques of shotgun proteomics and mass spectrometry to comprehensively describe a unique and highly sought after subcellular proteome of *M. tuberculosis*. From these efforts we characterized 528 proteins, 87 proteins with known or predicted secretion signal motifs and 23 lipoproteins. While bacterial cell walls have historically been thought of as structural scaffolding with little active interplay between the extracellular milieu and inner cytosolic environment, this perception of a static bacterial structure is fading and we are beginning to understand the true complexity and dynamic nature of the biological processes being facilitated within the cell wall. For instance, these proteins may contribute to bacterial cell wall remodeling events in response to environmental stressors such as nutrient depletion, antibiotic pressures, the host immune response and tissue remodeling events throughout the course of infection. Indeed several lipoproteins have been functionally associated with nutrient uptake and transport. For *M. tuberculosis*, the highly complex framework of the mAGP – outer lipid architecture can now be definitively complemented with a highly diverse protein population. This new perspective will provide further insight into cell wall remodeling processes during mycobacterial infection. In the past the provision of experimental data generated from descriptive studies have been purely data driven, but "information poor". One of the primary objectives of this study was borne of the observation that proteins within the cell wall were highly immunogenic for non-classically restricted CD8 T cells. The inclusion of this proteomic dataset in the design of a synthetic peptide library for the elucidation of novel antigens (discussed in the following chapter) was the first step in the shift to data rich, information rich

proteomic studies in which functional associations, interacting protein-protein networks and novel biomarkers of disease can be extrapolated and quantified among differing microenvironments and altered disease states.

3.6 LITERATURE CITED

1. Lederer, E., Adam, A., Ciorbaru, R., Petit, J. F., and Wietzerbin, J. (1975) *Mol Cell Biochem* 7, 87-104
2. Ribi, E., Granger, D. L., Milner, K. C., Yamamoto, K., Strain, S. M., Parker, R., Smith, R. W., Brehmer, W., and Azuma, I. (1982) *Immunology* 46, 297-305
3. Hirschfield, G. R., McNeil, M., and Brennan, P. J. (1990) *J Bacteriol* 172, 1005-1013
4. Young, D., Garbe, T., Lathigra, R., Abou-Zeid, C., and Zhang, Y. (1991) *Bull Int Union Tuberc Lung Dis* 66, 47-51
5. Briken, V., Porcelli, S. A., Besra, G. S., and Kremer, L. (2004) *Mol Microbiol* 53, 391-403
6. Rao, V., Gao, F., Chen, B., Jacobs, W. R., Jr., and Glickman, M. S. (2006) *J Clin Invest* 116, 1660-1667
7. Mahon, R. N., Rojas, R. E., Fulton, S. A., Franko, J., Harding, C. V., and Boom, W. H. (2009) *Infect. Immun.*, IAI.00222-00209
8. Kumar, D., and Rao, K. V. (2011) *Microbes Infect* 13, 121-133
9. Barnes, P. F., Mehra, V., Hirschfield, G. R., Fong, S. J., Abou-Zeid, C., Rook, G. A., Hunter, S. W., Brennan, P. J., and Modlin, R. L. (1989) *J Immunol* 143, 2656-2662
10. Dhiman, N., Verma, I., and Khuller, G. K. (1997) *Microbiol Immunol* 41, 495-502
11. Collins, F. M., Lamb, J. R., and Young, D. B. (1988) *Infect Immun* 56, 1260-1266
12. Rook, G. A., Steele, J., Barnass, S., Mace, J., and Stanford, J. L. (1986) *Clin Exp Immunol* 63, 105-110
13. da Silva, P. E., Von Groll, A., Martin, A., and Palomino, J. C. (2011) *FEMS Immunol Med Microbiol* 63, 1-9
14. Takayama, K., Schnoes, H. K., Armstrong, E. L., and Boyle, R. W. (1975) *J Lipid Res* 16, 308-317
15. Farrow, M. F., and Rubin, E. J. (2008) *J. Bacteriol.* 190, 1783-1791
16. Brennan, P. J. (2003) *Tuberculosis (Edinb)* 83, 91-97
17. Szczepan, J., Andrzej, S., and Katarzyna, K. (2008) *BioEssays* 30, 943-954
18. Jackson, M., Stadthagen, G., and Gicquel, B. (2007) *Tuberculosis (Edinb)* 87, 78-86
19. Neyrolles, O., Gould, K., Gares, M. P., Brett, S., Janssen, R., O'Gaora, P., Herrmann, J. L., Prevost, M. C., Perret, E., Thole, J. E., and Young, D. (2001) *J Immunol* 166, 447-457
20. Sutcliffe, I. C., and Harrington, D. J. (2004) *FEMS Microbiol Rev* 28, 645-659
21. Young, D. B., and Garbe, T. R. (1991) *Res Microbiol* 142, 55-65
22. Brightbill, H. D., Libraty, D. H., Krutzik, S. R., Yang, R. B., Belisle, J. T., Bleharski, J. R., Maitland, M., Norgard, M. V., Plevy, S. E., Smale, S. T., Brennan, P. J., Bloom, B. R., Godowski, P. J., and Modlin, R. L. (1999) *Science* 285, 732-736
23. Gehring, A. J., Dobos, K. M., Belisle, J. T., Harding, C. V., and Boom, W. H. (2004) *J Immunol* 173, 2660-2668
24. Pecora, N. D., Gehring, A. J., Canaday, D. H., Boom, W. H., and Harding, C. V. (2006) *J Immunol* 177, 422-429
25. Hertz, C. J., Kiertscher, S. M., Godowski, P. J., Bouis, D. A., Norgard, M. V., Roth, M. D., and Modlin, R. L. (2001) *J Immunol* 166, 2444-2450
26. Sanchez, A., Espinosa, P., Esparza, M. A., Colon, M., Bernal, G., and Mancilla, R. (2009) *Scand J Immunol* 69, 20-28
27. Malen, H., Softeland, T., and Wiker, H. G. (2008) *Scand J Immunol* 67, 245-252

28. Malen, H., Berven, F. S., Softeland, T., Arntzen, M. O., D'Santos, C. S., De Souza, G. A., and Wiker, H. G. (2008) *Proteomics* 8, 1859-1870
29. Malen, H., Berven, F. S., Fladmark, K. E., and Wiker, H. G. (2007) *Proteomics* 7, 1702-1718
30. Urquhart, B. L., Atsalos, T. E., Roach, D., Basseal, D. J., Bjellqvist, B., Britton, W. L., and Humphery-Smith, I. (1997) *Electrophoresis* 18, 1384-1392
31. Urquhart, B. L., Cordwell, S. J., and Humphery-Smith, I. (1998) *Biochem Biophys Res Commun* 253, 70-79
32. Rosenkrands, I., King, A., Weldingh, K., Moniatte, M., Moertz, E., and Andersen, P. (2000) *Electrophoresis* 21, 3740-3756
33. Jungblut, P. R., Muller, E. C., Mattow, J., and Kaufmann, S. H. (2001) *Infect Immun* 69, 5905-5907
34. Mattow, J., Jungblut, P. R., Schaible, U. E., Mollenkopf, H. J., Lamer, S., Zimny-Arndt, U., Hagens, K., Muller, E. C., and Kaufmann, S. H. (2001) *Electrophoresis* 22, 2936-2946
35. Gu, S., Chen, J., Dobos, K. M., Bradbury, E. M., Belisle, J. T., and Chen, X. (2003) *Mol Cell Proteomics* 2, 1284-1296
36. Schmidt, F., Donahoe, S., Hagens, K., Mattow, J., Schaible, U. E., Kaufmann, S. H., Aebersold, R., and Jungblut, P. R. (2004) *Mol Cell Proteomics* 3, 24-42
37. Mawuenyega, K. G., Forst, C. V., Dobos, K. M., Belisle, J. T., Chen, J., Bradbury, E. M., Bradbury, A. R., and Chen, X. (2005) *Mol Biol Cell* 16, 396-404
38. Xiong, Y., Chalmers, M. J., Gao, F. P., Cross, T. A., and Marshall, A. G. (2005) *J Proteome Res* 4, 855-861
39. Sinha, S., Kosalai, K., Arora, S., Namane, A., Sharma, P., Gaikwad, A. N., Brodin, P., and Cole, S. T. (2005) *Microbiology* 151, 2411-2419
40. Kruh, N. A., Troudt, J., Izzo, A., Prenni, J., and Dobos, K. M. (2010) *PLoS One* 5, e13938
41. Rosenkrands, I., Weldingh, K., Jacobsen, S., Hansen, C. V., Florio, W., Gianetri, I., and Andersen, P. (2000) *Electrophoresis* 21, 935-948
42. Isobe, T., Uchida, K., Taoka, M., Shinkai, F., Manabe, T., and Okuyama, T. (1991) *J Chromatogr* 588, 115-123
43. Mawuenyega, K. G., Kaji, H., Yamuchi, Y., Shinkawa, T., Saito, H., Taoka, M., Takahashi, N., and Isobe, T. (2003) *J Proteome Res* 2, 23-35
44. Mollenkopf, H. J., Jungblut, P. R., Raupach, B., Mattow, J., Lamer, S., Zimny-Arndt, U., Schaible, U. E., and Kaufmann, S. H. (1999) *Electrophoresis* 20, 2172-2180
45. Cole, S. T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S. V., Eiglmeier, K., Gas, S., Barry, C. E., 3rd, Tekaia, F., Badcock, K., Basham, D., Brown, D., Chillingworth, T., Connor, R., Davies, R., Devlin, K., Feltwell, T., Gentles, S., Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., Krogh, A., McLean, J., Moule, S., Murphy, L., Oliver, K., Osborne, J., Quail, M. A., Rajandream, M. A., Rogers, J., Rutter, S., Seeger, K., Skelton, J., Squares, R., Squares, S., Sulston, J. E., Taylor, K., Whitehead, S., and Barrell, B. G. (1998) *Nature* 393, 537-544
46. Mattow, J., Schaible, U. E., Schmidt, F., Hagens, K., Siejak, F., Brestrich, G., Haeselbarth, G., Muller, E. C., Jungblut, P. R., and Kaufmann, S. H. (2003) *Electrophoresis* 24, 3405-3420

47. Bendtsen, J. D., Nielsen, H., von Heijne, G., and Brunak, S. (2004) *J Mol Biol* 340, 783-795
48. Camus, J. C., Pryor, M. J., Medigue, C., and Cole, S. T. (2002) *Microbiology* 148, 2967-2973
49. Juncker, A. S., Willenbrock, H., Von Heijne, G., Brunak, S., Nielsen, H., and Krogh, A. (2003) *Protein Sci* 12, 1652-1662
50. Sartain, M. J., and Belisle, J. T. (2009) *Glycobiology* 19, 38-51
51. Sartain, M. J., Slayden, R. A., Singh, K. K., Laal, S., and Belisle, J. T. (2006) *Mol Cell Proteomics* 5, 2102-2113
52. Gonzalez-Zamorano, M., Mendoza-Hernandez, G., Xolalpa, W., Parada, C., Vallecillo, A. J., Bigi, F., and Espitia, C. (2009) *J Proteome Res* 8, 721-733
53. Drage, M. G., Tsai, H. C., Pecora, N. D., Cheng, T. Y., Arida, A. R., Shukla, S., Rojas, R. E., Seshadri, C., Moody, D. B., Boom, W. H., Sacchettini, J. C., and Harding, C. V. (2010) *Nat Struct Mol Biol* 17, 1088-1095
54. Driessen, A. J., Manting, E. H., and van der Does, C. (2001) *Nat Struct Biol* 8, 492-498
55. Dobos, K. M., Spencer, J. S., Orme, I. M., and Belisle, J. T. (2004) *Methods Mol Med* 94, 3-17
56. Mehaffy, C., Hess, A., Prenni, J. E., Mathema, B., Kreiswirth, B., and Dobos, K. M. (2010) *Proteomics*
57. Mattow, J., Jungblut, P. R., Muller, E. C., and Kaufmann, S. H. (2001) *Proteomics* 1, 494-507
58. Sonnenberg, M. G., and Belisle, J. T. (1997) *Infect Immun* 65, 4515-4524
59. Ciaramella, A., Cavone, A., Santucci, M. B., Garg, S. K., Sanarico, N., Bocchino, M., Galati, D., Martino, A., Auricchio, G., D'Orazio, M., Stewart, G. R., Neyrolles, O., Young, D. B., Colizzi, V., and Fraziano, M. (2004) *J Infect Dis* 190, 1167-1176
60. Drage, e. a. (2009) *Cellular Immunology* 258, 29-37
61. Sulzenbacher, G., Canaan, S., Bordat, Y., Neyrolles, O., Stadthagen, G., Roig-Zamboni, V., Rauzier, J., Maurin, D., Laval, F., Daffe, M., Cambillau, C., Gicquel, B., Bourne, Y., and Jackson, M. (2006) *EMBO J* 25, 1436-1444
62. Konigsberg, W. H., and Henderson, L. (1983) *Methods Enzymol* 91, 254-259
63. Radolf, J. D., Chamberlain, N. R., Clausell, A., and Norgard, M. V. (1988) *Infect Immun* 56, 490-498
64. Hellman, U., Wernstedt, C., Gonez, J., and Heldin, C. H. (1995) *Anal Biochem* 224, 451-455
65. Rosenfeld, J., Capdevielle, J., Guillemot, J. C., and Ferrara, P. (1992) *Anal Biochem* 203, 173-179
66. Turner, J., Dobos, K. M., Keen, M. A., Frank, A. A., Ehlers, S., Orme, I. M., Belisle, J. T., and Cooper, A. M. (2004) *Infect Immun* 72, 3759-3768
67. Keller, A., Nesvizhskii, A. I., Kolker, E., and Aebersold, R. (2002) *Anal Chem* 74, 5383-5392
68. Searle, B. C. *Proteomics* 10, 1265-1269
69. Nielsen, H., Engelbrecht, J., Brunak, S., and von Heijne, G. (1997) *Protein Eng* 10, 1-6
70. Nielsen, H., and Krogh, A. (1998) *Proc Int Conf Intell Syst Mol Biol* 6, 122-130

IV. CHAPTER IV: Mining the *M. tuberculosis* Proteome for CD8 T cell Antigens & Epitopes

In the sections below, we describe a large-scale effort to identify both classically and non-classically restricted CD8 T cell antigens. The identification of immunodominant epitopes is achieved through a T cell based approach, utilizing cells derived from *M. tuberculosis*-infected patients. Our primary role in this contract was to develop methods for the identification of CD8 specific antigens derived from the cell wall of *M. tuberculosis*. Specifically, our task was to identify the antigen presented in the context of the non-classically restricted molecule HLA-E.

4.1 Introduction

The cellular immune response to mycobacterial infection is essential to minimize dissemination and control the severity of disease. In the majority of immunocompetent patients lymphocytes and *M. tuberculosis* infected dendritic cells (DCs) and macrophages (MΦ) act to control the invading bacilli via a complex interplay of immune effectors (1). As summarized in Chapter 1, CD8 T cells are associated with a strong CD4 T cell response and influence the development of a Th1 phenotype. These events act to control primary infection and influence the development of latent disease when the bacteria take on a state of non-replicating persistence (2) (3). Unfortunately, the immune response alone is not able to completely eliminate the bacterial burden and reactivation of active disease can occur when the homeostatic balance between pathogen and host is perturbed. To better understand the dynamic lifecycle and physiological state of the bacilli during different disease states, we must work toward a comprehensive biochemical, molecular and immunological understanding of the antigens present during the cycles of *M. tuberculosis infection* (2,4). The immunodiagnostic tools (i.e. interferon-gamma release assays, IGRAs) that are currently available for tuberculosis are based only on the broadly defined repertoire of a set of known CD4 T cell antigens. While their accuracy in diagnosing *M.*

tuberculosis infection is much improved, the assays are not designed to distinguish between LTBI or active disease and their accuracy in diagnosing tuberculosis within highly endemic regions is still under evaluation (5,6). To circumvent the shortcomings of current diagnostic methods, researchers must gain further insight into the antigenic repertoire of the *M. tuberculosis* proteome during multiple stages of infection. The cellular envelope is highly immunogenic and we hypothesize that the large reservoir of protein antigens that reside within or that are secreted from the bacillus can be exploited in the pursuit of novel T cell diagnostics. Because the cell envelope is key to maintaining homeostasis during multiple infection and disease states, antigens deriving from this structure during different stages of infection could be utilized in the design of improved immunodiagnostics that are capable of identifying persons at risk for disease progression or that reflect the bacterial burden within individual patients. Indeed the study of the CD8 T cell immune response has shown that CD8 T cells are uniquely capable of recognizing highly infected host cells (7) and that CD8 T cells are required to prevent reactivation of disease in a mouse model of latency (8). Continued efforts in the study of the antigenic repertoire restricted to CD8 T cells in tuberculosis patients have successfully characterized several novel antigens and epitopes relating to the cell envelope of the bacillus (4).

In humans it is known that antigens presented at the host cell surface to CD8 T cells are presented in the context of either a classically MHC1a-restricted molecule (9-14) or non-classically restricted molecule such as CD1 (15-17), HLA-E (10,18) and MR1 (19,20). Several studies have identified classically restricted HLA-A2 presented peptides, however the frequency of the HLA-A haplotypes among *M. tuberculosis*-infected individuals was found to be quite low (21), suggesting that subsequent large-scale antigen discovery efforts needed to focus on the processing and presentation of antigens associated with the HLA haplotypes observed most

frequently during *M. tuberculosis* infection. Towards the identification of the new CD8 T cell antigens, two approaches were undertaken in a highly collaborative effort between the OHSU and CSU. Lewinsohn et al. utilized an unbiased alternative epitope discovery platform to identify disease-relevant immunodominant CD8 T cell epitopes. This strategy implemented the generation of a synthetic peptide library, in which CD8 T cells could be tested directly *ex vivo* from tuberculosis patients presenting with different stages of clinical disease. Antigen discovery efforts also continued for the identification of the *M. tuberculosis* antigen presented in the context of HLA-E. Initial characterization schemes for HLA-E, based on high-throughput methods such as a genomic peptide library or through the screening of synthetic peptides, were not successful at identifying the cognate antigen. Therefore we utilized an alternative, biochemical approach based on the purification of native proteins and peptide fragments to isolate biological activity to a particular subcellular fraction. Mass spectrometry was then utilized to identify protein species resident within such a fraction, facilitating the identification of the cognate epitope.

4.2 The High-Throughput CD8 T cell Antigen Discovery Project

4.2.1 Synthetic Peptide Libraries for Epitope Screening

Identification of the antigenic repertoire of human CD8 T cells in response to infection with *M. tuberculosis* can be accomplished in several complementary ways. As a high throughput strategy a synthetic peptide library was designed and developed by Dr. David Lewinsohn and his team at OHSU. Analysis of pathogen-derived T cell antigens were first described for the influenza A nucleoprotein. In these studies, it was observed that cytotoxic T-lymphocytes could not only respond to short peptide fragments within transfected cells, but that these CTLs were capable of recognizing short synthetic peptides on the cell surface of the antigen presenting cell

(22-24).. This work demonstrated that antigenic targets could be identified on the whole protein level and appropriate epitopes could be elucidated via synthetic means. Advancing this concept into a high throughput, microtiter based screening platform would require the integration of new technologies for the high throughput synthesis of peptides as well as computational interrogation of genomic and proteomic information to develop robust peptide libraries for ambitious screening efforts (25,26).

4.2.2 Integration of Descriptive Analyses into the Design of an *M. tuberculosis* Genomic Peptide Library

The development of the *M. tuberculosis* synthetic peptide library began with a comprehensive evaluation of available genomic and proteomic data that was both publically available (http://genolist.pasteur.fr/TubercuList/ and http://web.mpiib-berlin.mpg.de/cgi-bin/pdbs/2d-page/extern/index.cgi) and provided through our description of the cell wall proteome (see Chapter III for details). Also, lists of secreted and cell associated proteins identified through proteomic analyses done at CSU (unpublished data), was provided via personal communication with Dr. Dobos. From this information roughly 10% of the *M. tuberculosis* genome (389 genes) were represented in the final repertoire of ~39,000 peptides (Figure 4.1).

After a comprehensive analysis and strategic weighting scheme were applied (see 4.2.3 Discussion), the functional distribution of included genes (27) is visualized in Figure 4.2 (28).

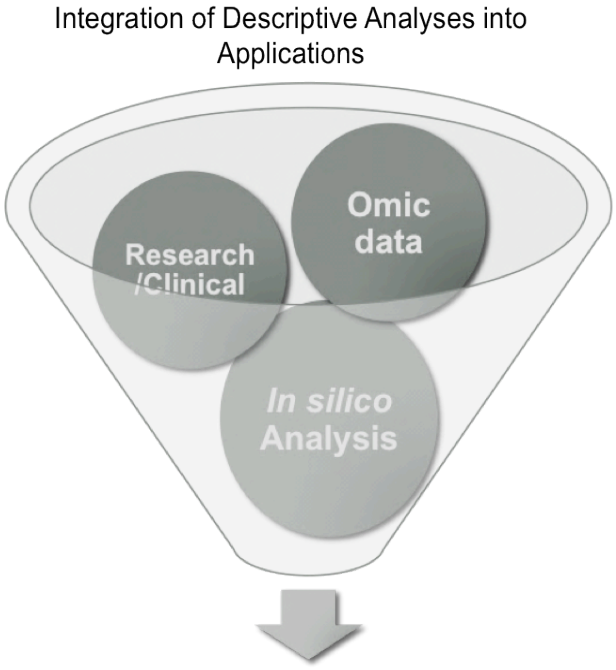

Figure 4.1 – Integration of Omics Data: Schematic diagram of the integration of genomic and proteomic data in the overall design of the synthetic peptide library, representing roughly 10% of the *M. tuberculosis* genome.

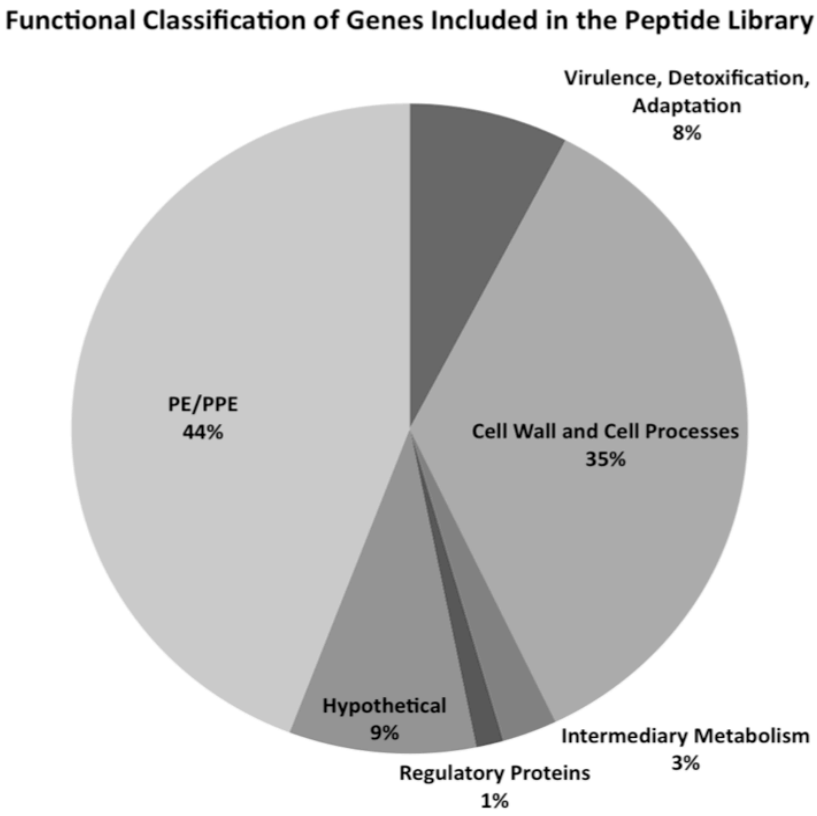

Figure 4.2 – Functional Distribution of the Genomic Peptide Library. Genomic and proteomic information was compiled and analyzed based on a number of factors which indicated the likelihood of a particular gene/protein of interest could be a CD8 T cell antigen. Genes meeting certain criteria – such as secretion status, known CD4 T cell antigens, functional category association, were assigned weights. Genes meeting a particular score chosen and a final list of 389 genes were selected to be included in the genomic peptide library. The 389 genes were represented in a library of 39,000 overlapping synthetic peptides. Their functional distribution (as defined in Tuberculist**)** is illustrated.

Numerically, the cell wall proteome contributed 56 total protein sequences to the final peptide library (Figure 4.3) and 8 of those proteins were selected for clinical validation as part of the OHSU-led T cell antigen discovery effort. These included members of the proline-glutamate rich proteins (PPE51, PPE18), possible exported proteins with unknown function (Rv1184c),

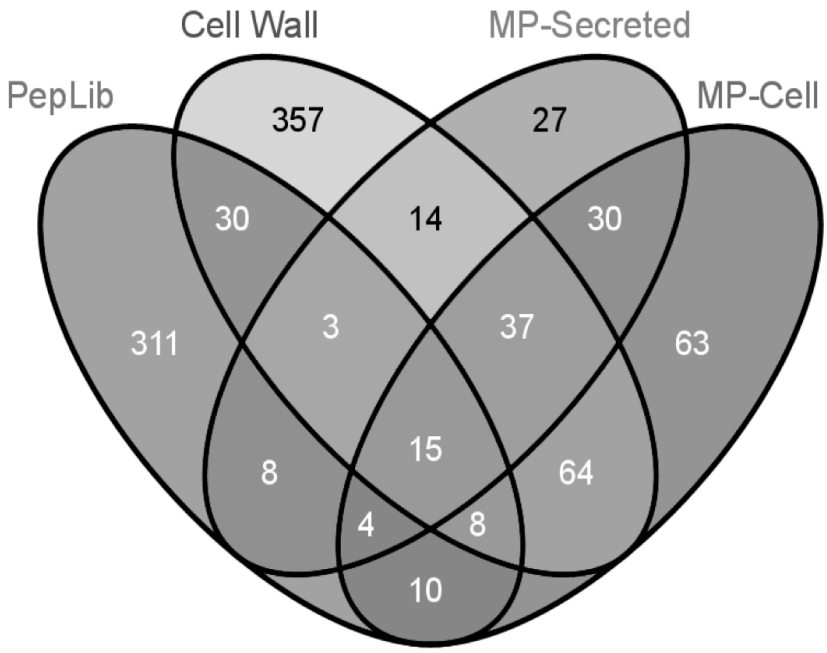

Figure 4.3 – Venn Diagram Comparison. Proteome datasets from the Max Plank Bioinformatics Center (MP) and from the Cell Wall proteome work were integrated into the design of the genomic peptide library (PepLib). After rigorous scoring schemes were applied, a total of 56 Cell Wall proteins were included in the final list of 389 representative genes/proteins of the PepLib.

ESAT-6 family proteins (EsxJ, EsxB). Others were the secreted antigens Mpt64 and Cfp21, as well as the iron-binding protein IdeR.

4.3 Discussion of the Genomic Peptide Library Screen

The proteomic scale prediction of immunological characteristics is rapidly developing with the exponential increase in the availability of proteomic data (29). A challenge for any large-scale descriptive study – like the cell wall proteome, is the integration and interpretation of the data into productive experimental and clinical utility. However, this proteomic information has the potential to provide bioinformaticists (such as those designing peptide libraries) and

clinicians, relevant patient and population-scale information. This holds especially true for host and/or pathogen proteomes derived from true states of infection. Therefore the integration of proteomic data can contribute useful information in the immunomic characterization of a disease state. This is especially true when considering that binding to MHC I/II molecules is a competitive process in which selected peptides bind to their MHC in the context of very complex peptide mixtures derived from processing events within the cell. Because of these dynamic interactions, Bremel and Homan present the idea that, "predictive determination of preferential epitope binding can only be made when considered in the context of the whole proteome" (29). In support of this idea, integration of *M. tuberculosis* proteomic data in the design of the first iteration of the OHSU peptide library significantly contributed to the identification of potent CD8 T cell antigens. Design analysis of the library was performed in order to understand which component of the library was more likely to predict CD8 T cell antigens. The parameters included the testing of proteomic data, genomic data and whether or not the proteins were expressed in *M. tuberculosis*, BCG. Together these parameters made up the "composite-based evidence" score, which was one of the major factors in determining the final protein representation of the library. This analysis, performed by Drs. Byung Park and Tomi Mori at OHSU determined that it was the association with the proteomic data that increased the likelihood of predicting a "good" CD8 T cell antigen (defined by an antigen producing IFN-γ responses among the top 5% of donor screens) and that this increased likelihood had the most influence on the value of the overall composite based score (30). Several predictive algorithms are now being developed in which proteomic data sets can be screened and analyzed for the presence of T cell epitopes (31). This further development will expand the utility of such large-scale descriptive proteomic platforms especially when looking at specific disease states and may

shift the success rate of identifying disease-relevant epitopes away from classical, biochemical based approaches in which antigens are isolated from whole cells, digested into peptide fragments and tested for T cell reactivity.

4.4 Biochemical Approaches to Define Non-Classically Restricted Epitopes

As a second, complementary approach to antigen discovery we utilized a biochemical proteomics-based screen to identify CD8 T cell antigens for a subset of non-classically restricted T cell clones. These clones had failed to recognize antigens represented in the genomic peptide library. The rationale behind this alternative approach included the recognized limitations of the synthetic peptide approach. For example, the scope of the peptide library was limited to only 10% of the annotated *M. tuberculosis*, H_{37}Rv genome and the initial characterization of HLA-E restricted T cells demonstrated activity with hydrophobic fractions of *M. tuberculosis* cell wall (18). This lead to the hypothesis that the non-classical peptide antigens could be post-translationally modified with motifs similar to other non-classically presented glycolipids (16,32). Proteomics and mass spectrometry studies provide robust and reliable methods in the discovery of disease state biomarkers and identification of novel drug targets. A comprehensive, full proteome screen for non-classically restricted CD8 T cell antigens provides a means to qualify the antigenic potential of naturally occurring proteins in their native state. In the field of mass spectrometry, the "bottom up" approach to protein identification involves the initial separation of complex protein mixtures and their subsequent proteolytic digestion into peptides. Peptides are then introduced into a mass spectrometer and their fragment spectra are identified and mapped to their corresponding protein based on sequence matches of a protein database (33). When coupled to immunological assays (i.e. cell proliferation (34,35), serum reactivity (36,37)

and cytokine secretion (38)) large-scale antigen discovery platforms and proteomics can be applied to the study of biologically active native protein fractions (39).

4.5 Rationale for Identification of HLA-E Antigens of *Mycobacterium tuberculosis*

The human CD8+ T cell response to *M. tuberculosis* infection has been shown to involve both the classical HLA-Ia restricted CD8+ T cells and those that are non-classically (neither HLA-Ia nor CD1) restricted (21). The HLA-Ib molecule HLA-E is the least polymorphic of the restricting antigen alleles, having only two human haplotypes – HLA-E0101 and HLA-E0103 (40). Further, the HLA-E restricted phenotype appears to be relatively common in those infected with *M. tuberculosis* (21) and recent work has identified HLA-E to be enriched in the MHC I processing phagosome (41). Therefore the identification of pathogen-specific HLA-E antigens could be novel vaccine candidates or diagnostic targets. Two CD8 T cell clones generated from a latently infected individual (D160 1-23 and 1-29) are known to be HLA-E restricted (18). We set out to identify the cognate antigen associated with this restricting allele using a biochemical proteomics based approach. Non-classically restricted antigens could have broad utility in the design of peptide vaccines and could lead to the further characterization of other HLA-E binding peptides. In the following sections we describe our efforts to identify the antigen associated with the non-classical HLA-E restricted T cell clone 1-23 and present substantial evidence that the post-translationally modified glycoprotein Mpt32 (Rv1860, Apa) is the cognate antigen. Limitations on the "bottom up" mass spectrometry approach to elucidate the presented peptide sequence are discussed and alternative approaches to epitope discovery are proposed for future work.

4.6 Results for HLA-E Antigen Discovery

4.6.1 Delipidation and Pronase Digestion of the *M. tuberculosis* Cell Wall

The HLA-E restricted T cell clone D160 1-23 was used to characterize the biological activity of subcellular fractions generated from mid-log phase grown *M. tuberculosis*, H_{37}Rv by measuring antigen specific IFN-γ release by ELISPOT assay. Original work done in collaboration between OHSU and CSU had isolated activity to the cell wall of *Mtb* (18). To facilitate the identification of cell wall proteins by mass spectrometry the fraction was subjected to delipidation by treating with organic solvents (42). The delipidated cell wall (dCW) remained active in the ELISPOT assay and was used as a positive control in subsequent assays. Pronase digestion of the dCW (dCW_pro) enhanced secretion of IFN-γ from the HLA-E restricted T cell clone and was also included as a point of reference for potential antigens to be screened. Trypsin digestion of dCW was not biologically active.

4.6.2 Characterization of X4-19 (HspX Knockout) Whole Cell Lysate

M. tuberculosis X4-19 (ΔHspX), was obtained as a generous gift from Dr. Fred Quinn at the University of Georgia. This strain was used primarily to avoid the high abundance of HspX in active fractions, as the presence of this protein confounded our proteomic analysis (discussed further in the Discussion). The HspX knockout strain was grown to mid-log phase to collect culture filtrate. Whole cells were harvested to generate whole cell lysate (WCL). An ELISPOT assay was performed on WCL with or without Pronase. From this analysis, it was determined that T cell reactivity was retained and enhanced in the absence of HspX, thus confirming that HspX was a confounding protein and not the cognate antigen. Pronase digests of ΔHspX_WCL were performed in triplicate and separated by RP-HPLC producing 30 peptide fractions per digest to produce a total of 90 fractions that could be assayed for T cell activity. Fractions I.9,

II.9, III.9 as well as fractions I/II/III.10 were reproducibly biologically active in these assays (Figure 4.4).

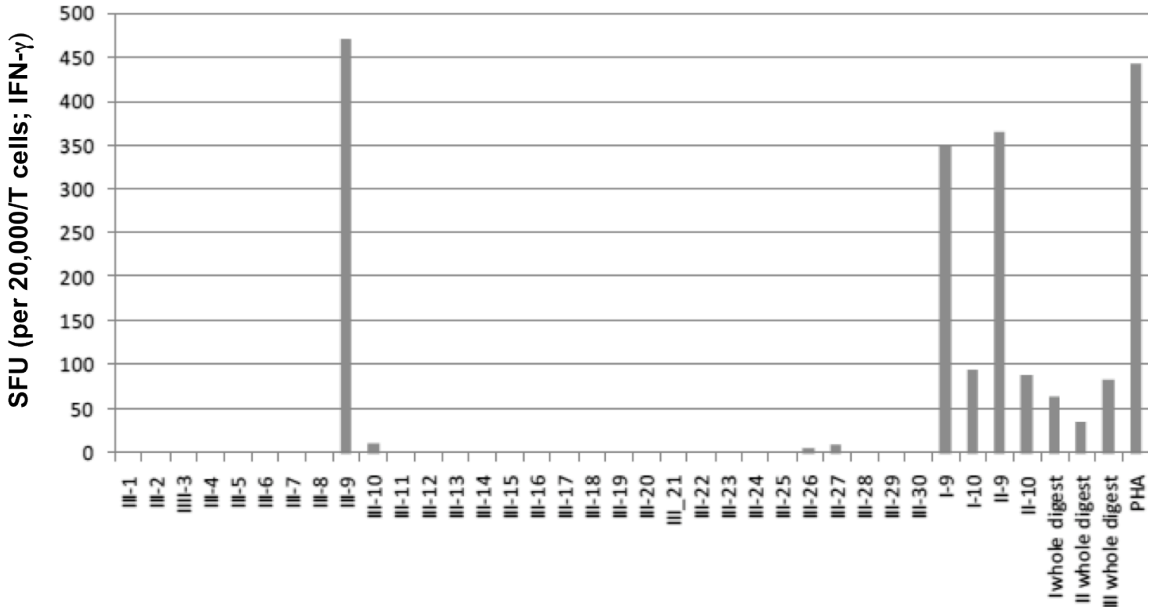

Figure 4.4 - IFN-γ ELISPOT Screen on ΔHspX Pronase RP-HPLC Fractions. Triplicate digests of ΔHspX whole cell lysate (labeled I. II. III.), were separated by C18 reverse phase chromatography to produce 30 fractions from each digest, for a total of 90 fractions. Replicate III, fractions 1-30 are representative of each digest and active fractions 9 and 10 from replicates I and II are also depicted, in comparison to unseparated whole digest. Phytohemagluttanin (PHA) is used a positive control.

Each of the 90 fractions was subjected to peptide and protein identification via liquid chromatography mass spectrometry (LC-MS). Several protein products were identified in each fraction, and results from this analysis were compared to the protein identifications found in previously positive samples (dCW and HspX-18). Results are summarized in Table 4.1.

Table 4.1 - Summary table of Proteins Identified by LC-MS/MS in three positive, pronase digested fractions - delipidated cell wall (dCW), native HspX (nHspX) fraction 18 and ΔHspX fraction 9.

Summary of Native and Recombinant Proteins Tested on D160 1-23

Positive Fractions			MS-IDs
dCW_Pronase	HspX_18	ΔHspX_9	
✔	✔		**HspX**
✔			CFP10
✔			ESAT6
✔			**38 kDa**
✔	✔	✔	*Ag85 Complex*
✔		✔	**Ag85 A**
✔	✔		**Ag85 B**
✔		✔	**Ag85 C**
✔			**GroES**
✔	✔	✔	GroEL2
✔		✔	**19kDa**

LC-MS identified several proteins within each fraction. All of these proteins, except ESAT6 (rec. only) were tested in both native and recombinant variations, as well as overlapping synthetic peptides. No single protein elicited IFN-γ secretion from the donor T cell clone. **Bold =** Both native and recombinant proteins were tested; Plain = Recombinant only (E. coli, 6xHis); *Italic* = Native only.

4.6.3 Further Definition of ΔHspX Fraction 9

Pronase digested ΔHspX generated a positive fraction when separated via RP-HPLC (fraction 9). To determine if the biological activity could be attributed to unique peptide interactions, secondary structure anomalies, or labile post-translational modifications, fraction 9 was boiled in 4% SDS, 1mM DTT for 30 minutes. Figure 4.5 demonstrates that reducing treatment of fraction 9 did not ablate IFN-γ secretion from D160 1-23 and that the individual

proteins identified via LC-MS (Table 4.1) were not responsible for activity. Fraction 9 was further resolved via RP-HPLC using an extended gradient of acetonitrile and subfractions 9A, 9B, 9C were generated and reactivity was measured for subfraction 9A only (data not shown). Peptide fragments ranging from 700 Da to 2100 Da were visualized within the active fraction using matrix-asssted laser desorption ionization, time-of-flight (MALDI-TOF) mass spectrometry to generate a peptide fragment map.

Through the complementary analysis of subfraction 9A with LC-MS/MS peptide sequences could be assigned for several of the ion species seen by MALDI-TOF (Figure 4.6), these included proteins of the Antigen 85 complex, as well as Mpt32 (Rv1860), a protein which had not been identified in other positive fractions.

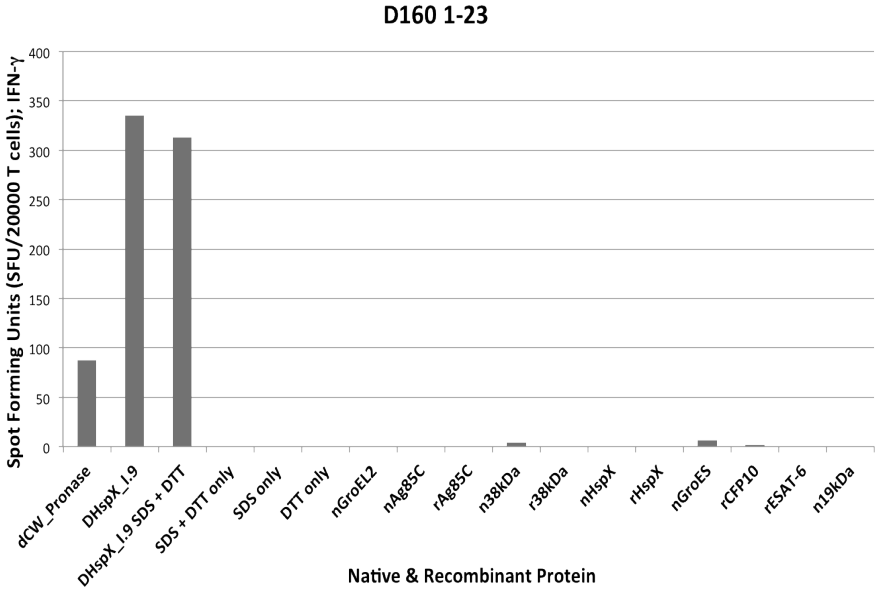

Figure 4.5 - IFN-γ ELISPOT Against ΔHspX Fraction 9 and Select Proteins. Despite subjecting fraction 9 to boiling in 4% SDS, 1mM DTT, biological activity was retained. SFU counts for proteins listed in Table 4.1 demonstrated no reactivity to the selected protein panel.

The T cell reactivity to the Ag85 complex had previously been ruled out, thus testing of both native and recombinant Mpt32 commenced (Figure 4.7).

4.6.4 Native Mpt32 Peptides are Presented in the Context of HLA-E

Mpt32 is a 45kDa alanine and proline-rich secreted glycoprotein (43) and a known CD4 T cell antigen (44). It is secreted into the culture filtrate of actively growing cells. To confirm the results observed from the testing of native whole protein, the ELISPOT assay was repeated with whole and pronase digested Mpt32, as well as whole and pronase digested CFP. Digested Mpt32 gave an overwhelming response in the T cell ELISPOT assay, as did Pronase digested CFP from which Mpt32 is purified from (see Materials & Methods). Mpt32 digested with trypsin induced a response equivalent to undigested protein (Figure 4.8). This was not unexpected as the primary sequence of Mpt32 contains very few Lys (K) or Arg (R) cut sites (Table 4.2).

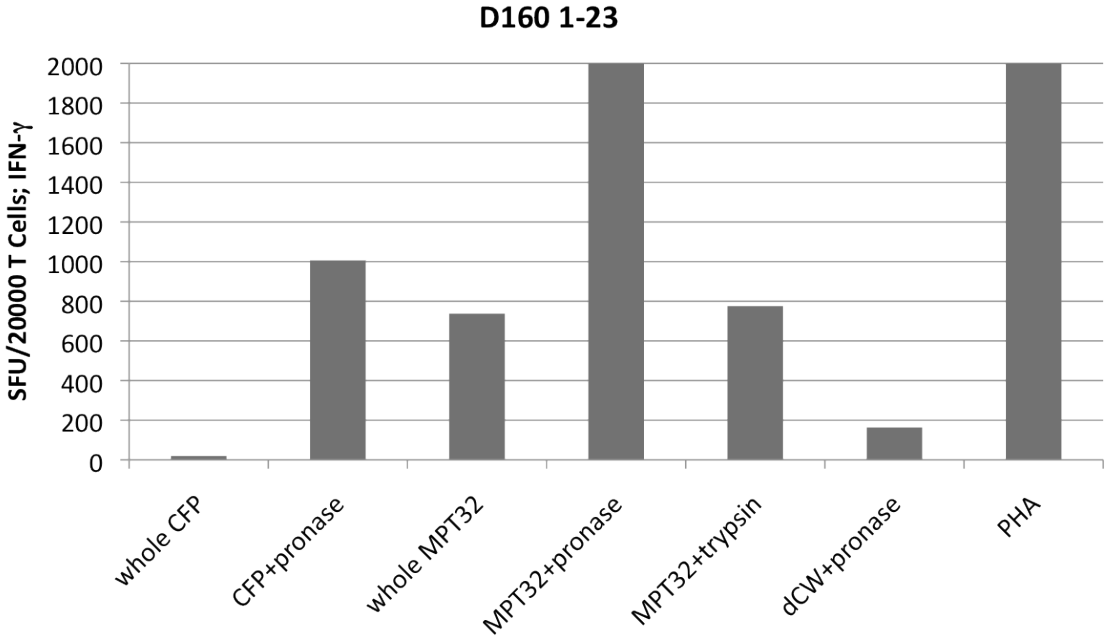

Figure 4.9 - Mpt32 is Presented in the Context of HLA-E. Culture Filtrate Proteins (CFP) from *Mtb*, $H_{37}Rv$ was tested as either a whole protein preparation or digested with pronase. Digestion of CFP with Pronase greatly enhances the T cell reactivity of this subcellular fraction. Along with CFP, native Mpt32 was digested with Pronase and trypsin. SFU for Mpt32+Pronase maxed out the readability of the assay and was determined to be too numerous to count. Mpt32+trypsin gave similar readings to undigested protein. Mpt32 is relatively resistant to trypsin, having less than 10 trypsin cut sites throughout its amino acid sequence**.**

Table 4.2 – Mpt32 Tryptic Peptides Assuming 0-2 Missed Cleavage Events

Missed Cleavage	Primary Sequence of Mpt32 Listed as Trypsin Cut Sites
0	DPEPAPPVPTTAASPPSTAAAPPAPATPVAPPPPAAANTPNAQPGDPN AAPPPADPNAPPPPVIAPNAPQPV**R**
1	DPEPAPPVPTTAASPPSTAAAPPAPATPVAPPPPAAANTPNAQPGDPN AAPPPADPNAPPPPVIAPNAPQPV**R**IDNPVGGFSFALPAGWVESDAAH FDYGSALLS**K**
2	DPEPAPPVPTTAASPPSTAAAPPAPATPVAPPPPAAANTPNAQPGDPN AAPPPADPNAPPPPVIAPNAPQPV**R**IDNPVGGFSFALPAGWVESDAAH FDYGSALLS**K**TTGDPPFPGQPPPVANDT**R**
0	(R)IDNPVGGFSFALPAGWVESDAAHFDYGSALLS**K**
1	(R) IDNPVGGFSFALPAGWVESDAAHFDYGSALLS**K**TTGDPPFPGQPPPVA NDT**R**
2	(R) IDNPVGGFSFALPAGWVESDAAHFDYGSALLS**K**TTGDPPFPGQPPPVA NDT**R**IVLG**R**
0	(K)TTGDPPFPGQPPPVANDT**R**
1	(K)TTGDPPFPGQPPPVANDT**R**IVLG**R**
2	(K)TTGDPPFPGQPPPVANDT**R**IVLG**R**LDQ**K**

The presentation of the Mpt32 epitope in the context of HLA-E was detectable in the ELISPOT assay down to 0.625μg/ml of sample (Figure 4.9). Overlapping synthetic peptides, the non-glycosylated recombinant Mpt32 from *E. coli*, nor the hyperglycosylated recombinant Mpt32 expressed in *M. smegmatis* resulted in a positive response. Therefore, characterization of the native protein and peptide products began with the isolation of a discrete peptide fraction from the Mpt32 Pronase digest (Fraction 8). Fraction 8 consistently induced IFN-γ secretion (Figure 4.10).

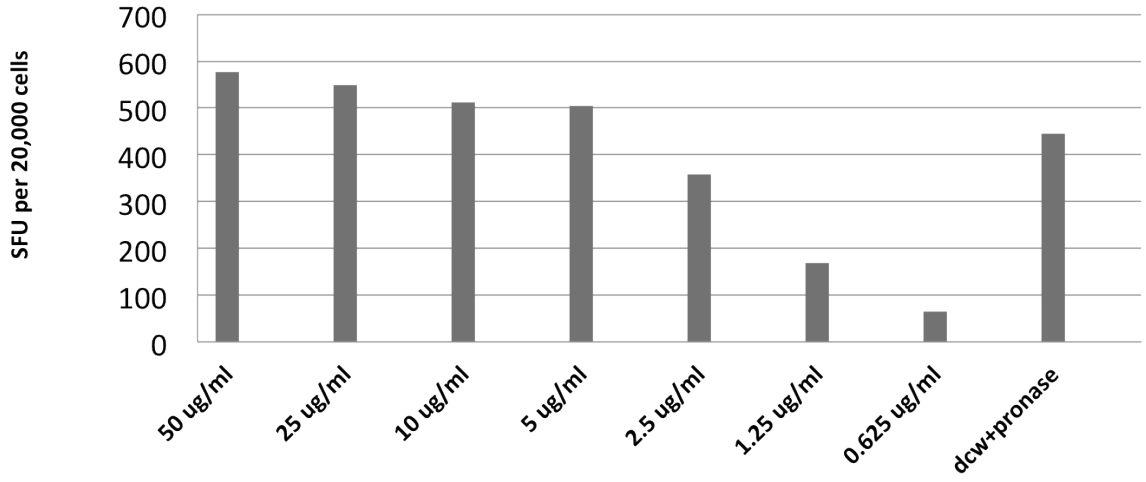

Figure 4.10 – Mpt32 Pronase Digest Presented by HLA-E. The HLA-E restricted T cell clone D160 1-23 was active in the presence of the Mpt32 Pronase antigen preparation, down to a concentration of 0.625 μg/ml.

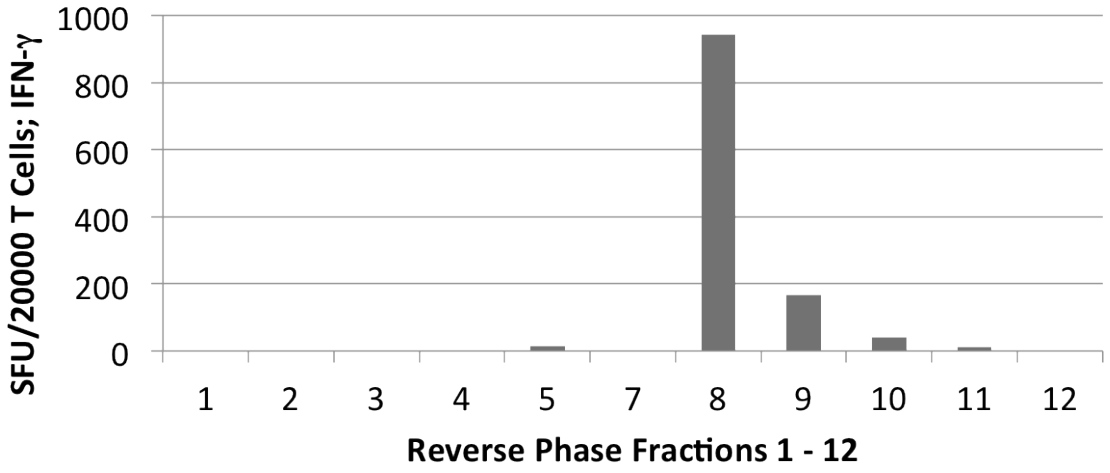

Figure 4.11 - Reverse Phase Fractions of Mpt32 Pronase Digest. Digested Mpt32 was separated over a C18 Reverse Phase column to elute discrete peptide fractions based on peptide hydrophobicity. Fraction 8, retained biological activity specific to the HLA-E T cell clone

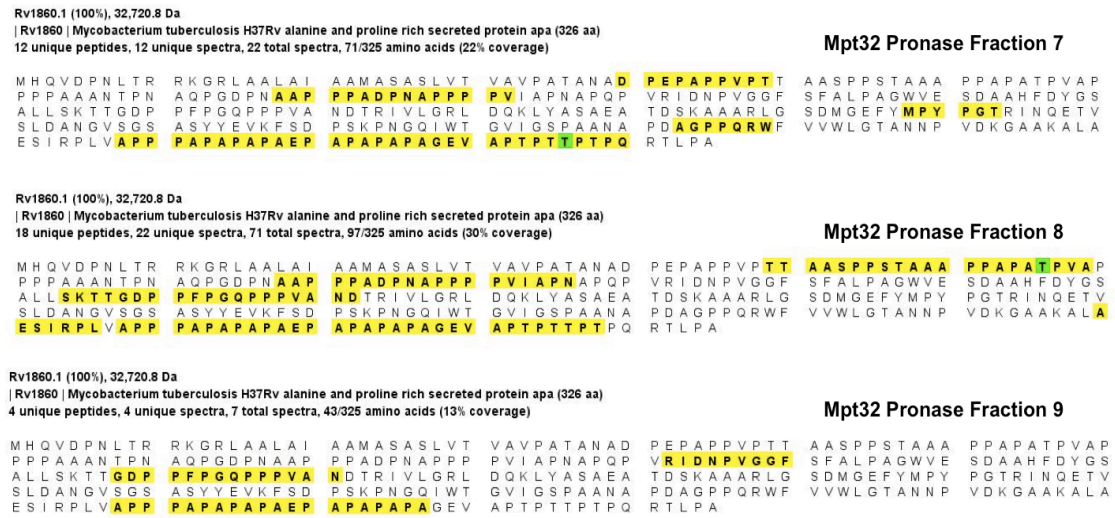

Figure 4.12 - Protein Sequence Coverage of Mpt32 Pronase Generated Peptides. Reverse phase fractions were collected and analyzed via LC-MS/MS to identify peptides present within each fraction. Fractions 7, 8 and 9 share many peptides, however the peptides identified in fraction 8 represent ~ 30% of the mature Mpt32 protein, compared with 22% (Fraction 7) and 13% (Fraction 9). Peptides unique to this fraction are located at the N-terminus where three glycosylated threonines reside (Thr10, Thr18 and Thr27).

4.6.5 Mass Spectrometry Analysis of Mpt32 Fractions 8

LC-MS analysis of fractions 7, 8 and 9 identified a total of 34 unique peptides. Those unique to the biologically active fraction 8, belong to the amino terminus of the protein, which contains several O-mannosylated threonines. Figure 4.11 highlights the protein coverage among the three reverse phase fractions, with Fraction 8 representing approximately 30% of the Mpt32 protein sequence. Unfortunately the full characterization of the N-terminal glycopeptides present in these native fractions has not been successful. Evidence of the N-terminal glycopeptide was detected by MALDI-TOF analysis of a trypsin digested Fraction 8. The parent masses of 3454.162, 3616.374 and 3778.577 ($[M+H]^{+1}$) were visualized and putatively represent the mature N-terminal glycopeptide: 1-DPEPAPPVPTTAASPPSTAAAPPAPATPVAPPPPAA-36, however convincing MS/MS of the peptide has yet to be confirmed by LC-MS/MS.

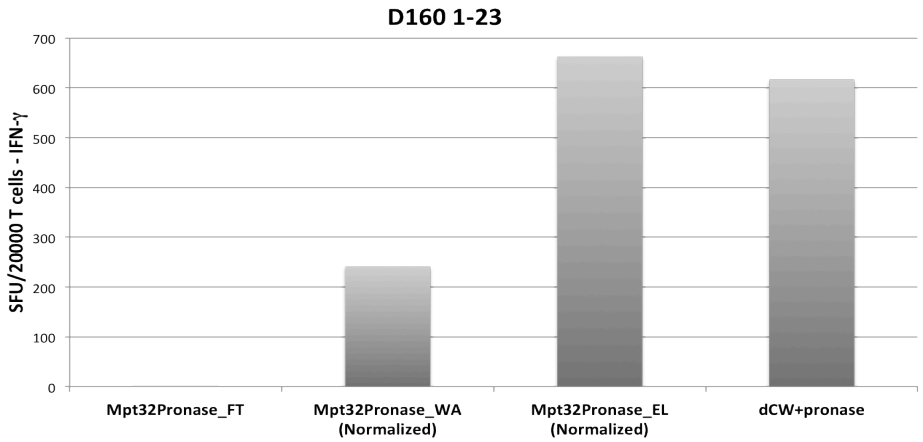

Figure 4.13 - Glycopeptide Enrichment of Mpt32 Pronase. Mpt32 Pronase digest was enriched for glycopeptides using ConA lectin chromatography. FT = flow through of unbound, non-glycosylated peptides, WA= wash of non-specific binders, EL= glycopeptide enriched eluate. Normalized peptide concentration against dCW+Pronase.

4.6.6 Mannosylation State of Mpt32

Overlapping synthetic peptides of those represented in Figure 4.11 were tested and found to be negative for the HLA-E restricted D160 1-23. This led to the hypothesis the mannosylation state of the N-terminal threonines contributed to the CD8 specific T cell response. Therefore, the glycopeptides present within the Mpt32 pronase digest were enriched over a Concavalin-A lectin column. This generated three fractions – flow through of non-glycosylated peptides, wash of excess non-specific binders and an eluate of mannosylated peptides (Figure 4.12). The response to the Con-A eluate was relatively modest in comparison to RP-HPLC purified fractions, however the reactivity to this fraction was consistent with our hypothesis that mycobacterial glycopeptides may be immunostimulatory for disease specific CD8 T cells. Further characterization of this phenomena and the defining epitope responsible for T cell stimulation remain to be elucidated.

4.7 Discussion

The initial efforts to identify the HLA-E restricted antigen began with characterization of the antigenic activity of crude subcellular fractions – cell wall, membrane, cytosol and culture filtrate, isolated from *M. tuberculosis*, H_{37}Rv grown to mid-log phase. Included in these original assays was a preparation of the hydrophobic TX-114 extracted cell wall. The biological activity of a particular fraction was measured and quantified via IFN-γ ELISPOT. These original studies concluded that biological activity was exclusive to the cell wall (either crude preparation or TX-114 soluble) and that activity was not affected if treated with nucleases such as DNase or RNase. Further, it was observed that cytokine secretion was greatly enhanced when the antigen source was pre-digested with the non-specific protease Pronase (18). Enzymatic treatment with additional proteases such as trypsin, chymotrypsion, endopeptidase Glu-C (V-8 protease), and proteinase K abolished T-cell reactivity. The ability of the different cleavage products to either induce INF-γ secretion or ablate it gave key evidence as to the biochemical properties of the HLA-E restricted antigen: 1. The antigen is a protein, 2. The antigen is resistant to specific proteolytic digestion and limited proteolysis with specific proteases such as trypsin, chymotrypsion and V-8 protease leaves large peptide fragments that may be insufficiently targeted to the proteosome for proper processing of the cognate epitope (Lewinsohn et al., unpublished observation) and 3. Biological activity observed in the cell wall (both untreated and TX-114 soluble) may be indicative of a modified protein product. Collaborative efforts between CSU and OHSU resulted in the identification of HspX (Acr, Rv2031c) as the HLA-E restricted antigen. However, synthetic overlapping peptides of this protein were generated and ELISPOT analysis showed no biological activity in response to this single protein library. Despite extensive effort to biochemically define the epitope associated with T cell reactivity, it was later concluded that HspX was not the HLA-E antigen.

Work continued on defining the antigenic protein in an HspX knockout strain of *M, tuberculosis*. Meticulous analysis of highly purified peptide fractions identified peptides belonging to the immunodominant, secreted glycoprotein Mpt32. Mpt32 is O-mannosylated at positions Threonine (Thr) 10, Thr18, Thr27 and Thr277. Thr10 and 18 are decorated with an α-D-(1,2) linked mannobiose, Thr27 with a single mannose and Thr277 was shown to have either a single α-D-Man, mannobiose or triose (43). Evidence presented in the preceding sections suggests that the HLA-E T cell receptor is capable of recognizing a post-translationally mannosylated glycopeptide of Mpt32. Future work remains in the full definition of the cognate epitope presented by HLA-E. In the absence of synthetically modified peptide products, we must continue to pursue alternative means of proving that the N-terminus of the protein elicits a specific biological activity and that this activity is dependent on the state of mannosylation of a particular threonine.

We are pursuing this line of work in several complementary ways: (1) Enzymatic treatment of the active fraction Mpt32 fraction 8, with α-mannosidase (2) Chemical alteration of the O-mannosylated N-terminus via acetylation or beta-elimination and (3) Genetic manipulation of the sites of glycosylation of Mpt32. Preliminary experiments for both enzymatic and chemical treatment (via acetylation) have been performed, however the precise characterization of the modifying treatments must be done to confirm proper alteration of the mannosyl moiety. We continue to pursue a genetic approach in which select mannosylation sites at the N and C-terminus of the protein are mutated to amino acids that cannot accept a mannose modification. The O-mannosylation of serine or threonine amongst mycobacterial proteins has been characterized for T cell antigens such as the 19kDa lipoglycoprotein (45) and the B cell antigen SodC (46) however many more glycosylated proteins are known to exist in mycobacteria (47).

Unfortunately for the purposes of epitope screening, the synthetic production of O-mannoslyated Thr is cost-prohibitive and the chemistry involved is not typically performed in a commercial setting, therefore its success rate unknown (personal communication, Proteomics Manager New England Peptide). However, similar to the broad characterization of CD8 T cell antigens, understanding the immune response to mycobacterial glycoproteins and the processing and presentation of glycopeptides to T cells could aid in the development of T cell-based diagnostic tools as well as supplement the search for novel vaccine candidates. Including the development of unique subunit or peptide vaccines designed specifically to target infected cells and promote the secretion of pro-inflammatory cytokines (IFN-γ, TNF-α).

Furthermore, knowledge of glycopeptide specific responses to mycobacterial infection would aid in the study of the unique CD8 T cell response in cases of active versus latent TB. Glycopeptide specific T cells could also be characterized by their *ex vivo* frequency as well as their cellular phenotype and function (45) within the infected host. Furthermore, we recognize the limitations of the bottom up approach to identify cognate epitopes within active fractions. One issue involved with this method is the generation of relatively complex peptide fractions that are only biologically active when digested with non-specific proteases such as pronase. This hinders the identification of peptides through automated search algorithms such as SEQUEST and Mascot and favors *de novo* or manual sequencing of the mass spectrometry data. Also related to this limitation is the reliance on often miss-annotated genomic information from which protein and peptide databases are created and searched. Bottom up approaches are also complicated with advancements in modern instrumentation in which the sensitivity has increased dramatically over time. Mass spectrometers are now able to detect and measure the presence of peptides over an expanded dynamic range. However this detection is not related to abundance of

the peptide in question. To overcome these limitations a more targeted approach would be most helpful in identifying cognate epitopes. With these methods the peptides being presented are eluted off of the antigen presenting molecule itself and directly analyzed via LC-MS (46-49).

4.8 Materials and Methods

4.8.1 Bacterial Growth and Isolation of Subcellular Fractions

M. tuberculosis, H_{37}Rv was cultured in glycerol-alanine-salts (GAS) media for 14 days at 37^{o}C, followed by harvest of supernatant and cells, lysis and subcellular fractionation as described previously(50). Briefly, culture filtrate proteins were separated from cells via filtration through a 0.22 μm membrane and concentrated using a stirred cell apparatus under N_2, followed by dialysis into 0.01 M NH_4HCO_3. Whole cells were inactivated by γ-irradiation and resuspended in breaking buffer (PBS, 1mM EDTA) supplemented with protease inhibitors, DNase and RNase for serial passage through a French pressure cell (Thermo Scientific, MA). Whole cell lysate was dialyzed into 0.01 M NH_4HCO_3, and centrifuged at 27,000 x g at 4^{o}C for 1h to generate cell wall. The supernatant was subsequently centrifuged at 100,000 x g at 4oC for 4 h to recover cell membrane. Cell wall, membrane and cytosol were all dialyzed and protein quantified via BCA (ThermoPierce, CA).

4.8.2 Delipidation of the Cell Wall

The cell wall fraction was subjected to organic extraction to remove non-covalently attached lipids and glycolipids (42). Lyophilized cell wall was subjected to two extractions of 2 h each followed by one 18 h extraction with chloroform/methanol (2:1, v/v) at a ratio of 30 mL/g of cell wall. Extractions were performed at 22 °C with agitation. Centrifugation at 27,000g for 30 min was performed to collect cell wall material. The 2:1 extracted cell wall was dried under N2 and further extracted twice for 2 h and one 18 h extraction with chloroform/methanol/water

(10:10:3, v/v/v) each at 22 °C. The fully delipidated cell wall was dried under N_2 and resuspended in PBS, 0.1% ASB-14 pH 7.4 to maximize solubility. Cell wall protein was quantified by bicinchoninic acid (BCA) assay (Thermo Pierce).

4.8.3 Purification of Native Proteins for Elispot Testing

4.8.3.1 Ag85 complex and subunit purification

Ammonium sulfate $(NH_4)_2SO_2$ is added to the culture filtrate proteins (CFP) to a final concentration of 40% and proteins allowed to separate for 1 hour at 4^oC. The CFP is then centrifuged at 16,000 x g for 30 minutes and supernatant removed for a second salt cut at 70%. The 40% pellet is resuspended in phenyl sepharose buffer A (10 mM KH2PO4 (pH 7.2), 1 mM EDTA, 1 mM DTT (1L) and a gradient is run through buffer B (10 mM Tris-Base (pH 8.9), 1 mM EDTA, 1 mM DTT (1L), and then through C (10 mM Tris-Base (pH 8.9), 1 mM EDTA, 1 mM DTT, 50% ethylene glycol (v/v)) to purify Ag85 complex. and Ag85A, B and C separately.

4.8.3.2 Mpt32

Unbound material from the phen-seph column (flow through) is run over concavalin-A lectin sepharose to purify Mpt32. Lyophilized starting material is resuspended in ConA binding buffer (50mM KH_2PO_4, 500mM NaCl, 1mM each of $MgCl_2$, $CaCl_2$, $MnCl_2$ and 1mM DTT). Bound protein is eluted with excess 1M Methyl α-D-mannopyranoside). Mpt32 is polished over a C18 reverse phase column in 20mM ammonium bicarbonate, 1mM DTT and eluted off with an increasing gradient of 20mM ammonium bicarbonate, 70% acetonitrile. This step removes contaminating glycolipids from the Con-A lectin preparation.

4.8.3.3 38kDa and GroES

The pellet generated from the centrifugation of the 70% ammonium sulfate cut of the CFP is also enriched over ConA to purify the 38kDa glycoprotein. Unbound ConA proteins are then used in the purification of GroES.

4.8.3.4 19kDa

The 19kDa glycolipoprotein partitions into the detergent layer of the TX-114 soluble proteins of the cell wall subcellular fraction. It is purified using preparative SDS-PAGE and whole gel electro-elution.

4.8.3.5 HspX

HspX is purified by differential extraction of a 40,000 x g pellet of the WCL with 1% N-octylthioglucoside in PBS. The detergent soluble fraction is dialyzed, lyophilized, and product resuspended in a buffer of 7.25 M urea, 0.4% 3-10 pharmalytes, 1.6% 4-7 pharmalytes, 1% N-octylthioglucoside, 2mM DTT and incubated. Soluble proteins are separated by preparative iso-electric focusing, and fractions (of ~ 0.5 pH units) collected, screened, and those enriched in a-crystallin are pooled. Polishing of the 16 kDa protein is achieved by size exclusion chromatography with a Sephadex-75 column and isocratic elution of the 16 kDa protein in a buffer of 3M Urea, 20mM Tris, 0.15M NaCl, 0.1% n-octylthioglucoside, 0.02% NaN3, pH 7.6. Purified a-crystallin is qualified and placed in inventory for distribution.

4.8.4 Pronase Digestion

Pronase (250 ku, EMD Millipore) was resuspended in digestion buffer - PBS, 0.1% ASB-14 for subcellular fractions or 0.2M NH_4HCO_3 for purified proteins, at a concentration of 5.0 mg/ml. Digestion was performed at 37^{o}C, 16h at an enzyme to substrate ratio of 1:20 for all

samples. Enzymatic reactivity was quenched with rapid freeze thaw cycles at -80°C. Crude peptide lysates were kept frozen and shipped on dry ice to OHSU for T cell reactivity.

4.8.5 γ-IFN ELISPOT

All ELISPOT analysis was performed at OHSU as previously described (18). Briefly, 96-well nitrocellulose backed plates were coated as recommended by the manufacturer with 10μg/ml capture mouse anti-IFN-γ overnight at 4°C. Plates were then washed six times with PBS/0.05% Tween 20 (Sigma-Aldrich), blocked with RPMI 1640/10% HS for 1 h at room temperature. T cell clones (200) and APC (20,000) are added, and the plate incubated overnight at 37°C. Plates are extensively washed and anti-IFN-γ secondary antibody conjugated to HRP is added. AEC developer substrate is added and the reaction stopped by washing with distilled water.

4.8.6 Reverse Phase HPLC

Whole protein digests (1mg to 5 mg) were separated on Waters Alliance 2695 HPLC (Waters Inc, Millford, MA) using a Grace Vydac Everest 300Å monomeric C18 column (4.6 x 150mm, 238EV5115). Fractions were collected every 2 minutes over a gradient of 0-50% B over 30 minutes, 0.5ml/min. Buffer A – 0.1% TFA in water, Buffer B – 0.1% TFA, 90% acetonitrile. For fraction 9A, the gradient was modified to begin at 15% B and increased to 30% B over 30 minutes. Fractions were collected every 3 minutes at a flow rate of 0.5 ml/min. For all peptide separations fractions were concentrated using vacuum centrifugation and resuspended at an approximate concentration of 0.1 ug/ul in Buffer A. Sample were kept frozen at -80°C and shipped on dry ice to OHSU for ELISPOT testing.

4.8.7 Concavalin-A Lectin Chromatography Of Peptides

3 mg of purified Mpt32 was subjected to Pronase digestion as described above (concentration of 1 mg/ml). Peptides were incubated with rocking, in the presence of 2 ml of Con-A conjugated sepharose slurry (i.e. 1ml settled resin) (GE Life Sciences) at 4^{o}C, for 16 h in a 15ml reaction tube. Resin-peptide mix was poured into an open column and allowed to settle. Flow through was collected. Resin was washed with 10 column volumes of binding buffer and collected. The peptides were eluted off with 10 CV of elution buffer. Collected fractions were exchanged back into 0.1% TFA after processing through solid-phase sep-pak columns (Waters) following manufacturer's instructions.

4.8.8 Mass Spectrometry

All mass spectrometry was performed in collaboration with the Proteomics and Metabolomics Facility located at Colorado State University. Liquid Chromatography Mass Spectrometry: Peptides were separated on a nanospray column (Zorbax C18, 5 mm, 75 mm ID 6 150 mm column). Samples were eluted into a LTQ linear ion trap mass spectrometer (Thermo) using a gradient of 0-100% B (A= 3% ACN, 0.1% formic acid; B =100% ACN, 0.1% formic acid) at a flow rate of 300 nL/min for 60 minutes. MALDI-TOF: 1 ml of peptide sample is mixed with 1 ml of a mixed matrix solution (10mg/ml α-Cyano-4-hydroxycinnamic acid (CHCA) and 10mg/ml 2,5-Dihydroxybenzoic acid (DHB) in 50% ACN, 0.1% TFA. The mixing of matrices enhanced intensity of ion peaks and allowed for the visualization of glycopeptides. The mixture is spotted on the MALDI target and allowed to air dry. The sample is analyzed by an Ultraflex-TOF/TOF mass spectrometer (Bruker Daltonics, Billerica, MA) in positive ion, reflector mode using a 25 kV accelerating voltage. External calibration is done using a peptide calibration mixture (4 to 6 peptides) on a spot adjacent to the sample. The raw data is processed in the FlexAnalysis software (version 3.3, Bruker Daltonics).

4.8.9 Peptide Identification

LC-MS Database Searching: Tandem mass spectra were extracted, charge state deconvoluted and deisotoped by Xcalibur version 1.7 SP2. All MS/MS samples were analyzed using Sequest (Thermo Fisher Scientific, San Jose, CA, USA; version v.27, rev. 11). Sequest was set up to search the MtbV3_Reverse database (7992 entries) assuming non-specific cleavage. Parameters for both search engines were set to a fragment ion mass tolerance of 1.5 Da and a parent ion tolerance of 3.0 Da. Oxidation of methionine was specified as variable modification. For glycopeptides, a modification of mannose (+162 da) and mannobiose (+324) were set as variable modification on threonine.

MALDI-TOF spectra were collected and select ions were chosen for fragmentation. Peptide sequences were elucidated via *de novo* calculation of amino acid fragments. M+H single charged ion masses were matched if identified using LC-MS protein identification.

4.8.10 Criteria For Protein Identification

Scaffold (version Scaffold_3.5.1, Proteome Software Inc., Portland, OR) was used to validate MS/MS based peptide and protein identifications. Peptide identifications were accepted if they exceeded specific database search engine thresholds. Sequest identifications required at least deltaCn scores of greater than 0.2 and XCorr scores of greater than 1.8, 2.0, 3.0 and 4.0 for singly, doubly, triply and quadruply charged peptides. Protein identifications assigned within each peptide fraction were accepted if they contained at least 2 identified peptides. All peptide spectra were manually inspected for sequence coverage and signal to noise.

4.9 Summary and Conclusion

Future work with Mpt32 will include specific manipulation of the glycosyl moieties located at the N-terminus of the protein through site-directed mutagenesis, and the overexpression of these mutants in an Mpt32 knockout strain of *M. tuberculosis*. Processing and

presentation of glycopeptides to T cells is known to occur and is of interest in the realms of autoimmunity (51), medical mycology (52) and cancer (53). Furthermore the glycosylation state of Mpt32 is known to affect its immunostimulatory properties for CD4 T cell activation (54,55). However the direct contribution of each glycopeptide in the generation of a cell-mediated immune response has never been elucidated. The genetic approach targeting the highly modified N-terminus of Mpt32, will unequivocally demonstrate the requirement of glycosylation for HLA-E restricted T cell activation.

The work described herein is the first description of a glycopeptide being presented in the context of HLA-E and the first mycobacterial protein to be presented in this manner. From this work, we also acknowledge the limitations of the "bottom up" proteomics approach to antigen discovery, especially when post-translational modifications and digestion with nonspecific proteases produce biologically active fractions. Mass spectrometry has been used quite successfully to characterize a number of both classically and non-classically restricted T cell epitopes (46,47). These approaches directly address peptides eluted off of the MHC molecule itself, and are aided with the use of synthetic peptide pools generated via analysis of predicted binders. Ultimately these studies identify epitopes with amino acid sequences similar to other, previously characterized epitopes (i.e. the MHC class 1a leader sequences in the case of HLA-E). However binding avidity of MHC Ia or b molecules is not necessarily an indicator of an immunostimulatory phenotype, therefore validation of defined epitopes must be performed. In our case, for this specific T cell clone, neither predictive algorithms nor the generation of synthetic peptides aided in the identification of the presented epitope. Ultimately, a proteomics based approach was successful in the identification of an HLA-E restricted *M. tuberculosis* antigen. However experimental approaches such as those described above are time consuming

and the characterization of one T cell clone, narrow in scope. A broader approach could involve the development of a library of T cell clones, irrespective of HLA haplotype, that are reactive against the glycoproteome of *M. tuberculsosis* grown under different physiological states. However, the continuing development and integration of proteome-scale prediction tools (29), which are trained to identify specific immune characteristics has the potential to increase the efficiency, accuracy and relevance of proteomics based antigen discovery efforts.

4.10 LITERATURE CITED

1. Flynn, J. L., and Chan, J. (2001) *Annu Rev Immunol* **19**, 93-129
2. Ernst, J. D. (2012) *Nat Rev Immunol* **12**, 581-591
3. Dannenberg, A. M., Jr. (1989) *Rev Infect Dis* **11 Suppl 2**, S369-378
4. Ottenhoff, T. H., Lewinsohn, D. A., and Lewinsohn, D. M. (2008) Human CD4 and CD8 T Cell Responses to Mycobacterium tuberculosis: Antigen Specificity, Function, Implications and Applications. in *Handbook of Tuberculosis: Immunology and Cell Biology* (Kaufmann, S. H., and Britton, W. J. eds.), Wiley-VCH, Germany. pp 119-143
5. Nyendak, M. R., Lewinsohn, D. A., and Lewinsohn, D. M. (2009) *Curr Opin Infect Dis* **22**, 174-182
6. Pai, M., Dheda, K., Cunningham, J., Scano, F., and O'Brien, R. (2007) *Lancet Infect Dis* **7**, 428-438
7. Lewinsohn, D. A., Heinzel, A. S., Gardner, J. M., Zhu, L., Alderson, M. R., and Lewinsohn, D. M. (2003) *Am J Respir Crit Care Med* **168**, 1346-1352
8. van Pinxteren, L. A., Cassidy, J. P., Smedegaard, B. H., Agger, E. M., and Andersen, P. (2000) *Eur J Immunol* **30**, 3689-3698
9. Lewinsohn, D. A., Lines, R. A., and Lewinsohn, D. M. (2002) *Am J Respir Crit Care Med* **166**, 843-848
10. Lewinsohn, D. M., Alderson, M. R., Briden, A. L., Riddell, S. R., Reed, S. G., and Grabstein, K. H. (1998) *J Exp Med* **187**, 1633-1640
11. Lalvani, A., Brookes, R., Wilkinson, R. J., Malin, A. S., Pathan, A. A., Andersen, P., Dockrell, H., Pasvol, G., and Hill, A. V. (1998) *Proc Natl Acad Sci U S A* **95**, 270-275
12. Mohagheghpour, N., Gammon, D., Kawamura, L. M., van Vollenhoven, A., Benike, C. J., and Engleman, E. G. (1998) *J Immunol* **161**, 2400-2406
13. Tan, J. S., Canaday, D. H., Boom, W. H., Balaji, K. N., Schwander, S. K., and Rich, E. A. (1997) *J Immunol* **159**, 290-297
14. Turner, J., and Dockrell, H. M. (1996) *Immunology* **87**, 339-342
15. Beckman, E. M., Melian, A., Behar, S. M., Sieling, P. A., Chatterjee, D., Furlong, S. T., Matsumoto, R., Rosat, J. P., Modlin, R. L., and Porcelli, S. A. (1996) *J Immunol* **157**, 2795-2803
16. Moody, D. B., Guy, M. R., Grant, E., Cheng, T. Y., Brenner, M. B., Besra, G. S., and Porcelli, S. A. (2000) *J Exp Med* **192**, 965-976
17. Rosat, J. P., Grant, E. P., Beckman, E. M., Dascher, C. C., Sieling, P. A., Frederique, D., Modlin, R. L., Porcelli, S. A., Furlong, S. T., and Brenner, M. B. (1999) *J Immunol* **162**, 366-371
18. Heinzel, A. S., Grotzke, J. E., Lines, R. A., Lewinsohn, D. A., McNabb, A. L., Streblow, D. N., Braud, V. M., Grieser, H. J., Belisle, J. T., and Lewinsohn, D. M. (2002) *J Exp Med* **196**, 1473-1481
19. Gold, M. C., Eid, T., Smyk-Pearson, S., Eberling, Y., Swarbrick, G. M., Langley, S. M., Streeter, P. R., Lewinsohn, D. A., and Lewinsohn, D. M. (2012) *Mucosal immunology*
20. Kjer-Nielsen, L., Patel, O., Corbett, A. J., Le Nours, J., Meehan, B., Liu, L., Bhati, M., Chen, Z., Kostenko, L., Reantragoon, R., Williamson, N. A., Purcell, A. W., Dudek, N. L., McConville, M. J., O'Hair, R. A., Khairallah, G. N., Godfrey, D. I., Fairlie, D. P., Rossjohn, J., and McCluskey, J. (2012) *Nature* **491**, 717-723

21. Lewinsohn, D. A., Winata, E., Swarbrick, G. M., Tanner, K. E., Cook, M. S., Null, M. D., Cansler, M. E., Sette, A., Sidney, J., and Lewinsohn, D. M. (2007) *PLoS Pathog* **3**, 1240-1249
22. Cease, K. B., Margalit, H., Cornette, J. L., Putney, S. D., Robey, W. G., Ouyang, C., Streicher, H. Z., Fischinger, P. J., Gallo, R. C., DeLisi, C., and et al. (1987) *Proc Natl Acad Sci U S A* **84**, 4249-4253
23. McMichael, A. J., Michie, C. A., Gotch, F. M., Smith, G. L., and Moss, B. (1986) *The Journal of general virology* **67 (Pt 4)**, 719-726
24. Townsend, A. R., Rothbard, J., Gotch, F. M., Bahadur, G., Wraith, D., and McMichael, A. J. (1986) *Cell* **44**, 959-968
25. Sospedra, M., Pinilla, C., and Martin, R. (2003) *Methods* **29**, 236-247
26. Frieder, M., and Lewinsohn, D. M. (2009) *Methods Mol Biol* **524**, 369-382
27. Camus, J. C., Pryor, M. J., Medigue, C., and Cole, S. T. (2002) *Microbiology* **148**, 2967-2973
28. Lewinsohn, D. M., Lewinsohn, D. A., and Swarbrick, G. M. (2009) Proposal: Large Scale T-Cell Epitope Discovery, RFP-BAA-NIHAI200832. Oregon Health & Science University, Oregon
29. Bremel, R. D., and Homan, E. J. (2010) *Immunome Res* **6**, 8
30. Park, B., and Mori, T. (2012) What Makes a Good CD8 T cell Antigen? in *T Cell Based Identification of Human CD8+ Tuberculosis Epitopes* (Lewinsohn, D. M., Lewinsohn, D. A., and Swarbrick, G. M. eds.), Portland, OR
31. Bremel, R. D., and Homan, E. J. (2010b) *Immunome Res* **6**, 7
32. Moody, D. B., Ulrichs, T., Muhlecker, W., Young, D. C., Gurcha, S. S., Grant, E., Rosat, J. P., Brenner, M. B., Costello, C. E., Besra, G. S., and Porcelli, S. A. (2000) *Nature* **404**, 884-888
33. Aebersold, R., and Mann, M. (2003) *Nature* **422**, 198-207
34. Andersen, P., Askgaard, D., Ljungqvist, L., Bentzon, M. W., and Heron, I. (1991) *Infect Immun* **59**, 1558-1563
35. Dhiman, N., Verma, I., and Khuller, G. K. (1997) *Microbiol Immunol* **41**, 495-502
36. Laal, S., Samanich, K. M., Sonnenberg, M. G., Zolla-Pazner, S., Phadtare, J. M., and Belisle, J. T. (1997) *Clin Diagn Lab Immunol* **4**, 49-56
37. Samanich, K. M., Belisle, J. T., Sonnenberg, M. G., Keen, M. A., Zolla-Pazner, S., and Laal, S. (1998) *J Infect Dis* **178**, 1534-1538
38. Covert, B. A., Spencer, J. S., Orme, I. M., and Belisle, J. T. (2001) *Proteomics* **1**, 574-586
39. Dobos, K. M., Spencer, J. S., Orme, I. M., and Belisle, J. T. (2004) *Methods Mol Med* **94**, 3-17
40. Pietra, G., Romagnani, C., Manzini, C., Moretta, L., and Mingari, M. C. *J Biomed Biotechnol* **2010**, 907092
41. Grotzke, J. E., Harriff, M. J., Siler, A. C., Nolt, D., Delepine, J., Lewinsohn, D. A., and Lewinsohn, D. M. (2009) *PLoS Pathog* **5**, e1000374
42. Wolfe, L. M., Mahaffey, S. B., Kruh, N. A., and Dobos, K. M. (2010) *J Proteome Res* **9**, 5816-5826
43. Dobos, K. M., Khoo, K. H., Swiderek, K. M., Brennan, P. J., and Belisle, J. T. (1996) *J Bacteriol* **178**, 2498-2506
44. Espitia, C., and Mancilla, R. (1989) *Clin Exp Immunol* **77**, 378-383

45. Caccamo, N., Guggino, G., Meraviglia, S., Gelsomino, G., Di Carlo, P., Titone, L., Bocchino, M., Galati, D., Matarese, A., Nouta, J., Klein, M. R., Salerno, A., Sanduzzi, A., Dieli, F., and Ottenhoff, T. H. (2009) *PLoS One* **4**, e5528
46. Hunt, D. F., Henderson, R. A., Shabanowitz, J., Sakaguchi, K., Michel, H., Sevilir, N., Cox, A. L., Appella, E., and Engelhard, V. H. (1992) *Science* **255**, 1261-1263
47. Lemmel, C., and Stevanovic, S. (2003) *Methods* **29**, 248-259
48. Millo, E., Pietra, G., Armirotti, A., Vacca, P., Mingari, M. C., Moretta, L., and Damonte, G. (2007) *Journal of immunological methods* **322**, 128-136
49. Millo, E., and Damonte, G. (2009) *Current pharmaceutical design* **15**, 3325-3335
50. Gu, S., Chen, J., Dobos, K. M., Bradbury, E. M., Belisle, J. T., and Chen, X. (2003) *Mol Cell Proteomics* **2**, 1284-1296
51. Corthay, A., Backlund, J., and Holmdahl, R. (2001) *Ann Med* **33**, 456-465
52. Fernandez-Alvarez, A., Marin-Menguiano, M., Lanver, D., Jimenez-Martin, A., Elias-Villalobos, A., Perez-Pulido, A. J., Kahmann, R., and Ibeas, J. I. *PLoS Pathog* **8**, e1002563
53. Vlad, A. M., Muller, S., Cudic, M., Paulsen, H., Otvos, L., Jr., Hanisch, F. G., and Finn, O. J. (2002) *J Exp Med* **196**, 1435-1446
54. Horn, C., Namane, A., Pescher, P., Riviere, M., Romain, F., Puzo, G., Barzu, O., and Marchal, G. (1999) *J Biol Chem* **274**, 32023-32030
55. Romain, F., Horn, C., Pescher, P., Namane, A., Riviere, M., Puzo, G., Barzu, O., and Marchal, G. (1999) *Infect Immun* **67**, 5567-5572

V. Chapter V: A Chemical Proteomics Approach to Profiling the ATPome of *Mycobacterium tuberculosis*

A modified version of this chapter entitled A Chemical Proteomics Approach to Profiling the ATPome of *Mycobacterium tuberculosis* was published March 5, 2013 in Molecular & Cellular Proteomics.

Lisa M. Wolfe[1], Usha Veeraraghavan[2], Susan Thomas[3], Stephan Schurer[3], Krister Wennerberg[4,5], Robert Reynolds[5], Gurdyal S. Besra[2], William Suling[5] Karen M. Dobos[1]
Affiliations:
1. Department of Microbiology, Immunology and Pathology, Colorado State University, Fort Collins, CO
2. School of Biosciences, College of Life and Environmental Sciences, University of Birmingham, Edgbaston, Birmingham, UK.
3. Miller School of Medicine, University of Miami, Miami ,FL
4. Institute for Molecular Medicine Finland FIMM, University of Helsinki, PO Box 20, Helsinki 00014, Finland
5. Department of Chemistry, University of Alabama, Tuskaloosa, AL

5.1 Introduction

M. tuberculosis is a highly adaptable organism, capable of surviving in multiple metabolic states. The cellular envelope of the bacillus provides integrity and protection from the insults encountered within the host environment and extracellular milieu. It is the dynamic nature of this structure that facilitates the survival of the bacterium throughout its diverse life cycle, regulating the physiological processes needed to acquire or expel the necessary molecules it needs to function in these altered capacities. Furthermore, *M. tuberculosis* infection remains a significant global health burden and the emergence of multi-drug resistant (MDR) and extensively drug resistant (XDR) cases continue to increase (1). In a state of non-replicating persistence, the bacilli are able to maintain their integrity during the course of drug treatment. Novel chemotherapeutics for the treatment of disease are needed in order to target the bacteria during multiple metabolic states. By elucidating the functionally linked protein networks that mediate transitions into and out of particular disease and drug resistant states, the effective treatment time and complexity of antibiotic regimens used against *M. tuberculosis* can be reduced.

5.2 ATP Pathways and Chemoproteomics

Beyond their function in cancer immune evasion and deregulated growth, ATP dependent catalytic pathways are being scrutinized for their roles in pathogen biology and their control over metabolic processes in multiple physiological stages. The recent description of ATP-competitive enzyme inhibitors as a novel class of antitubercular drugs(2-5) has bolstered interest in the identification of bacterial enzymes that utilize ATP - given that these enzymes may be feasible targets in the discovery and design of novel small molecule inhibitors of tuberculosis. Furthermore, elucidating ATP-dependent catalytic pathways present during different metabolic disease states is critical for understanding mechanisms of virulence and pathogenesis. Manipulation of these pathways via novel chemotherapeutic strategies could not only increase the effectiveness of drug treatment in MDR/XDR cases, but may possess potent *in vivo* efficacy against bacilli exhibiting multiple, often resistant phenotypes within the host (6). The study of kinases and other ATP-binding proteins (chaperones, ATPases, synthases and other metabolic enzymes) has become important in elucidating the roles of ATP-dependent pathways in the pathogenesis of cancer and other mechanisms of dysregulated growth. The large-scale profiling of such networks is facilitated with the use of active-site nucleotide probes (7). Traditionally studies have utilized this chemical proteomics approach to elucidate global kinase profiles of cell lines, as well as to map interaction networks using clinically approved kinase inhibitors (imatinib/dasatinib) (8,9), experimental inhibitor scaffolds (10) and/or broad spectrum kinase inhibitors such as staurosporine (7). Here we describe a chemical proteomics method that is designed to capture the full array of adenosine nucleotide-binding proteins, the ATPome, of *M. tuberculosis,* H_{37}Rv. ATP-binding proteins are labeled with desthiobiotin-conjugated ATP (Figure 5.1) and subsequently digested with trypsin. Labeled peptides are captured via

streptavidin affinity chromatography subjected to LC-MS/MS for the identification of ATP-labeled proteins.

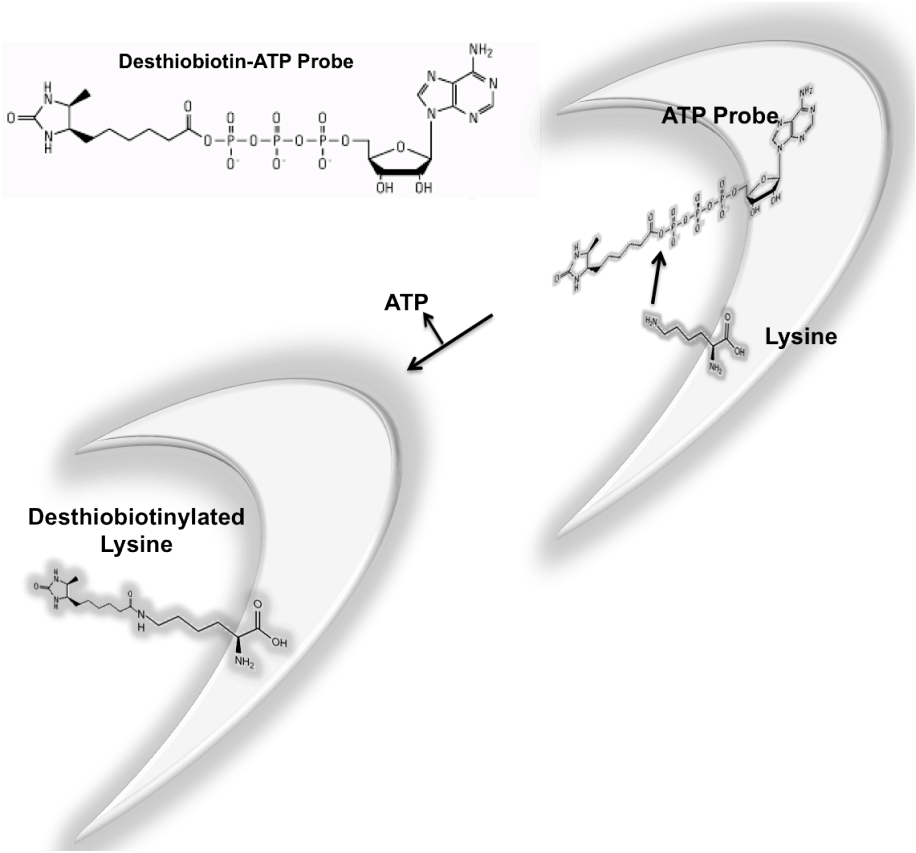

Figure 5.1 - Mechanism of Action of the Desthiobiotin Conjugated ATP Active Site Probe. This method utilizes a desthiobiotin-conjugated ATP (A) as a molecular probe in which target enzymes are covalently modified within characteristic active sites – in this case the nucleotide binding domains of kinases and other ATP-binding proteins (B). As the nucleotide probe binds, the phosphodiester bond linking the desthiobiotin affinity tag to the ATP molecule undergoes a nucleophilic attack initiated by the free amine of conserved active-site lysine residues. The desthiobiotin-linked lysine can then be considered covalently modified and detectable via mass spectrometry.

5.3 Functionally Linked Protein Networks Associated with the Cellular Envelope of *M. tuberculosis.*

The ATP-binding proteins of *M. tuberculosis* comprise a very unique and functional subset of the mycocaterial proteome. The proteins capable of nucleotide binding are targets for small-molecule inhibitors because they are functionally distributed among a variety of enzyme classes and play essential, often non-redundant roles in basic cellular metabolism. The functional nature of this sub-proteome supports the general hypothesis that proteins of the mycobacterial ATPome provide necessary mechanisms of adaptation utilized in the maintenance of growth under a variety of microenvironmental conditions. The physiology and molecular architecture of the *M. tuberculosis* cellular envelope is a dynamic and highly biologically active scaffold. This structure and the signaling processes associated with maintaining it, also work to sustain viability of the pathogen during intracellular infection within host macrophages. This stage of infection is successful because the bacillus is capable of manipulating its environment in such a way as to evade the host immune response. Physiochemically, its molecular composition is responsible for the organism's inherent tolerance to antimicrobial therapies, which prolongs effective treatment time. The molecular events that regulate the entrance into and exit out of different metabolic states remain poorly defined, although there are several hypotheses that remain testable (70,71). Approximately one-quarter of the *M. tuberculosis* genome remains functionally characterized as hypothetical (72). Recent re-annotation and prediction efforts concluded that the majority of hypothetical proteins could be redistributed among the categories of Small Molecule Metabolism, Cell Wall Processes, and Lipid Metabolism (54,55). As described in Chapter III, proteins extracted from the cell wall were significantly (70% of CWP) biased towards these categories. The utilization of ATP by these hypothetical proteins may provide further insight into their cellular roles and enzymatic functions and may lead to key insights in the study of

pathogen-specific mechanisms of persistence and reactivation. The idea of ATP binding and hydrolysis acting as a molecular switch controlling the transition into hypoxia has been observed in the study of mammalian models of low-oxygen conditions (56,57). For mycobacteria, one class of ATP-binding proteins, the Universal Stress Proteins (USPs), may be involved in the responses to changes in environmental and nutrient conditions leading to variations in virulence and adaptation. Several years ago, Drumm and Chan et al., investigated the nucleotide binding capabilities of Rv2623 and its role as a universal stress protein (USP)(58). Gene deletion mutants in *M. tuberculosis*, Erdmann demonstrated a hypervirulent phenotype that failed to enter into dormancy within susceptible Hartley guinea pigs. Disruption of the ATP-binding site of Rv2623 resulted in similar attenuated phenotypes described for the deletion mutants. It was hypothesized that the binding of ATP by similar USPs could be a regulatory mechanism utilized in the transition from normal growth to an oxygen-poor state of dormancy.

In this study we set out to (1) provide a comprehensive profile of ATP-binding proteins and catalytic pathways present within the native proteome of *M. tuberculosis,* H_{37}Rv. (2) Identify patterns of differential abundance of ATP-binding proteins present under normally growing and a hypoxia-induced, dormant state and (3) provide evidence of either competitive or transient ATP-binding states of nucleotide bound proteins for the identification of potential therapeutic targets of ATP-competitive kinase inhibitors. Through this work we identified essential gene products critical to survival, adaptation and the development of drug resistance. Many protein families were functionally linked with processes of cell wall integrity and metabolism. This work may lead to the identification of novel therapeutic targets and provide insight into the use of ATP-competitive small molecule inhibitors in the treatment of tuberculosis.

5.4 Results

5.4.1 Comprehensive Proteomic Analysis of the *M. tuberculosis* ATPome

A shotgun proteomics analysis was performed on the enriched subproteome of dethiobiotin-labeled ATP-binding proteins. We identified a total of 176 proteins, 122 were labeled via the nucleotide probe (Appendix III). Selective labeling was further validated by ranking the tagged proteins using a metric of ATP-labeling based on protein and peptide confidence levels greater than 90% as well as manual interpretation of spectral quality for each peptide sequence labeled with a desthiobiotin tag(11). This ranking accounted for the variation among identified peptides in signal to noise levels and sequence coverage. Proteins were listed by the quality of spectra demonstrating a desthiobiotin-labeled lysine (differential modification of lysine (K) of +196 Da). Low confidence peptide spectra exhibited no less than 90% peptide confidence with a minimum of one assigned peptide (number identified – 21). Medium confidence peptide spectra had a peptide score between 90 – 95% with two or more assigned peptides (number identified – 20). Proteins determined to have labeling in the high confidence range, exhibited greater than 95% peptide confidence and had two or more unique peptides assigned for identification (number identified – 81). A no-probe, streptavidin-only control was performed to account for proteins that have inherent biotin-like domains and may non-specifically interact with the streptavidin capture resin. Results from the control experiments revealed only a few proteins - chaperones GroEL2 (Rv0440), DnaK (Rv0350), HspX(Rv2031c), the acyl-carrier protein AcpM (Rv2244), the peptidyl-prolyl-cis-trans-isomerase PpiA (Rv0009) and the naturally biotinylated acetyl carboxylase AccA3 (Rv3285) - bound streptavidin non-specifically in addition to being confidently labeled with the desthiobiotin probe. In the case of GroEL2, DnaK and HspX we believe the promiscuous binding to the affinity resin was due to the high abundance of each protein and their chaperoning function. PpiA, while not present in

high abundance aids in protein folding as well (12). AccA3 most likely bound the streptavidin capture resin by virtue of its biotin-binding domain (13). AcpM is functionally associated with AccA3 as both proteins are involved in long chain fatty acid synthesis. Their association in this pathway is visualized via the STRING database (v. 9.0 2012; (14)), with curated pathway interactions in the BioCyc v. 16.1 pathway collection (http://biocyc.org/ECOLI/NEW-IMAGE?type=PATHWAY&object=PWY0-881). Overall, the utilization of the active site nucleotide probe to capture ATP-binding proteins resulted in a highly enriched proteome, with nearly 60% of the *M. tuberculosis* ATPome deemed essential to *in vitro* growth (15,16).

5.4.2 Functional Annotation of Labeled Proteins

Functional annotation of the proteins using web-based services as well as in-house implementation of a variety of tools were carried out on various levels in the laboratory of Dr. Stephan Shürer, with final analysis provided by Dr. Susan Idicula-Thomas, at the University of Miami, Miller School of Medicine. This work was performed in collaboration with Dr. Schürer in order to identify the functional domains and domain families that were selectively labeled and enriched via our chemical proteomic techniques. To begin, the amino acid sequence of each identified protein and covalently labeled peptide sequence was provided to Dr. Idicula-Thomas and subjected to an InterPro (http://www.ebi.ac.uk/interpro/) pattern search to identify functional domains and associate these regions with domain families (Pfams) (17).

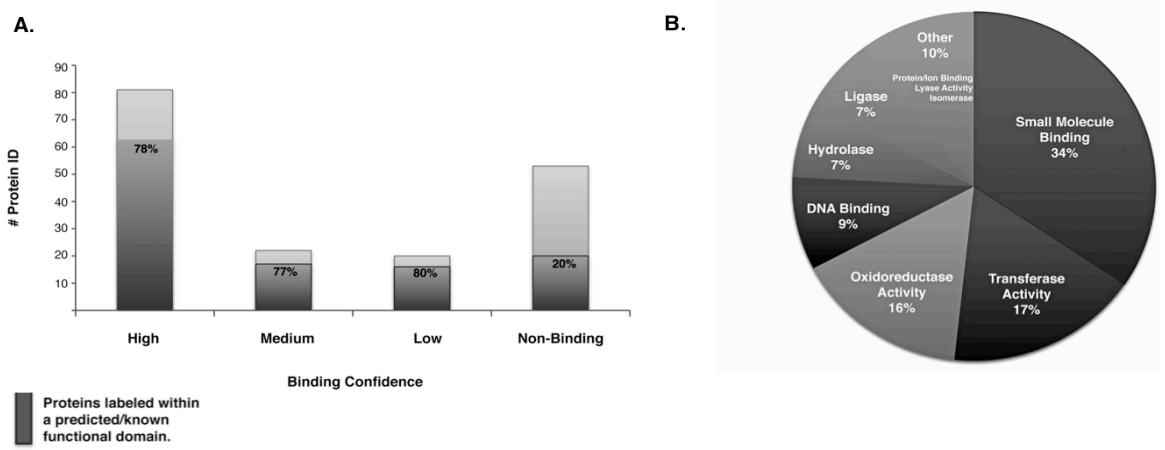

Figure 5.2 - Primary Sequence Analysis of ATP-Binding Peptides and Proteins. Each protein sequence was submitted for *in silico* analysis through InterPro (http://www.ebi.ac.uk/interpro/) and sorted via Gene Ontology (molecular function). A - Approximately 80% of the ATP-binding (i.e. desthiobiotin labeled) proteins were found to be associated with ATP within an annotated functional domain. Binding confidence - high, medium, low - was empirically determined associated with the quality of labeled peptide specta (i.e confident sequence coverage and low signal to noise). B – The functions of small molecule binding, transferase and oxidoreductase activity described the

It was determined that 13 ATP-associated Pfams were represented in the ATPome dataset across all ranges of labeling confidence (low to high, n=122) and none were represented in

proteins identified, but not labeled with nucleotide probe (n = 54).Among the ATP-associated Pfams were proteins involved with ATP synthesis (PF00006) and peptidoglycan synthesis (Mur Ligase (PF01225)), as well as protein kinases (Pkinase PF00069). Overall, approximately 80% of the ATPome had peptides that were labeled within a known or predicted ATP-associated Pfam domain (Figure 5.2A). Domain descriptions were then mapped to the GO term "ATP Binding" (GO ID:0005524) and associated with specific enzyme activities (Figure 5.2B). -

The majority of enzyme activities could be associated with small molecule binding (34%), transferase (17%) and oxidoreductase (16%) activity (Figure 5.2). The protein functional category Small Molecule Metabolism includes proteins involved in lipid metabolism, intermediary metabolism and regulatory proteins such as kinases. Categorization of the *M. tuberculosis* ATPome by broad protein function (i.e. Tuberculist Functional Categories - http://tuberculist.epfl.ch/) revealed that second and third to category 7 – Intermediary Metabolism, category 1 – Lipid Metabolism and category 10 – Conserved Hypotheticals,

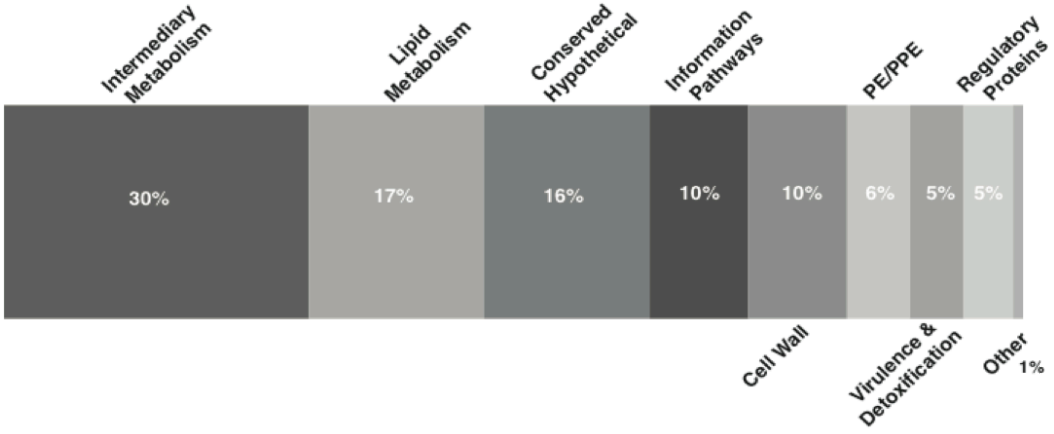

Figure 5.3 - Organization of the *M. tuberculosis*, $H_{37}Rv$ ATPome by Functional Category. Representing nearly 50% of the ATPome were proteins within functional category 7 – Intermediary Metabolism and functional category 2 – Lipid Metabolism. Conserved Hypotheticals (functional category 10) were the third most represented group.

respectively, represented a collective 33% of the enriched proteome (Figure 5.3). Proteins involved in fatty acid and mycolic acid biosynthesis (category 1) are of interest due to their key roles in maintenance of the cell envelope architecture and the essentiality of their encoding genes (16). A complete list of labeled and unlabeled protein IDs and their corresponding functional categories is provided in Appendix III.

5.4.3 Differential Abundance of Proteins in Normally Growing vs Hypoxic Bacteria:

We utilized the active site nucleotide probes to selectively capture and enrich for the *M. tuberculosis* ATPome under different growth conditions. Normally growing cells were compared to cells grown in limited oxygen conditions (see Materials and Methods). Overall, we identified 61 differentially abundant proteins. Results were calculated for protein abundance changes that had a p-value less than 0.05 (11,18). The log of the fold change (FC) values was plotted against the calculated p-values in order to visualize the distribution of proteins between the two growth states (Figure 5.4). Proteins were determined to be differentially abundant if they were only present within the hypoxic data set or had a calculated ration (log fold change, logFC) of spectral counts less than zero for proteins with higher abundance in hypoxic samples (Table 5.1). Proteins with a logFC greater that zero were in higher abundance during normal growth (Table 5.2)

Table 5.1 – Proteins with Increased Abundance During Hypoxia

Identified Proteins	Accession Number	Molecular Weight	Fisher's Exact Test (P-Value)	Normal NSAF*	Hypoxic NSAF	Log Fold Change
conserved hypothetical protein	Rv2623	32 kDa	(0.0000)	1	14	-3.81
acyl-[acyl-carrier protein] desaturase desA1	Rv0824c	39 kDa	(0.0000)	1	12	-3.58
hypothetical protein acg	Rv2032	37 kDa	(0.0003)	1	8	-3.00
phosphoribosylaminoimidazole-succinocarboxamide synthase purC	Rv0780	33 kDa	(0.0010)	1	7	-2.81
ketol-acid reductoisomerase ilvC	Rv3001c	36 kDa	(0.0010)	1	7	-2.81
acyl-[acyl-carrier protein] desaturase desA2	Rv1094	31 kDa	(0.0000)	2	12	-2.58
isocitrate lyase icl	Rv0467	47 kDa	(0.0033)	1	6	-2.58
10 kda culture filtrate antigen esxB	Rv3874	11 kDa	(0.0100)	1	5	-2.32
40 kda secreted L-alanine dehydrogenase ald	Rv2780	39 kDa	(0.0000)	29	134	-2.21
citrate synthase I gltA2	Rv0896	48 kDa	(0.0320)	1	4	-2.00
DNA gyrase subunit B gyrB	Rv0005	78 kDa	(0.0320)	1	4	-2.00
transcription termination factor rho	Rv1297	65 kDa	(0.0000)	4	15	-1.91
iron-regulated heparin binding hemagglutinin hbhA	Rv0475	22 kDa	(0.0004)	3	11	-1.87
ATP synthase alpha chain atpA	Rv1308	59 kDa	(0.0011)	4	11	-1.46
macrolide-transport ATP-binding protein ABC transporter	Rv2477c	62 kDa	(0.0005)	5	13	-1.38
6-phosphofructokinase pfkA	Rv3010c	37 kDa	(0.0140)	5	9	-0.85
fatty-acid oxidation protein fadB	Rv0860	76 kDa	(0.0010)	10	17	-0.77
heat shock protein hspX	Rv2031c	16 kDa	(0.0000)	133	223	-0.75
conserved hypothetical protein	Rv1738	11 kDa	(0.0019)	10	16	-0.68
bifunctional polyribonucleotide nucleotidyltransferase gpsI	Rv2783c	80 kDa	(0.0026)	13	18	-0.47
adenosylhomocysteinase sahH	Rv3248c	54 kDa	(0.0000)	45	53	-0.24
iron-regulated aconitate hydratase acn	Rv1475c	102 kDa	(0.0490)	10	11	-0.14

* A value of 1, indicates a normalized spectral abundance factor (NSAF) of 0 in hypoxic samples

Table 5.2 – Proteins with Increased Abundance During Normal

Identified Proteins	Accession Number	Molecular Weight	Fisher's Exact Test (P-Value)	Normal NSAF	Hypoxic NSAF*	Log Fold Change
polyphosphate kinase ppk	Rv2984	83 kDa	(0.0000)	66	1	6.04
phosphoglycerate kinase pgk	Rv1437	43 kDa	(0.0000)	57	2	4.83
acyl-CoA dehydrogenase fadE4	Rv0231	63 kDa	(0.0002)	23	1	4.52
dehydrogenase	Rv3389c	30 kDa	(0.0010)	19	1	4.25
19 kda lipoprotein antigen precursor lpqH	Rv3763	15 kDa	(0.0021)	17	1	4.09
immunogenic protein mpt64	Rv1980c	25 kDa	(0.0021)	17	1	4.09
hypothetical protein	Rv2319c	32 kDa	(0.0021)	17	1	4.09
aminomethyltransferase gcvT	Rv2211c	40 kDa	(0.0021)	17	1	4.09
superoxide dismutase sodA	Rv3846	23 kDa	(0.0031)	16	1	4.00
cysteinyl-tRNA synthetase 1 cysS1	Rv3580c	52 kDa	(0.0031)	16	1	4.00
DNA polymerase I polA	Rv1629	98 kDa	(0.0031)	16	1	4.00
transmembrane serine/threonine-protein kinase E pknE	Rv1743	61 kDa	(0.0031)	16	1	4.00
iron-regulated short-chain dehydrogenase/reductase	Rv3224	30 kDa	(0.0006)	26	2	3.70
fatty-acid-CoA ligase fadD23	Rv3826	63 kDa	(0.0099)	13	1	3.70
leucyl-tRNA synthetase leuS	Rv0041	108 kDa	(0.0150)	12	1	3.58
pyridoxamine 5-phosphate oxidase pdxH	Rv2607	25 kDa	(0.0150)	12	1	3.58
conserved hypothetical protein	Rv2624c	29 kDa	(0.0002)	34	3	3.50
aldehyde dehydrogenase	Rv0458	55 kDa	(0.0066)	19	2	3.25
acetyl-/propionyl-CoA carboxylase alpha subunit accA1	Rv2501c	71 kDa	(0.0001)	45	5	3.17
fatty-acid-CoA ligase fadD7	Rv0119	55 kDa	(0.0020)	27	3	3.17
electron transfer flavoprotein beta subunit fixA	Rv3029c	28 kDa	(0.0020)	27	3	3.17
fatty-acid-CoA ligase fadD36	Rv1193	50 kDa	(0.0460)	9	1	3.17
transferase	Rv1201c	33 kDa	(0.0460)	9	1	3.17
phosphoribosylamine-glycine ligase purD	Rv0772	44 kDa	(0.0028)	26	3	3.12
pyruvate kinase pykA	Rv1617	51 kDa	(0.0021)	31	4	2.95
low molecular weight protein antigen cfp2	Rv2376c	17 kDa	(0.0250)	15	2	2.91
glutamyl-tRNA	Rv3009c	55 kDa	(0.0000)	83	12	2.79
transmembrane serine/threonine-protein kinase H pknH	Rv1266c	67 kDa	(0.0036)	33	5	2.72
conserved alanine rich protein	Rv2744c	29 kDa	(0.0480)	13	2	2.70
transmembrane serine/threonine-protein kinase D pknD	Rv0931c	70 kDa	(0.0005)	51	8	2.67
conserved hypothetical protein	Rv2159c	36 kDa	(0.0320)	18	3	2.58
ATP-dependent protease ATP-binding subunit clpC1	Rv3596c	94 kDa	(0.0000)	181	32	2.50
glutamine synthetase glnA2	Rv2222c	50 kDa	(0.0280)	22	4	2.46
30S ribosomal protein S1 rpsA	Rv1630	53 kDa	(0.0360)	21	4	2.39
succinyl-CoA synthetase beta chain sucC	Rv0951	41 kDa	(0.0170)	40	9	2.15
endopeptidase ATP binding protein chain B clpB	Rv0384c	93 kDa	(0.0003)	131	33	1.99
glutamine synthetase glnA1	Rv2220	54 kDa	(0.0200)	65	18	1.85
10 kda chaperonin groES	Rv3418c	11 kDa	(0.0049)	136	41	1.73

* A value of 1, indicates a normalized spectral abundance factor (NSAF) of 0.

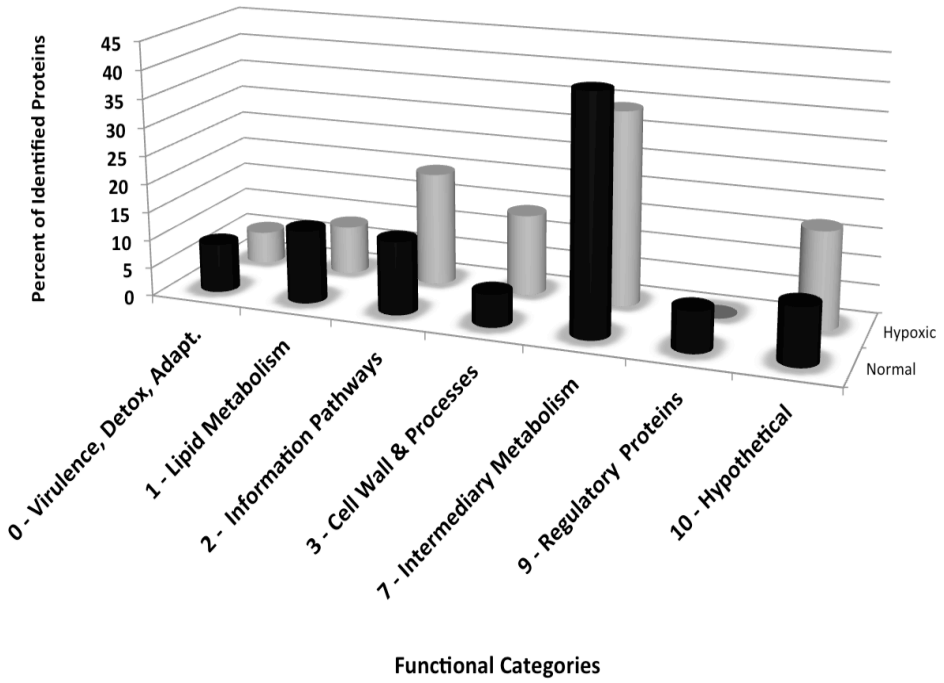

Figure 5.4 - Comparison of Functional Categories
Desthiobiotin-labeled proteins found to be differentially abundant between Normal and Hypoxic growth were sorted based on functional category (Tuberculist - http://tuberculist.epfl.ch/). Categories 4 (Stable RNAs), 5 (Insertion Sequences & Phages) and 6 (PE/PPE) were not represented in our ATPome dataset and were omitted for clarity.

During dormancy and hypoxic growth, *M. tuberculosis* undergoes changes in gene expression that typically involve the upregulation of enzymes involved in alternative metabolic pathways (i.e. Glyoxylate Bypass) and those observed to be under the control of the dormancy regulon DosR. The list of proteins in Table 5.2 includes the gene products HspX (Rv2031c), Acg (Rv2032c), TB31.7 (Rv2623), Rv2624c and Rv1738. These proteins are directly regulated or co-expressed with the response regulator DosR (Rv3133c) (19). Beyond the level of individual genes we observe patterns of protein function that change between Normal and Hypoxic growth (Functional Categories 2, 3, 9 and 10) and those that remain relatively stable (Categories 0, 1, and 7) (Figure 5.5)

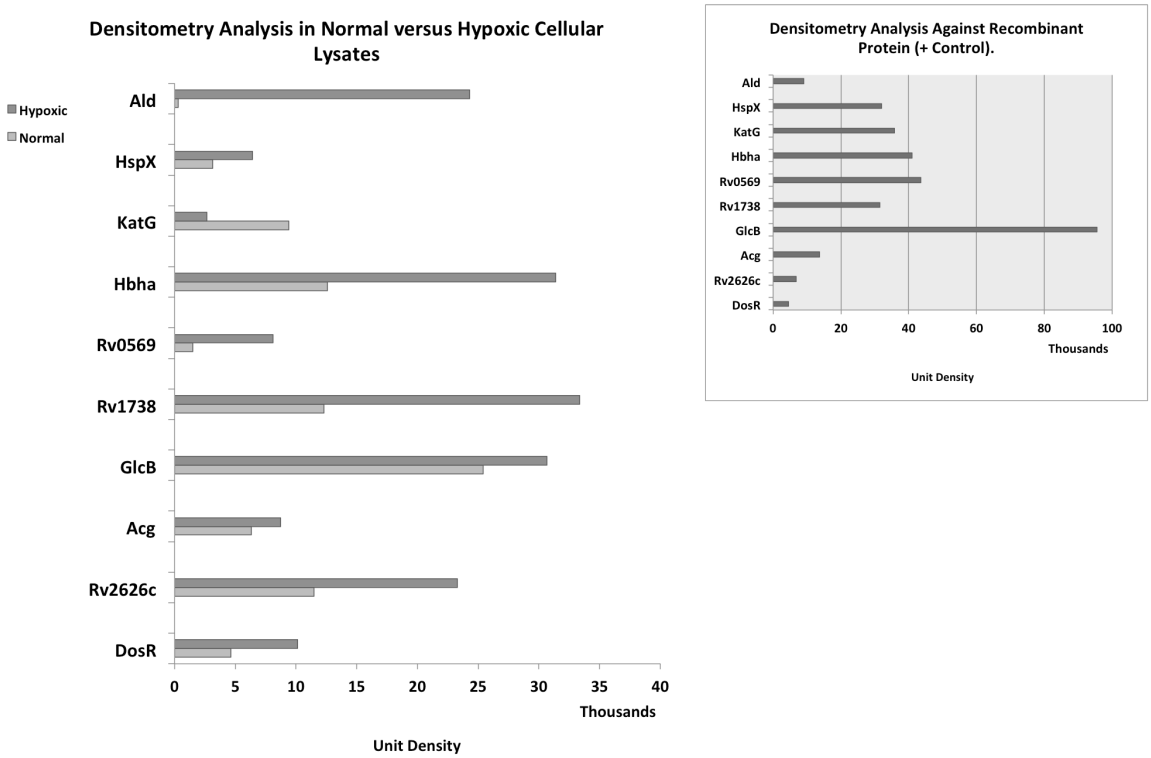

Figure V.5 – Densitometry Analysis. Several proteins found to be differentially abundant by NSAF were probed for immunoreactivity via western blot. The densitometry analysis of these blots corroborates the differences in protein levels between Normal and Hypoxic cells. The densitometry of each protein with its corresponding positive control (recombinant protein) is reported (top panel).

Indeed, isocitrate lyase (*icl*, Rv0467) the enzyme that catalyzes the reversible cleavage of isocitrate to glyoxylate and succinate (20) and has a role in the growth, survival and persistence of *M. tuberculosis* in macrophages and mice (21) was found to be labeled with desthiobiotin-ATP on Lys322 (PFAM: PF00463) and differentially abundant during hypoxic growth. The second enzyme involved in the glyoxylate cycle – malate synthase G (*glcB*, Rv1837c) was also labeled with our active site probe, however in this study its differencial abundance in hypoxic grown cultures was not significant (p-value > 0.60). We did, however confirm increased protein levels of GlcB via western blot (Figure 5.6). It is well known that the expression of alanine dehydrogenase (*ald, Rv2780*) is also upregulated during the growth of *M. tuberculosis* under low

oxygen conditions (20). It has recently been shown in *M. tuberculosis* and previously in *M. smegmatis* that alanine dehydrogenase is responsible for both glycine and alanine dehydrogenase activities (22,23). The main role of Ald is to generate L-alanine for peptidoglycan and protein synthesis (22). Both Icl and Ald are unique to bacteria with no human homologs, making them attractive drug targets. While no inhibitors for Ald have been reported, Icl inhibitors active against dormant and logarithmically grown mycobacteria are 3-nitro propionamides and 5-nitro-2-furoic acid hydrazones (24). Densitometry analysis of immunoblots of several proteins found to be in higher abundance during hypoxia via mass spectrometry, confirmed differences in protein levels (Figure 5.6).

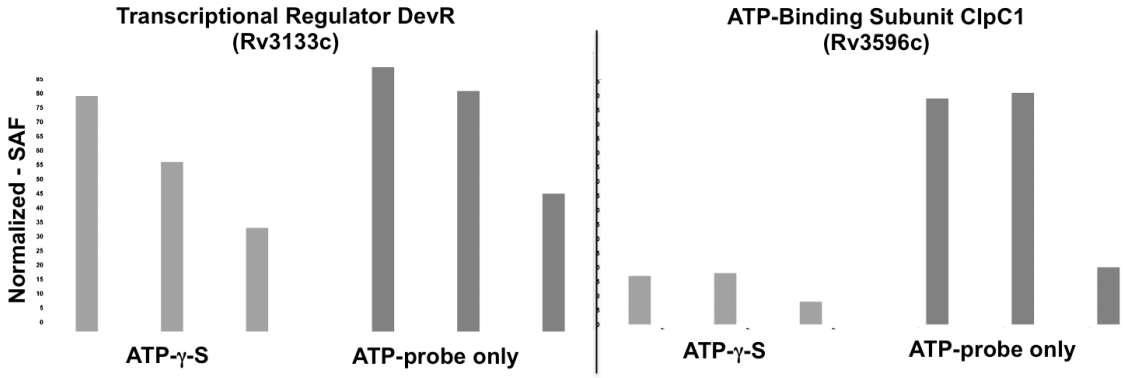

Figure V.6 – NSAF Profiles of ATP-labeled Peptides in the Presence/Absence of Excess ATP-γ-S. Transient binding of ATP was observed in many proteins, including the DNA-binding transcriptional regulator DevR (left). In the presence of excess ATP-g-S, the normalized spectral count profile of the ATP-binding subunit of ClpC1 were significantly reduced in cultures grown under normal conditions (right).

5.4.4 ATP-binding Properties of the *M. tuberculosis* ATPome

In addition to describing the use of ATP by essential enzymes in the bacterial proteome (sections 5.2.1-2), and identifying those proteins that demonstrated differential abundance patterns between normal and hypoxic states of growth, the third and final goal of this work was to characterize proteins whose ATP-binding function may be utilized in the development of novel ATP-competitive antibiotics. The desthiobiotin labeling of active sites has been used in the

study of many kinase inhibitor compounds that are clinically approved in the treatment of certain cancers. In these studies the binding of the nucleotide probe is quantified in the presence or absence of the drug of choice. These experiments work under the assumption that for a specific compound – target interaction, the ATP-competitive inhibitor compound will compete out the binding of the nucleotide probe. Our preliminary work into the chemical proteomics approach for identifying ATP-competitive inhibitors quantified the labeling of proteins in both normal and hypoxic states, in the presence of excess ATP (ATP-γ-S). Adenosine 5'-[γ-thio] triphosphate tetralithium (ATP-γ-S) is a non-hydrolysable analog of ATP. As the binding of ATP to various protein subunits and active sites can be very dynamic, it has been advantageous to utilize non-hydrolysable ATP analogs to identify true ATP-binding states of proteins (25). With this set of experiments we asked these two questions (Figure 5.7): 1. Which proteins bind ATP transiently (i.e. are the binding of ATP-γ-S and the binding of desthiobiotin-ATP interchangeable)? 2. Which proteins have a significantly reduced capacity to bind nucleotide probe in the presence of excess ATP-γ-S? Desthiobiotin-labeled peptides were quantified and those found to have a significant fold change differences between samples labeled in the presence of excess ATP-γ-S are listed in Tables 5.3 (Hypoxic) and 5.4 (Normal).

Table 5.3 – Competitive and Transient Binders of Desthiobiotin ATP in Hypoxic Cultures.

Competitive Binders of Desthiobiotin-ATP	Gene #	Mol.Wt.	P-Value	AveNSAF_ATP-γ-S*	AveNSAF_ATP**	Fold Change
fatty-acid oxidation protein fadB	Rv0860	76 kDa	(0.0039)	1	6.33	6.33
inosine-5-monophosphate dehydrogenase guaB3	Rv3410c	39 kDa	(0.0056)	1	6.00	6.00
transcription termination factor rho	Rv1297	65 kDa	(0.0079)	1	5.67	5.67
macrolide-transport ATP-binding protein ABC transporter	Rv2477c	62 kDa	(0.0160)	1	5.00	5.00
endopeptidase ATP binding protein chain B clpB	Rv0384c	93 kDa	(0.0031)	2.5	11.67	4.67
iron-regulated heparin binding hemagglutinin hbhA	Rv0475	22 kDa	(0.0320)	1	4.33	4.33
adenosylhomocysteinase sahH	Rv3248c	54 kDa	(0.0065)	5.5	18.33	3.33
chaperone protein dnaK	Rv0350	67 kDa	(0.0001)	16	50.00	3.13
adenylate kinase adk	Rv0733	20 kDa	(0.0400)	8	20.33	2.54
acyl-[acyl-carrier protein] desaturase desA1 ❖	Rv0824c	39 kDa	(0.2200)	2	4.67	2.33
ATP synthase beta chain atpD ❖	Rv1310	53 kDa	(0.1500)	5	11.33	2.27
bifunctional acetyl-/propionyl-coenzyme A carboxylase alpha chain accA3❖	Rv3285	64 kDa	(0.2600)	3	6.33	2.11
Transient Binders of Desthiobiotn-ATP						
50S ribosomal protein L7/L12 rplL	Rv0652	13 kDa	(0.3100)	4.5	8.67	1.93
10 kda chaperonin groES	Rv3418c	11 kDa	(0.3600)	8	14.33	1.79
ATP-dependent protease ATP-binding subunit clpC1	Rv3596c	94 kDa	(0.5000)	7	11.33	1.62
60 kda chaperonin 1 groEL1	Rv3417c	56 kDa	(0.5100)	3.5	5.67	1.62
conserved hypothetical protein	Rv2623	32 kDa	(0.5600)	3.5	5.33	1.52
40 kda secreted L-alanine dehydrogenase ald	Rv2780	39 kDa	(0.0520)	36	45.33	1.26
cold shock protein A cspA	Rv3648c	7 kDa	(0.4400)	4	5.00	1.25
60 kda chaperonin 2 groEL2	Rv0440	57 kDa	(0.0000)	135	168.67	1.25
heat shock protein hspX	Rv2031c	16 kDa	(0.0002)	67.5	75.00	1.11
transmembrane serine/threonine-protein kinase A pknA	Rv0015c	46 kDa	(0.3100)	4	4.33	1.08
isocitrate lyase icl	Rv0467	47 kDa	(0.4300)	2.5	2.67	1.07
integration host factor mihF	Rv1388	21 kDa	(0.0280)	19	20.00	1.05
6-phosphofructokinase pfkA	Rv3010c	37 kDa	(0.3200)	3.5	3.67	1.05
bifunctional polyribonucleotide nucleotidyltransferase gpsI	Rv2783c	80 kDa	(0.1600)	6.5	6.67	1.03
iron-regulated elongation factor tu tuf	Rv0685	44 kDa	(0.0120)	20.5	20.33	0.99
propionyl-CoA carboxylase beta chain 5 accD5	Rv3280	59 kDa	(0.5000)	1.5	1.33	0.89
conserved hypothetical protein	Rv3269	10 kDa	(0.1600)	4.5	4.00	0.89
electron transfer flavoprotein beta subunit fixA	Rv3029c	28 kDa	(0.3300)	2	1.67	0.83
conserved hypothetical protein	Rv1738	11 kDa	(0.0390)	7.5	6.00	0.80
iron-regulated conserved hypothetical protein	Rv1636	15 kDa	(0.0093)	10	7.33	0.73
transmembrane serine/threonine-protein kinase D pknD	Rv0931c	70 kDa	(0.0100)	6	3.33	0.56
conserved hypothetical protein	Rv3127	39 kDa	(0.0640)	3	1.67	0.56

* Average spectral count of desthiobiotin-labeled peptides in the presence of 500μM ATP-γ-S. A value of 1 indicates 0 spectral counts in the presence of excess ATP analog.

** Average spectral count of desthiobiotin-labeled peptides with no excess ATP-γ-S

❖ P-value > 0.05

Table 5.4 – Competitive and Transient Binders of Desthiobiotin ATP in Normal Cultures

Competitive Binders of Desthiobiotin-ATP	Gene #	Mol. Wt.	P-Value	AveNSAF_ATP-γ-S*	AveNSAF_ATP**	Fold Change
acetyl-/propionyl-CoA carboxylase alpha subunit accA1	Rv2501c	71 kDa	(0.0000)	1.00	15.67	15.67
glutamyl-tRNA amidotransferase subunit B gatB	Rv3009c	55 kDa	(0.0000)	2.33	28.33	12.14
inosine-5-monophosphate dehydrogenase guaB3	Rv3410c	39 kDa	(0.0000)	1.00	11.67	11.67
transmembrane serine/threonine-protein kinase H pknH	Rv1266c	67 kDa	(0.0000)	1.00	11.33	11.33
fatty-acid-CoA ligase fadD7	Rv0119	55 kDa	(0.0000)	1.00	9.67	9.67
phosphoribosylamine-glycine ligase purD	Rv0772	44 kDa	(0.0000)	1.00	9.33	9.33
tryptophanyl-tRNA synthetase trpS	Rv3336c	36 kDa	(0.0001)	1.00	7.00	7.00
polyphosphate kinase ppk	Rv2984	83 kDa	(0.0000)	3.33	22.67	6.80
phosphoglycerate kinase pgk	Rv1437	43 kDa	(0.0000)	3.00	19.67	6.56
cysteinyl-tRNA synthetase 1 cysS1	Rv3580c	52 kDa	(0.0005)	1.00	5.67	5.67
fatty-acid-CoA ligase fadD23	Rv3826	63 kDa	(0.0014)	1.00	5.00	5.00
universal stress protein	Rv2319c	32 kDa	(0.0012)	1.33	6.33	4.75
CTP synthase pyrG	Rv1699	64 kDa	(0.0024)	1.00	4.67	4.67
phosphofructokinase pfkB	Rv2029c	35 kDa	(0.0073)	1.00	4.00	4.00
ATP-dependent protease ATP-binding subunit clpC1	Rv3596c	94 kDa	(0.0000)	15.33	61.00	3.98
anchored-membrane serine/threonine-protein kinase pknF	Rv1746	51 kDa	(0.0130)	1.00	3.67	3.67
chaperone protein htpG	Rv2299c	73 kDa	(0.0130)	1.00	3.67	3.67
DNA polymerase I polA	Rv1629	98 kDa	(0.0078)	1.67	6.00	3.60
acetate kinase ackA	Rv0409	41 kDa	(0.0220)	1.00	3.33	3.33
UDP-N-acetylmuramoylalanyl-D-glutamyl-2,6-diaminopimelate-D-alanyl-D-alanyl ligase murF	Rv2157c	52 kDa	(0.0220)	1.00	3.33	3.33
conserved hypothetical protein	Rv3282	23 kDa	(0.0220)	1.00	3.33	3.33
conserved hypothetical protein	Rv2510c	57 kDa	(0.0220)	1.00	3.33	3.33
fatty-acid-CoA ligase fadD28	Rv2941	63 kDa	(0.0077)	2.33	7.67	3.29
nucleoside diphosphate kinase ndkA	Rv2445c	15 kDa	(0.0150)	2.00	6.33	3.17
40 kda secreted L-alanine dehydrogenase ald	Rv2780	39 kDa	(0.0066)	3.00	9.33	3.11
adenosylhomocysteinase sahH	Rv3248c	54 kDa	(0.0017)	5.00	15.00	3.00
pyruvate kinase pykA	Rv1617	51 kDa	(0.0053)	3.67	11.00	3.00
malate synthase G glcB	Rv1837c	80 kDa	(0.0420)	1.67	4.67	2.80
aldehyde dehydrogenase	Rv0458	55 kDa	(0.0350)	2.67	7.00	2.63
adenylate kinase adk	Rv0733	20 kDa	(0.0035)	17.33	36.67	2.12
transmembrane serine/threonine-protein kinase D pknD	Rv0931c	70 kDa	(0.0400)	8.67	17.67	2.04
ATP synthase beta chain atpD	Rv1310	53 kDa	(0.0250)	13.33	26.33	1.98
Transient Binders of Desthiobiotin-ATP						
endopeptidase ATP binding protein chain B clpB	Rv0384c	93 kDa	(0.0410)	25.33	44.33	1.75
3-hydroxyacyl-thioester dehydrogenase htdY	Rv3389c	30 kDa	(0.4000)	4.67	7.00	1.50
iron-regulated short-chain dehydrogenase/reductase	Rv3224	30 kDa	(0.4100)	6.33	9.33	1.47
aminomethyltransferase gcvT	Rv2211c	40 kDa	(0.4300)	4.33	6.33	1.46
phosphopantetheine adenylyltransferase kdtB	Rv2965c	18 kDa	(0.4700)	2.33	3.33	1.43
chaperone protein dnaK	Rv0350	67 kDa	(0.3800)	40.00	57.00	1.43
cold shock protein A cspA	Rv3648c	7 kDa	(0.5300)	5.67	7.67	1.35
secreted fibronectin-binding protein antigen 85-B fbpB	Rv1886c	35 kDa	(0.5400)	2.00	2.67	1.33
iron-regulated peptidyl-prolyl-cis-trans-isomerase A ppiA	Rv0009	19 kDa	(0.5400)	3.00	4.00	1.33
glutamine synthetase glnA2	Rv2222c	50 kDa	(0.5500)	6.00	8.00	1.33
pyruvate dehydrogenase E2 component sucB	Rv2215	57 kDa	(0.5500)	7.33	9.67	1.32
acyl-CoA dehydrogenase fadE5	Rv0244c	66 kDa	(0.5700)	3.33	4.33	1.30
pyruvate dehydrogenase E1 component aceE	Rv2241	100 kDa	(0.5700)	4.67	6.00	1.29
60 kda chaperonin 2 groEL2	Rv0440	57 kDa	(0.0180)	279.00	352.00	1.26
iron-regulated conserved hypothetical protein	Rv1636	15 kDa	(0.3500)	17.00	21.00	1.24
propionyl-CoA carboxylase beta chain 5 accD5	Rv3280	59 kDa	(0.5800)	3.00	3.67	1.22
low molecular weight protein antigen cfp2	Rv2376c	17 kDa	(0.5000)	4.67	5.67	1.21
glutamine synthetase glnA1	Rv2220	54 kDa	(0.3000)	17.67	21.33	1.21
heat shock protein hspX	Rv2031c	16 kDa	(0.1800)	36.00	43.33	1.20
succinyl-CoA synthetase beta chain sucC	Rv0951	41 kDa	(0.3500)	11.67	14.00	1.20
oligoribonuclease orn	Rv2511	23 kDa	(0.4800)	5.00	6.00	1.20
protein transport protein secE2	Rv0379	8 kDa	(0.5400)	3.33	4.00	1.20
L-lactate dehydrogenase lldD2	Rv1872c	45 kDa	(0.4500)	5.33	6.33	1.19
conserved hypothetical protein	Rv2159c	36 kDa	(0.4300)	5.67	6.67	1.18
bifunctional coenzyme A carboxylase alpha chain accA3	Rv3285	64 kDa	(0.2800)	13.67	16.00	1.17
50S ribosomal protein L7/L12 rplL	Rv0652	13 kDa	(0.2500)	16.00	18.67	1.17
conserved hypothetical protein	Rv3269	10 kDa	(0.4800)	4.00	4.67	1.17
ribosome recycling factor frr	Rv2882c	21 kDa	(0.5700)	2.33	2.67	1.14
superoxide dismutase sodA	Rv3846	23 kDa	(0.3900)	5.33	6.00	1.13
transmembrane serine/threonine-protein kinase E pknE	Rv1743	61 kDa	(0.3900)	5.33	6.00	1.13
two component system transcriptional regulator devR	Rv3133c	23 kDa	(0.4600)	3.33	3.67	1.10
tetrahydrodipicolinate N-succinyltransferase	Rv1201c	33 kDa	(0.4600)	3.33	3.67	1.10
leucyl-tRNA synthetase leuS	Rv0041	108 kDa	(0.3800)	4.33	4.67	1.08
iron-regulated elongation factor tu tuf	Rv0685	44 kDa	(0.0300)	40.33	43.33	1.07
immunogenic protein mpt64	Rv1980c	25 kDa	(0.3000)	6.00	6.33	1.06
30S ribosomal protein S1 rpsA	Rv1630	53 kDa	(0.2500)	7.33	7.67	1.05
transmembrane serine/threonine-protein kinase A pknA	Rv0015c	46 kDa	(0.2100)	8.67	9.00	1.04
transcriptional regulator	Rv0023	27 kDa	(0.5600)	1.67	1.67	1.00
hypothetical protein wag31	Rv2145c	28 kDa	(0.4200)	2.67	2.67	1.00
hypothetical protein	Rv3818	58 kDa	(0.5000)	2.00	2.00	1.00
conserved alanine rich protein	Rv2744c	29 kDa	(0.2800)	5.00	5.00	1.00
enoyl-CoA hydratase echA9	Rv1071c	36 kDa	(0.3900)	3.00	3.00	1.00
conserved hypothetical protein	Rv2140c	19 kDa	(0.5000)	2.00	2.00	1.00

thiosulfate sulfurtransferase cysA2	Rv0815c	31 kDa	(0.5600)	1.67	1.67	1.00
zinc-type alcohol dehydrogenase NAD dependent adhB	Rv0761c	40 kDa	(0.2300)	4.67	4.33	0.93
transcriptional regulator moxR1	Rv1479	41 kDa	(0.2800)	3.33	3.00	0.90
60 kda chaperonin 1 groEL1	Rv3417c	56 kDa	(0.0270)	16.33	14.67	0.90
meromycolate extension acyl carrier protein acpM	Rv2244	13 kDa	(0.1400)	6.33	5.67	0.89
haloalkane dehalogenase	Rv2296	33 kDa	(0.1500)	6.00	5.33	0.89
succinyl-CoA synthetase alpha chain sucD	Rv0952	31 kDa	(0.1600)	5.67	5.00	0.88
ATP synthase alpha chain atpA	Rv1308	59 kDa	(0.3300)	2.33	2.00	0.86
integration host factor mihF	Rv1388	21 kDa	(0.0001)	41.33	34.67	0.84
acyl-CoA dehydrogenase fadE25	Rv3274c	42 kDa	(0.3500)	2.00	1.67	0.83
DNA polymerase III beta chain dnaN	Rv0002	42 kDa	(0.3500)	2.00	1.67	0.83
pyridoxamine 5-phosphate oxidase pdxH	Rv2607	25 kDa	(0.1100)	5.67	4.67	0.82
10 kda chaperonin groES	Rv3418c	11 kDa	(0.0000)	50.33	40.67	0.81
iron-regulated aconitate hydratase acn	Rv1475c	102 kDa	(0.1200)	5.00	4.00	0.80
30S ribosomal protein S16 rpsP	Rv2909c	17 kDa	(0.3800)	1.67	1.33	0.80
hypothetical protein cfp17	Rv1827	17 kDa	(0.2000)	3.33	2.67	0.80
fructose-bisphosphate aldolase fba	Rv0363c	37 kDa	(0.2000)	3.00	2.33	0.78
short-chain type dehydrogenase/reductase	Rv0148	30 kDa	(0.2000)	3.00	2.33	0.78
transcriptional regulator, tetR-family	Rv0144	31 kDa	(0.2000)	3.00	2.33	0.78
19 kda lipoprotein antigen precursor lpqH	Rv3763	15 kDa	(0.0250)	8.67	6.33	0.73
conserved hypothetical protein	Rv2406c	15 kDa	(0.2100)	2.33	1.67	0.71
electron transfer flavoprotein beta subunit fixA	Rv3029c	28 kDa	(0.0039)	13.67	9.67	0.71
enolase eno	Rv1023	45 kDa	(0.1300)	3.33	2.33	0.70

* Average spectral count of desthiobiotin-labeled peptides in the presence of 500μM ATP-γ-S. A value of 1 indicates 0 spectral counts in the presence of excess ATP analog.

** Average spectral count of desthiobiotin-labeled peptides with no excess ATP-γ-S

5.4.5 ATP-binding proteins and Associated Biochemical Pathways

In order to find clusters of protein families functionally linked in relevant biochemical pathways we utilized the list of 81 confidently labeled proteins and expanded our dataset to include non-labeled proteins that were confidently identified by mass spectrometry (i.e proteins with total spectral counts across biological replicates >5 with 90% Peptide Probability) irrespective of ATP-labeling (Appendix III). Functional association networks using the web based Search Tool for the Retrieval of Interacting Genes(14) (STRING v 9.0) were generated from the 81 ATP-binding proteins combined with the 54 unlabeled proteins. Emerging from this data set we visualized clusters of associated protein families (Figure 5.8) including members of the polyketide synthase family (Lipid Metabolism), ribosomal protein synthesis (Information Pathways), peptidoglycan and mycolic acid synthesis (Cell Wall and Processes). Select pathways and protein families are discussed in the following sections.

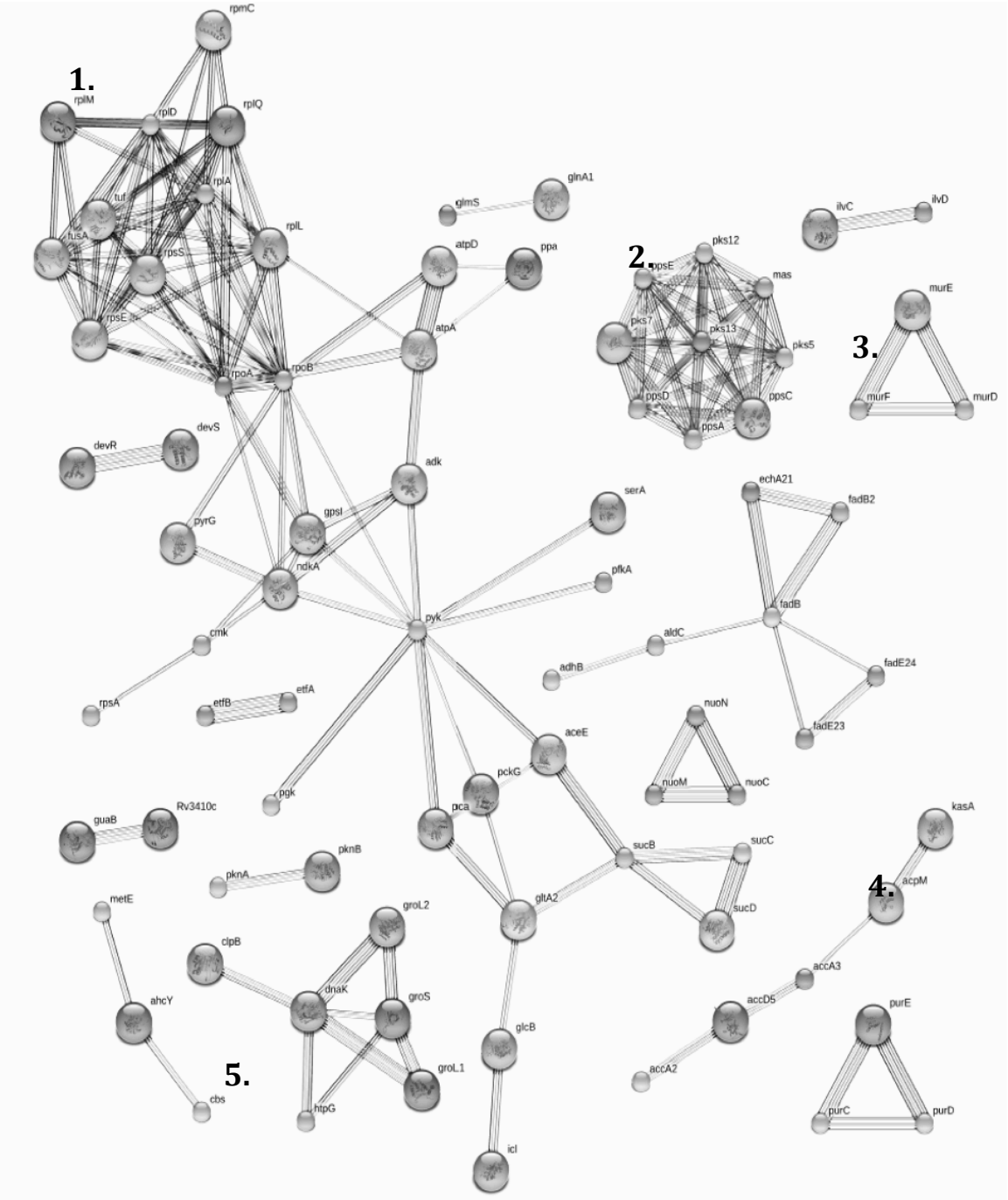

Figure 5.7 - Protein-Protein Interaction networks of the *M. tuberculosis* ATPome. The list of protein IDs from our MS analysis was input into the STRING database (STRING v. 9.0; www.string-db.org) to identify known and predicted functional networks. 48% of the proteins in our shotgun analysis were shown to be functionally associated with at least one known interacting partner. Emerging protein clusters are functionally relevant in basic metabolism (i.e. respiratory chain, protein synthesis), cell wall biosynthesis (i.e. fatty acid, peptidoglycan synthesis), and virulence (i.e. lipid synthesis). Protein groups are numbered as follows: 1. Ribosomal Protein Synthesis, 2. Lipid Biosynthesis, 3. Cell Wall –Peptidoglycan, 4. Mycolic Acid Biosynthesis, 5. Chaperones/Heat-Shock.

5.5 Discussion

5.5.1 *M. tuberculosis* Kinases

M. tuberculosis signal transduction is co-mediated by 3 main families of kinases and 2 families of protein phosphatases. Of the kinases, the first family comprises eleven complete two component systems (TCS) each consisting of a membrane bound histidine kinase and a cytosolic response regulator, which directly interacts with DNA to affect gene transcription. The second family of kinases is comprised of 11 eukaryotic-like ser/thr kinases (STPKs) which phosphorylate various cellular targets (26-28). The third class of kinase is a newly discovered protein tyrosine kinase (PtkA) (28) .

5.5.2 Ser/Thr Kinases

Three of the eleven STPKs, PknA, PknB and PknG, are deemed necessary for successful intracellular infection and growth. PknB has been previously cited as a promising bacterial signaling target, as it is one of the most widely distributed STPKs among bacterial kinases (29) and contains a highly conserved active site (30). We identified 6 STPK gene products, PknABDHEF. PknA and PknB are essential transmembrane proteins responsible for cell growth and morphology(31,32). They are a part of an operon and are predominantly expressed during exponential growth *in vitro*. Protein level analysis of the *M. tuberculosis* proteome within granulomas of infected guinea pigs demonstrated protein levels of PknA decreasing during the course of 90-day infection, while levels of PknB begin to increase (33). While deemed non-essential for growth *in vitro*, PknD, F and H have been implicated in a wide variety of important biological roles. For instance, phosphorylation of MmpL7 by PknD, affects the transport of a key virulence factor in *M. tuberculosis* pathogenesis, phthiocerol dimycocerosate (PDIM) (34). PknF has been shown to be critical for regulating cell growth, septum formation and glucose transport *in vitro* (35). In our ATPome, the Ser/Thr kinases were more abundant during active growth and

showed no trend towards increased abundance in hypoxic cultures. Monitoring these proteins using the active site labeling tools may need further optimization to provide more quantitative evidence of their differential presence during different states of growth.

5.5.3 Two-Component Sensor Kinases and Response Regulators

The sensor histidine kinase PrrB was the only two-component sensor kinase whose spectra exceeded the required protein and peptide probability thresholds set for confident identification in our ATPome. The PrrA-PrrB TCS is required for early intracellular multiplication of *M. tuberculosis* in human macrophages (36). The *prrAB* genes are present in an operon and are co-transcribed during logarithmic growth. During stationary phase and under conditions of hypoxia, transcript levels are decreased however protein levels remain relatively stable throughout growth phases. TCS's are unique to bacteria and are absent in mammals and PrrAB is now known to be essential for mycobacterial survival and pathogenesis (37).

5.5.4 Lipid Biosynthesis and Metabolism

Polyketide synthases are large multi-domain proteins involved in lipid and mycolic acid biosynthesis. Pks5, Pks12 and Pks13 as well as the phthiocerol dimycocerate synthases PpsABCDE and mycocerosic acid synthase (Mas) work in coordination to synthesize the cell wall associated and virulence determinant phthiocerol dimycocerosate (PDIM)(38,39). Within this group of proteins PpsC was found to bind ATP. PpsC catalyzes the complete reduction of malonyl CoA in the synthesis of phthiocerol. The localization of the ATP nucleotide probe was not within any of the annotated domains of PpsC. Although PpsC is a non-essential enzyme, its associated protein partners identified in this study do play essential roles (pks12/13) (39).

Overall, 10 proteins involved in the synthesis of fatty acids and mycolic acids are listed in our ATPome as nucleotide binding: AccA1, AccA3, AcpM, DesA1, DesA2, FadD21, FadD25,

FadD29, FadD32 and FadE23. The *M. tuberculosis* genome contains three biotin dependent essential acyl CoA carboxylases (AccA1 to 3) which are involved in the first step of fatty acid biosynthesis (i.e. the production of malonyl CoA and methyl malonyl CoA from acyl coenzyme A (CoA) (13). The biotin binding domain of these enzymes did allow for the non-specific attachment to the streptavidin capture affinity resin (as discussed in section 5.2.1) however their ATP-binding function is essential to their enzymatic activity, and labeling with nucleotide probe was located within an annotated nucleotide binding domain (K116) (40).

The acyl carrier protein AcpM (Rv2244) is a versatile carrier of fatty acyl chains through the biosynthetic processes of FAS II (41). The essential acyl (ACP) membrane bound desaturases, DesA1 and DesA2 catalyze the introduction of the first double bond in saturated C16 and C18 fatty acids. *DesA1* and *desA2* are essential genes for mycobacterial survival and DesA1 is predicted to be a relevant drug target based on interactome and genome-scale structural analysis (42). A third member of this family DesA3 is a putative target of the thiourea drug isoxyl (43).

5.5.5 Information Pathways

A known target for the antimycobacterial drug rifampin (RpoB) and other peripheral ribosomal proteins were keenly represented in our interaction data set (Figure 5.8). We identified 16 gene products involved in the synthesis of proteins and 14 of them are essential. Protein synthesis is a highly conserved, cyclic process across all domains of life. A variety of soluble protein cofactors interact with the ribosome machinery. For example the essential and ubiquitously conserved GTP-binding proteins Tuf (EF-Tu, Rv0685) and FusA1 (EF-G, Rv0684) bind nucleotide as a mechanism of conformational change to facilitate translation (i.e. GTP hydrolysis by EF-Tu releases aminoacyl-tRNA). It should be noted that the phosphate-binding

loop of GTP-binding proteins resembles that of ATP-binding proteins (conserved P-loop NTPase family of proteins) with a conserved lysine essential for nucleotide binding (44).

5.5.6 Cell Wall and Cell Wall Processes

The processes involved in the formation of peptidoglycan (PG) are very well characterized in most bacteria. The mode of action for the beta-lactam class of antibiotics targets the formation of PG by irreversibly inhibiting the cellular activity of the penicillin binding proteins. The Mur ligases MurCDEF, located within the cytoplasm of mycobacteria assemble the peptide stem of PG. Specifically within our ATPome, we identified the D-glutamic acid ligase (MurD), the meso-diaminopimelic acid ligase (MurE) and the dipeptide D-Ala-D-Ala ligase (MurF). Dihyrodipicolinate reductase (DapB) belongs to the **L**-lysine biosynthetic pathway and in mycobacteria is also used in the synthesis of meso-diaminopimeleic acid. DapB catalyzes the reduction of dihydrodipicolinate to yield tetrahydrodipicolinate (45) and has been classified as a slow growth mutant (46).

5.5.7 Intermediary metabolism and respiration

Several conserved proteins involved in glycolysis (Pgk, PykA and PfkA), glyoxylate bypass (Icl, GlcB), TCA cycle (SucB,C and D), electron transport (FixA and B) and purine ribonucleotide biosynthesis (PurC and D) were identified in our study. *pykA* encoding pyruvate kinase, catalyzes the transphosphorylation of phosphoenol pyruvate (PEP) and ADP to pyruvate and ATP, and is essential for *M. tuberculosis* growth. The knockout of the *pykA* gene in *M. tuberculosis* resulted in the utilization of fatty acids for energy production (47,48)

5.6 Material and Methods

5.6.1 Bacterial Growth:

Mycobacterium tuberculosis $H_{37}Rv$ seed culture was grown to log phase ($O.D_{600}$ 1.2) in Middlebrook 7H9, ADC. Normally growing cultures: Three cultures (195 mL in 500 mL vented cap flasks) were inoculated with 2.5 ml of seed culture. A magnetic stir bar was added and cultures were incubated at 37°C with stirring (200 RPM). Final $O.D_{600}$ 1.2-1.6. Dormant cultures: Six cultures (paired-replicates of 390 mL in 500 mL sealed cap flasks) were inoculated with 5.0 mL seed culture. Cultures were incubated at 37°C with stirring (100 RPM). Final $O.D_{600}$ 0.65-0.70. Cell pellets were harvested on days 7 (Normal) & 14 (Dormant).

5.6.2 Sample Preparation:

Cell pellets were resuspended at a concentration of 0.5 g/mL in IP/Lysis Buffer (25mM Tris-HCl pH 7.4, 150mM NaCl, 1mM EDTA, 1% NP-40 and 5% glycerol) containing HALT™ Protease/Phosphatase Inhibitor Cocktail (ThermoPierce). Resuspended cells were placed in Lysing Matrix B bead beater vials (pre-filled with 0.1 mm silica; MP Biomedicals). Lysis of cells occurred over 12 bead-beat cycles (30 sec lyse/45 sec rest on ice). Cell lysates were cleared of silica and cellular debris via centrifugation at 4,000 x g for 10 minutes. Supernatant was transferred to new microcentrifuge tube and centrifuged again for 10 minutes at 13,000 x g. Cleared lysate was then filtered through a 0.8mm/0.2mm syringe filter to sterilize lysate for working under BSL-2 conditions. The sterile lysate was desalted using 7K Thermo Scientific Zeba™ Spin Desalting columns and protein amount quantified by BCA (ThermoPierce). 500 µg aliquots of whole cell lysate were labeled with 5 mM Desthiobiotin-ATP for 10 minutes as per manufacturer's instructions (ThermoPierce).

5.6.3 Active Site Peptide Capture:

Desthiobiotin-ATP labeled proteins were reduced in 1mM DTT and alkylated in 1mM iodoacetamide before buffer exchange into digestion buffer (20 mM Tris pH 8.0, 2M urea). Each sample was digested with trypsin (1μg/μl) at an enzyme to substrate ratio of 1:50 for 2 hr at 37^{o}C. Peptide capture with streptavidin agarose resin and elution using 50% acetonitrile/0.1% TFA was followed as per manufacturer instructions.

LC-MS: Peptides were separated on a nanospray column (Zorbax C18, 5 μm x 75 μm, 150 μm column). Samples were eluted into a LTQ linear ion trap mass spectrometer (Thermo) using a gradient of 0-100% B (A= 3% ACN, 0.1% formic acid; B =100% ACN, 0.1% formic acid) at a flow rate of 300 nL/min for 103 minutes. All samples were run in triplicate.

5.6.4 Database Searching:

Tandem mass spectra were extracted, charge state deconvoluted and deisotoped by Xcalibur version 2.2 SP1. All MS/MS samples were analyzed using Mascot (Matrix Science, London, UK; version 2.3.02) and Sequest (Thermo Fisher Scientific, San Jose, CA, USA; version v.27, rev. 11). Mascot and Sequest were set up to search the MtbReverse041712 database (7992 entries) assuming the digestion enzyme Trypsin. Parameters for both search engines were set to a fragment ion mass tolerance of 1.5 Da and a parent ion tolerance of 0.5 Da. Oxidation of methionine, iodoacetamide derivative of cysteine and the desthiobiotin modification of lysine were specified in Mascot and Sequest as variable modifications.

5.6.5 Criteria For Protein Identification:

Scaffold (version Scaffold_3.6.1, Proteome Software Inc., Portland, OR) was used to validate MS/MS based peptide and protein identifications. Peptide identifications were accepted if they exceeded specific database search engine thresholds. Mascot identifications required at

least ion scores to be greater than both the associated identity scores and 50, 65, 65 and 65 for singly, doubly, triply and quadruply charged peptides. Sequest identifications required at least deltaCn scores of greater than 0.2 and XCorr scores of greater than 1.8, 2.0, 3.0 and 4.0 for singly, doubly, triply and quadruply charged peptides. Protein identifications were accepted if they contained at least 1 identified peptide. Peptide spectra meeting the most minimum requirement were manually inspected for quality. Proteins that contained similar peptides and could not be differentiated based on MS/MS analysis alone were grouped to satisfy the principles of parsimony.

5.6.6 Statistical Analysis:

The design for each experimental condition consisted of three biological replicates per sample group (Normal/Hypoxic-ATP, Normal/Hypoxic-ATP-γ-S and Normal/Hypoxic-Streptavidin only). In the case of hypoxic cultures, 6 biological replicates were grown to set time points and cell material was pooled into 3-paired replicates and subsequently treated as triplicate replications. Each replicate was injected into the mass spectrometer 3 times for a total of 9 injections per sample. Spectral count data, as visualized in Scaffold (Proteome Software, v. 3.6.1) was normalized to Quantitative Values (http://www.proteomesoftware.com/pdf_files/Normalizing_Spectral_Counting-Scaffold.pdf) (49) and analysis was performed using Fisher's Exact Test (FET) in the comparison of two groups (i.e Normal-ATP vs. Hypoxic-ATP or Normal-ATP vs Normal-ATP–γ-S). Fisher's Exact Test is a valid method of identifying differences in protein abundance (i.e spectral counts) in shotgun proteomics data sets using experimental designs of at least 3 biological replicates and performs with similar power to more complex generalized linear modeling strategies (50).

5.6.7 Blast-Description:

The most relevant description for each of the sequences was acquired based on the significant BLAST results. The homologs for the sequences were retrieved using the blastp algorithm and the nonredundant database of NCBI. The Blast2GO suite (51) was used for this purpose, since it can annotate several sequences in one session. Pfam-domain based description: The domains present in the sequences, along with their Pfam and InterPro descriptions, were identified using the InterProscan Web service, which was accessed via the Pipeline Pilot (Accelrys) implementation in the sequence analysis collection. Mapping the domain and the labeled peptide sequences retained the information for the relevant domains.

5.6.8 Gene Ontology Annotation:

The Pfam domains were mapped to Gene Ontology (GO) terms using the lookup table provided by Pfam2go (http://www.geneontology.org/external2go/pfam2go). GO terms are hierarchical and inter-related in nature. All the GO terms originate from three distinct subsumption hierarchy trees, namely -cellular component, biological process and molecular function. Thus each domain can have multiple GO terms based on the level and type of annotation. An in-house script was written to retrieve GO annotations based on the root term as "molecular function" and their distance from the root term.

5.6.9 Immunoblots:

5 μg of normal and hypoxic lysate were separated on 4-12% Bis-Tris SDS-PAGE gel (Life Technologies (Invitrogen), Carlsbad, CA). Primary antibodies were either mouse monoclonal (HspX, Ald, Hbha, GlcB, KatG) or rabbit polyclonal (Rv0569, Rv1738, Rv2626c, Rv2032, Rv3133c) and diluted to suggested titers. All antibodies were provided through BEI Resources (www.beiresources.org). Control samples were recombinant proteins generated from E. coli or, if

unavailable, whole cell lysate from *M. tuberculosis,* $H_{37}Rv$ was used (Hbha, KatG, Ald). Protein bands were visualized via alkaline phosphatase conjugated IgG (Sigma). Densitometry analysis was performed via the ImageJ suite (http://rsbweb.nih.gov/ij/).

5.7 Summary

The results obtained in this study represent a first look into the ATP-binding proteome of the pathogenic organism, *M. tuberculosis*. The majority of these proteins are essential gene products and may be relevant therapeutic targets. We quantitatively measured the differences in protein levels between normally growing and hypoxic state bacteria and provided preliminary data into the binding characteristics and utilization of ATP across multiple classes of functional enzymes. Using desthiobiotin-conjugate nucleotide probes in competition with ATP-analogs provides the framework necessary to pursue antimicrobial inhibitors whose mode of action relies on competition within the ATP-binding site of select protein targets (4,52). The utilization of ATP in lipid and cell wall biosynthesis pathways makes it tempting to speculate that a broad-spectrum nucleotide competitive compound may affect these critical processes in such as way as to alter cell wall architecture and heterogeneity. Because the cell wall of *M. tuberculosis* is the primary defense mechanism against effective chemotherapies, breakdown of this structure using a novel class of antimicrobial could help circumvent the problems of multi-drug and extensively drug resistant bacteria (53). In our profile of ATP-binding proteins of normally growing and hypoxic state *M. tuberculosis*, we identified several proteins known to be under the control of the dormancy regulon *dosR*. The DevR-DevS TCS is implicated in virulence and mediates the expression of ~48 dormancy associated genes when *M. tuberculosis* adapts to hypoxia and is exposed to other stress factors like nitric oxide, carbon monoxide and ascorbic acid (59). The ~48 genes which comprise the *devR* dormancy regulon has some well known genes like *hspX*

(the alpha crystallin like chaperone), the nitrate reductase *acg, and* several uncharacterized hypothetical proteins such as Rv0569, Rv1738 and Rv2626c. A derivative of phenylcoumarin had reduced the survival of hypoxically adapted *M. tuberculosis* and also inhibited DevR binding to target DNA (60). We would expect to have identified the sensor kinase DevS, however it did not meet criteria for final inclusion, most likely due to its subcellular location with the plasma membrane. The response regulator DevR/DosR (Rv3133c) was shown to be confidently labeled with the ATP probe at the C-terminal DNA binding domain. In our competition experiments with 100-fold excess ATP-γS, the desthiobiotin nucleotide probe was still able to bind and label two lysines (K179 and K182) within the helix-turn-helix (HTH) DNA binding domain of DevR (UniProt ID >sp|P951931|167-186). The physiological implications of nucleotide binding to DevR remain to be elucidated. Furthermore, we identified several proteins belonging to the universal stress protein family (Rv2005c, Rv2028c, Rv2319c, Rv2623, Rv2624, and Rv2626c). This class of nucleotide binding proteins has been implicated in the events required for the activation of dormancy (4).

5.8 Conclusion

Genomic scale surveys of essential genes of *M. tuberculosis* and transcriptome-wide analyses of the bacterial response to environmental or metabolic conditions mimicking the host environment have been carried out (19,61,62). Additionally, large-scale proteomic profiling under simulatory *in vitro* conditions or *in vivo* have also been performed (33,63-65). Together these studies have shaped our understanding of possible roles of individual genes as well as functional protein groups in the context of various stages of pathogenesis. With the enormous amount of available data, various studies have attempted to gain insightful information using sophisticated models of computational analyses (66-68). However, understanding models of gene

regulation, protein-protein interactions and unique metabolic pathways at the systems level remains incomplete, especially for functional changes that mediate the switch into dormancy (69). Through the study of the *M. tuberculosis* ATPome, we have defined a functionally linked analysis among gene products of the mycobacterial proteome. Future work on the absolute quantification and monitoring of these proteins during transitions into and out of dormancy can further define, at the proteome level, a link between members of the DosR regulon and the central regulatory circuit that controls the transition into the states of non-replicating persistence and resuscitation into normal growth or active disease.

5.9 LITERATURE CITED

1. WHO. (2011) WHO Report 2011: Global Tuberculosis Control. World Health Organization, Geneva
2. Ananthan, S., Faaleolea, E. R., Goldman, R. C., Hobrath, J. V., Kwong, C. D., Laughon, B. E., Maddry, J. A., Mehta, A., Rasmussen, L., Reynolds, R. C., Secrist, J. A., 3rd, Shindo, N., Showe, D. N., Sosa, M. I., Suling, W. J., and White, E. L. (2009) *Tuberculosis (Edinb)* **89**, 334-353
3. Magnet, S., Hartkoorn, R. C., Szekely, R., Pato, J., Triccas, J. A., Schneider, P., Szantai-Kis, C., Orfi, L., Chambon, M., Banfi, D., Bueno, M., Turcatti, G., Keri, G., and Cole, S. T. (2010) *Tuberculosis (Edinb)* **90**, 354-360
4. Reynolds, R. C., Ananthan, S., Faaleolea, E., Hobrath, J. V., Kwong, C. D., Maddox, C., Rasmussen, L., Sosa, M. I., Thammasuvimol, E., White, E. L., Zhang, W., and Secrist, J. A., 3rd. (2011) *Tuberculosis (Edinb)* **92**, 72-83
5. Triola, G., Wetzel, S., Ellinger, B., Koch, M. A., Hubel, K., Rauh, D., and Waldmann, H. (2009) *Bioorg Med Chem* **17**, 1079-1087
6. Kumari, S., and Ram, V. J. (2004) *Drugs Today (Barc)* **40**, 487-500
7. Patricelli, M. P., Szardenings, A. K., Liyanage, M., Nomanbhoy, T. K., Wu, M., Weissig, H., Aban, A., Chun, D., Tanner, S., and Kozarich, J. W. (2007) *Biochemistry* **46**, 350-358
8. Bantscheff, M., Eberhard, D., Abraham, Y., Bastuck, S., Boesche, M., Hobson, S., Mathieson, T., Perrin, J., Raida, M., Rau, C., Reader, V., Sweetman, G., Bauer, A., Bouwmeester, T., Hopf, C., Kruse, U., Neubauer, G., Ramsden, N., Rick, J., Kuster, B., and Drewes, G. (2007) *Nat Biotechnol* **25**, 1035-1044
9. Li, J., Rix, U., Fang, B., Bai, Y., Edwards, A., Colinge, J., Bennett, K. L., Gao, J., Song, L., Eschrich, S., Superti-Furga, G., Koomen, J., and Haura, E. B. (2010) *Nat Chem Biol* **6**, 291-299
10. Bantscheff, M., Hopf, C., Kruse, U., and Drewes, G. (2007) *Ernst Schering Found Symp Proc*, 1-28
11. Keller, A., Nesvizhskii, A. I., Kolker, E., and Aebersold, R. (2002) *Anal Chem* **74**, 5383-5392
12. Fischer, G., and Schmid, F. X. (1990) *Biochemistry* **29**, 2205-2212
13. Norman, E., De Smet, K. A., Stoker, N. G., Ratledge, C., Wheeler, P. R., and Dale, J. W. (1994) *J Bacteriol* **176**, 2525-2531
14. Szklarczyk, D., Franceschini, A., Kuhn, M., Simonovic, M., Roth, A., Minguez, P., Doerks, T., Stark, M., Muller, J., Bork, P., Jensen, L. J., and von Mering, C. (2011) *Nucleic Acids Res* **39**, D561-568
15. Griffin, J. E., Gawronski, J. D., Dejesus, M. A., Ioerger, T. R., Akerley, B. J., and Sassetti, C. M. (2011) *PLoS Pathog* **7**, e1002251
16. Sassetti, C. M., Boyd, D. H., and Rubin, E. J. (2003) *Mol Microbiol* **48**, 77-84
17. Punta, M., Coggill, P. C., Eberhardt, R. Y., Mistry, J., Tate, J., Boursnell, C., Pang, N., Forslund, K., Ceric, G., Clements, J., Heger, A., Holm, L., Sonnhammer, E. L., Eddy, S. R., Bateman, A., and Finn, R. D. *Nucleic Acids Res* **40**, D290-301
18. Fischer, J. J., Graebner Baessler, O. Y., Dalhoff, C., Michaelis, S., Schrey, A. K., Ungewiss, J., Andrich, K., Jeske, D., Kroll, F., Glinski, M., Sefkow, M., Dreger, M., and Koester, H. (2009) *J Proteome Res* **9**, 806-817

19. Sherman, D. R., Voskuil, M., Schnappinger, D., Liao, R., Harrell, M. I., and Schoolnik, G. K. (2001) *Proc Natl Acad Sci U S A* **98**, 7534-7539
20. Wayne, L. G., and Lin, K. Y. (1982) *Infect Immun* **37**, 1042-1049
21. Munoz-Elias, E. J., and McKinney, J. D. (2005) *Nat Med* **11**, 638-644
22. Giffin, M. M., Modesti, L., Raab, R. W., Wayne, L. G., and Sohaskey, C. D. (2012) *J Bacteriol* **194**, 1045-1054
23. Usha, V., Jayaraman, R., Toro, J. C., Hoffner, S. E., and Das, K. S. (2002) *Can J Microbiol* **48**, 7-13
24. Sriram, D., Yogeeswari, P., Vyas, D. R., Senthilkumar, P., Bhat, P., and Srividya, M. (2010) *Bioorg Med Chem Lett* **20**, 4313-4316
25. Smith, D. M., Fraga, H., Reis, C., Kafri, G., and Goldberg, A. L. (2011) *Cell* **144**, 526-538
26. Av-Gay, Y., and Deretic, V. (2004) Two-Component Systems, Protein Kinases and Signal Transduction in Mycobacterium tuberculosis. in *Tuberculosis and the Tubercle Bacillus* (Cole, S. T. ed.), ASM Press, Washington, D.C. pp
27. Greenstein, A. E., Grundner, C., Echols, N., Gay, L. M., Lombana, T. N., Miecskowski, C. A., Pullen, K. E., Sung, P. Y., and Alber, T. (2005) *J Mol Microbiol Biotechnol* **9**, 167-181
28. Chao, J., Wong, D., Zheng, X., Poirier, V., Bach, H., Hmama, Z., and Av-Gay, Y. (2010b) *Biochim Biophys Acta* **1804**, 620-627
29. Wehenkel, A., Bellinzoni, M., Grana, M., Duran, R., Villarino, A., Fernandez, P., Andre-Leroux, G., England, P., Takiff, H., Cervenansky, C., Cole, S. T., and Alzari, P. M. (2008) *Biochim Biophys Acta* **1784**, 193-202
30. Mieczkowski, C., Iavarone, A. T., and Alber, T. (2008) *EMBO J* **27**, 3186-3197
31. Parikh, A., Verma, S. K., Khan, S., Prakash, B., and Nandicoori, V. K. (2009) *J Mol Biol* **386**, 451-464
32. Chao, J. D., Papavinasasundaram, K. G., Zheng, X., Chavez-Steenbock, A., Wang, X., Lee, G. Q., and Av-Gay, Y. (2010a) *J Biol Chem* **285**, 29239-29246
33. Kruh, N. A., Troudt, J., Izzo, A., Prenni, J., and Dobos, K. M. (2010) *PLoS One* **5**, e13938
34. Perez, J., Garcia, R., Bach, H., de Waard, J. H., Jacobs, W. R., Jr., Av-Gay, Y., Bubis, J., and Takiff, H. E. (2006) *Biochem Biophys Res Commun* **348**, 6-12
35. Deol, P., Vohra, R., Saini, A. K., Singh, A., Chandra, H., Chopra, P., Das, T. K., Tyagi, A. K., and Singh, Y. (2005) *J Bacteriol* **187**, 3415-3420
36. Ewann, F., Jackson, M., Pethe, K., Cooper, A., Mielcarek, N., Ensergueix, D., Gicquel, B., Locht, C., and Supply, P. (2002) *Infect Immun* **70**, 2256-2263
37. Haydel, S. E., Malhotra, V., Cornelison, G. L., and Clark-Curtiss, J. E. (2012) *J Bacteriol* **194**, 354-361
38. Pedelacq, J. D., Nguyen, H. B., Cabantous, S., Mark, B. L., Listwan, P., Bell, C., Friedland, N., Lockard, M., Faille, A., Mourey, L., Terwilliger, T. C., and Waldo, G. S. (2011) *Nucleic Acids Res* **39**, e125
39. Gavalda, S., Leger, M., van der Rest, B., Stella, A., Bardou, F., Montrozier, H., Chalut, C., Burlet-Schiltz, O., Marrakchi, H., Daffe, M., and Quemard, A. (2009) *J Biol Chem* **284**, 19255-19264
40. Gago, G., Kurth, D., Diacovich, L., Tsai, S. C., and Gramajo, H. (2006) *J Bacteriol* **188**, 477-486

41. Kremer, L., Nampoothiri, K. M., Lesjean, S., Dover, L. G., Graham, S., Betts, J., Brennan, P. J., Minnikin, D. E., Locht, C., and Besra, G. S. (2001) *J Biol Chem* **276**, 27967-27974
42. Raman, K., and Chandra, N. (2008) *BMC Microbiol* **8**, 234
43. Phetsuksiri, B., Jackson, M., Scherman, H., McNeil, M., Besra, G. S., Baulard, A. R., Slayden, R. A., DeBarber, A. E., Barry, C. E., 3rd, Baird, M. S., Crick, D. C., and Brennan, P. J. (2003) *J Biol Chem* **278**, 53123-53130
44. Fischer, J. (2007) Initial Characterization of the Ribosome-Associated ATP Binding Cassette (ABC) Protein of YHIH from E. coli. in *School of Graduate Studies*, University of Lethbridge, Alberta
45. Pavelka, M. S., Jr., Weisbrod, T. R., and Jacobs, W. R., Jr. (1997) *J Bacteriol* **179**, 2777-2782
46. Janowski, R., Kefala, G., and Weiss, M. S. (2009) *Acta Crystallogr D Biol Crystallogr* **66**, 61-72
47. Keating, L. A., Wheeler, P. R., Mansoor, H., Inwald, J. K., Dale, J., Hewinson, R. G., and Gordon, S. V. (2005) *Mol Microbiol* **56**, 163-174
48. Chavadi, S., Wooff, E., Coldham, N. G., Sritharan, M., Hewinson, R. G., Gordon, S. V., and Wheeler, P. R. (2009) *J Bacteriol* **191**, 7545-7553
49. Eng, J. K., Searle, B. C., Clauser, K. R., and Tabb, D. L. (2011) *Mol Cell Proteomics* **10**, R111 009522
50. Li, M., Gray, W., Zhang, H., Chung, C. H., Billheimer, D., Yarbrough, W. G., Liebler, D. C., Shyr, Y., and Slebos, R. J. (2012) *J Proteome Res* **9**, 4295-4305
51. Conesa, A., Gotz, S., Garcia-Gomez, J. M., Terol, J., Talon, M., and Robles, M. (2005) *Bioinformatics* **21**, 3674-3676
52. Lougheed, K. E., Osborne, S. A., Saxty, B., Whalley, D., Chapman, T., Bouloc, N., Chugh, J., Nott, T. J., Patel, D., Spivey, V. L., Kettleborough, C. A., Bryans, J. S., Taylor, D. L., Smerdon, S. J., and Buxton, R. S. (2011) *Tuberculosis (Edinb)* **91**, 277-286
53. Dalton, T., Cegielski, P., Akksilp, S., Asencios, L., Caoili, J. C., Cho, S. N., Erokhin, V. V., Ershova, J., Gler, M. T., Kazennyy, B. Y., Kim, H. J., Kliiman, K., Kurbatova, E., Kvasnovsky, C., Leimane, V., van der Walt, M., Via, L. E., Volchenkov, G. V., Yagui, M. A., and Kang, H. (2012) *Lancet*
54. Doerks, T., van Noort, V., Minguez, P., and Bork, P. (2011) *PLoS One* **7**, e34302
55. Mazandu, G. K., and Mulder, N. J. (2012) *Int J Mol Sci* **13**, 7283-7302
56. De Palma, S., Ripamonti, M., Vigano, A., Moriggi, M., Capitanio, D., Samaja, M., Milano, G., Cerretelli, P., Wait, R., and Gelfi, C. (2007) *J Proteome Res* **6**, 1974-1984
57. Li, X., Arslan, F., Ren, Y., Adav, S. S., Poh, K. K., Sorokin, V., Lee, C. N., de Kleijn, D., Lim, S. K., and Sze, S. K. (2012) *J Proteome Res* **11**, 2331-2346
58. Drumm, J. E., Mi, K., Bilder, P., Sun, M., Lim, J., Bielefeldt-Ohmann, H., Basaraba, R., So, M., Zhu, G., Tufariello, J. M., Izzo, A. A., Orme, I. M., Almo, S. C., Leyh, T. S., and Chan, J. (2009) *PLoS Pathog* **5**, e1000460
59. Gautam, U. S., Sikri, K., and Tyagi, J. S. (2011) *J Bacteriol* **193**, 4849-4858
60. Gupta, M., Sajid, A., Arora, G., Tandon, V., and Singh, Y. (2009) *J Biol Chem* **284**, 34723-34734
61. Muttucumaru, D. G., Roberts, G., Hinds, J., Stabler, R. A., and Parish, T. (2004) *Tuberculosis (Edinb)* **84**, 239-246

62. Voskuil, M. I., Visconti, K. C., and Schoolnik, G. K. (2004) *Tuberculosis (Edinb)* **84**, 218-227
63. Rosenkrands, I., Slayden, R. A., Crawford, J., Aagaard, C., Barry, C. E., 3rd, and Andersen, P. (2002) *J Bacteriol* **184**, 3485-3491
64. Schmidt, F., Donahoe, S., Hagens, K., Mattow, J., Schaible, U. E., Kaufmann, S. H., Aebersold, R., and Jungblut, P. R. (2004) *Mol Cell Proteomics* **3**, 24-42
65. Mawuenyega, K. G., Forst, C. V., Dobos, K. M., Belisle, J. T., Chen, J., Bradbury, E. M., Bradbury, A. R., and Chen, X. (2005) *Mol Biol Cell* **16**, 396-404
66. Strong, M., Graeber, T. G., Beeby, M., Pellegrini, M., Thompson, M. J., Yeates, T. O., and Eisenberg, D. (2003) *Nucleic Acids Res* **31**, 7099-7109
67. Balazsi, G., Heath, A. P., Shi, L., and Gennaro, M. L. (2008) *Mol Syst Biol* **4**, 225
68. Cui, T., Zhang, L., Wang, X., and He, Z. G. (2009) *BMC Genomics* **10**, 118
69. Hegde, S. R., Rajasingh, H., Das, C., Mande, S. S., and Mande, S. C. (2012) *PLoS One* **7**, e33893
70. Zhang, Y. (2004) *Frontiers in bioscience : a journal and virtual library* **9**, 1136-1156
71. Gomez, J. E., and McKinney, J. D. (2004) *Tuberculosis (Edinb)* **84**, 29-44
72. Lew, J. M., Kapopoulou, A., Jones, L. M., and Cole, S. T. (2011) *Tuberculosis (Edinb)* **91**, 1-7

VI. CHAPTER VI: Concluding Remarks and Future Directions

Robert Koch first identified *M. tuberculosis* as the causative agent of human tuberculosis in 1882 [1]. During that time it was estimated that TB was responsible for one-quarter of all deaths among developed nations and historical accounts estimate that "70-90% of the urban populations of Europe and North America were infected with the TB bacillus, and about 80% of those individuals who developed active tuberculosis died of it" (http://ocp.hul.harvard.edu/contagion/tuberculosis.html). One-hundred thirty years later the disease is mostly under control in first-world countries however vigilance towards the eradication of disease worldwide must be maintained if we are to see the reversal of incident cases in our children's lifetime. Having evolved for thousands of years within animal species and man, *M. tuberculosis* has exquisite control over the immune mechanisms of its hosts. Despite great successes in the implementation of global control programs such as directly observed therapy, short course (DOTS) and epidemiological surveillance committees, attempts to lessen the control of the bacillus through the use of vaccine and multi-drug chemotherapies have not curtailed the successfulness of the pathogen. In the case of MDR-TB, human control over the disease is even lessened.

Critical to the advancement of eradication and control programs will be the identification of disease-state biomarkers that will allow for the procurement of second-generation diagnostics. These simple, rapid and inexpensive methods will be designed to better detect true infection (versus exposure or past vaccination). Also, information must be provided about the drug-susceptibility profile within the pathogen and the effectiveness of such treatments within the host. In addition, optimal targets for new anti-mycobacterial drugs must be identified in disease-relevant states of infection. In this regard, drug, vaccine and molecular diagnostic targets can be

most efficiently identified using global biomarker-based discovery platforms [2]. Gene expression profiles and elucidation of the transcriptome under various biological states continues to be a very informative tool when digging into the systems-level biology of an organism. Studies of the proteome focus on the characterization of functional gene products and take into account post-translational events such as processing and cleavage, modification, secretion and subcellular localization. Advances in mass spectrometry-based proteomics have driven the successful application of these platforms in all areas of disease research.

For *Mycobacterium tuberculosis*, thorough descriptions of both its gene expression and protein profiles have been utilized in the greater understanding of the patho-physiology of the organism. The cellular envelope of *M. tuberculosis*, recognized for decades as a key modulator of host-pathogen interactions is chemically well-defined [3] and is a major reservoir of both protein and non-protein based antigens [4]. Key to its survival within the host is its ability to exist in a state of metabolic dormancy, induced to some extent, by the hypoxic environment encountered within a granuloma. Understanding this switch into and out of dormancy at the protein level will give us unique bio-signatures of functionally linked protein networks that can be further developed and studied quantitatively to assess their contribution to this state of metabolic growth. Further, proteins within these functional networks may be relevant drug targets for new drugs with novel modes of action.

6.1 Relevance of the Cell Wall Proteome

Physiologically, the cell envelope of *M. tuberculosis* is highly hydrophobic and contains many proteins not amenable to traditional techniques of protein purification. Elucidation of the cell wall proteome was undertaken based from the observation that cell wall proteins were responsible for TB-specific immunoreactivity in non-classically restricted T cells (Chapter IV). Also a full characterization using both gel-based and non-gel based separation techniques had also not been previously published for *M. tuberculosis*. Our results demonstrated complimentary approaches to elucidating proteins non-covalently associated with the cell wall core. Extraction methods with SDS or guanidine hydrochloride for the most part, resolved soluble proteins and many antigens, known to be secreted into the extracellular milieu. The identification of these antigenic proteins with the cell wall lent support to the observation that cell wall proteins could be retained within the core to elicit an immune response. Extraction using Triton-X 100 detergents generated a hydrophobic protein fraction enriched with many lipoproteins. The majority of these protein species could not be adequately resolved using gel-based methods, and a multi-dimensional liquid chromatography strategy was used to identify proteins within this fraction. From a biochemical standpoint the protein profiles generated in this study could be further refined to address multiple biologically relevant questions broadly pertaining to cell wall remodeling events. These include elucidation of lipid and glycolipid export systems and the changes in protein profiles in response to drug treatment – perhaps increased abundance of small molecule export systems or lipoproteins which may function to shuttle molecules across the hydrophobic barriers.

Immunologically, the proteins of the cell wall could be more specifically defined as immune modulators of the host response to infection. Studies performed on *in vivo* grown bacilli demonstrated an almost complete absence of known secreted protein antigens. It was

hypothesized that these proteins were immediately detected and shuttled to the lymph node for presentation to antigen-specific T cells [5]. In the *in vivo* model, we can hypothesize that the remaining resident proteins of the cell wall must stay the course and mediate localized responses necessary to maintain infection as well as to maintain cell wall integrity.

During the original collaboration in the identification of cell wall antigen-specific CD8 T cells, we began the process of elucidating whole protein fractions of cell wall proteins. In essence creating a cell wall protein library in which extracted proteins were purified using multi-dimensional separation schemes. A diagram of this is depicted below (Figure 6.1). The ultimate goal would then be to generate *M. tuberculosis*-specific T cell clones reactive to specific proteins within distinct cell wall fractions. From this information the minimal epitopes could be gleaned (if not post-translationally modified) and this information could be provided to repositories such as the Immune Epitope Database (www.iedb.org). Post-translationally modified (PTM) protein antigens would be of special interest as demonstration of PTM-specific T cell reactivity to protein antigens has thus far only been peripherally demonstrated for a few modified proteins [6].

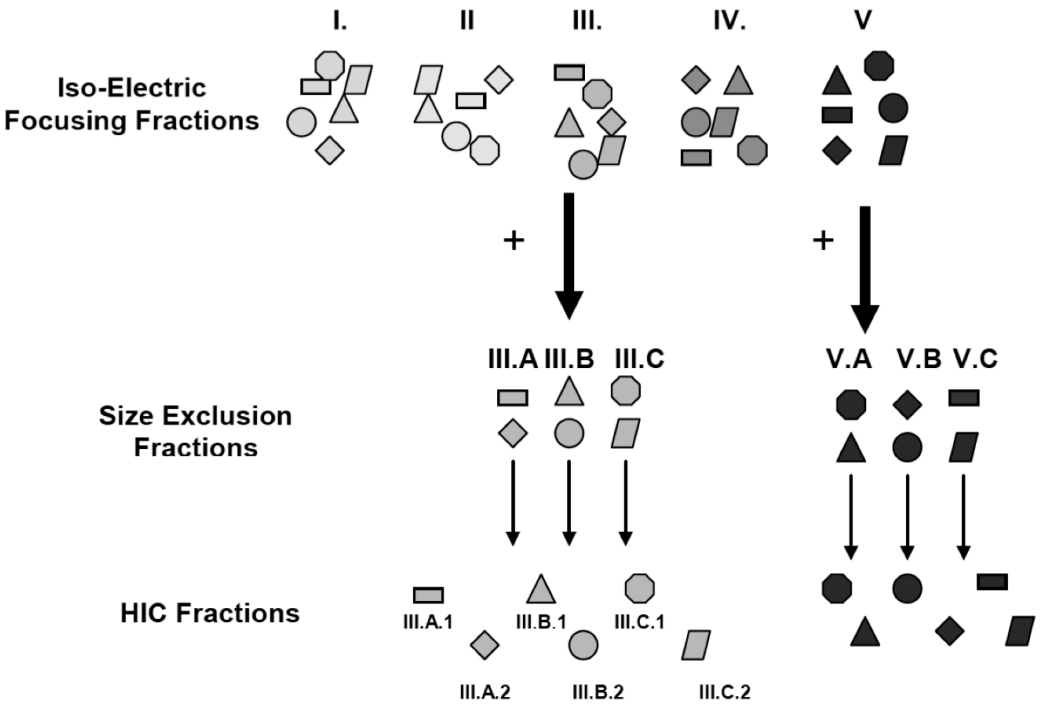

Figure 6.1 – Generation of a Cell Wall Protein Library. In this schematic, *M. tuberculosis* proteins are separated via iso-electric point (Fractions I – V); gray scale indicates similar pH. Biologically active fractions are then further separated by size-exclusion chromatography, thereby creating a non-gel 2D separation scheme. From here, proteins can be further separated by additional HPLC chemistries (i.e. hydrophobic interaction chromatography). All protein fractions can be characterized via LC-MS/MS and MALDI-TOF mass spectrometry and fractions used in the generation of M. tuberculosis-specific CD8 T cell clones.

6.2 Non-Classically Restricted T cell Antigens

The primary objective in the proteomic approach to antigen discovery was the elucidation of the HLA-E restricted *M. tuberculosis* antigen. We identified the glycoprotein Mpt32 as being the cognate antigen and have evidence to suggest the glycosyl modifications at the N-terminus of the protein are responsible for T cell reactivity. Future work to definitively describe this glyco-specific epitope will involve the generation of site-directed mutants unable to be glycosylated at known sites of modification. Furthermore, because the generation of synthetic modified peptides is cost-prohibitive, we plan to adapt an *in vitro* enzymatic assay in which enzymatically active

membrane preparations will glycosylate pools of overlapping synthetic peptides, spanning the N-terminus of the protein. A brief description of this proposal is described below.

6.2.1 Mpt32 Glycosylation Mutants and *In vitro* Glycosylation Assay

Mpt32 (Apa, 45/47kDa) was the first glycoprotein to be characterized in mycobacteria [7] and alteration of its mannosylation state is known to down-regulate both delayed-type hypersensitivity (DTH) reactions *in vivo*, and T cell proliferation *in vitro* [8,9]. Since these original studies, important cellular functions of several mycobacterial glycoproteins and glycolipoproteins have been predicted or described [10-12]. In the immunopathogenesis of cancer, glycosylation and specific glycosylation patterns affect biologic activity of antigen-specific T cells [13]. Previous studies at CSU studied the role of glycosylation using Mpt32 specific murine T cell hybridomas and observed that glycosylation was required to retain T cell reactivity. However the results remained incomplete and the precise epitope responsible for reactivity was not fully confirmed.

Our goal for Mpt32 in the CD8 T cell antigen study, will be to confidently identify sites of glycosylation that are required for the generation of a productive T cell response for the HLA-E restricted human T cell clone D160 1-23. This would be the first description of an *M. tuberculosis* antigen presented in the context of HLA-E and would for the first time describe a glycosylation-dependent T cell response for CD8 T cells isolated from human patients. We have designed and generated glyco-mutant constructs with altered states of glycosylation, expressed and purified from an Mpt32 knockout strain of *M. tuberculosis.* Characterization of the sites of altered modification will be determined via mass spectrometry using a Thermo Scientific Orbi-TRAP Velos, equipped with an ETD source. This instrument will provide the exquisite sensitivity and precise mass accuracy we will need to characterize glycopeptides present within

active fractions generated from native or recombinant protein preparations. We would expect to see an alteration of activity if our hypotheses that specific glycosyl moieties are required for T cell activation are correct. Briefly, recombinant constructs will be purified from cell lysate or culture filtrate and then be introduced to D160 1-23 to assess the differential ability of the glycan moieties to elicit biological activity, via γ-IFN ELISPOT. We expect that removal of specific threonine residues, and thus specific mannosylation patterns, will perturb the T cell response. These series of experiments will establish the role of adjacent glycosyl moieties in antigen recognition of an HLA-E restricted T cell antigen.

Complementary to these genetic studies, overlapping synthetic peptides of the N-terminus of Mpt32 have been designed and will carry a cleavable affinity FLAG tag (Sigma-Aldrich) at each C-terminus. Tagged peptides will be used in a modified *in vitro* glycosylation assay in which membrane preparations from native bacilli will be used to generate O-mannosylated glycopeptides [14]. The affinity tags will be cleaved off and newly modified peptides tested for T cell reactivity. Modification will be confirmed by MALDI-TOF MS and peptides will be assayed for activity by ELISPOT. As control, Mpt32 specific peptides will also be generated from preparations of *M. smegmatis* membrane. The protein mannosylation function in M. smegmatis results in a hyper-glycosylated phenotype in comparison to M. tuberculosis – derived proteins. Recombinant Mpt32 purified from M. smegmatis does not activate HLA-E specific T cells. Future work may then further address questions of antigen processing and presentation of glycoprotein antigens something that has not been assessed for bacterial glycoproteins [15]. It is well documented that processing of eukaryotic glycoproteins through the endocytic MHC II pathway does not alter or remove the carbohydrate moieties from the surrogate peptide and that the retention of the covalently linked glycan activates glycopeptide

specific T cell populations [16,17]. For eubacterial glycoproteins, the processing pathways are less defined and it is unknown if the glycans are essential for TCR recognition, or initial processing and binding to MHC. Further it is unknown whether the addition of a glycan to a peptide will ablate a productive T cell response. Given the promiscuity of glycosylation in *M. tuberculosis*, both questions need to be addressed in order to understand the cell mediated immune response during disease. The antigen processing mechanisms required for presentation within the context of HLA-E have been characterized [18,19]. Using these studies as framework, we can look at processing of native Mpt32 and presentation of its glycosylated peptides. Perhaps these modified synthetic peptide products, which are specifically presented in the context of the non-polymorphic HLA-E can be used to test for their capacity to detect TB disease in real-world clinical settings.

6.3 Assessing ATP-inhibitors against cellular targets in *M. tuberculosis*

In this section I propose a research plan for the identification of cellular targets of ATP-inhibitor antimicrobial compounds, using quantitative chemical proteomics and the techniques utilized in the definition of the *Mycobacterial* ATPome (Chapter V).

As was discussed in Chapter V, we identified broad classes of ATP-binding proteins in *M. tuberculosis*, and measured their differential abundance in both normal, active growth and during hypoxia. In this work, we selectively enriched for and purified a unique subproteome using active site directed nucleotide probes in both actively growing and hypoxic state bacteria. The ATP-probes bind within conserved active sites and a desthiobiotin compound covalently attaches to adjacent lysines. This tagging facilitates purification and identification of binding sites via LC-MS/MS. Using label-free quantitative proteomics we identified 122 ATP-binding proteins in either metabolic state, including several proteins under control of the dormancy regulon (DosR)

as well as five of ten, universal stress proteins. Primary sequence analysis of the ATP-binding proteins identified in this study revealed ATP-binding associated within a functional domain in 80% of our proteome. Further, biosynthetic enzymes for the synthesis of the virulence lipid phthiocerol dimycoserate (PDIM) and mycocerosic acid synthase (Pps-ACDE, Mas) were also enriched using the nucleotide affinity approach. Because we identified proteins that bind small molecule ligands (ATP), we expect several of these proteins to be relevant therapeutic targets for a novel class of ATP-inhibitor antibiotics in the treatment of tuberculosis.

6.3.1 Universal Stress Proteins and Dormancy

The *M. tuberculosis* dormancy regulon encodes 48 genes that are up regulated during hypoxia, during increased nitric oxide and carbon monoxide conditions. Encompassed in this regulon are 6 universal stress proteins, which are also over-expressed during multiple environmental stressors. Several of these proteins have ATP-binding domains and at least for one USP, Rv2623, disrupting the ATP-binding capacity of this protein prevented latency in experimentally infected guinea pigs. It has been proposed by Wayne and others that if M. tuberculosis is prevented from entering into its non-replicative dormant state, that it may be more susceptible to treatment with currently available antibiotics. I hypothesize that by inhibiting the ATP-binding function of Rv2623 and the other ATP-binding domain USPs using a competitive inhibitor, *M. tuberculosis* will remain in an actively growing state *in vivo*. The USPs are a conserved protein family observed in archaea, eubacteria, plants and fungi. They are absent in higher eukaryotes such as humans, thus making them a viable anti-bacterial target. There is ample evidence of the upregulation of the USP proteins by gene expression microarray analysis and molecular manipulation of a select number of USP proteins has demonstrated their redundant functions *in vitro*. No proteomic analysis has looked at this protein family under

varying *in vitro* or *in vivo* conditions. Targeted mass spectrometry analysis of members of this protein family during normal growth and during hypoxia is needed in order to better characterize these proteins and their function in regulating the transition into and out of dormancy.

6.3.2 *M. tuberculosis* kinases as drug targets

Signal transduction in *Mycobacterium tuberculosis* is regulated by the activities of two main families of kinases. The first class comprises members of classical two-component systems (TCSs) and includes eleven membrane-bound sensor histidine kinases and their respective response regulators[20]. Phosphotransfer events mediated by the TCSs in bacteria are thought to be "short-term/transient" signals that are initiated in response to stimuli of the external environment. Each of these systems are known to play integral roles in mycobacterial physiology, virulence and growth. Tyagi, Av-Gay and others have done extensive review of TCSs and their effects on cellular physiology[21-24]. Two-component systems as novel drug targets have had limited success, which may be attributable their unconventional ATP-binding active site [25,26], however several compounds have been characterized [27]. The second major class of kinases encoded within the mycobacterial genome, are the eukaryotic-like Ser/Thr kinases (STPKs). There are 11 members of this group and three are deemed essential for growth *in vitro* [28]. Not surprisingly, studies of the STPKs in terms of bacterial physiology have revealed their vital roles in regulation of cell division and cell wall formation [24]. Additionally, roles for STPKs in the context of successful intracellular infections have also been established [29]. This makes the identification and validation of these and other ATP-binding proteins as novel drug targets a valuable priority [25,30].

6.3.3 ATP-competitive Inhibitors

Kinase inhibitor compounds have been utilized for many years primarily in the treatment of cancer [31]. The majority of these compounds work in competition with ATP-binding. Bacterial

kinases and other ATP-binding proteins have the potential to be relevant therapeutic targets due to their roles in survival and pathogenesis. Thus, the evaluation of inhibitors for bacterial enzymes that utilize ATP is worthwhile [27]. Several groups have turned their focus on targeting signaling networks of *M. tuberculosis* - either through direct inhibition of specific kinases such as the STPK PknB or through phenotypic screening of chemically relevant (kinase/ATP-inhibitor) libraries in which specific protein targets are unknown [32-35]. Nearly a decade ago, Fernandez et al used *Mycobacterium smegmatis* to over express *M. tuberculosis* PknB to demonstrate that an increase in PknB levels was able to increase resistance to their K252a compound, however they were only able to see a modest 2-fold increase in resistance compared to controls [36]. More recently, Lougheed et al used a target-based approach in which a small library of compounds was chosen for optimization against PknB. They used biochemical assays as a measure of drug effectiveness. However they were unable to correlate *in vitro* potencies with activity against whole cells. They were also unable to identify the *in vivo* target(s) of their inhibitors as PknB over expression in *M. tuberculosis* leads to undesirable phenotypic affects (i.e. slow growth, altered morphology) [37] which makes assessment of compound specificity difficult to interpret by traditional drug discovery efforts.

Chemical proteomics techniques for the large-scale, unbiased identification of target proteins can identify cellular targets that are either directly or indirectly associated with sets of anti-mycobacterial compounds. These proteins and signaling cascades are currently not targeted by available drug regimens in the treatment of tuberculosis.

Current chemotherapeutic regimens for the treatment of pulmonary tuberculosis involve multiple antibiotics. Isoniazid and Rifampin are the standard-of-care first line drugs. These two are then coupled with Ethambutol and Pyrazinamide. The recommended treatment duration for patients

afflicted with active disease is no shorter than 6 months [38]. The increase in prevalence of multidrug-resistant (MDR) and extensively drug-resistant (XDR) TB is multifactorial, however TB drug discovery efforts need to identify compounds that can circumvent current mechanisms of drug resistance. As mentioned above, kinase inhibitors have had substantial success in the fields of cancer and inflammation. Kinase inhibitors to treat bacterial diseases are of interest despite the reduced repertoire of kinases encoded by bacterial genomes. Our methodologies for target deconvolution have the potential to bridge the transition between pre-clinical research of novel anti-TB active compounds and pre-clinical development (toxicology, formulation, binding profiles, pharmacokinetics) of such drugs using innovative, non-biased chemical proteomic techniques.

The use of mass spectrometry and chemical proteomics in the drug discovery pipeline is in its infancy, but in the last decade has shown repeated success and proof-of-relevance concepts for many drug formulations. To date these techniques have not been applied in the study of inhibitors for infectious disease [27,39,40]. Several high-throughput compound screenings based on ATP-inhibitor scaffolds have identified compounds with acceptable activity against *M. tuberculosis*. [32,41-43]. The cellular targets of these compounds remain unknown (with the exception of the D-ala-D-ala ligase). Proteomics-based strategies for inhibitors of mycobacterial kinases and other ATP-binding proteins can provide comprehensive characterization of several promising anti-TB compounds. Through these methods we can identify specific protein-drug interactions as well as characterize, protein profile changes in response to drug treatment. For a bacterium such as *M. tuberculosis*, where the development of novel antimicrobials is urgently needed, a focused and innovative approach in drug discovery efforts will have a great impact on how anti-TB drugs can be assessed beyond the pre-clinical stage.

6.4 LITERATURE CITED

1. Koch, R. An Address on Bacteriological Research. *Br Med J* **2**, 380-3 (1890).
2. Maertzdorf J Fau - Weiner Iii, J., Weiner Iii J Fau - Kaufmann, S.H.E. & Kaufmann, S.H. Enabling biomarkers for tuberculosis control. (2012).
3. Kaur, D., Guerin, M.E., Skovierova, H., Brennan, P.J. & Jackson, M. Chapter 2: Biogenesis of the cell wall and other glycoconjugates of Mycobacterium tuberculosis. *Adv Appl Microbiol* **69**, 23-78 (2009).
4. Dhiman, N., Verma, I. & Khuller, G.K. Cellular immune responses to cell wall peptidoglycan associated protein antigens in tuberculosis patients and healthy subjects. *Microbiol Immunol* **41**, 495-502 (1997).
5. Kruh, N.A., Troudt, J., Izzo, A., Prenni, J. & Dobos, K.M. Portrait of a pathogen: the Mycobacterium tuberculosis proteome in vivo. *PLoS One* **5**, e13938 (2010).
6. Sutcliffe, I.C. & Harrington, D.J. Lipoproteins of Mycobacterium tuberculosis: an abundant and functionally diverse class of cell envelope components. *FEMS Microbiol Rev* **28**, 645-59 (2004).
7. Dobos, K.M., Khoo, K.H., Swiderek, K.M., Brennan, P.J. & Belisle, J.T. Definition of the full extent of glycosylation of the 45-kilodalton glycoprotein of Mycobacterium tuberculosis. *J Bacteriol* **178**, 2498-506 (1996).
8. Horn, C. et al. Decreased capacity of recombinant 45/47-kDa molecules (Apa) of Mycobacterium tuberculosis to stimulate T lymphocyte responses related to changes in their mannosylation pattern. *J Biol Chem* **274**, 32023-30 (1999).
9. Romain, F. et al. Deglycosylation of the 45/47-kilodalton antigen complex of Mycobacterium tuberculosis decreases its capacity to elicit in vivo or in vitro cellular immune responses. *Infect Immun* **67**, 5567-72 (1999).
10. Drage, M.G. et al. Mycobacterium tuberculosis lipoprotein LprG (Rv1411c) binds triacylated glycolipid agonists of Toll-like receptor 2. *Nat Struct Mol Biol* **17**, 1088-95 (2010).
11. Gehring, A.J. et al. The Mycobacterium tuberculosis 19-kilodalton lipoprotein inhibits gamma interferon-regulated HLA-DR and Fc gamma R1 on human macrophages through Toll-like receptor 2. *Infect Immun* **71**, 4487-97 (2003).
12. Sieling, P.A. et al. Conserved mycobacterial lipoglycoproteins activate TLR2 but also require glycosylation for MHC class II-restricted T cell activation. *J Immunol* **180**, 5833-42 (2008).
13. Hanisch, F.G. & Ninkovic, T. Immunology of O-glycosylated proteins: approaches to the design of a MUC1 glycopeptide-based tumor vaccine. *Curr Protein Pept Sci* **7**, 307-15 (2006).
14. Cooper, H.N. et al. Characterization of mycobacterial protein glycosyltransferase activity using synthetic peptide acceptors in a cell-free assay. *Glycobiology* **12**, 427-34 (2002).
15. Hanisch, F.G., Schwientek, T., Von Bergwelt-Baildon, M.S., Schultze, J.L. & Finn, O. O-Linked glycans control glycoprotein processing by antigen-presenting cells: a biochemical approach to the molecular aspects of MUC1 processing by dendritic cells. *Eur J Immunol* **33**, 3242-54 (2003).
16. Carbone, F.R. & Gleeson, P.A. Carbohydrates and antigen recognition by T cells. *Glycobiology* **7**, 725-30 (1997).

17. Harding, C.V., Kihlberg, J., Elofsson, M., Magnusson, G. & Unanue, E.R. Glycopeptides bind MHC molecules and elicit specific T cell responses. *J Immunol* **151**, 2419-25 (1993).
18. Grotzke, J.E. et al. The Mycobacterium tuberculosis phagosome is a HLA-I processing competent organelle. *PLoS Pathog* **5**, e1000374 (2009).
19. Heinzel, A.S. et al. HLA-E-dependent presentation of M. tuberculosis-derived antigen to human CD8+ T cells. *J Exp Med* **196**, 1473-81 (2002).
20. Chao, J. et al. Protein kinase and phosphatase signaling in Mycobacterium tuberculosis physiology and pathogenesis. *Biochim Biophys Acta* **1804**, 620-7 (2010).
21. Av-Gay, Y. & Deretic, V. Two-Component Systems, Protein Kinases and Signal Transduction in Mycobacterium tuberculosis. in *Tuberculosis and the Tubercle Bacillus*, Vol. 2 (ed. Cole, S.T.) (ASM Press, Washington, D.C., 2004).
22. Bagchi, G., Chauhan, S., Sharma, D. & Tyagi, J.S. Transcription and autoregulation of the Rv3134c-devR-devS operon of Mycobacterium tuberculosis. *Microbiology* **151**, 4045-53 (2005).
23. Chao, J.D. et al. Convergence of Ser/Thr and two-component signaling to coordinate expression of the dormancy regulon in Mycobacterium tuberculosis. *J Biol Chem* **285**, 29239-46 (2010).
24. Tyagi, A.K. The Role of Mycobacterial Kinases and Phosphatases in Growth, Pathogenesis and Cell Wall Metabolism. in *The Mycobacterial Cell Envelope* (ed. Reyrat, M.D.a.J.M.) (ASM Press, Washington, DC, 2008).
25. Av-Gay, Y. & Alber, T. Prospects for TB Therapeutics Targeting Mycobacterium tuberculosis Phosphosignaling Networks. in *Protein Kinases as Drug Targets* 349-364 (Wiley-VCH Verlag GmbH & Co. KGaA, 2011).
26. Gotoh, Y. et al. Two-component signal transduction as potential drug targets in pathogenic bacteria. *Curr Opin Microbiol* **13**, 232-9.
27. Kurosu, M. Bacterial Protein Kinase Inhibitors. *Drug Development Research* **71**, 168-187 (2010).
28. Sassetti, C.M., Boyd, D.H. & Rubin, E.J. Genes required for mycobacterial growth defined by high density mutagenesis. *Mol Microbiol* **48**, 77-84 (2003).
29. Walburger, A. et al. Protein kinase G from pathogenic mycobacteria promotes survival within macrophages. *Science* **304**, 1800-4 (2004).
30. Wehenkel, A. et al. Mycobacterial Ser/Thr protein kinases and phosphatases: physiological roles and therapeutic potential. *Biochim Biophys Acta* **1784**, 193-202 (2008).
31. Zhang, J., Yang, P.L. & Gray, N.S. Targeting cancer with small molecule kinase inhibitors. *Nat Rev Cancer* **9**, 28-39 (2009).
32. Hegymegi-Barakonyi, B. et al. Signalling inhibitors against Mycobacterium tuberculosis--early days of a new therapeutic concept in tuberculosis. *Curr Med Chem* **15**, 2760-70 (2008).
33. Lougheed, K.E. et al. Effective inhibitors of the essential kinase PknB and their potential as anti-mycobacterial agents. *Tuberculosis (Edinb)* **91**, 277-86 (2011).
34. Ananthan, S. et al. High-throughput screening for inhibitors of Mycobacterium tuberculosis H37Rv. *Tuberculosis (Edinb)* **89**, 334-53 (2009).

35. Wehenkel, A. et al. The structure of PknB in complex with mitoxantrone, an ATP-competitive inhibitor, suggests a mode of protein kinase regulation in mycobacteria. *FEBS Lett* **580**, 3018-22 (2006).
36. Fernandez, P. et al. The Ser/Thr protein kinase PknB is essential for sustaining mycobacterial growth. *J Bacteriol* **188**, 7778-84 (2006).
37. Kang, C.M. et al. The Mycobacterium tuberculosis serine/threonine kinases PknA and PknB: substrate identification and regulation of cell shape. *Genes Dev* **19**, 1692-704 (2005).
38. Kumari, S. & Ram, V.J. Advances in molecular targets and chemotherapy of tuberculosis. *Drugs Today (Barc)* **40**, 487-500 (2004).
39. Drewes, G. Chemical proteomics in drug discovery. *Methods Mol Biol* **803**, 15-21.
40. Mehaffy, M.C., Kruh-Garcia, N.A. & Dobos, K.M. Prospective on Mycobacterium tuberculosis proteomics. *J Proteome Res* **11**, 17-25.
41. Triola, G. et al. ATP competitive inhibitors of D-alanine-D-alanine ligase based on protein kinase inhibitor scaffolds. *Bioorg Med Chem* **17**, 1079-87 (2009).
42. Magnet, S. et al. Leads for antitubercular compounds from kinase inhibitor library screens. *Tuberculosis (Edinb)* **90**, 354-60 (2010).
43. Reynolds, R.C. et al. High throughput screening of a library based on kinase inhibitor scaffolds against Mycobacterium tuberculosis H37Rv. *Tuberculosis (Edinb)* **92**, 72-83 (2011).

APPENDIX I

Appendix I. Mass Spectrometry Data Report for Cell Wall Proteome

Biological sample name	Protein name	Protein accession numbers	MW (Da)	Protein ID probability	Number of unique peptides	Number of unique spectra	Number of total spectra	Percentage of total spectra	Percentage sequence coverage	Peptide sequence	Best Peptide identification probability	Best SEQUEST XCorr score	Best SEQUEST ΔCn score	Number of ID +1H spectra	Number of ID +2H spectra	Number of ID +3H spectra	Number of ID +4H spectra	Number of enzymatic termini	Calculated +1H Peptide Mass (AMU)	Peptide start index	Peptide stop index
GuHCl	hypo protein Rv2476c	Rv2476c	176885	99.70%	2	2	2	0.01%	1.72%	FSVAAMNADVLEIPTISRR	95.00%	1.99	0.173	0	0	1	0	2	2,091.48	316	334
GuHCl	hypo protein Rv2476c	Rv2476c	176885	99.70%	2	2	2	0.01%	1.72%	YEIFVYSPR	95.00%	1.14	0.0725	1	0	0	0	2	1,174.39	788	796
GuHCl	hypo protein Rv1777	Rv1777	47169	99.40%	2	2	2	0.01%	9.45%	LWVDGMQDGCIEWMGAM ANRLPMMVVAEL	95.00%	1.83	0.255	0	0	1	0	2	3,300.84	158	186
GuHCl	hypo protein Rv1777	Rv1777	47169	99.40%	2	2	2	0.01%	9.45%	VMEQFTVQAADRLWVDGM QDGCIEWMGAMANR	89.00%	2.15	0.209	0	0	1	0	2	3,764.08	146	177
GuHCl	hypo protein Rv3592	Rv3592	11166	99.70%	2	2	3	0.01%	16.20%	INAIEVPAGAGPELEK	95.00%	2.09	0.423	0	2	0	0	2	1,608.87	6	21
GuHCl	hypo protein Rv3592	Rv3592	11166	99.70%	2	2	3	0.01%	16.20%	INAIEVPAGAGPELEKR	95.00%	3.35	0.446	0	1	0	0	2	1,765.07	6	22
GuHCl	PE_PGRS	Rv1450c	107389	100.00%	3	3	3	0.01%	6.55%	GGAGGNGGAGGNGGQGV VSDGAGGAGGAGGDGGA PGDGANGGNGQGAGAFAG	83.40%	1.22	0.215	0	0	1	0	2	4,299.06	939	995
GuHCl	PE_PGRS	Rv1450c	107389	100.00%	3	3	3	0.01%	6.55%	GGDGAFGGMSANATNPGE NGPNGN	95.00%	1.73	0.322	0	1	0	0	2	2,164.18	859	882
GuHCl	PE_PGRS	Rv1450c	107389	100.00%	3	3	3	0.01%	6.55%	GGGEFV	95.00%	0.952	0.352	1	0	0	0	2	565.5384	1324	1329
GuHCl	groEL2 (groEL2)	Rv0440	56709	100.00%	28	35	466	1.39%	52.20%	AAVEEGIVAGGGVTLLQAA PTLDELK	95.00%	5.37	0.574	0	19	0	0	2	2,523.85	403	428
GuHCl	groEL2 (groEL2)	Rv0440	56709	100.00%	28	35	466	1.39%	52.20%	AGAATEVELK	95.00%	1.69	0.433	2	1	0	0	2	989.0784	379	388
GuHCl	groEL2 (groEL2)	Rv0440	56709	100.00%	28	35	466	1.39%	52.20%	ASVPGGGDMGGMDF	84.20%	1.06	0.108	1	0	0	0	2	1,330.42	527	540
GuHCl	groEL2 (groEL2)	Rv0440	56709	100.00%	28	35	466	1.39%	52.20%	DETTIVEGAGDTDAIAGR	95.00%	5.57	0.373	1	8	0	0	2	1,791.88	326	343
GuHCl	groEL2 (groEL2)	Rv0440	56709	100.00%	28	35	466	1.39%	52.20%	DLLPLLEK	95.00%	2.2	0.423	24	0	0	0	2	941.3384	230	237
GuHCl	groEL2 (groEL2)	Rv0440	56709	100.00%	28	35	466	1.39%	52.20%	EIELEDPYEK	95.00%	3.28	0.399	7	0	0	0	2	1,265.44	58	67
GuHCl	groEL2 (groEL2)	Rv0440	56709	100.00%	28	35	466	1.39%	52.20%	EQIAATAAISAGDQSIGDLI AEAMDK	95.00%	5.78	0.718	0	117	0	0	2	2,607.04	141	166
GuHCl	groEL2 (groEL2)	Rv0440	56709	100.00%	28	35	466	1.39%	52.20%	FDKGYISGYFVTDPER	95.00%	3.89	0.611	0	1	0	0	2	1,895.19	193	208
GuHCl	groEL2 (groEL2)	Rv0440	56709	100.00%	28	35	466	1.39%	52.20%	GLNALADAVK	95.00%	1.98	0.434	13	0	0	0	2	972.1784	18	27
GuHCl	groEL2 (groEL2)	Rv0440	56709	100.00%	28	35	466	1.39%	52.20%	GYISGYFVTDPER	95.00%	4.09	0.561	0	1	0	0	2	1,504.69	196	208
GuHCl	groEL2 (groEL2)	Rv0440	56709	100.00%	28	35	466	1.39%	52.20%	GYISGYFVTDPERQEAVLED PYILLVSSK	95.00%	4.52	0.686	0	3	2	0	2	3,290.79	196	224
GuHCl	groEL2 (groEL2)	Rv0440	56709	100.00%	28	35	466	1.39%	52.20%	IGAELVK	95.00%	1.76	0.133	2	0	0	0	2	729.9184	68	74
GuHCl	groEL2 (groEL2)	Rv0440	56709	100.00%	28	35	466	1.39%	52.20%	KTDDVAGDGTTTATVLAQA LVR	95.00%	6.11	0.65	0	87	0	0	2	2,204.35	79	100
GuHCl	groEL2 (groEL2)	Rv0440	56709	100.00%	28	35	466	1.39%	52.20%	KVVVTK	95.00%	2.03	0.271	1	0	0	0	2	673.7284	320	325
GuHCl	groEL2 (groEL2)	Rv0440	56709	100.00%	28	35	466	1.39%	52.20%	KVVVTKDETTIVEGAGDTD AIAGR	95.00%	6.33	0.671	0	71	4	0	2	2,446.59	320	343
GuHCl	groEL2 (groEL2)	Rv0440	56709	100.00%	28	35	466	1.39%	52.20%	KWGAPTITNDGVSIAK	95.00%	4.78	0.58	0	2	0	0	2	1,658.95	42	57
GuHCl	groEL2 (groEL2)	Rv0440	56709	100.00%	28	35	466	1.39%	52.20%	LAGGVAVIK	95.00%	2.03	0.0469	1	0	0	0	2	828.0184	370	378
GuHCl	groEL2 (groEL2)	Rv0440	56709	100.00%	28	35	466	1.39%	52.20%	LEGDEATGANIVK	95.00%	1.87	0.29	1	0	0	0	2	1,317.45	429	441
GuHCl	groEL2 (groEL2)	Rv0440	56709	100.00%	28	35	466	1.39%	52.20%	NVAAGANPLGLK	95.00%	3.66	0.458	13	2	0	0	2	1,125.35	105	116
GuHCl	groEL2 (groEL2)	Rv0440	56709	100.00%	28	35	466	1.39%	52.20%	NVVLEK	95.00%	2.28	0.215	1	0	0	0	2	701.7584	36	41
GuHCl	groEL2 (groEL2)	Rv0440	56709	100.00%	28	35	466	1.39%	52.20%	QEAVLEDPYILLVSSK	95.00%	3.99	0.573	0	6	0	0	2	1,805.12	209	224
GuHCl	groEL2 (groEL2)	Rv0440	56709	100.00%	28	35	466	1.39%	52.20%	QIAFNSGLEPGVVAEK	95.00%	4.43	0.607	6	16	0	0	2	1,659.82	450	465
GuHCl	groEL2 (groEL2)	Rv0440	56709	100.00%	28	35	466	1.39%	52.20%	TDDVAGDGTTTATVLAQAL VR	95.00%	5.16	0.517	0	21	0	0	2	2,076.15	80	100
GuHCl	groEL2 (groEL2)	Rv0440	56709	100.00%	28	35	466	1.39%	52.20%	VALEAPLK	95.00%	2.2	0.431	9	0	0	0	2	841.0684	442	449
GuHCl	groEL2 (groEL2)	Rv0440	56709	100.00%	28	35	466	1.39%	52.20%	VTETLLK	95.00%	1.25	0.262	3	0	0	0	2	803.9884	126	132
GuHCl	groEL2 (groEL2)	Rv0440	56709	100.00%	28	35	466	1.39%	52.20%	VTLGPK	95.00%	1.57	0.17	1	0	0	0	2	614.7584	28	33
GuHCl	groEL2 (groEL2)	Rv0440	56709	100.00%	28	35	466	1.39%	52.20%	VVVTKDETTIVEGAGDTDAI AGR	95.00%	5.26	0.672	0	13	0	0	2	2,318.39	321	343
GuHCl	groEL2 (groEL2)	Rv0440	56709	100.00%	28	35	466	1.39%	52.20%	WGAPTITNDGVSIAK	95.00%	4.09	0.553	1	5	0	0	2	1,530.75	43	57

GuHCl	hypo protein Rv2296	Rv2296	33340	100.00%	9	12	20	0.06%	35.00%	AGYDAPFPDK	95.00%	1.88	0.256	1	0	0	0	2	1,081.27	194	203
GuHCl	hypo protein Rv2296	Rv2296	33340	100.00%	9	12	20	0.06%	35.00%	AGYDAPFPDKTYQAGAR	95.00%	2.91	0.559	0	1	0	0	2	1,829.08	194	210
GuHCl	hypo protein Rv2296	Rv2296	33340	100.00%	9	12	20	0.06%	35.00%	ASHFIQEDSGTELAER	95.00%	3.79	0.559	0	3	0	0	2	1,790.89	277	292
GuHCl	hypo protein Rv2296	Rv2296	33340	100.00%	9	12	20	0.06%	35.00%	DPILGQADGPLIK	95.00%	3.61	0.43	1	1	0	0	2	1,337.74	250	262
GuHCl	hypo protein Rv2296	Rv2296	33340	100.00%	9	12	20	0.06%	35.00%	FEHLVGYPFAPHYVDVTAG DTQPLR	95.00%	4.72	0.718	0	2	2	0	2	2,831.05	11	35
GuHCl	hypo protein Rv2296	Rv2296	33340	100.00%	9	12	20	0.06%	35.00%	IEDYTYLR	95.00%	1.59	0.25	1	0	0	0	2	1,073.32	93	100
GuHCl	hypo protein Rv2296	Rv2296	33340	100.00%	9	12	20	0.06%	35.00%	SAAGHRVLAPDLIGFGR	95.00%	2.05	0.268	0	0	1	0	2	1,738.08	70	86
GuHCl	hypo protein Rv2296	Rv2296	33340	100.00%	9	12	20	0.06%	35.00%	VLAPDLIGFGR	95.00%	2.43	0.432	3	3	0	0	2	1,158.49	76	86
GuHCl	hypo protein Rv2296	Rv2296	33340	100.00%	9	12	20	0.06%	35.00%	YSPVLPAGR	95.00%	1.6	0.208	1	0	0	0	2	960.1384	170	178
GuHCl	hspX (hspX)	Rv2031c	16210	100.00%	16	21	524	1.56%	76.40%	AELPGVDPDKDVDIMVR	95.00%	3.43	0.605	12	95	1	0	2	1,870.10	55	71
GuHCl	hspX (hspX)	Rv2031c	16210	100.00%	16	21	524	1.56%	76.40%	DFDGR	95.00%	1.17	0.0982	1	0	0	0	2	609.6684	86	90
GuHCl	hspX (hspX)	Rv2031c	16210	100.00%	16	21	524	1.56%	76.40%	DGQLTIK	95.00%	1.39	0.417	8	0	0	0	2	774.9684	72	78
GuHCl	hspX (hspX)	Rv2031c	16210	100.00%	16	21	524	1.56%	76.40%	FAAFPSFAGLRPTFDTR	95.00%	2.17	0.389	0	6	0	0	2	1,902.33	22	38
GuHCl	hspX (hspX)	Rv2031c	16210	100.00%	16	21	524	1.56%	76.40%	FPEFSELFAAFPSFAGLRPTF DTR	95.00%	3.6	0.472	0	0	1	0	2	2,752.33	15	38
GuHCl	hspX (hspX)	Rv2031c	16210	100.00%	16	21	524	1.56%	76.40%	GILTVSVAVSEGKPTEK	95.00%	4.03	0.687	50	183	0	0	2	1,715.89	120	136
GuHCl	hspX (hspX)	Rv2031c	16210	100.00%	16	21	524	1.56%	76.40%	HIQIR	95.00%	1.62	0.141	4	0	0	0	2	666.8184	137	141
GuHCl	hspX (hspX)	Rv2031c	16210	100.00%	16	21	524	1.56%	76.40%	PGVDPDKDVDIMVR	95.00%	2.72	0.497	0	2	0	0	2	1,572.72	58	71
GuHCl	hspX (hspX)	Rv2031c	16210	100.00%	16	21	524	1.56%	76.40%	PVGADEDDIK	95.00%	2.36	0.3	14	0	0	0	2	1,059.14	105	114
GuHCl	hspX (hspX)	Rv2031c	16210	100.00%	16	21	524	1.56%	76.40%	SEFAYGSFVR	95.00%	3.2	0.561	18	8	0	0	2	1,163.28	91	100
GuHCl	hspX (hspX)	Rv2031c	16210	100.00%	16	21	524	1.56%	76.40%	SLFPEFSEL	95.00%	1.64	0.0877	3	0	0	0	2	1,069.30	13	21
GuHCl	hspX (hspX)	Rv2031c	16210	100.00%	16	21	524	1.56%	76.40%	SLFPEFSELFAAFPSFAGLR	95.00%	5.1	0.628	0	16	0	0	2	2,234.79	13	32
GuHCl	hspX (hspX)	Rv2031c	16210	100.00%	16	21	524	1.56%	76.40%	SLFPEFSELFAAFPSFAGLRP TFDTR	95.00%	3.36	0.582	0	0	44	0	2	2,952.61	13	38
GuHCl	hspX (hspX)	Rv2031c	16210	100.00%	16	21	524	1.56%	76.40%	TVSLPVGADEDDIK	95.00%	3.35	0.531	18	25	0	0	2	1,459.59	101	114
GuHCl	hspX (hspX)	Rv2031c	16210	100.00%	16	21	524	1.56%	76.40%	TVSLPVGADEDDIKATYDK	95.00%	4.62	0.607	0	13	0	0	2	2,038.27	101	119
GuHCl	hspX (hspX)	Rv2031c	16210	100.00%	16	21	524	1.56%	76.40%	YEVR	95.00%	1.1	0.123	2	0	0	0	2	566.5884	51	54
GuHCl	pra (pra)	Rv1078	25058	99.70%	2	2	2	0.01%	14.20%	VLAAFIDWAPYVVL	95.00%	1.97	0.265	0	0	1	0	2	1,577.89	95	108
GuHCl	pra (pra)	Rv1078	25058	99.70%	2	2	2	0.01%	14.20%	WDAKRQTLADKIMTTVCVP I	95.00%	1.68	0.269	0	1	0	0	2	2,290.74	221	240
GuHCl	ppsA (ppsA)	Rv2931	198816	99.70%	2	2	2	0.01%	4.10%	DEPIAVVGMGCRFPGGISC PEALWDFLCER	95.00%	1.71	0.305	0	0	1	0	2	3,285.76	101	130
GuHCl	ppsA (ppsA)	Rv2931	198816	99.70%	2	2	2	0.01%	4.10%	QVMWPKIGGSQVLHDAFP PGSVDFFYLTASAAGIFGIP GQGSYAAAN	95.00%	1.35	0.308	0	0	1	0	2	4,843.54	1611	1657
GuHCl	kdpA (kdpA)	Rv1029	60146	99.70%	2	2	2	0.01%	8.23%	MVHDRRQGWAVLAIMGII WIGMSVAAMSFEAK	95.00%	2.53	0.304	0	0	1	0	2	3,609.34	278	309
GuHCl	kdpA (kdpA)	Rv1029	60146	99.70%	2	2	2	0.01%	8.23%	TPEYLGKKIQATEMK	95.00%	1.31	0.334	0	1	0	0	2	1,754.17	413	427
GuHCl	drrC (drrC)	Rv2938	29576	99.90%	3	3	4	0.01%	14.50%	GFAMGGPVLSPMIGML GFAMGGPVLSPMIGMLVWT	89.20%	1.74	0.202	0	0	1	0	2	1,595.09	237	252
GuHCl	drrC (drrC)	Rv2938	29576	99.90%	3	3	4	0.01%	14.50%	AGICVVCAVPLAIGYR VWTAGICVVCAVPLAIGYR	95.00%	1.79	0.431	0	0	2	0	2	3,639.41	237	271
GuHCl	drrC (drrC)	Rv2938	29576	99.90%	3	3	4	0.01%	14.50%	RASTH SQVEEIIGQGQQAPSGHIP	85.80%	1.43	0.245	0	0	1	0	2	2,615.90	253	276
GuHCl	clpC (clpC)	Rv3596c	93537	99.70%	2	2	2	0.01%	5.19%	FTPR TAVVEGLAQAIVHGEVPETL K	95.00%	3.39	0.602	0	1	0	0	2	2,477.64	61	83
GuHCl	clpC (clpC)	Rv3596c	93537	99.70%	2	2	2	0.01%	5.19%	AAAADGAAVGCGHAGDGN	95.00%	1.97	0.194	0	1	0	0	2	2,162.26	223	243
GuHCl	hypo protein Rv2280	Rv2280	48035	99.20%	2	2	8	0.02%	11.50%	VHMAIACKDPEK	85.70%	1.49	0.214	0	0	1	0	2	2,896.03	363	392
GuHCl	hypo protein Rv2280	Rv2280	48035	99.20%	2	2	8	0.02%	11.50%	GSGCGLSGAARPVEGGLLI SFDR	95.00%	2.41	0.331	0	7	0	0	2	2,291.69	70	92
GuHCl	pta (pta)	Rv0408	72912	97.00%	2	2	2	0.01%	7.54%	AAELGVDLDGATVIEPCASE L	87.30%	1.27	0.277	0	1	0	0	2	2,145.36	419	439
GuHCl	pta (pta)	Rv0408	72912	97.00%	2	2	2	0.01%	7.54%	TVPGISTVSSIFLMCLPDRV LAYGDCAIIPNI	79.50%	1.89	0.156	0	0	1	0	2	3,410.14	505	535
GuHCl	accD6	Rv2247	50117	99.70%	2	2	2	0.01%	4.86%	SKAEAGDTDIH	95.00%	1.34	0.0492	1	0	0	0	2	1,144.21	221	231
GuHCl	accD6	Rv2247	50117	99.70%	2	2	2	0.01%	4.86%	TYGGAYIAMNSR	95.00%	1.57	0.208	1	0	0	0	2	1,304.56	363	374
GuHCl	kasA (kasA)	Rv2245	43266	99.70%	2	2	2	0.01%	9.62%	AWRQIVMGDADVAVCGGV EGPIEALPIAAFSMMRAMST R	95.00%	1.44	0.249	0	0	1	0	2	4,125.76	181	219
GuHCl	kasA (kasA)	Rv2245	43266	99.70%	2	2	2	0.01%	9.62%	QIVMGDADVAVCGGVEGPI EALPIAAFSMMRAMSTRN	95.00%	1.92	0.242	0	0	1	0	2	3,897.46	184	220
GuHCl	tsf (tsf)	Rv2889c	28737	100.00%	3	3	5	0.01%	16.60%	GDDAAAAHAVALQIAALR	95.00%	4.91	0.458	0	3	0	0	2	1,734.98	167	184
GuHCl	tsf (tsf)	Rv2889c	28737	100.00%	3	3	5	0.01%	16.60%	NAEFQTLADQVVAAAAAAK	95.00%	2.21	0.259	0	1	0	0	2	1,890.00	82	100
GuHCl	tsf (tsf)	Rv2889c	28737	100.00%	3	3	5	0.01%	16.60%	NAEFQTLADQVVAAAAAAK PADVDALK	95.00%	1.34	0.416	0	1	0	0	2	2,699.95	82	108
GuHCl	hypo protein Rv2488c	Rv2488c	120802	100.00%	3	3	6	0.02%	6.77%	AEVLAWQAYVGFAGEGDP GATRAAGEEAR	87.90%	1.28	0.285	0	1	0	0	2	2,951.07	740	768
GuHCl	hypo protein Rv2488c	Rv2488c	120802	100.00%	3	3	6	0.02%	6.77%	AGSMLASMSKIHPGVDVVP VDWSADGVSELVPTGTVTL	84.10%	1.24	0.28	0	0	1	0	2	3,841.11	22	59
GuHCl	hypo protein Rv2488c	Rv2488c	120802	100.00%	3	3	6	0.02%	6.77%	LALVTEICHR	95.00%	1.23	-0.0399	4	0	0	0	2	1,155.37	414	423
GuHCl	hypo protein Rv1627c	Rv1627c	42368	99.70%	2	2	3	0.01%	8.96%	LLDICATSDGAAALIVASKS FTEK	94.90%	1.4	0.291	0	1	0	0	2	2,425.98	208	231
GuHCl	hypo protein Rv1627c	Rv1627c	42368	99.70%	2	2	3	0.01%	8.96%	YGATVEDFAQVK HDVDPNEVVSVGDEVEALV	95.00%	1.79	0.107	2	0	0	0	2	1,328.37	161	172
GuHCl	rpsA (rpsA)	Rv1630	53184	100.00%	6	6	13	0.04%	18.70%	LTK	95.00%	4.66	0.627	0	2	0	0	2	2,365.27	72	93
GuHCl	rpsA (rpsA)	Rv1630	53184	100.00%	6	6	13	0.04%	18.70%	HVEVPDQVVAVGDDAMVK	95.00%	4.24	0.615	0	4	0	0	2	1,924.77	330	347
GuHCl	rpsA (rpsA)	Rv1630	53184	100.00%	6	6	13	0.04%	18.70%	ISLSLK	95.00%	0.884	0.224	1	0	0	0	2	660.9784	358	363
GuHCl	rpsA (rpsA)	Rv1630	53184	100.00%	6	6	13	0.04%	18.70%	TAGGSLASDAQLAALR	95.00%	3.61	0.518	0	2	0	0	2	1,502.78	459	474
GuHCl	rpsA (rpsA)	Rv1630	53184	100.00%	6	6	13	0.04%	18.70%	THAIGQIVPGK	95.00%	1.68	0.474	0	1	0	0	2	1,121.29	290	300
GuHCl	rpsA (rpsA)	Rv1630	53184	100.00%	6	6	13	0.04%	18.70%	VEEGIEGLVHISELAER	95.00%	2.92	0.439	0	3	0	0	2	1,881.02	313	329

GuHCl	hypo protein (Rv0926c)	Rv0926c	37734	99.70%	2	2	2	0.01%	19.30%	CMEIVDYATVDSAVVMFDV MGFGKPMDQIPML	94.60%	1.81	0.243	0	0	1	0	2	3,635.28	164	195
GuHCl	hypo protein (Rv0926c)	Rv0926c	37734	99.70%	2	2	2	0.01%	19.30%	RADLCPEWPQPAQPGGSY RIEISGEPCYAMDICLSSR	95.00%	1.79	0.286	0	0	1	0	2	4,169.81	274	310
GuHCl	hypo protein (Rv2880c)	Rv2880c	29581	100.00%	3	3	3	0.01%	21.50%	AAAVRFRDFGPRGDGVDG GSGPCYPQPCRR	95.00%	2.24	0.155	0	0	1	0	2	3,308.61	196	225
GuHCl	hypo protein (Rv2880c)	Rv2880c	29581	100.00%	3	3	3	0.01%	21.50%	NTVCISSQAGCGMACPFCA TGQGGL	95.00%	1.84	0.293	0	0	1	0	2	2,606.79	116	140
GuHCl	hypo protein (Rv2880c)	Rv2880c	29581	100.00%	3	3	3	0.01%	21.50%	RNTVCISSQAGCGMACPFC ATGQGGLTRN	95.00%	2.01	0.222	0	0	1	0	2	2,992.24	115	143
GuHCl	hypo protein (Rv3894c)	Rv3894c	153685	99.80%	2	2	2	0.01%	3.30%	DVKLMVVTRN	95.00%	1.52	0.138	1	0	0	0	2	1,191.33	248	257
GuHCl	hypo protein (Rv3894c)	Rv3894c	153685	99.80%	2	2	2	0.01%	3.30%	TTTLMALMCSAATMYTPER VTFFCIGGATMAQIGSL	90.50%	1.74	0.257	0	0	1	0	2	3,823.62	892	927
GuHCl	whiB3 (whiB3)	Rv3416	11594	98.40%	2	2	2	0.01%	25.50%	GMDSSMFFHPDGERGRAR TQR	87.70%	1.44	0.235	0	1	0	0	2	2,439.73	25	45
GuHCl	whiB3	Rv3416	11594	98.40%	2	2	2	0.01%	25.50%	QGLCR	89.00%	0.855	0.102	1	0	0	0	2	576.6684	20	24
GuHCl	hypo protein (Rv2576c)	Rv2576c	15742	99.70%	2	2	3	0.01%	42.90%	QVTYTVTTTSDLMANIRYM SADPPSMAAFNADSSK	95.00%	2.28	0.218	0	0	2	0	2	3,819.30	51	85
GuHCl	hypo protein (Rv2576c)	Rv2576c	15742	99.70%	2	2	3	0.01%	42.90%	TPIAGGQPLVYTATLANPSQ WAIVTASGGLR	95.00%	1.97	0.317	0	0	1	0	2	3,112.58	92	122
GuHCl	hypo protein (Rv0386)	Rv0386	116775	99.90%	2	2	2	0.01%	5.81%	RLAVFVGCFFVDDAQAVAC SGDVQRYQVL	93.90%	1.59	0.257	0	0	1	0	2	3,249.42	447	475
GuHCl	hypo protein (Rv0386)	Rv0386	116775	99.90%	2	2	2	0.01%	5.81%	WCLGYAQLMRGELAAAAA QFGEVVDEAEASQEVL	95.00%	2.34	0.199	0	0	1	0	2	3,698.98	741	774
GuHCl	pks2 (pks2)	Rv3825c	225737	100.00%	4	4	4	0.01%	6.82%	FEAMLLLDDQTPVSTVATVT SPGVVDFAVEALQEGVGH	80.00%	0.906	0.312	0	0	1	0	2	3,945.08	983	1020
GuHCl	pks2 (pks2)	Rv3825c	225737	100.00%	4	4	4	0.01%	6.82%	MRETIEAIAAMGSEVVVEC GDIAQPGTAERLVATAVATG LPVR	95.00%	1.94	0.248	0	0	1	0	2	4,416.80	1811	1853
GuHCl	pks2 (pks2)	Rv3825c	225737	100.00%	4	4	4	0.01%	6.82%	TPGDMQTIELAAFHRVPPG PGQIEVAVRASSVN	95.00%	1.29	0.302	0	0	1	0	2	3,446.71	1434	1466
GuHCl	pks2 (pks2)	Rv3825c	225737	100.00%	4	4	4	0.01%	6.82%	VYTGYAPVIGAPWLVAFAER SRFFEVFSSSN	91.50%	1.74	0.217	0	0	1	0	2	3,469.88	1993	2023
GuHCl	wag31 (wag31)	Rv2145c	28260	100.00%	9	13	26	0.08%	37.70%	ADALQADAER	95.00%	1.61	0.218	1	0	0	0	2	1,060.14	180	189
GuHCl	wag31 (wag31)	Rv2145c	28260	100.00%	9	13	26	0.08%	37.70%	ANAEQILGEAR	95.00%	3.75	0.54	2	3	0	0	2	1,172.31	134	144
GuHCl	wag31 (wag31)	Rv2145c	28260	100.00%	9	13	26	0.08%	37.70%	GSAAPVDSNADAGGFDQF NR	95.00%	5.1	0.647	0	6	0	0	2	1,997.04	238	257
GuHCl	wag31 (wag31)	Rv2145c	28260	100.00%	9	13	26	0.08%	37.70%	HSEIMGTINQQR	95.00%	1.81	0.336	0	1	0	0	2	1,430.55	191	202
GuHCl	wag31 (wag31)	Rv2145c	28260	100.00%	9	13	26	0.08%	37.70%	HTADATVAEAR	95.00%	2.11	0.488	1	1	0	0	2	1,142.11	145	155
GuHCl	wag31 (wag31)	Rv2145c	28260	100.00%	9	13	26	0.08%	37.70%	KHSEIMGTINQQR	95.00%	1.79	0.416	0	1	0	0	2	1,558.75	190	202
GuHCl	wag31 (wag31)	Rv2145c	28260	100.00%	9	13	26	0.08%	37.70%	LIEENSDLR	95.00%	2.09	0.213	0	1	0	0	2	1,089.30	42	50
GuHCl	wag31 (wag31)	Rv2145c	28260	100.00%	9	13	26	0.08%	37.70%	TYLESQLEELGQR	95.00%	4.68	0.598	2	5	0	0	2	1,566.75	225	237
GuHCl	wag31 (wag31)	Rv2145c	28260	100.00%	9	13	26	0.08%	37.70%	VLSLAQDTADR	95.00%	2.83	0.418	1	1	0	0	2	1,189.33	106	116
GuHCl	alaS (alaS)	Rv2555c	97307	100.00%	3	3	3	0.01%	9.62%	DTYGFPIELTLEMAAETGLQ VDEIGFR	95.00%	1.84	0.247	0	0	1	0	2	3,033.52	401	427
GuHCl	alaS (alaS)	Rv2555c	97307	100.00%	3	3	3	0.01%	9.62%	PAERIQRRGMADNYWSMG IPGPCGPSSEIYYDR	95.00%	1.99	0.251	0	0	1	0	2	3,862.48	146	178
GuHCl	alaS (alaS)	Rv2555c	97307	100.00%	3	3	3	0.01%	9.62%	TFTEQLDKAKAMGAIALFG ESYPDEVR	95.00%	2.2	0.251	0	0	1	0	2	2,989.51	648	674
GuHCl	hypo protein (Rv0858c)	Rv0858c	42191	100.00%	3	3	4	0.01%	15.90%	LRPYATTVFAEMSAL	95.00%	1.79	0.264	2	0	0	0	2	1,671.03	6	20
GuHCl	hypo protein (Rv0858c)	Rv0858c	42191	100.00%	3	3	4	0.01%	15.90%	TITISSAAKMFN	95.00%	1.46	0.219	1	0	0	0	2	1,284.64	224	235
GuHCl	hypo protein (Rv0858c)	Rv0858c	42191	100.00%	3	3	4	0.01%	15.90%	VGVAAIPMSAFCDPAAGQA SQQADVWNHLVRFTFCK	82.00%	1.82	0.166	0	0	1	0	2	3,909.24	341	376
GuHCl	hypo protein (Rv2326c)	Rv2326c	72818	99.70%	2	2	4	0.01%	6.17%	QRLALAAALAR	95.00%	1.11	0.172	1	0	0	0	2	1,154.52	380	390
GuHCl	hypo protein (Rv2326c)	Rv2326c	72818	99.70%	2	2	4	0.01%	6.17%	RGRGTPTVVVSSLIGGFVF GAAMVGMLAAMVR	95.00%	1.79	0.332	0	0	3	0	2	3,241.74	121	152
GuHCl	rmlA (rmlA)	Rv0334	31475	99.70%	2	2	2	0.01%	17.00%	DIQLITPHDAPGFHR RFQSISGGAIFAYWVANPS	95.00%	4.21	0.515	0	1	0	0	2	1,819.09	48	63
GuHCl	rmlA (rmlA)	Rv0334	31475	99.70%	2	2	2	0.01%	17.00%	AYGVVEFGAEGMAL	95.00%	1.85	0.242	0	0	1	0	2	3,483.92	122	154
GuHCl	hypo protein (Rv0798c)	Rv0798c	28813	99.70%	2	2	3	0.01%	16.60%	TIFEGYSAASIEGIRSASSN PALTL	95.00%	1.97	0.237	0	0	1	0	2	2,557.06	118	142
GuHCl	hypo protein (Rv0798c)	Rv0798c	28813	99.70%	2	2	3	0.01%	16.60%	VVDVSDPGGPVTAAVSTGR HPAVIDGGVQALRIWGAGA	95.00%	2.54	0.596	0	2	0	0	2	1,784.68	36	54
GuHCl	bioF (bioF)	Rv1569	40009	100.00%	3	3	8	0.02%	15.50%	TGSR	95.00%	2.46	0.225	0	6	0	0	2	2,290.53	56	78
GuHCl	bioF (bioF)	Rv1569	40009	100.00%	3	3	8	0.02%	15.50%	IWGAGATGSRL MCGVAAVPDSAMVSVILGE	95.00%	1.33	0	1	0	0	0	2	1,089.31	69	79
GuHCl	bioF (bioF)	Rv1569	40009	100.00%	3	3	8	0.02%	15.50%	PESAVAAAAACLDAGVK	80.30%	1.8	0.161	0	0	1	0	2	3,478.75	306	341
GuHCl	pabB (pabB)	Rv1005c	47862	100.00%	3	3	3	0.01%	14.40%	CENPQFCQAR GIYCGTVGLASPVAGCELN	94.60%	1.31	0.0753	1	0	0	0	2	1,196.22	20	29
GuHCl	pabB (pabB)	Rv1005c	47862	100.00%	3	3	3	0.01%	14.40%	VAIRTVEFDTAGN TVEFDTAGNAVLGVGGGIT	84.30%	1.09	0.303	0	0	1	0	2	3,270.55	379	410
GuHCl	pabB (pabB)	Rv1005c	47862	100.00%	3	3	3	0.01%	14.40%	ADSDPDAEWAECLH	95.00%	1.43	0.301	0	0	1	0	2	3,319.38	402	434
GuHCl	plsB1 (plsB1)	Rv1551	69176	99.80%	2	2	2	0.01%	6.44%	RPEDLRFLVR	94.80%	1.06	0.0837	1	0	0	0	2	1,301.61	262	271

GuHCl	plsB1 (plsB1)	Rv1551	69176	99.80%	2	2	2	0.01%	6.44%	WTLHQMVASGVVSVYDAG TEAVWGIGEDQH	87.70%	1.62	0.235	0	0	1	0	2	3,244.17	382	411
GuHCl	hypo protein Rv0614	Rv0614	34626	99.90%	2	2	2	0.01%	28.20%	DTGVRLVDVKQDGGTSPPP PFAPWDTPDGTPPPGTGLS PTL	89.80%	2.02	0.238	0	0	1	0	2	4,142.57	174	214
GuHCl	hypo protein Rv0614	Rv0614	34626	99.90%	2	2	2	0.01%	28.20%	THCSTNDWIDRTAQFASGA IVFSIGVLTAETGAGVVAAA AGGVGAATAGVSL	95.00%	1.63	0.245	0	0	1	0	2	4,993.30	271	322
GuHCl	dnaE1	Rv1547	129310	100.00%	3	3	3	0.01%	3.55%	AGAFDSLGH	87.10%	0.951	-0.0264	1	0	0	0	2	874.9584	911	919
GuHCl	dnaE1	Rv1547	129310	100.00%	3	3	3	0.01%	3.55%	IALIAN	95.00%	0.952	0.197	1	0	0	0	2	614.8784	1091	1096
GuHCl	dnaE1 (dnaE1)	Rv1547	129310	100.00%	3	3	3	0.01%	3.55%	TYSSYGADIVDDAVVLVNA KVAVRDDR	92.90%	2.32	0.137	0	0	1	0	2	2,913.07	1064	1090
GuHCl	echA18 (echA18)	Rv3373	22645	99.90%	2	2	2	0.01%	12.20%	AMTKMDEASNPCGGDIEA EMCQL	95.00%	2.13	0.204	0	1	0	0	2	2,548.84	5	27
GuHCl	echA18 (echA18)	Rv3373	22645	99.90%	2	2	2	0.01%	12.20%	RAMTKMDEASNPCGGDIE AEMCQLMR	95.00%	1.25	0.269	0	0	1	0	2	2,937.36	4	29
GuHCl	fixB (fixB)	Rv3028c	31672	100.00%	10	10	18	0.05%	59.10%	AAVDSGYYPGQFQVGQTG K	95.00%	4.17	0.624	0	1	0	0	2	1,974.02	233	251
GuHCl	fixB (fixB)	Rv3028c	31672	100.00%	10	10	18	0.05%	59.10%	AGAVEAEPAAGAGEQVSVE VPAAAENAAR	95.00%	1.72	0.545	0	2	0	0	2	2,693.55	152	180
GuHCl	fixB (fixB)	Rv3028c	31672	100.00%	10	10	18	0.05%	59.10%	AGMQTSKTIVAVNKDEEAP IFEIADYGVVGDLFK	95.00%	2.31	0.319	0	0	1	0	2	3,658.17	271	304
GuHCl	fixB (fixB)	Rv3028c	31672	100.00%	10	10	18	0.05%	59.10%	ALGEPAAVVVGVPGTAAPLV DGLK	95.00%	4.22	0.534	0	4	0	0	2	2,202.43	28	51
GuHCl	fixB (fixB)	Rv3028c	31672	100.00%	10	10	18	0.05%	59.10%	EPAVAGDRPELTEATIVVAG GR	95.00%	2.1	0.485	0	3	0	0	2	2,209.34	185	206
GuHCl	fixB (fixB)	Rv3028c	31672	100.00%	10	10	18	0.05%	59.10%	GVGSAENFSVVEALADSLG AAVGASR	95.00%	6.4	0.633	0	3	0	0	2	2,435.50	207	232
GuHCl	fixB (fixB)	Rv3028c	31672	100.00%	10	10	18	0.05%	59.10%	IGSGLLVDVVDVR	95.00%	1.88	0.265	1	0	0	0	2	1,342.48	109	121
GuHCl	fixB (fixB)	Rv3028c	31672	100.00%	10	10	18	0.05%	59.10%	IYVAESDLVDK	95.00%	1.85	0.226	1	0	0	0	2	1,252.42	58	68
GuHCl	fixB (fixB)	Rv3028c	31672	100.00%	10	10	18	0.05%	59.10%	TIVAVNKDEEAPIFEIADYG VVGDLFK	95.00%	2.68	0.477	0	1	0	0	2	2,954.36	278	304
GuHCl	fixB (fixB)	Rv3028c	31672	100.00%	10	10	18	0.05%	59.10%	VAPQLTEAIK	95.00%	1.65	0.283	1	0	0	0	2	1,070.27	305	314
GuHCl	ctpV (ctpV)	Rv0969	80083	99.70%	2	2	3	0.01%	5.84%	AQVNGGPVVVGR TAVFVGQDGQVVGVLAVAD	95.00%	1.18	0.102	2	0	0	0	2	1,153.05	532	543
GuHCl	ctpV (ctpV)	Rv0969	80083	99.70%	2	2	3	0.01%	5.84%	TVKDDAADVVGRLH	95.00%	1.6	0.283	0	0	1	0	2	3,323.25	571	603
GuHCl	sucD (sucD)	Rv0952	31212	99.70%	2	2	2	0.01%	10.90%	LPVFGSVAEAMEK TNVSKPVVGYVAGFTAPEG	95.00%	2.8	0.514	0	1	0	0	2	1,394.57	63	75
GuHCl	sucD (sucD)	Rv0952	31212	99.70%	2	2	2	0.01%	10.90%	K	95.00%	4.87	0.524	0	1	0	0	2	2,022.13	236	255
GuHCl	hypo protein Rv0944	Rv0944	16445	99.00%	2	2	2	0.01%	24.10%	PCPVCGDTVREVSFADKSF QYCPTCQTGGKALADR	95.00%	2.22	0.31	0	0	1	0	2	3,823.12	117	151
GuHCl	hypo protein Rv0944	Rv0944	16445	99.00%	2	2	2	0.01%	24.10%	TGLPCPVCGDTVREVSFAD KSFQYCPTCQTGGKAL	82.60%	1.23	0.288	0	0	1	0	2	3,823.16	114	148
GuHCl	hypo protein Rv1984c	Rv1984c	21764	99.70%	2	2	2	0.01%	22.10%	IVLGGYSQGATVIDLSTSA MPPAVADH	95.00%	2.1	0.249	0	0	1	0	2	2,687.98	112	138
GuHCl	hypo protein Rv1984c	Rv1984c	21764	99.70%	2	2	2	0.01%	22.10%	TINLCAPDDPICTGGGNIM AH	95.00%	3.05	0.216	0	1	0	0	2	2,185.55	173	193
GuHCl	mbtB (mbtB)	Rv2383c	151584	99.90%	2	2	2	0.01%	4.38%	AYVLFTSGSTGEPKGVEVA HDAAMNTVETFIR	95.00%	2.42	0.296	0	0	1	0	2	3,415.65	693	724
GuHCl	mbtB (mbtB)	Rv2383c	151584	99.90%	2	2	2	0.01%	4.38%	NACRVVADSGDDCPDWVA GELWVSGRGIAR	94.50%	1.73	0.189	0	0	1	0	2	3,247.38	873	902
GuHCl	echA6 (echA6)	Rv0905	26012	98.90%	2	2	2	0.01%	20.60%	AGDGSARAIVLTGQGTAFC AGADLSGDAFAADYPDR	80.10%	1.12	0.252	0	0	1	0	2	3,560.97	39	74
GuHCl	echA6	Rv0905	26012	98.90%	2	2	2	0.01%	20.60%	VVAPDAFFQFPTSK	95.00%	3.14	0.479	0	1	0	0	2	1,554.74	112	125
GuHCl	fadD13 (fadD13)	Rv3089	54442	99.10%	2	2	2	0.01%	12.50%	LRSAAADEPAVECGGDDNL FIMYTSGTTGH	91.20%	2.16	0.204	0	0	1	0	2	3,116.39	141	170
GuHCl	fadD13 (fadD13)	Rv3089	54442	99.10%	2	2	2	0.01%	12.50%	WGEIAAAIVVADQNEVSEQ QIVEYCGTRLARYK	91.30%	1.78	0.159	0	0	1	0	2	3,783.06	437	469
GuHCl	fadH (fadH)	Rv1175c	72884	99.90%	3	3	3	0.01%	7.86%	RARSVAVVGAGPAGLATAA N	79.90%	1.8	0.283	0	1	0	0	2	1,809.94	375	394
GuHCl	fadH (fadH)	Rv1175c	72884	99.90%	3	3	3	0.01%	7.86%	SVAVVGAGPAGLATAAN	95.00%	1.08	0.179	1	0	0	0	2	1,426.46	378	394
GuHCl	fadH (fadH)	Rv1175c	72884	99.90%	3	3	3	0.01%	7.86%	VPTIVTSVPGGAFVDISSAV AEHVTIPVVASNR	95.00%	1.9	0.298	0	0	1	0	2	3,291.38	259	291
GuHCl	ftsK (ftsK)	Rv2748c	94390	99.10%	2	2	2	0.01%	8.15%	CSDDGPPAWPSAEVPQDD TATIPEASAGRGSGRRGRR	95.00%	1.64	0.26	0	0	1	0	2	3,825.06	344	380
GuHCl	ftsK (ftsK)	Rv2748c	94390	99.10%	2	2	2	0.01%	8.15%	PWQRGEQIAVTSVADGVD GVIATRLAVMSSKTVAR	84.00%	1.32	0.262	0	0	1	0	2	3,670.95	26	60
GuHCl	lppC (lppC)	Rv1911c	19780	99.70%	2	2	3	0.01%	19.40%	ATTVDATTPAPAAEPLTIASP MFADGAPIPVQFSCK	95.00%	2.29	0.306	0	0	2	0	2	3,605.13	41	76
GuHCl	lppC (lppC)	Rv1911c	19780	99.70%	2	2	3	0.01%	19.40%	ATTVDATTPAPAAEPLTIASP MFADGAPIPVQFSCKGAN	95.00%	1.72	0.313	0	0	1	0	2	3,902.44	41	79
GuHCl	hypo protein Rv0480c	Rv0480c	35145	99.90%	2	2	4	0.01%	15.90%	ARALDSMSYVAAAGQADP GDARTGVGASSAAPTGVG GSLVASPL	95.00%	2.08	0.161	0	0	1	0	2	4,032.37	248	291
GuHCl	hypo protein Rv0480c	Rv0480c	35145	99.90%	2	2	4	0.01%	15.90%	VVFPEATMCR	95.00%	1.21	0.18	3	0	0	0	2	1,153.26	96	105
GuHCl	35kd_ag (35kd_ag)	Rv2744c	29240	100.00%	12	12	57	0.17%	69.60%	ATEYNNAAEAFAAQLVTAE QSVEDLK	95.00%	5.02	0.643	0	7	0	0	2	2,784.90	95	120
GuHCl	35kd_ag (35kd_ag)	Rv2744c	29240	100.00%	12	12	57	0.17%	69.60%	GEALPAGGTTATPRPATETS GGAIAEQPYGQ	95.00%	2.26	0.378	0	3	0	0	2	2,958.16	240	270
GuHCl	35kd_ag (35kd_ag)	Rv2744c	29240	100.00%	12	12	57	0.17%	69.60%	LLSQLEQAK	95.00%	1.82	0.0279	3	0	0	0	2	1,030.28	154	162
GuHCl	35kd_ag (35kd_ag)	Rv2744c	29240	100.00%	12	12	57	0.17%	69.60%	MLEVEQAGIQMAGHSR	95.00%	3.87	0.485	0	4	0	0	2	1,789.93	215	230
GuHCl	35kd_ag (35kd_ag)	Rv2744c	29240	100.00%	12	12	57	0.17%	69.60%	QALTLADQATAAGDAAK	95.00%	3.97	0.601	0	5	0	0	2	1,616.83	78	94
GuHCl	35kd_ag (35kd_ag)	Rv2744c	29240	100.00%	12	12	57	0.17%	69.60%	RYANAIGSAELAESSVQGR	95.00%	2.06	0.446	0	2	0	0	2	1,980.15	196	214
GuHCl	35kd_ag (35kd_ag)	Rv2744c	29240	100.00%	12	12	57	0.17%	69.60%	SMSELAAPGNTPSLDEVR	95.00%	3.57	0.556	0	7	0	0	2	1,891.08	173	190
GuHCl	35kd_ag (35kd_ag)	Rv2744c	29240	100.00%	12	12	57	0.17%	69.60%	THQALTQQAAQVIGNQR	95.00%	5.62	0.673	0	8	0	0	2	1,864.88	40	56
GuHCl	35kd_ag (35kd_ag)	Rv2744c	29240	100.00%	12	12	57	0.17%	69.60%	TLHDQALSAAAQAK	95.00%	4.52	0.641	0	3	0	0	2	1,425.60	121	134
GuHCl	35kd_ag (35kd_ag)	Rv2744c	29240	100.00%	12	12	57	0.17%	69.60%	VQIQQAIEEAQR	95.00%	4.54	0.531	0	4	0	0	2	1,413.45	28	39
GuHCl	35kd_ag (35kd_ag)	Rv2744c	29240	100.00%	12	12	57	0.17%	69.60%	YANAIGSAELAESSVQGR	95.00%	4.32	0.547	0	10	0	0	2	1,823.95	197	214
GuHCl	35kd_ag (35kd_ag)	Rv2744c	29240	100.00%	12	12	57	0.17%	69.60%	YLMALFSSK	86.40%	1.3	0	1	0	0	0	2	1,076.46	11	19

GuHCl	ppiA (ppiA)	Rv0009	19222	100.00%	4	4	10	0.03%	35.20%	DYSTQNASGGPSGPFYDG AVFHR	95.00%	2.64	0.504	0	3	0	0	2	2,431.53	51	73
GuHCl	ppiA (ppiA)	Rv0009	19222	100.00%	4	4	10	0.03%	35.20%	HTIFGEVIDAESQR	95.00%	4.11	0.595	0	3	0	0	2	1,602.70	141	154
GuHCl	ppiA (ppiA)	Rv0009	19222	100.00%	4	4	10	0.03%	35.20%	IALFGNHAPK	95.00%	2.02	0.336	0	1	0	0	2	1,068.35	28	37
GuHCl	ppiA (ppiA)	Rv0009	19222	100.00%	4	4	10	0.03%	35.20%	VIQGFMIQGGDPTGTGR	95.00%	3.25	0.565	0	3	0	0	2	1,750.96	74	90
GuHCl	sigA (sigA)	Rv2703	57752	99.20%	2	2	2	0.01%	10.80%	EDDVAPPADDDLDSGDDE DHEDLEAEAAVAPGQTADD DEEIAEPTEK	95.00%	1.5	0.309	0	0	1	0	2	4,986.84	149	195
GuHCl	sigA (sigA)	Rv2703	57752	99.20%	2	2	2	0.01%	10.80%	EMDITPEKVL	86.80%	0.973	0	1	0	0	0	2	1,191.41	402	411
GuHCl	hypo protein Rv2005c	Rv2005c	30965	100.00%	3	3	4	0.01%	16.30%	GLLGSVSSSLVR	95.00%	2.22	0.503	0	1	0	0	2	1,175.38	127	138
GuHCl	hypo protein Rv2005c	Rv2005c	30965	100.00%	3	3	4	0.01%	16.30%	SELVFSTPVPTMVEISNEAE MVVLGSSGR	95.00%	3.08	0.591	0	2	0	0	2	3,099.29	93	121
GuHCl	hypo protein Rv2005c	Rv2005c	30965	100.00%	3	3	4	0.01%	16.30%	VPVIVAR	95.00%	1.3	0.238	1	0	0	0	2	753.8284	287	293
GuHCl	hypo protein Rv1245c	Rv1245c	29197	100.00%	3	3	5	0.01%	21.00%	NATAAEGLDQAELAETFDK R	95.00%	3.08	0.389	0	2	0	0	2	2,151.37	197	216
GuHCl	hypo protein Rv1245c	Rv1245c	29197	100.00%	3	3	5	0.01%	21.00%	VAISDVDTDGLADTEHR	95.00%	2.06	0.363	0	2	0	0	2	1,814.85	33	49
GuHCl	hypo protein Rv1245c	Rv1245c	29197	100.00%	3	3	5	0.01%	21.00%	VAVVTGAGSGIGQALAIELA R	95.00%	3.41	0.455	0	1	0	0	2	1,954.21	8	28
GuHCl	ilvX (ilvX)	Rv3509c	52022	100.00%	3	3	3	0.01%	17.70%	KYDAPL	87.50%	0.967	0	1	0	0	0	2	706.9184	106	111
GuHCl	ilvX (ilvX)	Rv3509c	52022	100.00%	3	3	3	0.01%	17.70%	PADVCWSDGAHAAAGVPA QAAAAPVDVGPVAGVLRSG EPAMMLIGGDATR	95.00%	1.37	0.279	0	0	1	0	2	4,843.17	158	207
GuHCl	ilvX (ilvX)	Rv3509c	52022	100.00%	3	3	3	0.01%	17.70%	SPVSFFAYPGMPSDLVPAG CEVHVLAEPGGAADAL	94.60%	1.59	0.247	0	0	1	0	2	3,473.82	270	304
GuHCl	mas (mas)	Rv2940c	224360	99.70%	2	2	3	0.01%	2.84%	AIVEEAPAEASAPESSPGDA EVGPRLFMLSSTSSDALR	95.00%	1.21	0.362	0	0	1	0	2	3,847.26	424	461
GuHCl	mas (mas)	Rv2940c	224360	99.70%	2	2	3	0.01%	2.84%	WRRAQGLPVSAIAWGAWG EVGR	95.00%	2.01	0.219	0	0	2	0	2	2,424.68	1925	1946
GuHCl	hypo protein Rv2314c	Rv2314c	48694	99.90%	2	2	2	0.01%	15.50%	ADTGEPDDWDAPVPGTGA GVFTGIAGSLSRGFR	95.00%	1.69	0.261	0	0	1	0	2	3,278.55	106	138
GuHCl	hypo protein Rv2314c	Rv2314c	48694	99.90%	2	2	2	0.01%	15.50%	TVELPAGRYQTIMPPSTVAD MMIYLGWSMAGRGAQEGR	95.00%	1.89	0.287	0	0	1	0	2	4,175.85	215	252
GuHCl	moxR	Rv1479	40744	100.00%	21	23	35	0.10%	59.70%	AIFEVK	86.30%	1.33	0.0881	1	0	0	0	2	706.8684	46	51
GuHCl	moxR	Rv1479	40744	100.00%	21	23	35	0.10%	59.70%	ASLGIIAAAR	95.00%	1.4	0.149	1	0	0	0	2	943.2684	288	297
GuHCl	moxR (moxR)	Rv1479	40744	100.00%	21	23	35	0.10%	59.70%	AVPQQGHSVPPVMQAAAA ASGR	95.00%	1.47	0.33	0	1	0	0	2	2,147.13	356	377
GuHCl	moxR	Rv1479	40744	100.00%	21	23	35	0.10%	59.70%	DYVIPQDVIEVIPDVLR	95.00%	4.01	0.534	0	1	0	0	2	1,984.24	306	322
GuHCl	moxR	Rv1479	40744	100.00%	21	23	35	0.10%	59.70%	EIIYR	95.00%	1.3	0.0766	1	0	0	0	2	693.9184	217	221
GuHCl	moxR	Rv1479	40744	100.00%	21	23	35	0.10%	59.70%	FLFK	79.80%	1.03	0	1	0	0	0	2	554.8184	201	204
GuHCl	moxR	Rv1479	40744	100.00%	21	23	35	0.10%	59.70%	GHVLLEGVPGVAK	95.00%	3.72	0.488	3	2	0	0	2	1,276.38	72	84
GuHCl	moxR	Rv1479	40744	100.00%	21	23	35	0.10%	59.70%	GRDYVIPQDVIEVIPDVLR	95.00%	5.41	0.654	0	2	0	0	2	2,197.49	304	322
GuHCl	moxR	Rv1479	40744	100.00%	21	23	35	0.10%	59.70%	IIVGQDQLVER	95.00%	1.77	0.138	1	0	0	0	2	1,270.41	53	63
GuHCl	moxR	Rv1479	40744	100.00%	21	23	35	0.10%	59.70%	INVGYPSPEEEREIIYR	95.00%	2.44	0.381	0	1	0	0	2	2,065.36	205	221
GuHCl	moxR	Rv1479	40744	100.00%	21	23	35	0.10%	59.70%	IQFTPDLVPTDIIGTR	95.00%	3.32	0.482	0	3	0	0	2	1,787.18	102	117
GuHCl	moxR	Rv1479	40744	100.00%	21	23	35	0.10%	59.70%	LVLTYDALADEISPEIVINR	95.00%	2.4	0.392	0	1	0	0	2	2,245.72	325	344
GuHCl	moxR	Rv1479	40744	100.00%	21	23	35	0.10%	59.70%	MLVGLLSK	95.00%	1.54	0.327	1	0	0	0	2	877.2178	64	71
GuHCl	moxR	Rv1479	40744	100.00%	21	23	35	0.10%	59.70%	QILSRGDLLR	95.00%	1.85	0.237	0	1	0	0	2	1,116.45	231	240
GuHCl	moxR	Rv1479	40744	100.00%	21	23	35	0.10%	59.70%	RIIVGQDQLVER	95.00%	2.67	0.474	0	2	0	0	2	1,426.61	52	63
GuHCl	moxR	Rv1479	40744	100.00%	21	23	35	0.10%	59.70%	SLALVR	95.00%	1.64	0.117	1	0	0	0	2	658.8484	298	303
GuHCl	moxR (moxR)	Rv1479	40744	100.00%	21	23	35	0.10%	59.70%	TFPMPSPFLVMATQNPIEHE GVYPLPEAQR	95.00%	2.36	0.417	0	1	1	0	2	3,430.87	169	198
GuHCl	moxR	Rv1479	40744	100.00%	21	23	35	0.10%	59.70%	TLAVETFAR	95.00%	1.6	0.308	2	0	0	0	2	1,008.15	85	93
GuHCl	moxR	Rv1479	40744	100.00%	21	23	35	0.10%	59.70%	VLQTVALPQVNAVPQQGH	95.00%	2.49	0.389	0	1	0	0	2	1,899.85	345	362
GuHCl	moxR (moxR)	Rv1479	40744	100.00%	21	23	35	0.10%	59.70%	VLQTVALPQVNAVPQQGHS VPPVMQAAAAASGR	95.00%	3.76	0.609	0	0	4	0	2	3,310.34	345	377
GuHCl	moxR	Rv1479	40744	100.00%	21	23	35	0.10%	59.70%	VQSALLEVMQER	95.00%	3.78	0.432	0	2	0	0	2	1,419.52	150	161
GuHCl	hypo protein Rv1424c	Rv1424c	27410	99.70%	2	2	2	0.01%	26.10%	MTVVPGAPSRPASAVSRPS YRQCVQASAQTSAR	95.00%	1.42	0.349	0	0	1	0	2	3,434.59	1	33
GuHCl	hypo protein Rv1424c	Rv1424c	27410	99.70%	2	2	2	0.01%	26.10%	VGYCYPAVTTPGKPPSVLLY APSACDESLPSPR	95.00%	2.51	0.458	0	0	1	0	2	3,437.93	157	189
GuHCl	bisC (bisC)	Rv1442	83373	100.00%	4	4	4	0.01%	12.80%	EWLEHMYDKWSAELDFPV PSFAEFWR	92.90%	1.75	0.194	0	0	1	0	2	3,317.75	526	551
GuHCl	bisC (bisC)	Rv1442	83373	100.00%	4	4	4	0.01%	12.80%	MLGGYTASR	95.00%	1.43	0.161	1	0	0	0	2	972.1778	130	138
GuHCl	bisC (bisC)	Rv1442	83373	100.00%	4	4	4	0.01%	12.80%	MQVYTSATHWGVFTARVH GGDIAAVAALASDTN	83.40%	1.83	0.159	0	0	1	0	2	3,435.57	1	33
GuHCl	bisC (bisC)	Rv1442	83373	100.00%	4	4	4	0.01%	12.80%	VDTIVVHEQYWTAMAKHA DIVVPTTTSFER	95.00%	1.63	0.4	0	0	1	0	2	3,446.61	444	473
GuHCl	hypo protein Rv3282	Rv3282	22992	99.70%	2	2	4	0.01%	20.70%	AAAKAAQVATTLTGTQRIVA ADCVVVACDSMLYIEGR	95.00%	1.96	0.297	0	0	3	0	2	3,911.32	54	90
GuHCl	hypo protein Rv3282	Rv3282	22992	99.70%	2	2	4	0.01%	20.70%	PLLRSLVQR	95.00%	1.56	0.248	1	0	0	0	2	1,082.39	193	201
GuHCl	hypo protein Rv3224	Rv3224	29797	99.80%	2	2	2	0.01%	11.70%	KPEVYADAAYVIVNKPATEY TGK	95.00%	3.36	0.515	0	1	0	0	2	2,528.84	216	238
GuHCl	hypo protein Rv3224	Rv3224	29797	99.80%	2	2	2	0.01%	11.70%	LPGTVFTAAK	90.00%	1.09	0.00827	1	0	0	0	2	1,005.22	45	54
GuHCl	fbpC2 (fbpC2)	Rv0129c	36754	100.00%	4	4	5	0.01%	18.80%	ADIQHVLNGATPPAAPAAPA A	95.00%	3.45	0.622	0	2	0	0	2	1,954.16	320	340
GuHCl	fbpC2	Rv0129c	36754	100.00%	4	4	5	0.01%	18.80%	FLEGLTLR	95.00%	1.17	0.124	1	0	0	0	2	949.2684	272	279
GuHCl	fbpC2	Rv0129c	36754	100.00%	4	4	5	0.01%	18.80%	PGLPVEYLQVPSASMGR	95.00%	3.13	0.572	0	1	0	0	2	1,818.06	50	66
GuHCl	fbpC2	Rv0129c	36754	100.00%	4	4	5	0.01%	18.80%	VQFQGGGPHAVYLLDGLR	95.00%	2.64	0.405	0	1	0	0	2	1,928.16	70	87
GuHCl	hypo protein Rv2376c	Rv2376c	16617	99.00%	2	2	4	0.01%	28.60%	MKMVKSIAAGLTAAAAIGA AAAGVTSIMAGGPVVYQM QPVVFGAPL	82.50%	1.37	0.189	0	0	2	0	2	4,406.22	1	46
GuHCl	hypo protein Rv2376c	Rv2376c	16617	99.00%	2	2	4	0.01%	28.60%	MVKSIAAGLTAAAAIGAAAA GVTSIMAGGPVVYQMQPV VFGAPLPL	95.00%	1.81	0.372	0	0	2	0	2	4,357.14	3	48
GuHCl	PPE (PPE)	Rv1361c	39965	99.00%	2	2	2	0.01%	10.60%	GFAPAAAQAVETAAQNGVQ AMSSL	86.10%	1.72	0.288	0	1	0	0	2	2,291.38	268	291
GuHCl	PPE (PPE)	Rv1361c	39965	99.00%	2	2	2	0.01%	10.60%	GQLTNSGGGFGGVSNALR	93.60%	1.4	0.0627	1	0	0	0	2	1,692.83	364	381

GuHCl	hypo protein Rv0830	Rv0830	33399	99.70%	2	2	2	0.01%	19.90%	AWPVGSIVYEVDMPEVIEF KTATLSDLGAEPATER	95.00%	1.12	0.347	0	0	1	0	2	3,823.24	116	150
GuHCl	hypo protein Rv0830	Rv0830	33399	99.70%	2	2	2	0.01%	19.90%	WDLATSVGATATMVAAQRA LAADPR	95.00%	1.86	0.274	0	1	0	0	2	2,560.85	9	33
GuHCl	glf (glf)	Rv3809c	45797	98.50%	2	2	2	0.01%	5.26%	TEADRALLATYR	88.20%	1.94	0.133	0	1	0	0	2	1,380.66	332	343
GuHCl	glf (glf)	Rv3809c	45797	98.50%	2	2	2	0.01%	5.26%	TVIMREYSR	88.80%	1.22	-0.0265	1	0	0	0	2	1,155.37	310	318
GuHCl	hypo protein Rv2623	Rv2623	31633	100.00%	8	10	36	0.11%	32.00%	AGPPTVHSEIVPAAAVPTLV DMSK	95.00%	5.61	0.672	0	11	1	0	2	2,404.61	86	109
GuHCl	hypo protein Rv2623	Rv2623	31633	100.00%	8	10	36	0.11%	32.00%	GGYAGMLVGSVGETVAQLA R	95.00%	4.56	0.654	0	10	0	0	2	1,953.10	267	286
GuHCl	hypo protein Rv2623	Rv2623	31633	100.00%	8	10	36	0.11%	32.00%	HLIDDALK	95.00%	2.08	0.282	3	0	0	0	2	925.1984	70	77
GuHCl	hypo protein Rv2623	Rv2623	31633	100.00%	8	10	36	0.11%	32.00%	LLGSVSSGLLR	95.00%	3.21	0.458	0	2	0	0	2	1,102.43	128	138
GuHCl	hypo protein Rv2623	Rv2623	31633	100.00%	8	10	36	0.11%	32.00%	QLVQR	92.20%	1.19	0.189	1	0	0	0	2	643.6884	248	252
GuHCl	hypo protein Rv2623	Rv2623	31633	100.00%	8	10	36	0.11%	32.00%	SEEAQLVVVGSR	95.00%	3.46	0.541	4	1	0	0	2	1,274.22	253	264
GuHCl	hypo protein Rv2623	Rv2623	31633	100.00%	8	10	36	0.11%	32.00%	TPVIVAR	95.00%	1.68	0.218	1	0	0	0	2	755.8584	287	293
GuHCl	hypo protein Rv2623	Rv2623	31633	100.00%	8	10	36	0.11%	32.00%	YPNVAITR	95.00%	2.38	0.274	0	2	0	0	2	934.0884	231	238
GuHCl	hypo protein Rv1038c	Rv1197,Rv2347c	10959	100.00%	5	6	24	0.07%	48.00%	DANNYEQQEQASQQILSS	95.00%	2.79	0.613	0	11	0	0	2	2,054.02	81	98
GuHCl	hypo protein Rv1038c	Rv1197,Rv2347c	10959	100.00%	5	6	24	0.07%	48.00%	DGLVR	95.00%	1.1	0.0548	1	0	0	0	2	559.6384	76	80
GuHCl	hypo protein Rv1038c	Rv1197,Rv2347c	10959	100.00%	5	6	24	0.07%	48.00%	FEVHAQTVEDEAR	95.00%	4.9	0.615	1	7	0	0	2	1,531.42	19	31
GuHCl	hypo protein Rv1038c	Rv1197,Rv2347c	10959	100.00%	5	6	24	0.07%	48.00%	FEVHAQTVEDEARR	95.00%	1.72	0.325	0	1	0	0	2	1,687.62	19	32
GuHCl	hypo protein Rv1038c	Rv1197,Rv2347c	10959	100.00%	5	6	24	0.07%	48.00%	NIVNMLHGVR	95.00%	2.5	0.358	0	3	0	0	2	1,169.31	66	75
GuHCl	hypo protein Rv0148	Rv0148	29760	100.00%	10	10	13	0.04%	59.40%	AVANYDSVATEDGAANIIK	95.00%	3.06	0.524	0	1	0	0	2	1,923.09	66	84
GuHCl	hypo protein Rv0148	Rv0148	29760	100.00%	10	10	13	0.04%	59.40%	DGTGAGSAMADEVVAEIR	95.00%	3.86	0.46	0	1	0	0	2	1,765.81	43	60
GuHCl	hypo protein Rv0148	Rv0148	29760	100.00%	10	10	13	0.04%	59.40%	EYALTLAGEGASVVVNDLG GAR	95.00%	2.72	0.424	0	1	0	0	2	2,163.33	21	42
GuHCl	hypo protein Rv0148	Rv0148	29760	100.00%	10	10	13	0.04%	59.40%	IAGFKL	95.00%	1.65	0.0764	1	0	0	0	2	648.9484	281	286
GuHCl	hypo protein Rv0148	Rv0148	29760	100.00%	10	10	13	0.04%	59.40%	MTQDILPPEVLEK	95.00%	2.95	0.357	0	2	0	0	2	1,529.83	200	212
GuHCl	hypo protein Rv0148	Rv0148	29760	100.00%	10	10	13	0.04%	59.40%	TALDEFGAVHGVVSNAGIL R	95.00%	4.94	0.524	0	1	0	0	2	2,027.20	85	104
GuHCl	hypo protein Rv0148	Rv0148	29760	100.00%	10	10	13	0.04%	59.40%	VALFGNDGANFDKPPSVQD VAAR	95.00%	4.38	0.594	0	2	0	0	2	2,389.57	247	269
GuHCl	hypo protein Rv0148	Rv0148	29760	100.00%	10	10	13	0.04%	59.40%	VVVATSTSGLFGNFGQTNY GAAK	95.00%	2.77	0.398	0	1	0	0	2	2,290.43	146	168
GuHCl	hypo protein Rv0148	Rv0148	29760	100.00%	10	10	13	0.04%	59.40%	WAEITDLSGAK	95.00%	1.86	0.207	1	0	0	0	2	1,191.41	270	280
GuHCl	hypo protein Rv0148	Rv0148	29760	100.00%	10	10	13	0.04%	59.40%	YNIHANALAPIAATR	95.00%	2.84	0.303	0	2	0	0	2	1,596.94	185	199
GuHCl	hypo protein Rv0142	Rv0142	33004	99.80%	3	3	3	0.01%	22.10%	CVAWGSGAEEFVDMAPAM LGAADDASDFVPL	88.70%	2.1	0.173	0	0	1	0	2	3,247.52	69	99
GuHCl	hypo protein Rv0142	Rv0142	33004	99.80%	3	3	3	0.01%	22.10%	CVAWGSGAEEFVDMAPAM LGAADDASDFVPLHPAVAA AH	88.60%	1.11	0.268	0	0	1	0	2	3,899.15	69	107
GuHCl	hypo protein Rv0142	Rv0142	33004	99.80%	3	3	3	0.01%	22.10%	GQQDPCFRVPGDGTIWRT SLLPTGPVTAR	87.80%	1.59	0.219	0	0	1	0	2	3,198.58	30	58
GuHCl	pks13 (pks13)	Rv3800c	186428	99.20%	2	2	2	0.01%	3.69%	AVAAGKQAPNVFSVDGPVT TGPVWVLAGFGAQHR	85.90%	1.13	0.275	0	0	1	0	2	3,406.51	691	724
GuHCl	pks13 (pks13)	Rv3800c	186428	99.20%	2	2	2	0.01%	3.69%	FDEFGNIITDSAVAEEPEPE LPGVTEEALR	95.00%	1.63	0.3	0	0	1	0	2	3,276.54	577	606
GuHCl	hypo protein Rv0139	Rv0139	36958	99.70%	2	2	2	0.01%	17.40%	TRFHGDVFDTATVAEAMAG CDDVYYCVVDTR	95.00%	0.957	0.465	0	0	1	0	2	3,445.49	52	82
GuHCl	hypo protein Rv0139	Rv0139	36958	99.70%	2	2	2	0.01%	17.40%	VLQYAHDAGLPAVAMCVST TYGGGDWGR	83.30%	1.28	0.26	0	0	1	0	2	2,913.10	156	183
GuHCl	pepA (pepA)	Rv0125	34907	99.70%	2	2	2	0.01%	20.60%	GQVVGMNTAASDNFQLSQ GGQGFAIPIGQAMAIAGQI R	95.00%	1.16	0.397	0	0	1	0	2	3,838.27	217	254
GuHCl	pepA (pepA)	Rv0125	34907	99.70%	2	2	2	0.01%	20.60%	TQDVAVLQLRGAGGLPSAA IGGGVAVGEPVVAMGN	95.00%	1.33	0.336	0	0	1	0	2	3,249.44	124	158
GuHCl	hypo protein Rv0111	Rv0111	74323	99.30%	2	2	2	0.01%	6.42%	FAARCAATVVLAGASWWLI EQPIRR	92.90%	1.36	0.272	0	0	1	0	2	2,787.29	374	398
GuHCl	hypo protein Rv0111	Rv0111	74323	99.30%	2	2	2	0.01%	6.42%	IVVAGAAVIVVASVAMEQR	91.10%	1.85	0.242	0	1	0	0	2	1,899.97	310	328
GuHCl	hypo protein Rv1666c	Rv1666c	47816	99.00%	2	2	2	0.01%	3.72%	LYSPGVISARR	94.40%	1.34	0.0537	1	0	0	0	2	1,219.50	294	304
GuHCl	hypo protein Rv1666c	Rv1666c	47816	99.00%	2	2	2	0.01%	3.72%	SAPAQ	84.60%	0.517	0.39	1	0	0	0	2	473.4784	426	430
GuHCl	pdhB (pdhB)	Rv2496c	38046	99.60%	2	2	2	0.01%	17.00%	PARPDETLAVAVSDITQSLT MVQAINR	93.60%	1.63	0.285	0	1	0	0	2	2,914.25	8	34
GuHCl	pdhB (pdhB)	Rv2496c	38046	99.60%	2	2	2	0.01%	17.00%	SGTDVTVVTYGNLVSTALS SADTAEQQHDWSL	95.00%	2.18	0.429	0	0	1	0	2	3,354.34	222	253
GuHCl	hypo protein Rv3874	Rv3874	10776	100.00%	5	6	15	0.04%	61.00%	ADEEQQQALSSQMGF	95.00%	2.51	0.266	0	1	0	0	2	1,685.69	86	100
GuHCl	hypo protein Rv3874	Rv3874	10776	100.00%	5	6	15	0.04%	61.00%	QELDEISTNIR	95.00%	1.66	0.271	1	0	0	0	2	1,318.50	67	77
GuHCl	hypo protein Rv3874	Rv3874	10776	100.00%	5	6	15	0.04%	61.00%	QKQELDEISTNIR	95.00%	3.74	0.576	0	3	0	0	2	1,574.80	65	77
GuHCl	hypo protein Rv3874	Rv3874	10776	100.00%	5	6	15	0.04%	61.00%	TDAATLAQEAGNFER	95.00%	5.22	0.586	3	6	0	0	2	1,594.69	6	20
GuHCl	hypo protein Rv3874	Rv3874	10776	100.00%	5	6	15	0.04%	61.00%	TQIDQVESTAGSLQGQWR	95.00%	4.57	0.572	0	1	0	0	2	2,005.03	27	44

GuHCl	hypo protein Rv0831c	Rv0831c	30171	100.00%	4	4	6	0.02%	33.20%	DALVSTFQDLYGPAQVVFQ EMITSR	95.00%	1.66	0.271	0	1	0	0	2	2,833.12	239	263
GuHCl	hypo protein Rv0831c	Rv0831c	30171	100.00%	4	4	6	0.02%	33.20%	FTPGGLVLTEWQGAAVYR	95.00%	2.75	0.531	0	1	0	0	2	1,966.19	163	180
GuHCl	hypo protein Rv0831c	Rv0831c	30171	100.00%	4	4	6	0.02%	33.20%	LINDLPIERQAQDVSWGMT APGGAPTPVADR	89.40%	1.35	0.241	0	0	1	0	2	3,277.69	43	73
GuHCl	hypo protein Rv0831c	Rv0831c	30171	100.00%	4	4	6	0.02%	33.20%	YGPGMGQALDPNYHLR	95.00%	3.72	0.452	0	3	0	0	2	1,806.09	192	207
GuHCl	mmpL1 (mmpL1)	Rv0402c	104792	99.60%	2	2	2	0.01%	3.34%	GPHDAPSLIAMK	81.70%	1.27	0.164	1	0	0	0	2	1,237.55	54	65
GuHCl	mmpL1 (mmpL1)	Rv0402c	104792	99.60%	2	2	2	0.01%	3.34%	TFDAANNDDSFYLPPEAFQ N	95.00%	1.52	0.287	0	1	0	0	2	2,277.48	667	686
GuHCl	hypo protein Rv3835	Rv3835	47026	99.90%	3	3	3	0.01%	14.70%	FEVAESIDMR	90.50%	0.998	0	1	0	0	0	2	1,197.35	122	131
GuHCl	hypo protein Rv3835	Rv3835	47026	99.90%	3	3	3	0.01%	14.70%	RFEVAESIDMRTFPGMEYG QNAAPPSPAR	90.00%	2.17	0.222	0	0	1	0	2	3,242.65	121	149
GuHCl	hypo protein Rv3835	Rv3835	47026	99.90%	3	3	3	0.01%	14.70%	VWPAGTCLGIDATTNQPID VPVDCAAPHAMEVSGTVN	89.20%	1.79	0.166	0	0	1	0	2	3,910.03	224	260
GuHCl	hypo protein Rv1179c	Rv1179c	100624	97.70%	2	2	2	0.01%	3.62%	ESRAFAGNWR	83.30%	1.14	0.049	1	0	0	0	2	1,194.31	8	17
GuHCl	hypo protein Rv1179c	Rv1179c	100624	97.70%	2	2	2	0.01%	3.62%	LVTGQTFACPAAIEDDLIAF CAER	87.70%	1.61	0.261	0	1	0	0	2	2,627.04	470	493
GuHCl	atpG (atpG)	Rv1309	33874	100.00%	3	3	4	0.01%	17.70%	IAPMVVEYVEEDIGPR	95.00%	2.81	0.498	0	1	0	0	2	1,834.00	202	217
GuHCl	atpG (atpG)	Rv1309	33874	100.00%	3	3	4	0.01%	17.70%	QAQITQEISEIVGGANALAE AR	95.00%	2.5	0.437	0	1	0	0	2	2,270.47	284	305
GuHCl	atpG (atpG)	Rv1309	33874	100.00%	3	3	4	0.01%	17.70%	VYAALLESAASELASR	95.00%	3.81	0.473	0	2	0	0	2	1,651.93	242	257
GuHCl	atpA (atpA)	Rv1308	59271	100.00%	3	3	3	0.01%	8.38%	ASEEEILTEIR	95.00%	1.24	0.262	1	0	0	0	2	1,290.48	477	487
GuHCl	atpA (atpA)	Rv1308	59271	100.00%	3	3	3	0.01%	8.38%	LSDDLGGGSLTGLPIIETK	95.00%	4.89	0.59	0	1	0	0	2	1,887.40	315	333
GuHCl	atpA (atpA)	Rv1308	59271	100.00%	3	3	3	0.01%	8.38%	TGEVLSVPVGDGFLGR	95.00%	2.37	0.463	0	1	0	0	2	1,603.73	94	109
GuHCl	fadE23 (fadE23)	Rv3140	43327	99.70%	2	2	2	0.01%	8.98%	ILDIFEGTQQIQQLVVAR	95.00%	4.47	0.525	0	1	0	0	2	2,072.39	373	390
GuHCl	fadE23 (fadE23)	Rv3140	43327	99.70%	2	2	2	0.01%	8.98%	TVELAGTTGYSEQSLLEK	95.00%	1.27	0.333	0	1	0	0	2	1,927.13	349	366
GuHCl	cysA2	Rv0815c,Rv3117	30997	99.70%	2	2	2	0.01%	8.30%	AFRDEVLAAINVK	95.00%	3.63	0.36	0	1	0	0	2	1,446.70	146	158
GuHCl	cysA2	Rv0815c,Rv3117	30997	99.70%	2	2	2	0.01%	8.30%	DFVDAQQFSK	95.00%	1.67	0.283	1	0	0	0	2	1,185.25	58	67
GuHCl	ilvB (ilvB)	Rv3003c	66072	99.90%	2	2	2	0.01%	6.80%	HHHIPGTIEMADWWAYL SWLNSGGLGTMGFAIPAA	95.00%	2.39	0.201	0	1	0	0	2	2,094.35	369	385
GuHCl	ilvB (ilvB)	Rv3003c	66072	99.90%	2	2	2	0.01%	6.80%	MGAKIAL	95.00%	1.73	0.301	0	1	0	0	2	2,452.15	445	469
GuHCl	hypo protein Rv1969	Rv1969	44371	99.70%	2	2	2	0.01%	18.20%	AIQLTPPYAGGPTMADGTVI PQERTVVPVEWDDLRAQL	95.00%	1.62	0.21	0	0	1	0	2	4,123.57	103	140
GuHCl	hypo protein Rv1969	Rv1969	44371	99.70%	2	2	2	0.01%	18.20%	PNEVTYSEDWLRPDYVAPV ADTPPDPAAAVTVDPATGL R	94.40%	1.56	0.225	0	0	1	0	2	4,168.42	375	413
GuHCl	hypo protein Rv1941	Rv1941	25799	97.70%	2	2	2	0.01%	21.90%	CADIDGDAADAAATKIGCG AAACR	86.70%	1.88	0.18	0	0	1	0	2	2,424.72	37	60
GuHCl	hypo protein Rv1941	Rv1941	25799	97.70%	2	2	2	0.01%	21.90%	LPAFVDTPMQQTAMAMFD GALGAGGARSMIAR	85.10%	2.16	0.161	0	0	1	0	2	3,300.97	183	214
GuHCl	hypo protein Rv2721c	Rv2721c	72322	99.70%	2	2	2	0.01%	7.58%	GAKKGGQYPIGGDGIAQDF VGGK	95.00%	1.22	0.323	0	0	1	0	2	2,221.57	241	263
GuHCl	hypo protein Rv2721c	Rv2721c	72322	99.70%	2	2	2	0.01%	7.58%	VVPETFVPQAFVPEAVAPEA VPPDVHAADL	95.00%	2.09	0.249	0	0	1	0	2	3,113.11	576	605
GuHCl	hypo protein Rv1937	Rv1937	93412	99.90%	3	3	3	0.01%	7.27%	ECQSGICGTCVATCTAGRY QMGRTEGLSDVER	84.10%	1.31	0.232	0	0	1	0	2	3,596.76	40	71
GuHCl	hypo protein Rv1937	Rv1937	93412	99.90%	3	3	3	0.01%	7.27%	TMPVRCDQTVLDAAEEHGV AIVNECQSGICGTCVATCTA GR	90.50%	0.892	0.369	0	0	1	0	2	4,440.50	17	57
GuHCl	hypo protein Rv1937	Rv1937	93412	99.90%	3	3	3	0.01%	7.27%	YGAGHR	95.00%	1.24	0.232	1	0	0	0	2	660.6984	710	715
GuHCl	glcB (glcB)	Rv1837c	80386	100.00%	5	5	32	0.10%	16.50%	FLDDSVPLSSGSFGDATGF TVQDGQLVVALPDK	95.00%	3.19	0.523	0	2	0	0	2	3,384.72	175	207
GuHCl	glcB (glcB)	Rv1837c	80386	100.00%	5	5	32	0.10%	16.50%	GMWTMTELMADMVETK NVGHLMTNDAIVDTDGSE	95.00%	3.1	0.394	0	2	0	0	2	1,875.22	514	529
GuHCl	glcB (glcB)	Rv1837c	80386	100.00%	5	5	32	0.10%	16.50%	VFEGIMDAL	95.00%	3.08	0.482	0	2	0	0	2	2,865.12	340	366
GuHCl	glcB (glcB)	Rv1837c	80386	100.00%	5	5	32	0.10%	16.50%	SQPWILAYEDHNVDAGLAA GFSGR	95.00%	3.53	0.591	0	24	0	0	2	2,575.84	485	508
GuHCl	glcB (glcB)	Rv1837c	80386	100.00%	5	5	32	0.10%	16.50%	WGSLYDALYGTDVIPETDG AEK	95.00%	3.49	0.474	0	2	0	0	2	2,401.70	135	156
GuHCl	thrA (thrA)	Rv1294	45502	100.00%	3	3	3	0.01%	12.50%	GVPIELLTDDIEEL LMFYGQGAGGAPTASAVTG	95.00%	2.16	0.282	0	1	0	0	2	1,556.86	56	69
GuHCl	thrA (thrA)	Rv1294	45502	100.00%	3	3	3	0.01%	12.50%	DLVMAAR	95.00%	3.2	0.582	0	1	0	0	2	2,545.89	301	326
GuHCl	thrA (thrA)	Rv1294	45502	100.00%	3	3	3	0.01%	12.50%	YAQLPVAPMGFIETR	89.10%	1.59	0.223	0	1	0	0	2	1,710.04	340	354
GuHCl	hypo protein Rv0338c	Rv0338c	95436	99.70%	2	2	2	0.01%	3.51%	FWVGCAGAYDDKAKK	95.00%	1.26	0.198	1	0	0	0	2	1,731.01	479	493
GuHCl	hypo protein Rv0338c	Rv0338c	95436	99.70%	2	2	2	0.01%	3.51%	ILGMSMTAVVGVFALR	95.00%	1.52	0.0706	1	0	0	0	2	1,682.07	10	25
GuHCl	deaD (deaD)	Rv1253	61436	99.60%	2	2	2	0.01%	5.68%	HNGGKPR	95.00%	0.831	0.281	1	0	0	0	2	765.8384	552	558
GuHCl	deaD (deaD)	Rv1253	61436	99.60%	2	2	2	0.01%	5.68%	VDFLVLDEADEMLTMGFAD DVERIL	92.90%	1.31	0.326	0	0	1	0	2	2,874.34	157	181
GuHCl	hypo protein Rv1593c	Rv1593c	25890	99.40%	2	2	2	0.01%	16.10%	MAHGSTAHEVL	95.00%	1.09	0	1	0	0	0	2	1,153.18	1	11
GuHCl	hypo protein Rv1593c	Rv1593c	25890	99.40%	2	2	2	0.01%	16.10%	PGIRMIASTYLGVVPSPATP ELPADTR	89.40%	1.88	0.25	0	0	1	0	2	2,827.36	93	119
GuHCl	hypo protein Rv3243c	Rv3243c	30671	99.50%	2	3	3	0.01%	10.00%	MSPRVPRLRWDDPFR	91.00%	1.34	0.288	0	1	0	0	2	1,945.33	1	15
GuHCl	hypo protein Rv3243c	Rv3243c	30671	99.50%	2	3	3	0.01%	10.00%	TLFTTLQQLLIGK	95.00%	1.5	0.0286	1	1	0	0	2	1,476.97	46	58
GuHCl	fadE21 (fadE21)	Rv2789c	44726	99.70%	2	2	2	0.01%	10.50%	LYSSEAATDVAMEAVQLFG GNGYMAEYRVEQL	95.00%	1.23	0.314	0	0	1	0	2	3,530.84	346	377

GuHCl	fadE21 (fadE21)	Rv2789c	44726	99.70%	2	2	2	0.01%	10.50%	NIGQFQLIQLK	94.30%	0.878	0.185	1	0	0	0	2	1,302.67	298	308
GuHCl	tuf (tuf)	Rv0685	43543	100.00%	23	30	266	0.79%	70.70%	ADAVDDEELLELVEMEVR	95.00%	5.99	0.651	0	29	0	0	2	2,092.19	140	157
GuHCl	tuf (tuf)	Rv0685	43543	100.00%	23	30	266	0.79%	70.70%	AFDQIDNAPEER	95.00%	2.9	0.518	1	2	0	0	2	1,405.50	48	59
GuHCl	tuf (tuf)	Rv0685	43543	100.00%	23	30	266	0.79%	70.70%	EHVLLAR	95.00%	1.43	0.308	5	0	0	0	2	837.9684	120	126
GuHCl	tuf (tuf)	Rv0685	43543	100.00%	23	30	266	0.79%	70.70%	ELLAAQEFDEDAPVVR	95.00%	4.2	0.601	5	28	0	0	2	1,802.92	158	173
GuHCl	tuf (tuf)	Rv0685	43543	100.00%	23	30	266	0.79%	70.70%	ETDKPFLMPVEDVFTITGR	95.00%	3.79	0.489	0	26	1	0	2	2,212.55	207	225
GuHCl	tuf (tuf)	Rv0685	43543	100.00%	23	30	266	0.79%	70.70%	FAIR	94.80%	1.22	0	1	0	0	0	2	506.6984	377	380
GuHCl	tuf (tuf)	Rv0685	43543	100.00%	23	30	266	0.79%	70.70%	GITINIAHVEYQTDK	95.00%	2.74	0.515	0	4	0	0	2	1,702.92	62	76
GuHCl	tuf (tuf)	Rv0685	43543	100.00%	23	30	266	0.79%	70.70%	GITINIAHVEYQTDKR	95.00%	4.51	0.66	0	13	0	0	2	1,859.12	62	77
GuHCl	tuf (tuf)	Rv0685	43543	100.00%	23	30	266	0.79%	70.70%	GQVVTKPGTTTPHTEFEGQ VYIL	79.80%	1.35	0.289	0	1	0	0	2	2,503.62	292	314
GuHCl	tuf (tuf)	Rv0685	43543	100.00%	23	30	266	0.79%	70.70%	GQVVTKPGTTTPHTEFEGQ VYILSK	95.00%	4.12	0.57	0	2	2	0	2	2,718.90	292	316
GuHCl	tuf (tuf)	Rv0685	43543	100.00%	23	30	266	0.79%	70.70%	GQVVTKPGTTTPHTEFEGQ VYILSKDEGGR	95.00%	2.07	0.393	0	0	1	0	2	3,233.40	292	321
GuHCl	tuf (tuf)	Rv0685	43543	100.00%	23	30	266	0.79%	70.70%	GVINVNEEVEIVGIRPSTTK	95.00%	4.13	0.54	0	24	0	0	2	2,155.30	236	255
GuHCl	tuf (tuf)	Rv0685	43543	100.00%	23	30	266	0.79%	70.70%	HTPFFNNYR	95.00%	1.56	0.212	1	0	0	0	2	1,196.34	322	330
GuHCl	tuf (tuf)	Rv0685	43543	100.00%	23	30	266	0.79%	70.70%	HYAHVDAPGHADYIK	95.00%	3.48	0.561	1	6	0	0	2	1,694.80	78	92
GuHCl	tuf (tuf)	Rv0685	43543	100.00%	23	30	266	0.79%	70.70%	IAHVEYQTDKR	95.00%	2.67	0.523	0	2	0	0	2	1,360.47	67	77
GuHCl	tuf (tuf)	Rv0685	43543	100.00%	23	30	266	0.79%	70.70%	KADAVDDEELLELVEMEVR	95.00%	2.94	0.507	0	2	0	0	2	2,220.39	139	157
GuHCl	tuf (tuf)	Rv0685	43543	100.00%	23	30	266	0.79%	70.70%	LIQPVAMDEGLR	95.00%	3.11	0.461	0	5	0	0	2	1,358.64	365	376
GuHCl	tuf (tuf)	Rv0685	43543	100.00%	23	30	266	0.79%	70.70%	NMITGAAQMDGAILVVAAT DGPMPQTR	95.00%	5.18	0.643	0	53	6	0	2	2,779.15	93	119
GuHCl	tuf (tuf)	Rv0685	43543	100.00%	23	30	266	0.79%	70.70%	PQFYFR	95.00%	1.99	0.418	3	4	0	0	2	858.0384	331	336
GuHCl	tuf (tuf)	Rv0685	43543	100.00%	23	30	266	0.79%	70.70%	TTDVIGQVVTLPEGTEMVMP GDNTNISVK	95.00%	2.7	0.505	0	14	0	0	2	2,939.04	337	364
GuHCl	tuf (tuf)	Rv0685	43543	100.00%	23	30	266	0.79%	70.70%	TTLTAAITK	95.00%	1.3	0.378	1	0	0	0	2	920.1784	26	34
GuHCl	tuf (tuf)	Rv0685	43543	100.00%	23	30	266	0.79%	70.70%	TTVTGVEMFR	95.00%	2.52	0.432	0	1	0	0	2	1,157.21	256	265
GuHCl	tuf (tuf)	Rv0685	43543	100.00%	23	30	266	0.79%	70.70%	WVASVEELMNAVDESIPDP VR	95.00%	5.2	0.662	0	22	0	0	2	2,373.46	186	206
GuHCl	fadE6 (fadE6)	Rv0271c	78460	99.90%	2	2	2	0.01%	6.43%	AASPVEQIIIAQEFTAGR ALDDAGESSSDVEFAAAVA	95.00%	1.21	-0.00674	1	0	0	0	2	1,902.16	412	429
GuHCl	nadC (nadC)	Rv1596	29933	99.70%	2	2	4	0.01%	13.70%	ATLAPATAQR	95.00%	1.88	0.252	0	1	0	0	2	2,806.97	277	305
GuHCl	nadC (nadC)	Rv1596	29933	99.70%	2	2	4	0.01%	13.70%	VPPGEALMTLEAQTR YGPDVTTLATVPASATTTAS LVTR	95.00%	3.02	0.577	0	2	0	0	2	1,629.84	84	98
GuHCl	mmpL4 (mmpL4)	Rv0450c	105188	99.90%	2	2	6	0.02%	5.48%	DAPSFEAMGRIGMVFKEGD SDSFAMVIIEGN	95.00%	4.33	0.57	0	2	0	0	2	2,394.60	25	48
GuHCl	mmpL4 (mmpL4)	Rv0450c	105188	99.90%	2	2	6	0.02%	5.48%	EMDRLMPQMVAQIPPQIEA MEN	95.00%	2.2	0.371	0	0	1	0	2	3,338.86	64	94
GuHCl	hypo protein Rv2477c	Rv2477c	61876	99.70%	2	2	2	0.01%	9.50%	GRAYPYEGNYSTYLEK	95.00%	2.75	0.221	0	0	0	0	2	1,912.20	232	247
GuHCl	hypo protein Rv2477c	Rv2477c	61876	99.70%	2	2	2	0.01%	9.50%	SYVDQARAGIDPRKTVWEV VSDGLDY1QVGQTEVPSR	95.00%	2.02	0.275	0	0	1	0	2	4,136.33	387	423
GuHCl	hypo protein Rv2712c	Rv2712c	36939	99.00%	2	2	3	0.01%	16.20%	CVDGCGCSGVIDDPSASPL AMAAVLDGRQLYPR	82.50%	2.7	0.113	0	0	1	0	2	3,567.98	127	159
GuHCl	hypo protein Rv2712c	Rv2712c	36939	99.00%	2	2	3	0.01%	16.20%	PAAAIAVIVDANGAQCPRC NEEYR	95.00%	2.12	0.285	0	0	2	0	2	2,603.81	70	93
GuHCl	recC (recC)	Rv0631c	119486	99.60%	2	2	2	0.01%	3.83%	GAKPFTFDPAALAAAQAAA GKRCPPTAFISGR	92.30%	2.15	0.182	0	0	1	0	2	3,161.83	757	788
GuHCl	recC (recC)	Rv0631c	119486	99.60%	2	2	2	0.01%	3.83%	RRYGLDTVVH	95.00%	1.45	0.216	1	0	0	0	2	1,216.31	483	492
GuHCl	ponA' (ponA')	Rv3682	84619	99.50%	2	2	3	0.01%	6.05%	FQAKGLGSGGAK PEDGGPPPPVGSQVVEIPG	95.00%	1.36	0.161	2	0	0	0	2	1,121.36	425	436
GuHCl	ponA' (ponA')	Rv3682	84619	99.50%	2	2	3	0.01%	6.05%	LPPITIPLLAPPPPAPPP SMHTTMSVFVTSTVSMFGI	91.80%	1.9	0.171	0	0	1	0	2	3,609.29	774	810
GuHCl	mmpL11 (mmpL11)	Rv0202c	103485	98.80%	2	2	3	0.01%	5.59%	ALAVDYSLFILMR TQLPRVAGAAQVDVGGPTA	90.70%	1.6	0.192	0	0	1	0	2	3,606.31	225	256
GuHCl	mmpL11 (mmpL11)	Rv0202c	103485	98.80%	2	2	3	0.01%	5.59%	LIK	89.20%	1.84	0.236	0	2	0	0	2	2,163.44	497	518
GuHCl	ephE (ephE)	Rv3670	36316	99.20%	2	2	2	0.01%	8.87%	LVRAIALISSPHPAAL	94.50%	1.58	0.221	0	0	1	0	2	1,630.11	143	158
GuHCl	ephE (ephE)	Rv3670	36316	99.20%	2	2	2	0.01%	8.87%	QAIQIPAAAHCAL	86.80%	1.3	0.0524	1	0	0	0	2	1,307.54	224	236
GuHCl	dnaK (dnaK)	Rv0350	66813	100.00%	19	22	43	0.13%	40.30%	AALGGSDISAIK DAEAYLGEDITDAVITTPAY	95.00%	1.34	0.126	1	0	0	0	2	1,103.42	553	564
GuHCl	dnaK (dnaK)	Rv0350	66813	100.00%	19	22	43	0.13%	40.30%	FNDAQR	95.00%	2.93	0.518	0	2	0	0	2	2,861.16	102	127
GuHCl	dnaK (dnaK)	Rv0350	66813	100.00%	19	22	43	0.13%	40.30%	DAGQIAGLNVLR EAEGGSKVPEDTLNKVDAA	95.00%	2.95	0.482	0	3	0	0	2	1,227.45	132	143
GuHCl	dnaK (dnaK)	Rv0350	66813	100.00%	19	22	43	0.13%	40.30%	VAEAK	95.00%	4.55	0.513	0	6	1	0	2	2,429.53	529	552
GuHCl	dnaK (dnaK)	Rv0350	66813	100.00%	19	22	43	0.13%	40.30%	GTSGIDLTK GVNPDEVVAVGAALQAGVL	95.00%	1.05	0.25	1	0	0	0	2	892.0984	218	226
GuHCl	dnaK (dnaK)	Rv0350	66813	100.00%	19	22	43	0.13%	40.30%	K	95.00%	4.66	0.636	0	2	0	0	2	1,907.96	338	357
GuHCl	dnaK (dnaK)	Rv0350	66813	100.00%	19	22	43	0.13%	40.30%	IQEGSGLSKEDIDR	95.00%	3.67	0.44	0	1	0	0	2	1,547.78	475	488
GuHCl	dnaK (dnaK)	Rv0350	66813	100.00%	19	22	43	0.13%	40.30%	IVNEPTAAALAYGLDK	95.00%	4.36	0.559	1	3	0	0	2	1,646.98	144	159
GuHCl	dnaK (dnaK)	Rv0350	66813	100.00%	19	22	43	0.13%	40.30%	IVNEPTAAALAYGLDKGEK	95.00%	4.08	0.622	0	4	0	0	2	1,961.33	144	162
GuHCl	dnaK (dnaK)	Rv0350	66813	100.00%	19	22	43	0.13%	40.30%	KYTAPEISAR	95.00%	2.48	0.365	0	1	0	0	2	1,136.38	85	94
GuHCl	dnaK (dnaK)	Rv0350	66813	100.00%	19	22	43	0.13%	40.30%	LLGSFELTGIPPAPR	95.00%	2.96	0.521	0	2	0	0	2	1,569.04	427	441
GuHCl	dnaK (dnaK)	Rv0350	66813	100.00%	19	22	43	0.13%	40.30%	MPAVTDLVK	95.00%	2.46	0.429	2	1	0	0	2	990.1578	319	327
GuHCl	dnaK (dnaK)	Rv0350	66813	100.00%	19	22	43	0.13%	40.30%	NGEVLVGQPAK	95.00%	2.72	0.43	1	0	0	0	2	1,112.16	45	55
GuHCl	dnaK (dnaK)	Rv0350	66813	100.00%	19	22	43	0.13%	40.30%	NQAETLVYQTEK RDAEAYLGEDITDAVITTPA	95.00%	3.67	0.587	0	1	0	0	2	1,424.47	511	522
GuHCl	dnaK (dnaK)	Rv0350	66813	100.00%	19	22	43	0.13%	40.30%	YFNDAQRQATK SETFTTADDNQPSVQIQVY	90.60%	2.16	0.152	0	0	1	0	2	3,445.84	101	131
GuHCl	dnaK (dnaK)	Rv0350	66813	100.00%	19	22	43	0.13%	40.30%	QGER SVVSVLEGGDPVVVANSEG	95.00%	4.18	0.55	0	4	0	0	2	2,614.57	397	419
GuHCl	dnaK (dnaK)	Rv0350	66813	100.00%	19	22	43	0.13%	40.30%	SR	95.00%	3.8	0.588	0	1	0	0	2	2,057.91	14	34
GuHCl	dnaK (dnaK)	Rv0350	66813	100.00%	19	22	43	0.13%	40.30%	TTPSIVAFAR	95.00%	1.37	0.27	2	0	0	0	2	1,063.25	35	44
GuHCl	dnaK (dnaK)	Rv0350	66813	100.00%	19	22	43	0.13%	40.30%	VPEDTLNKVDAAVAEAK AGLVGKLSVPQGWTVAAPE	95.00%	3.69	0.53	0	2	0	0	2	1,770.87	536	552
GuHCl	PPE (PPE)	Rv1801	41460	99.10%	2	2	2	0.01%	16.30%	IPSPAAALQATR TGAWSGPSSTSMASAAAP	84.30%	2.2	0.153	0	0	1	0	2	3,058.48	324	354
GuHCl	PPE (PPE)	Rv1801	41460	99.10%	2	2	2	0.01%	16.30%	YVAWMSATAVHAELAGAQ AR	95.00%	2.44	0.266	0	0	1	0	2	3,754.99	53	90
GuHCl	linB (linB)	Rv2579	33710	99.70%	2	2	3	0.01%	12.00%	NLPIDGEPAEVVALVNEYR	95.00%	4.16	0.674	0	1	0	0	2	2,099.28	211	229
GuHCl	linB (linB)	Rv2579	33710	99.70%	2	2	3	0.01%	12.00%	SPQGEPMALEHNIFVER	95.00%	3.14	0.554	0	2	0	0	2	1,971.14	157	173
GuHCl	pks4 (pks4)	Rv1181	168064	99.80%	2	2	3	0.01%	3.29%	ELAELPGEPPRL GPVWVFSGQGSQWPGMG	89.30%	1.27	0.0956	2	0	0	0	2	1,321.61	786	797
GuHCl	pks4 (pks4)	Rv1181	168064	99.80%	2	2	3	0.01%	3.29%	TELLVAEPVFAATVAAMEPV IAR	95.00%	1.42	0.328	0	0	1	0	2	4,190.51	33	72
GuHCl	narH (narH)	Rv1162	62020	100.00%	3	3	3	0.01%	13.80%	PRICEHCLNPSCVASCPSG AMYK	95.00%	1.84	0.203	0	0	1	0	2	2,538.88	182	204

GuHCl	narH (narH)	Rv1162	62020	100.00%	3	3	3	0.01%	13.80%	TEDGIVLVDQDRCRGWRM CVSGCPYK	95.00%	1.35	0.321	0	0	1	0	2	3,130.43	206	231
GuHCl	narH (narH)	Rv1162	62020	100.00%	3	3	3	0.01%	13.80%	YEERYVIPTSYAGELPAAAM TDDMGCSL	87.60%	2.2	0.177	0	0	1	0	2	3,087.51	456	483
GuHCl	hypo protein Rv3237c	Rv3237c	17216	99.70%	2	2	4	0.01%	25.60%	AGDVLIVIGTEDGIAGVEQI VEKG	95.00%	1.91	0.21	0	1	0	0	2	2,383.61	137	160
GuHCl	hypo protein Rv3237c	Rv3237c	17216	99.70%	2	2	4	0.01%	25.60%	LTDEEAEAVAQILGAPR	95.00%	4.03	0.532	0	3	0	0	2	1,783.98	54	70
GuHCl	glnE (glnE)	Rv2221c	109089	99.90%	2	2	2	0.01%	5.84%	GYGSDADVMFVCEPATGV DDARAVKWSTSIAERVR	95.00%	1.84	0.247	0	0	1	0	2	3,832.06	726	760
GuHCl	glnE (glnE)	Rv2221c	109089	99.90%	2	2	2	0.01%	5.84%	PDDDRAPAAIAVIGMGRLG GAELGVGSDADVMFVCEPA TGVDDAR	95.00%	1.26	0.27	0	0	1	0	2	4,627.11	703	747
GuHCl	hypo protein Rv1130	Rv1130	57732	98.80%	2	2	2	0.01%	7.98%	DSYTKQHSAEYQSQAPIDL ACR	95.00%	1.6	0.311	0	1	0	0	2	2,583.80	303	324
GuHCl	hypo protein Rv1130	Rv1130	57732	98.80%	2	2	2	0.01%	7.98%	VPDQDTKVRFFRVFCWCPV L	79.80%	1.37	0.232	0	0	1	0	2	2,527.82	1	20
GuHCl	groES	Rv3418c	10754	100.00%	4	4	9	0.03%	47.00%	DVLAVVSK	95.00%	1.75	0.436	2	0	0	0	2	830.8884	93	100
GuHCl	groES (groES)	Rv3418c	10754	100.00%	4	4	9	0.03%	47.00%	ILVQANEAETTTASGLVIPDT AK	95.00%	1.9	0.536	0	1	0	0	2	2,343.63	13	35
GuHCl	groES	Rv3418c	10754	100.00%	4	4	9	0.03%	47.00%	IPLDVAEGDTVIYSK	95.00%	4.24	0.49	0	4	0	0	2	1,620.89	59	73
GuHCl	groES	Rv3418c	10754	100.00%	4	4	9	0.03%	47.00%	RIPLDVAEGDTVIYSK	95.00%	3.33	0.566	0	2	0	0	2	1,777.09	58	73
GuHCl	groEL1 (groEL1)	Rv3417c	55859	100.00%	9	9	18	0.05%	33.40%	AADAVSEALLASATPVSGK	95.00%	2.17	0.385	0	1	0	0	2	1,758.95	122	140
GuHCl	groEL1 (groEL1)	Rv3417c	55859	100.00%	9	9	18	0.05%	33.40%	AAVEEGIVPGGGASLIHQA R	95.00%	2.5	0.392	0	4	0	0	2	1,933.08	403	422
GuHCl	groEL1 (groEL1)	Rv3417c	55859	100.00%	9	9	18	0.05%	33.40%	AFGGPTVTNDGVTVAR	94.60%	1.5	0.267	0	1	0	0	2	1,562.56	42	57
GuHCl	groEL1 (groEL1)	Rv3417c	55859	100.00%	9	9	18	0.05%	33.40%	AFLEDLAVVTGGQVVNPDA GMVLR	82.10%	1.73	0.222	0	1	0	0	2	2,488.68	285	308
GuHCl	groEL1 (groEL1)	Rv3417c	55859	100.00%	9	9	18	0.05%	33.40%	DEQIGDLVGEAMSK	85.70%	1.61	0.256	0	1	0	0	2	1,508.65	153	166
GuHCl	groEL1 (groEL1)	Rv3417c	55859	100.00%	9	9	18	0.05%	33.40%	EIELEDPFEDLGAQLVK	95.00%	2.55	0.315	0	2	0	0	2	1,946.24	58	74
GuHCl	groEL1 (groEL1)	Rv3417c	55859	100.00%	9	9	18	0.05%	33.40%	LAGGVAVIK	95.00%	2.03	0.0469	1	0	0	0	2	828.0184	370	378
GuHCl	groEL1 (groEL1)	Rv3417c	55859	100.00%	9	9	18	0.05%	33.40%	LVAAGVNPIALGVGIGK	95.00%	3.2	0.51	0	1	0	0	2	1,549.89	105	121
GuHCl	groEL1 (groEL1)	Rv3417c	55859	100.00%	9	9	18	0.05%	33.40%	RVVVSKDDTVIVDGGGTAE AVAN	95.00%	2.44	0.274	0	5	0	0	2	2,273.21	320	342
GuHCl	groEL1 (groEL1)	Rv3417c	55859	100.00%	9	9	18	0.05%	33.40%	TNDVAGDGTTTATILAQALI K	95.00%	2.43	0.375	0	2	0	0	2	2,075.41	80	100
GuHCl	hypo protein Rv1151c	Rv1151c	25643	99.50%	2	2	2	0.01%	18.60%	CARCGVPYTDALPEMPEPAI EVEPPVCDCGGLIRPDIVW FGEPL	95.00%	1.39	0.256	0	0	1	0	2	4,899.55	112	155
GuHCl	hypo protein Rv1151c	Rv1151c	25643	99.50%	2	2	2	0.01%	18.60%	CGVPYTDALPEMPEPAIEVE PPVCDCGGLIR	91.50%	2.06	0.199	0	0	1	0	2	3,414.81	115	145
GuHCl	nirA (nirA)	Rv2391	62981	99.70%	2	2	2	0.01%	4.44%	FTPYQKLVIL	95.00%	1.35	0.196	1	0	0	0	2	1,222.61	383	392
GuHCl	nirA (nirA)	Rv2391	62981	99.70%	2	2	2	0.01%	4.44%	WRRNLMACSGIEFCK FNVVTFQAEDEIGGICAALG	95.00%	2.07	0.216	0	1	0	0	2	1,886.31	418	432
GuHCl	hypo protein Rv2455c	Rv2455c	69134	99.10%	2	2	2	0.01%	7.81%	AAYGGALGVTSTSGPGISL K	84.90%	1.89	0.135	0	0	1	0	2	3,844.37	298	337
GuHCl	hypo protein Rv2455c	Rv2455c	69134	99.10%	2	2	2	0.01%	7.81%	VVAPELNLGQL	94.70%	1.13	0.211	1	0	0	0	2	1,153.31	579	589
GuHCl	mmpL9 (mmpL9)	Rv2339	104598	97.90%	2	2	2	0.01%	5.61%	MADMLTMSDEMLVAIN	89.50%	1.61	0.167	1	0	0	0	2	1,818.23	513	528
GuHCl	mmpL9 (mmpL9)	Rv2339	104598	97.90%	2	2	2	0.01%	5.61%	TMRTMMLTMRSTISGVQD QMADMQDHATAMGQAFDT AK	82.10%	1.72	0.219	0	0	1	0	2	4,242.85	641	678
GuHCl	fbpB (fbpB)	Rv1886c	34564	99.70%	2	2	6	0.02%	12.00%	AVKPTGSAAIGLSMAGSSA MIL	95.00%	4.15	0.5	0	1	0	0	2	2,066.58	154	175
GuHCl	fbpB (fbpB)	Rv1886c	34564	99.70%	2	2	6	0.02%	12.00%	PGLPVEYLQVPSPSMGR	95.00%	4.02	0.575	0	5	0	0	2	1,844.10	44	60
GuHCl	hypo protein Rv3786c	Rv3786c	44927	99.30%	2	2	2	0.01%	14.30%	ELLYR	95.00%	1.3	0.0766	1	0	0	0	2	693.9184	75	79
GuHCl	hypo protein Rv3786c	Rv3786c	44927	99.30%	2	2	2	0.01%	14.30%	MADFCRPDWVMMVDADW LVETDIDLR	87.20%	2.35	0.173	0	0	1	0	2	3,247.69	80	105
GuHCl	hypo protein Rv3786c	Rv3786c	44927	99.30%	2	2	2	0.01%	14.30%	MCPMVSRWDDPEYPDLIPV MGTAEALR	95.00%	1.35	0.295	0	0	1	0	2	3,213.70	119	145
SDS	hypo protein Rv3107c	Rv3107c	56503	100.00%	3	3	3	0.01%	9.30%	GVIAETFETACTWDGFDTL	95.00%	1.9	0.319	0	1	0	0	2	2,148.33	392	410
SDS	hypo protein Rv3107c	Rv3107c	56503	100.00%	3	3	3	0.01%	9.30%	GVIAETFETACTWDGFDTL HAAVTDAARTAIWKVCGTG VVTCR	95.00%	1.82	0.228	0	0	1	0	2	4,687.98	392	434
SDS	hypo protein Rv3107c	Rv3107c	56503	100.00%	3	3	3	0.01%	9.30%	YPANCR	95.00%	1.24	0.149	1	0	0	0	2	794.8963	299	304
SDS	PE_PGRS	Rv1450c	107389	97.60%	2	2	2	0.01%	4.29%	GGDGAFGGMSANATNPGE NGPNGN	87.80%	2.29	0.169	0	1	0	0	2	2,164.18	859	882
SDS	PE_PGRS	Rv1450c	107389	97.60%	2	2	2	0.01%	4.29%	GGDGGNAGNAGAGGPGG TGSTAGKAGPAGSILH	83.70%	1.82	0.168	0	0	1	0	2	2,649.80	996	1028
SDS	aspB (aspB)	Rv3565	41031	99.50%	2	2	2	0.01%	12.40%	GCEVVEIPCGPQTRFQPTA QMLAEIDPPLR	92.20%	2.35	0.175	0	0	1	0	2	3,368.80	136	165
SDS	aspB (aspB)	Rv3565	41031	99.50%	2	2	2	0.01%	12.40%	VAMASPGYPCYRNILSALG CEVVEIPCGPQTR	95.00%	2.58	0.247	0	0	1	0	2	3,412.86	118	149
SDS	groEL2 (groEL2)	Rv0440	56709	100.00%	29	33	177	0.70%	53.30%	AASIAGLFLTTEAVVADKPE K	95.00%	3.58	0.58	0	1	0	0	2	2,132.51	504	524
SDS	groEL2 (groEL2)	Rv0440	56709	100.00%	29	33	177	0.70%	53.30%	AQTGVYEDLLAAGVADPVK	95.00%	2.89	0.415	0	1	0	0	2	1,918.07	477	495
SDS	groEL2 (groEL2)	Rv0440	56709	100.00%	29	33	177	0.70%	53.30%	ASVPGGGDMGGMDF	95.00%	1.27	0.219	1	0	0	0	2	1,314.42	527	540
SDS	groEL2 (groEL2)	Rv0440	56709	100.00%	29	33	177	0.70%	53.30%	DETTIVEGAGDTDAIAGR	95.00%	3.74	0.364	0	5	0	0	2	1,791.88	326	343
SDS	groEL2 (groEL2)	Rv0440	56709	100.00%	29	33	177	0.70%	53.30%	DLLPLLEK	95.00%	1.67	0.262	20	0	0	0	2	941.3384	230	237

SDS	groEL2 (groEL2)	Rv0440	56709	100.00%	29	33	177	0.70%	53.30%	EIELEDPYEK	95.00%	2.83	0.331	6	0	0	0	2	1,265.44	58	67
SDS	groEL2 (groEL2)	Rv0440	56709	100.00%	29	33	177	0.70%	53.30%	EQIAATAAISAGDQSIGDLI AEAMDK	95.00%	4.67	0.672	0	17	0	0	2	2,591.04	141	166
SDS	groEL2 (groEL2)	Rv0440	56709	100.00%	29	33	177	0.70%	53.30%	FLTTEAVVADKPEK	95.00%	2.82	0.444	0	1	0	0	2	1,548.74	511	524
SDS	groEL2 (groEL2)	Rv0440	56709	100.00%	29	33	177	0.70%	53.30%	GLNALADAVK	95.00%	2.07	0.49	5	0	0	0	2	972.1784	18	27
SDS	groEL2 (groEL2)	Rv0440	56709	100.00%	29	33	177	0.70%	53.30%	GLNAQTGVYEDLLAAGVAD PVK	95.00%	3.91	0.537	0	3	0	0	2	2,202.42	474	495
SDS	groEL2 (groEL2)	Rv0440	56709	100.00%	29	33	177	0.70%	53.30%	GYISGYFVTDPER	95.00%	3.72	0.441	0	1	0	0	2	1,504.69	196	208
SDS	groEL2 (groEL2)	Rv0440	56709	100.00%	29	33	177	0.70%	53.30%	GYISGYFVTDPERQEAVLED PYILLVSSK	95.00%	3.26	0.487	0	2	1	0	2	3,290.79	196	224
SDS	groEL2 (groEL2)	Rv0440	56709	100.00%	29	33	177	0.70%	53.30%	IGAELVK	95.00%	1.5	0.1	1	0	0	0	2	729.9184	68	74
SDS	groEL2 (groEL2)	Rv0440	56709	100.00%	29	33	177	0.70%	53.30%	KTDDVAGDGTTTATVLAQA LVR	95.00%	6.1	0.616	0	44	1	0	2	2,204.35	79	100
SDS	groEL2 (groEL2)	Rv0440	56709	100.00%	29	33	177	0.70%	53.30%	KVVVTKDETTIVEGAGDTD AIAGR	95.00%	5.8	0.665	0	17	0	0	2	2,446.59	320	343
SDS	groEL2 (groEL2)	Rv0440	56709	100.00%	29	33	177	0.70%	53.30%	KWGAPTITNDGVSIAK	95.00%	4.38	0.448	0	2	0	0	2	1,658.95	42	57
SDS	groEL2 (groEL2)	Rv0440	56709	100.00%	29	33	177	0.70%	53.30%	LEGDEATGANIVK	95.00%	1.59	0.266	1	0	0	0	2	1,317.45	429	441
SDS	groEL2 (groEL2)	Rv0440	56709	100.00%	29	33	177	0.70%	53.30%	LQAAPTLDELK	95.00%	1.3	0.26	1	0	0	0	2	1,199.50	418	428
SDS	groEL2 (groEL2)	Rv0440	56709	100.00%	29	33	177	0.70%	53.30%	NAQTGVYEDLLAAGVADPV K	95.00%	2.5	0.25	0	1	0	0	2	2,032.17	476	495
SDS	groEL2 (groEL2)	Rv0440	56709	100.00%	29	33	177	0.70%	53.30%	NVAAGANPLGLK	95.00%	2.69	0.433	5	1	0	0	2	1,125.35	105	116
SDS	groEL2 (groEL2)	Rv0440	56709	100.00%	29	33	177	0.70%	53.30%	NVVLEK	95.00%	2.05	0.209	2	0	0	0	2	701.7584	36	41
SDS	groEL2 (groEL2)	Rv0440	56709	100.00%	29	33	177	0.70%	53.30%	QEAVLEDPYILLVSSK	95.00%	3.93	0.547	0	7	0	0	2	1,805.12	209	224
SDS	groEL2 (groEL2)	Rv0440	56709	100.00%	29	33	177	0.70%	53.30%	QIAFNSGLEPGVVAEK	95.00%	3.96	0.643	1	9	0	0	2	1,659.82	450	465
SDS	groEL2 (groEL2)	Rv0440	56709	100.00%	29	33	177	0.70%	53.30%	TDDVAGDGTTTATVLAQAL VR	95.00%	5.02	0.513	0	10	0	0	2	2,076.15	80	100
SDS	groEL2 (groEL2)	Rv0440	56709	100.00%	29	33	177	0.70%	53.30%	VALEAPLK	95.00%	1.44	0.453	4	0	0	0	2	841.0684	442	449
SDS	groEL2 (groEL2)	Rv0440	56709	100.00%	29	33	177	0.70%	53.30%	VTETLLK	95.00%	1.17	0.0794	1	0	0	0	2	803.9884	126	132
SDS	groEL2 (groEL2)	Rv0440	56709	100.00%	29	33	177	0.70%	53.30%	VTLGPK	95.00%	1.5	0.192	1	0	0	0	2	614.7584	28	33
SDS	groEL2 (groEL2)	Rv0440	56709	100.00%	29	33	177	0.70%	53.30%	VVVTKDETTIVEGAGDTDAI AGR	95.00%	4.33	0.59	0	2	0	0	2	2,318.39	321	343
SDS	groEL2 (groEL2)	Rv0440	56709	100.00%	29	33	177	0.70%	53.30%	WGAPTITNDGVSIAK	95.00%	3.89	0.584	0	2	0	0	2	1,530.75	43	57
SDS	desA2	Rv1094	31342	99.70%	2	2	3	0.01%	12.00%	LQNVADAGIFGKPQLR	95.00%	3.35	0.561	0	2	0	0	2	1,728.07	236	251
SDS	desA2	Rv1094	31342	99.70%	2	2	3	0.01%	12.00%	NLAAQIEEPILAGLIDR	95.00%	3.82	0.476	0	1	0	0	2	1,837.33	168	184
SDS	clpP2 (clpP2)	Rv2460c	23491	99.00%	2	2	2	0.01%	26.60%	ILTAEEAKDYGIIDTVLEYR	95.00%	2.69	0.48	0	1	0	0	2	2,313.80	188	207
SDS	clpP2 (clpP2)	Rv2460c	23491	99.00%	2	2	2	0.01%	26.60%	MAIYDTMQYVRADIQTVCL GQAASAAAVLLAAGTPGK	82.90%	1.05	0.345	0	0	1	0	2	3,771.40	87	123
SDS	hypo protein Rv1201c	Rv1201c	32591	99.60%	2	2	2	0.01%	11.00%	LAVSDVPVELAALIGR	93.00%	1.94	0.177	0	1	0	0	2	1,623.92	42	57
SDS	hypo protein Rv1201c	Rv1201c	32591	99.60%	2	2	2	0.01%	11.00%	TVIGSLDDVAADPYDAYLR	95.00%	2.23	0.385	0	1	0	0	2	2,055.35	70	88
SDS	hspX (hspX)	Rv2031c	16210	100.00%	17	22	402	1.58%	77.10%	AELPGVDPDK	95.00%	1.22	0.172	1	0	0	0	2	1,041.16	55	64
SDS	hspX (hspX)	Rv2031c	16210	100.00%	17	22	402	1.58%	77.10%	AELPGVDPDKDVDIMVR	95.00%	4.35	0.654	9	57	3	0	2	1,870.10	55	71
SDS	hspX (hspX)	Rv2031c	16210	100.00%	17	22	402	1.58%	77.10%	AELPGVDPDKDVDIMVRD GQLTIK	95.00%	2.7	0.494	0	2	0	0	2	2,626.05	55	78
SDS	hspX (hspX)	Rv2031c	16210	100.00%	17	22	402	1.58%	77.10%	DGQLTIK	95.00%	1.32	0.228	5	0	0	0	2	774.9684	72	78
SDS	hspX (hspX)	Rv2031c	16210	100.00%	17	22	402	1.58%	77.10%	FAAFPSFAGLRPTFDTR	95.00%	2.03	0.285	0	1	0	0	2	1,902.33	22	38
SDS	hspX (hspX)	Rv2031c	16210	100.00%	17	22	402	1.58%	77.10%	GILTVSVAVSEGKPTEK	95.00%	4.34	0.615	35	200	0	0	2	1,715.89	120	136
SDS	hspX (hspX)	Rv2031c	16210	100.00%	17	22	402	1.58%	77.10%	HIQIR	95.00%	1.47	0.121	5	0	0	0	2	666.8184	137	141
SDS	hspX (hspX)	Rv2031c	16210	100.00%	17	22	402	1.58%	77.10%	LEDEMK	95.00%	1.08	0.188	3	0	0	0	2	780.9178	42	47
SDS	hspX (hspX)	Rv2031c	16210	100.00%	17	22	402	1.58%	77.10%	PGVDPDKDVDIMVR	95.00%	3.2	0.542	0	3	0	0	2	1,572.72	58	71
SDS	hspX (hspX)	Rv2031c	16210	100.00%	17	22	402	1.58%	77.10%	PVGADEDDIK	95.00%	2.04	0.379	8	0	0	0	2	1,059.14	105	114
SDS	hspX (hspX)	Rv2031c	16210	100.00%	17	22	402	1.58%	77.10%	SEFAYGSFVR	95.00%	2.71	0.565	9	4	0	0	2	1,163.28	91	100
SDS	hspX (hspX)	Rv2031c	16210	100.00%	17	22	402	1.58%	77.10%	SLFPEFSEL	95.00%	1.6	0.245	1	0	0	0	2	1,069.30	13	21
SDS	hspX (hspX)	Rv2031c	16210	100.00%	17	22	402	1.58%	77.10%	SLFPEFSELFAAFPSFAGLR	95.00%	3.83	0.572	0	4	0	0	2	2,234.79	13	32
SDS	hspX (hspX)	Rv2031c	16210	100.00%	17	22	402	1.58%	77.10%	SLFPEFSELFAAFPSFAGLRP TFDTR	95.00%	2.87	0.58	0	0	24	0	2	2,952.61	13	38
SDS	hspX (hspX)	Rv2031c	16210	100.00%	17	22	402	1.58%	77.10%	TVSLPVGADEDDIK	95.00%	3.28	0.523	7	14	0	0	2	1,459.59	101	114
SDS	hspX (hspX)	Rv2031c	16210	100.00%	17	22	402	1.58%	77.10%	TVSLPVGADEDDIKATYDK	95.00%	4.98	0.539	0	6	0	0	2	2,038.27	101	119
SDS	hspX (hspX)	Rv2031c	16210	100.00%	17	22	402	1.58%	77.10%	YEVR	94.70%	1.04	0.164	1	0	0	0	2	566.5884	51	54
SDS	hypo protein Rv2030c	Rv2030c	74880	99.70%	2	2	3	0.01%	7.93%	ARYACFDHACADDGQAYGF AAAFGAGPSCER	95.00%	2.01	0.22	0	0	1	0	2	3,413.67	403	433
SDS	hypo protein Rv2030c	Rv2030c	74880	99.70%	2	2	3	0.01%	7.93%	VAIDAPGGVPTHEVLAELVG DAR	95.00%	2.54	0.312	0	2	0	0	2	2,287.41	246	268
SDS	drrC (drrC)	Rv2938	29576	99.70%	2	2	2	0.01%	17.00%	GFAMGGPVLSPMIGMLVWT AGICVVCAVPLAIGYRR	95.00%	1.22	0.302	0	0	1	0	2	3,811.61	237	272
SDS	drrC (drrC)	Rv2938	29576	99.70%	2	2	2	0.01%	17.00%	QPVSYAIAAMRGFAMGGPV LSPMIGMLVWTAGICVVCA VPLAIGYR	95.00%	1.94	0.258	0	0	1	0	2	4,930.89	226	271
SDS	hypo protein Rv1006	Rv1006	61280	100.00%	3	3	4	0.02%	8.29%	GPLNAMGSPAIPSTAQEIPN PLR	95.00%	3.43	0.54	0	2	0	0	2	2,348.82	35	57
SDS	hypo protein Rv1006	Rv1006	61280	100.00%	3	3	4	0.02%	8.29%	TLPAAVDLK	95.00%	1.33	0.106	1	0	0	0	2	928.1684	468	476
SDS	hypo protein Rv1006	Rv1006	61280	100.00%	3	3	4	0.02%	8.29%	YSVEAQPGSQALAGK	91.90%	1.72	0.25	0	1	0	0	2	1,506.61	418	432
SDS	hypo protein Rv3778c	Rv3778c	41833	100.00%	4	4	4	0.02%	16.60%	DAAREAVADLVNADPGGVV LGADR	95.00%	2.07	0.305	0	0	1	0	2	2,352.45	64	87
SDS	hypo protein Rv3778c	Rv3778c	41833	100.00%	4	4	4	0.02%	16.60%	DPSVMNSFGSVSTNPYATG PAR	91.70%	1.5	0.315	0	1	0	0	2	2,272.40	232	253

SDS	hypo protein Rv3778c	Rv3778c	41833	100.00%	4	4	4	0.02%	16.60%	EAVADLVNADPGGVVLGADR	95.00%	3.66	0.459	0	1	0	0	2	1,938.99	68	87
SDS	hypo protein Rv3778c	Rv3778c	41833	100.00%	4	4	4	0.02%	16.60%	LVHDVGALVVVDHSAAAPYR	95.00%	5.8	0.439	0	1	0	0	2	2,090.14	182	201
SDS	pknH (pknH)	Rv1266c	66735	98.80%	2	2	2	0.01%	14.20%	GSPPTITVIDTQEGAEGWECQRAMSVANNVVVDVNACGYR	80.20%	0.907	0.264	0	0	1	0	2	4,310.34	568	607
SDS	pknH (pknH)	Rv1266c	66735	98.80%	2	2	2	0.01%	14.20%	PGPTGQPSWAPNSGPMPASGPTPTPQYYQGGGWGAP	95.00%	2.08	0.19	0	0	1	0	2	4,888.22	349	397
SDS	kasB (kasB)	Rv2246	46371	100.00%	3	3	5	0.02%	10.50%	AVSPLTVQK	95.00%	1.44	0.369	1	0	0	0	2	943.0384	156	164
SDS	kasB (kasB)	Rv2246	46371	100.00%	3	3	5	0.02%	10.50%	IMGASITSDGFHMVAPDPNGER	95.00%	2.79	0.345	0	1	0	0	2	2,319.60	286	307
SDS	kasB (kasB)	Rv2246	46371	100.00%	3	3	5	0.02%	10.50%	TLDDPFVEEFDLPVR	95.00%	3.87	0.524	0	3	0	0	2	1,793.00	65	79
SDS	kasA (kasA)	Rv2245	43266	99.10%	2	2	3	0.01%	10.10%	QIVMGDADVAVCGGVEGPIEALPIAAFSMMRAMSTRN	95.00%	2.48	0.331	0	0	2	0	2	3,842.38	184	220
SDS	kasA (kasA)	Rv2245	43266	99.10%	2	2	3	0.01%	10.10%	QIVMGDADVAVCGGVEGPIEALPIAAFSMMRAMSTRNDEPER	85.90%	1.4	0.266	0	0	1	0	2	4,453.00	184	225
SDS	tkt (tkt)	Rv1449c	75539	99.70%	2	2	2	0.01%	3.43%	AHGAAL	95.00%	0.887	0.188	1	0	0	0	2	539.6084	282	287
SDS	tkt (tkt)	Rv1449c	75539	99.70%	2	2	2	0.01%	3.43%	NGSGPVGLILTRQGVPVL	95.00%	2.06	0.24	0	0	1	0	2	1,778.05	541	558
SDS	sucB (sucB)	Rv2215	57070	99.70%	2	2	4	0.02%	9.40%	EGVNLTFLPFFAK	95.00%	2.24	0.272	0	1	0	0	2	1,483.84	368	380
SDS	sucB (sucB)	Rv2215	57070	99.70%	2	2	4	0.02%	9.40%	PAPEPPPVQPTSGAPAGGDAKPVLMPELGESVTEGTVIR	95.00%	3.09	0.539	0	0	3	0	2	3,867.21	102	140
SDS	tsf (tsf)	Rv2889c	28737	100.00%	6	7	13	0.05%	32.50%	NAEFQTLADQVVAAAAAAK	95.00%	4.16	0.587	0	1	0	0	2	1,890.00	82	100
SDS	tsf (tsf)	Rv2889c	28737	100.00%	6	7	13	0.05%	32.50%	NAEFQTLADQVVAAAAAAK	95.00%	3.13	0.582	0	3	3	0	2	2,699.95	82	108
SDS	tsf (tsf)	Rv2889c	28737	100.00%	6	7	13	0.05%	32.50%	PADVDALK	95.00%	3.66	0.494	0	1	0	0	2	1,936.16	25	42
SDS	tsf (tsf)	Rv2889c	28737	100.00%	6	7	13	0.05%	32.50%	NALAETDGDFDKAVEALR	95.00%	2.93	0.541	0	2	0	0	2	1,586.76	152	166
SDS	tsf (tsf)	Rv2889c	28737	100.00%	6	7	13	0.05%	32.50%	SADLPPAVGVLVEYR	95.00%	1.65	0.212	1	0	0	0	2	1,260.41	116	127
SDS	tsf (tsf)	Rv2889c	28737	100.00%	6	7	13	0.05%	32.50%	TVEQAIAELSAK	95.00%	2.47	0.378	0	2	0	0	2	1,865.02	187	202
SDS	tyrS (tyrS)	Rv1689	46313	98.50%	2	2	2	0.01%	9.67%	YLSRDDVPEDIVASERAGGATGMIGDPRDVGERSL	95.00%	2.47	0.378	0	2	0	0	2	1,865.02	187	202
SDS	tyrS (tyrS)	Rv1689	46313	98.50%	2	2	2	0.01%	9.67%	NEADTVAEWTER	92.70%	1.54	0.249	0	0	1	0	2	3,278.43	71	101
SDS	accD4 (accD4)	Rv3799c	56151	99.70%	2	2	2	0.01%	7.54%	HFSVNVMLARINGRPVGVFSHDQTVFGGTVGEMFGRK	83.30%	1.15	0	1	0	0	0	2	1,190.32	143	152
SDS	accD4	Rv3799c	56151	99.70%	2	2	2	0.01%	7.54%	SYGGAYAVMGSK	95.00%	0.994	0.302	1	0	0	0	2	1,191.36	407	418
SDS	hypo protein Rv1230c	Rv1230c	41368	99.40%	2	2	2	0.01%	20.40%	ICIGSHGPPVGAPFPPTAPPPFLPAAPPPPDPLAGPPGDAGLAPPAPAPAG	90.30%	2.04	0.162	0	0	1	0	2	4,883.86	359	411
SDS	hypo protein Rv1230c	Rv1230c	41368	99.40%	2	2	2	0.01%	20.40%	PGTWARYATDGDDDGVADPQNLFDSTLAAAR	95.00%	1	0.458	0	1	0	0	2	3,268.54	205	235
SDS	fbpC2 (fbpC2)	Rv0129c	36754	100.00%	3	3	8	0.03%	16.50%	ADIQHVLNGATPPAAPAAPAA	95.00%	3.68	0.576	0	4	0	0	2	1,954.16	320	340
SDS	fbpC2	Rv0129c	36754	100.00%	3	3	8	0.03%	16.50%	PGLPVEYLQVPSASMGR	95.00%	3.1	0.532	0	1	0	0	2	1,818.06	50	66
SDS	fbpC2	Rv0129c	36754	100.00%	3	3	8	0.03%	16.50%	VQFQGGGPHAVYLLDGLRQRVGAEVEVTADTDAAMTADGLVVPGVGAFAACMAGLRK	95.00%	5.06	0.565	0	3	0	0	2	1,928.16	70	87
SDS	hisH (hisH)	Rv1602	21368	99.00%	2	2	2	0.01%	18.90%	VGAEVEVTADTDAAMTADGLVVPGVGAFAACMAGLR	95.00%	1.46	0.262	0	0	1	0	2	3,882.13	24	62
SDS	hisH (hisH)	Rv1602	21368	99.00%	2	2	2	0.01%	18.90%	LVVPGVGAFAACMAGLR	83.30%	1.03	0.313	0	0	1	0	2	3,469.63	26	61
SDS	hypo protein Rv3038c	Rv3038c	36031	98.60%	2	2	2	0.01%	10.40%	IAHNAGAVEVQTATEEFTAAML	94.40%	2.17	0.239	0	1	0	0	2	2,291.39	247	268
SDS	hypo protein Rv3038c	Rv3038c	36031	98.60%	2	2	2	0.01%	10.40%	VVTNATKLPGLR	80.00%	1.09	0.219	1	0	0	0	2	1,269.51	201	212
SDS	hypo protein Rv3894c	Rv3894c	153685	99.90%	3	3	3	0.01%	4.08%	GKTTTLMALMCSAATMYTPERVTFFCIGGATMAQIGSLMCSAATMYTPERVTFFCIG	79.80%	1.05	0.304	0	0	1	0	2	4,095.94	890	927
SDS	hypo protein Rv3894c	Rv3894c	153685	99.90%	3	3	3	0.01%	4.08%	GATMAQIGSLPHVTDIVSPK	86.20%	1.76	0.235	0	0	1	0	2	4,220.88	899	937
SDS	hypo protein Rv3894c	Rv3894c	153685	99.90%	3	3	3	0.01%	4.08%	VTFFCIGGATMAQIGSLPHVTDIVSPKDAEGIERIL	95.00%	2.23	0.367	0	0	1	0	2	3,875.55	911	946
SDS	hypo protein Rv2205c	Rv2205c	48556	99.70%	2	2	2	0.01%	13.10%	GGSACTDGGKGMIAELGGLDAARR	95.00%	1.95	0.254	0	0	1	0	2	2,335.75	257	280
SDS	hypo protein Rv2205c	Rv2205c	48556	99.70%	2	2	2	0.01%	13.10%	VRPGPSWCGDGIGCAWAWAGPWGRWFATRAMCPTRTCVR	95.00%	1.46	0.389	0	0	1	0	2	4,466.97	1	39
SDS	fusA2 (fusA2)	Rv0120c	75612	99.90%	2	2	2	0.01%	7.98%	AQMDKGVHAGYPVVDIRVTL	93.50%	2.68	0.124	0	1	0	0	2	2,170.38	569	588
SDS	fusA2 (fusA2)	Rv0120c	75612	99.90%	2	2	2	0.01%	7.98%	PGSVTEGTTVCDFDEAEIRQQRSVGLAVASLAYDGIK	94.90%	1.13	0.312	0	0	1	0	2	3,956.26	53	89
SDS	wag31 (wag31)	Rv2145c	28260	100.00%	4	4	11	0.04%	21.20%	ANAEQILGEAR	90.30%	1.16	0.127	1	0	0	0	2	1,172.31	134	144
SDS	wag31 (wag31)	Rv2145c	28260	100.00%	4	4	11	0.04%	21.20%	GSAAPVDSNADAGGFDQFNR	95.00%	4.33	0.58	0	7	0	0	2	1,997.04	238	257
SDS	wag31 (wag31)	Rv2145c	28260	100.00%	4	4	11	0.04%	21.20%	TYLESQLEELGQR	95.00%	5.03	0.556	0	2	0	0	2	1,566.75	225	237
SDS	wag31 (wag31)	Rv2145c	28260	100.00%	4	4	11	0.04%	21.20%	VLSLAQDTADR	95.00%	2.65	0.354	0	1	0	0	2	1,189.33	106	116
SDS	hypo protein Rv0248c	Rv0248c	70663	99.70%	2	2	3	0.01%	7.12%	ALSSRPAVSAEAIDAAAQQALSPFEGPK	95.00%	1.85	0.438	0	1	0	0	2	2,784.16	449	476
SDS	hypo protein Rv0248c	Rv0248c	70663	99.70%	2	2	3	0.01%	7.12%	FIVFEAPAVVLATGGIGKAAEGMEESQIEDLIDDPEQ	95.00%	2.77	0.546	0	2	0	0	2	1,790.14	197	214
SDS	hsdM (hsdM)	Rv2756c	60067	99.20%	2	2	2	0.01%	14.30%	YQGYGVFVVPVSARWKVYDPCCGSGGMFVQTEKFIYEHDGDPKDVSIYGQESIEETWR	86.80%	1.28	0.264	0	0	1	0	2	3,974.19	67	101
SDS	hsdM (hsdM)	Rv2756c	60067	99.20%	2	2	2	0.01%	14.30%	AVGIDLGTTNSVVSVLEGGDPVVANSEGSR	95.00%	1.68	0.32	0	0	1	0	2	4,818.16	224	265
SDS	dnaK (dnaK)	Rv0350	66813	100.00%	6	6	10	0.04%	17.90%	DPVVVANSEGSR	95.00%	4.62	0.576	0	1	0	0	2	2,999.96	4	34

SDS	dnaK (dnaK)	Rv0350	66813	100.00%	6	6	10	0.04%	17.90%	DAEAYLGEDITDAVITTPAY FNDAQR	95.00%	2.04	0.323	0	1	0	0	2	2,861.16	102	127
SDS	dnaK (dnaK)	Rv0350	66813	100.00%	6	6	10	0.04%	17.90%	GVNPDEVVAVGAALQAGVL K	95.00%	4.08	0.633	0	3	0	0	2	1,907.96	338	357
SDS	dnaK (dnaK)	Rv0350	66813	100.00%	6	6	10	0.04%	17.90%	IVNEPTAAALAYGLDKGEK	95.00%	2.6	0.447	0	2	0	0	2	1,961.33	144	162
SDS	dnaK (dnaK)	Rv0350	66813	100.00%	6	6	10	0.04%	17.90%	KGVNPDEVVAVGAALQAG VLK	95.00%	4.7	0.507	0	1	0	0	2	2,036.16	337	357
SDS	dnaK (dnaK)	Rv0350	66813	100.00%	6	6	10	0.04%	17.90%	LLGSFELTGIPPAPR	95.00%	3.15	0.485	0	2	0	0	2	1,569.04	427	441
SDS	hypo protein Rv2191	Rv2191	69130	99.70%	2	2	4	0.02%	6.05%	IVGVSNDAAGLASPVRSAG PWAAWAATAR	95.00%	1.79	0.438	0	0	3	0	2	2,824.08	562	590
SDS	hypo protein Rv2191	Rv2191	69130	99.70%	2	2	4	0.02%	6.05%	RKRMTEMVML	95.00%	1.01	0.141	1	0	0	0	2	1,311.69	285	294
SDS	hypo protein Rv2326c	Rv2326c	72818	100.00%	3	3	6	0.02%	6.17%	GRGTPTVVVSSLIGGFVFG AAMVGML	94.10%	2.02	0.241	0	0	1	0	2	2,556.91	122	147
SDS	hypo protein Rv2326c	Rv2326c	72818	100.00%	3	3	6	0.02%	6.17%	QRLALAAALAR	92.20%	1.37	0.09	1	0	0	0	2	1,154.52	380	390
SDS	hypo protein Rv2326c	Rv2326c	72818	100.00%	3	3	6	0.02%	6.17%	RGRGTPTVVVSSLIGGFVF GAAMVGMLAAMVR	95.00%	1.96	0.316	0	0	4	0	2	3,241.74	121	152
SDS	hypo protein Rv1075c	Rv1075c	32829	99.70%	2	2	2	0.01%	14.60%	GVCGQVDAMFVVGPPPDA AVIMIGANDITALN	95.00%	1.3	0.32	0	0	1	0	2	3,157.53	124	155
SDS	hypo protein Rv1075c	Rv1075c	32829	99.70%	2	2	2	0.01%	14.60%	IRLSTKAIVGATSKGVCGQ VDAMFVVGPPPDAAVIMIG AN	95.00%	2.09	0.216	0	0	1	0	2	3,988.59	110	149
SDS	hypo protein Rv1256c	Rv1256c	44562	99.70%	2	2	3	0.01%	18.30%	GVPEEDWTQFDGWTQAIV AANAVDGATTGALDAVGS MMAYFTGLIER	95.00%	1.55	0.445	0	0	2	0	2	4,935.34	157	203
SDS	hypo protein Rv1256c	Rv1256c	44562	99.70%	2	2	3	0.01%	18.30%	MTSVMSHEFQLATAETWPN PWPMYRAL	95.00%	1.72	0.264	0	0	1	0	2	3,244.65	1	27
SDS	hypo protein Rv3099c	Rv3099c	30458	100.00%	4	5	9	0.04%	30.70%	AAVQRPSSDDELGGPASPL VLDEIAATMGFVVGK	95.00%	5.91	0.658	0	1	3	0	2	3,415.80	161	194
SDS	hypo protein Rv3099c	Rv3099c	30458	100.00%	4	5	9	0.04%	30.70%	HLGTESGVGLLER	95.00%	2.87	0.529	0	1	0	0	2	1,368.52	96	108
SDS	hypo protein Rv3099c	Rv3099c	30458	100.00%	4	5	9	0.04%	30.70%	VLASLSDDEWNAPTTTPSG PDSYGR	95.00%	2.52	0.479	0	1	0	0	2	2,637.83	119	143
SDS	hypo protein Rv3099c	Rv3099c	30458	100.00%	4	5	9	0.04%	30.70%	VVDDFGGPAPTATIR	95.00%	3.63	0.461	0	3	0	0	2	1,516.66	226	240
SDS	nadC (nadC)	Rv1596	29933	99.70%	2	2	7	0.03%	13.70%	VPPGEALMTLEAQTR	95.00%	3.04	0.544	0	3	0	0	2	1,629.84	84	98
SDS	nadC (nadC)	Rv1596	29933	99.70%	2	2	7	0.03%	13.70%	YGPDVTTLATVPASATTTAS LVTR	95.00%	3.34	0.507	0	4	0	0	2	2,394.60	25	48
SDS	nadA (nadA)	Rv1594	37390	99.70%	2	2	2	0.01%	14.30%	MITPAALLRCL	95.00%	1.55	0.215	1	0	0	0	2	1,218.70	305	315
SDS	nadA (nadA)	Rv1594	37390	99.70%	2	2	2	0.01%	14.30%	PGAVVVSYVNTTAAVKALT DICCTSSNAVDVVASIDPD R	95.00%	2.26	0.276	0	0	1	0	2	3,924.07	133	171
SDS	helZ (helZ)	Rv2101	111615	100.00%	3	3	3	0.01%	6.02%	RTACPSGLELDADGAYR	95.00%	2.41	0.164	0	0	1	0	2	1,796.06	349	365
SDS	helZ (helZ)	Rv2101	111615	100.00%	3	3	3	0.01%	6.02%	SGGMRLWAEDSDLLVK TTEQASLYQAVVADMMEKI	94.30%	1.14	0.359	0	1	0	0	2	1,794.13	11	26
SDS	helZ (helZ)	Rv2101	111615	100.00%	3	3	3	0.01%	6.02%	ENTEGIERR	95.00%	1.7	0.242	0	0	1	0	2	3,246.53	774	801
SDS	hypo protein Rv2100	Rv2100	58918	99.60%	2	2	2	0.01%	3.82%	PSNLKCACRK	93.60%	1.02	0	1	0	0	0	2	1,191.48	438	447
SDS	hypo protein Rv2100	Rv2100	58918	99.60%	2	2	2	0.01%	3.82%	VDARGLMMPRR	95.00%	1.42	0.00739	1	0	0	0	2	1,302.64	508	518
SDS	fadD19 (fadD19)	Rv3515c	59691	99.50%	2	2	2	0.01%	6.20%	AEHAIDAVPDRVAVICGDE QLTYAQLEDK	95.00%	1.82	0.343	0	0	1	0	2	3,242.48	11	39
SDS	fadD19 (fadD19)	Rv3515c	59691	99.50%	2	2	2	0.01%	6.20%	VTSGG	91.40%	0.764	0.22	1	0	0	0	2	420.3684	544	548
SDS	bioF (bioF)	Rv1569	40009	99.70%	2	2	4	0.02%	14.00%	HPAVIDGGVQALRIWGAGA TGSR	95.00%	2.17	0.244	0	3	0	0	2	2,290.53	56	78
SDS	bioF (bioF)	Rv1569	40009	99.70%	2	2	4	0.02%	14.00%	MCGVAAVPDSAMVSVILGE PESAVAAAAACL	95.00%	1.41	0.296	0	0	1	0	2	3,079.32	306	336
SDS	hypo protein Rv0036c	Rv0036c	27528	100.00%	5	5	5	0.02%	24.90%	AESDDLDALVAHLPADR	95.00%	3.65	0.42	0	1	0	0	2	1,809.01	13	29
SDS	hypo protein Rv0036c	Rv0036c	27528	100.00%	5	5	5	0.02%	24.90%	ALSTLDVNAVGEDAQR	95.00%	2.93	0.495	0	1	0	0	2	1,659.73	226	241
SDS	hypo protein Rv0036c	Rv0036c	27528	100.00%	5	5	5	0.02%	24.90%	LHEELLAVPDGR	95.00%	3.46	0.532	0	1	0	0	2	1,349.54	106	117
SDS	hypo protein Rv0036c	Rv0036c	27528	100.00%	5	5	5	0.02%	24.90%	SIAHL	95.00%	0.906	0.164	1	0	0	0	2	540.6784	162	166
SDS	hypo protein Rv0036c	Rv0036c	27528	100.00%	5	5	5	0.02%	24.90%	WLTIAQAFAGPPGR	95.00%	3.91	0.487	0	1	0	0	2	1,485.80	242	255
SDS	hypo protein Rv3371	Rv3371	48830	99.90%	2	2	4	0.02%	4.93%	MIHALAR	95.00%	1.35	0.0382	3	0	0	0	2	828.0778	318	324
SDS	hypo protein Rv3371	Rv3371	48830	99.90%	2	2	4	0.02%	4.93%	PSAARRGRPSVPTAR	94.80%	1.12	0.312	0	0	1	0	2	1,579.80	429	443
SDS	hypo protein Rv2625c	Rv2625c	41463	99.70%	2	2	4	0.02%	6.36%	DAIPLGRIAGFVVNVHWSV L	95.00%	2.58	0.213	0	3	0	0	2	2,164.46	3	22
SDS	hypo protein Rv2625c	Rv2625c	41463	99.70%	2	2	4	0.02%	6.36%	FSDAG	95.00%	1.15	0.0767	1	0	0	0	2	496.5284	389	393
SDS	sucD (sucD)	Rv0952	31212	99.70%	2	2	3	0.01%	10.90%	LPVFGSVAEAMEK	95.00%	2.38	0.5	0	1	0	0	2	1,394.57	63	75
SDS	sucD (sucD)	Rv0952	31212	99.70%	2	2	3	0.01%	10.90%	TNVSKPVVGYVAGFTAPEG K	95.00%	4.94	0.513	0	2	0	0	2	2,022.13	236	255
SDS	hypo protein Rv1128c	Rv1128c	49258	99.70%	2	2	2	0.01%	8.20%	AAPGMCNPEQK	95.00%	1.22	0.306	1	0	0	0	2	1,162.27	229	239
SDS	hypo protein Rv1128c	Rv1128c	49258	99.70%	2	2	2	0.01%	8.20%	IMLYAKDSGCSAPGCDVPG YYCEVHH	95.00%	2.4	0.299	0	0	1	0	2	2,888.15	354	379
SDS	phoS1	Rv0934	38194	100.00%	4	4	12	0.05%	23.80%	ASFLDQVHFQPLPPAVVK	95.00%	3.18	0.535	0	2	0	0	2	1,994.23	346	363

SDS	phoS1	Rv0934	38194	100.00%	4	4	12	0.05%	23.80%	SDGSGDTFLFTQYLSK	95.00%	3.11	0.544	0	2	0	0	2	1,767.06	186	201
SDS	phoS1 (phoS1)	Rv0934	38194	100.00%	4	4	12	0.05%	23.80%	TPANQAISMIDGPAPDGYPI INYEYAIVNNR	95.00%	4.18	0.678	0	5	0	0	2	3,395.97	292	322
SDS	phoS1 (phoS1)	Rv0934	38194	100.00%	4	4	12	0.05%	23.80%	TWDDPQIAALNPGVNLPGT AVVPL	95.00%	2.01	0.313	0	3	0	0	2	2,459.75	160	183
SDS	hypo protein Rv0284	Rv0284	145124	99.70%	2	2	3	0.01%	1.50%	DYGGGQLR	95.00%	1.39	0.137	2	0	0	0	2	865.9684	861	868
SDS	hypo protein Rv0284	Rv0284	145124	99.70%	2	2	3	0.01%	1.50%	FASQTLDVGKIK	95.00%	1.59	0.0315	1	0	0	0	2	1,307.60	630	641
SDS	lprG (lprG)	Rv1411c	24530	100.00%	3	3	6	0.02%	14.40%	IPGLSLK	95.00%	2.08	0.374	2	0	0	0	2	728.0684	68	74
SDS	lprG (lprG)	Rv1411c	24530	100.00%	3	3	6	0.02%	14.40%	TLSGDLTTNPTAATGNVK	95.00%	2.8	0.583	0	3	0	0	2	1,761.95	75	92
SDS	lprG (lprG)	Rv1411c	24530	100.00%	3	3	6	0.02%	14.40%	VLANFADAK	91.50%	1.08	0.119	1	0	0	0	2	949.1284	145	153
SDS	hypo protein Rv2713	Rv2713	50736	99.80%	2	2	2	0.01%	8.55%	AGLEVQGRGRIFVDDRFQT K	95.00%	1.93	0.258	0	0	1	0	2	2,293.59	282	301
SDS	hypo protein Rv2713	Rv2713	50736	99.80%	2	2	2	0.01%	8.55%	IVGHGRFIDPHTILVEDQAR	88.60%	1.84	0.206	0	1	0	0	2	2,274.56	111	130
SDS	35kd_ag (35kd_ag)	Rv2744c	29240	100.00%	9	9	21	0.08%	62.60%	ATEYNNAAEAFAAQLVTAE QSVEDLK	95.00%	4.29	0.564	0	4	0	0	2	2,784.90	95	120
SDS	35kd_ag (35kd_ag)	Rv2744c	29240	100.00%	9	9	21	0.08%	62.60%	GEALPAGGTTATPRPATETS GGAIAEQPYGQ	95.00%	1.46	0.462	0	1	0	0	2	2,958.16	240	270
SDS	35kd_ag (35kd_ag)	Rv2744c	29240	100.00%	9	9	21	0.08%	62.60%	MLEVEQAGIQMAGHSR	95.00%	2.92	0.504	0	2	0	0	2	1,757.93	215	230
SDS	35kd_ag (35kd_ag)	Rv2744c	29240	100.00%	9	9	21	0.08%	62.60%	QALTLADQATAAGDAAK	95.00%	3.03	0.444	0	1	0	0	2	1,616.83	78	94
SDS	35kd_ag (35kd_ag)	Rv2744c	29240	100.00%	9	9	21	0.08%	62.60%	SMSELAAPGNTPSLDEVR	95.00%	2.4	0.363	0	3	0	0	2	1,891.08	173	190
SDS	35kd_ag (35kd_ag)	Rv2744c	29240	100.00%	9	9	21	0.08%	62.60%	THQALTQQAAQVIGNQR	95.00%	5.53	0.684	0	3	0	0	2	1,864.88	40	56
SDS	35kd_ag (35kd_ag)	Rv2744c	29240	100.00%	9	9	21	0.08%	62.60%	TLHDQALSAAAQAK	95.00%	2.61	0.542	0	2	0	0	2	1,425.60	121	134
SDS	35kd_ag (35kd_ag)	Rv2744c	29240	100.00%	9	9	21	0.08%	62.60%	VQIQQAIEEAQR	95.00%	4.72	0.503	0	2	0	0	2	1,413.45	28	39
SDS	(35kd_ag)	Rv2744c	29240	100.00%	9	9	21	0.08%	62.60%	YANAIGSAELAESSVQGR	95.00%	3.89	0.496	0	3	0	0	2	1,823.95	197	214
SDS	sucA (sucA)	Rv1248c	133954	99.70%	2	2	3	0.01%	3.38%	SVPRLMPGQGAIIGVGAME YPAEFQGASEER	95.00%	2.31	0.232	0	0	2	0	2	3,249.65	251	281
SDS	sucA (sucA)	Rv1248c	133954	99.70%	2	2	3	0.01%	3.38%	TNPGTIGTVHSVPRLMPGQ GAIIGVGAMEYPAEFQGAS EER	95.00%	1.69	0.221	0	0	1	0	2	4,243.64	241	281
SDS	tpx (tpx)	Rv1932	16878	98.80%	2	2	6	0.02%	35.20%	TFDERAAASGATVLCVSKD LPFAQK	80.40%	2.02	0.161	0	0	1	0	2	2,626.99	66	90
SDS	tpx (tpx)	Rv1932	16878	98.80%	2	2	6	0.02%	35.20%	TVGELPAVGSPAPAFTLTGG DLGVISSDQFRGK	95.00%	1.99	0.255	0	0	5	0	2	3,246.67	13	45
SDS	hypo protein Rv2258c	Rv2258c	37496	99.70%	2	2	3	0.01%	11.60%	ALRPGGVLLMVDIK	94.80%	1.98	0.252	0	2	0	0	2	1,498.96	264	277
SDS	hypo protein Rv2258c	Rv2258c	37496	99.70%	2	2	3	0.01%	11.60%	SGADVADFGCGSGRAVKL MAQAFGASR	95.00%	1.84	0.261	0	0	1	0	2	2,702.01	171	197
SDS	hypo protein Rv2005c	Rv2005c	30965	100.00%	3	3	4	0.02%	25.40%	GLLGSVSSSLVR	95.00%	2.18	0.464	0	1	0	0	2	1,175.38	127	138
SDS	hypo protein Rv2005c	Rv2005c	30965	100.00%	3	3	4	0.02%	25.40%	IPLTVVHVVNADVATWPPM PYPETWGVWQEDEGR	95.00%	1.6	0.281	0	0	1	0	2	3,907.00	36	69
SDS	hypo protein Rv2005c	Rv2005c	30965	100.00%	3	3	4	0.02%	25.40%	SELVFSTPVPTMVEISNEAE MVVLGSSGR	95.00%	3.58	0.628	0	2	0	0	2	3,099.29	93	121
SDS	hypo protein Rv2223c	Rv2223c	55060	100.00%	3	3	3	0.01%	19.20%	AQDCVDRMGFSFLANIGTA SVARDMDMVR	82.00%	1.68	0.227	0	0	1	0	2	3,281.69	190	218
SDS	hypo protein Rv2223c	Rv2223c	55060	100.00%	3	3	3	0.01%	19.20%	DADGHYSNDQDAFNAVRC VDAPTPADPAAWVAADQR	94.70%	1.86	0.192	0	0	1	0	2	3,831.84	374	409
SDS	hypo protein Rv2223c	Rv2223c	55060	100.00%	3	3	3	0.01%	19.20%	SCREFIADTSEIRTARCATV SVPVDYDQPGGTQAK	95.00%	1.39	0.381	0	0	1	0	2	3,845.11	57	91
SDS	mutA (mutA)	Rv1492	64695	99.70%	2	3	3	0.01%	7.15%	RDRPAPPIEEVVAVASR	95.00%	1.51	0.259	0	0	1	0	2	1,863.01	192	208
SDS	mutA (mutA)	Rv1492	64695	99.70%	2	3	3	0.01%	7.15%	YRDEVADIVQAARAAGVSR VYLAGPEK	95.00%	2.39	0.355	0	1	1	0	2	2,906.18	557	583
SDS	moxR (moxR)	Rv1479	40744	100.00%	17	20	44	0.17%	57.80%	AVPQQGHSVPPVMQAAAA ASGR	81.90%	1.66	0.188	0	1	0	0	2	2,147.13	356	377
SDS	moxR	Rv1479	40744	100.00%	17	20	44	0.17%	57.80%	DYVPQDVIEVIPDVLR	95.00%	3.92	0.466	0	4	0	0	2	1,984.24	306	322
SDS	moxR	Rv1479	40744	100.00%	17	20	44	0.17%	57.80%	GHVLLEGVPGVAK	95.00%	3.73	0.512	2	3	0	0	2	1,276.38	72	84
SDS	moxR	Rv1479	40744	100.00%	17	20	44	0.17%	57.80%	GRDYVIPQDVIEVIPDVLR	95.00%	4.48	0.637	0	3	0	0	2	2,197.49	304	322
SDS	moxR	Rv1479	40744	100.00%	17	20	44	0.17%	57.80%	IIVGQDQLVER	95.00%	2.04	0.174	2	2	0	0	2	1,270.41	53	63
SDS	moxR	Rv1479	40744	100.00%	17	20	44	0.17%	57.80%	IQFTPDLVPTDIIGTR	95.00%	3.13	0.508	0	3	0	0	2	1,787.18	102	117
SDS	moxR	Rv1479	40744	100.00%	17	20	44	0.17%	57.80%	LVLTYDALADEISPEIVINR	95.00%	4.54	0.539	0	2	0	0	2	2,245.72	325	344
SDS	moxR	Rv1479	40744	100.00%	17	20	44	0.17%	57.80%	MLVGLLSK	95.00%	1.17	0.241	3	0	0	0	2	877.2178	64	71
SDS	moxR	Rv1479	40744	100.00%	17	20	44	0.17%	57.80%	QGREEFDTELGPVVAN	95.00%	2.55	0.41	0	1	0	0	2	1,761.76	121	136
SDS	moxR (moxR)	Rv1479	40744	100.00%	17	20	44	0.17%	57.80%	SASPAHEAPPGGAEGLAAE VHTLER	95.00%	2.4	0.436	0	0	1	0	2	2,455.54	21	45
SDS	moxR	Rv1479	40744	100.00%	17	20	44	0.17%	57.80%	SLALVR	95.00%	1.32	0	2	0	0	0	2	658.8484	298	303
SDS	moxR (moxR)	Rv1479	40744	100.00%	17	20	44	0.17%	57.80%	TFPMPSPFLVMATQNPIEHE GVYPLPEAQR	95.00%	2.67	0.475	0	3	2	0	2	3,430.87	169	198
SDS	moxR	Rv1479	40744	100.00%	17	20	44	0.17%	57.80%	TLAVETFAR	95.00%	1.6	0.297	2	0	0	0	2	1,008.15	85	93
SDS	moxR	Rv1479	40744	100.00%	17	20	44	0.17%	57.80%	VLQTVALPQVN	95.00%	1.89	0.188	3	0	0	0	2	1,182.23	345	355
SDS	moxR	Rv1479	40744	100.00%	17	20	44	0.17%	57.80%	VLQTVALPQVNAVPQQGH	95.00%	2.24	0.381	0	2	0	0	2	1,899.85	345	362
SDS	moxR (moxR)	Rv1479	40744	100.00%	17	20	44	0.17%	57.80%	VLQTVALPQVNAVPQQGHS VPPVMQAAAAASGR	95.00%	2.7	0.385	0	0	2	0	2	3,310.34	345	377
SDS	moxR	Rv1479	40744	100.00%	17	20	44	0.17%	57.80%	VQSALLEVMQER	95.00%	3.75	0.445	0	1	0	0	2	1,419.52	150	161
SDS	ctaB (ctaB)	Rv1451	33489	100.00%	4	4	6	0.02%	25.00%	QDYQVAGVPMLPAVATERQ VTKQILIYTWL	95.00%	2.18	0.34	0	0	1	0	2	3,433.93	203	232
SDS	ctaB (ctaB)	Rv1451	33489	100.00%	4	4	6	0.02%	25.00%	QVTK	88.40%	1.01	0.0243	3	0	0	0	2	475.4884	221	224
SDS	ctaB (ctaB)	Rv1451	33489	100.00%	4	4	6	0.02%	25.00%	RRTSQNVVWGGAAGCMPV MIGWSAITGTIAWPAL	91.00%	1.59	0.273	0	0	1	0	2	3,560.08	147	180
SDS	ctaB (ctaB)	Rv1451	33489	100.00%	4	4	6	0.02%	25.00%	TSQNVVWGGAAGCMPVMI GWSAITGTIAWPALAMFAII	95.00%	1.13	0.32	0	0	1	0	2	4,910.76	149	193
SDS	accD5	Rv3280	59337	97.10%	2	2	2	0.01%	8.21%	FFWTPPH	95.00%	1.04	0	1	0	0	0	2	866.9384	443	450
SDS	accD5 (accD5)	Rv3280	59337	97.10%	2	2	2	0.01%	8.21%	DMGCDVNL GDGVVTGYGTIDGRDVCIF SQDATVFGGSLGEVYGEK	86.00%	1.44	0.25	0	0	1	0	2	3,770.96	95	131

SDS	hypo protein Rv2864c	Rv2864c	62996	99.70%	2	2	2	0.01%	4.98%	TIPNYGGFDLGVVPMSRAF ASSCN	95.00%	2.25	0.251	0	0	1	0	2	2,520.85	396	419
SDS	hypo protein Rv2864c	Rv2864c	62996	99.70%	2	2	2	0.01%	4.98%	TIPNYGGFDLGVVPMSRAF ASSCNTTFAEL	95.00%	2.46	0.349	0	0	1	0	2	3,254.71	396	425
SDS	hypo protein Rv0892	Rv0892	55021	100.00%	3	3	3	0.01%	9.29%	CPTVAVVGAGMSGMCVAIT LLSAGITDVCIYEK	95.00%	1.8	0.389	0	0	1	0	2	3,361.85	5	37
SDS	hypo protein Rv0892	Rv0892	55021	100.00%	3	3	3	0.01%	9.29%	MTGRCPTVAVVGAGMSGM CVAITLLSAGITDVCIYEK	92.60%	1.82	0.237	0	0	1	0	2	3,823.40	1	37
SDS	hypo protein Rv0892	Rv0892	55021	100.00%	3	3	3	0.01%	9.29%	MTGRCPTVAVVGAGMSGM CVAITLLSAGITDVCIYEKA DDVGGTWR	95.00%	1.8	0.419	0	0	1	0	2	4,820.42	1	46
SDS	fmu (fmu)	Rv1407	48476	99.60%	2	2	2	0.01%	6.56%	DERSWVGELAPDAQNDPI GH	93.90%	1.59	0.246	0	0	1	0	2	2,207.27	140	159
SDS	fmu (fmu)	Rv1407	48476	99.60%	2	2	2	0.01%	6.56%	PVELLRVDGR	95.00%	1.5	0.213	1	0	0	0	2	1,154.33	318	327
SDS	fadA (fadA)	Rv0859	42397	99.60%	2	2	3	0.01%	10.40%	AAVLASGMPVTSGGVQLNR IVATATSGADPVIMLTGPTP	94.20%	1.93	0.248	0	1	0	0	2	1,829.00	72	90
SDS	fadA (fadA)	Rv0859	42397	99.60%	2	2	3	0.01%	10.40%	ATR	95.00%	4.91	0.423	0	2	0	0	2	2,257.62	286	308
SDS	atpG (atpG)	Rv1309	33874	99.30%	2	2	4	0.02%	9.51%	LESARPYAFEITR	89.20%	1.45	0.276	0	1	0	0	2	1,553.88	40	52
SDS	atpG (atpG)	Rv1309	33874	99.30%	2	2	4	0.02%	9.51%	VYAALLESAASELASR	95.00%	3.28	0.351	0	3	0	0	2	1,651.93	242	257
SDS	hypo protein Rv3389c	Rv3389c	30278	100.00%	3	3	7	0.03%	15.20%	LHAPLPAAGK	95.00%	1.47	0.266	2	0	0	0	2	975.2484	87	96
SDS	hypo protein Rv3389c	Rv3389c	30278	100.00%	3	3	7	0.03%	15.20%	LSVVTEVADIQDKGEGK	95.00%	3.91	0.488	0	3	0	0	2	1,788.89	97	113
SDS	hypo protein Rv3389c	Rv3389c	30278	100.00%	3	3	7	0.03%	15.20%	VGTFNPAALLHGSQGIR	95.00%	2.64	0.479	0	2	0	0	2	1,739.00	70	86
SDS	hypo protein Rv0194	Rv0194	129234	99.40%	2	2	2	0.01%	1.59%	EVPALAGINLR	89.80%	1.12	-0.0759	1	0	0	0	2	1,153.42	955	965
SDS	hypo protein Rv0194	Rv0194	129234	99.40%	2	2	2	0.01%	1.59%	PPGPDPLR PGIVRASYAMPDVHWGYG	95.00%	0.793	0.324	1	0	0	0	2	849.0484	610	617
SDS	hypo protein Rv2631	Rv2631	27844	98.80%	2	2	2	0.01%	30.60%	FPIGGVAATDVDNDGVVSP GGVGFDISCGVR	79.70%	0.924	0.348	0	0	1	0	2	4,937.22	18	66
SDS	hypo protein Rv2631	Rv2631	27844	98.80%	2	2	2	0.01%	30.60%	QMEQAMGRYGIAVPDRQL ACVPVHSPDGQAYL	95.00%	1.47	0.446	0	0	1	0	2	3,573.92	218	249
SDS	katG (katG)	Rv1908c	80556	100.00%	3	3	4	0.02%	9.59%	DIEEVMTTSQPWWPADYG HYGPLFIRMAWH	95.00%	1.74	0.266	0	0	1	0	2	3,636.09	79	108
SDS	katG (katG)	Rv1908c	80556	100.00%	3	3	4	0.02%	9.59%	SSYGTGTGKDAITSGIEVV WTN	82.70%	1.35	0.327	0	1	0	0	2	2,244.38	302	323
SDS	katG (katG)	Rv1908c	80556	100.00%	3	3	4	0.02%	9.59%	TDASQEQTDVESFAVLEPK	95.00%	3.9	0.49	0	2	0	0	2	2,095.10	572	590
SDS	hypo protein Rv2623	Rv2623	31633	100.00%	7	7	37	0.15%	30.00%	AGPPTVHSEIVPAAAVPTLV DMSK	95.00%	5.07	0.651	0	10	0	0	2	2,404.61	86	109
SDS	hypo protein Rv2623	Rv2623	31633	100.00%	7	7	37	0.15%	30.00%	AGPPTVHSEIVPAAAVPTLV DMSKDAVLMVVGCL	95.00%	2.37	0.223	0	0	1	0	2	3,476.83	86	119
SDS	hypo protein Rv2623	Rv2623	31633	100.00%	7	7	37	0.15%	30.00%	GGYAGMLVGSVGETVAQLA R	95.00%	4.07	0.634	0	16	0	0	2	1,953.10	267	286
SDS	hypo protein Rv2623	Rv2623	31633	100.00%	7	7	37	0.15%	30.00%	HLIDDALK	95.00%	1.52	0.107	4	0	0	0	2	925.1984	70	77
SDS	hypo protein Rv2623	Rv2623	31633	100.00%	7	7	37	0.15%	30.00%	SEEAQLVVVGSR	95.00%	1.91	0.347	2	0	0	0	2	1,274.22	253	264
SDS	hypo protein Rv2623	Rv2623	31633	100.00%	7	7	37	0.15%	30.00%	TPVIVAR	95.00%	1.92	0.341	3	0	0	0	2	755.8584	287	293
SDS	hypo protein Rv2623	Rv2623	31633	100.00%	7	7	37	0.15%	30.00%	YPNVAITR	95.00%	1.42	0.132	1	0	0	0	2	934.0884	231	238
SDS	hypo protein Rv2622	Rv2622	29524	99.00%	2	2	2	0.01%	14.70%	VTDAADTGLSDASADVVIG EAMLTMQGN	95.00%	1.98	0.232	0	0	1	0	2	2,769.94	105	132
SDS	hypo protein Rv2622	Rv2622	29524	99.00%	2	2	2	0.01%	14.70%	VVTASMALLQPR	83.70%	1.36	0.316	0	1	0	0	2	1,286.52	206	217
SDS	hypo protein Rv1038c	Rv1197,Rv2347c	10959	100.00%	3	5	13	0.05%	41.80%	DANNYEQQEQASQQILSS	95.00%	3.46	0.497	0	3	0	0	2	2,054.02	81	98
SDS	hypo protein Rv1038c	Rv1197,Rv2347c	10959	100.00%	3	5	13	0.05%	41.80%	FEVHAQTVEDEAR	95.00%	4.24	0.48	1	5	0	0	2	1,531.42	19	31
SDS	hypo protein Rv1038c	Rv1197,Rv2347c	10959	100.00%	3	5	13	0.05%	41.80%	NIVNMLHGVR	95.00%	2.45	0.361	2	2	0	0	2	1,153.31	66	75
SDS	fbpA (fbpA)	Rv3804c	35668	98.90%	2	2	3	0.01%	9.76%	FLEGFVR	81.70%	1.48	0.197	1	0	0	0	2	868.0384	271	277
SDS	fbpA (fbpA)	Rv3804c	35668	98.90%	2	2	3	0.01%	9.76%	NDPLLNVGK	95.00%	1.62	0.18	2	0	0	0	2	970.1584	234	242
SDS	fbpA (fbpA)	Rv3804c	35668	98.90%	2	2	3	0.01%	9.76%	PGLPVEYLQVPSPSMGR	95.00%	4.58	0.556	0	12	0	0	2	1,828.10	47	63
SDS	hypo protein Rv1907c	Rv1907c	24017	97.30%	2	2	2	0.01%	23.70%	GAEMCWMCDHPEATAEEY LDEVYGIML	85.30%	1.23	0.255	0	0	1	0	2	3,266.60	56	82
SDS	hypo protein Rv1907c	Rv1907c	24017	97.30%	2	2	2	0.01%	23.70%	LAGCWCLPGAICQTPRAW WSQARR	84.40%	1.87	0.17	0	0	1	0	2	2,732.16	21	44
SDS	hypo protein Rv2307c	Rv2307c	29653	99.50%	2	2	2	0.01%	16.40%	ALPVVAIVALVASGVIMFIW SQQR	95.00%	2.04	0.266	0	0	1	0	2	2,570.02	8	31
SDS	hypo protein Rv2307c	Rv2307c	29653	99.50%	2	2	2	0.01%	16.40%	RLIYFPSAGPVPSASSVLPA GR	92.40%	2.18	0.172	0	0	1	0	2	2,243.70	32	53
SDS	hypo protein Rv0148	Rv0148	29760	100.00%	4	4	5	0.02%	28.70%	AVANYDSVATEDGAANIIK	95.00%	3.94	0.514	0	2	0	0	2	1,923.09	66	84
SDS	hypo protein Rv0148	Rv0148	29760	100.00%	4	4	5	0.02%	28.70%	DGTGAGSAMADEVVAEIR	95.00%	1.72	0.349	0	1	0	0	2	1,765.81	43	60
SDS	hypo protein Rv0148	Rv0148	29760	100.00%	4	4	5	0.02%	28.70%	EYALTLAGEGASVVVNDLG GAR	95.00%	2.48	0.362	0	1	0	0	2	2,163.33	21	42
SDS	hypo protein Rv0148	Rv0148	29760	100.00%	4	4	5	0.02%	28.70%	VALFGNDGANFDKPPSVQD VAAR	95.00%	4.59	0.542	0	1	0	0	2	2,389.57	247	269
SDS	hypo protein Rv0142	Rv0142	33004	99.70%	2	2	2	0.01%	12.70%	CVAWGSGAEEFVDMAPAM LGAADDASDFVPL	95.00%	1.4	0.312	0	0	1	0	2	3,247.52	69	99
SDS	hypo protein Rv0142	Rv0142	33004	99.70%	2	2	2	0.01%	12.70%	CVAWGSGAEEFVDMAPAM LGAADDASDFVPLHPAVAA AH	95.00%	1.64	0.304	0	0	1	0	2	3,915.15	69	107
SDS	hypo protein Rv0111	Rv0111	74323	99.80%	2	2	2	0.01%	6.42%	FAARCAATVVLAGASWWLI EQPIRR	95.00%	1.86	0.288	0	0	1	0	2	2,787.29	374	398

SDS	hypo protein Rv0111	Rv0111	74323	99.80%	2	2	2	0.01%	6.42%	IVVAGAAVIVVASVAMEQR	91.70%	1.81	0.281	0	1	0	0	2	1,899.97	310	328
SDS	hypo protein Rv3874	Rv3874	10776	99.90%	2	2	2	0.01%	28.00%	QKQELDEISTNIR	95.00%	3.22	0.503	0	1	0	0	2	1,574.80	65	77
SDS	hypo protein Rv3874	Rv3874	10776	99.90%	2	2	2	0.01%	28.00%	TDAATLAQEAGNFER	95.00%	4.39	0.564	0	1	0	0	2	1,594.69	6	20
SDS	hypo protein Rv3877	Rv3877	53964	99.30%	2	2	2	0.01%	9.39%	VSPVVKRTLELIDGAMIAAI IPML	91.60%	1.73	0.251	0	1	0	0	2	2,568.34	473	496
SDS	hypo protein Rv3877	Rv3877	53964	99.30%	2	2	2	0.01%	9.39%	YAERWCAWALLAATVAIPT GLTAK	82.00%	1.19	0.226	0	0	1	0	2	2,578.12	414	437
SDS	pdhC (pdhC)	Rv2495c	41042	98.80%	2	2	2	0.01%	12.00%	LLVIALK	95.00%	1.67	0	1	0	0	0	2	770.1684	225	231
SDS	pdhC (pdhC)	Rv2495c	41042	98.80%	2	2	2	0.01%	12.00%	SFPVPDLGEGLQEVTVTCW SVAVGDDVEINQTLCSVET AK	80.40%	1.41	0.227	0	0	1	0	2	4,309.38	9	48
SDS	nrp (nrp)	Rv0101	269359	100.00%	4	4	4	0.02%	3.90%	AVLTAPAPTPVSIPQMLAAQ VARIPEAEAVCCGDASMTY R	93.80%	0.974	0.308	0	0	1	0	2	4,117.63	1482	1521
SDS	nrp (nrp)	Rv0101	269359	100.00%	4	4	4	0.02%	3.90%	DRGYVKAVR	89.10%	1.62	0.0782	1	0	0	0	2	1,064.19	303	311
SDS	nrp (nrp)	Rv0101	269359	100.00%	4	4	4	0.02%	3.90%	PIDPANPPPR VMLNAYGPTETTICAAISAP	95.00%	1.9	0.203	1	0	0	0	2	1,074.30	1571	1580
SDS	nrp (nrp)	Rv0101	269359	100.00%	4	4	4	0.02%	3.90%	LRPGSGMPPIGVPVSGAAL	95.00%	1.31	0.445	0	0	1	0	2	3,853.60	1764	1802
SDS	hypo protein Rv0831c	Rv0831c	30171	99.70%	2	2	5	0.02%	12.50%	FTPGGLVLTEWQGAAVYR	95.00%	3.95	0.594	0	2	0	0	2	1,966.19	163	180
SDS	hypo protein Rv0831c	Rv0831c	30171	99.70%	2	2	5	0.02%	12.50%	YGPGMGQALDPNYHLR	95.00%	2.78	0.442	0	3	0	0	2	1,790.09	192	207
SDS	hypo protein Rv2063c	Rv2063c	12394	98.60%	2	2	2	0.01%	29.50%	DTGTTTTSSSSIPVSSSSTE TTTAGRCLPGSPAR	81.20%	2.02	0.152	0	0	1	0	2	3,302.48	7	40
SDS	hypo protein Rv2063c	Rv2063c	12394	98.60%	2	2	2	0.01%	29.50%	ELDTGTTTTSSSSIPVSSSS TETTTAGRCL	93.60%	1.62	0.318	0	1	0	0	2	2,979.13	5	34
SDS	hypo protein Rv1179c	Rv1179c	100624	99.70%	2	2	3	0.01%	4.90%	CGTER	95.00%	0.687	0.32	1	0	0	0	2	565.5684	94	98
SDS	hypo protein Rv1179c	Rv1179c	100624	99.70%	2	2	3	0.01%	4.90%	GAVAAVSVVGASTATAVAS ANLGMLAGAGTAGAIVAAG VGL	95.00%	2.74	0.305	0	0	2	0	2	3,442.67	718	758
SDS	atpD (atpD)	Rv1310	53077	100.00%	11	14	31	0.12%	40.10%	DEQGQDVLLFIDNIFR	95.00%	3.21	0.537	0	4	0	0	2	1,923.24	254	269
SDS	atpD (atpD)	Rv1310	53077	100.00%	11	14	31	0.12%	40.10%	DTALVFGQMDEPPGTR GDFDHVPEQAFFLIGGLDD LAK	95.00%	2.02	0.377	0	1	0	0	2	1,750.91	223	238
SDS	atpD (atpD)	Rv1310	53077	100.00%	11	14	31	0.12%	40.10%	GIFPAVDPLASSSTILDPSV VGDEHYR	95.00%	3.63	0.589	0	5	0	0	2	2,405.82	456	477
SDS	atpD (atpD)	Rv1310	53077	100.00%	11	14	31	0.12%	40.10%	IALFGGAGVGK	95.00%	3.67	0.534	0	4	3	0	2	2,844.17	352	378
SDS	atpD (atpD)	Rv1310	53077	100.00%	11	14	31	0.12%	40.10%	KPPAFEELEPR MPSAVGYQPTLADEMGELQ ER	95.00%	1.96	0.209	1	0	0	0	2	990.2484	167	177
SDS	atpD (atpD)	Rv1310	53077	100.00%	11	14	31	0.12%	40.10%	NFGGTSVFAGVGER	95.00%	2.9	0.385	1	1	0	0	2	1,313.56	134	144
SDS	atpD (atpD)	Rv1310	53077	100.00%	11	14	31	0.12%	40.10%	SITSMQAVYVPADDYTDPA PATTFAHLDATTELSR	95.00%	3.32	0.523	0	2	0	0	2	1,398.42	191	204
SDS	atpD (atpD)	Rv1310	53077	100.00%	11	14	31	0.12%	40.10%	VTGPVVDVEFPR	95.00%	4.28	0.544	0	0	1	0	2	3,774.14	312	346
SDS	atpD (atpD)	Rv1310	53077	100.00%	11	14	31	0.12%	40.10%	VVDLLTPYVR	95.00%	3.07	0.522	2	1	0	0	2	1,315.29	26	37
SDS	atpD (atpD)	Rv1310	53077	100.00%	11	14	31	0.12%	40.10%	HPLDCPMCDK QQVGIYADEPFESYFSGNT	95.00%	1.37	0.206	1	0	0	0	2	1,175.35	154	163
SDS	nuoG (nuoG)	Rv3151	85374	99.70%	2	2	2	0.01%	5.71%	VQICPVGALTGTAYRFR	95.00%	0.819	0.25	1	0	0	0	2	1,246.44	111	120
SDS	nuoG (nuoG)	Rv3151	85374	99.70%	2	2	2	0.01%	5.71%	ALELQAPSVVHR	80.70%	1.97	0.145	0	0	1	0	2	3,987.37	192	227
SDS	atpA (atpA)	Rv1308	59271	100.00%	16	19	41	0.16%	41.30%	AMKEVAGSL	95.00%	2.96	0.395	2	3	0	0	2	1,320.42	131	142
SDS	atpA (atpA)	Rv1308	59271	100.00%	16	19	41	0.16%	41.30%	ASEEEILTEIR	95.00%	1.1	0	2	0	0	0	2	922.0778	385	393
SDS	atpA (atpA)	Rv1308	59271	100.00%	16	19	41	0.16%	41.30%	EAYPGDVFYLHSR	95.00%	1.62	0.0601	2	1	0	0	2	1,290.48	477	487
SDS	atpA (atpA)	Rv1308	59271	100.00%	16	19	41	0.16%	41.30%	ELEAFAAFASDLDAASK GFAATGGGSVVPDEHVEAL	95.00%	3.49	0.56	0	3	0	0	2	1,554.72	295	307
SDS	atpA (atpA)	Rv1308	59271	100.00%	16	19	41	0.16%	41.30%	DEDKLAK	95.00%	1.87	0.281	0	1	0	0	2	1,757.06	402	418
SDS	atpA (atpA)	Rv1308	59271	100.00%	16	19	41	0.16%	41.30%	LSDDLGGGSLTGLPIIETK	95.00%	4.36	0.582	0	3	0	0	2	2,613.75	510	535
SDS	atpA (atpA)	Rv1308	59271	100.00%	16	19	41	0.16%	41.30%	LTEEAADKLTEVIK	95.00%	4.01	0.497	0	3	0	0	2	1,887.40	315	333
SDS	atpA (atpA)	Rv1308	59271	100.00%	16	19	41	0.16%	41.30%	LVELLK	95.00%	4.66	0.399	0	3	0	0	2	1,560.85	492	505
SDS	atpA (atpA)	Rv1308	59271	100.00%	16	19	41	0.16%	41.30%	QGVKEPLQTGIK QPQSQPMPVEEQVVSIFLG	95.00%	1.9	0.157	3	0	0	0	2	714.9884	427	432
SDS	atpA (atpA)	Rv1308	59271	100.00%	16	19	41	0.16%	41.30%	TGGHLDSVPVEDVR	95.00%	1.59	0.287	0	1	0	0	2	1,298.51	143	154
SDS	atpA (atpA)	Rv1308	59271	100.00%	16	19	41	0.16%	41.30%	RALELQAPSVVHR	95.00%	3.71	0.572	0	0	3	0	2	3,592.61	433	465
SDS	atpA (atpA)	Rv1308	59271	100.00%	16	19	41	0.16%	41.30%	RTGEVLSVPVGDGFLGR	95.00%	2.7	0.373	0	1	0	0	2	1,476.62	130	142
SDS	atpA (atpA)	Rv1308	59271	100.00%	16	19	41	0.16%	41.30%	TGEVLSVPVGDGFLGR TLEEGGAMDYTTIVAAAAS	95.00%	3.89	0.391	0	2	0	0	2	1,759.93	93	109
SDS	atpA (atpA)	Rv1308	59271	100.00%	16	19	41	0.16%	41.30%	ESAGFK	95.00%	2.58	0.429	0	2	0	0	2	1,603.73	94	109
SDS	atpA (atpA)	Rv1308	59271	100.00%	16	19	41	0.16%	41.30%	VVNPLGQPIDGR	95.00%	4.3	0.644	0	3	0	0	2	2,507.78	222	246
SDS	atpA (atpA)	Rv1308	59271	100.00%	16	19	41	0.16%	41.30%	AEQQVASTLQTAHEQLK	95.00%	3.16	0.406	2	1	0	0	2	1,265.40	110	121
SDS	atpF (atpF)	Rv1306	18307	99.70%	2	2	5	0.02%	10.50%	AEQQVASTLQTAHEQLK	95.00%	4.07	0.543	0	3	0	0	2	1,882.91	118	134
SDS	atpF (atpF)	Rv1306	18307	99.70%	2	2	5	0.02%	10.50%	AEQQVASTLQTAHEQLKR GFAGVMETFDNTRPIVAAM	95.00%	2.3	0.497	0	2	0	0	2	2,039.11	118	135
SDS	(fadE23) fadE23	Rv3140	43327	100.00%	5	5	7	0.03%	24.40%	AVGIGRAALEEIR	91.90%	1.9	0.161	0	0	1	0	2	3,364.93	243	274
SDS	(fadE23) fadE23	Rv3140	43327	100.00%	5	5	7	0.03%	24.40%	ILDIFEGTQQIQQLVVAR	95.00%	2.46	0.235	0	1	0	0	2	2,072.39	373	390
SDS	(fadE23) fadE23	Rv3140	43327	100.00%	5	5	7	0.03%	24.40%	LLGLSSSELK	89.90%	1.11	0.121	1	0	0	0	2	1,047.41	392	401
SDS	(fadE23) fadE23	Rv3140	43327	100.00%	5	5	7	0.03%	24.40%	QAIIVKTHQGAAEMMRPIA R	95.00%	1.64	0.282	0	0	1	0	2	2,238.68	12	31
SDS	(fadE23)	Rv3140	43327	100.00%	5	5	7	0.03%	24.40%	TVELAGTTGYSEQSLLEK	95.00%	3.24	0.517	0	3	0	0	2	1,927.13	349	366
SDS	cysA2	Rv0815c,Rv3117	30997	100.00%	7	7	9	0.04%	36.50%	AANEDGTFKSDEELAK	95.00%	2.17	0.2	0	1	0	0	2	1,725.89	201	216
SDS	cysA2	Rv0815c,Rv3117	30997	100.00%	7	7	9	0.04%	36.50%	DFVDAQQFSK	95.00%	2.3	0.399	3	0	0	0	2	1,185.25	58	67
SDS	cysA2	Rv0815c,Rv3117	30997	100.00%	7	7	9	0.04%	36.50%	ILAPAHLPQEQSQR	95.00%	2.35	0.333	0	1	0	0	2	1,588.80	173	186
SDS	cysA2	Rv0815c,Rv3117	30997	100.00%	7	7	9	0.04%	36.50%	LYADAGLDNSK	95.00%	1.15	0.27	1	0	0	0	2	1,167.41	217	227
SDS	cysA2 (cysA2) cysA2	Rv0815c,Rv3117	30997	100.00%	7	7	9	0.04%	36.50%	NLIDVR NYDGSWTEYGSLVGAPIELGS	87.30%	1.53	0.105	1	0	0	0	2	729.8884	159	164
SDS	(cysA2) cysA2	Rv0815c,Rv3117	30997	100.00%	7	7	9	0.04%	36.50%	VVFVEVDEDTSAYDRDHIA GAIK	95.00%	1.87	0.404	0	1	0	0	2	2,216.43	257	277
SDS	hypo protein Rv0712	Rv0712	32693	99.50%	2	2	3	0.01%	16.10%	CPGAMVFCPTAGPVDLRD WR	88.70%	1.3	0.239	0	0	1	0	2	2,550.67	22	44
SDS	hypo protein Rv0712	Rv0712	32693	99.50%	2	2	3	0.01%	16.10%	VDLPGGSFRMGSTRFYPEE APIHTVTVR	92.50%	2.46	0.151	0	2	0	0	2	2,279.56	78	97
SDS	ctpA (ctpA)	Rv0092	78832	100.00%	3	3	3	0.01%	6.70%	DRGHVVAMVGDGIN	95.00%	1.26	0.156	1	0	0	0	2	1,440.46	627	640

										HAESLGETAVFVEVDGEPC											
SDS	ctpA (ctpA)	Rv0092	78832	100.00%	3	3	3	0.01%	6.70%	GVIAVADAVK	88.10%	1.85	0.242	0	0	1	0	2	2,913.89	541	569
SDS	ctpA (ctpA)	Rv0092	78832	100.00%	3	3	3	0.01%	6.70%	VEQAQAQK	95.00%	0.932	0.267	1	0	0	0	2	901.8484	341	348
	hypo protein									GGDAASAPPPGAGGPAPPQ											
SDS	Rv0064	Rv0064	107352	100.00%	3	3	3	0.01%	12.30%	AVPPPRTTQPPAAPPRGPDV	95.00%	1.71	0.317	0	0	1	0	2	4,354.73	907	953
	hypo protein									PPATVAEL											
SDS	Rv0064	Rv0064	107352	100.00%	3	3	3	0.01%	12.30%	VGYAPTLAESLDQVFGPGT	95.00%	1.71	0.317	0	0	1	0	2	4,354.73	907	953
	hypo protein									GRVATARGGDAASAPPPGA											
SDS	Rv0064	Rv0064	107352	100.00%	3	3	3	0.01%	12.30%	GGPAPPQAVPPPR	95.00%	1.3	0.383	0	0	1	0	2	4,806.23	881	931
	hypo protein									VYYGEVIAQADPDYAIVGG											
SDS	Rv0064	Rv0064	107352	100.00%	3	3	3	0.01%	12.30%	APGSAPREYDTDTSKYTYT	79.60%	1.27	0.247	0	0	1	0	2	4,819.18	493	539
	hypo protein									FFGASAAAIINDPTSSAQIS											
SDS	Rv2488c	Rv2488c	120802	100.00%	2	2	6	0.02%	3.25%	CARGDLN	95.00%	1.95	0.241	0	0	1	0	2	2,770.22	888	914
	hypo protein																				
SDS	Rv2488c	Rv2488c	120802	100.00%	2	2	6	0.02%	3.25%	LALVTEICHR	95.00%	1.53	0.0362	5	0	0	0	2	1,155.37	414	423
SDS	pckA (pckA)	Rv0211	67236	99.90%	2	2	3	0.01%	4.62%	TYICSAKEIDAGPTNNWMD PGEMRSIMK	95.00%	1.97	0.258	0	0	2	0	2	3,247.86	86	113
SDS	pckA (pckA)	Rv0211	67236	99.90%	2	2	3	0.01%	4.62%	WMDPGEMR	94.10%	1.09	0.0000118	1	0	0	0	2	1,038.19	102	109
	hypo protein									GVAIVNECQSGICGTCVAT											
SDS	Rv1937	Rv1937	93412	99.70%	2	2	2	0.01%	4.89%	CTAGR	95.00%	1.49	0.345	0	0	1	0	2	2,385.43	34	57
	hypo protein									TMPVRCDQTVLDAAEEHGV											
SDS	Rv1937	Rv1937	93412	99.70%	2	2	2	0.01%	4.89%	AIVNECQSGICGTCVATCTA GR	95.00%	1.77	0.254	0	0	1	0	2	4,353.42	17	57
SDS	glcB (glcB)	Rv1837c	80386	100.00%	2	2	2	0.01%	2.83%	GDLAAAVDKDGTAFLR	95.00%	2.06	0.456	0	1	0	0	2	1,620.91	297	312
SDS	glcB (glcB)	Rv1837c	80386	100.00%	2	2	2	0.01%	2.83%	NVGHL	95.00%	0.962	0.227	1	0	0	0	2	539.5384	340	344
	hypo protein																				
SDS	Rv3725	Rv3725	32707	99.50%	2	2	3	0.01%	5.83%	MLSTNMAGAQN	95.00%	1.32	0.219	1	0	0	0	2	1,154.31	94	104
	hypo protein																				
SDS	Rv3725	Rv3725	32707	99.50%	2	2	3	0.01%	5.83%	MQNATMR	92.40%	1.57	0.0627	2	0	0	0	2	867.9978	1	7
	hypo protein									ATGPDDFHLAYAAALASTG											
SDS	Rv0282	Rv0282	68089	100.00%	3	3	4	0.02%	13.50%	GPEEFAK	95.00%	4.92	0.59	0	2	0	0	2	2,608.95	134	159
	hypo protein									LGMFAPALSYLEEPDGPVAV											
SDS	Rv0282	Rv0282	68089	100.00%	3	3	4	0.02%	13.50%	AAVDGALAK	95.00%	1.8	0.364	0	1	0	0	2	2,890.38	226	254
	hypo protein									MAGVGEGDSGGVERDDIG											
SDS	Rv0282	Rv0282	68089	100.00%	3	3	4	0.02%	13.50%	MVAASPVASRVN	86.50%	1.45	0.208	0	0	1	0	2	2,920.95	1	30
	hypo protein																				
SDS	Rv1279	Rv1279	57313	99.70%	2	2	2	0.01%	2.65%	GHTH	85.40%	0.668	0.35	1	0	0	0	2	451.3684	509	512
	hypo protein																				
SDS	Rv1279	Rv1279	57313	99.70%	2	2	2	0.01%	2.65%	TLYHPMGTCR	93.10%	0.907	0.0127	1	0	0	0	2	1,195.39	465	474
	hypo protein																				
SDS	Rv1278	Rv1278	93301	100.00%	3	3	4	0.02%	9.14%	EDAVPVLIDDAL	95.00%	0.95	0.216	2	0	0	0	2	1,270.44	815	826
	hypo protein									GFTDPERLAKMGEVFDTIG											
SDS	Rv1278	Rv1278	93301	100.00%	3	3	4	0.02%	9.14%	ADGQVIVLTCSPTR	95.00%	1.98	0.257	0	0	1	0	2	3,542.01	827	859
	hypo protein									PTGEWSAAVSRLAAAEAAV											
SDS	Rv1278	Rv1278	93301	100.00%	3	3	4	0.02%	9.14%	ADCAAAVAEVDDGVRR	95.00%	1.86	0.249	0	0	1	0	2	3,528.59	191	225
										CDMGATILEGCMPDGSFRF											
SDS	adhD (adhD)	Rv3086	38323	99.70%	2	2	4	0.02%	10.30%	HSQGTDFGAMCMLGTFAE R	95.00%	1.98	0.196	0	0	3	0	2	4,266.89	105	142
										EGCMPDGSFRFHSQGTDF											
SDS	adhD (adhD)	Rv3086	38323	99.70%	2	2	4	0.02%	10.30%	GAMCMLGTFAER	95.00%	1.66	0.326	0	0	1	0	2	3,303.71	113	142
	hypo protein																				
SDS	Rv3071	Rv3071	40448	99.70%	2	2	2	0.01%	10.00%	EGLPIHR	95.00%	1.03	0.113	1	0	0	0	2	821.9884	260	266
	hypo protein									SEDPPVTVAAVDIESKIRSL											
SDS	Rv3071	Rv3071	40448	99.70%	2	2	2	0.01%	10.00%	VDDVTAMTDR	95.00%	1.7	0.289	0	0	1	0	2	3,247.49	66	95
										VEITGENVTECLGGAQDIS											
SDS	aroG (aroG)	Rv2178c	50623	99.70%	2	2	2	0.01%	7.79%	ETDLAGR	95.00%	2.1	0.188	0	1	0	0	2	2,749.88	410	435
SDS	aroG (aroG)	Rv2178c	50623	99.70%	2	2	2	0.01%	7.79%	VSRMGNHKVR	95.00%	1.17	0.191	1	0	0	0	2	1,200.29	335	344
										VTPVPGAFGRMPEQTAAAV											
SDS	lpqB (lpqB)	Rv3244c	61133	98.60%	2	2	2	0.01%	7.38%	SRSGRQVASVVTL	91.50%	1.71	0.219	0	0	1	0	2	3,241.39	333	364
SDS	lpqB (lpqB)	Rv3244c	61133	98.60%	2	2	2	0.01%	7.38%	VVFVETRSAEK	86.60%	1.23	0.0248	1	0	0	0	2	1,265.29	93	103
SDS	tuf (tuf)	Rv0685	43543	100.00%	21	27	179	0.70%	71.70%	ADAVDDEELLELVEMEVR	95.00%	5.15	0.699	0	22	0	0	2	2,076.19	140	157
SDS	tuf (tuf)	Rv0685	43543	100.00%	21	27	179	0.70%	71.70%	AFDQIDNAPEER	95.00%	3.57	0.406	3	2	0	0	2	1,405.50	48	59
SDS	tuf (tuf)	Rv0685	43543	100.00%	21	27	179	0.70%	71.70%	EHVLLAR	95.00%	1.55	0.295	5	0	0	0	2	837.9684	120	126
SDS	tuf (tuf)	Rv0685	43543	100.00%	21	27	179	0.70%	71.70%	ELLAAQEFDEDAPVVR	95.00%	3.57	0.571	3	19	0	0	2	1,802.92	158	173
SDS	tuf (tuf)	Rv0685	43543	100.00%	21	27	179	0.70%	71.70%	ETDKPFLMPVEDVFTITGR	95.00%	3.44	0.556	0	22	0	0	2	2,212.55	207	225
SDS	tuf (tuf)	Rv0685	43543	100.00%	21	27	179	0.70%	71.70%	FAIR	95.00%	1.27	0	3	0	0	0	2	506.6984	377	380
SDS	tuf (tuf)	Rv0685	43543	100.00%	21	27	179	0.70%	71.70%	FPDLNETK	95.00%	1.65	0.151	2	0	0	0	2	964.1384	40	47
SDS	tuf (tuf)	Rv0685	43543	100.00%	21	27	179	0.70%	71.70%	GITINIAHVEYQTDKR	95.00%	3.58	0.583	0	9	0	0	2	1,859.12	62	77
										GQVVTKPGTTTPHTEFEGQ											
SDS	tuf (tuf)	Rv0685	43543	100.00%	21	27	179	0.70%	71.70%	VYILSK	95.00%	3.53	0.57	0	2	1	0	2	2,718.90	292	316
SDS	tuf (tuf)	Rv0685	43543	100.00%	21	27	179	0.70%	71.70%	GVINVNEEVEIVGIRPSTTK	95.00%	4.31	0.672	0	14	0	0	2	2,155.30	236	255
SDS	tuf (tuf)	Rv0685	43543	100.00%	21	27	179	0.70%	71.70%	HTPFFNNYR	95.00%	1.16	0.112	2	0	0	0	2	1,196.34	322	330
SDS	tuf (tuf)	Rv0685	43543	100.00%	21	27	179	0.70%	71.70%	HYAHVDAPGHADYIK	95.00%	4.43	0.57	1	5	0	0	2	1,694.80	78	92
SDS	tuf (tuf)	Rv0685	43543	100.00%	21	27	179	0.70%	71.70%	IAHVEYQTDKR	95.00%	3.22	0.493	0	5	0	0	2	1,360.47	67	77
SDS	tuf (tuf)	Rv0685	43543	100.00%	21	27	179	0.70%	71.70%	IGTIGHVDHGK	95.00%	1.68	0.22	1	0	0	0	2	1,134.24	15	25
SDS	tuf (tuf)	Rv0685	43543	100.00%	21	27	179	0.70%	71.70%	LIQPVAMDEGLR	95.00%	3.01	0.49	4	9	0	0	2	1,342.64	365	376
										NMITGAAQMDGAILVVAAT											
SDS	tuf (tuf)	Rv0685	43543	100.00%	21	27	179	0.70%	71.70%	DGPMPQTR	95.00%	2.02	0.522	0	8	0	0	2	2,763.15	93	119
SDS	tuf (tuf)	Rv0685	43543	100.00%	21	27	179	0.70%	71.70%	PQFYFR	95.00%	1.87	0.348	3	1	0	0	2	858.0384	331	336
										TTDVTGVVTLPEGTEMVMP											
SDS	tuf (tuf)	Rv0685	43543	100.00%	21	27	179	0.70%	71.70%	GDNTN	95.00%	1.86	0.272	0	1	0	0	2	2,479.49	337	360
										TTDVTGVVTLPEGTEMVMP											
SDS	tuf (tuf)	Rv0685	43543	100.00%	21	27	179	0.70%	71.70%	GDNTNISVK	95.00%	3.07	0.478	0	15	0	0	2	2,907.04	337	364
SDS	tuf (tuf)	Rv0685	43543	100.00%	21	27	179	0.70%	71.70%	TTVTGVEMFR	95.00%	1.52	0.234	1	0	0	0	2	1,141.21	256	265
										WVASVEELMNAVDESIPDP											
SDS	tuf (tuf)	Rv0685	43543	100.00%	21	27	179	0.70%	71.70%	VR	95.00%	5.04	0.563	0	16	0	0	2	2,373.46	186	206
										DTTTGDTLSDPNQQIVLES											
SDS	fusA (fusA)	Rv0684	77153	100.00%	9	10	18	0.07%	24.00%	MTFPDPVIEVAIEPK	95.00%	3.06	0.426	0	0	2	0	2	3,719.10	387	420
										EYIPSVDAGAQDAMQYGVL											
SDS	fusA (fusA)	Rv0684	77153	100.00%	9	10	18	0.07%	24.00%	AGYPLVNLK	95.00%	3.47	0.638	0	3	0	0	2	3,000.42	544	571
SDS	fusA (fusA)	Rv0684	77153	100.00%	9	10	18	0.07%	24.00%	ILYYTGINYK	95.00%	1.59	0.264	2	0	0	0	2	1,248.67	33	42
SDS	fusA (fusA)	Rv0684	77153	100.00%	9	10	18	0.07%	24.00%	LAEEDPTFK	95.00%	1.63	0.12	1	0	0	0	2	1,050.22	435	443
										LGETYDTVEIPADLAEQAEE											
SDS	fusA (fusA)	Rv0684	77153	100.00%	9	10	18	0.07%	24.00%	YR	95.00%	3.39	0.556	0	2	0	0	2	2,513.70	195	216
SDS	fusA (fusA)	Rv0684	77153	100.00%	9	10	18	0.07%	24.00%	VIINLEPFTGEEGATYEFESK	95.00%	4.47	0.637	0	2	0	0	2	2,374.67	515	535
SDS	fusA (fusA)	Rv0684	77153	100.00%	9	10	18	0.07%	24.00%	VLDGAVAVFDGK	95.00%	2.28	0.249	2	0	0	0	2	1,191.29	101	112
SDS	fusA (fusA)	Rv0684	77153	100.00%	9	10	18	0.07%	24.00%	VTLLDGAYHEVDSSEMAFK	95.00%	3.58	0.516	0	2	0	0	2	2,129.33	572	590
SDS	fusA (fusA)	Rv0684	77153	100.00%	9	10	18	0.07%	24.00%	YLGGEELTVDEIK	95.00%	1.17	0.308	1	1	0	0	2	1,466.69	234	246

SDS	echA4 (echA4)	Rv0673	33523	99.50%	2	2	2	0.01%	13.10%	LFTGDCITGAQAAEWGLAV EAPEPADLDER	92.30%	1.9	0.198	0	0	1	0	2	3,218.54	198	227
SDS	echA4	Rv0673	33523	99.50%	2	2	2	0.01%	13.10%	MVSTVFDGAAR	95.00%	0.867	0.0533	1	0	0	0	0	2	1,154.25	264	274
SDS	pks5 (pks5)	Rv1527c	223837	99.50%	2	2	2	0.01%	0.66%	DLLRDMGIEH	92.50%	1.14	0.192	1	0	0	0	0	2	1,215.47	1597	1606
SDS	pks5 (pks5)	Rv1527c	223837	99.50%	2	2	2	0.01%	0.66%	MCDK	95.00%	1.02	0.165	1	0	0	0	2	496.6184	2094	2097
SDS	hypo protein Rv0622	Rv0622	36094	99.20%	2	2	2	0.01%	9.38%	CTGSDDAAMSFCVYCGAEL ADPTR	86.00%	1.12	0.306	0	0	1	0	2	2,555.77	19	42
SDS	hypo protein Rv0622	Rv0622	36094	99.20%	2	2	2	0.01%	9.38%	CTGSDDAAMSFCVYCGAEL ADPTRCGACGAYK	95.00%	2.23	0.274	0	0	1	0	2	3,467.78	19	50
SDS	mmpL12 (mmpL12)	Rv1522c	122384	99.40%	2	2	2	0.01%	5.24%	DADKPSMAGFNIPPQIFSR DEFKK	92.60%	2.2	0.206	0	0	1	0	2	2,756.35	728	751
SDS	mmpL12 (mmpL12)	Rv1522c	122384	99.40%	2	2	2	0.01%	5.24%	DESPHGIRIGVIRTVGSTG GVITSAGLIFAASMFGL	93.00%	1.8	0.255	0	0	1	0	2	3,589.26	908	943
SDS	adi (adi)	Rv2531c	106013	99.40%	2	2	2	0.01%	4.54%	RYATPFFDALR	95.00%	1.34	0.121	1	0	0	0	2	1,357.70	287	297
SDS	adi (adi)	Rv2531c	106013	99.40%	2	2	2	0.01%	4.54%	TLVSTTFVVPYPPGFPVLVP GQLVSKEIIYFL	89.40%	1.09	0.314	0	0	1	0	2	3,523.20	845	876
SDS	hypo protein Rv1866	Rv1866	83165	99.60%	2	2	2	0.01%	5.91%	ERVMFGPAVTA	93.70%	1.3	0.0494	1	0	0	0	2	1,178.29	768	778
SDS	hypo protein Rv1866	Rv1866	83165	99.60%	2	2	2	0.01%	5.91%	SELVAAVSAWTSTRTPVQA AGALQAAGVAAGPMNR	95.00%	2.12	0.266	0	0	1	0	2	3,411.63	658	692
SDS	rpsA (rpsA)	Rv1630	53184	100.00%	5	5	10	0.04%	16.40%	HDVDPNEVVSVGDEVEALV LTK	95.00%	3.36	0.525	0	4	0	0	2	2,365.27	72	93
SDS	rpsA (rpsA)	Rv1630	53184	100.00%	5	5	10	0.04%	16.40%	HVEVPDQVVAVGDDAMVK	95.00%	4.56	0.602	0	2	0	0	2	1,924.77	330	347
SDS	rpsA (rpsA)	Rv1630	53184	100.00%	5	5	10	0.04%	16.40%	SEFLNNLQK	95.00%	1.26	0.192	1	0	0	0	2	1,093.30	200	208
SDS	rpsA (rpsA)	Rv1630	53184	100.00%	5	5	10	0.04%	16.40%	TAGGSLASDAQLAALR	95.00%	3.48	0.406	0	2	0	0	2	1,502.78	459	474
SDS	rpsA (rpsA)	Rv1630	53184	100.00%	5	5	10	0.04%	16.40%	YFNDGDIVEGTIVK	95.00%	2.6	0.485	0	1	0	0	2	1,570.76	32	45
SDS	pks12 (pks12)	Rv2048c	431553	100.00%	4	4	4	0.02%	2.96%	DIMITLGMFTH	95.00%	1.01	0.204	1	0	0	0	2	1,295.67	1401	1411
SDS	pks12 (pks12)	Rv2048c	431553	100.00%	4	4	4	0.02%	2.96%	LLLPMPADWSYAEAAAISA VFTTAYYAFIHL	92.50%	2.13	0.181	0	0	1	0	2	3,419.23	1457	1487
SDS	pks12 (pks12)	Rv2048c	431553	100.00%	4	4	4	0.02%	2.96%	SSEPIAIVGMSCRFPGGVD SPEGLWQMVADARDVMSE FPTDR	95.00%	1.22	0.296	0	0	1	0	2	4,646.10	33	74
SDS	pks12 (pks12)	Rv2048c	431553	100.00%	4	4	4	0.02%	2.96%	TPAVRTTSEDPIAIVGMACR YPGGVNSPDDMWDMLIQG R	95.00%	1.69	0.319	0	0	1	0	2	4,238.81	2048	2086
SDS	mmpL11 (mmpL11)	Rv0202c	103485	99.70%	2	2	4	0.02%	2.90%	MMRLSR	95.00%	1.62	0.0863	1	0	0	0	2	810.0978	1	6
SDS	mmpL11 (mmpL11)	Rv0202c	103485	99.70%	2	2	4	0.02%	2.90%	TQLPRVAGAAQVDVGGPTA LIK	95.00%	1.87	0.276	0	3	0	0	2	2,163.44	497	518
SDS	hypo protein Rv1815	Rv1815	22768	99.00%	2	2	3	0.01%	26.70%	FGVSTGETCGTVESVNNG WFTMSHGVLSEK	95.00%	2.3	0.415	0	0	2	0	2	3,249.25	140	169
SDS	hypo protein Rv1815	Rv1815	22768	99.00%	2	2	3	0.01%	26.70%	SVWGGFPAAVSWRSTSEQ VHADLGVTPLA	83.70%	1.79	0.221	0	0	1	0	2	3,027.13	193	221
SDS	linB (linB)	Rv2579	33710	100.00%	3	3	6	0.02%	14.70%	NLPIDGEPAEVVALVNEYR	95.00%	4.22	0.631	0	4	0	0	2	2,099.28	211	229
SDS	linB (linB)	Rv2579	33710	100.00%	3	3	6	0.02%	14.70%	SPQGEPMALEHNIFVER	95.00%	2.43	0.369	0	1	0	0	2	1,971.14	157	173
SDS	linB (linB)	Rv2579	33710	100.00%	3	3	6	0.02%	14.70%	VLPGAILR	95.00%	1.07	0.1	1	0	0	0	2	839.1384	174	181
SDS	hypo protein Rv3903c	Rv3903c	88316	99.40%	2	2	2	0.01%	6.86%	LRAAAVAWRSAGTQFALTEI QSTAGPMGVIR	89.30%	2	0.203	0	0	1	0	2	3,247.75	151	181
SDS	hypo protein Rv3903c	Rv3903c	88316	99.40%	2	2	2	0.01%	6.86%	SNPVGVVIDGTGQQLKGFG EEAFGMAK	95.00%	1.7	0.269	0	0	1	0	2	2,737.99	312	338
SDS	hypo protein Rv1165	Rv1165	67530	99.30%	2	2	2	0.01%	2.71%	AGKVGGHKLTAR	95.00%	1.12	0.0659	1	0	0	0	2	1,195.40	339	350
SDS	hypo protein Rv1165	Rv1165	67530	99.30%	2	2	2	0.01%	2.71%	GITIL	89.10%	0.864	0.117	1	0	0	0	2	516.7684	51	55
SDS	narG (narG)	Rv1161	136906	99.90%	3	3	4	0.02%	3.25%	AVNGDRSAELVTLPSFDTID GH	90.20%	1.91	0.263	0	1	0	0	2	2,315.50	429	450
SDS	narG (narG)	Rv1161	136906	99.90%	3	3	4	0.02%	3.25%	IMSIDFR	84.60%	0.895	0.123	1	0	0	0	2	898.1978	765	771
SDS	narG (narG)	Rv1161	136906	99.90%	3	3	4	0.02%	3.25%	TPDAHWMAEAR	95.00%	1.67	0.172	2	0	0	0	2	1,301.38	272	282
SDS	hypo protein Rv0327c	Rv0327c	49994	99.70%	2	2	3	0.01%	9.13%	APEIQSQVVWAAVGGDDG FLEAVLKEGMRR	95.00%	2.06	0.286	0	0	2	0	2	3,246.52	281	310
SDS	hypo protein Rv0327c	Rv0327c	49994	99.70%	2	2	3	0.01%	9.13%	NITTVPKHGAH	95.00%	1.28	0.0985	1	0	0	0	2	1,175.24	421	431
SDS	groES	Rv3418c	10754	100.00%	6	6	9	0.04%	71.00%	DVLAVVSK	95.00%	1.34	0.282	1	0	0	0	2	830.8884	93	100
SDS	groES	Rv3418c	10754	100.00%	6	6	9	0.04%	71.00%	EKPQEGTVVAVGPGR	95.00%	2.89	0.443	0	1	0	0	2	1,524.50	36	50
SDS	groES (groES)	Rv3418c	10754	100.00%	6	6	9	0.04%	71.00%	ILVQANEAETTTASGLVIPDT AK	95.00%	3.2	0.485	0	4	0	0	2	2,343.63	13	35
SDS	groES	Rv3418c	10754	100.00%	6	6	9	0.04%	71.00%	IPLDVAEGDTVIYSK	95.00%	4.46	0.515	0	1	0	0	2	1,620.89	59	73
SDS	groES	Rv3418c	10754	100.00%	6	6	9	0.04%	71.00%	RIPLDVAEGDTVIYSK	95.00%	3.96	0.591	0	1	0	0	2	1,777.09	58	73
SDS	groES	Rv3418c	10754	100.00%	6	6	9	0.04%	71.00%	VNIKPLEDK	94.30%	1.51	0.0813	1	0	0	0	2	1,056.31	4	12
SDS	groEL1 (groEL1)	Rv3417c	55859	100.00%	16	18	57	0.22%	43.60%	AADAVSEALLASATPVSGK	95.00%	3.16	0.447	1	2	0	0	2	1,758.95	122	140
SDS	groEL1 (groEL1)	Rv3417c	55859	100.00%	16	18	57	0.22%	43.60%	AAVEEGIVPGGGASLIHQA R	95.00%	3.88	0.578	0	5	0	0	2	1,933.08	403	422
SDS	groEL1 (groEL1)	Rv3417c	55859	100.00%	16	18	57	0.22%	43.60%	AEIDKSDSDWDREK	95.00%	2.73	0.381	0	1	0	0	2	1,694.86	349	362
SDS	groEL1 (groEL1)	Rv3417c	55859	100.00%	16	18	57	0.22%	43.60%	AFGGPTVTNDGVTVAR	95.00%	2.64	0.462	0	4	0	0	2	1,562.56	42	57
SDS	groEL1 (groEL1)	Rv3417c	55859	100.00%	16	18	57	0.22%	43.60%	AFLEDLAVVTGGQVVNPDA GMVLR	95.00%	4.99	0.605	0	7	0	0	2	2,472.68	285	308
SDS	groEL1 (groEL1)	Rv3417c	55859	100.00%	16	18	57	0.22%	43.60%	DEQIGDLVGEAMSK	95.00%	3.22	0.498	0	3	0	0	2	1,508.65	153	166
SDS	groEL1 (groEL1)	Rv3417c	55859	100.00%	16	18	57	0.22%	43.60%	EIELEDPFEDLGAQLVK	95.00%	3.55	0.476	0	7	0	0	2	1,946.24	58	74
SDS	groEL1 (groEL1)	Rv3417c	55859	100.00%	16	18	57	0.22%	43.60%	EVGLEVLGSAR	95.00%	1.74	0.276	2	0	0	0	2	1,130.22	309	319
SDS	groEL1 (groEL1)	Rv3417c	55859	100.00%	16	18	57	0.22%	43.60%	LVAAGVNPIALGVGIGK	95.00%	4.69	0.624	1	3	0	0	2	1,549.89	105	121
SDS	groEL1 (groEL1)	Rv3417c	55859	100.00%	16	18	57	0.22%	43.60%	PDLLPLLEK	95.00%	1.59	0.285	1	0	0	0	2	1,038.46	229	237
SDS	groEL1 (groEL1)	Rv3417c	55859	100.00%	16	18	57	0.22%	43.60%	RVVVSKDDTVIVDGGGTAE AVAN	95.00%	2.1	0.277	0	2	0	0	2	2,273.21	320	342
SDS	groEL1 (groEL1)	Rv3417c	55859	100.00%	16	18	57	0.22%	43.60%	TGIAQVATVSSR	95.00%	3.03	0.601	0	1	0	0	2	1,190.23	141	152
SDS	groEL1 (groEL1)	Rv3417c	55859	100.00%	16	18	57	0.22%	43.60%	TGIAQVATVSSRDEQIGDL VGEAMSK	95.00%	3.89	0.561	0	2	0	0	2	2,663.86	141	166
SDS	groEL1 (groEL1)	Rv3417c	55859	100.00%	16	18	57	0.22%	43.60%	TLSYGDLAADGVIDPVK	95.00%	3.18	0.579	0	2	0	0	2	1,735.02	482	498
SDS	groEL1 (groEL1)	Rv3417c	55859	100.00%	16	18	57	0.22%	43.60%	TNDVAGDGTTTATILAQALI K	95.00%	4.33	0.59	0	11	0	0	2	2,075.41	80	100

SDS	groEL1 (groEL1)	Rv3417c	55859	100.00%	16	18	57	0.22%	43.60%	VVVSKDDTVIVDGGGTAEA VANR	95.00%	3.3	0.46	0	2	0	0	2	2,273.21	321	343
SDS	hypo protein Rv0540	Rv0540	22922	99.30%	2	2	2	0.01%	18.20%	DCCAPGSR	90.60%	1.19	0.0403	1	0	0	0	2	808.8484	203	210
SDS	hypo protein Rv0540	Rv0540	22922	99.30%	2	2	2	0.01%	18.20%	DTLDAVAAAPVTARAVALTG DLDSAADSAEIR	93.80%	1.2	0.281	0	0	1	0	2	3,128.46	43	74
SDS	hypo protein Rv1330c	Rv1330c	54284	99.80%	2	2	2	0.01%	3.73%	LVEVDGVPVQK	87.20%	1.48	0.153	1	0	0	0	2	1,183.17	403	413
SDS	hypo protein Rv1330c	Rv1330c	54284	99.80%	2	2	2	0.01%	3.73%	MVSAAGGR	94.60%	0.989	0.236	1	0	0	0	2	764.8278	217	224
SDS	mmpL2 (mmpL2)	Rv0507	106188	99.50%	2	2	2	0.01%	5.06%	GTPLQGAAIYLGGTAATFK	95.00%	1.42	0.417	0	1	0	0	2	1,838.26	737	755
SDS	mmpL2 (mmpL2)	Rv0507	106188	99.50%	2	2	2	0.01%	5.06%	PLGTTIKHTSIPFIISMQGV NSSEQMEFMK	92.10%	1.23	0.265	0	0	1	0	2	3,386.05	476	505
SDS	fas (fas)	Rv2524c	326201	99.90%	2	2	2	0.01%	1.27%	LLATYSELRSR	95.00%	1.53	0.267	1	0	0	0	2	1,309.66	584	594
SDS	fas (fas)	Rv2524c	326201	99.90%	2	2	2	0.01%	1.27%	RDEIIAAMAKTAKPYFGDVA DMTYLQWL	95.00%	1.15	0.331	0	0	1	0	2	3,235.96	713	740
SDS	fbpB (fbpB)	Rv1886c	34564	100.00%	3	3	16	0.06%	15.10%	AADMWGPSSDPAWER	95.00%	3.12	0.371	0	2	0	0	2	1,676.81	216	230
SDS	fbpB (fbpB)	Rv1886c	34564	100.00%	3	3	16	0.06%	15.10%	PGLPVEYLQVPSPSMGR	95.00%	4.58	0.556	0	12	0	0	2	1,828.10	44	60
SDS	fbpB (fbpB)	Rv1886c	34564	100.00%	3	3	16	0.06%	15.10%	WETFLTSELPQWLSANR	95.00%	2.42	0.497	0	2	0	0	2	2,079.38	137	153
SDS	hypo protein Rv1273c	Rv1273c	62098	100.00%	3	3	8	0.03%	16.30%	FSGTVADNLRYGGGPDQV VTEQEMWEALR	95.00%	2.05	0.35	0	0	6	0	2	3,243.39	422	450
SDS	hypo protein Rv1273c	Rv1273c	62098	100.00%	3	3	8	0.03%	16.30%	IDSGQMQVGSLIAFLSYFA QILMAVLMATMTL	95.00%	1.82	0.196	0	0	1	0	2	3,466.42	264	295
SDS	hypo protein Rv1273c	Rv1273c	62098	100.00%	3	3	8	0.03%	16.30%	QILFLVQMTATVLVTAPIMC VGGIIMAIHQEAAL	95.00%	1.15	0.318	0	0	1	0	2	3,614.42	124	157
SDS	hypo protein Rv1794	Rv1794	32381	99.70%	2	2	6	0.02%	10.30%	LLYGVIDDENQPPGSR	95.00%	3.32	0.398	0	3	0	0	2	1,774.01	94	109
SDS	hypo protein Rv1794	Rv1794	32381	99.70%	2	2	6	0.02%	10.30%	VLAAPDLEVVALLSR	95.00%	2.64	0.429	0	3	0	0	2	1,566.87	77	91
SDS	hypo protein Rv3786c	Rv3786c	44927	99.50%	2	2	3	0.01%	7.13%	MADFCRPDWVMMVDADW LVETDIDLR	91.90%	2.11	0.204	0	0	2	0	2	3,247.69	80	105
SDS	hypo protein Rv3786c	Rv3786c	44927	99.50%	2	2	3	0.01%	7.13%	MADFCRPDWVMMVDADW LVETDIDLRAVL	95.00%	1.67	0.338	0	0	1	0	2	3,427.96	80	108
TX114	lprA (lprA)	Rv1270c	24856	99.70%	2	4	8	0.09%	11.50%	ITGNSSADDIATLAGSR	95.00%	3.96	0.556	1	4	0	0	2	1,649.90	175	191
TX114	lprA (lprA)	Rv1270c	24856	99.70%	2	4	8	0.09%	11.50%	LAVTGDVPNLR	95.00%	2.36	0.316	2	1	0	0	2	1,155.31	67	77
TX114	hypo protein Rv3534c	Rv3534c	36425	99.70%	2	2	2	0.02%	9.54%	EARDNGGSICRIATH	95.00%	1.42	0.234	0	1	0	0	2	1,600.76	99	113
TX114	hypo protein Rv3534c	Rv3534c	36425	99.70%	2	2	2	0.02%	9.54%	KLIGGQEDQLIDIALEIK	95.00%	1.55	0.301	0	1	0	0	2	1,997.60	318	335
TX114	hypo protein Rv3786c	Rv3786c	44927	99.60%	2	2	3	0.04%	13.00%	MADFCRPDWVMMVDADW LVETDIDLR	95.00%	1.42	0.29	0	0	2	0	2	3,247.69	80	105
TX114	hypo protein Rv3786c	Rv3786c	44927	99.60%	2	2	3	0.04%	13.00%	MCPMVSRWDDPEYPDLIPV MGTAEALR	93.30%	1.24	0.222	0	0	1	0	2	3,213.70	119	145
TX114	groEL2 (groEL2)	Rv0440	56709	100.00%	19	19	33	0.39%	42.20%	AAVEEGIVAGGGVTLLQAA PTLDELK	95.00%	3.21	0.448	0	2	0	0	2	2,523.85	403	428
TX114	groEL2 (groEL2)	Rv0440	56709	100.00%	19	19	33	0.39%	42.20%	ASVPGGGDMGGMDF	95.00%	1.52	0.305	1	0	0	0	2	1,298.42	527	540
TX114	groEL2 (groEL2)	Rv0440	56709	100.00%	19	19	33	0.39%	42.20%	AVEKVTETLLK	95.00%	1.45	0.254	1	0	0	0	2	1,231.44	122	132
TX114	groEL2 (groEL2)	Rv0440	56709	100.00%	19	19	33	0.39%	42.20%	DETTIVEGAGDTDAIAGR	95.00%	2.76	0.321	0	1	0	0	2	1,791.88	326	343
TX114	groEL2 (groEL2)	Rv0440	56709	100.00%	19	19	33	0.39%	42.20%	DLLPLLEK	95.00%	1.89	0.349	2	0	0	0	2	941.3384	230	237
TX114	groEL2 (groEL2)	Rv0440	56709	100.00%	19	19	33	0.39%	42.20%	EIELEDPYEK	95.00%	2.65	0.353	1	0	0	0	2	1,265.44	58	67
TX114	groEL2 (groEL2)	Rv0440	56709	100.00%	19	19	33	0.39%	42.20%	EQIAATAAISAGDQSIGDLI AEAMDK	95.00%	3.09	0.591	0	1	0	0	2	2,607.04	141	166
TX114	groEL2 (groEL2)	Rv0440	56709	100.00%	19	19	33	0.39%	42.20%	GLNALADAVK	95.00%	1.31	0.352	3	0	0	0	2	972.1784	18	27
TX114	groEL2 (groEL2)	Rv0440	56709	100.00%	19	19	33	0.39%	42.20%	KTDDVAGDGTTTATVLAQA LVR	95.00%	5.59	0.62	0	4	0	0	2	2,204.35	79	100
TX114	groEL2 (groEL2)	Rv0440	56709	100.00%	19	19	33	0.39%	42.20%	KVVVTKDETTIVEGAGDTD AIAGR	95.00%	6.6	0.69	0	4	0	0	2	2,446.59	320	343
TX114	groEL2 (groEL2)	Rv0440	56709	100.00%	19	19	33	0.39%	42.20%	KWGAPTITNDGVSIAK	95.00%	2.39	0.318	0	1	0	0	2	1,658.95	42	57
TX114	groEL2 (groEL2)	Rv0440	56709	100.00%	19	19	33	0.39%	42.20%	LAGGVAVIK	95.00%	1.44	0.197	1	0	0	0	2	828.0184	370	378
TX114	groEL2 (groEL2)	Rv0440	56709	100.00%	19	19	33	0.39%	42.20%	NVAAGANPLGLK	95.00%	2.21	0.302	1	0	0	0	2	1,125.35	105	116
TX114	groEL2 (groEL2)	Rv0440	56709	100.00%	19	19	33	0.39%	42.20%	QEAVLEDPYILLVSSK	95.00%	3.71	0.507	0	1	0	0	2	1,805.12	209	224
TX114	groEL2 (groEL2)	Rv0440	56709	100.00%	19	19	33	0.39%	42.20%	QIAFNSGLEPGVVAEK	95.00%	3.72	0.611	0	3	0	0	2	1,659.82	450	465
TX114	groEL2 (groEL2)	Rv0440	56709	100.00%	19	19	33	0.39%	42.20%	TDDVAGDGTTTATVLAQAL VR	95.00%	4.17	0.507	0	2	0	0	2	2,076.15	80	100
TX114	groEL2 (groEL2)	Rv0440	56709	100.00%	19	19	33	0.39%	42.20%	VALEAPLK	95.00%	1.74	0.237	1	0	0	0	2	841.0684	442	449
TX114	groEL2 (groEL2)	Rv0440	56709	100.00%	19	19	33	0.39%	42.20%	VVVTKDETTIVEGAGDTDAI AGR	95.00%	4.86	0.639	0	1	0	0	2	2,318.39	321	343
TX114	groEL2 (groEL2)	Rv0440	56709	100.00%	19	19	33	0.39%	42.20%	WGAPTITNDGVSIAK	95.00%	3.38	0.613	0	2	0	0	2	1,530.75	43	57
TX114	hypo protein Rv0475	Rv0475	21517	100.00%	6	8	15	0.18%	29.10%	AAEGYLEAATSR	95.00%	2.31	0.467	0	1	0	0	2	1,239.37	83	94
TX114	hypo protein Rv0475	Rv0475	21517	100.00%	6	8	15	0.18%	29.10%	AEGYVDQAVELTQEALGTV ASQTR	95.00%	6.5	0.559	0	4	1	0	2	2,537.53	122	145
TX114	hypo protein Rv0475	Rv0475	21517	100.00%	6	8	15	0.18%	29.10%	KAAEGYLEAATSR	95.00%	2.65	0.566	0	1	0	0	2	1,367.57	82	94
TX114	hypo protein Rv0475	Rv0475	21517	100.00%	6	8	15	0.18%	29.10%	LQEDLPEQLTELR	95.00%	4.34	0.601	1	4	0	0	2	1,584.84	60	72
TX114	hypo protein Rv0475	Rv0475	21517	100.00%	6	8	15	0.18%	29.10%	LVGIELPK	95.00%	1.8	0.179	2	0	0	0	2	869.1584	154	161

TX114	hypo protein Rv0475	Rv0475	21517	100.00%	6	8	15	0.18%	29.10%	PEQLTELR	95.00%	1.38	0.167	1	0	0	0	2	986.1384	65	72
TX114	hspX (hspX)	Rv2031c	16210	99.90%	2	2	7	0.08%	18.80%	AELPGVDPDKDVDIMVR	95.00%	2.74	0.471	0	4	0	0	2	1,886.10	55	71
TX114	hspX (hspX)	Rv2031c	16210	99.90%	2	2	7	0.08%	18.80%	SEFAYGSFVR	95.00%	1.41	0.271	3	0	0	0	2	1,163.28	91	100
TX114	hypo protein Rv1393c	Rv1393c	55234	100.00%	3	3	3	0.04%	15.20%	IPGGKAVHR	95.00%	0.832	0.253	1	0	0	0	2	935.0884	224	232
TX114	hypo protein Rv1393c	Rv1393c	55234	100.00%	3	3	3	0.04%	15.20%	TGKRVGIIGTGASAVQVIPE MAPIVSHLTVFQR	95.00%	1.88	0.273	0	0	1	0	2	3,434.91	169	201
TX114	hypo protein Rv1393c	Rv1393c	55234	100.00%	3	3	3	0.04%	15.20%	VTEPIDKITPTAVATTDGAS HEIDVLVLATGFK	95.00%	1.75	0.202	0	0	1	0	2	3,411.84	309	341
TX114	sodC (sodC)	Rv0432	23826	100.00%	4	4	13	0.15%	27.50%	GDGSAMLVTTTDAFTMDDL LSGAK	95.00%	4.13	0.603	0	7	0	0	2	2,434.84	166	189
TX114	sodC (sodC)	Rv0432	23826	100.00%	4	4	13	0.15%	27.50%	TAIIIHAGADNFANIPPER YVQVNGTPGPDETTLTTGD	95.00%	4.51	0.687	0	4	0	0	2	2,021.43	190	208
TX114	sodC (sodC)	Rv0432	23826	100.00%	4	4	13	0.15%	27.50%	AGK YVQVNGTPGPDETTLTTGD	95.00%	5.63	0.631	0	1	0	0	2	2,222.28	209	230
TX114	sodC (sodC)	Rv0432	23826	100.00%	4	4	13	0.15%	27.50%	AGKR	95.00%	2.53	0.368	0	1	0	0	2	2,378.48	209	231
TX114	lppX (lppX)	Rv2945c	24122	100.00%	4	4	5	0.06%	29.60%	ASIDLGSGSIQLTQSK	95.00%	3.66	0.625	0	1	0	0	2	1,605.92	210	225
TX114	lppX (lppX)	Rv2945c	24122	100.00%	4	4	5	0.06%	29.60%	LFDDWSNLGSISELSTSR	95.00%	3.39	0.474	0	2	0	0	2	2,028.37	122	139
TX114	lppX (lppX)	Rv2945c	24122	100.00%	4	4	5	0.06%	29.60%	PATVWIAQDGSHHLVR	95.00%	4.85	0.512	0	1	0	0	2	1,787.87	194	209
TX114	lppX (lppX)	Rv2945c	24122	100.00%	4	4	5	0.06%	29.60%	TTGKVDSLLGITSADVDVR HPMVPTISPKKSWEGFAGS	95.00%	2.6	0.292	0	1	0	0	2	1,948.17	73	91
TX114	cdsA (cdsA)	Rv2881c	31985	99.50%	2	2	2	0.02%	21.90%	LVCGITATIITATFL IGGQAAVWLTWPFGAVGAL	90.70%	1.08	0.258	0	0	1	0	2	3,647.43	203	236
TX114	cdsA (cdsA)	Rv2881c	31985	99.50%	2	2	2	0.02%	21.90%	AGFGGMVVVCMIWR	95.00%	1.9	0.137	0	0	1	0	2	3,438.94	92	124
TX114	hypo protein Rv3894c	Rv3894c	153685	99.70%	2	2	3	0.04%	3.08%	DVKLMVVTRN	95.00%	1.23	0.0884	1	0	0	0	2	1,191.33	248	257
TX114	hypo protein Rv3894c	Rv3894c	153685	99.70%	2	2	3	0.04%	3.08%	GKRGAWTPPTVASPTAMGS ALETGQVGVDLGPH	95.00%	1.82	0.24	0	0	2	0	2	3,247.49	304	336
TX114	pks2 (pks2)	Rv3825c	225737	99.70%	2	2	2	0.02%	2.40%	NSRTQPTQKMRETIEAIAA R	95.00%	2.52	0.349	0	0	1	0	2	4,192.48	1802	1840
TX114	pks2 (pks2)	Rv3825c	225737	99.70%	2	2	2	0.02%	2.40%	MGSEVVVECGDIAQPGTAE SPGGSTVAVHPL	95.00%	1.27	0.313	1	0	0	0	2	1,122.14	905	916
TX114	mmpL8 (mmpL8)	Rv3823c	115984	100.00%	3	3	3	0.04%	8.26%	DTATTPAMAGFYIPPELLSYA TGESVKAETMPSEYR	90.70%	1.31	0.258	0	0	1	0	2	3,929.51	750	785
TX114	mmpL8 (mmpL8)	Rv3823c	115984	100.00%	3	3	3	0.04%	8.26%	GDTIDADASFVANNTEWAS PVLGALDSSPMCTADPACA SAR	91.90%	1.3	0.21	0	0	1	0	2	4,187.52	596	636
TX114	mmpL8 (mmpL8)	Rv3823c	115984	100.00%	3	3	3	0.04%	8.26%	MGFGLGEASAFLL	83.30%	1.57	0.033	1	0	0	0	2	1,329.71	734	746
TX114	bioB (bioB)	Rv1589	37500	99.70%	2	2	2	0.02%	9.46%	MRWCGPEVEVEGIISLKTG GCPEDCHFCSQSGL	95.00%	2.49	0.24	0	0	1	0	2	3,783.12	65	97
TX114	bioB (bioB)	Rv1589	37500	99.70%	2	2	2	0.02%	9.46%	TGGCPEDCH	95.00%	1.05	0.193	1	0	0	0	2	1,060.99	82	90
TX114	hypo protein Rv1520	Rv1520	39528	99.80%	2	2	4	0.05%	11.30%	HVRHAVGGEIAMLPETMAV YR	81.60%	1.32	0.244	0	0	1	0	2	2,370.59	181	201
TX114	hypo protein Rv1520	Rv1520	39528	99.80%	2	2	4	0.05%	11.30%	QPSYDDIPANVMPIDWYLH AGAVEAEPAAGAGEQVSVE	95.00%	2.04	0.21	0	3	0	0	2	2,275.61	163	181
TX114	fixB (fixB)	Rv3028c	31672	99.70%	2	2	2	0.02%	16.70%	VPAAAENAAR ALGEPAAVVVGVPGTAAPLV	95.00%	3.4	0.549	0	1	0	0	2	2,693.55	152	180
TX114	fixB (fixB)	Rv3028c	31672	99.70%	2	2	2	0.02%	16.70%	DGLK	95.00%	3.66	0.585	0	1	0	0	2	2,202.43	28	51
TX114	phoS1	Rv0934	38194	100.00%	5	7	17	0.20%	26.50%	ASFLDQVHFQPLPPAVVK	95.00%	4.11	0.616	0	6	0	0	2	1,994.23	346	363
TX114	phoS1	Rv0934	38194	100.00%	5	7	17	0.20%	26.50%	SDGSGDTFLFTQYLSK	95.00%	3.54	0.521	1	5	0	0	2	1,767.06	186	201
TX114	phoS1 (phoS1)	Rv0934	38194	100.00%	5	7	17	0.20%	26.50%	SSGNFLLPDAQSIQAAAAG FASK	95.00%	3.17	0.433	0	1	0	0	2	2,252.64	269	291
TX114	phoS1 (phoS1)	Rv0934	38194	100.00%	5	7	17	0.20%	26.50%	TPANQAISMIDGPAPDGYPI INYEYAIVNNR	95.00%	5.02	0.637	0	2	1	0	2	3,395.97	292	322
TX114	phoS1	Rv0934	38194	100.00%	5	7	17	0.20%	26.50%	VLAAMYQGTIK	88.90%	1.31	0	1	0	0	0	2	1,195.50	149	159
TX114	35kd_ag (35kd_ag)	Rv2744c	29240	100.00%	10	10	61	0.72%	61.10%	ATEYNNAAEAFAAQLVTAE QSVEDLK	95.00%	4.2	0.567	0	4	0	0	2	2,784.90	95	120
TX114	35kd_ag (35kd_ag)	Rv2744c	29240	100.00%	10	10	61	0.72%	61.10%	GEALPAGGTTATPRPATETS GGAIAEQPYGQ	95.00%	2.55	0.545	0	7	0	0	2	2,958.16	240	270
TX114	35kd_ag (35kd_ag)	Rv2744c	29240	100.00%	10	10	61	0.72%	61.10%	MLEVEQAGIQMAGHSR	95.00%	4	0.421	0	5	0	0	2	1,757.93	215	230
TX114	35kd_ag (35kd_ag)	Rv2744c	29240	100.00%	10	10	61	0.72%	61.10%	QALTLADQATAAGDAAK	95.00%	4.06	0.656	0	5	0	0	2	1,616.83	78	94
TX114	35kd_ag (35kd_ag)	Rv2744c	29240	100.00%	10	10	61	0.72%	61.10%	RYANAIGSAELAESSVQGR	95.00%	3.65	0.473	0	4	0	0	2	1,980.15	196	214
TX114	35kd_ag (35kd_ag)	Rv2744c	29240	100.00%	10	10	61	0.72%	61.10%	SMSELAAPGNTPSLDEVR	95.00%	2.9	0.453	0	16	0	0	2	1,891.08	173	190
TX114	35kd_ag (35kd_ag)	Rv2744c	29240	100.00%	10	10	61	0.72%	61.10%	THQALTQQAAQVIGNQR	95.00%	5.11	0.697	0	6	0	0	2	1,864.88	40	56
TX114	35kd_ag (35kd_ag)	Rv2744c	29240	100.00%	10	10	61	0.72%	61.10%	VQIQQAIEEAQR	95.00%	3.55	0.499	0	2	0	0	2	1,413.45	28	39
TX114	35kd_ag (35kd_ag)	Rv2744c	29240	100.00%	10	10	61	0.72%	61.10%	YANAIGSAELAESSVQGR	95.00%	4.28	0.599	0	11	0	0	2	1,823.95	197	214
TX114	35kd_ag (35kd_ag)	Rv2744c	29240	100.00%	10	10	61	0.72%	61.10%	YLMALFSSK	95.00%	1.26	0.298	1	0	0	0	2	1,060.46	11	19
TX114	ppiA (ppiA)	Rv0009	19222	100.00%	5	6	15	0.18%	48.40%	DYSTQNASGGPSGPFYDG AVFHR	95.00%	3.46	0.541	0	3	0	0	2	2,431.53	51	73
TX114	ppiA (ppiA)	Rv0009	19222	100.00%	5	6	15	0.18%	48.40%	HTIFGEVIDAESQR	95.00%	5.08	0.667	0	3	0	0	2	1,602.70	141	154
TX114	ppiA (ppiA)	Rv0009	19222	100.00%	5	6	15	0.18%	48.40%	TATDGNDRPTDPVVIESITI S	95.00%	2.28	0.315	0	3	0	0	2	2,202.39	162	182
TX114	ppiA (ppiA)	Rv0009	19222	100.00%	5	6	15	0.18%	48.40%	TVANFVGLAQGTK	95.00%	1.45	0.248	1	1	0	0	2	1,306.42	38	50
TX114	ppiA (ppiA)	Rv0009	19222	100.00%	5	6	15	0.18%	48.40%	VIQGFMIQGGDPTGTGR	95.00%	2.26	0.396	0	4	0	0	2	1,750.96	74	90
TX114	sucA (sucA)	Rv1248c	133954	100.00%	3	3	3	0.04%	7.00%	LTTEDFAGVTISLTNPGTIGT VH	91.70%	1.05	0.252	0	0	1	0	2	2,345.59	228	250
TX114	sucA (sucA)	Rv1248c	133954	100.00%	3	3	3	0.04%	7.00%	SVPRLMPGQGAIIGVGAME YPAEFQGASEER	95.00%	2.09	0.233	0	0	1	0	2	3,249.65	251	281
TX114	sucA (sucA)	Rv1248c	133954	100.00%	3	3	3	0.04%	7.00%	YVGQKRFSLEGAESVIPMM DAAIDQCAEHGL	95.00%	1.56	0.255	0	0	1	0	2	3,367.81	485	515
TX114	pckA (pckA)	Rv0211	67236	99.70%	2	2	3	0.04%	9.41%	GAEDPKLGVEITDSEYVVV SMRTMTRMGK	95.00%	2.21	0.182	0	0	1	0	2	3,249.61	136	164
TX114	pckA (pckA)	Rv0211	67236	99.70%	2	2	3	0.04%	9.41%	PGEMRSIMK	95.00%	1.29	0.229	0	0	2	0	2	3,247.86	86	113
TX114	hypo protein Rv3421c	Rv3421c	21553	98.50%	2	2	2	0.02%	24.60%	AGMASAAAYGHALGIPVYG VCSLDAIGGQTIGDTLVVT DAR	91.30%	1.25	0.224	0	0	1	0	2	3,980.47	82	122

TX114	hypo protein Rv3421c	Rv3421c	21553	98.50%	2	2	2	0.02%	24.60%	AHAERLTPNVL	83.90%	1.42	0.0493	1	0	0	0	2	1,221.37	42	52
TX114	hypo protein Rv2623	Rv2623	31633	100.00%	4	4	9	0.11%	21.50%	AGPPTVHSEIVPAAAVPTLV DMSK	95.00%	3.6	0.624	0	2	0	0	2	2,388.61	86	109
TX114	hypo protein Rv2623	Rv2623	31633	100.00%	4	4	9	0.11%	21.50%	GGYAGMLVGSVGETVAQLA R	95.00%	3.7	0.568	0	3	0	0	2	1,953.10	267	286
TX114	hypo protein Rv2623	Rv2623	31633	100.00%	4	4	9	0.11%	21.50%	HLIDDALK	95.00%	1.52	0.413	2	0	0	0	2	925.1984	70	77
TX114	hypo protein Rv2623	Rv2623	31633	100.00%	4	4	9	0.11%	21.50%	SEEAQLVVVGSR	95.00%	1.35	0.295	2	0	0	0	2	1,274.22	253	264
TX114	lprF (lprF)	Rv1368	26833	100.00%	3	3	3	0.04%	18.40%	GLGAVVGQVQNPTIQGR	95.00%	4.7	0.607	0	1	0	0	2	1,694.73	161	177
TX114	lprF (lprF)	Rv1368	26833	100.00%	3	3	3	0.04%	18.40%	IYDPGILDKDR	95.00%	1.49	0.386	0	1	0	0	2	1,418.89	149	160
TX114	lprF (lprF)	Rv1368	26833	100.00%	3	3	3	0.04%	18.40%	VSGTIDAAVIDPIVPQLGK	95.00%	2.36	0.355	0	1	0	0	2	1,894.21	188	206
TX114	atpF (atpF)	Rv1306	18307	100.00%	7	7	20	0.24%	29.80%	AEQQVASTLQTAHEQLK	95.00%	4.79	0.61	0	3	0	0	2	1,882.91	118	134
TX114	atpF (atpF)	Rv1306	18307	100.00%	7	7	20	0.24%	29.80%	AEQQVASTLQTAHEQLKR	95.00%	3.15	0.453	0	3	0	0	2	2,039.11	118	135
TX114	atpF (atpF)	Rv1306	18307	100.00%	7	7	20	0.24%	29.80%	ERDAVELDLR	95.00%	1.9	0.16	2	0	0	0	2	1,216.37	136	145
TX114	atpF (atpF)	Rv1306	18307	100.00%	7	7	20	0.24%	29.80%	KSDEQFAAAQADYDEAMTE AR	95.00%	3.87	0.61	0	4	0	0	2	2,364.48	73	93
TX114	atpF (atpF)	Rv1306	18307	100.00%	7	7	20	0.24%	29.80%	QTAHEQLKRER	95.00%	1.07	0.0504	1	0	0	0	2	1,396.50	127	137
TX114	atpF (atpF)	Rv1306	18307	100.00%	7	7	20	0.24%	29.80%	SDEQFAAAQADYDEAMTEA R	95.00%	4.95	0.584	0	4	0	0	2	2,220.28	74	93
TX114	atpF (atpF)	Rv1306	18307	100.00%	7	7	20	0.24%	29.80%	VRAEQQVASTLQTAHEQLK	95.00%	4.77	0.604	0	3	0	0	2	2,138.18	116	134
TX114	fadE18 (fadE18)	Rv1933c	38459	97.70%	2	2	2	0.02%	10.50%	ACADMMIEIDATRATVMFA AMSAANGDELQTVAPLAK	82.90%	1.19	0.266	0	0	1	0	2	3,878.51	265	301
TX114	fadE18 (fadE18)	Rv1933c	38459	97.70%	2	2	2	0.02%	10.50%	HACADMMIEIDATRATVMF AAMSAANGDEL	87.20%	2.2	0.124	0	0	1	0	2	3,277.73	264	293
TX114	hypo protein Rv3792	Rv3792	69499	99.70%	2	2	3	0.04%	5.44%	AAVDPLCRLAVVGAIAAAIG STTWL	95.00%	1.53	0.361	0	1	0	0	2	2,511.97	301	325
TX114	hypo protein Rv3792	Rv3792	69499	99.70%	2	2	3	0.04%	5.44%	YYPAIDAAIR	95.00%	1.29	0.18	2	0	0	0	2	1,153.48	497	506
TX114	lpqH (lpqH)	Rv3763	15096	100.00%	3	3	17	0.20%	34.00%	IAIGGAATGIAAVLTDGNPP EVK	95.00%	3.68	0.539	0	3	0	0	2	2,136.50	76	98
TX114	lpqH (lpqH)	Rv3763	15096	100.00%	3	3	17	0.20%	34.00%	ITGTATGVDMANPMSPVNK	95.00%	3.37	0.499	0	12	0	0	2	1,937.14	132	150
TX114	lpqH (lpqH)	Rv3763	15096	100.00%	3	3	17	0.20%	34.00%	SVGLGNVNGVTL	95.00%	1.27	0.231	2	0	0	0	2	1,130.16	99	110
TX114	hypo protein Rv3496c	Rv3496c	47183	99.50%	2	2	2	0.02%	12.40%	SETPVPELTWVPAGGGAPV GNPADLQSL	95.00%	1.02	0.346	0	0	1	0	2	2,760.93	398	425
TX114	hypo protein Rv3496c	Rv3496c	47183	99.50%	2	2	2	0.02%	12.40%	SSDVKITMSVSKDVKVPVD VQAVIMSPN	89.90%	1.74	0.174	0	0	1	0	2	2,991.23	75	102
TX114	hypo protein Rv3083	Rv3083	55466	99.70%	2	2	2	0.02%	4.44%	CAVPNGDLFK	95.00%	1.04	0.251	1	0	0	0	2	1,135.32	298	307
TX114	hypo protein Rv3083	Rv3083	55466	99.70%	2	2	2	0.02%	4.44%	IAIQRKLYQACR	95.00%	1.24	0.332	0	0	1	0	2	1,534.96	248	259
TX114	tuf (tuf)	Rv0685	43543	100.00%	6	6	8	0.09%	23.70%	EHVLLAR	95.00%	1.56	0.182	1	0	0	0	2	837.9684	120	126
TX114	tuf (tuf)	Rv0685	43543	100.00%	6	6	8	0.09%	23.70%	ELLAAQEFDEDAPVVR	95.00%	4.1	0.584	0	1	0	0	2	1,802.92	158	173
TX114	tuf (tuf)	Rv0685	43543	100.00%	6	6	8	0.09%	23.70%	ETDKPFLMPVEDVFTITGR	95.00%	3.41	0.491	0	2	0	0	2	2,212.55	207	225
TX114	tuf (tuf)	Rv0685	43543	100.00%	6	6	8	0.09%	23.70%	HYAHVDAPGHADYIK	95.00%	2.96	0.593	0	1	0	0	2	1,694.80	78	92
TX114	tuf (tuf)	Rv0685	43543	100.00%	6	6	8	0.09%	23.70%	TTDVTGVVTLPEGTEMVMP GDNTNISVK	95.00%	2.55	0.311	0	2	0	0	2	2,923.04	337	364
TX114	tuf (tuf)	Rv0685	43543	100.00%	6	6	8	0.09%	23.70%	TTLTAAITK	91.20%	0.991	0	1	0	0	0	2	920.1784	26	34
TX114	hypo protein Rv2471	Rv2471	60426	99.10%	2	2	2	0.02%	8.24%	IPPFGFSTCPDTWLPMPPE WAALTAEKQR	95.00%	2.02	0.292	0	0	1	0	2	3,288.91	433	461
TX114	hypo protein Rv2471	Rv2471	60426	99.10%	2	2	2	0.02%	8.24%	YFFRDGRGPDGSLPPN	82.30%	1.16	0.22	0	0	1	0	2	1,796.11	153	168
TX114	pks12 (pks12)	Rv2048c	431553	99.50%	2	2	2	0.02%	1.54%	SYARTGGFVDGVADFDPAF FGISPSEALAMDPQHRML	95.00%	1.82	0.319	0	0	1	0	2	3,991.56	92	128
TX114	pks12 (pks12)	Rv2048c	431553	99.50%	2	2	2	0.02%	1.54%	VDSDAPLDDSALAAVVTTG EPQVLWRR	90.00%	1.3	0.263	0	0	1	0	2	2,883.07	3335	3361
TX114	mmpL11 (mmpL11)	Rv0202c	103485	99.70%	2	2	3	0.04%	2.90%	MMRLSR	95.00%	1.41	0.139	1	0	0	0	2	810.0978	1	6
TX114	mmpL11 (mmpL11)	Rv0202c	103485	99.70%	2	2	3	0.04%	2.90%	TQLPRVAGAAQVDVGGPTA LIK	95.00%	2.26	0.141	0	2	0	0	2	2,163.44	497	518
TX114	lpqG (lpqG)	Rv3623	24818	99.70%	2	2	3	0.04%	17.90%	VISISEASGAAPTHEAPAPP R	95.00%	4.14	0.621	0	1	0	0	2	2,059.26	195	215
TX114	lpqG (lpqG)	Rv3623	24818	99.70%	2	2	3	0.04%	17.90%	VTVAPQYSNPEPAGTATITG YR	95.00%	4.11	0.655	0	2	0	0	2	2,294.44	99	120
TX114	narG (narG)	Rv1161	136906	99.70%	2	2	2	0.02%	1.87%	AGKHLVCTVFDL	95.00%	1.67	0.0876	1	0	0	0	2	1,303.49	463	474
TX114	narG (narG)	Rv1161	136906	99.70%	2	2	2	0.02%	1.87%	TPPAHWMAAR	95.00%	1.7	0.138	1	0	0	0	2	1,301.38	272	282
TX114	dppA (dppA)	Rv3666c	58365	99.50%	2	2	2	0.02%	4.07%	FGGEEGRLRR	90.90%	1.1	0.175	1	0	0	0	2	1,177.37	307	316
TX114	dppA (dppA)	Rv3666c	58365	99.50%	2	2	2	0.02%	4.07%	SRTTMSGLRVVN	95.00%	1.46	0.195	1	0	0	0	2	1,321.47	150	161
TX114	hemB (hemB)	Rv0512	34853	99.10%	2	2	2	0.02%	13.70%	AGADIVLTYWAVDAAGWLT	95.00%	2	0.217	0	0	1	0	2	1,994.26	311	329
TX114	hemB (hemB)	Rv0512	34853	99.10%	2	2	2	0.02%	13.70%	AVAQAESGAHVVGPSGMM DGQVAAIR	83.60%	1.47	0.183	0	0	1	0	2	2,542.56	158	183
TX114	fas (fas)	Rv2524c	326201	98.90%	2	2	2	0.02%	1.50%	IRWEDDPQPGWYDTESGE MVDESELVQRYHDAVVQR	95.00%	1.87	0.255	0	0	1	0	2	4,353.38	2482	2517
TX114	fas (fas)	Rv2524c	326201	98.90%	2	2	2	0.02%	1.50%	SLVHLDHAAR	79.50%	1.14	0.0204	1	0	0	0	2	1,119.23	1116	1125
2DLC TX114	hypo protein Rv1771	Rv1771	48027	100.00%	4	4	6	0.03%	15.40%	ALGPQLAQR	95.00%	2.54	0.339	0	2	0	0	2	954.1484	90	98
2DLC TX114	hypo protein Rv1771	Rv1771	48027	100.00%	4	4	6	0.03%	15.40%	LGLENQGDVDPQSITGATA TATHGTGVR	95.00%	4.39	0.528	0	0	1	0	2	2,767.85	100	127
2DLC TX114	hypo protein Rv1771	Rv1771	48027	100.00%	4	4	6	0.03%	15.40%	SLPIMFPIEVR	95.00%	2.8	0.46	0	2	0	0	2	1,302.71	323	333
2DLC TX114	hypo protein Rv1771	Rv1771	48027	100.00%	4	4	6	0.03%	15.40%	VLDVDQPTGLVTVEGGAK	95.00%	3.51	0.287	0	0	1	0	2	1,798.85	70	87
2DLC TX114	lprA (lprA)	Rv1270c	24856	100.00%	3	5	13	0.08%	11.50%	ITGNSSADDIATLAGSR	95.00%	3.88	0.656	0	4	4	0	2	1,649.90	175	191
2DLC TX114	lprA (lprA)	Rv1270c	24856	100.00%	3	5	13	0.08%	11.50%	LAVTGDVPNLR	95.00%	1.58	0.308	2	1	0	0	2	1,155.31	67	77
2DLC TX114	lprA (lprA)	Rv1270c	24856	100.00%	3	5	13	0.08%	11.50%	SSADDIATLAGSR	95.00%	2.36	0.426	0	2	0	0	2	1,264.45	179	191
2DLC TX114	hypo protein Rv3718c	Rv3718c	15643	99.80%	2	3	3	0.02%	16.30%	ILSPHYSEYQVLEGGK	95.00%	3.39	0.499	0	0	1	0	2	1,821.07	33	48
2DLC TX114	hypo protein Rv3718c	Rv3718c	15643	99.80%	2	3	3	0.02%	16.30%	TFAPLGLK	95.00%	1.29	0.437	1	1	0	0	2	847.1684	122	129
2DLC TX114	prfB (prfB)	Rv3105c	42107	99.90%	2	2	2	0.01%	6.61%	SGAGGVDAADWAEMLMR	95.00%	4.2	0.597	0	1	0	0	2	1,737.94	142	158
2DLC TX114	prfB (prfB)	Rv3105c	42107	99.90%	2	2	2	0.01%	6.61%	VLDVEGLR	95.00%	2.37	0.338	0	1	0	0	2	901.0084	29	36
2DLC TX114	hypo protein Rv1748	Rv1748	27050	100.00%	4	4	11	0.06%	14.40%	LLDFPVMTDMR	95.00%	2.96	0.568	0	3	0	0	2	1,338.71	109	119

2DLC TX114	hypo protein Rv1748	Rv1748	27050	100.00%	4	4	11	0.06%	14.40%	TELDGLIVLPDR	95.00%	3.89	0.444	0	4	0	0	2	1,341.66	217	228
2DLC TX114	hypo protein Rv1748	Rv1748	27050	100.00%	4	4	11	0.06%	14.40%	TELDGLIVLPDRTR	95.00%	2.14	0.271	0	0	2	0	2	1,598.96	217	230
2DLC TX114	hypo protein Rv1748	Rv1748	27050	100.00%	4	4	11	0.06%	14.40%	WLEYELDAAK	95.00%	3.26	0.454	0	2	0	0	2	1,238.48	99	108
2DLC TX114	hypo protein Rv1747	Rv1747	92104	99.40%	2	2	2	0.01%	4.39%	GVTWTIDGDKTLLDGISLTAR	87.10%	2.09	0.149	0	0	1	0	2	2,233.70	323	343
2DLC TX114	hypo protein Rv1747	Rv1747	92104	99.40%	2	2	2	0.01%	4.39%	VHIGNPDGPALDFEVGR	95.00%	2.99	0.533	0	0	1	0	2	1,793.93	89	105
2DLC TX114	hypo protein Rv2466c	Rv2466c	23017	99.80%	2	2	4	0.02%	8.70%	EVAKVR	95.00%	1.17	0.197	1	0	0	0	2	701.7384	31	36
2DLC TX114	hypo protein Rv2466c	Rv2466c	23017	99.80%	2	2	4	0.02%	8.70%	VLDPLYTAMGNR	95.00%	3.52	0.564	0	3	0	0	2	1,350.64	86	97
2DLC TX114	lpqE (lpqE)	Rv3584	18769	99.80%	2	3	12	0.07%	18.10%	AGQGSVMVPISAGLATPHE	95.00%	3.95	0.606	0	4	0	0	2	1,822.95	164	182
2DLC TX114	lpqE (lpqE)	Rv3584	18769	99.80%	2	3	12	0.07%	18.10%	IQAVQTSDFIQPGK	95.00%	3.44	0.445	0	6	2	0	2	1,532.72	62	75
2DLC TX114	clpB (clpB)	Rv0384c	92520	100.00%	6	7	15	0.09%	9.91%	ALADFLFDDER	95.00%	2.99	0.543	0	3	0	0	2	1,312.58	619	629
2DLC TX114	clpB (clpB)	Rv0384c	92520	100.00%	6	7	15	0.09%	9.91%	FLPDKAIDLVDEAASR	95.00%	4.04	0.584	0	3	1	0	2	1,761.13	386	401
2DLC TX114	clpB (clpB)	Rv0384c	92520	100.00%	6	7	15	0.09%	9.91%	KLDAALPQAQAR	89.30%	1.84	0.188	0	0	1	0	2	1,282.56	513	524
2DLC TX114	clpB (clpB)	Rv0384c	92520	100.00%	6	7	15	0.09%	9.91%	LVGATTLDEYR	95.00%	3.14	0.517	0	4	0	0	2	1,238.42	311	321
2DLC TX114	clpB (clpB)	Rv0384c	92520	100.00%	6	7	15	0.09%	9.91%	LVQQAIGDQLAK	95.00%	3.05	0.557	0	2	0	0	2	1,284.50	810	821
2DLC TX114	clpB (clpB)	Rv0384c	92520	100.00%	6	7	15	0.09%	9.91%	NNPVLIGEPGVGK	95.00%	2.5	0.502	0	4	0	0	2	1,294.45	200	212
2DLC TX114	clpB (clpB)	Rv0384c	92520	100.00%	6	7	15	0.09%	9.91%	NPVLIGEPGVGK	95.00%	2.59	0.556	0	2	0	0	2	1,180.35	201	212
2DLC TX114	clpB (clpB)	Rv0384c	92520	100.00%	6	7	15	0.09%	9.91%	TKNNPVLIGEPGVGK	95.00%	3.18	0.287	0	0	2	0	2	1,523.75	198	212
2DLC TX114	clpB (clpB)	Rv0384c	92520	100.00%	6	7	15	0.09%	9.91%	VVQVLSR	95.00%	1.99	0.318	0	1	0	0	2	800.8084	190	196
2DLC TX114	groEL2 (groEL2)	Rv0440	56709	100.00%	42	66	364	2.10%	47.40%	AGAATEVELK	95.00%	3.15	0.448	2	17	0	0	2	989.0784	379	388
2DLC TX114	groEL2 (groEL2)	Rv0440	56709	100.00%	42	66	364	2.10%	47.40%	AGAATEVELKER	95.00%	3.14	0.521	0	3	7	0	2	1,274.38	379	390
2DLC TX114	groEL2 (groEL2)	Rv0440	56709	100.00%	42	66	364	2.10%	47.40%	AGGVAVIK	95.00%	1.59	0.304	2	1	0	0	2	714.8184	371	378
2DLC TX114	groEL2 (groEL2)	Rv0440	56709	100.00%	42	66	364	2.10%	47.40%	ALADAVK	95.00%	1.62	0.187	2	0	0	0	2	687.8284	21	27
2DLC TX114	groEL2 (groEL2)	Rv0440	56709	100.00%	42	66	364	2.10%	47.40%	APGFGDR	95.00%	2.04	0.604	0	13	0	0	2	719.8184	276	282
2DLC TX114	groEL2 (groEL2)	Rv0440	56709	100.00%	42	66	364	2.10%	47.40%	AQTGVYEDLLAAGVADPVKVTR	95.00%	2.24	0.346	0	1	0	0	2	2,274.44	477	498
2DLC TX114	groEL2 (groEL2)	Rv0440	56709	100.00%	42	66	364	2.10%	47.40%	ASVPGGGDMGGMDF	95.00%	1.75	0.491	6	4	0	0	2	1,314.42	527	540
2DLC TX114	groEL2 (groEL2)	Rv0440	56709	100.00%	42	66	364	2.10%	47.40%	AVEKVTETLLK	95.00%	2.69	0.517	6	18	19	0	2	1,231.44	122	132
2DLC TX114	groEL2 (groEL2)	Rv0440	56709	100.00%	42	66	364	2.10%	47.40%	DLLPLLEK	95.00%	1.83	0.399	8	1	0	0	2	941.3384	230	237
2DLC TX114	groEL2 (groEL2)	Rv0440	56709	100.00%	42	66	364	2.10%	47.40%	EIELEDPYEK	95.00%	1.33	0.0984	1	0	0	0	2	1,265.44	58	67
2DLC TX114	groEL2 (groEL2)	Rv0440	56709	100.00%	42	66	364	2.10%	47.40%	EIELEDPYEKIGAELVK	95.00%	4.57	0.488	0	4	3	0	2	1,976.34	58	74
2DLC TX114	groEL2 (groEL2)	Rv0440	56709	100.00%	42	66	364	2.10%	47.40%	EIELEDPYEKIGAELVKEVAK	95.00%	2.8	0.319	0	0	1	0	2	2,403.79	58	78
2DLC TX114	groEL2 (groEL2)	Rv0440	56709	100.00%	42	66	364	2.10%	47.40%	GLNALADAVK	95.00%	2.13	0.394	13	17	0	0	2	972.1784	18	27
2DLC TX114	groEL2 (groEL2)	Rv0440	56709	100.00%	42	66	364	2.10%	47.40%	IEDAVR	95.00%	1.44	0.245	2	3	0	0	2	702.7684	394	399
2DLC TX114	groEL2 (groEL2)	Rv0440	56709	100.00%	42	66	364	2.10%	47.40%	IGAELVK	95.00%	1.53	0.0906	4	8	0	0	2	729.9184	68	74
2DLC TX114	groEL2 (groEL2)	Rv0440	56709	100.00%	42	66	364	2.10%	47.40%	IGAELVKEVAK	95.00%	3.93	0.471	0	4	3	0	2	1,157.37	68	78
2DLC TX114	groEL2 (groEL2)	Rv0440	56709	100.00%	42	66	364	2.10%	47.40%	KTDDVAGDGTTTATVL	94.60%	1.77	0.261	0	1	0	0	2	1,565.62	79	94
2DLC TX114	groEL2 (groEL2)	Rv0440	56709	100.00%	42	66	364	2.10%	47.40%	KTDDVAGDGTTTATVLAQAL	95.00%	4.48	0.625	0	3	0	0	2	1,949.08	79	98
2DLC TX114	groEL2 (groEL2)	Rv0440	56709	100.00%	42	66	364	2.10%	47.40%	KTDDVAGDGTTTATVLAQALVR	95.00%	5.96	0.684	0	11	11	0	2	2,204.35	79	100
2DLC TX114	groEL2 (groEL2)	Rv0440	56709	100.00%	42	66	364	2.10%	47.40%	KVVVTKDETTIVEGAGDTDAIAGR	95.00%	5.16	0.607	0	2	13	0	2	2,446.59	320	343
2DLC TX114	groEL2 (groEL2)	Rv0440	56709	100.00%	42	66	364	2.10%	47.40%	KWGAPTITNDGVSIAK	95.00%	3.95	0.521	0	2	2	0	2	1,658.95	42	57
2DLC TX114	groEL2 (groEL2)	Rv0440	56709	100.00%	42	66	364	2.10%	47.40%	LAGGVAVIK	95.00%	1.86	0.247	5	15	0	0	2	828.0184	370	378
2DLC TX114	groEL2 (groEL2)	Rv0440	56709	100.00%	42	66	364	2.10%	47.40%	LPLLEK	95.00%	1.77	0.368	3	1	0	0	2	713.0384	232	237
2DLC TX114	groEL2 (groEL2)	Rv0440	56709	100.00%	42	66	364	2.10%	47.40%	NALADAVK	95.00%	1.88	0.381	3	0	0	0	2	801.9284	20	27
2DLC TX114	groEL2 (groEL2)	Rv0440	56709	100.00%	42	66	364	2.10%	47.40%	NAQTGVYEDLLAAGVADPVKVTR	95.00%	1.91	0.246	0	0	1	0	2	2,388.54	476	498
2DLC TX114	groEL2 (groEL2)	Rv0440	56709	100.00%	42	66	364	2.10%	47.40%	NVAAGANPL	95.00%	1.78	0.273	4	2	0	0	2	826.8984	105	113
2DLC TX114	groEL2 (groEL2)	Rv0440	56709	100.00%	42	66	364	2.10%	47.40%	NVAAGANPLGL	95.00%	2.27	0.438	2	0	0	0	2	997.1484	105	115
2DLC TX114	groEL2 (groEL2)	Rv0440	56709	100.00%	42	66	364	2.10%	47.40%	NVAAGANPLGLK	95.00%	2.97	0.522	6	21	1	0	2	1,125.35	105	116
2DLC TX114	groEL2 (groEL2)	Rv0440	56709	100.00%	42	66	364	2.10%	47.40%	NVVLEK	95.00%	1.68	0.227	4	4	0	0	2	701.7584	36	41
2DLC TX114	groEL2 (groEL2)	Rv0440	56709	100.00%	42	66	364	2.10%	47.40%	NVVLEKK	95.00%	2.29	0.387	0	1	0	0	2	829.9584	36	42
2DLC TX114	groEL2 (groEL2)	Rv0440	56709	100.00%	42	66	364	2.10%	47.40%	QIAFNSGLEPGVVAEK	95.00%	3.55	0.622	0	8	7	0	2	1,659.82	450	465
2DLC TX114	groEL2 (groEL2)	Rv0440	56709	100.00%	42	66	364	2.10%	47.40%	SGLEPGVVAEK	95.00%	2.12	0.438	0	1	0	0	2	1,086.14	455	465
2DLC TX114	groEL2 (groEL2)	Rv0440	56709	100.00%	42	66	364	2.10%	47.40%	TIAYDEEAR	95.00%	1.84	0.457	0	1	0	0	2	1,068.18	4	12
2DLC TX114	groEL2 (groEL2)	Rv0440	56709	100.00%	42	66	364	2.10%	47.40%	TIAYDEEARR	95.00%	1.92	0.301	0	1	0	0	2	1,224.38	4	13
2DLC TX114	groEL2 (groEL2)	Rv0440	56709	100.00%	42	66	364	2.10%	47.40%	VAAGANPLGLK	95.00%	3.1	0.611	0	4	0	0	2	1,011.25	106	116
2DLC TX114	groEL2 (groEL2)	Rv0440	56709	100.00%	42	66	364	2.10%	47.40%	VALEAPLK	95.00%	1.8	0.372	2	16	0	0	2	841.0684	442	449
2DLC TX114	groEL2 (groEL2)	Rv0440	56709	100.00%	42	66	364	2.10%	47.40%	VAQIRQEIEN	95.00%	1.9	0.285	0	1	0	0	2	1,200.27	344	353
2DLC TX114	groEL2 (groEL2)	Rv0440	56709	100.00%	42	66	364	2.10%	47.40%	VSTVKDLLPLL	89.70%	1.93	0.207	0	1	0	0	2	1,198.56	225	235

2DLC TX114	groEL2 (groEL2)	Rv0440	56709	100.00%	42	66	364	2.10%	47.40%	VTETLLK	95.00%	1.11	0.227	2	11	0	0	2	803.9884	126	132
2DLC TX114	groEL2 (groEL2)	Rv0440	56709	100.00%	42	66	364	2.10%	47.40%	VTLGPK	95.00%	1.19	0.209	3	0	0	0	2	614.7584	28	33
2DLC TX114	groEL2 (groEL2)	Rv0440	56709	100.00%	42	66	364	2.10%	47.40%	WGAPTITNDGVSIAK	95.00%	4.2	0.636	0	11	0	0	2	1,530.75	43	57
2DLC TX114	groEL2 (groEL2)	Rv0440	56709	100.00%	42	66	364	2.10%	47.40%	WGAPTITNDGVSIAKEIELE DPYEK	95.00%	4.26	0.613	0	0	4	0	2	2,777.17	43	67
2DLC TX114	hypo protein Rv3547	Rv3547	17354	99.80%	2	2	7	0.04%	12.60%	IPVALLTTTGR	95.00%	2.82	0.274	0	1	0	0	2	1,142.44	44	54
2DLC TX114	hypo protein Rv3547	Rv3547	17354	99.80%	2	2	7	0.04%	12.60%	VNPLYFLR	95.00%	2.24	0.428	0	6	0	0	2	1,022.31	61	68
2DLC TX114	tig (tig)	Rv2462c	50566	100.00%	8	8	20	0.12%	22.50%	AAFDAEAR	95.00%	2.91	0.454	0	2	0	0	2	850.9384	337	344
2DLC TX114	tig (tig)	Rv2462c	50566	100.00%	8	8	20	0.12%	22.50%	ASLSDQVR	95.00%	2.38	0.257	0	1	0	0	2	875.9284	264	271
2DLC TX114	tig (tig)	Rv2462c	50566	100.00%	8	8	20	0.12%	22.50%	FNELLVEQGSSR	95.00%	2.78	0.509	0	4	0	0	2	1,379.50	325	336
2DLC TX114	tig (tig)	Rv2462c	50566	100.00%	8	8	20	0.12%	22.50%	NQLPTMFADVR	95.00%	2.68	0.416	0	5	0	0	2	1,292.49	399	409
2DLC TX114	tig (tig)	Rv2462c	50566	100.00%	8	8	20	0.12%	22.50%	QYGIEPQQLFGYLQER	95.00%	3.98	0.7	0	3	0	0	2	1,970.24	383	398
2DLC TX114	tig (tig)	Rv2462c	50566	100.00%	8	8	20	0.12%	22.50%	STVEQLSPTR	95.00%	2.77	0.497	0	2	0	0	2	1,118.17	3	12
2DLC TX114	tig (tig)	Rv2462c	50566	100.00%	8	8	20	0.12%	22.50%	VSAGEAEEAEPADEGAAR YGQAVAESDVQPLGRPNIE	88.20%	1.86	0.204	0	0	1	0	2	1,759.67	441	458
2DLC TX114	tig (tig)	Rv2462c	50566	100.00%	8	8	20	0.12%	22.50%	VTK	95.00%	3.38	0.585	0	0	2	0	2	2,372.51	76	97
2DLC TX114	rpsB (rpsB)	Rv2890c	31072	99.80%	2	2	6	0.03%	6.97%	SAALLTR	95.00%	2.05	0.373	0	4	0	0	2	731.9584	209	215
2DLC TX114	rpsB (rpsB)	Rv2890c	31072	99.80%	2	2	6	0.03%	6.97%	VIASAVAEGLQAR	95.00%	3	0.428	0	2	0	0	2	1,285.41	216	228
2DLC TX114	hypo protein Rv2296	Rv2296	33340	100.00%	6	6	9	0.05%	24.30%	AAWEALGR	95.00%	2.44	0.206	0	1	0	0	2	874.0084	230	237
2DLC TX114	hypo protein Rv2296	Rv2296	33340	100.00%	6	6	9	0.05%	24.30%	AGYDAPFPDKTYQAGAR	95.00%	3.9	0.65	0	0	2	0	2	1,829.08	194	210
2DLC TX114	hypo protein Rv2296	Rv2296	33340	100.00%	6	6	9	0.05%	24.30%	ASHFIQEDSGTELAER	95.00%	3.05	0.379	0	0	1	0	2	1,790.89	277	292
2DLC TX114	hypo protein Rv2296	Rv2296	33340	100.00%	6	6	9	0.05%	24.30%	LVPTSPDDPAVPANR	95.00%	2.03	0.236	0	0	1	0	2	1,549.68	215	229
2DLC TX114	hypo protein Rv2296	Rv2296	33340	100.00%	6	6	9	0.05%	24.30%	SAAGHRVLAPDLIGFGR	94.80%	1.95	0.206	0	0	1	0	2	1,738.08	70	86
2DLC TX114	hypo protein Rv2296	Rv2296	33340	100.00%	6	6	9	0.05%	24.30%	VLAPDLIGFGR	95.00%	2.88	0.502	0	3	0	0	2	1,158.49	76	86
2DLC TX114	hypo protein Rv1096	Rv1096	31059	99.80%	2	2	5	0.03%	8.25%	LVTVSELLGPR	95.00%	3.18	0.441	0	4	0	0	2	1,184.41	223	233
2DLC TX114	hypo protein Rv1096	Rv1096	31059	99.80%	2	2	5	0.03%	8.25%	YRPAGGLSNDAVR	95.00%	2.08	0.25	0	0	1	0	2	1,376.55	141	153
2DLC TX114	desA2	Rv1094	31342	99.80%	2	2	3	0.02%	8.73%	DWDPSGSTLPR	95.00%	2.61	0.254	0	1	0	0	2	1,302.42	53	63
2DLC TX114	desA2	Rv1094	31342	99.80%	2	2	3	0.02%	8.73%	ITAWGLAGEPSLK	95.00%	3.44	0.537	0	2	0	0	2	1,343.68	258	270
2DLC TX114	glyA (glyA)	Rv1093	45012	100.00%	3	3	4	0.02%	6.10%	IAATPEFADR	95.00%	1.88	0.291	0	1	0	0	2	1,091.28	273	282
2DLC TX114	glyA (glyA)	Rv1093	45012	100.00%	3	3	4	0.02%	6.10%	LMAPDVAK	95.00%	1.94	0.244	0	1	0	0	2	845.0684	297	304
2DLC TX114	glyA (glyA)	Rv1093	45012	100.00%	3	3	4	0.02%	6.10%	VLDFAAFR	95.00%	2.56	0.484	0	2	0	0	2	939.1484	178	185
2DLC TX114	hypo protein Rv3046c	Rv3046c	13333	100.00%	3	3	9	0.05%	29.80%	STPLTEEQIGEVVR	95.00%	2.06	0.265	0	2	0	0	2	1,558.61	40	53
2DLC TX114	hypo protein Rv3046c	Rv3046c	13333	100.00%	3	3	9	0.05%	29.80%	VALLSLLR	95.00%	2.51	0.338	0	4	0	0	2	885.2484	32	39
2DLC TX114	hypo protein Rv3046c	Rv3046c	13333	100.00%	3	3	9	0.05%	29.80%	WLQVGYPEGVPGPDR	95.00%	3.03	0.459	0	3	0	0	2	1,670.77	17	31
2DLC TX114	hspX (hspX)	Rv2031c	16210	100.00%	18	28	214	1.23%	71.50%	AELPGVDPDK	95.00%	1.95	0.415	0	1	0	0	2	1,041.16	55	64
2DLC TX114	hspX (hspX)	Rv2031c	16210	100.00%	18	28	214	1.23%	71.50%	AELPGVDPDKDVDIMVR	95.00%	4.62	0.686	1	22	36	0	2	1,870.10	55	71
2DLC TX114	hspX (hspX)	Rv2031c	16210	100.00%	18	28	214	1.23%	71.50%	DFDGRSEFAYGSFVR	95.00%	2.97	0.349	0	0	1	0	2	1,753.93	86	100
2DLC TX114	hspX (hspX)	Rv2031c	16210	100.00%	18	28	214	1.23%	71.50%	DGQLTIK	95.00%	2.43	0.478	12	26	0	0	2	774.9684	72	78
2DLC TX114	hspX (hspX)	Rv2031c	16210	100.00%	18	28	214	1.23%	71.50%	DGQLTIKAER	95.00%	2.18	0.375	0	1	2	0	2	1,131.35	72	81
2DLC TX114	hspX (hspX)	Rv2031c	16210	100.00%	18	28	214	1.23%	71.50%	GILTVSVAVSEGK	95.00%	2.39	0.402	0	0	9	0	2	1,260.37	120	132
2DLC TX114	hspX (hspX)	Rv2031c	16210	100.00%	18	28	214	1.23%	71.50%	GILTVSVAVSEGKPTEK	95.00%	3.23	0.587	0	2	3	0	2	1,715.89	120	136
2DLC TX114	hspX (hspX)	Rv2031c	16210	100.00%	18	28	214	1.23%	71.50%	HIQIR	95.00%	1.38	0.206	4	0	0	0	2	666.8184	137	141
2DLC TX114	hspX (hspX)	Rv2031c	16210	100.00%	18	28	214	1.23%	71.50%	LEDEMKEGR	95.00%	2.01	0.249	0	2	0	0	2	1,107.27	42	50
2DLC TX114	hspX (hspX)	Rv2031c	16210	100.00%	18	28	214	1.23%	71.50%	LEDEMKEGRYEVR	95.00%	2.55	0.227	0	0	2	0	2	1,654.84	42	54
2DLC TX114	hspX (hspX)	Rv2031c	16210	100.00%	18	28	214	1.23%	71.50%	LMRLEDEMK	95.00%	3.48	0.422	1	10	13	0	2	1,165.52	39	47
2DLC TX114	hspX (hspX)	Rv2031c	16210	100.00%	18	28	214	1.23%	71.50%	LMRLEDEMKEGR	95.00%	2.63	0.377	0	0	1	0	2	1,523.87	39	50
2DLC TX114	hspX (hspX)	Rv2031c	16210	100.00%	18	28	214	1.23%	71.50%	MRLEDEMK	95.00%	1.76	0.258	0	1	0	0	2	1,052.32	40	47
2DLC TX114	hspX (hspX)	Rv2031c	16210	100.00%	18	28	214	1.23%	71.50%	PGVDPDKDVDIMVR	95.00%	4	0.493	0	2	2	0	2	1,556.72	58	71
2DLC TX114	hspX (hspX)	Rv2031c	16210	100.00%	18	28	214	1.23%	71.50%	SEFAYGSFVR	95.00%	3.45	0.583	10	41	0	0	2	1,163.28	91	100
2DLC TX114	hspX (hspX)	Rv2031c	16210	100.00%	18	28	214	1.23%	71.50%	TEQKDFDGR	95.00%	2.18	0.256	0	1	0	0	2	1,096.17	82	90
2DLC TX114	hspX (hspX)	Rv2031c	16210	100.00%	18	28	214	1.23%	71.50%	TVSLPVGADEDDIK	95.00%	3.62	0.53	0	2	3	0	2	1,459.59	101	114
2DLC TX114	hspX (hspX)	Rv2031c	16210	100.00%	18	28	214	1.23%	71.50%	TVSLPVGADEDDIKATYDK	95.00%	5.44	0.678	0	3	0	0	2	2,038.27	101	119
2DLC TX114	cysM2	Rv1077	48617	100.00%	3	3	5	0.03%	8.62%	LADAVSAHMSPPLR	95.00%	3.01	0.512	0	0	1	0	2	1,465.73	402	415
2DLC TX114	cysM2	Rv1077	48617	100.00%	3	3	5	0.03%	8.62%	MIGAGELVSAAGK	95.00%	2.99	0.526	0	3	0	0	2	1,204.46	416	428
2DLC TX114	cysM2	Rv1077	48617	100.00%	3	3	5	0.03%	8.62%	YDLLGFLSEGAGR	95.00%	2.66	0.331	0	1	0	0	2	1,398.73	451	463
2DLC TX114	ephD (ephD)	Rv2214c	64011	99.50%	2	3	13	0.08%	3.04%	ETALAFAR	90.90%	1.59	0.237	0	1	0	0	2	879.0584	339	346
2DLC TX114	ephD (ephD)	Rv2214c	64011	99.50%	2	3	13	0.08%	3.04%	VADAIVSAVK	95.00%	1.92	0.086	3	9	0	0	2	973.0484	548	557
2DLC TX114	eno (eno)	Rv1023	44912	100.00%	6	9	16	0.09%	16.80%	AAADSAELPLFR	95.00%	3	0.449	0	2	1	0	2	1,261.54	120	131
2DLC TX114	eno (eno)	Rv1023	44912	100.00%	6	9	16	0.09%	16.80%	DLISR	95.00%	1.59	0.152	2	0	0	0	2	603.7984	219	223
2DLC TX114	eno (eno)	Rv1023	44912	100.00%	6	9	16	0.09%	16.80%	GVANALLVK	95.00%	2.76	0.359	0	3	0	0	2	885.0684	327	335
2DLC TX114	eno (eno)	Rv1023	44912	100.00%	6	9	16	0.09%	16.80%	IEEALGDAAR	95.00%	3.21	0.427	0	3	0	0	2	1,045.21	405	414
2DLC TX114	eno (eno)	Rv1023	44912	100.00%	6	9	16	0.09%	16.80%	LVDQALVDLDGTPDKSR	95.00%	4.07	0.574	0	1	2	0	2	1,843.09	88	104
2DLC TX114	eno (eno)	Rv1023	44912	100.00%	6	9	16	0.09%	16.80%	WGAEVYHALK	95.00%	3.32	0.498	0	2	1	0	2	1,174.30	180	189
2DLC TX114	eno (eno)	Rv1023	44912	100.00%	6	9	16	0.09%	16.80%	YAGDLAFPR	95.00%	3.06	0.564	0	1	0	0	2	1,010.25	415	423
2DLC TX114	hypo protein Rv0283	Rv0283	55924	100.00%	4	5	11	0.06%	9.11%	ADALLAHDTLVPDSR	95.00%	3.57	0.51	0	3	0	0	2	1,594.83	512	526
2DLC TX114	hypo protein Rv0283	Rv0283	55924	100.00%	4	5	11	0.06%	9.11%	GNLIGIPGAPER	95.00%	2.83	0.527	0	1	0	0	2	1,194.49	156	167
2DLC TX114	hypo protein Rv0283	Rv0283	55924	100.00%	4	5	11	0.06%	9.11%	LGADEVAKLPVSR	95.00%	3.44	0.434	0	1	2	0	2	1,355.57	351	363
2DLC TX114	hypo protein Rv0283	Rv0283	55924	100.00%	4	5	11	0.06%	9.11%	STELDQFPR	95.00%	2.71	0.439	0	4	0	0	2	1,093.22	147	155
2DLC TX114	gcvT (gcvT)	Rv2211c	39579	99.80%	2	3	4	0.02%	8.18%	GPGAAQFVNSALTNDLGR	95.00%	4.42	0.599	0	2	0	0	2	1,788.98	76	93
2DLC TX114	gcvT (gcvT)	Rv2211c	39579	99.80%	2	3	4	0.02%	8.18%	TAVGLFDVSHLGK	95.00%	3.91	0.546	0	1	1	0	2	1,344.52	59	71
2DLC TX114	fadD15 (fadD15)	Rv2187	63985	99.80%	2	2	4	0.02%	3.67%	IADDGELLVR	95.00%	3.16	0.238	0	2	0	0	2	1,101.32	408	417
2DLC TX114	fadD15 (fadD15)	Rv2187	63985	99.80%	2	2	4	0.02%	3.67%	NVAPAVLEDQLR	95.00%	1.83	0.466	0	2	0	0	2	1,325.44	472	483
2DLC TX114	hypo protein Rv2971	Rv2971	30314	100.00%	3	4	7	0.04%	12.40%	LIDTAYAYGNEAAVGR	95.00%	4.03	0.491	0	2	2	0	2	1,684.91	50	65
2DLC TX114	hypo protein Rv2971	Rv2971	30314	100.00%	3	4	7	0.04%	12.40%	TPAQVLLR	95.00%	2.93	0.196	0	2	0	0	2	898.0884	218	225
2DLC TX114	hypo protein Rv2971	Rv2971	30314	100.00%	3	4	7	0.04%	12.40%	VREDPLTYAGT	93.20%	1.28	0.37	0	1	0	0	2	1,222.34	272	282

2DLC TX114	hypo protein Rv2159c	Rv2159c	36358	100.00%	8	9	20	0.12%	27.30%	AGGLVFAR	95.00%	2.32	0.366	0	2	0	0	2	790.9484	180	187
2DLC TX114	hypo protein Rv2159c	Rv2159c	36358	100.00%	8	9	20	0.12%	27.30%	ETLLVGQVPR	95.00%	2.57	0.347	0	4	0	0	2	1,112.23	56	65
2DLC TX114	hypo protein Rv2159c	Rv2159c	36358	100.00%	8	9	20	0.12%	27.30%	FVNHIEPVAPR	95.00%	3.09	0.592	0	2	4	0	2	1,279.38	3	13
2DLC TX114	hypo protein Rv2159c	Rv2159c	36358	100.00%	8	9	20	0.12%	27.30%	IGTWIGAAAEGQVSR	95.00%	4.16	0.545	0	1	0	0	2	1,516.66	325	339
2DLC TX114	hypo protein Rv2159c	Rv2159c	36358	100.00%	8	9	20	0.12%	27.30%	MRRAGGLVFARK	89.90%	2.32	0.17	0	1	0	0	2	1,378.75	177	188
2DLC TX114	hypo protein Rv2159c	Rv2159c	36358	100.00%	8	9	20	0.12%	27.30%	RAGGAVAEVYAEAR	95.00%	3.29	0.438	0	0	4	0	2	1,420.46	14	27
2DLC TX114	hypo protein Rv2159c	Rv2159c	36358	100.00%	8	9	20	0.12%	27.30%	TGLAPHQVTDDDVAAAR	95.00%	3.56	0.487	0	0	1	0	2	1,737.75	285	301
2DLC TX114	hypo protein Rv2159c	Rv2159c	36358	100.00%	8	9	20	0.12%	27.30%	VVGSWHGEPMPMSSR	94.50%	2.49	0.145	0	0	1	0	2	1,657.74	248	262
2DLC TX114	hypo protein Rv2953	Rv2953	45088	99.80%	2	2	2	0.01%	4.07%	IALAGR	95.00%	1.08	0.0072	1	0	0	0	2	600.8284	36	41
2DLC TX114	hypo protein Rv2953	Rv2953	45088	99.80%	2	2	2	0.01%	4.07%	SNALQEWAYGR	95.00%	1.74	0.337	0	1	0	0	2	1,295.41	261	271
2DLC TX114	secD (secD)	Rv2587c	60218	100.00%	3	5	12	0.07%	5.58%	EALAQAQQIISAR	95.00%	2.77	0.375	2	3	2	0	2	1,399.62	65	77
2DLC TX114	secD (secD)	Rv2587c	60218	100.00%	3	5	12	0.07%	5.58%	LGIDLQGGTR	95.00%	2.63	0.389	0	2	0	0	2	1,030.27	40	49
2DLC TX114	secD (secD)	Rv2587c	60218	100.00%	3	5	12	0.07%	5.58%	TPDGSAPSR	95.00%	2.37	0.335	0	3	0	0	2	887.9484	56	64
2DLC TX114	hypo protein Rv0497	Rv0497	33073	100.00%	3	3	5	0.03%	14.50%	QISVAELLAR	95.00%	2.51	0.42	0	3	0	0	2	1,100.33	15	24
2DLC TX114	hypo protein Rv0497	Rv0497	33073	100.00%	3	3	5	0.03%	14.50%	TVYWSQPEPR	95.00%	2.41	0.508	0	1	0	0	2	1,263.31	105	114
2DLC TX114	hypo protein Rv0497	Rv0497	33073	100.00%	3	3	5	0.03%	14.50%	VQVGEAAPQSPAEPVAEQV AEEPTR	95.00%	3.25	0.472	0	0	1	0	2	2,590.41	80	104
2DLC TX114	sigH (sigH)	Rv3223c	24208	99.70%	2	2	2	0.01%	8.80%	GLLADVAR	94.70%	2.11	0.19	0	1	0	0	2	814.9984	193	200
2DLC TX114	sigH (sigH)	Rv3223c	24208	99.70%	2	2	2	0.01%	8.80%	ILTNTYINSYR	95.00%	3.24	0.468	0	1	0	0	2	1,358.70	85	95
2DLC TX114	drrA (drrA)	Rv2936	35801	99.80%	2	2	3	0.02%	5.74%	IELADIALR	95.00%	3.62	0.295	0	2	0	0	2	1,014.38	295	303
2DLC TX114	drrA (drrA)	Rv2936	35801	99.80%	2	2	3	0.02%	5.74%	IVALDDVSFK	95.00%	2.32	0.506	0	1	0	0	2	1,107.32	22	31
2DLC TX114	ppsE (ppsE)	Rv2935	158695	100.00%	4	4	6	0.03%	3.16%	DAGVSDKTLADPAYVR	95.00%	3.29	0.29	0	0	2	0	2	1,678.85	47	62
2DLC TX114	ppsE (ppsE)	Rv2935	158695	100.00%	4	4	6	0.03%	3.16%	LSDVAYTLAR	95.00%	2.62	0.424	0	1	0	0	2	1,109.33	487	496
2DLC TX114	ppsE (ppsE)	Rv2935	158695	100.00%	4	4	6	0.03%	3.16%	LTEAGIPVR	95.00%	2.23	0.346	0	2	0	0	2	956.1384	723	731
2DLC TX114	ppsE (ppsE)	Rv2935	158695	100.00%	4	4	6	0.03%	3.16%	RIPAATAEALER	95.00%	2.09	0.359	0	0	1	0	2	1,298.56	1433	1444
2DLC TX114	hypo protein Rv1006	Rv1006	61280	99.80%	2	2	3	0.02%	7.41%	DGRDDSGFYPVATLDIPRD AQTAVNAS	95.00%	2.48	0.442	0	0	1	0	2	2,853.08	541	567
2DLC TX114	hypo protein Rv1006	Rv1006	61280	99.80%	2	2	3	0.02%	7.41%	SSDQPTPASVAETVR	95.00%	2.48	0.427	0	2	0	0	2	1,545.50	487	501
2DLC TX114	hypo protein Rv0475	Rv0475	21517	100.00%	14	22	92	0.53%	44.20%	AAEGYLEAATSR	95.00%	1.72	0.245	2	4	1	0	2	1,239.37	83	94
2DLC TX114	hypo protein Rv0475	Rv0475	21517	100.00%	14	22	92	0.53%	44.20%	AAKLVGIELPK	95.00%	3.45	0.365	0	3	0	0	2	1,139.52	151	161
2DLC TX114	hypo protein Rv0475	Rv0475	21517	100.00%	14	22	92	0.53%	44.20%	ATVNELITNLR	95.00%	3.17	0.468	0	3	0	0	2	1,244.47	26	36
2DLC TX114	hypo protein Rv0475	Rv0475	21517	100.00%	14	22	92	0.53%	44.20%	FTAEELR	95.00%	1.42	0.241	1	11	0	0	2	865.9984	75	81
2DLC TX114	hypo protein Rv0475	Rv0475	21517	100.00%	14	22	92	0.53%	44.20%	FTAEELRK	95.00%	1.96	0.468	0	0	2	0	2	994.1984	75	82
2DLC TX114	hypo protein Rv0475	Rv0475	21517	100.00%	14	22	92	0.53%	44.20%	KAAAKK	95.00%	1.09	0.0867	1	0	0	0	2	616.8584	180	185
2DLC TX114	hypo protein Rv0475	Rv0475	21517	100.00%	14	22	92	0.53%	44.20%	KAAEGYLEAATSR	95.00%	4	0.65	0	5	2	0	2	1,367.57	82	94
2DLC TX114	hypo protein Rv0475	Rv0475	21517	100.00%	14	22	92	0.53%	44.20%	LQEDLPEQLTELR	95.00%	3.94	0.57	1	8	1	0	2	1,584.84	60	72
2DLC TX114	hypo protein Rv0475	Rv0475	21517	100.00%	14	22	92	0.53%	44.20%	LRSQQSFEEVSAR	95.00%	3.78	0.442	0	0	1	0	2	1,537.61	109	121
2DLC TX114	hypo protein Rv0475	Rv0475	21517	100.00%	14	22	92	0.53%	44.20%	LVGIELPK	95.00%	1.89	0.265	4	6	0	0	2	869.1584	154	161
2DLC TX114	hypo protein Rv0475	Rv0475	21517	100.00%	14	22	92	0.53%	44.20%	SQQSFEEVSAR	95.00%	2.86	0.512	0	6	0	0	2	1,268.21	111	121
2DLC TX114	hypo protein Rv0475	Rv0475	21517	100.00%	14	22	92	0.53%	44.20%	VGIELPK	95.00%	1.85	0.166	2	0	0	0	2	755.9584	155	161
2DLC TX114	hypo protein Rv0475	Rv0475	21517	100.00%	14	22	92	0.53%	44.20%	YNELVER	95.00%	2.5	0.499	0	8	0	0	2	922.9884	95	101
2DLC TX114	hypo protein Rv0475	Rv0475	21517	100.00%	14	22	92	0.53%	44.20%	YNELVERGEAALER	95.00%	3.87	0.547	0	6	14	0	2	1,649.80	95	108
2DLC TX114	hypo protein Rv0474	Rv0474	15346	100.00%	3	3	8	0.05%	25.00%	ASDVASDIGSFIR	95.00%	3.48	0.61	0	1	0	0	2	1,338.54	15	27
2DLC TX114	hypo protein Rv0474	Rv0474	15346	100.00%	3	3	8	0.05%	25.00%	SGVSNPYLSQVER	95.00%	3.9	0.553	0	4	0	0	2	1,436.47	44	56
2DLC TX114	hypo protein Rv0474	Rv0474	15346	100.00%	3	3	8	0.05%	25.00%	VSAEVLYVR	95.00%	3.12	0.447	0	3	0	0	2	1,036.09	75	83
2DLC TX114	fadB2	Rv0468	30678	99.80%	2	2	3	0.02%	5.59%	GFYTY	95.00%	0.783	0.0972	1	0	0	0	2	650.7684	282	286
2DLC TX114	fadB2	Rv0468	30678	99.80%	2	2	3	0.02%	5.59%	LSDLVGLDTLK	95.00%	3.51	0.539	0	2	0	0	2	1,174.52	235	245
2DLC TX114	lsr2 (lsr2)	Rv3597c	12081	100.00%	3	4	12	0.07%	24.10%	GAIDREQSAAIR	95.00%	2.86	0.393	0	0	2	0	2	1,287.49	73	84
2DLC TX114	lsr2 (lsr2)	Rv3597c	12081	100.00%	3	4	12	0.07%	24.10%	GRIPADVIDAYHAAT	95.00%	2.18	0.347	0	4	2	0	2	1,570.78	98	112
2DLC TX114	lsr2 (lsr2)	Rv3597c	12081	100.00%	3	4	12	0.07%	24.10%	IPADVIDAYHAAT	95.00%	3	0.482	0	4	0	0	2	1,357.53	100	112
2DLC TX114	clpC (clpC)	Rv3596c	93537	100.00%	12	12	23	0.13%	12.40%	AIDLIDEAGAR	95.00%	3.92	0.327	0	1	0	0	2	1,144.41	399	409
2DLC TX114	clpC (clpC)	Rv3596c	93537	100.00%	12	12	23	0.13%	12.40%	AIDLIDEAGARMRIR	89.60%	2.39	0.176	0	0	1	0	2	1,717.21	399	413
2DLC TX114	clpC (clpC)	Rv3596c	93537	100.00%	12	12	23	0.13%	12.40%	FLPDKAIDLIDEAGAR	95.00%	3.8	0.518	0	2	0	0	2	1,745.23	394	409
2DLC TX114	clpC (clpC)	Rv3596c	93537	100.00%	12	12	23	0.13%	12.40%	GELQTIGATTLDEYR	95.00%	2.72	0.576	0	2	0	0	2	1,667.90	315	329
2DLC TX114	clpC (clpC)	Rv3596c	93537	100.00%	12	12	23	0.13%	12.40%	GGESGSPSTSLVLDQFGR	95.00%	3.89	0.641	0	3	0	0	2	1,794.93	159	176
2DLC TX114	clpC (clpC)	Rv3596c	93537	100.00%	12	12	23	0.13%	12.40%	NLTAAAMEGK	92.90%	1.63	0.246	0	2	0	0	2	1,006.21	177	186
2DLC TX114	clpC (clpC)	Rv3596c	93537	100.00%	12	12	23	0.13%	12.40%	NLTAAAMEGKLDPVIGR	85.70%	2.16	0.175	0	0	1	0	2	1,773.15	177	193
2DLC TX114	clpC (clpC)	Rv3596c	93537	100.00%	12	12	23	0.13%	12.40%	NNPVLIGEPGVGK	95.00%	2.5	0.502	0	4	0	0	2	1,294.45	210	222
2DLC TX114	clpC (clpC)	Rv3596c	93537	100.00%	12	12	23	0.13%	12.40%	NPVLIGEPGVGK	95.00%	2.59	0.556	0	2	0	0	2	1,180.35	211	222
2DLC TX114	clpC (clpC)	Rv3596c	93537	100.00%	12	12	23	0.13%	12.40%	SLESLGISLEGVR	95.00%	4.15	0.448	0	2	0	0	2	1,360.63	48	60
2DLC TX114	clpC (clpC)	Rv3596c	93537	100.00%	12	12	23	0.13%	12.40%	TKNNPVLIGEPGVGK	95.00%	3.18	0.287	0	0	2	0	2	1,523.75	208	222
2DLC TX114	clpC (clpC)	Rv3596c	93537	100.00%	12	12	23	0.13%	12.40%	VMQVLSR	95.00%	2.11	0.375	0	1	0	0	2	832.9384	200	206
2DLC TX114	hypo protein Rv2280	Rv2280	48035	100.00%	3	3	3	0.02%	10.00%	ALGADDIIDTVVPR	95.00%	2.58	0.212	0	1	0	0	2	1,455.69	336	349

2DLC TX114	hypo protein Rv2280	Rv2280	48035	100.00%	3	3	3	0.02%	10.00%	IKQSFDPAGILNPGVVFGDT	95.00%	2.71	0.417	0	1	0	0	2	2,076.41	440	459
2DLC TX114	hypo protein Rv2280	Rv2280	48035	100.00%	3	3	3	0.02%	10.00%	TGYFLELEDPVK	95.00%	4.09	0.41	0	1	0	0	2	1,411.66	422	433
2DLC TX114	hypo protein Rv1269c	Rv1269c	12532	100.00%	3	5	12	0.07%	23.40%	AAAEATAVK	95.00%	2.22	0.374	0	2	0	0	2	831.8884	60	68
2DLC TX114	hypo protein Rv1269c	Rv1269c	12532	100.00%	3	5	12	0.07%	23.40%	AYQGGVGPTLAAAMK	95.00%	3.01	0.628	0	4	2	0	2	1,435.68	93	107
2DLC TX114	hypo protein Rv1269c	Rv1269c	12532	100.00%	3	5	12	0.07%	23.40%	AYQGGVGPTLAAAMKDALT K	95.00%	3.63	0.61	0	2	2	0	2	1,964.36	93	112
2DLC TX114	sodC (sodC)	Rv0432	23826	99.80%	2	2	4	0.02%	9.58%	YVQVNGTPGPDETTLTTGD AGK	95.00%	4.01	0.637	0	2	0	0	2	2,222.28	209	230
2DLC TX114	sodC (sodC)	Rv0432	23826	99.80%	2	2	4	0.02%	9.58%	YVQVNGTPGPDETTLTTGD AGKR	95.00%	2.14	0.431	0	0	2	0	2	2,378.48	209	231
2DLC TX114	hypo protein Rv0431	Rv0431	16887	99.80%	2	4	10	0.06%	13.40%	TLGAAVELR	95.00%	1.34	0.262	1	3	0	0	2	930.0984	139	147
2DLC TX114	hypo protein Rv0431	Rv0431	16887	99.80%	2	4	10	0.06%	13.40%	VYNISGTEGAAAR	95.00%	3.08	0.479	0	5	1	0	2	1,309.41	81	93
2DLC TX114	lppS (lppS)	Rv2518c	43347	99.80%	2	2	5	0.03%	6.62%	WSTTEQLGYNR	95.00%	2.68	0.562	0	2	0	0	2	1,355.45	112	122
2DLC TX114	lppS (lppS)	Rv2518c	43347	99.80%	2	2	5	0.03%	6.62%	YTLNATALGLGGAATR	95.00%	5	0.631	0	3	0	0	2	1,550.89	124	139
2DLC TX114	hypo protein Rv2260	Rv2260	22119	99.70%	2	2	2	0.01%	13.70%	GTALDAPVLMHPGDAVLWR	93.70%	2.28	0.201	0	1	0	0	2	2,020.34	76	94
2DLC TX114	hypo protein Rv2260	Rv2260	22119	99.70%	2	2	2	0.01%	13.70%	SYSDFPTILR	95.00%	2.13	0.316	0	1	0	0	2	1,199.50	162	171
2DLC TX114	embR	Rv1267c	41916	99.30%	2	2	2	0.01%	6.19%	ASAVIAELEAL	85.10%	2.49	0.137	0	1	0	0	2	1,087.29	187	197
2DLC TX114	embR	Rv1267c	41916	99.30%	2	2	2	0.01%	6.19%	TTAAGTVTVLDQR	95.00%	2.87	0.35	0	1	0	0	2	1,333.37	268	280
2DLC TX114	lppX (lppX)	Rv2945c	24122	100.00%	6	7	14	0.08%	33.90%	ANPLAAK	95.00%	1.16	0.271	1	0	0	0	2	684.8784	92	98
2DLC TX114	lppX (lppX)	Rv2945c	24122	100.00%	6	7	14	0.08%	33.90%	ASIDLGSGSIQLTQSK	95.00%	4.77	0.695	0	3	0	0	2	1,605.92	210	225
2DLC TX114	lppX (lppX)	Rv2945c	24122	100.00%	6	7	14	0.08%	33.90%	ITGTIPASSVK	95.00%	1.88	0.428	0	1	0	0	2	1,074.30	173	183
2DLC TX114	lppX (lppX)	Rv2945c	24122	100.00%	6	7	14	0.08%	33.90%	LFDDWSNLGSISELSTSR	95.00%	6.1	0.6	0	5	0	0	2	2,028.37	122	139
2DLC TX114	lppX (lppX)	Rv2945c	24122	100.00%	6	7	14	0.08%	33.90%	TSVHVAVR	95.00%	2.33	0.33	0	1	0	0	2	868.7884	65	72
2DLC TX114	lppX (lppX)	Rv2945c	24122	100.00%	6	7	14	0.08%	33.90%	TTGKVDSLLGITSADVDVR	95.00%	5.41	0.519	0	2	1	0	2	1,948.17	73	91
2DLC TX114	hypo protein Rv0407	Rv0407	36940	100.00%	3	3	4	0.02%	7.44%	FLELFQSDLAPR	95.00%	3.99	0.548	0	2	0	0	2	1,436.80	320	331
2DLC TX114	hypo protein Rv0407	Rv0407	36940	100.00%	3	3	4	0.02%	7.44%	FWAPLSLTAEQK	95.00%	2.3	0.214	0	1	0	0	2	1,391.68	248	259
2DLC TX114	hypo protein Rv0407	Rv0407	36940	100.00%	3	3	4	0.02%	7.44%	RFLELFQSDLAPR	91.00%	1.77	0.179	0	0	1	0	2	1,593.00	319	331
2DLC TX114	pknH (pknH)	Rv1266c	66735	100.00%	3	3	4	0.02%	7.51%	FGPLTPPR	95.00%	1.76	0.294	0	1	0	0	2	885.1284	108	115
2DLC TX114	pknH (pknH)	Rv1266c	66735	100.00%	3	3	4	0.02%	7.51%	LMTAEFSKDPVFR	95.00%	2.31	0.345	0	0	1	0	2	1,541.87	46	58
2DLC TX114	pknH (pknH)	Rv1266c	66735	100.00%	3	3	4	0.02%	7.51%	SQESTLPAPPKPVPPPTMPA TAMAPR	95.00%	3.84	0.473	0	0	2	0	2	2,671.15	291	316
2DLC TX114	accD6	Rv2247	50117	100.00%	4	4	9	0.05%	10.40%	ALHAVGQVFEAMIR	95.00%	3.35	0.443	0	0	1	0	2	1,542.75	107	120
2DLC TX114	accD6	Rv2247	50117	100.00%	4	4	9	0.05%	10.40%	SGVLAAAGTVNGVR	95.00%	2.86	0.521	0	2	0	0	2	1,272.30	40	53
2DLC TX114	accD6	Rv2247	50117	100.00%	4	4	9	0.05%	10.40%	TVGVLANNPLR	95.00%	3.25	0.502	0	2	0	0	2	1,154.31	280	290
2DLC TX114	accD6	Rv2247	50117	100.00%	4	4	9	0.05%	10.40%	VFVTGPDVVR	95.00%	3.01	0.333	0	4	0	0	2	1,089.07	157	166
2DLC TX114	kasB (kasB)	Rv2246	46371	100.00%	5	5	8	0.05%	14.20%	IEAVPIAGFAQMR	95.00%	2.73	0.612	0	2	0	0	2	1,403.70	222	234
2DLC TX114	kasB (kasB)	Rv2246	46371	100.00%	5	5	8	0.05%	14.20%	IMGASITSDGFHMVAPDPN GER	95.00%	2.7	0.377	0	0	2	0	2	2,303.60	286	307
2DLC TX114	kasB (kasB)	Rv2246	46371	100.00%	5	5	8	0.05%	14.20%	MGYLQR	86.90%	1.64	0.213	0	1	0	0	2	767.9684	100	105
2DLC TX114	kasB (kasB)	Rv2246	46371	100.00%	5	5	8	0.05%	14.20%	MSTVLSR	95.00%	2.08	0.288	0	2	0	0	2	793.9484	106	112
2DLC TX114	kasB (kasB)	Rv2246	46371	100.00%	5	5	8	0.05%	14.20%	YMPNGAAAAVGLER	95.00%	3.11	0.533	0	1	0	0	2	1,420.63	165	178
2DLC TX114	kasA (kasA)	Rv2245	43266	100.00%	3	6	17	0.10%	13.20%	GLLAGESGIHALEDEFVTK	95.00%	3.98	0.647	0	1	1	0	2	1,987.28	35	53
2DLC TX114	kasA (kasA)	Rv2245	43266	100.00%	3	6	17	0.10%	13.20%	IVESYDLMNAGGPR	95.00%	3.18	0.481	0	2	2	0	2	1,522.77	122	135
2DLC TX114	kasA (kasA)	Rv2245	43266	100.00%	3	6	17	0.10%	13.20%	LLGAGITSDAFHMVAPAAD GVR	95.00%	6.75	0.688	0	4	7	0	2	2,170.51	265	286
2DLC TX114	acpM (acpM)	Rv2244	12474	100.00%	7	13	42	0.24%	47.80%	AKIESENPDAVANVQAR	95.00%	4.73	0.583	0	4	8	0	2	1,812.88	93	109
2DLC TX114	acpM (acpM)	Rv2244	12474	100.00%	7	13	42	0.24%	47.80%	IESENPDAVANVQAR	95.00%	4.77	0.399	0	2	3	0	2	1,613.60	95	109
2DLC TX114	acpM (acpM)	Rv2244	12474	100.00%	7	13	42	0.24%	47.80%	IPPDLAGLR	95.00%	1.56	0.349	2	4	0	0	2	1,099.37	59	68
2DLC TX114	acpM (acpM)	Rv2244	12474	100.00%	7	13	42	0.24%	47.80%	LEEENPEAAQALR	95.00%	3.96	0.382	0	1	1	0	2	1,470.58	80	92
2DLC TX114	acpM (acpM)	Rv2244	12474	100.00%	7	13	42	0.24%	47.80%	TVGDVVAYIQK TVGDVVAYIQKLEEENPEAA	95.00%	4.31	0.511	2	5	0	0	2	1,193.26	69	79
2DLC TX114	acpM (acpM)	Rv2244	12474	100.00%	7	13	42	0.24%	47.80%	QALR	95.00%	5.8	0.643	0	7	2	0	2	2,644.82	69	92
2DLC TX114	acpM (acpM)	Rv2244	12474	100.00%	7	13	42	0.24%	47.80%	YGVKIPDEDLAGLR	95.00%	4.5	0.341	0	0	1	0	2	1,546.89	55	68
2DLC TX114	aceE (aceE)	Rv2241	100168	100.00%	4	4	6	0.03%	4.44%	ALGGFVPER	95.00%	2.27	0.32	0	1	0	0	2	946.0884	462	470
2DLC TX114	aceE (aceE)	Rv2241	100168	100.00%	4	4	6	0.03%	4.44%	GFVLGATAGR	95.00%	2.85	0.429	0	1	0	0	2	949.0984	626	635
2DLC TX114	aceE (aceE)	Rv2241	100168	100.00%	4	4	6	0.03%	4.44%	GPVLAVSDWMR	95.00%	3.03	0.515	0	3	0	0	2	1,231.39	804	814
2DLC TX114	aceE (aceE)	Rv2241	100168	100.00%	4	4	6	0.03%	4.44%	IIQELESFFR	95.00%	3.41	0.486	0	1	0	0	2	1,282.60	275	284
2DLC TX114	hsp (hsp)	Rv0251c	17768	100.00%	5	5	8	0.05%	32.70%	APAETQAQR LPAHVTSEAIAASYDAGVLT	95.00%	2.84	0.554	0	3	0	0	2	971.9784	146	154
2DLC TX114	hsp (hsp)	Rv0251c	17768	100.00%	5	5	8	0.05%	32.70%	VR	95.00%	2.83	0.489	0	0	1	0	2	2,242.46	118	139
2DLC TX114	hsp (hsp)	Rv0251c	17768	100.00%	5	5	8	0.05%	32.70%	PVWDVEPWDR	95.00%	2.71	0.474	0	2	0	0	2	1,299.30	10	19
2DLC TX114	hsp (hsp)	Rv0251c	17768	100.00%	5	5	8	0.05%	32.70%	VAGAYKAPAETQAQR	95.00%	3.13	0.504	0	0	1	0	2	1,561.66	140	154
2DLC TX114	hsp (hsp)	Rv0251c	17768	100.00%	5	5	8	0.05%	32.70%	YGSFR	83.60%	1.22	0.337	0	1	0	0	2	629.7484	109	113
2DLC TX114	hypo protein Rv2095c	Rv2095c	33747	99.80%	2	2	2	0.01%	6.33%	LLLGFGSDVR	95.00%	2.8	0.4	0	1	0	0	2	1,077.37	279	288
2DLC TX114	hypo protein Rv2095c	Rv2095c	33747	99.80%	2	2	2	0.01%	6.33%	VLAPESLAQR	95.00%	2.25	0.294	0	1	0	0	2	1,084.25	289	298
2DLC TX114	greA (greA)	Rv1080c	17837	99.60%	2	2	3	0.02%	13.40%	QLQDLLSNAK	95.00%	1.79	0.379	0	1	0	0	2	1,130.38	66	75
2DLC TX114	greA (greA)	Rv1080c	17837	99.60%	2	2	3	0.02%	13.40%	QSGVALPGSVVK	91.20%	2.02	0.223	0	2	0	0	2	1,142.19	82	93
2DLC TX114	glnA1	Rv2220	53522	99.80%	2	2	3	0.02%	6.49%	LVPGYEAPINLVYSQR	95.00%	3.38	0.492	0	0	1	0	2	1,820.11	330	345
2DLC TX114	glnA1	Rv2220	53522	99.80%	2	2	3	0.02%	6.49%	SVFDDGLAFDGSSIR	95.00%	3.62	0.578	0	2	0	0	2	1,586.81	45	59
2DLC TX114	hypo protein Rv2216	Rv2216	31654	100.00%	3	3	4	0.02%	12.30%	AAFGEFADEGLLIGQR	95.00%	2.44	0.259	0	1	0	0	2	1,695.01	256	271
2DLC TX114	hypo protein Rv2216	Rv2216	31654	100.00%	3	3	4	0.02%	12.30%	AVNRPTPL	95.00%	1.81	0.251	0	2	0	0	2	868.0084	240	247
2DLC TX114	hypo protein Rv2216	Rv2216	31654	100.00%	3	3	4	0.02%	12.30%	TGVVLSPAGGMLR AEPAPSPPAAQPAGAAEGA	95.00%	2.31	0.418	0	1	0	0	2	1,258.49	162	174
2DLC TX114	sucB (sucB)	Rv2215	57070	100.00%	6	8	15	0.09%	18.40%	PYVTPLVR	95.00%	3.69	0.532	0	0	2	0	2	2,586.82	223	249
2DLC TX114	sucB (sucB)	Rv2215	57070	100.00%	6	8	15	0.09%	18.40%	AIADIAAR	95.00%	1.96	0.165	2	3	0	0	2	801.0384	440	447
2DLC TX114	sucB (sucB)	Rv2215	57070	100.00%	6	8	15	0.09%	18.40%	AKAPAPAAQAAAAPAPK	95.00%	3.73	0.485	0	0	1	0	2	1,502.80	286	302
2DLC TX114	sucB (sucB)	Rv2215	57070	100.00%	6	8	15	0.09%	18.40%	APPAPAPALAHLR IGVAADIGAAPAPKPAPKPV	95.00%	1.94	0.472	0	2	2	0	2	1,282.60	303	315
2DLC TX114	sucB (sucB)	Rv2215	57070	100.00%	6	8	15	0.09%	18.40%	PEPAPTPK	95.00%	3.02	0.484	0	0	2	0	2	2,659.20	195	222
2DLC TX114	sucB (sucB)	Rv2215	57070	100.00%	6	8	15	0.09%	18.40%	LIDGADAGR	95.00%	2.12	0.309	0	1	0	0	2	888.0784	525	533
2DLC TX114	pepB (pepB)	Rv2213	53431	100.00%	4	4	6	0.03%	10.30%	AGYGGVIGVGQGSSRPPR	95.00%	3.82	0.579	0	0	1	0	2	1,715.84	236	253
2DLC TX114	pepB (pepB)	Rv2213	53431	100.00%	4	4	6	0.03%	10.30%	IPGVMGSDEFRDR	95.00%	2.29	0.189	0	0	1	0	2	1,479.69	401	413

2DLC TX114	pepB (pepB)	Rv2213	53431	100.00%	4	4	6	0.03%	10.30%	LILADAIVR	95.00%	3.07	0.488	0	3	0	0	2	984.3484	366	374
2DLC TX114	pepB (pepB)	Rv2213	53431	100.00%	4	4	6	0.03%	10.30%	STVADLANVSGQR	95.00%	2.68	0.388	0	1	0	0	2	1,318.33	440	452
2DLC TX114	hypo protein Rv2629	Rv2629	40822	100.00%	3	4	12	0.07%	9.89%	DAKQELIDSLEEAVR	95.00%	4.18	0.425	0	6	2	0	2	1,716.93	49	63
2DLC TX114	hypo protein Rv2629	Rv2629	40822	100.00%	3	4	12	0.07%	9.89%	IAPLDGVGALLR	95.00%	2.23	0.572	0	3	0	0	2	1,195.57	351	362
2DLC TX114	Rv2629	Rv2629	40822	100.00%	3	4	12	0.07%	9.89%	TDLLSTLPQR	95.00%	2.81	0.261	0	1	0	0	2	1,144.42	206	215
2DLC TX114	tsf (tsf)	Rv2889c	28737	99.80%	2	2	4	0.02%	10.30%	ALLDVAGVTVTR	95.00%	4.41	0.54	0	2	0	0	2	1,215.34	251	262
2DLC TX114	tsf (tsf)	Rv2889c	28737	99.80%	2	2	4	0.02%	10.30%	YLSRDDVPEDIVASER	95.00%	3.18	0.489	0	0	2	0	2	1,865.02	187	202
2DLC TX114	prsA (prsA)	Rv1017c	35442	99.80%	2	2	2	0.01%	6.75%	LIADLLK	95.00%	2	0.28	0	1	0	0	2	786.1984	119	125
2DLC TX114	prsA (prsA)	Rv1017c	35442	99.80%	2	2	2	0.01%	6.75%	WADALGGVPLAFIHK	95.00%	2.83	0.258	0	0	1	0	2	1,595.95	188	202
2DLC TX114	hypo protein Rv1261c	Rv1261c	16738	99.80%	2	2	4	0.02%	8.05%	PIPVVVLTR	95.00%	2.06	0.628	0	2	0	0	2	994.1684	140	148
2DLC TX114	hypo protein Rv1261c	Rv1261c	16738	99.80%	2	2	4	0.02%	8.05%	TSRPIPVVVLTR	95.00%	3.22	0.291	0	0	2	0	2	1,338.55	137	148
2DLC TX114	amiC (amiC)	Rv2888c	50898	100.00%	3	3	4	0.02%	4.86%	PLAIMR	92.30%	1.06	0.159	1	0	0	0	2	701.0184	355	360
2DLC TX114	amiC (amiC)	Rv2888c	50898	100.00%	3	3	4	0.02%	4.86%	QLTLK	95.00%	1.49	0.188	1	0	0	0	2	602.8184	276	280
2DLC TX114	amiC (amiC)	Rv2888c	50898	100.00%	3	3	4	0.02%	4.86%	VVLGTGLVSLGK	95.00%	2.71	0.413	0	2	0	0	2	1,143.36	113	124
2DLC TX114	hypo protein Rv2091c	Rv2091c	26001	100.00%	3	3	4	0.02%	7.38%	GTYEVGRPQ	95.00%	2.6	0.541	0	2	0	0	2	1,007.01	236	244
2DLC TX114	hypo protein Rv2091c	Rv2091c	26001	100.00%	3	3	4	0.02%	7.38%	VTVTFQDNK	95.00%	2.75	0.602	0	1	0	0	2	1,052.06	227	235
2DLC TX114	hypo protein Rv2091c	Rv2091c	26001	100.00%	3	3	4	0.02%	7.38%	VTVTFQDNKGTYEVGRPQ	95.00%	2.67	0.196	0	0	1	0	2	2,040.05	227	244
2DLC TX114	lpqT (lpqT)	Rv1016c	24045	99.80%	2	2	4	0.02%	10.60%	GESYPTAMLMVFK	95.00%	2.63	0.588	0	1	0	0	2	1,474.82	110	122
2DLC TX114	lpqT (lpqT)	Rv1016c	24045	99.80%	2	2	4	0.02%	10.60%	IVFPTGAPPAK	95.00%	2.51	0.412	0	3	0	0	2	1,098.36	181	191
2DLC TX114	ask (ask)	Rv3709c	44411	99.80%	2	2	4	0.02%	5.70%	VSLIGAGMR	95.00%	2.69	0.455	0	2	0	0	2	904.1484	347	355
2DLC TX114	ask (ask)	Rv3709c	44411	99.80%	2	2	4	0.02%	5.70%	VTIVGLPDIPGYAAK	95.00%	1.96	0.352	0	2	0	0	2	1,514.86	267	281
2DLC TX114	secA (secA)	Rv3240c	105974	100.00%	5	5	12	0.07%	7.38%	FNGAALETLLTR	95.00%	3.3	0.328	0	1	0	0	2	1,306.63	594	605
2DLC TX114	secA (secA)	Rv3240c	105974	100.00%	5	5	12	0.07%	7.38%	NVLLNVIDR	95.00%	3.17	0.342	0	4	0	0	2	1,056.26	769	777
2DLC TX114	secA (secA)	Rv3240c	105974	100.00%	5	5	12	0.07%	7.38%	SAQTQVEQQNFEVR	95.00%	3.57	0.447	0	0	2	0	2	1,664.52	625	638
2DLC TX114	secA (secA)	Rv3240c	105974	100.00%	5	5	12	0.07%	7.38%	TLYPVGITADSLTR	95.00%	3.22	0.509	0	4	0	0	2	1,507.82	707	720
2DLC TX114	secA (secA)	Rv3240c	105974	100.00%	5	5	12	0.07%	7.38%	VADYVGTLSDDVEKLTDAE LR	95.00%	4.35	0.634	0	0	1	0	2	2,310.52	20	40
2DLC TX114	lldD2 (lldD2)	Rv1872c	45325	100.00%	6	7	15	0.09%	13.30%	AIEILQTGVIR	95.00%	4.44	0.524	0	4	0	0	2	1,213.52	373	383
2DLC TX114	lldD2 (lldD2)	Rv1872c	45325	100.00%	6	7	15	0.09%	13.30%	DIEFHPTILR	95.00%	2.75	0.312	0	2	0	0	2	1,241.54	75	84
2DLC TX114	lldD2 (lldD2)	Rv1872c	45325	100.00%	6	7	15	0.09%	13.30%	GVDGIVLSNHGGR	95.00%	2.91	0.429	0	0	2	0	2	1,281.34	292	304
2DLC TX114	lldD2 (lldD2)	Rv1872c	45325	100.00%	6	7	15	0.09%	13.30%	LGPIGAPT	95.00%	1.13	0.465	1	1	0	0	2	725.9384	407	414
2DLC TX114	lldD2 (lldD2)	Rv1872c	45325	100.00%	6	7	15	0.09%	13.30%	RLGAALTIQDLR	95.00%	1.95	0.198	0	0	1	0	2	1,327.73	30	41
2DLC TX114	lldD2 (lldD2)	Rv1872c	45325	100.00%	6	7	15	0.09%	13.30%	RLGPIGAPT	95.00%	1.73	0.399	0	4	0	0	2	882.1384	406	414
2DLC TX114	asd (asd)	Rv3708c	36212	99.80%	2	3	5	0.03%	7.54%	GLALFVSGDNLR	95.00%	3.39	0.529	0	2	0	0	2	1,262.55	316	327
2DLC TX114	asd (asd)	Rv3708c	36212	99.80%	2	3	5	0.03%	7.54%	TLLDERDFPASAVR	95.00%	3.59	0.424	0	1	2	0	2	1,590.85	20	33
2DLC TX114	hypo protein Rv1871c	Rv1871c	14645	100.00%	4	4	6	0.03%	34.90%	AMGTDLLTIR	95.00%	2.91	0.367	0	1	0	0	2	1,091.45	116	125
2DLC TX114	hypo protein Rv1871c	Rv1871c	14645	100.00%	4	4	6	0.03%	34.90%	FVVNPIGR	94.40%	1.98	0.202	0	1	0	0	2	902.0284	18	25
2DLC TX114	hypo protein Rv1871c	Rv1871c	14645	100.00%	4	4	6	0.03%	34.90%	QLPMTMLETIGR	95.00%	2.04	0.448	0	2	0	0	2	1,422.79	26	37
2DLC TX114	hypo protein Rv1871c	Rv1871c	14645	100.00%	4	4	6	0.03%	34.90%	SGTAYLLPDDDPRQR	95.00%	1.82	0.403	0	0	2	0	2	1,704.97	88	102
2DLC TX114	argB (argB)	Rv1654	30918	99.80%	2	2	3	0.02%	7.14%	AVIGGVPSAHIIDGR	95.00%	2.61	0.57	0	2	0	0	2	1,462.67	258	272
2DLC TX114	argB (argB)	Rv1654	30918	99.80%	2	2	3	0.02%	7.14%	VEACLRAVIGGVPSAHIIDG R	95.00%	1.84	0.263	0	0	1	0	2	2,192.45	252	272
2DLC TX114	otsA (otsA)	Rv3490	55845	99.80%	2	2	5	0.03%	5.80%	SPGGLVTALEPVLR	95.00%	4.21	0.612	0	4	0	0	2	1,409.66	46	59
2DLC TX114	otsA (otsA)	Rv3490	55845	99.80%	2	2	5	0.03%	5.80%	WAQSFLDALAGAHPR	95.00%	3.32	0.427	0	0	1	0	2	1,640.89	483	497
2DLC TX114	hypo protein Rv2028c	Rv2028c	29422	99.10%	2	2	2	0.01%	12.50%	LLYAIEPDDPGYAAHGAAAR	95.00%	2.64	0.442	0	0	1	0	2	2,072.44	40	59
2DLC TX114	hypo protein Rv2028c	Rv2028c	29422	99.10%	2	2	2	0.01%	12.50%	QEAGCTLLVVGQQYL	81.40%	1.88	0.197	0	0	1	0	2	1,679.79	265	279
2DLC TX114	hypo protein Rv0559c	Rv0559c	12099	99.50%	2	2	4	0.02%	19.60%	FGIYGPQDYNAWLAK	95.00%	3.36	0.633	0	3	0	0	2	1,744.10	39	53
2DLC TX114	hypo protein Rv0559c	Rv0559c	12099	99.50%	2	2	4	0.02%	19.60%	SATFLQR	89.60%	1.62	0.287	0	1	0	0	2	822.9784	70	76
2DLC TX114	hypo protein Rv1636	Rv1636	15294	100.00%	4	6	13	0.08%	34.20%	LIIASAYLPQHEDAR	95.00%	4.56	0.62	0	4	1	0	2	1,698.06	34	48
2DLC TX114	hypo protein Rv1636	Rv1636	15294	100.00%	4	6	13	0.08%	34.20%	LLGSVPANVSR	95.00%	2.82	0.216	0	3	0	0	2	1,113.27	124	134
2DLC TX114	hypo protein Rv1636	Rv1636	15294	100.00%	4	6	13	0.08%	34.20%	VDVLIVHTT	95.00%	2.89	0.608	0	2	0	0	2	997.0284	138	146
2DLC TX114	hypo protein Rv1636	Rv1636	15294	100.00%	4	6	13	0.08%	34.20%	VTGTAPIYEILHDAK	95.00%	2.79	0.461	0	2	1	0	2	1,628.92	60	74
2DLC TX114	polA (polA)	Rv1629	98425	100.00%	4	4	9	0.05%	4.31%	ALNEAQLASR	95.00%	2.62	0.341	0	2	0	0	2	1,073.24	841	850
2DLC TX114	polA (polA)	Rv1629	98425	100.00%	4	4	9	0.05%	4.31%	FTPEAIVEK	94.90%	1.47	0.306	0	1	0	0	2	1,020.06	174	182
2DLC TX114	polA (polA)	Rv1629	98425	100.00%	4	4	9	0.05%	4.31%	SFTLDDLSLR	95.00%	3.01	0.515	0	4	0	0	2	1,167.48	433	442
2DLC TX114	polA (polA)	Rv1629	98425	100.00%	4	4	9	0.05%	4.31%	SLVDNVDAVR	95.00%	2.81	0.422	0	2	0	0	2	1,088.09	227	236
2DLC TX114	hypo protein Rv1626	Rv1626	22652	99.80%	2	3	5	0.03%	12.20%	RDGIDAASEIASK	95.00%	2.61	0.451	0	1	2	0	2	1,333.57	71	83
2DLC TX114	hypo protein Rv1626	Rv1626	22652	99.80%	2	3	5	0.03%	12.20%	RVLIAEDEALIR	95.00%	3.9	0.374	0	2	0	0	2	1,398.75	15	26
2DLC TX114	hypo protein Rv2030c	Rv2030c	74880	100.00%	8	8	18	0.10%	12.60%	AIGVVYLPATER	95.00%	2.37	0.463	0	2	0	0	2	1,289.49	623	634
2DLC TX114	hypo protein Rv2030c	Rv2030c	74880	100.00%	8	8	18	0.10%	12.60%	ALEPLEVTSR	95.00%	2.51	0.51	0	4	0	0	2	1,115.27	658	667
2DLC TX114	hypo protein Rv2030c	Rv2030c	74880	100.00%	8	8	18	0.10%	12.60%	DAEAPLDVVR	95.00%	2.67	0.396	0	2	0	0	2	1,085.14	610	619

2DLC TX114	hypo protein Rv2030c	Rv2030c	74880	100.00%	8	8	18	0.10%	12.60%	DFVEWLR	95.00%	2.12	0.237	0	2	0	0	2	965.0884	349	355
2DLC TX114	hypo protein Rv2030c	Rv2030c	74880	100.00%	8	8	18	0.10%	12.60%	LLATPTAGPSLR	95.00%	2.88	0.435	0	3	0	0	2	1,197.55	221	232
2DLC TX114	hypo protein Rv2030c	Rv2030c	74880	100.00%	8	8	18	0.10%	12.60%	LSRDAEAPLDVVR	95.00%	3.38	0.334	0	0	2	0	2	1,441.62	607	619
2DLC TX114	hypo protein Rv2030c	Rv2030c	74880	100.00%	8	8	18	0.10%	12.60%	VAIDAPGGVPTHEVLAELVGDAR	95.00%	3.96	0.548	0	0	2	0	2	2,287.41	246	268
2DLC TX114	hypo protein Rv2030c	Rv2030c	74880	100.00%	8	8	18	0.10%	12.60%	VVVNDDVVR	95.00%	2	0.274	0	1	0	0	2	1,014.87	92	100
2DLC TX114	fbpC2	Rv0129c	36754	99.50%	2	2	2	0.01%	8.53%	ADIQHVLNGATPPAAPAAPA	90.00%	1.79	0.253	0	0	1	0	2	1,954.16	320	340
2DLC TX114	(fbpC2)	Rv0129c	36754	99.50%	2	2	2	0.01%	8.53%	A	90.00%	1.79	0.253	0	0	1	0	2	1,954.16	320	340
2DLC TX114	fbpC2	Rv0129c	36754	99.50%	2	2	2	0.01%	8.53%	FLEGLTLR	95.00%	2.3	0.183	0	1	0	0	2	949.2684	272	279
2DLC TX114	frr (frr)	Rv2882c	20810	100.00%	3	4	8	0.05%	16.80%	ANPGMFSR	95.00%	1.74	0.417	0	2	0	0	2	880.0484	32	39
2DLC TX114	frr (frr)	Rv2882c	20810	100.00%	3	4	8	0.05%	16.80%	LVVIKPYEANQLR	95.00%	2.05	0.336	0	2	1	0	2	1,543.86	64	76
2DLC TX114	frr (frr)	Rv2882c	20810	100.00%	3	4	8	0.05%	16.80%	VAVPQLTEER	95.00%	2.47	0.5	0	3	0	0	2	1,142.16	100	109
2DLC TX114	appC (appC)	Rv1623c	53809	99.80%	2	2	3	0.02%	4.12%	MFQQQPMK	95.00%	2.18	0.28	0	1	0	0	2	1,054.24	251	258
2DLC TX114	appC (appC)	Rv1623c	53809	99.80%	2	2	3	0.02%	4.12%	PTGDQLVRLTVK	95.00%	1.25	0.334	0	0	2	0	2	1,327.53	408	419
2DLC TX114	trpC (trpC)	Rv1611	28005	100.00%	3	3	4	0.02%	11.00%	ASVSIPVLR	95.00%	2	0.223	0	2	0	0	2	942.1184	110	118
2DLC TX114	trpC (trpC)	Rv1611	28005	100.00%	3	3	4	0.02%	11.00%	FQGSLDDLDAVR	95.00%	2.86	0.451	0	1	0	0	2	1,336.50	98	109
2DLC TX114	trpC (trpC)	Rv1611	28005	100.00%	3	3	4	0.02%	11.00%	IVSVVTEQR	95.00%	2.78	0.39	0	1	0	0	2	1,031.01	88	96
2DLC TX114	hypo protein Rv3448	Rv3448	46713	98.60%	2	2	2	0.01%	4.28%	LSGALAASCITAGGGLMLVR	79.60%	1.86	0.2	0	0	1	0	2	1,878.37	120	139
2DLC TX114	hypo protein Rv3448	Rv3448	46713	98.60%	2	2	2	0.01%	4.28%	SGALAASCITAGGGLMLVR	93.30%	1.85	0.263	0	0	1	0	2	1,806.22	121	139
2DLC TX114	hypo protein Rv0569	Rv0569	9504	100.00%	3	5	13	0.08%	35.20%	FGAVQSAILHAR	95.00%	3.5	0.436	0	2	3	0	2	1,270.46	75	86
2DLC TX114	hypo protein Rv0569	Rv0569	9504	100.00%	3	5	13	0.08%	35.20%	SSDGSPPYVVR	95.00%	2.53	0.426	0	2	0	0	2	1,164.19	30	40
2DLC TX114	hypo protein Rv0569	Rv0569	9504	100.00%	3	5	13	0.08%	35.20%	VGDWLVIK	95.00%	1.87	0.272	2	4	0	0	2	930.1084	5	12
2DLC TX114	sahH (sahH)	Rv3248c	54306	100.00%	6	6	13	0.08%	13.90%	GVTEETTTGVLR	95.00%	2.33	0.275	0	2	0	0	2	1,263.26	214	225
2DLC TX114	sahH (sahH)	Rv3248c	54306	100.00%	6	6	13	0.08%	13.90%	IADLSLADFGR	95.00%	3.62	0.502	0	2	0	0	2	1,178.51	23	33
2DLC TX114	sahH (sahH)	Rv3248c	54306	100.00%	6	6	13	0.08%	13.90%	IAEHEMPGLMSLR	95.00%	2.77	0.373	0	0	2	0	2	1,484.85	38	50
2DLC TX114	sahH (sahH)	Rv3248c	54306	100.00%	6	6	13	0.08%	13.90%	SIIVLSEGR	95.00%	2.77	0.378	0	2	0	0	2	974.1984	398	406
2DLC TX114	sahH (sahH)	Rv3248c	54306	100.00%	6	6	13	0.08%	13.90%	VFLNLLR	95.00%	2.07	0.374	0	4	0	0	2	875.1884	189	195
2DLC TX114	sahH (sahH)	Rv3248c	54306	100.00%	6	6	13	0.08%	13.90%	VNVKPQVDLWTFGDTGR	95.00%	2.82	0.359	0	0	1	0	2	1,933.05	381	397
2DLC TX114	infA (infA)	Rv3462c	8471	100.00%	4	5	9	0.05%	53.40%	IELENGHKVL	83.60%	2.06	0.114	0	0	1	0	2	1,152.34	25	34
2DLC TX114	infA (infA)	Rv3462c	8471	100.00%	4	5	9	0.05%	53.40%	ILPEDRVVVELSPYDLSR	95.00%	2.34	0.27	0	0	1	0	2	2,101.43	48	65
2DLC TX114	infA (infA)	Rv3462c	8471	100.00%	4	5	9	0.05%	53.40%	VVEPLPNAMFR	95.00%	3.31	0.459	0	4	1	0	2	1,273.48	14	24
2DLC TX114	infA (infA)	Rv3462c	8471	100.00%	4	5	9	0.05%	53.40%	VVVELSPYDLSR	95.00%	2.3	0.535	0	2	0	0	2	1,377.51	54	65
2DLC TX114	topA (topA)	Rv3646c	102303	100.00%	5	5	6	0.03%	6.75%	ALELGIGIRPSTYSSIIK	95.00%	2.47	0.288	0	0	1	0	2	1,935.46	525	542
2DLC TX114	topA (topA)	Rv3646c	102303	100.00%	5	5	6	0.03%	6.75%	KPRIPVK	90.60%	1.75	0.235	0	1	0	0	2	838.1284	128	134
2DLC TX114	topA (topA)	Rv3646c	102303	100.00%	5	5	6	0.03%	6.75%	LVGINLEGIDAR	95.00%	3.69	0.437	0	2	0	0	2	1,270.57	628	639
2DLC TX114	topA (topA)	Rv3646c	102303	100.00%	5	5	6	0.03%	6.75%	MVFHEITEPAIR	95.00%	2.65	0.219	0	0	1	0	2	1,443.69	136	147
2DLC TX114	topA (topA)	Rv3646c	102303	100.00%	5	5	6	0.03%	6.75%	SMDLQTVTLEDALR	95.00%	2.31	0.377	0	1	0	0	2	1,592.85	769	782
2DLC TX114	rpsM (rpsM)	Rv3460c	14333	97.80%	2	2	2	0.01%	15.30%	DLTEEQLIHLR	80.30%	1.69	0.255	0	1	0	0	2	1,367.62	47	57
2DLC TX114	rpsM (rpsM)	Rv3460c	14333	97.80%	2	2	2	0.01%	15.30%	LVGVDLPR	88.70%	2.67	0.0992	0	1	0	0	2	869.0284	4	11
2DLC TX114	metE (metE)	Rv1133c	81532	100.00%	7	9	15	0.09%	9.62%	AGEIDEAEYVRR	95.00%	2.17	0.399	0	0	3	0	2	1,408.50	455	466
2DLC TX114	metE (metE)	Rv1133c	81532	100.00%	7	9	15	0.09%	9.62%	EVVVLAR	95.00%	2.02	0.186	0	1	0	0	2	785.8084	359	365
2DLC TX114	metE (metE)	Rv1133c	81532	100.00%	7	9	15	0.09%	9.62%	LHLPPLPTTTIGSYPQTSAIR	95.00%	3.29	0.533	0	0	2	0	2	2,264.79	427	447
2DLC TX114	metE (metE)	Rv1133c	81532	100.00%	7	9	15	0.09%	9.62%	SELEAVAATLR	95.00%	1.28	0.231	2	1	0	0	2	1,160.31	41	51
2DLC TX114	metE (metE)	Rv1133c	81532	100.00%	7	9	15	0.09%	9.62%	TSRSELEAVAATLR	95.00%	2.98	0.44	0	1	3	0	2	1,504.69	38	51
2DLC TX114	metE (metE)	Rv1133c	81532	100.00%	7	9	15	0.09%	9.62%	VPSAEEMADSLR	95.00%	3.11	0.577	0	1	0	0	2	1,305.43	704	715
2DLC TX114	metE (metE)	Rv1133c	81532	100.00%	7	9	15	0.09%	9.62%	WAVGAFR	95.00%	1.97	0.33	0	1	0	0	2	806.8984	626	632
2DLC TX114	icd2 (icd2)	Rv0066c	82534	100.00%	3	3	7	0.04%	6.04%	EASDISVAAR	95.00%	3.34	0.502	0	2	0	0	2	1,132.29	39	49
2DLC TX114	icd2 (icd2)	Rv0066c	82534	100.00%	3	3	7	0.04%	6.04%	TFNAALEAVQG	95.00%	2.4	0.447	0	3	0	0	2	1,121.18	735	745
2DLC TX114	icd2 (icd2)	Rv0066c	82534	100.00%	3	3	7	0.04%	6.04%	TIVLKPEVPLDDGDVIDSMFMSK	95.00%	3.51	0.61	0	0	2	0	2	2,551.08	209	231
2DLC TX114	pks2 (pks2)	Rv3825c	225737	99.90%	2	2	3	0.02%	1.36%	ELSDLPGAVPR	95.00%	2.68	0.464	0	2	0	0	2	1,154.34	1317	1327
2DLC TX114	pks2 (pks2)	Rv3825c	225737	99.90%	2	2	3	0.02%	1.36%	LGLFPFR	95.00%	1.87	0.319	0	2	0	0	2	850.1884	1674	1680
2DLC TX114	pks2 (pks2)	Rv3825c	225737	99.90%	2	2	3	0.02%	1.36%	NMGIEHVYDSR	95.00%	1.54	0.267	0	0	1	0	2	1,321.42	1608	1618
2DLC TX114	hypo protein Rv0386	Rv0386	116775	99.40%	2	2	2	0.01%	1.84%	ADENAR	93.50%	1.1	-0.0216	1	0	0	0	2	675.6784	603	608
2DLC TX114	hypo protein Rv0386	Rv0386	116775	99.40%	2	2	2	0.01%	1.84%	SEAGEVDAVFARHR	90.30%	1.86	0.182	0	0	1	0	2	1,544.53	513	526
2DLC TX114	sec (sec)	Rv0379	7948	100.00%	3	3	4	0.02%	29.60%	IKLEVSFK	95.00%	2.24	0.313	0	0	1	0	2	964.2684	57	64
2DLC TX114	sec (sec)	Rv0379	7948	100.00%	3	3	4	0.02%	29.60%	LEVSFK	95.00%	1.48	0.353	0	1	0	0	2	722.8684	59	64
2DLC TX114	sec (sec)	Rv0379	7948	100.00%	3	3	4	0.02%	29.60%	VIEQDMAVDSAGK	95.00%	2.91	0.568	0	2	0	0	2	1,363.45	40	52
2DLC TX114	wag31 (wag31)	Rv2145c	28260	100.00%	8	11	31	0.18%	39.60%	ADAMLADAQSR	95.00%	3.33	0.356	0	2	0	0	2	1,149.32	158	168
2DLC TX114	wag31 (wag31)	Rv2145c	28260	100.00%	8	11	31	0.18%	39.60%	ANAEQILGEAR	95.00%	4.16	0.451	1	7	0	0	2	1,172.31	134	144
2DLC TX114	wag31 (wag31)	Rv2145c	28260	100.00%	8	11	31	0.18%	39.60%	GSAAPVDSNADAGGFDQFNR	95.00%	4.71	0.551	0	3	6	0	2	1,997.04	238	257
2DLC TX114	wag31 (wag31)	Rv2145c	28260	100.00%	8	11	31	0.18%	39.60%	HSEIMGTINQQR	95.00%	2.39	0.401	0	0	1	0	2	1,414.55	191	202
2DLC TX114	wag31 (wag31)	Rv2145c	28260	100.00%	8	11	31	0.18%	39.60%	HTADATVAEAR	95.00%	2.3	0.558	0	1	0	0	2	1,142.11	145	155
2DLC TX114	wag31 (wag31)	Rv2145c	28260	100.00%	8	11	31	0.18%	39.60%	LEQLR	95.00%	1.49	0.0791	3	0	0	0	2	658.8184	209	213
2DLC TX114	wag31 (wag31)	Rv2145c	28260	100.00%	8	11	31	0.18%	39.60%	LIEENSDLR	95.00%	2.74	0.34	0	2	0	0	2	1,089.30	42	50
2DLC TX114	wag31 (wag31)	Rv2145c	28260	100.00%	8	11	31	0.18%	39.60%	TYLESQLEELGQR	95.00%	4.36	0.48	0	2	3	0	2	1,566.75	225	237
2DLC TX114	wag31 (wag31)	Rv2145c	28260	100.00%	8	11	31	0.18%	39.60%	VLSLAQDTADR	95.00%	2.92	0.527	0	3	0	0	2	1,189.33	106	116
2DLC TX114	hypo protein Rv0361	Rv0361	29964	99.80%	2	3	6	0.03%	3.64%	FTAPGFDAK	95.00%	1.68	0.515	0	4	0	0	2	954.1484	68	76
2DLC TX114	hypo protein Rv0361	Rv0361	29964	99.80%	2	3	6	0.03%	3.64%	RFTAPGFDAK	95.00%	2.18	0.292	0	1	1	0	2	1,110.35	67	76
2DLC TX114	hypo protein Rv1926c	Rv1926c	16496	99.80%	2	3	8	0.05%	13.20%	GSVTPAVSQFNAR	95.00%	2.52	0.479	0	3	3	0	2	1,334.35	87	99
2DLC TX114	hypo protein Rv1926c	Rv1926c	16496	99.80%	2	3	8	0.05%	13.20%	TADGINYR	95.00%	1.96	0.32	0	2	0	0	2	910.0484	100	107

2DLC TX114	hypo protein Rv0247c	Rv0247c	28607	99.80%	2	2	2	0.01%	9.68%	IAELEMHPLDTR	95.00%	2.89	0.392	0	0	1	0	2	1,425.72	176	187
2DLC TX114	hypo protein Rv0247c	Rv0247c	28607	99.80%	2	2	2	0.01%	9.68%	LQQTQTPDLAVR	95.00%	2.76	0.507	0	1	0	0	2	1,370.49	39	50
2DLC TX114	dnaJ (dnaJ)	Rv0352	41327	100.00%	4	4	10	0.06%	12.20%	DLHPDANPGNPAAGER	95.00%	2.81	0.343	0	0	1	0	2	1,631.72	36	51
2DLC TX114	dnaJ (dnaJ)	Rv0352	41327	100.00%	4	4	10	0.06%	12.20%	LAGQGEAGLR	95.00%	2.93	0.435	0	1	0	0	2	972.1284	261	270
2DLC TX114	dnaJ (dnaJ)	Rv0352	41327	100.00%	4	4	10	0.06%	12.20%	LFAGGGFGGR	95.00%	2.88	0.436	0	6	0	0	2	939.1484	77	86
2DLC TX114	dnaJ (dnaJ)	Rv0352	41327	100.00%	4	4	10	0.06%	12.20%	SGGSGDLLVTVK	95.00%	2.9	0.503	0	2	0	0	2	1,133.27	347	358
2DLC TX114	qcrA (qcrA)	Rv2195	46906	99.80%	2	2	4	0.02%	3.96%	NPVMLIR	95.00%	2.31	0.345	0	2	0	0	2	843.1084	317	323
2DLC TX114	qcrA (qcrA)	Rv2195	46906	99.80%	2	2	4	0.02%	3.96%	YQGETIYLAR	95.00%	2.63	0.373	0	2	0	0	2	1,214.45	254	263
2DLC TX114	lppM (lppM)	Rv2171	23796	99.80%	2	2	3	0.02%	10.60%	LKPGVVSTMSAQAR	95.00%	3.1	0.414	0	0	1	0	2	1,445.65	164	177
2DLC TX114	lppM (lppM)	Rv2171	23796	99.80%	2	2	3	0.02%	10.60%	NGNIVILEGR	95.00%	3.16	0.303	0	2	0	0	2	1,085.29	115	124
2DLC TX114	hypo protein Rv3651	Rv3651	37950	99.70%	2	2	2	0.01%	6.38%	APAVPVDDLAQR	95.00%	2.01	0.209	0	0	1	0	2	1,252.34	212	223
2DLC TX114	hypo protein Rv3651	Rv3651	37950	99.70%	2	2	2	0.01%	6.38%	KLVPITTFLR	94.90%	1.86	0.206	0	1	0	0	2	1,188.61	26	35
2DLC TX114	fadA3 (fadA3)	Rv1074c	42638	100.00%	3	3	6	0.03%	13.80%	AGMAITDIDLVEINEAFAVQ VLGSAREL	95.00%	2.39	0.176	0	0	1	0	2	2,963.41	315	342
2DLC TX114	fadA3	Rv1074c	42638	100.00%	3	3	6	0.03%	13.80%	AKELGLTPLAR	95.00%	2.97	0.461	0	4	0	0	2	1,169.55	276	286
2DLC TX114	fadA3	Rv1074c	42638	100.00%	3	3	6	0.03%	13.80%	GSLVGMRPDDLAVQMVR	95.00%	3.59	0.531	0	0	1	0	2	1,845.11	20	36
2DLC TX114	fabG4	Rv0242c	46781	100.00%	8	12	26	0.15%	23.60%	ADILVNNAGITR	95.00%	2.53	0.282	0	2	0	0	2	1,257.50	289	300
2DLC TX114	fabG4	Rv0242c	46781	100.00%	8	12	26	0.15%	23.60%	AGEPPLTGSLLIGGAGR	95.00%	3.06	0.371	0	2	2	0	2	1,566.95	39	55
2DLC TX114	fabG4	Rv0242c	46781	100.00%	8	12	26	0.15%	23.60%	ALEGFTR	95.00%	1.92	0.378	0	2	0	0	2	793.9484	140	146
2DLC TX114	fabG4 (fabG4)	Rv0242c	46781	100.00%	8	12	26	0.15%	23.60%	GATTALVYLSPDAKPAATGL ESTMR	95.00%	3.21	0.547	0	0	3	0	2	2,522.99	155	179
2DLC TX114	fabG4	Rv0242c	46781	100.00%	8	12	26	0.15%	23.60%	GIGATLAEVFAR	95.00%	3.23	0.45	1	4	2	0	2	1,205.43	224	235
2DLC TX114	fabG4	Rv0242c	46781	100.00%	8	12	26	0.15%	23.60%	GLHEFFTPVLR	95.00%	2.64	0.484	0	3	1	0	2	1,316.56	101	111
2DLC TX114	fabG4	Rv0242c	46781	100.00%	8	12	26	0.15%	23.60%	VAIVTGAAR	95.00%	2.39	0.354	0	1	0	0	2	857.9484	215	223
2DLC TX114	fabG4	Rv0242c	46781	100.00%	8	12	26	0.15%	23.60%	VIGLSSIAGIAGNR	95.00%	3.29	0.584	0	3	0	0	2	1,328.66	342	355
2DLC TX114	hypo protein Rv0309	Rv0309	22510	99.80%	2	2	3	0.02%	16.50%	TGITTHIGSAGMAPEAK	95.00%	3.54	0.378	0	0	1	0	2	1,642.91	71	87
2DLC TX114	hypo protein Rv0309	Rv0309	22510	99.80%	2	2	3	0.02%	16.50%	WWSGDDNSPTFNSMQVC QK	95.00%	1.93	0.369	0	2	0	0	2	2,289.33	124	142
2DLC TX114	echA8	Rv1070c	27256	100.00%	3	3	5	0.03%	12.80%	AFESSLSEGLLYER	95.00%	3.07	0.484	0	1	0	0	2	1,601.89	214	227
2DLC TX114	echA8	Rv1070c	27256	100.00%	3	3	5	0.03%	12.80%	FGQPEIK	95.00%	2.16	0.319	0	2	0	0	2	818.9884	127	133
2DLC TX114	echA8	Rv1070c	27256	100.00%	3	3	5	0.03%	12.80%	LGVLPGMGGSQR	95.00%	2.1	0.416	0	2	0	0	2	1,172.39	134	145
2DLC TX114	hypo protein Rv3099c	Rv3099c	30458	100.00%	3	3	9	0.05%	12.00%	LDGLQFTR	95.00%	2.8	0.334	0	3	0	0	2	950.1684	241	248
2DLC TX114	hypo protein Rv3099c	Rv3099c	30458	100.00%	3	3	9	0.05%	12.00%	VLLELTGPLSR	95.00%	2.87	0.507	0	4	0	0	2	1,198.54	204	214
2DLC TX114	hypo protein Rv3099c	Rv3099c	30458	100.00%	3	3	9	0.05%	12.00%	VVDDFGGPAPTATIR	95.00%	2.52	0.521	0	2	0	0	2	1,516.66	226	240
2DLC TX114	hisD (hisD)	Rv1599	45326	99.70%	2	2	2	0.01%	7.31%	GAELTAAELR	95.00%	2.37	0.315	0	1	0	0	2	1,031.21	9	18
2DLC TX114	hisD (hisD)	Rv1599	45326	99.70%	2	2	2	0.01%	7.31%	TPVDMITGPGNIYVTAAKRL CR	93.70%	1.65	0.279	0	0	1	0	2	2,377.86	211	232
2DLC TX114	hypo protein Rv1558	Rv1558	16329	100.00%	3	4	7	0.04%	29.10%	AIWWQR	87.50%	1.39	0.262	0	1	0	0	2	859.9984	112	117
2DLC TX114	hypo protein Rv1558	Rv1558	16329	100.00%	3	4	7	0.04%	29.10%	AVAVWPDYASYQTK	95.00%	3.23	0.435	0	2	1	0	2	1,599.70	118	131
2DLC TX114	hypo protein Rv1558	Rv1558	16329	100.00%	3	4	7	0.04%	29.10%	SGGTEGTQLQGKPVILLTTV GAK	95.00%	3.98	0.648	0	0	3	0	2	2,256.59	25	47
2DLC TX114	hypo protein Rv1544	Rv1544	28182	100.00%	3	3	4	0.02%	9.74%	GFPLMLVAR	95.00%	2.53	0.38	0	1	0	0	2	1,004.34	33	41
2DLC TX114	hypo protein Rv1544	Rv1544	28182	100.00%	3	3	4	0.02%	9.74%	SLLLPAIR	94.20%	1.56	0.268	0	1	0	0	2	883.2984	249	256
2DLC TX114	hypo protein Rv1544	Rv1544	28182	100.00%	3	3	4	0.02%	9.74%	TVVPGIVPK	95.00%	1.56	0.4	0	2	0	0	2	910.0184	229	237
2DLC TX114	hypo protein Rv1543	Rv1543	36803	99.80%	2	2	5	0.03%	5.57%	FTTVHMALVR	95.00%	1.63	0.371	0	0	2	0	2	1,175.34	232	241
2DLC TX114	hypo protein Rv1543	Rv1543	36803	99.80%	2	2	5	0.03%	5.57%	NLLDAVQGR	95.00%	2.52	0.445	0	3	0	0	2	986.1184	37	45
2DLC TX114	rne (rne)	Rv2444c	103341	100.00%	3	4	6	0.03%	4.09%	EVVPSDAGVIIR	95.00%	2.52	0.379	0	2	0	0	2	1,255.36	490	501
2DLC TX114	rne (rne)	Rv2444c	103341	100.00%	3	4	6	0.03%	4.09%	FLVYVPGASSTGISR	95.00%	2.84	0.451	0	2	1	0	2	1,554.80	460	474
2DLC TX114	rne (rne)	Rv2444c	103341	100.00%	3	4	6	0.03%	4.09%	RVLDALTALDGR	95.00%	2.81	0.483	0	1	0	0	2	1,300.60	38	49
2DLC TX114	fba (fba)	Rv0363c	36527	99.30%	2	2	3	0.02%	9.01%	LRPDILAQGQQVAAAK	95.00%	3.57	0.518	0	0	2	0	2	1,679.98	222	237
2DLC TX114	fba (fba)	Rv0363c	36527	99.30%	2	2	3	0.02%	9.01%	PIAGHMFTNYDGVLK	85.30%	1.67	0.237	0	0	1	0	2	1,679.99	286	300
2DLC TX114	fadE25)	Rv3274c	41706	99.80%	2	3	9	0.05%	5.66%	ITQIYEGTNQIQR	95.00%	4.1	0.485	0	2	2	0	2	1,564.77	368	380
2DLC TX114	fadE25 (fadE25)	Rv3274c	41706	99.80%	2	3	9	0.05%	5.66%	LGTMGLILR	95.00%	2.44	0.28	0	5	0	0	2	974.4184	102	110
2DLC TX114	oxcA (oxcA)	Rv0118c	61154	99.70%	2	2	2	0.01%	6.87%	AAYRIGQVQDIGRGVARAI R	84.80%	1.63	0.288	0	0	1	0	2	2,171.53	138	157
2DLC TX114	oxcA (oxcA)	Rv0118c	61154	99.70%	2	2	2	0.01%	6.87%	SASPCTVLTDGCHLVVDAL K	95.00%	2.25	0.218	0	1	0	0	2	2,030.22	5	24
2DLC TX114	fixA (fixA)	Rv3029c	28063	100.00%	3	4	7	0.04%	16.20%	AVEEALQIR	95.00%	1.84	0.406	2	2	0	0	2	1,029.15	43	51
2DLC TX114	fixA (fixA)	Rv3029c	28063	100.00%	3	4	7	0.04%	16.20%	LTDGDFTLDR	89.60%	2.88	0.0738	0	1	0	0	2	1,153.37	20	29
2DLC TX114	fixA (fixA)	Rv3029c	28063	100.00%	3	4	7	0.04%	16.20%	TAGEKVTDEGEGGNQIVQY LVAQK	95.00%	4.55	0.6	0	0	2	0	2	2,535.59	241	264
2DLC TX114	rpoA (rpoA)	Rv3457c	37689	100.00%	6	9	26	0.15%	20.50%	EGVHTVGELVAR	95.00%	1.88	0.474	0	2	0	0	2	1,267.21	267	278
2DLC TX114	rpoA (rpoA)	Rv3457c	37689	100.00%	6	9	26	0.15%	20.50%	GYVPAVQNR	95.00%	1.74	0.452	0	2	0	0	2	1,004.01	145	153
2DLC TX114	rpoA (rpoA)	Rv3457c	37689	100.00%	6	9	26	0.15%	20.50%	IPVDSIYSPVLK	95.00%	3.62	0.447	0	3	1	0	2	1,331.66	162	173
2DLC TX114	rpoA (rpoA)	Rv3457c	37689	100.00%	6	9	26	0.15%	20.50%	TDFDKLILDVETK	95.00%	3.84	0.41	0	1	3	0	2	1,537.89	187	199
2DLC TX114	rpoA (rpoA)	Rv3457c	37689	100.00%	6	9	26	0.15%	20.50%	TLLSSIPGAAVTSIR	95.00%	3.37	0.553	0	6	0	0	2	1,486.86	41	55
2DLC TX114	rpoA (rpoA)	Rv3457c	37689	100.00%	6	9	26	0.15%	20.50%	TLVELFGLAR	95.00%	1.48	0.155	2	6	0	0	2	1,119.42	214	223
2DLC TX114	fixB (fixB)	Rv3028c	31672	100.00%	4	7	19	0.11%	16.40%	AAVDSGYYPGQFQVGQTG K	95.00%	4.77	0.646	0	1	1	0	2	1,974.02	233	251
2DLC TX114	fixB (fixB)	Rv3028c	31672	100.00%	4	7	19	0.11%	16.40%	IGSGLLVDVVDVR	95.00%	3.51	0.458	1	10	0	0	2	1,342.48	109	121
2DLC TX114	fixB (fixB)	Rv3028c	31672	100.00%	4	7	19	0.11%	16.40%	VAPQLTEAIK	95.00%	1.76	0.449	2	2	0	0	2	1,070.27	305	314
2DLC TX114	fixB (fixB)	Rv3028c	31672	100.00%	4	7	19	0.11%	16.40%	VSAELITAAR	95.00%	2.56	0.437	0	2	0	0	2	1,031.21	18	27
2DLC TX114	hypo protein Rv2626c	Rv2626c	15500	99.80%	2	3	6	0.03%	21.00%	GLAAGLDPNTATAGELAR	95.00%	3.78	0.598	0	3	1	0	2	1,698.99	58	75
2DLC TX114	hypo protein Rv2626c	Rv2626c	15500	99.80%	2	3	6	0.03%	21.00%	LVGIVTEADIAR	95.00%	2.87	0.544	0	2	0	0	2	1,257.47	110	121
2DLC TX114	hypo protein Rv0981	Rv0981	25844	99.80%	2	2	3	0.02%	10.90%	FSDLTLDPVTR	95.00%	2.84	0.512	0	2	0	0	2	1,264.49	134	144

2DLC TX114	hypo protein Rv0981	Rv0981	25844	99.80%	2	2	3	0.02%	10.90%	GTGDDLPILVLTAR	95.00%	1.94	0.278	0	1	0	0	2	1,441.79	70	83
2DLC TX114	rpmA (rpmA)	Rv2441c	8950	99.80%	2	2	4	0.02%	16.30%	AGEILVR	95.00%	2.59	0.213	0	1	0	0	2	757.9184	33	39
2DLC TX114	rpmA (rpmA)	Rv2441c	8950	99.80%	2	2	4	0.02%	16.30%	GASSSRN	95.00%	0.952	0.129	3	0	0	0	2	678.6884	6	12
2DLC TX114	ctpV (ctpV)	Rv0969	80083	100.00%	3	4	5	0.03%	3.51%	AQVNGGPVVVGR	84.20%	1.02	0.102	2	0	0	0	2	1,153.05	532	543
2DLC TX114	ctpV (ctpV)	Rv0969	80083	100.00%	3	4	5	0.03%	3.51%	VLAEVLPQDK	95.00%	2.32	0.33	0	1	0	0	2	1,112.26	632	641
2DLC TX114	ctpV (ctpV)	Rv0969	80083	100.00%	3	4	5	0.03%	3.51%	VLAEVLPQDKVAEVR	95.00%	3.23	0.482	0	1	1	0	2	1,666.78	632	646
2DLC TX114	hypo protein Rv0968	Rv0968	10245	99.80%	2	3	16	0.09%	32.70%	AGEEVPPLAVAGSDDGHD H	95.00%	1.99	0.266	0	0	1	0	2	1,873.77	80	98
2DLC TX114	hypo protein Rv0968	Rv0968	10245	99.80%	2	3	16	0.09%	32.70%	LMFADVLAEASER	95.00%	5.15	0.582	0	11	4	0	2	1,468.71	67	79
2DLC TX114	hypo protein Rv0902c	Rv0902c	47812	99.20%	2	2	2	0.01%	3.59%	IEATLSALER	95.00%	3.44	0.337	0	1	0	0	2	1,103.36	280	289
2DLC TX114	hypo protein Rv0902c	Rv0902c	47812	99.20%	2	2	2	0.01%	3.59%	IWNEQN	84.90%	0.948	-0.013	1	0	0	0	2	803.8184	217	222
2DLC TX114	hypo protein Rv0954	Rv0954	30185	99.80%	2	2	5	0.03%	3.63%	FDPYGQYGR	95.00%	1.51	0.336	0	4	0	0	2	1,103.24	168	176
2DLC TX114	hypo protein Rv0954	Rv0954	30185	99.80%	2	2	5	0.03%	3.63%	PKFDPYGQYGR	95.00%	2.25	0.338	0	0	1	0	2	1,328.56	166	176
2DLC TX114	sucD (sucD)	Rv0952	31212	100.00%	7	7	12	0.07%	27.70%	AADFIK	95.00%	1.19	0.384	0	1	0	0	2	664.8784	230	235
2DLC TX114	sucD (sucD)	Rv0952	31212	100.00%	7	7	12	0.07%	27.70%	AGTQIVGGVNAR	95.00%	3.07	0.308	0	2	0	0	2	1,143.17	34	45
2DLC TX114	sucD (sucD)	Rv0952	31212	100.00%	7	7	12	0.07%	27.70%	EILLSL	95.00%	1.54	0.001	2	0	0	0	2	687.9984	298	303
2DLC TX114	sucD (sucD)	Rv0952	31212	100.00%	7	7	12	0.07%	27.70%	LPVFGSVAEAMEK TNVSKPVVGYVAGFTAPEG K	95.00%	2.92	0.569	0	1	0	0	2	1,378.57	63	75
2DLC TX114	sucD (sucD)	Rv0952	31212	100.00%	7	7	12	0.07%	27.70%	TPSATAALAR	95.00%	2.29	0.455	0	2	0	0	2	959.1384	288	297
2DLC TX114	sucD (sucD)	Rv0952	31212	100.00%	7	7	12	0.07%	27.70%	VIVQGITGSEATVHTAR	95.00%	2.51	0.408	0	0	2	0	2	1,739.77	14	30
2DLC TX114	sucC (sucC)	Rv0951	40908	100.00%	7	10	16	0.09%	19.10%	AIATEIGRPVMVK	95.00%	3.43	0.36	0	1	0	0	2	1,385.69	33	45
2DLC TX114	sucC (sucC)	Rv0951	40908	100.00%	7	10	16	0.09%	19.10%	GVDLDFAR	95.00%	1.55	0.234	2	2	0	0	2	893.0184	146	153
2DLC TX114	sucC (sucC)	Rv0951	40908	100.00%	7	10	16	0.09%	19.10%	ILALDAK	95.00%	1.89	0.25	0	1	0	0	2	744.0784	201	207
2DLC TX114	sucC (sucC)	Rv0951	40908	100.00%	7	10	16	0.09%	19.10%	MDLFEYQAK	95.00%	2.75	0.408	0	1	0	0	2	1,145.40	1	9
2DLC TX114	sucC (sucC)	Rv0951	40908	100.00%	7	10	16	0.09%	19.10%	NILGLDIK SIAEQGHLPAEVLDTAAVTI AK	95.00%	2.38	0.414	1	1	0	0	2	886.2684	74	81
2DLC TX114	sucC (sucC)	Rv0951	40908	100.00%	7	10	16	0.09%	19.10%	VPVNAVK	95.00%	3.94	0.606	0	2	4	0	2	2,235.51	154	175
2DLC TX114	sucC (sucC)	Rv0951	40908	100.00%	7	10	16	0.09%	19.10%	GLSTHEGALLSER	95.00%	2.03	0.302	0	1	0	0	2	726.7284	139	145
2DLC TX114	ald (ald)	Rv2780	38695	99.90%	2	3	7	0.04%	7.01%	VAITPAGVAELTR	95.00%	2.54	0.348	0	0	3	0	2	1,370.56	342	354
2DLC TX114	ald (ald)	Rv2780	38695	99.90%	2	3	7	0.04%	7.01%	VAITPAGVAELTR	95.00%	2.65	0.508	0	3	1	0	2	1,298.47	16	28
2DLC TX114	hypo protein Rv3880c	Rv3880c	12150	100.00%	4	6	12	0.07%	20.90%	ALTLAAR	95.00%	0.981	0.195	2	0	0	0	2	715.9584	13	19
2DLC TX114	hypo protein Rv3880c	Rv3880c	12150	100.00%	4	6	12	0.07%	20.90%	IEDGLLK	95.00%	1.71	0.312	2	2	0	0	2	788.0684	60	66
2DLC TX114	hypo protein Rv3880c	Rv3880c	12150	100.00%	4	6	12	0.07%	20.90%	KLGAEAVAQR	95.00%	3.03	0.397	0	1	0	0	2	1,043.18	67	76
2DLC TX114	hypo protein Rv3880c	Rv3880c	12150	100.00%	4	6	12	0.07%	20.90%	LGAEAVAQR	95.00%	3.71	0.502	2	3	0	0	2	914.9784	68	76
2DLC TX114	phoS1	Rv0934	38194	100.00%	6	10	70	0.40%	21.10%	ASFLDQVHFQPLPPAVVK	95.00%	5.47	0.664	0	19	27	0	2	1,994.23	346	363
2DLC TX114	phoS1	Rv0934	38194	100.00%	6	10	70	0.40%	21.10%	LSDALIATISS	95.00%	1.43	0.535	4	0	0	0	2	1,091.42	364	374
2DLC TX114	phoS1	Rv0934	38194	100.00%	6	10	70	0.40%	21.10%	PGVNLPGTAVVPLHR	95.00%	4.81	0.403	0	0	2	0	2	1,527.67	171	185
2DLC TX114	phoS1 phoS1	Rv0934	38194	100.00%	6	10	70	0.40%	21.10%	SDGSGDTFLFTQYLSK SDGSGDTFLFTQYLSKQDP	95.00%	4.54	0.622	0	3	3	0	2	1,767.06	186	201
2DLC TX114	(phoS1)	Rv0934	38194	100.00%	6	10	70	0.40%	21.10%	EGWGK	95.00%	3.79	0.587	0	0	2	0	2	2,664.98	186	209
2DLC TX114	phoS1 phoS2	Rv0934	38194	100.00%	6	10	70	0.40%	21.10%	VLAAMYQGTIK AFLQSTIGAGQSGLGDNGY	95.00%	3.17	0.41	2	7	1	0	2	1,211.50	149	159
2DLC TX114	(phoS2)	Rv0928	37917	99.40%	2	2	2	0.01%	10.30%	IPIPDEFKSR	95.00%	2.01	0.289	0	0	1	0	2	3,040.61	333	361
2DLC TX114	phoS2	Rv0928	37917	99.40%	2	2	2	0.01%	10.30%	LSTAVNAIA	88.60%	1.51	0.258	0	1	0	0	2	860.0084	362	370
2DLC TX114	moeZ	Rv3206c	42124	99.80%	2	2	3	0.02%	5.36%	IDGAQLIPK	95.00%	1.52	-0.056	1	0	0	0	2	955.2684	320	328
2DLC TX114	moeZ	Rv3206c	42124	99.80%	2	2	3	0.02%	5.36%	LITGIGETLLGR	95.00%	3.45	0.531	0	2	0	0	2	1,243.67	227	238
2DLC TX114	serA (serA)	Rv2996c	54504	100.00%	9	12	18	0.10%	22.00%	AGTDVAESVR	95.00%	2.61	0.453	0	1	0	0	2	1,004.95	291	300
2DLC TX114	serA (serA)	Rv2996c	54504	100.00%	9	12	18	0.10%	22.00%	AGVGLDNVDVDAATAR	95.00%	4.57	0.601	0	5	1	0	2	1,544.55	72	87
2DLC TX114	serA (serA)	Rv2996c	54504	100.00%	9	12	18	0.10%	22.00%	GELAAEEVEVLR	95.00%	2.69	0.28	0	1	0	0	2	1,315.37	354	365
2DLC TX114	serA (serA)	Rv2996c	54504	100.00%	9	12	18	0.10%	22.00%	GGLVDEAALADAITGGHVR	95.00%	4.32	0.607	0	1	1	0	2	1,822.98	233	251
2DLC TX114	serA (serA)	Rv2996c	54504	100.00%	9	12	18	0.10%	22.00%	IGQLVAQR	95.00%	2.71	0.375	0	3	0	0	2	885.0184	152	159
2DLC TX114	serA (serA)	Rv2996c	54504	100.00%	9	12	18	0.10%	22.00%	SATTVDAEVLAAAPK	95.00%	3.31	0.385	0	1	1	0	2	1,444.56	51	65
2DLC TX114	serA (serA)	Rv2996c	54504	100.00%	9	12	18	0.10%	22.00%	TKPGVIIVNAAR	95.00%	3.25	0.243	0	0	1	0	2	1,239.49	221	232
2DLC TX114	serA (serA)	Rv2996c	54504	100.00%	9	12	18	0.10%	22.00%	TPETAGLIDKEALAK	95.00%	2.92	0.412	0	0	1	0	2	1,557.93	206	220
2DLC TX114	serA (serA)	Rv2996c	54504	100.00%	9	12	18	0.10%	22.00%	TVGVVGLGR	95.00%	2.06	0.289	0	1	0	0	2	857.8784	143	151
2DLC TX114	hypo protein Rv3205c	Rv3205c	31334	99.80%	2	2	2	0.01%	9.59%	STAEAFPGLAHTAALVR	95.00%	2.32	0.245	0	0	1	0	2	1,712.94	273	289
2DLC TX114	hypo protein Rv3205c	Rv3205c	31334	99.80%	2	2	2	0.01%	9.59%	VRETLFVDGVR	95.00%	2.86	0.341	0	0	1	0	2	1,291.38	62	72
2DLC TX114	echA6	Rv0905	26012	99.70%	2	2	3	0.02%	7.82%	AMLLSAEK	94.90%	1.65	0.294	0	1	0	0	2	863.1584	149	156
2DLC TX114	echA6	Rv0905	26012	99.70%	2	2	3	0.02%	7.82%	YGLALDNWSIR	95.00%	3.67	0.479	0	2	0	0	2	1,308.63	126	136
2DLC TX114	hypo protein Rv0901	Rv0901	18891	100.00%	3	4	14	0.08%	34.90%	AGADGSGPQGWLVK	95.00%	2.83	0.485	0	4	0	0	2	1,343.45	116	129
2DLC TX114	hypo protein Rv0901	Rv0901	18891	100.00%	3	4	14	0.08%	34.90%	IPVVPYAPYGPGSAR TKVATESAAEQIPVAGEPAA	95.00%	3.24	0.467	0	4	4	0	2	1,544.78	101	115
2DLC TX114	hypo protein Rv0901	Rv0901	18891	100.00%	3	4	14	0.08%	34.90%	EPIPVAGEPAAR	95.00%	5.09	0.636	0	0	2	0	2	3,129.33	69	100
2DLC TX114	mpt64 (mpt64)	Rv1980c	24805	99.80%	2	2	4	0.02%	7.02%	GTQAVVLK	95.00%	2.14	0.224	0	2	0	0	2	815.8884	108	115
2DLC TX114	mpt64 (mpt64)	Rv1980c	24805	99.80%	2	2	4	0.02%	7.02%	SAIDSMLA	95.00%	1.39	0.23	2	0	0	0	2	808.0384	221	228
2DLC TX114	hypo protein Rv3816c	Rv3816c	28404	99.70%	2	2	2	0.01%	7.34%	DRPKKL	94.10%	1.03	0.0752	1	0	0	0	2	757.0384	160	165
2DLC TX114	hypo protein Rv3816c	Rv3816c	28404	99.70%	2	2	2	0.01%	7.34%	LGGGAPSLAEAAR	95.00%	2.8	0.521	0	1	0	0	2	1,170.39	222	234
2DLC TX114	guaA (guaA)	Rv3396c	56009	100.00%	4	5	7	0.04%	6.67%	ELGLPEEIVAR	95.00%	1.9	0.424	0	2	0	0	2	1,226.44	389	399
2DLC TX114	guaA (guaA)	Rv3396c	56009	100.00%	4	5	7	0.04%	6.67%	FTLVEPLR	95.00%	2.39	0.291	0	1	0	0	2	975.2084	369	376
2DLC TX114	guaA (guaA)	Rv3396c	56009	100.00%	4	5	7	0.04%	6.67%	LLFKDEVR	95.00%	2.27	0.316	0	1	1	0	2	1,020.29	377	384
2DLC TX114	guaA (guaA)	Rv3396c	56009	100.00%	4	5	7	0.04%	6.67%	VPYEVLER	95.00%	2.57	0.343	0	2	0	0	2	1,005.08	488	495
2DLC TX114	hypo protein Rv0292	Rv0292	35915	99.50%	2	3	8	0.05%	8.46%	IPLQETAQVAAR	95.00%	3.18	0.526	0	4	3	0	2	1,297.45	154	165

2DLC TX114	hypo protein Rv0292	Rv0292	35915	99.50%	2	3	8	0.05%	8.46%	RRRIAEPATCTRTTVL	89.00%	1.79	0.274	0	0	1	0	2	1,903.20	71	86
2DLC TX114	hypo protein Rv0291	Rv0291	46108	100.00%	3	3	4	0.02%	7.59%	MVDQAALGAAIR	95.00%	3.13	0.397	0	1	0	0	2	1,216.46	211	222
2DLC TX114	hypo protein Rv0291	Rv0291	46108	100.00%	3	3	4	0.02%	7.59%	NVAFAGAAAL	95.00%	1.47	0.344	1	0	0	0	2	905.0384	430	439
2DLC TX114	hypo protein Rv0291	Rv0291	46108	100.00%	3	3	4	0.02%	7.59%	SRYPGLNATEVVR	95.00%	3.51	0.201	0	0	2	0	2	1,462.59	357	369
2DLC TX114	hypo protein Rv0290	Rv0290	47927	100.00%	3	3	10	0.06%	5.93%	EILPAVQR	95.00%	1.4	0.371	0	1	0	0	2	926.0884	32	39
2DLC TX114	hypo protein Rv0290	Rv0290	47927	100.00%	3	3	10	0.06%	5.93%	LLEDLPR	95.00%	2.28	0.246	0	4	0	0	2	856.1384	312	318
2DLC TX114	hypo protein Rv0290	Rv0290	47927	100.00%	3	3	10	0.06%	5.93%	LTEMALPAELPLR	95.00%	3.14	0.593	0	5	0	0	2	1,454.92	19	31
2DLC TX114	hypo protein Rv0284	Rv0284	145124	100.00%	18	22	53	0.31%	13.50%	ALQDLAHVGSVASALEPER	95.00%	4.08	0.62	0	1	4	0	2	1,964.11	869	887
2DLC TX114	hypo protein Rv0284	Rv0284	145124	100.00%	18	22	53	0.31%	13.50%	AVVVQSVPEPK	95.00%	3.27	0.433	0	2	0	0	2	1,153.10	713	723
2DLC TX114	hypo protein Rv0284	Rv0284	145124	100.00%	18	22	53	0.31%	13.50%	DPLVFDAR	95.00%	2.63	0.398	0	1	0	0	2	933.0884	812	819
2DLC TX114	hypo protein Rv0284	Rv0284	145124	100.00%	18	22	53	0.31%	13.50%	FADTLR	95.00%	1.7	0.369	0	2	0	0	2	722.8984	544	549
2DLC TX114	hypo protein Rv0284	Rv0284	145124	100.00%	18	22	53	0.31%	13.50%	GASFTTLLGIEDASR	95.00%	4.18	0.61	0	2	0	0	2	1,538.84	416	430
2DLC TX114	hypo protein Rv0284	Rv0284	145124	100.00%	18	22	53	0.31%	13.50%	GVGFLVPAPGATPIR	89.60%	1.92	0.204	0	1	0	0	2	1,452.73	681	695
2DLC TX114	hypo protein Rv0284	Rv0284	145124	100.00%	18	22	53	0.31%	13.50%	HTVPLATTLR	90.50%	1.3	0.322	0	1	0	0	2	1,109.29	163	172
2DLC TX114	hypo protein Rv0284	Rv0284	145124	100.00%	18	22	53	0.31%	13.50%	KLIATIGEQLAR	95.00%	4.87	0.462	0	4	5	0	2	1,313.73	752	763
2DLC TX114	hypo protein Rv0284	Rv0284	145124	100.00%	18	22	53	0.31%	13.50%	LDVPALWAPR	95.00%	2.65	0.401	0	4	0	0	2	1,138.39	431	440
2DLC TX114	hypo protein Rv0284	Rv0284	145124	100.00%	18	22	53	0.31%	13.50%	LIATIGEQLAR	95.00%	3.3	0.336	0	1	0	0	2	1,185.53	753	763
2DLC TX114	hypo protein Rv0284	Rv0284	145124	100.00%	18	22	53	0.31%	13.50%	LIFEAR	95.00%	1.27	0.123	2	0	0	0	2	748.9984	4	9
2DLC TX114	hypo protein Rv0284	Rv0284	145124	100.00%	18	22	53	0.31%	13.50%	NPLLAR	95.00%	1.94	0.243	0	2	0	0	2	683.9184	947	952
2DLC TX114	hypo protein Rv0284	Rv0284	145124	100.00%	18	22	53	0.31%	13.50%	SIGDVPTGIDLTK	95.00%	1.9	0.278	0	1	0	0	2	1,316.59	199	211
2DLC TX114	hypo protein Rv0284	Rv0284	145124	100.00%	18	22	53	0.31%	13.50%	SIGDVPTGIDLTKVSPITVL GER	95.00%	2.6	0.346	0	0	3	0	2	2,368.78	199	221
2DLC TX114	hypo protein Rv0284	Rv0284	145124	100.00%	18	22	53	0.31%	13.50%	STYVDGIYEPPQTAK	95.00%	3.21	0.456	0	3	2	0	2	1,669.84	698	712
2DLC TX114	hypo protein Rv0284	Rv0284	145124	100.00%	18	22	53	0.31%	13.50%	VSPITVLGER	95.00%	2.37	0.434	0	2	0	0	2	1,071.21	212	221
2DLC TX114	hypo protein Rv0284	Rv0284	145124	100.00%	18	22	53	0.31%	13.50%	WLPHVDIPGR	95.00%	2.46	0.495	0	2	0	0	2	1,190.38	263	272
2DLC TX114	hypo protein Rv0284	Rv0284	145124	100.00%	18	22	53	0.31%	13.50%	WPLGEIDRPFEMR	95.00%	3.11	0.514	0	2	6	0	2	1,647.01	798	810
2DLC TX114	hypo protein Rv0282	Rv0282	68089	100.00%	7	7	15	0.09%	11.30%	ADLIGQHIGETEAK	95.00%	2.65	0.393	0	0	2	0	2	1,482.68	419	432
2DLC TX114	hypo protein Rv0282	Rv0282	68089	100.00%	7	7	15	0.09%	11.30%	ALGIAVYQR	95.00%	3.31	0.516	0	4	0	0	2	991.1984	47	55
2DLC TX114	hypo protein Rv0282	Rv0282	68089	100.00%	7	7	15	0.09%	11.30%	GLGLSVR	89.30%	1.88	0.147	0	4	0	0	2	701.8684	624	630
2DLC TX114	hypo protein Rv0282	Rv0282	68089	100.00%	7	7	15	0.09%	11.30%	LAAESTPDTNGISR	95.00%	2.78	0.578	0	1	0	0	2	1,432.61	550	563
2DLC TX114	hypo protein Rv0282	Rv0282	68089	100.00%	7	7	15	0.09%	11.30%	SGFAALTR	95.00%	2.83	0.469	0	1	0	0	2	823.0084	66	73
2DLC TX114	hypo protein Rv0282	Rv0282	68089	100.00%	7	7	15	0.09%	11.30%	SLDIAGNGR	95.00%	2	0.327	0	1	0	0	2	903.0784	565	573
2DLC TX114	hypo protein Rv0282	Rv0282	68089	100.00%	7	7	15	0.09%	11.30%	TATTAGVLQR	95.00%	2.13	0.383	0	2	0	0	2	1,018.10	102	111
2DLC TX114	lexA (lexA)	Rv2720	23045	99.80%	2	2	4	0.02%	10.60%	ELVGEGTLFLLK	95.00%	3.3	0.54	0	3	0	0	2	1,319.69	125	136
2DLC TX114	lexA (lexA)	Rv2720	23045	99.80%	2	2	4	0.02%	10.60%	VLSADSALTER	95.00%	2.91	0.553	0	1	0	0	2	1,162.31	1	11
2DLC TX114	lprG (lprG)	Rv1411c	24530	100.00%	5	7	22	0.13%	23.70%	DTINGQNTIR	95.00%	2.88	0.347	0	3	0	0	2	1,132.27	158	167
2DLC TX114	lprG (lprG)	Rv1411c	24530	100.00%	5	7	22	0.13%	23.70%	GSGNSVQMTLSK	95.00%	3.39	0.594	0	7	0	0	2	1,209.33	212	223
2DLC TX114	lprG (lprG)	Rv1411c	24530	100.00%	5	7	22	0.13%	23.70%	IPGLSLK	95.00%	1.95	0.374	2	2	0	0	2	728.0684	68	74
2DLC TX114	lprG (lprG)	Rv1411c	24530	100.00%	5	7	22	0.13%	23.70%	TLSGDLTTNPTAATGNVK	95.00%	2.52	0.507	0	2	4	0	2	1,761.95	75	92
2DLC TX114	lprG (lprG)	Rv1411c	24530	100.00%	5	7	22	0.13%	23.70%	VQVTKPPVS	95.00%	2.16	0.379	0	2	0	0	2	954.9484	228	236
2DLC TX114	fadD2	Rv0270	59876	100.00%	6	6	9	0.05%	12.10%	AGDGVAILAR	95.00%	2.64	0.356	0	2	0	0	2	943.1284	110	119
2DLC TX114	fadD2	Rv0270	59876	100.00%	6	6	9	0.05%	12.10%	AVSLAQPPLGK	95.00%	3.15	0.405	0	1	0	0	2	1,081.32	171	181
2DLC TX114	fadD2	Rv0270	59876	100.00%	6	6	9	0.05%	12.10%	IFVGNAFPFEGYTGGGGK	95.00%	2.91	0.453	0	2	0	0	2	1,819.09	417	434
2DLC TX114	fadD2	Rv0270	59876	100.00%	6	6	9	0.05%	12.10%	SIGQYVAR	86.00%	1.5	0.272	0	1	0	0	2	893.9984	18	25
2DLC TX114	fadD2	Rv0270	59876	100.00%	6	6	9	0.05%	12.10%	VIIYDDEYTK	95.00%	1.75	0.428	0	2	0	0	2	1,259.49	161	170
2DLC TX114	fadD2	Rv0270	59876	100.00%	6	6	9	0.05%	12.10%	WFVIANYGAAR	95.00%	2.45	0.36	0	1	0	0	2	1,268.48	123	133
2DLC TX114	lysX (lysX)	Rv1640c	128192	99.80%	2	2	2	0.01%	2.99%	ETLYRSNMK	95.00%	2.24	0.21	0	1	0	0	2	1,142.40	587	595
2DLC TX114	lysX (lysX)	Rv1640c	128192	99.80%	2	2	2	0.01%	2.99%	HRDIAEDEMAQTITRADSW RDTETER	95.00%	2.16	0.271	0	0	1	0	2	3,150.34	444	469
2DLC TX114	ideR (ideR)	Rv2711	25214	100.00%	4	6	11	0.06%	21.70%	ALAIAVMR	95.00%	1.92	0.301	0	1	0	0	2	861.1278	70	77
2DLC TX114	ideR (ideR)	Rv2711	25214	100.00%	4	6	11	0.06%	21.70%	LTELPAGSPVAVVVR	95.00%	3.72	0.636	0	2	2	0	2	1,508.63	154	168
2DLC TX114	ideR (ideR)	Rv2711	25214	100.00%	4	6	11	0.06%	21.70%	MNELVDTTEMYLR	95.00%	4.19	0.449	0	2	0	0	2	1,615.89	1	13
2DLC TX114	ideR (ideR)	Rv2711	25214	100.00%	4	6	11	0.06%	21.70%	TIYDLEEEGVTPLR	95.00%	2.78	0.451	0	2	2	0	2	1,635.86	14	27
2DLC TX114	hypo protein Rv1919c	Rv1919c	16785	99.80%	2	2	5	0.03%	12.30%	FFAGQLVK	95.00%	2.25	0.331	0	3	0	0	2	910.1184	138	145
2DLC TX114	hypo protein Rv1919c	Rv1919c	16785	99.80%	2	2	5	0.03%	12.30%	TSSAPAATLFR	95.00%	2.14	0.521	0	2	0	0	2	1,122.34	13	23
2DLC TX114	hypo protein Rv3720	Rv3720	46835	99.80%	2	2	2	0.01%	5.95%	SIGVEHILPIAPPPQEAR	95.00%	2.54	0.275	0	0	1	0	2	1,925.26	101	118
2DLC TX114	hypo protein Rv3720	Rv3720	46835	99.80%	2	2	2	0.01%	5.95%	VLANVVR	95.00%	1.93	0.226	0	1	0	0	2	770.8084	94	100
2DLC TX114	nirB (nirB)	Rv0252	90335	99.80%	2	3	7	0.04%	2.58%	IKLGVSGCAR	95.00%	1.44	0.244	5	1	0	0	2	1,004.25	687	696
2DLC TX114	nirB (nirB)	Rv0252	90335	99.80%	2	3	7	0.04%	2.58%	LYVAGNGGMTPK	95.00%	0.951	0.259	1	0	0	0	2	1,224.44	717	728
2DLC TX114	35kd_ag (35kd_ag)	Rv2744c	29240	100.00%	19	28	112	0.65%	71.50%	AIGSAELAESSVQGR	95.00%	3.21	0.496	0	1	1	0	2	1,475.57	200	214
2DLC TX114	35kd_ag (35kd_ag)	Rv2744c	29240	100.00%	19	28	112	0.65%	71.50%	GEALPAGGTTATPR	95.00%	2.55	0.546	0	2	0	0	2	1,299.45	240	253
2DLC TX114	35kd_ag (35kd_ag)	Rv2744c	29240	100.00%	19	28	112	0.65%	71.50%	GEALPAGGTTATPRPATETS GGAIAEQPYGQ	95.00%	4.63	0.72	0	0	2	0	2	2,958.16	240	270

2DLC TX114	35kd_ag (35kd_ag)	Rv2744c	29240	100.00%	19	28	112	0.65%	71.50%	LEQIR	95.00%	1.49	0.0791	3	0	0	0	2	658.8184	231	235
2DLC TX114	35kd_ag (35kd_ag)	Rv2744c	29240	100.00%	19	28	112	0.65%	71.50%	LLSQLEQAK	95.00%	2.68	0.447	0	12	0	0	2	1,030.28	154	162
2DLC TX114	35kd_ag (35kd_ag)	Rv2744c	29240	100.00%	19	28	112	0.65%	71.50%	LQVNVR	95.00%	1.65	0.231	2	0	0	0	2	728.7584	72	77
2DLC TX114	35kd_ag (35kd_ag)	Rv2744c	29240	100.00%	19	28	112	0.65%	71.50%	MLEVEQAGIQMAGHSR	95.00%	4.2	0.54	0	7	19	0	2	1,757.93	215	230
2DLC TX114	35kd_ag (35kd_ag)	Rv2744c	29240	100.00%	19	28	112	0.65%	71.50%	MQEQVSASLR	95.00%	3.41	0.435	0	4	0	0	2	1,149.23	163	172
2DLC TX114	35kd_ag (35kd_ag)	Rv2744c	29240	100.00%	19	28	112	0.65%	71.50%	NAMVLQQK	95.00%	2.89	0.442	0	5	0	0	2	948.0678	140	147
2DLC TX114	35kd_ag (35kd_ag)	Rv2744c	29240	100.00%	19	28	112	0.65%	71.50%	PATETSGGAIAEQPYGQ	95.00%	4.14	0.446	0	0	1	0	2	1,677.73	254	270
2DLC TX114	35kd_ag (35kd_ag)	Rv2744c	29240	100.00%	19	28	112	0.65%	71.50%	QALTLADQATAAGDAAK	95.00%	3.02	0.56	0	1	1	0	2	1,616.83	78	94
2DLC TX114	35kd_ag (35kd_ag)	Rv2744c	29240	100.00%	19	28	112	0.65%	71.50%	RYANAIGSAELAESSVQGR	95.00%	3.32	0.523	0	2	2	0	2	1,980.15	196	214
2DLC TX114	35kd_ag (35kd_ag)	Rv2744c	29240	100.00%	19	28	112	0.65%	71.50%	SMSELAAPGNTPSLDEVR	95.00%	3.05	0.442	0	6	0	0	2	1,875.08	173	190
2DLC TX114	35kd_ag (35kd_ag)	Rv2744c	29240	100.00%	19	28	112	0.65%	71.50%	SMSELAAPGNTPSLDEVRD K	95.00%	2.6	0.525	0	3	5	0	2	2,134.38	173	192
2DLC TX114	35kd_ag (35kd_ag)	Rv2744c	29240	100.00%	19	28	112	0.65%	71.50%	THQALTQQAAQVIGNQR	95.00%	6.02	0.561	0	4	1	0	2	1,864.88	40	56
2DLC TX114	35kd_ag (35kd_ag)	Rv2744c	29240	100.00%	19	28	112	0.65%	71.50%	TLHDQALSAAAQAK	95.00%	3.32	0.508	0	2	4	0	2	1,425.60	121	134
2DLC TX114	35kd_ag (35kd_ag)	Rv2744c	29240	100.00%	19	28	112	0.65%	71.50%	VQIQQAIEEAQR	95.00%	3.16	0.567	0	5	0	0	2	1,413.45	28	39
2DLC TX114	35kd_ag (35kd_ag)	Rv2744c	29240	100.00%	19	28	112	0.65%	71.50%	YANAIGSAELAESSVQGR	95.00%	4.77	0.591	0	6	4	0	2	1,823.95	197	214
2DLC TX114	35kd_ag (35kd_ag)	Rv2744c	29240	100.00%	19	28	112	0.65%	71.50%	YLMALFSSK	95.00%	2.72	0.638	2	5	0	0	2	1,076.46	11	19
2DLC TX114	lpqI (lpqI)	Rv0237	39500	100.00%	3	3	3	0.02%	7.73%	AAVNLLR	95.00%	2.35	0.174	0	1	0	0	2	756.9484	288	294
2DLC TX114	lpqI (lpqI)	Rv0237	39500	100.00%	3	3	3	0.02%	7.73%	EVPAVLDRLEQALR	95.00%	1.55	0.317	0	0	1	0	2	1,609.84	347	360
2DLC TX114	lpqI (lpqI)	Rv0237	39500	100.00%	3	3	3	0.02%	7.73%	FGVSEAVLR	95.00%	3.57	0.397	0	1	0	0	2	978.0684	323	331
2DLC TX114	hypo protein Rv2074	Rv2074	14974	100.00%	4	5	8	0.05%	36.50%	AAVNSDIDAVRDAELR	95.00%	2.57	0.464	0	0	2	0	2	1,715.86	88	103
2DLC TX114	hypo protein Rv2074	Rv2074	14974	100.00%	4	5	8	0.05%	36.50%	HLAMLTTLR	95.00%	2.38	0.417	0	1	0	0	2	1,056.40	22	30
2DLC TX114	hypo protein Rv2074	Rv2074	14974	100.00%	4	5	8	0.05%	36.50%	LSDDALAFLSER	95.00%	3.86	0.531	0	3	1	0	2	1,337.64	10	21
2DLC TX114	hypo protein Rv2074	Rv2074	14974	100.00%	4	5	8	0.05%	36.50%	SGLAVLSQVDGAR	95.00%	2.88	0.48	0	1	0	0	2	1,273.38	68	80
2DLC TX114	hypo protein Rv2050	Rv2050	12954	100.00%	3	3	9	0.05%	25.20%	NGMEGTLIEGDLPEPK	95.00%	3.66	0.552	0	2	0	0	2	1,701.01	58	73
2DLC TX114	hypo protein Rv2050	Rv2050	12954	100.00%	3	3	9	0.05%	25.20%	SIEELEELLK	95.00%	3.47	0.342	0	3	0	0	2	1,203.50	90	99
2DLC TX114	hypo protein Rv2050	Rv2050	12954	100.00%	3	3	9	0.05%	25.20%	SIEELEELLKER	95.00%	3.6	0.41	0	4	0	0	2	1,488.80	90	101
2DLC TX114	hypo protein Rv2005c	Rv2005c	30965	100.00%	3	4	13	0.08%	11.50%	GLLGSVSSSLVR	95.00%	3.14	0.528	3	8	0	0	2	1,175.38	127	138
2DLC TX114	hypo protein Rv2005c	Rv2005c	30965	100.00%	3	4	13	0.08%	11.50%	LAGWQER	91.30%	1.83	0.172	0	1	0	0	2	859.9484	224	230
2DLC TX114	hypo protein Rv2005c	Rv2005c	30965	100.00%	3	4	13	0.08%	11.50%	LAGWQERYPDVPVSR	95.00%	3.45	0.33	0	0	1	0	2	1,773.91	224	238
2DLC TX114	hypo protein Rv2005c	Rv2005c	30965	100.00%	3	4	13	0.08%	11.50%	VPVIVAR	95.00%	2.14	0.295	0	1	0	0	2	753.8284	287	293
2DLC TX114	hypo protein Rv2096c	Rv2096c	35282	100.00%	3	3	7	0.04%	11.40%	DLGIPLEVGR	95.00%	2.06	0.507	0	4	0	0	2	1,069.31	57	66
2DLC TX114	hypo protein Rv2096c	Rv2096c	35282	100.00%	3	3	7	0.04%	11.40%	SSVAGYSDSPSVEAFSR	95.00%	3.54	0.654	0	2	0	0	2	1,746.77	30	46
2DLC TX114	hypo protein Rv2096c	Rv2096c	35282	100.00%	3	3	7	0.04%	11.40%	VSALEPTEGYR	95.00%	1.71	0.258	0	1	0	0	2	1,222.32	67	77
2DLC TX114	hypo protein Rv2032	Rv2032	36542	100.00%	4	4	6	0.03%	16.00%	ADAILLR	95.00%	2.43	0.194	0	1	0	0	2	772.0784	114	120
2DLC TX114	hypo protein Rv2032	Rv2032	36542	100.00%	4	4	6	0.03%	16.00%	IDVIADDMRPELAAASK	95.00%	1.98	0.317	0	0	1	0	2	1,816.21	151	167
2DLC TX114	hypo protein Rv2032	Rv2032	36542	100.00%	4	4	6	0.03%	16.00%	LRADAILLR	95.00%	2.72	0.186	0	2	0	0	2	1,041.48	112	120
2DLC TX114	hypo protein Rv2032	Rv2032	36542	100.00%	4	4	6	0.03%	16.00%	VGLAPEMEEPPPATPR	95.00%	2.77	0.273	0	0	1	0	2	1,691.90	301	316
2DLC TX114	Rv2032	Rv2032	36542	100.00%	4	4	6	0.03%	16.00%	VLVLSTYDNER	95.00%	2.19	0.258	0	2	0	0	2	1,309.44	235	245
2DLC TX114	ilvX (ilvX)	Rv3509c	52022	99.80%	2	3	3	0.02%	6.02%	ALDLLDISRPTMDFVK	95.00%	3.98	0.535	0	1	1	0	2	1,835.37	457	472
2DLC TX114	ilvX (ilvX)	Rv3509c	52022	99.80%	2	3	3	0.02%	6.02%	SGEPAMMLIGGDATR	95.00%	2.95	0.444	0	1	0	0	2	1,506.83	193	207
2DLC TX114	mas (mas)	Rv2940c	224360	100.00%	14	14	24	0.14%	6.82%	AGAALDYSALYPAGR	95.00%	3.97	0.493	0	3	0	0	2	1,496.82	848	862
2DLC TX114	mas (mas)	Rv2940c	224360	100.00%	14	14	24	0.14%	6.82%	AGLELLAFGGR	95.00%	3.89	0.527	0	1	0	0	2	1,104.43	1642	1652
2DLC TX114	mas (mas)	Rv2940c	224360	100.00%	14	14	24	0.14%	6.82%	ANLEQAGLR	95.00%	2.29	0.338	0	1	0	0	2	972.1284	1337	1345
2DLC TX114	mas (mas)	Rv2940c	224360	100.00%	14	14	24	0.14%	6.82%	ELAEFEGELPR	95.00%	3.16	0.442	0	2	0	0	2	1,290.47	1307	1317
2DLC TX114	mas (mas)	Rv2940c	224360	100.00%	14	14	24	0.14%	6.82%	GLLTELHR	95.00%	2.63	0.248	0	1	0	0	2	939.1684	840	847
2DLC TX114	mas (mas)	Rv2940c	224360	100.00%	14	14	24	0.14%	6.82%	GRDGVVIVYGPR	95.00%	2.1	0.229	0	0	1	0	2	1,288.40	1271	1282
2DLC TX114	mas (mas)	Rv2940c	224360	100.00%	14	14	24	0.14%	6.82%	ITRELAEFEGELPR	95.00%	3.83	0.462	0	0	2	0	2	1,660.97	1304	1317
2DLC TX114	mas (mas)	Rv2940c	224360	100.00%	14	14	24	0.14%	6.82%	LFMLSSTSSDALR	95.00%	4.07	0.58	0	4	0	0	2	1,428.82	449	461
2DLC TX114	mas (mas)	Rv2940c	224360	100.00%	14	14	24	0.14%	6.82%	LGLFPFR	95.00%	1.87	0.319	0	2	0	0	2	850.1884	1667	1673
2DLC TX114	mas (mas)	Rv2940c	224360	100.00%	14	14	24	0.14%	6.82%	LLTLGWQQR	95.00%	2.29	0.388	0	1	0	0	2	1,115.37	1190	1198
2DLC TX114	mas (mas)	Rv2940c	224360	100.00%	14	14	24	0.14%	6.82%	LVDAPLPAWTHAR	95.00%	1.93	0.318	0	0	2	0	2	1,447.67	863	875
2DLC TX114	mas (mas)	Rv2940c	224360	100.00%	14	14	24	0.14%	6.82%	LVLDVPR	95.00%	1.89	0.295	0	1	0	0	2	811.9784	1739	1745
2DLC TX114	mas (mas)	Rv2940c	224360	100.00%	14	14	24	0.14%	6.82%	SVAVTPEQAPLYR	95.00%	3.16	0.377	0	1	0	0	2	1,431.54	1750	1762
2DLC TX114	mas (mas)	Rv2940c	224360	100.00%	14	14	24	0.14%	6.82%	VGDPDEHSLLAGR	95.00%	2.08	0.366	0	0	2	0	2	1,366.47	1283	1295
2DLC TX114	ftsY (ftsY)	Rv2921c	43981	100.00%	3	4	7	0.04%	13.70%	IDTSGLPAVGDDATVPR TIADVHLPEFEPEPQAPEVP	95.00%	2.85	0.39	0	2	2	0	2	1,684.84	70	86
2DLC TX114	ftsY (ftsY)	Rv2921c	43981	100.00%	3	4	7	0.04%	13.70%	EADAIAPPEGR	95.00%	4.59	0.59	0	0	2	0	2	3,323.55	92	122
2DLC TX114	ftsY (ftsY)	Rv2921c	43981	100.00%	3	4	7	0.04%	13.70%	VVLGAADTFR	95.00%	2.46	0.516	0	1	0	0	2	1,049.17	254	263
2DLC TX114	hypo protein Rv1488	Rv1488	41232	100.00%	9	11	21	0.12%	27.30%	AGRPTPEMLAYQYLQTLPE MAR	95.00%	2.22	0.326	0	0	1	0	2	2,554.07	266	287

hypo protein																			
2DLC TX114 Rv1488	Rv1488	41232	100.00%	9	11	21	0.12%	27.30%	AMILTAEGTR	95.00%	3.02	0.502	0	4	0	0	2	1,079.33	190	199
hypo protein																			
2DLC TX114 Rv1488	Rv1488	41232	100.00%	9	11	21	0.12%	27.30%	AVATAEAIAR	95.00%	2.89	0.444	0	4	0	0	2	973.0884	357	366
hypo protein																			
2DLC TX114 Rv1488	Rv1488	41232	100.00%	9	11	21	0.12%	27.30%	DQINAQLR	95.00%	2.22	0.435	0	2	0	0	2	958.0984	139	146
hypo protein																			
2DLC TX114 Rv1488	Rv1488	41232	100.00%	9	11	21	0.12%	27.30%	LLGKPGEDGVFR	95.00%	1.85	0.33	0	1	0	0	2	1,288.56	310	321
hypo protein																			
2DLC TX114 Rv1488	Rv1488	41232	100.00%	9	11	21	0.12%	27.30%	QAAILAAEADR	95.00%	2.4	0.315	0	1	0	0	2	1,129.32	222	232
hypo protein																			
2DLC TX114 Rv1488	Rv1488	41232	100.00%	9	11	21	0.12%	27.30%	QAQILAAEGAK	95.00%	1.9	0.389	0	2	0	0	2	1,100.29	211	221
hypo protein																			
2DLC TX114 Rv1488	Rv1488	41232	100.00%	9	11	21	0.12%	27.30%	SIDPPPSIQASMEK	95.00%	2	0.253	0	1	1	0	2	1,500.80	167	180
hypo protein																			
2DLC TX114 Rv1488	Rv1488	41232	100.00%	9	11	21	0.12%	27.30%	TFAAIK	95.00%	1.39	0.288	2	2	0	0	2	650.8784	260	265
2DLC TX114 inhA (inhA)	Rv1484	28511	99.80%	2	2	2	0.01%	6.32%	ALLPIMN	95.00%	0.988	0.142	1	0	0	0	2	788.1178	133	139
2DLC TX114 inhA (inhA)	Rv1484	28511	99.80%	2	2	2	0.01%	6.32%	SNLVAAGPIR	95.00%	2.96	0.494	0	1	0	0	2	998.1984	186	195
2DLC TX114 gcpE (gcpE)	Rv2868c	40431	99.90%	2	2	3	0.02%	5.43%	AAGAAGIPIR	95.00%	2.7	0.457	0	2	0	0	2	897.1584	134	143
2DLC TX114 gcpE (gcpE)	Rv2868c	40431	99.90%	2	2	3	0.02%	5.43%	SAVAFGALLSR	95.00%	3.36	0.458	0	1	0	0	2	1,092.34	230	240
2DLC TX114 moxR	Rv1479	40744	100.00%	8	10	19	0.11%	21.50%	ASLGIIAAAR	95.00%	3.2	0.524	0	2	0	0	2	943.2684	288	297
2DLC TX114 moxR	Rv1479	40744	100.00%	8	10	19	0.11%	21.50%	GHVLLEGVPGVAK	95.00%	4.01	0.594	0	2	0	0	2	1,276.38	72	84
2DLC TX114 moxR	Rv1479	40744	100.00%	8	10	19	0.11%	21.50%	IIVGQDQLVER	95.00%	3.15	0.436	0	2	1	0	2	1,270.41	53	63
2DLC TX114 moxR	Rv1479	40744	100.00%	8	10	19	0.11%	21.50%	MLVGLLSK	95.00%	1.84	0.411	3	2	0	0	2	861.2184	64	71
2DLC TX114 moxR	Rv1479	40744	100.00%	8	10	19	0.11%	21.50%	QILSTGDLLR	95.00%	2.16	0.343	0	1	0	0	2	1,116.45	231	240
2DLC TX114 moxR	Rv1479	40744	100.00%	8	10	19	0.11%	21.50%	TLAVETFAR	95.00%	2.42	0.444	0	4	0	0	2	1,008.15	85	93
2DLC TX114 moxR	Rv1479	40744	100.00%	8	10	19	0.11%	21.50%	VQSALLEVMQER	95.00%	3.65	0.553	0	1	0	0	2	1,403.52	150	161
2DLC TX114 moxR	Rv1479	40744	100.00%	8	10	19	0.11%	21.50%	VVGGTFSR	82.60%	1.69	0.195	0	1	0	0	2	822.8384	94	101
hypo protein																			
2DLC TX114 Rv1473	Rv1473	58375	99.80%	2	2	2	0.01%	4.43%	AGEIGYLPQDPKVGDL	95.00%	1.65	0.208	0	0	1	0	2	1,672.96	63	78
hypo protein																			
2DLC TX114 Rv1473	Rv1473	58375	99.80%	2	2	2	0.01%	4.43%	RTLSGGQR	95.00%	1.45	0.11	1	0	0	0	2	874.9984	160	167
2DLC TX114 trxC (trxC)	Rv3914	12526	99.80%	2	2	4	0.02%	19.80%	ELSDVVPNLN	95.00%	1.56	0.223	2	0	0	0	2	1,100.16	107	116
2DLC TX114 trxC (trxC)	Rv3914	12526	99.80%	2	2	4	0.02%	19.80%	MVAPVLEEIATER	95.00%	2.43	0.48	0	2	0	0	2	1,474.64	42	54
hypo protein																			
2DLC TX114 Rv1466	Rv1466	12366	99.80%	2	2	9	0.05%	16.50%	ALGFTV	95.00%	1.31	0.483	5	0	0	0	2	607.7184	110	115
hypo protein																			
2DLC TX114 Rv1466	Rv1466	12366	99.80%	2	2	9	0.05%	16.50%	SALVGSGLVDDIR	95.00%	4.01	0.57	0	4	0	0	2	1,302.50	74	86
2DLC TX114 bisC (bisC)	Rv1442	83373	99.70%	2	2	3	0.02%	3.13%	DPDSMCVHGN	94.30%	1.31	0.365	0	1	0	0	2	1,132.09	712	721
2DLC TX114 bisC (bisC)	Rv1442	83373	99.70%	2	2	3	0.02%	3.13%	IAMPAMVPPYANAR	95.00%	2.58	0.205	0	0	2	0	2	1,518.87	486	499
2DLC TX114 accA3 accA3	Rv3285	63732	100.00%	5	5	8	0.05%	12.20%	DLGDKVTAR GAHTGAAASGDAVTAPMQ	95.00%	2.26	0.338	0	2	0	0	2	975.1184	122	130
2DLC TX114 (accA3)	Rv3285	63732	100.00%	5	5	8	0.05%	12.20%	GTVK	82.20%	2.29	0.193	0	0	1	0	2	2,098.11	522	544
2DLC TX114 accA3	Rv3285	63732	100.00%	5	5	8	0.05%	12.20%	IANGEKLDITEDPTPR	95.00%	3.75	0.359	0	0	2	0	2	1,770.09	325	340
2DLC TX114 accA3	Rv3285	63732	100.00%	5	5	8	0.05%	12.20%	NFLPAPGPVTK	95.00%	1.9	0.308	0	2	0	0	2	1,141.38	356	366
2DLC TX114 accA3	Rv3285	63732	100.00%	5	5	8	0.05%	12.20%	TIDEIPELYESAVR	95.00%	4.16	0.483	0	1	0	0	2	1,635.87	183	196
hypo protein																			
2DLC TX114 Rv3281	Rv3281	18995	100.00%	3	4	13	0.08%	15.30%	ITLQEMTHMR	95.00%	2.87	0.308	0	1	1	0	2	1,260.52	167	176
hypo protein																			
2DLC TX114 Rv3281	Rv3281	18995	100.00%	3	4	13	0.08%	15.30%	WGLPVDQLR	95.00%	2.56	0.491	0	6	0	0	2	1,084.26	150	158
hypo protein																			
2DLC TX114 Rv3281	Rv3281	18995	100.00%	3	4	13	0.08%	15.30%	YPVFSWQR	95.00%	2.64	0.352	0	5	0	0	2	1,083.19	159	166
2DLC TX114 accD5	Rv3280	59337	99.80%	2	2	4	0.02%	2.55%	GYIGTALR	95.00%	2.03	0.421	0	2	0	0	2	851.0984	522	529
2DLC TX114 accD5	Rv3280	59337	99.80%	2	2	4	0.02%	2.55%	ITVITR	95.00%	1.91	0	2	0	0	0	2	702.8884	424	429
2DLC TX114 gap (gap)	Rv1436	35907	100.00%	8	12	43	0.25%	31.60%	AAALNIVPTSTGAAK	95.00%	2.28	0.461	0	2	0	0	2	1,385.64	206	220
2DLC TX114 gap (gap)	Rv1436	35907	100.00%	8	12	43	0.25%	31.60%	AIGLVMPQLK	95.00%	3.33	0.498	0	3	0	0	2	1,070.44	221	230
2DLC TX114 gap (gap)	Rv1436	35907	100.00%	8	12	43	0.25%	31.60%	ASVDEINAAFK	95.00%	2.69	0.57	1	9	0	0	2	1,165.31	258	268
2DLC TX114 gap (gap)	Rv1436	35907	100.00%	8	12	43	0.25%	31.60%	EGPAALPWGDL	95.00%	2.01	0.306	0	1	0	0	2	1,126.32	85	95
2DLC TX114 gap (gap)	Rv1436	35907	100.00%	8	12	43	0.25%	31.60%	LVDLVTLVGK	95.00%	3.36	0.483	4	8	0	0	2	1,057.28	328	337
2DLC TX114 gap (gap)	Rv1436	35907	100.00%	8	12	43	0.25%	31.60%	VLDDEFGIVK	95.00%	1.49	0.363	2	2	0	0	2	1,135.31	169	178
2DLC TX114 gap (gap)	Rv1436	35907	100.00%	8	12	43	0.25%	31.60%	VVSWYDNEWGYSNR YYDAPIVSSDIVTDPHSSIF	95.00%	4.81	0.595	0	3	1	0	2	1,775.77	314	327
2DLC TX114 gap (gap)	Rv1436	35907	100.00%	8	12	43	0.25%	31.60%	DSGLTK	95.00%	4.58	0.655	0	0	7	0	2	2,829.23	281	306
hypo protein																			
2DLC TX114 Rv0730	Rv0730	25962	99.80%	2	5	10	0.06%	5.37%	EIADALLAALER	95.00%	3.24	0.414	1	4	1	0	2	1,285.64	26	37
hypo protein																			
2DLC TX114 Rv0730	Rv0730	25962	99.80%	2	5	10	0.06%	5.37%	REIADALLAALER	95.00%	3.75	0.529	0	3	1	0	2	1,441.84	25	37
hypo protein																			
2DLC TX114 Rv1184c	Rv1184c	37801	100.00%	5	6	13	0.08%	20.60%	GLDPLTAPGSIVEGAR	95.00%	2.42	0.485	0	1	0	0	2	1,553.82	334	349
hypo protein																			
2DLC TX114 Rv1184c	Rv1184c	37801	100.00%	5	6	13	0.08%	20.60%	GLLGSPAFGG	95.00%	1.28	0.484	4	2	0	0	2	876.0984	350	359
hypo protein									LANDPTAPAPDKLQFTTFG										
2DLC TX114 Rv1184c	Rv1184c	37801	100.00%	5	6	13	0.08%	20.60%	DPTGR	95.00%	4.08	0.684	0	0	2	0	2	2,531.94	154	177
hypo protein																			
2DLC TX114 Rv1184c	Rv1184c	37801	100.00%	5	6	13	0.08%	20.60%	NDNWFTRPVSVDPVR	95.00%	3.44	0.423	0	0	2	0	2	1,802.85	319	333
hypo protein																			
2DLC TX114 Rv1184c	Rv1184c	37801	100.00%	5	6	13	0.08%	20.60%	VVYALGGAR	95.00%	2.38	0.43	0	2	0	0	2	906.0184	52	60
hypo protein																			
2DLC TX114 Rv3269	Rv3269	9732	99.80%	2	3	4	0.02%	11.80%	LKVADVMAEAR	95.00%	3.24	0.516	0	1	2	0	2	1,203.40	63	73
hypo protein																			
2DLC TX114 Rv3269	Rv3269	9732	99.80%	2	3	4	0.02%	11.80%	VADVMAEAR	95.00%	3	0.461	0	1	0	0	2	961.9984	65	73
hypo protein																			
2DLC TX114 Rv1423	Rv1423	34915	99.80%	2	2	3	0.02%	8.62%	ALEILGDTVPEHLASAGK	95.00%	3.37	0.621	0	1	0	0	2	1,822.13	247	264
hypo protein																			
2DLC TX114 Rv1423	Rv1423	34915	99.80%	2	2	3	0.02%	8.62%	LANFDDANLR	95.00%	3.01	0.422	0	2	0	0	2	1,149.38	222	231
mmpS3 2DLC TX114 (mmpS3)	Rv2198c	30936	99.80%	2	3	6	0.03%	12.40%	APGDIISVTYVDAAGR	95.00%	3.66	0.608	0	2	0	0	2	1,605.80	223	238
mmpS3 2DLC TX114 (mmpS3)	Rv2198c	30936	99.80%	2	3	6	0.03%	12.40%	AYSAPESEHVTGGPYVPAD LR	95.00%	3.78	0.541	0	2	2	0	2	2,217.32	56	76
2DLC TX114 rpsD (rpsD)	Rv3458c	23458	99.80%	2	2	2	0.01%	10.40%	TDLVGGDQAFEK	95.00%	2.32	0.382	0	1	0	0	2	1,280.37	17	28
2DLC TX114 rpsD (rpsD)	Rv3458c	23458	99.80%	2	2	2	0.01%	10.40%	VLIHQLPER	95.00%	1.78	0.373	0	0	1	0	2	1,105.31	175	183
hypo protein																			
2DLC TX114 Rv3520c	Rv3520c	37567	100.00%	4	4	7	0.04%	12.10%	DLAALV	95.00%	1.13	0.256	2	0	0	0	2	601.7484	342	347
hypo protein																			
2DLC TX114 Rv3520c	Rv3520c	37567	100.00%	4	4	7	0.04%	12.10%	LGTSVIQLSAR	95.00%	2.52	0.331	0	1	0	0	2	1,145.38	63	73

2DLC TX114	hypo protein Rv3520c	Rv3520c	37567	100.00%	4	4	7	0.04%	12.10%	LPLTGEGTTGLGK	95.00%	2.65	0.478	0	2	0	0	2	1,244.54	146	158
2DLC TX114	hypo protein Rv3520c	Rv3520c	37567	100.00%	4	4	7	0.04%	12.10%	MGYTQVVDEVTK	95.00%	3.87	0.569	0	2	0	0	2	1,370.38	273	284
2DLC TX114	hypo protein Rv0108c	Rv0108c	7554	99.80%	2	2	3	0.02%	24.60%	VNHQLQVIEK	95.00%	2.58	0.162	0	0	1	0	2	1,208.26	44	53
2DLC TX114	hypo protein Rv0108c	Rv0108c	7554	99.80%	2	2	3	0.02%	24.60%	YIVLESK	95.00%	2.03	0.341	0	2	0	0	2	852.0684	23	29
2DLC TX114	ompA (ompA)	Rv0899	33525	99.80%	2	2	4	0.02%	11.00%	GVAGDHIATVGLGSVNPIA SNATPEGR	95.00%	4.24	0.645	0	0	3	0	2	2,561.70	289	315
2DLC TX114	ompA	Rv0899	33525	99.80%	2	2	4	0.02%	11.00%	IVADYLVAR	95.00%	2.69	0.369	0	1	0	0	2	1,020.22	280	288
2DLC TX114	PE (PE)	Rv2431c	10669	99.80%	2	2	3	0.02%	36.40%	AATFLVEYAR	95.00%	2.46	0.202	0	2	0	0	2	1,141.33	51	60
2DLC TX114	PE (PE)	Rv2431c	10669	99.80%	2	2	3	0.02%	36.40%	AIQSDAQVAPMTTAVRPPA ADLVSEK	95.00%	3.48	0.604	0	0	1	0	2	2,667.93	25	50
2DLC TX114	fadE10 (fadE10)	Rv0873	70728	100.00%	8	8	12	0.07%	14.20%	AAQIPDEYVK	95.00%	1.6	0.344	0	1	0	0	2	1,134.27	79	88
2DLC TX114	fadE10 (fadE10)	Rv0873	70728	100.00%	8	8	12	0.07%	14.20%	EFGALATHLR	95.00%	2.15	0.398	0	1	0	0	2	1,115.33	488	497
2DLC TX114	fadE10 (fadE10)	Rv0873	70728	100.00%	8	8	12	0.07%	14.20%	FLTVSPSSEAKL	95.00%	2.38	0.421	0	2	0	0	2	1,279.53	639	650
2DLC TX114	fadE10 (fadE10)	Rv0873	70728	100.00%	8	8	12	0.07%	14.20%	FPLGLIHPFPKPSDAEEAR	95.00%	5.3	0.568	0	0	2	0	2	2,122.59	38	56
2DLC TX114	fadE10 (fadE10)	Rv0873	70728	100.00%	8	8	12	0.07%	14.20%	IFEGSSEIMR	95.00%	3.52	0.374	0	2	0	0	2	1,169.43	420	429
2DLC TX114	fadE10 (fadE10)	Rv0873	70728	100.00%	8	8	12	0.07%	14.20%	LANDVLEGR	95.00%	2.87	0.452	0	1	0	0	2	987.1184	595	603
2DLC TX114	fadE10 (fadE10)	Rv0873	70728	100.00%	8	8	12	0.07%	14.20%	LREFLDTVDGSVIER	95.00%	3.88	0.465	0	0	2	0	2	1,749.99	64	78
2DLC TX114	fadE10 (fadE10)	Rv0873	70728	100.00%	8	8	12	0.07%	14.20%	TEAFLVK	95.00%	1.71	0.276	0	1	0	0	2	807.9684	57	63
2DLC TX114	hypo protein Rv3224	Rv3224	29797	100.00%	4	4	8	0.05%	17.00%	DGANIALIAK	95.00%	3.35	0.512	0	3	0	0	2	986.3084	28	37
2DLC TX114	hypo protein Rv3224	Rv3224	29797	100.00%	4	4	8	0.05%	17.00%	FDLMNGIQVR	95.00%	2.92	0.536	0	3	0	0	2	1,193.44	113	122
2DLC TX114	hypo protein Rv3224	Rv3224	29797	100.00%	4	4	8	0.05%	17.00%	TMFISGASR	95.00%	2.31	0.531	0	1	0	0	2	970.2084	7	15
2DLC TX114	hypo protein Rv3224	Rv3224	29797	100.00%	4	4	8	0.05%	17.00%	WLRPTAYMMAKYGMTLCAL	95.00%	2.37	0.215	0	0	1	0	2	2,253.93	156	174
2DLC TX114	fadB (fadB)	Rv0860	76088	100.00%	6	6	11	0.06%	9.44%	FLPPESLLS	95.00%	1.07	0.356	2	0	0	0	2	1,003.32	712	720
2DLC TX114	fadB (fadB)	Rv0860	76088	100.00%	6	6	11	0.06%	9.44%	SDALLAR	95.00%	2.06	0.161	0	1	0	0	2	745.9584	386	392
2DLC TX114	fadB (fadB)	Rv0860	76088	100.00%	6	6	11	0.06%	9.44%	SGLWPGLR TMIQARPEDAGDVFNTVET	82.30%	2.04	0.271	0	1	0	0	2	886.1184	613	620
2DLC TX114	fadB (fadB)	Rv0860	76088	100.00%	6	6	11	0.06%	9.44%	IKR	95.00%	6.02	0.587	0	0	2	0	2	2,492.79	72	93
2DLC TX114	fadB (fadB)	Rv0860	76088	100.00%	6	6	11	0.06%	9.44%	VFDYTLAIGK	95.00%	3.17	0.599	0	3	0	0	2	1,127.42	489	498
2DLC TX114	fadB (fadB)	Rv0860	76088	100.00%	6	6	11	0.06%	9.44%	YFASLVTGQVAK	95.00%	3.05	0.654	0	2	0	0	2	1,284.45	286	297
2DLC TX114	rhlE (rhlE)	Rv3211	56686	99.40%	2	2	2	0.01%	4.74%	RRDTDCDGQKTAQH	87.80%	1.5	0.278	0	0	1	0	2	1,689.68	430	443
2DLC TX114	rhlE (rhlE)	Rv3211	56686	99.40%	2	2	2	0.01%	4.74%	TFAFGVPLLQR	95.00%	2.48	0.422	0	1	0	0	2	1,249.54	61	71
2DLC TX114	fadA (fadA)	Rv0859	42397	100.00%	3	3	5	0.03%	7.20%	AAEAWSGGYFAK	95.00%	3.98	0.556	0	2	0	0	2	1,258.42	178	189
2DLC TX114	fadA (fadA)	Rv0859	42397	100.00%	3	3	5	0.03%	7.20%	FQKDLNIPDEKL	95.00%	2.59	0.191	0	2	0	0	2	1,460.84	337	348
2DLC TX114	fadA (fadA)	Rv0859	42397	100.00%	3	3	5	0.03%	7.20%	RSQQKAAEAWSGGYFAK	85.00%	1.93	0.182	0	0	1	0	2	1,886.10	173	189
2DLC TX114	hypo protein Rv0854	Rv0854	16328	99.80%	2	3	4	0.02%	14.30%	GTIDTATEALR	95.00%	0.932	0.302	1	1	0	0	2	1,148.33	127	137
2DLC TX114	hypo protein Rv0854	Rv0854	16328	99.80%	2	3	4	0.02%	14.30%	WTLVSSTQQR	95.00%	2.67	0.43	0	2	0	0	2	1,206.25	78	87
2DLC TX114	rplM (rplM)	Rv3443c	16287	99.80%	2	2	4	0.02%	12.20%	AILGMLPK	95.00%	2.18	0.292	0	3	0	0	2	843.2684	104	111
2DLC TX114	rplM (rplM)	Rv3443c	16287	99.80%	2	2	4	0.02%	12.20%	LAVAAANLLR	95.00%	3.06	0.482	0	1	0	0	2	1,012.31	28	37
2DLC TX114	trkA (trkA)	Rv2691	24222	99.90%	2	2	4	0.02%	8.37%	FGTGVLPEPK	95.00%	1.9	0.403	0	2	0	0	2	1,045.23	179	188
2DLC TX114	trkA (trkA)	Rv2691	24222	99.90%	2	2	4	0.02%	8.37%	VYIAAISGR	95.00%	2.76	0.483	0	2	0	0	2	950.1784	197	205
2DLC TX114	gatC (gatC)	Rv3012c	10698	99.40%	2	2	2	0.01%	27.30%	DVNVTRPDETVPCL	88.40%	1.81	0.212	0	0	1	0	2	1,558.57	59	72
2DLC TX114	gatC (gatC)	Rv3012c	10698	99.40%	2	2	2	0.01%	27.30%	QVLDQAPDAVDGR	95.00%	2.19	0.49	0	1	0	0	2	1,384.39	76	88
2DLC TX114	hupB (hupB)	Rv2986c	22169	100.00%	6	8	22	0.13%	21.50%	AELIDVLTQK	95.00%	2.95	0.485	2	4	0	0	2	1,130.37	4	13
2DLC TX114	hupB (hupB)	Rv2986c	22169	100.00%	6	8	22	0.13%	21.50%	AVVSGAQR	95.00%	1.8	0.259	0	1	0	0	2	787.7484	87	94
2DLC TX114	hupB (hupB)	Rv2986c	22169	100.00%	6	8	22	0.13%	21.50%	LPAEGPAVK	95.00%	2.25	0.414	0	3	0	0	2	882.0384	95	103
2DLC TX114	hupB (hupB)	Rv2986c	22169	100.00%	6	8	22	0.13%	21.50%	MNKAELIDVLTQK	95.00%	4.61	0.521	0	2	5	0	2	1,503.87	1	13
2DLC TX114	hupB (hupB)	Rv2986c	22169	100.00%	6	8	22	0.13%	21.50%	QATAAVENVVDTIVR	95.00%	4.53	0.577	0	3	0	0	2	1,586.54	20	34
2DLC TX114	hupB (hupB)	Rv2986c	22169	100.00%	6	8	22	0.13%	21.50%	RQATAAVENVVDTIVR	95.00%	3.21	0.366	0	0	2	0	2	1,742.74	19	34
2DLC TX114	hypo protein Rv2129c	Rv2129c	31314	99.80%	2	2	4	0.02%	5.80%	LTVVADVR	95.00%	1.69	0.319	0	1	0	0	2	872.9084	57	64
2DLC TX114	hypo protein Rv2129c	Rv2129c	31314	99.80%	2	2	4	0.02%	5.80%	VVFITGAAR	95.00%	2.69	0.429	0	3	0	0	2	934.0684	8	16
2DLC TX114	dnaJ2	Rv2373c	40438	99.80%	2	2	3	0.02%	5.76%	ELHPDVNPDEAAQAK	95.00%	2.84	0.296	0	0	2	0	2	1,634.67	30	44
2DLC TX114	dnaJ2	Rv2373c	40438	99.80%	2	2	3	0.02%	5.76%	SLLGQML	95.00%	1.57	0.256	1	0	0	0	2	778.0478	177	183
2DLC TX114	hypo protein Rv0048c	Rv0048c	30768	99.80%	2	2	3	0.02%	5.54%	AVTLGDLQR	95.00%	2.57	0.279	0	2	0	0	2	973.1184	55	63
2DLC TX114	hypo protein Rv0048c	Rv0048c	30768	99.80%	2	2	3	0.02%	5.54%	TAVGIMR	95.00%	1.99	0.3	0	1	0	0	2	747.9184	221	227
2DLC TX114	hypo protein Rv2556c	Rv2556c	13874	100.00%	3	3	6	0.03%	12.40%	IVDLTDAVR	95.00%	2.9	0.504	0	4	0	0	2	1,002.14	11	19
2DLC TX114	hypo protein Rv2556c	Rv2556c	13874	100.00%	3	3	6	0.03%	12.40%	LSFVEG	95.00%	1.85	0.481	1	0	0	0	2	651.7184	124	129
2DLC TX114	hypo protein Rv2556c	Rv2556c	13874	100.00%	3	3	6	0.03%	12.40%	RIVDLTDAVR	95.00%	2.39	0.297	0	1	0	0	2	1,158.34	10	19
2DLC TX114	hypo protein Rv0183	Rv0183	34889	99.80%	2	3	4	0.02%	7.12%	LGAAGLVTYALDHR	95.00%	3.37	0.463	0	1	2	0	2	1,457.73	97	110
2DLC TX114	hypo protein Rv0183	Rv0183	34889	99.80%	2	3	4	0.02%	7.12%	NFAGIGDVR	95.00%	2.91	0.351	0	1	0	0	2	949.0684	53	61
2DLC TX114	katG (katG)	Rv1908c	80556	100.00%	6	6	10	0.06%	10.10%	DGAGAGTIPDPFGGPGR	95.00%	1.99	0.457	0	1	0	0	2	1,542.74	357	373
2DLC TX114	katG (katG)	Rv1908c	80556	100.00%	6	6	10	0.06%	10.10%	FAPLNSWPDNASLDK	95.00%	2.19	0.37	0	1	0	0	2	1,675.98	129	143
2DLC TX114	katG (katG)	Rv1908c	80556	100.00%	6	6	10	0.06%	10.10%	FVQDFVAAWDK	95.00%	3.55	0.538	0	1	0	0	2	1,326.42	720	730
2DLC TX114	katG (katG)	Rv1908c	80556	100.00%	6	6	10	0.06%	10.10%	SPTMLATDLSLR	95.00%	2.7	0.439	0	2	0	0	2	1,305.68	374	385
2DLC TX114	katG (katG)	Rv1908c	80556	100.00%	6	6	10	0.06%	10.10%	VDLVFGSNSELR	95.00%	2.83	0.55	0	3	0	0	2	1,336.47	694	705
2DLC TX114	katG (katG)	Rv1908c	80556	100.00%	6	6	10	0.06%	10.10%	YLGPLVPK	95.00%	1.91	0.401	0	2	0	0	2	887.1784	426	433

2DLC TX114	hypo protein Rv2623	Rv2623	31633	100.00%	10	12	41	0.24%	30.00%	AGPPTVHSEIVPAAAVPTLV DMSK	95.00%	2.68	0.422	0	0	1	0	2	2,388.61	86	109
2DLC TX114	hypo protein Rv2623	Rv2623	31633	100.00%	10	12	41	0.24%	30.00%	HLIDDALK	95.00%	2.93	0.535	0	5	0	0	2	925.1984	70	77
2DLC TX114	hypo protein Rv2623	Rv2623	31633	100.00%	10	12	41	0.24%	30.00%	LAGWQER	91.30%	1.83	0.172	0	1	0	0	2	859.9484	224	230
2DLC TX114	hypo protein Rv2623	Rv2623	31633	100.00%	10	12	41	0.24%	30.00%	LAGWQERYPNVAITR	95.00%	4.51	0.362	0	0	2	0	2	1,775.02	224	238
2DLC TX114	hypo protein Rv2623	Rv2623	31633	100.00%	10	12	41	0.24%	30.00%	LLGSVSSGLLR	95.00%	1.54	0.275	2	10	0	0	2	1,102.43	128	138
2DLC TX114	hypo protein Rv2623	Rv2623	31633	100.00%	10	12	41	0.24%	30.00%	SEEAQLVVVGSR	95.00%	3.19	0.464	0	3	0	0	2	1,274.22	253	264
2DLC TX114	hypo protein Rv2623	Rv2623	31633	100.00%	10	12	41	0.24%	30.00%	TPVIVAR	95.00%	1.55	0.347	2	6	0	0	2	755.8584	287	293
2DLC TX114	hypo protein Rv2623	Rv2623	31633	100.00%	10	12	41	0.24%	30.00%	TPVIVARESLT	95.00%	2.86	0.511	0	2	0	0	2	1,186.34	287	297
2DLC TX114	hypo protein Rv2623	Rv2623	31633	100.00%	10	12	41	0.24%	30.00%	VVEQASLR	95.00%	2.71	0.429	0	2	0	0	2	901.9184	78	85
2DLC TX114	hypo protein Rv2623	Rv2623	31633	100.00%	10	12	41	0.24%	30.00%	YPNVAITR	95.00%	2.67	0.399	0	5	0	0	2	934.0884	231	238
2DLC TX114	hypo protein Rv1038c	Rv1197,Rv2347c	10959	100.00%	5	8	21	0.12%	51.00%	DANNYEQQEQASQQILSS	95.00%	3.42	0.587	0	3	1	0	2	2,054.02	81	98
2DLC TX114	hypo protein Rv1038c	Rv1197,Rv2347c	10959	100.00%	5	8	21	0.12%	51.00%	FEVHAQTVEDEAR	95.00%	1.89	0.204	0	0	1	0	2	1,531.42	19	31
2DLC TX114	hypo protein Rv1038c	Rv1197,Rv2347c	10959	100.00%	5	8	21	0.12%	51.00%	FMTDPHAMR	95.00%	2.07	0.541	0	2	1	0	2	1,106.32	5	13
2DLC TX114	hypo protein Rv1038c	Rv1197,Rv2347c	10959	100.00%	5	8	21	0.12%	51.00%	NIVNML	95.00%	1.36	0.213	1	0	0	0	2	719.8878	66	71
2DLC TX114	hypo protein Rv1038c	Rv1197,Rv2347c	10959	100.00%	5	8	21	0.12%	51.00%	NIVNMLHGVR	95.00%	1.57	0.212	2	10	0	0	2	1,153.31	66	75
2DLC TX114	fbpA (fbpA)	Rv3804c	35668	100.00%	3	4	8	0.05%	9.17%	ALGATPNTGPAPQGA	95.00%	1.79	0.442	0	2	0	0	2	1,323.45	324	338
2DLC TX114	fbpA (fbpA)	Rv3804c	35668	100.00%	3	4	8	0.05%	9.17%	FLEGFVR	95.00%	1.64	0.324	2	0	0	0	2	868.0384	271	277
2DLC TX114	fbpA (fbpA)	Rv3804c	35668	100.00%	3	4	8	0.05%	9.17%	NDPLLNVGK	95.00%	2.16	0.0152	2	2	0	0	2	970.1584	234	242
2DLC TX114	hypo protein Rv0174	Rv0174	53988	100.00%	3	3	4	0.02%	10.90%	DPYVPLGTNPWFGDPNQIL TCPAPGAR	91.50%	1.91	0.22	0	0	1	0	2	2,898.32	384	410
2DLC TX114	hypo protein Rv0174	Rv0174	53988	100.00%	3	3	4	0.02%	10.90%	GTVPSEIGPALDNSNR	95.00%	2.27	0.505	0	2	0	0	2	1,627.77	130	145
2DLC TX114	hypo protein Rv0174	Rv0174	53988	100.00%	3	3	4	0.02%	10.90%	LVDSTQAIVGDFK	95.00%	3.56	0.608	0	1	0	0	2	1,393.57	176	188
2DLC TX114	lprK (lprK)	Rv0173	42223	99.80%	2	3	5	0.03%	6.15%	LLAYVGGR	95.00%	2.09	0.456	0	2	0	0	2	849.0684	219	226
2DLC TX114	lprK (lprK)	Rv0173	42223	99.80%	2	3	5	0.03%	6.15%	YTPNPNDAPGGPLVER	95.00%	2.94	0.328	0	1	2	0	2	1,697.85	370	385
2DLC TX114	hypo protein Rv0172	Rv0172	56481	100.00%	3	3	8	0.05%	6.04%	GDFFAVVR	95.00%	2.12	0.423	0	2	0	0	2	910.9884	207	214
2DLC TX114	hypo protein Rv0172	Rv0172	56481	100.00%	3	3	8	0.05%	6.04%	GGPVLADNAVIPVER	95.00%	2.18	0.505	0	1	0	0	2	1,507.63	124	138
2DLC TX114	hypo protein Rv0172	Rv0172	56481	100.00%	3	3	8	0.05%	6.04%	NIQLEPPYR	95.00%	2.79	0.397	0	5	0	0	2	1,130.36	115	123
2DLC TX114	hypo protein Rv2047c	Rv2047c	91335	99.80%	2	2	2	0.01%	3.28%	CGRDTAPAAGPELVSAR	90.90%	1.81	0.305	0	1	0	0	2	1,729.86	514	530
2DLC TX114	hypo protein Rv2047c	Rv2047c	91335	99.80%	2	2	2	0.01%	3.28%	RGVGVCGGRVR	91.20%	1.72	0.249	0	1	0	0	2	1,116.13	737	747
2DLC TX114	mce1 (mce1)	Rv0169	47770	100.00%	3	3	5	0.03%	9.69%	DTYSPELFCTIRNFYDADPL AK	95.00%	2.16	0.262	0	0	2	0	2	2,638.15	289	310
2DLC TX114	mce1 (mce1)	Rv0169	47770	100.00%	3	3	5	0.03%	9.69%	FILDVDPR	95.00%	1.92	0.269	0	1	0	0	2	975.2084	90	97
2DLC TX114	mce1 (mce1)	Rv0169	47770	100.00%	3	3	5	0.03%	9.69%	YIHLIPANVNADIK	95.00%	3.94	0.411	0	0	2	0	2	1,581.97	98	111
2DLC TX114	hypo protein Rv1390	Rv1390	11793	99.80%	2	3	8	0.05%	20.00%	EIHADLLEHTEGE	95.00%	1.85	0.566	0	2	2	0	2	1,493.55	98	110
2DLC TX114	hypo protein Rv1390	Rv1390	11793	99.80%	2	3	8	0.05%	20.00%	YALVIYAAK	95.00%	2.89	0.612	0	4	0	0	2	1,012.33	51	59
2DLC TX114	hypo protein Rv0227c	Rv0227c	45510	99.80%	2	2	3	0.02%	6.18%	FTQNVGYTPEGK	95.00%	2.99	0.479	0	1	0	0	2	1,341.41	197	208
2DLC TX114	hypo protein Rv0227c	Rv0227c	45510	99.80%	2	2	3	0.02%	6.18%	GLNDENPPTAIPLR	95.00%	2.15	0.441	0	2	0	0	2	1,507.81	138	151
2DLC TX114	fadD32 (fadD32)	Rv3801c	69214	100.00%	14	15	34	0.20%	23.40%	AAYLDGSLR	95.00%	1.93	0.451	0	2	0	0	2	966.2084	616	624
2DLC TX114	fadD32 (fadD32)	Rv3801c	69214	100.00%	14	15	34	0.20%	23.40%	DILWSDFSAR	95.00%	2.64	0.423	0	4	0	0	2	1,210.46	61	70
2DLC TX114	fadD32 (fadD32)	Rv3801c	69214	100.00%	14	15	34	0.20%	23.40%	DVLLVSAGTIPR	95.00%	2.39	0.516	0	2	0	0	2	1,241.49	592	603
2DLC TX114	fadD32 (fadD32)	Rv3801c	69214	100.00%	14	15	34	0.20%	23.40%	FFEAFAPYGLK	95.00%	2.63	0.516	0	3	0	0	2	1,290.65	332	342
2DLC TX114	fadD32 (fadD32)	Rv3801c	69214	100.00%	14	15	34	0.20%	23.40%	FLDFSTER	95.00%	2.61	0.232	0	1	0	0	2	1,015.20	48	55
2DLC TX114	fadD32 (fadD32)	Rv3801c	69214	100.00%	14	15	34	0.20%	23.40%	FLDFSTERDGVAR	95.00%	2.17	0.504	0	0	2	0	2	1,513.70	48	60
2DLC TX114	fadD32 (fadD32)	Rv3801c	69214	100.00%	14	15	34	0.20%	23.40%	GILNGSEPVSPASMR	95.00%	2.5	0.646	0	7	0	0	2	1,515.75	316	330
2DLC TX114	fadD32 (fadD32)	Rv3801c	69214	100.00%	14	15	34	0.20%	23.40%	GVPRDDEPPLDLSNVK	95.00%	2.86	0.408	0	0	2	0	2	1,751.95	300	315
2DLC TX114	fadD32 (fadD32)	Rv3801c	69214	100.00%	14	15	34	0.20%	23.40%	IKDLVIIDGR	95.00%	2.89	0.182	0	1	0	0	2	1,142.54	491	500
2DLC TX114	fadD32 (fadD32)	Rv3801c	69214	100.00%	14	15	34	0.20%	23.40%	IRFPANTNLVR	95.00%	2.9	0.257	0	0	1	0	2	1,301.59	21	31
2DLC TX114	fadD32 (fadD32)	Rv3801c	69214	100.00%	14	15	34	0.20%	23.40%	ISESRAEGAPDDALWVR	95.00%	3.92	0.393	0	0	2	0	2	1,873.06	457	473
2DLC TX114	fadD32 (fadD32)	Rv3801c	69214	100.00%	14	15	34	0.20%	23.40%	KFFEAFAPYGLK	95.00%	4.15	0.425	0	2	2	0	2	1,418.85	331	342
2DLC TX114	fadD32 (fadD32)	Rv3801c	69214	100.00%	14	15	34	0.20%	23.40%	LDHQPIVDDIR	95.00%	2.92	0.235	0	0	2	0	2	1,321.51	569	579
2DLC TX114	fadD32 (fadD32)	Rv3801c	69214	100.00%	14	15	34	0.20%	23.40%	SGVGSPTVFATSD	84.00%	0.901	0	1	0	0	0	2	1,225.20	625	637
2DLC TX114	hypo protein Rv0148	Rv0148	29760	100.00%	4	4	10	0.06%	18.50%	IAGFKL	91.40%	1.84	0.176	0	1	0	0	2	648.9484	281	286

2DLC TX114	hypo protein Rv0148	Rv0148	29760	100.00%	4	4	10	0.06%	18.50%	VALFGNDGANFDKPPSVQD VAAR	95.00%	4.14	0.637	0	0	6	0	2	2,389.57	247	269
2DLC TX114	hypo protein Rv0148	Rv0148	29760	100.00%	4	4	10	0.06%	18.50%	VIVVTGAGGGLGR	95.00%	2.97	0.0762	0	1	0	0	2	1,156.26	8	20
2DLC TX114	hypo protein Rv0148	Rv0148	29760	100.00%	4	4	10	0.06%	18.50%	WAEITDLSGAK	95.00%	2.6	0.261	0	2	0	0	2	1,191.41	270	280
2DLC TX114	hypo protein Rv0147	Rv0147	55018	99.50%	2	2	2	0.01%	5.14%	KQQLRALQK	89.60%	1.44	0.198	1	0	0	0	2	1,113.40	71	79
2DLC TX114	hypo protein Rv0147	Rv0147	55018	99.50%	2	2	2	0.01%	5.14%	LPFGGVGASGMGAYHGR	95.00%	3.97	0.603	0	0	1	0	2	1,634.85	449	465
2DLC TX114	hisI (hisI)	Rv2122c	10225	99.80%	2	2	2	0.01%	24.70%	GLSLLDDVYR	95.00%	2.46	0.464	0	1	0	0	2	1,038.22	83	91
2DLC TX114	hisI (hisI)	Rv2122c	10225	99.80%	2	2	2	0.01%	24.70%	LLEEAGEVWLAAEH	95.00%	2.12	0.228	0	1	0	0	2	1,567.68	45	58
2DLC TX114	bfrB (bfrB)	Rv3841	20424	100.00%	5	5	9	0.05%	34.30%	AGANLFELENFVAR	95.00%	3.33	0.5	0	3	0	0	2	1,551.78	149	162
2DLC TX114	bfrB (bfrB)	Rv3841	20424	100.00%	5	5	9	0.05%	34.30%	EALALALDQER	95.00%	2.87	0.438	0	2	0	0	2	1,229.46	90	100
2DLC TX114	bfrB (bfrB)	Rv3841	20424	100.00%	5	5	9	0.05%	34.30%	EVDVAPAASGAPHAAGGRL	95.00%	1.44	0.361	0	0	1	0	2	1,746.81	163	181
2DLC TX114	bfrB (bfrB)	Rv3841	20424	100.00%	5	5	9	0.05%	34.30%	TVTDQVGR	95.00%	2.21	0.385	0	2	0	0	2	875.8084	101	108
2DLC TX114	bfrB (bfrB)	Rv3841	20424	100.00%	5	5	9	0.05%	34.30%	VEIPGVDTVR	95.00%	1.93	0.295	0	1	0	0	2	1,085.10	73	82
2DLC TX114	pks13	Rv3800c	186428	100.00%	15	16	33	0.19%	12.10%	AAAAHVLPDAQLTAAEQR	95.00%	3.86	0.45	0	0	4	0	2	1,833.97	1112	1129
2DLC TX114	pks13 (pks13)	Rv3800c	186428	100.00%	15	16	33	0.19%	12.10%	AEGPITAEDVLTSSNIEALA DKVR	95.00%	3.4	0.549	0	0	1	0	2	2,500.81	1421	1444
2DLC TX114	pks13	Rv3800c	186428	100.00%	15	16	33	0.19%	12.10%	AGEEIPQTKEEIR	95.00%	2.66	0.333	0	0	1	0	2	1,500.67	1564	1576
2DLC TX114	pks13	Rv3800c	186428	100.00%	15	16	33	0.19%	12.10%	ALGQIEADRTSEVGKQ	95.00%	3.01	0.521	0	0	3	0	2	1,702.83	1718	1733
2DLC TX114	pks13	Rv3800c	186428	100.00%	15	16	33	0.19%	12.10%	DGNVDLAALVR	95.00%	3.73	0.433	0	5	0	0	2	1,143.27	1101	1111
2DLC TX114	pks13 (pks13)	Rv3800c	186428	100.00%	15	16	33	0.19%	12.10%	DIKGFDSEFFAVAKTEADNI DPQQRMAL	95.00%	1.78	0.309	0	0	1	0	2	3,174.66	182	209
2DLC TX114	pks13	Rv3800c	186428	100.00%	15	16	33	0.19%	12.10%	FVGLIDAVR	95.00%	1.97	0.174	2	2	0	0	2	990.1884	1555	1563
2DLC TX114	pks13	Rv3800c	186428	100.00%	15	16	33	0.19%	12.10%	MITTPTDWPR	95.00%	1.77	0.415	0	1	0	0	2	1,218.46	507	516
2DLC TX114	pks13	Rv3800c	186428	100.00%	15	16	33	0.19%	12.10%	SEGGGMLVLK	95.00%	2.87	0.43	0	1	0	0	2	991.2184	352	361
2DLC TX114	pks13	Rv3800c	186428	100.00%	15	16	33	0.19%	12.10%	SNGLIAPNQDAQADVLR	95.00%	1.87	0.282	0	1	0	0	2	1,782.98	388	404
2DLC TX114	pks13	Rv3800c	186428	100.00%	15	16	33	0.19%	12.10%	SPGGIFNELPR	95.00%	3.95	0.459	0	1	0	0	2	1,187.44	1394	1404
2DLC TX114	pks13	Rv3800c	186428	100.00%	15	16	33	0.19%	12.10%	SPGGIFNELPRLDDEAAAK	95.00%	3.07	0.427	0	0	3	0	2	2,001.38	1394	1412
2DLC TX114	pks13	Rv3800c	186428	100.00%	15	16	33	0.19%	12.10%	TYLEAGQIDGFVR	95.00%	3.86	0.566	0	3	0	0	2	1,469.67	1445	1457
2DLC TX114	pks13	Rv3800c	186428	100.00%	15	16	33	0.19%	12.10%	WSEFLEEPR	95.00%	2.59	0.313	0	2	0	0	2	1,193.32	157	165
2DLC TX114	pks13	Rv3800c	186428	100.00%	15	16	33	0.19%	12.10%	WSPDSGETIAER	95.00%	1.94	0.352	0	2	0	0	2	1,348.43	1226	1237
2DLC TX114	hypo protein Rv3134c	Rv3134c	27989	99.80%	2	2	2	0.01%	9.70%	AATHAALWAVDEAVNR	95.00%	2.97	0.167	0	0	1	0	2	1,695.74	19	34
2DLC TX114	hypo protein Rv3134c	Rv3134c	27989	99.80%	2	2	2	0.01%	9.70%	AVVVGIDGSR	95.00%	2.62	0.521	0	1	0	0	2	972.9884	9	18
2DLC TX114	hypo protein Rv3870	Rv3870	80897	99.80%	2	2	2	0.01%	3.21%	AGVTVITLGNHR	95.00%	2.1	0.239	0	0	1	0	2	1,238.34	337	348
2DLC TX114	hypo protein Rv3870	Rv3870	80897	99.80%	2	2	2	0.01%	3.21%	MGEVLTGELDRR	95.00%	1.86	0.332	0	0	1	0	2	1,376.59	544	555
2DLC TX114	hypo protein Rv3133c	Rv3133c	23245	100.00%	8	9	19	0.11%	27.20%	AAAALMAK	95.00%	2.65	0.319	0	1	0	0	2	747.0184	129	136
2DLC TX114	hypo protein Rv3133c	Rv3133c	23245	100.00%	8	9	19	0.11%	27.20%	DLLSR	95.00%	1.59	0.152	2	0	0	0	2	603.7984	68	72
2DLC TX114	hypo protein Rv3133c	Rv3133c	23245	100.00%	8	9	19	0.11%	27.20%	MFLAEK	95.00%	2.03	0.27	0	2	0	0	2	738.9984	174	179
2DLC TX114	hypo protein Rv3133c	Rv3133c	23245	100.00%	8	9	19	0.11%	27.20%	MFLAEKTVKN	95.00%	2.14	0.185	1	0	0	0	2	1,197.47	174	183
2DLC TX114	hypo protein Rv3133c	Rv3133c	23245	100.00%	8	9	19	0.11%	27.20%	PDVAVLDVR	95.00%	3.42	0.389	0	3	0	0	2	984.0284	48	56
2DLC TX114	hypo protein Rv3133c	Rv3133c	23245	100.00%	8	9	19	0.11%	27.20%	TQAAVFATELK	95.00%	3.39	0.522	0	2	0	0	2	1,179.33	198	208
2DLC TX114	hypo protein Rv3133c	Rv3133c	23245	100.00%	8	9	19	0.11%	27.20%	VFLVDDHEVVR	95.00%	1.63	0.256	0	0	1	0	2	1,328.30	4	14
2DLC TX114	hypo protein Rv3133c	Rv3133c	23245	100.00%	8	9	19	0.11%	27.20%	VPAARPDVAVLDVR	95.00%	3.07	0.497	0	2	5	0	2	1,478.58	43	56
2DLC TX114	hypo protein Rv3874	Rv3874	10776	100.00%	6	11	41	0.24%	80.00%	ADEEQQQALSSQMGF	95.00%	3.71	0.499	0	5	5	0	2	1,685.69	86	100
2DLC TX114	hypo protein Rv3874	Rv3874	10776	100.00%	6	11	41	0.24%	80.00%	GAAGTAAQAAVVR	95.00%	4.34	0.371	0	5	1	0	2	1,143.14	45	57
2DLC TX114	hypo protein Rv3874	Rv3874	10776	100.00%	6	11	41	0.24%	80.00%	ISGDLK	95.00%	1.3	0.0999	1	0	0	0	2	632.8484	21	26
2DLC TX114	hypo protein Rv3874	Rv3874	10776	100.00%	6	11	41	0.24%	80.00%	QKQELDEISTNIR	95.00%	3.08	0.443	0	2	6	0	2	1,574.80	65	77
2DLC TX114	hypo protein Rv3874	Rv3874	10776	100.00%	6	11	41	0.24%	80.00%	TDAATLAQEAGNFER	95.00%	4.46	0.587	0	3	4	0	2	1,594.69	6	20
2DLC TX114	hypo protein Rv3874	Rv3874	10776	100.00%	6	11	41	0.24%	80.00%	TQIDQVESTAGSLQGQWR	95.00%	4.64	0.528	0	5	4	0	2	2,005.03	27	44
2DLC TX114	hypo protein Rv3871	Rv3871	64543	99.80%	2	2	2	0.01%	3.89%	EVVLDQLGTAESR	95.00%	3.27	0.539	0	1	0	0	2	1,417.47	12	24
2DLC TX114	hypo protein Rv3871	Rv3871	64543	99.80%	2	2	2	0.01%	3.89%	FALGIMDEPR	95.00%	3.62	0.414	0	1	0	0	2	1,149.47	53	62
2DLC TX114	hypo protein Rv3132c	Rv3132c	62222	99.80%	2	3	5	0.03%	3.63%	AIVHSATSLVDAR	95.00%	3	0.557	0	2	1	0	2	1,340.46	70	82
2DLC TX114	hypo protein Rv3132c	Rv3132c	62222	99.80%	2	3	5	0.03%	3.63%	TFLGVPVR	95.00%	1.76	0.47	0	2	0	0	2	889.0284	154	161
2DLC TX114	hypo protein Rv3869	Rv3869	51075	100.00%	3	3	6	0.03%	8.12%	FVALQSPDPR	95.00%	2.4	0.357	0	3	0	0	2	1,130.29	394	403
2DLC TX114	hypo protein Rv3869	Rv3869	51075	100.00%	3	3	6	0.03%	8.12%	LVLGNPANPATVK	95.00%	2.26	0.212	0	1	0	0	2	1,294.51	103	115
2DLC TX114	hypo protein Rv3869	Rv3869	51075	100.00%	3	3	6	0.03%	8.12%	SLGLSSPQNAPWEIVR	95.00%	2.64	0.513	0	2	0	0	2	1,755.00	428	443
2DLC TX114	hypo protein Rv3865	Rv3865	10600	100.00%	3	3	7	0.04%	35.90%	ATDTVAGISGR	95.00%	2.96	0.329	0	1	0	0	2	1,048.13	29	39
2DLC TX114	hypo protein Rv3865	Rv3865	10600	100.00%	3	3	7	0.04%	35.90%	FNDTLQEFETTR	95.00%	2.41	0.458	0	2	0	0	2	1,501.62	51	62
2DLC TX114	hypo protein Rv3865	Rv3865	10600	100.00%	3	3	7	0.04%	35.90%	VLAGMHNEIVGDIK	95.00%	3.09	0.457	0	0	4	0	2	1,496.74	14	27

2DLC TX114	hns (hns)	Rv3852	13806	100.00%	5	6	28	0.16%	35.10%	IETNGQLAAAAK	95.00%	2.86	0.479	0	8	0	0	2	1,187.39	70	81
2DLC TX114	hns (hns)	Rv3852	13806	100.00%	5	6	28	0.16%	35.10%	IETNGQLAAAAKDAAAQAK	95.00%	4.64	0.585	0	2	0	0	2	1,843.11	70	88
2DLC TX114	hns (hns)	Rv3852	13806	100.00%	5	6	28	0.16%	35.10%	PAEAPVSLQQR	95.00%	3.06	0.469	0	2	0	0	2	1,196.27	59	69
2DLC TX114	hns (hns)	Rv3852	13806	100.00%	5	6	28	0.16%	35.10%	SAPPKPAEAPVSLQQR	95.00%	2.26	0.46	0	2	8	0	2	1,676.87	54	69
2DLC TX114	hns (hns)	Rv3852	13806	100.00%	5	6	28	0.16%	35.10%	STVEGANDALAR	95.00%	2.69	0.439	0	6	0	0	2	1,204.26	89	100
2DLC TX114	hypo protein Rv3849	Rv3849	14691	100.00%	4	4	12	0.07%	37.90%	AEGITMSAPYLSQLR	95.00%	4.22	0.599	0	4	0	0	2	1,638.01	35	49
2DLC TX114	hypo protein Rv3849	Rv3849	14691	100.00%	4	4	12	0.07%	37.90%	LFDTVYPPGR	95.00%	2.21	0.493	0	5	0	0	2	1,165.38	12	21
2DLC TX114	hypo protein Rv3849	Rv3849	14691	100.00%	4	4	12	0.07%	37.90%	TNPSGATMAALANFFR	95.00%	4.31	0.632	0	1	0	0	2	1,685.99	54	69
2DLC TX114	hypo protein Rv3849	Rv3849	14691	100.00%	4	4	12	0.07%	37.90%	VLDRIDELR	95.00%	1.86	0.273	0	0	2	0	2	1,129.39	117	125
2DLC TX114	hypo protein Rv0831c	Rv0831c	30171	99.80%	2	2	3	0.02%	8.49%	AQVSSIVGLER	95.00%	2.89	0.223	0	1	0	0	2	1,159.25	118	128
2DLC TX114	hypo protein Rv0831c	Rv0831c	30171	99.80%	2	2	3	0.02%	8.49%	NQAIVVETTAYR	95.00%	3.11	0.459	0	2	0	0	2	1,365.42	90	101
2DLC TX114	metK (metK)	Rv1392	42997	99.80%	2	3	7	0.04%	7.44%	AIGEVFDLRPGAIIR	95.00%	2.84	0.398	0	2	1	0	2	1,628.07	349	363
2DLC TX114	metK (metK)	Rv1392	42997	99.80%	2	3	7	0.04%	7.44%	FVLGGPMGDAGLTGR	95.00%	3.56	0.481	0	4	0	0	2	1,464.74	251	265
2DLC TX114	mIHF (mIHF)	Rv1388	20800	100.00%	7	8	22	0.13%	25.30%	GGTNLTQVLK	95.00%	2.21	0.416	0	2	0	0	2	1,031.19	120	129
2DLC TX114	mIHF (mIHF)	Rv1388	20800	100.00%	7	8	22	0.13%	25.30%	KALLEK	95.00%	1.85	0.31	0	1	0	0	2	701.9984	181	186
2DLC TX114	mIHF (mIHF)	Rv1388	20800	100.00%	7	8	22	0.13%	25.30%	KALLEKFGSA	95.00%	2.83	0.323	0	1	0	0	2	1,064.41	181	190
2DLC TX114	mIHF (mIHF)	Rv1388	20800	100.00%	7	8	22	0.13%	25.30%	LEALPK	95.00%	1.16	0.151	2	0	0	0	2	670.9184	146	151
2DLC TX114	mIHF (mIHF)	Rv1388	20800	100.00%	7	8	22	0.13%	25.30%	MKVSALLEALPK	95.00%	3.84	0.484	0	1	0	0	2	1,300.75	140	151
2DLC TX114	mIHF (mIHF)	Rv1388	20800	100.00%	7	8	22	0.13%	25.30%	VKAQEIMTELEIAPTR	95.00%	4.43	0.519	0	3	0	0	2	1,830.17	155	170
2DLC TX114	mIHF (mIHF)	Rv1388	20800	100.00%	7	8	22	0.13%	25.30%	VSALLEALPK	95.00%	3.72	0.389	4	8	0	0	2	1,041.35	142	151
2DLC TX114	pyrR (pyrR)	Rv1379	20609	99.80%	2	2	2	0.01%	10.90%	VPTSRSESVHVR	95.00%	1.28	0.341	0	0	1	0	2	1,354.29	168	179
2DLC TX114	pyrR (pyrR)	Rv1379	20609	99.80%	2	2	2	0.01%	10.90%	VVLLGIPTR	95.00%	1.96	0.4	0	1	0	0	2	968.2284	50	58
2DLC TX114	hypo protein Rv3688c	Rv3688c	16684	99.80%	2	2	3	0.02%	14.30%	MLLAAIQTEEVSGK	95.00%	2.3	0.398	0	1	0	0	2	1,490.78	29	42
2DLC TX114	hypo protein Rv3688c	Rv3688c	16684	99.80%	2	2	3	0.02%	14.30%	SDLTQAMK	95.00%	2.47	0.519	0	2	0	0	2	894.0784	10	17
2DLC TX114	hypo protein Rv3075c	Rv3075c	33083	99.80%	2	3	6	0.03%	7.82%	LAFGIGDFR	95.00%	2.34	0.403	0	3	0	0	2	996.2984	156	164
2DLC TX114	hypo protein Rv3075c	Rv3075c	33083	99.80%	2	3	6	0.03%	7.82%	RLPNVPIVALVETAR	95.00%	3.57	0.398	0	1	2	0	2	1,648.93	125	139
2DLC TX114	lprF (lprF)	Rv1368	26833	100.00%	5	7	19	0.11%	26.10%	DAIDGLATVK	95.00%	1.54	0.39	2	4	0	0	2	1,003.20	178	187
2DLC TX114	lprF (lprF)	Rv1368	26833	100.00%	5	7	19	0.11%	26.10%	GGDYVSVGPAEK	95.00%	3.34	0.55	0	2	0	0	2	1,179.19	137	148
2DLC TX114	lprF (lprF)	Rv1368	26833	100.00%	5	7	19	0.11%	26.10%	GLGAVVGQVQNPTIQGR	95.00%	4.3	0.63	0	2	0	0	2	1,694.73	161	177
2DLC TX114	lprF (lprF)	Rv1368	26833	100.00%	5	7	19	0.11%	26.10%	IYDPGILDKDR	95.00%	2.63	0.54	0	2	4	0	2	1,418.89	149	160
2DLC TX114	lprF (lprF)	Rv1368	26833	100.00%	5	7	19	0.11%	26.10%	MKPNTPVVATEFLVTNK	95.00%	3.82	0.396	0	0	3	0	2	1,890.15	114	130
2DLC TX114	hypo protein Rv1324	Rv1324	32120	100.00%	5	6	11	0.06%	17.80%	KSYQAILDANPGSVEAK	95.00%	4.31	0.449	0	0	1	0	2	1,792.06	190	206
2DLC TX114	hypo protein Rv1324	Rv1324	32120	100.00%	5	6	11	0.06%	17.80%	LIALVR	95.00%	1.6	0.0444	2	1	0	0	2	684.9684	259	264
2DLC TX114	hypo protein Rv1324	Rv1324	32120	100.00%	5	6	11	0.06%	17.80%	NLANALY	95.00%	1.46	0.459	2	0	0	0	2	778.9784	298	304
2DLC TX114	hypo protein Rv1324	Rv1324	32120	100.00%	5	6	11	0.06%	17.80%	QQLEDGDFVAAR	95.00%	1.98	0.457	0	2	0	0	2	1,349.40	178	189
2DLC TX114	hypo protein Rv1324	Rv1324	32120	100.00%	5	6	11	0.06%	17.80%	WVDSLLSATAGK	95.00%	3.22	0.559	0	3	0	0	2	1,248.46	145	156
2DLC TX114	atpC (atpC)	Rv1311	13117	100.00%	3	4	11	0.06%	25.60%	FLFTR	95.00%	1.24	0.112	3	0	0	0	2	683.9184	22	26
2DLC TX114	atpC (atpC)	Rv1311	13117	100.00%	3	4	11	0.06%	25.60%	HIPLVAQLVDDAMVR	95.00%	3.63	0.469	0	0	1	0	2	1,693.91	38	52
2DLC TX114	atpC (atpC)	Rv1311	13117	100.00%	3	4	11	0.06%	25.60%	TTVGEIGILPR	95.00%	1.38	0.199	2	5	0	0	2	1,156.41	27	37
2DLC TX114	atpD (atpD)	Rv1310	53077	100.00%	14	21	59	0.34%	34.20%	FTQAGSEVSTLLGR	95.00%	3.08	0.512	0	7	0	0	2	1,466.63	270	283
2DLC TX114	atpD (atpD)	Rv1310	53077	100.00%	14	21	59	0.34%	34.20%	GIFPAVDPLASSSTILDPSV VGDEHYR	95.00%	3.01	0.543	0	1	5	0	2	2,844.17	352	378
2DLC TX114	atpD (atpD)	Rv1310	53077	100.00%	14	21	59	0.34%	34.20%	GVEVIDTGR	95.00%	2.69	0.328	0	2	0	0	2	945.9584	87	95
2DLC TX114	atpD (atpD)	Rv1310	53077	100.00%	14	21	59	0.34%	34.20%	IALFGGAGVGK	95.00%	2.07	0.403	1	2	0	0	2	990.2484	167	177
2DLC TX114	atpD (atpD)	Rv1310	53077	100.00%	14	21	59	0.34%	34.20%	KAESLGAKL	95.00%	2.95	0.411	0	2	0	0	2	917.2084	478	486
2DLC TX114	atpD (atpD)	Rv1310	53077	100.00%	14	21	59	0.34%	34.20%	KPPAFEELEPR	95.00%	3.23	0.134	0	0	1	0	2	1,313.56	134	144
2DLC TX114	atpD (atpD)	Rv1310	53077	100.00%	14	21	59	0.34%	34.20%	NFGSTSVFAGVGER	95.00%	2.51	0.315	0	1	0	0	2	1,398.42	191	204
2DLC TX114	atpD (atpD)	Rv1310	53077	100.00%	14	21	59	0.34%	34.20%	SISVPVGEGVK	95.00%	1.75	0.378	0	3	0	0	2	1,072.11	96	106
2DLC TX114	atpD (atpD)	Rv1310	53077	100.00%	14	21	59	0.34%	34.20%	TEMLETGLK	95.00%	2.62	0.51	0	2	0	0	2	1,022.27	145	153
2DLC TX114	atpD (atpD)	Rv1310	53077	100.00%	14	21	59	0.34%	34.20%	TISLQPTDGLVR	95.00%	2.33	0.455	0	9	0	0	2	1,300.54	75	86
2DLC TX114	atpD (atpD)	Rv1310	53077	100.00%	14	21	59	0.34%	34.20%	TVLIQEMINR	95.00%	3.24	0.419	3	5	1	0	2	1,233.49	178	187
2DLC TX114	atpD (atpD)	Rv1310	53077	100.00%	14	21	59	0.34%	34.20%	VAQEVIR	95.00%	1.87	0.241	0	1	0	0	2	814.8384	379	385
2DLC TX114	atpD (atpD)	Rv1310	53077	100.00%	14	21	59	0.34%	34.20%	VTGPVVDVEFPR	95.00%	3.39	0.501	2	7	2	0	2	1,315.29	26	37
2DLC TX114	atpD (atpD)	Rv1310	53077	100.00%	14	21	59	0.34%	34.20%	VVDLLTPYVR	95.00%	2.12	0.454	1	1	0	0	2	1,175.35	154	163
2DLC TX114	nuoI (nuoI)	Rv3153	23417	99.50%	2	2	3	0.02%	14.20%	PEMAAPPHPR TPGATDKDYYLGNVTAEGL	90.00%	1.83	0.277	0	1	0	0	2	1,103.26	170	179
2DLC TX114	nuoI (nuoI)	Rv3153	23417	99.50%	2	2	3	0.02%	14.20%	R	95.00%	4.35	0.588	0	0	2	0	2	2,142.42	180	199
2DLC TX114	atpG (atpG)	Rv1309	33874	100.00%	5	6	14	0.08%	18.70%	ALTLMANR	95.00%	2.21	0.529	0	2	0	0	2	890.1784	274	281
2DLC TX114	atpG (atpG)	Rv1309	33874	100.00%	5	6	14	0.08%	18.70%	LESARPYAFEITR	95.00%	2.52	0.393	0	1	0	0	2	1,553.88	40	52
2DLC TX114	atpG (atpG)	Rv1309	33874	100.00%	5	6	14	0.08%	18.70%	SATDNADDLIK	95.00%	2	0.386	0	2	0	0	2	1,163.36	263	273
2DLC TX114	atpG (atpG)	Rv1309	33874	100.00%	5	6	14	0.08%	18.70%	SEELFSLLR	95.00%	3.06	0.321	0	2	0	0	2	1,094.38	100	108
2DLC TX114	atpG (atpG)	Rv1309	33874	100.00%	5	6	14	0.08%	18.70%	VYALLESAASELASR	95.00%	3.83	0.44	0	6	1	0	2	1,651.93	242	257
2DLC TX114	atpA (atpA)	Rv1308	59271	100.00%	12	17	47	0.27%	21.30%	AIDAMTPIGR	95.00%	2.33	0.375	0	5	0	0	2	1,045.35	155	164
2DLC TX114	atpA (atpA)	Rv1308	59271	100.00%	12	17	47	0.27%	21.30%	AISLLLR	95.00%	1.55	0.159	2	2	0	0	2	786.1784	283	289
2DLC TX114	atpA (atpA)	Rv1308	59271	100.00%	12	17	47	0.27%	21.30%	ALELQAPSVVHR	95.00%	2.46	0.449	0	3	5	0	2	1,320.42	131	142
2DLC TX114	atpA (atpA)	Rv1308	59271	100.00%	12	17	47	0.27%	21.30%	DSQKLTEEAADKLTEVIK	95.00%	3.6	0.551	0	0	1	0	2	2,019.33	488	505
2DLC TX114	atpA (atpA)	Rv1308	59271	100.00%	12	17	47	0.27%	21.30%	EAYPGDVFYLHSR	95.00%	1.96	0.353	0	2	0	0	2	1,554.72	295	307
2DLC TX114	atpA (atpA)	Rv1308	59271	100.00%	12	17	47	0.27%	21.30%	LDLSQYR	95.00%	1.84	0.346	0	4	0	0	2	895.0984	395	401
2DLC TX114	atpA (atpA)	Rv1308	59271	100.00%	12	17	47	0.27%	21.30%	LTEEAADKLTEVIK	95.00%	3.28	0.368	0	2	0	0	2	1,560.85	492	505
2DLC TX114	atpA (atpA)	Rv1308	59271	100.00%	12	17	47	0.27%	21.30%	LVELLK	95.00%	1.75	0.147	1	2	0	0	2	714.9884	427	432
2DLC TX114	atpA (atpA)	Rv1308	59271	100.00%	12	17	47	0.27%	21.30%	TGEVLSVPVGDGFLGR	95.00%	3.64	0.571	0	3	1	0	2	1,603.73	94	109
2DLC TX114	atpA (atpA)	Rv1308	59271	100.00%	12	17	47	0.27%	21.30%	VGGAAQIK	91.10%	2.08	0.155	0	1	0	0	2	743.8484	377	384
2DLC TX114	atpA (atpA)	Rv1308	59271	100.00%	12	17	47	0.27%	21.30%	VVNPLGQPIDGR VVNPLGQPIDGRGDVDSDT	95.00%	2.84	0.575	0	6	3	0	2	1,265.40	110	121
2DLC TX114	atpA (atpA)	Rv1308	59271	100.00%	12	17	47	0.27%	21.30%	R	95.00%	1.81	0.486	0	0	4	0	2	2,111.20	110	129
2DLC TX114	atpH (atpH)	Rv1307	48787	100.00%	6	7	12	0.07%	14.60%	ILDVQPR	95.00%	2.05	0.302	0	2	0	0	2	841.0084	302	308
2DLC TX114	atpH (atpH)	Rv1307	48787	100.00%	6	7	12	0.07%	14.60%	IAAAEARLPD	95.00%	3.17	0.444	0	1	0	0	2	1,027.26	437	446
2DLC TX114	atpH (atpH)	Rv1307	48787	100.00%	6	7	12	0.07%	14.60%	LLDREAVVTR	95.00%	2.09	0.224	0	0	1	0	2	1,172.34	212	221
2DLC TX114	atpH (atpH)	Rv1307	48787	100.00%	6	7	12	0.07%	14.60%	VGAPTLEVLR	95.00%	3.08	0.435	0	2	0	0	2	1,055.21	245	254
2DLC TX114	atpH (atpH)	Rv1307	48787	100.00%	6	7	12	0.07%	14.60%	WSANSDLIDAIEHVSR	95.00%	3.43	0.457	0	2	2	0	2	1,813.99	261	276
2DLC TX114	atpH (atpH)	Rv1307	48787	100.00%	6	7	12	0.07%	14.60%	YLTVPAEDATPR	95.00%	1.91	0.475	0	2	0	0	2	1,333.49	222	233

2DLC TX114	atpF (atpF)	Rv1306	18307	100.00%	7	10	28	0.16%	35.10%	AEQQVASTLQTAHEQLK	95.00%	4.09	0.696	0	2	5	0	2	1,882.91	118	134
2DLC TX114	atpF (atpF)	Rv1306	18307	100.00%	7	10	28	0.16%	35.10%	AHVGTMSATL	95.00%	1.92	0.275	0	1	0	0	2	988.0784	146	155
2DLC TX114	atpF (atpF)	Rv1306	18307	100.00%	7	10	28	0.16%	35.10%	AHVGTMSATLASR	95.00%	2.63	0.377	0	0	4	0	2	1,302.44	146	158
2DLC TX114	atpF (atpF)	Rv1306	18307	100.00%	7	10	28	0.16%	35.10%	DAVELDLR	95.00%	1.66	0.184	2	0	0	0	2	931.0684	138	145
2DLC TX114	atpF (atpF)	Rv1306	18307	100.00%	7	10	28	0.16%	35.10%	ERDAVELDLR	95.00%	2.8	0.259	0	1	3	0	2	1,216.37	136	145
2DLC TX114	atpF (atpF)	Rv1306	18307	100.00%	7	10	28	0.16%	35.10%	ILGVDLTASAATR	95.00%	3.91	0.612	2	6	0	0	2	1,288.56	159	171
2DLC TX114	atpF (atpF)	Rv1306	18307	100.00%	7	10	28	0.16%	35.10%	VQASSLR	95.00%	1.47	0.241	2	0	0	0	2	760.8284	94	100
2DLC TX114	hypo protein Rv3143	Rv3143	14336	100.00%	3	3	5	0.03%	25.60%	AEAAVPHPVDPIVLGR	95.00%	1.87	0.226	0	0	1	0	2	1,641.78	107	122
2DLC TX114	hypo protein Rv3143	Rv3143	14336	100.00%	3	3	5	0.03%	25.60%	ILVYSDNVQTR	95.00%	2.45	0.564	0	2	0	0	2	1,308.44	10	20
2DLC TX114	hypo protein Rv3143	Rv3143	14336	100.00%	3	3	5	0.03%	25.60%	TVLSLLR	95.00%	1.22	0.151	2	0	0	0	2	802.0684	123	129
2DLC TX114	fadE23 (fadE23)	Rv3140	43327	100.00%	4	4	7	0.04%	8.23%	AALEEIR	83.60%	1.89	0.172	0	1	0	0	2	801.9784	268	274
2DLC TX114	fadE23 (fadE23)	Rv3140	43327	100.00%	4	4	7	0.04%	8.23%	IFVTAGSR	95.00%	2.29	0.37	0	2	0	0	2	850.9984	163	170
2DLC TX114	fadE23 (fadE23)	Rv3140	43327	100.00%	4	4	7	0.04%	8.23%	KLQAIIVK	95.00%	2.7	0.421	0	2	0	0	2	913.2684	10	17
2DLC TX114	fadE23 (fadE23)	Rv3140	43327	100.00%	4	4	7	0.04%	8.23%	LLGLSSSELK	95.00%	2.54	0.429	0	2	0	0	2	1,047.41	392	401
2DLC TX114	fadE24 (fadE24)	Rv3139	49629	99.80%	2	2	2	0.01%	6.62%	AAELGITAINIPEDFDGIAE HR	95.00%	3.54	0.524	0	0	1	0	2	2,353.76	149	170
2DLC TX114	fadE24 (fadE24)	Rv3139	49629	99.80%	2	2	2	0.01%	6.62%	AEQGLPFAR	95.00%	2.58	0.376	0	1	0	0	2	989.1484	406	414
2DLC TX114	PPE (PPE)	Rv3136	37962	100.00%	3	3	8	0.05%	8.16%	IQLLAL	95.00%	1.36	0.0083	1	0	0	0	2	670.9984	115	120
2DLC TX114	PPE (PPE)	Rv3136	37962	100.00%	3	3	8	0.05%	8.16%	LTVMAHPPAAG	95.00%	2.3	0.491	0	5	0	0	2	1,065.22	370	380
2DLC TX114	PPE (PPE)	Rv3136	37962	100.00%	3	3	8	0.05%	8.16%	MDFALLPPEVNSAR	95.00%	1.67	0.317	0	2	0	0	2	1,560.87	1	14
2DLC TX114	hypo protein Rv2185c	Rv2185c	16276	99.90%	2	2	4	0.02%	13.20%	LIDGALKDLK	95.00%	2.66	0.175	0	2	0	0	2	1,086.55	130	139
2DLC TX114	hypo protein Rv2185c	Rv2185c	16276	99.90%	2	2	4	0.02%	13.20%	MLMDAAIFK	95.00%	2.54	0.498	0	2	0	0	2	1,040.48	54	62
2DLC TX114	hypo protein Rv0771	Rv0771	16056	99.80%	2	2	5	0.03%	11.10%	ANVAEVLAK	95.00%	2.15	0.431	0	2	0	0	2	914.9984	130	138
2DLC TX114	hypo protein Rv0771	Rv0771	16056	99.80%	2	2	5	0.03%	11.10%	LNSAIER	95.00%	1.73	0.278	0	3	0	0	2	802.9784	105	111
2DLC TX114	cspB (cspB)	Rv0871	14971	100.00%	3	3	5	0.03%	23.00%	LIEPPPSLSR	95.00%	1.6	0.429	0	2	0	0	2	1,109.44	64	73
2DLC TX114	cspB (cspB)	Rv0871	14971	100.00%	3	3	5	0.03%	23.00%	SSALPTGVEALK	95.00%	2.35	0.441	0	2	0	0	2	1,173.38	30	41
2DLC TX114	cspB (cspB)	Rv0871	14971	100.00%	3	3	5	0.03%	23.00%	VEFGIASGR	95.00%	2.18	0.504	0	1	0	0	2	936.0484	46	54
2DLC TX114	hypo protein Rv3127	Rv3127	38470	100.00%	6	6	14	0.08%	23.80%	AQAILQR	95.00%	2.2	0.255	0	2	0	0	2	799.9784	109	115
2DLC TX114	hypo protein Rv3127	Rv3127	38470	100.00%	6	6	14	0.08%	23.80%	FPQPNQPDQLATVEFSPID HVTAGQR	95.00%	6.89	0.715	0	0	2	0	2	2,894.03	81	106
2DLC TX114	hypo protein Rv3127	Rv3127	38470	100.00%	6	6	14	0.08%	23.80%	LVVASQLSEVLR	95.00%	4.28	0.538	0	3	0	0	2	1,314.47	157	168
2DLC TX114	hypo protein Rv3127	Rv3127	38470	100.00%	6	6	14	0.08%	23.80%	RPLDSVLQIR	95.00%	2.84	0.132	0	1	0	0	2	1,197.49	312	321
2DLC TX114	hypo protein Rv3127	Rv3127	38470	100.00%	6	6	14	0.08%	23.80%	VGIAPPLAAVPAPTPR	95.00%	2.81	0.445	0	2	0	0	2	1,527.83	296	311
2DLC TX114	hypo protein Rv3127	Rv3127	38470	100.00%	6	6	14	0.08%	23.80%	VLVLSTPSDTR	95.00%	2.44	0.474	0	4	0	0	2	1,188.34	230	240
2DLC TX114	hypo protein Rv1109c	Rv1109c	22940	100.00%	3	3	10	0.06%	21.70%	APFQTLLANR	95.00%	2.81	0.435	0	4	0	0	2	1,131.40	196	205
2DLC TX114	hypo protein Rv1109c	Rv1109c	22940	100.00%	3	3	10	0.06%	21.70%	FALYSVSDTPETTTASR	95.00%	3.64	0.571	0	4	0	0	2	1,847.01	113	129
2DLC TX114	hypo protein Rv1109c sseC2	Rv1109c	22940	100.00%	3	3	10	0.06%	21.70%	TILMYPMTLASQAAHVVMR ALSAAGNGDAVVQPSGAGI	95.00%	1.58	0.303	0	0	2	0	2	2,182.60	28	46
2DLC TX114	(sseC2)	Rv0814c,Rv3118	10147	100.00%	4	4	15	0.09%	71.00%	HEVDVK	95.00%	3.39	0.592	0	0	4	0	2	2,363.38	74	98
2DLC TX114	sseC2 sseC2	Rv0814c,Rv3118	10147	100.00%	4	4	15	0.09%	71.00%	FFAAPGSWTLR	95.00%	2.77	0.62	0	4	0	0	2	1,253.53	63	73
2DLC TX114	(sseC2)	Rv0814c,Rv3118	10147	100.00%	4	4	15	0.09%	71.00%	QGLTLPASVDLEKETVITGR	95.00%	2.48	0.504	0	0	2	0	2	2,128.44	7	26
2DLC TX114	sseC2	Rv0814c,Rv3118	10147	100.00%	4	4	15	0.09%	71.00%	VVDGDGQAVGGAFVR	95.00%	4.95	0.548	0	5	0	0	2	1,447.36	27	41
2DLC TX114	cysA2	Rv0815c,Rv3117	30997	100.00%	4	7	12	0.07%	12.30%	AFRDEVLAAINVK	95.00%	2.84	0.272	0	1	4	0	2	1,446.70	146	158
2DLC TX114	cysA2	Rv0815c,Rv3117	30997	100.00%	4	7	12	0.07%	12.30%	DFVDAQQFSK	95.00%	3.37	0.487	1	3	0	0	2	1,185.25	58	67
2DLC TX114	cysA2	Rv0815c,Rv3117	30997	100.00%	4	7	12	0.07%	12.30%	LLSER	95.00%	1.37	0.0014	1	0	0	0	2	617.7984	68	72
2DLC TX114	cysA2	Rv0815c,Rv3117	30997	100.00%	4	7	12	0.07%	12.30%	NLIDVR	95.00%	1.47	0.271	1	1	0	0	2	729.8884	159	164
2DLC TX114	fprA (fprA)	Rv3106	49323	99.80%	2	2	2	0.01%	3.95%	FLTSPIEIK	95.00%	1.77	0.274	0	1	0	0	2	1,048.42	271	279
2DLC TX114	fprA (fprA)	Rv3106	49323	99.80%	2	2	2	0.01%	3.95%	LASLAELLR	95.00%	2.45	0.231	0	1	0	0	2	986.3584	444	452
2DLC TX114	hypo protein Rv2978c	Rv2978c	51387	97.60%	2	2	2	0.01%	4.79%	DATEMLRQK	86.50%	1.88	0.208	0	1	0	0	2	1,108.30	315	323
2DLC TX114	hypo protein Rv2978c	Rv2978c	51387	97.60%	2	2	2	0.01%	4.79%	WQTSRAGKRAGKR	82.00%	1.88	0.176	0	0	1	0	2	1,502.76	113	125
2DLC TX114	adk (adk)	Rv0733	20076	100.00%	4	4	11	0.06%	26.00%	LGIPQISTGELFR	95.00%	2.47	0.312	0	3	0	0	2	1,431.82	24	36
2DLC TX114	adk (adk)	Rv0733	20076	100.00%	4	4	11	0.06%	26.00%	TVDAVGTMDEVFAR	95.00%	2.45	0.441	0	3	0	0	2	1,511.54	161	174
2DLC TX114	adk (adk)	Rv0733	20076	100.00%	4	4	11	0.06%	26.00%	VLLLGPPGAGK	95.00%	2.37	0.396	0	3	0	0	2	1,022.36	3	13
2DLC TX114	adk (adk)	Rv0733	20076	100.00%	4	4	11	0.06%	26.00%	VSEEVLLER	95.00%	2.52	0.292	0	2	0	0	2	1,074.14	115	123
2DLC TX114	hypo protein Rv3615c	Rv3615c	10777	100.00%	3	4	18	0.10%	17.50%	AIDGLFT	95.00%	0.98	0.295	5	0	0	0	2	736.9484	97	103
2DLC TX114	hypo protein Rv3615c	Rv3615c	10777	100.00%	3	4	18	0.10%	17.50%	IYSEADEAWR	95.00%	2.34	0.549	0	5	0	0	2	1,240.36	86	95
2DLC TX114	hypo protein Rv3615c	Rv3615c	10777	100.00%	3	4	18	0.10%	17.50%	KAIDGLFT	95.00%	1.36	0.364	4	4	0	0	2	865.1484	96	103
2DLC TX114	ilvN (ilvN)	Rv3002c	18169	100.00%	6	6	9	0.05%	23.80%	EALLR	95.00%	1.42	0	1	0	0	0	2	601.7984	135	139
2DLC TX114	ilvN (ilvN)	Rv3002c	18169	100.00%	6	6	9	0.05%	23.80%	GKLEALLR	95.00%	2.18	0.289	0	1	0	0	2	900.2484	132	139
2DLC TX114	ilvN (ilvN)	Rv3002c	18169	100.00%	6	6	9	0.05%	23.80%	LINVIK	95.00%	2.06	0.202	0	1	0	0	2	699.9884	70	75
2DLC TX114	ilvN (ilvN)	Rv3002c	18169	100.00%	6	6	9	0.05%	23.80%	SQVIEAVNLFR	95.00%	4.38	0.551	0	3	0	0	2	1,276.42	102	112
2DLC TX114	ilvN (ilvN)	Rv3002c	18169	100.00%	6	6	9	0.05%	23.80%	VAALFSR	95.00%	1.96	0.264	0	2	0	0	2	763.9284	22	28
2DLC TX114	ilvN (ilvN)	Rv3002c	18169	100.00%	6	6	9	0.05%	23.80%	VLEPFGIR	85.30%	2	0.164	0	1	0	0	2	931.1584	140	147
2DLC TX114	sppA (sppA)	Rv0724	65918	100.00%	5	5	11	0.06%	9.15%	GLVDELGGLR	95.00%	3.19	0.355	0	4	0	0	2	1,029.24	533	542
2DLC TX114	sppA (sppA)	Rv0724	65918	100.00%	5	5	11	0.06%	9.15%	IVSYPGSSLWDMVRPR	85.10%	1.46	0.258	0	0	1	0	2	1,864.19	562	577
2DLC TX114	sppA (sppA)	Rv0724	65918	100.00%	5	5	11	0.06%	9.15%	LGVGSDAVR	95.00%	1.81	0.513	0	3	0	0	2	873.9184	459	467
2DLC TX114	sppA (sppA)	Rv0724	65918	100.00%	5	5	11	0.06%	9.15%	VLAGLDEDTEVR	95.00%	2.64	0.475	0	2	0	0	2	1,317.39	550	561

2DLC TX114	sppA (sppA)	Rv0724	65918	100.00%	5	5	11	0.06%	9.15%	VWTGADALDRGL	95.00%	1.43	0.292	1	0	0	0	2	1,274.45	523	534
2DLC TX114	hypo protein Rv3614c	Rv3614c	19784	100.00%	4	5	11	0.06%	31.50%	AQSAQYAFILDR	95.00%	3.05	0.491	0	3	0	0	2	1,383.64	116	127
2DLC TX114	hypo protein Rv3614c	Rv3614c	19784	100.00%	4	5	11	0.06%	31.50%	IDHVELSAR	95.00%	2.79	0.387	0	1	3	0	2	1,040.15	84	92
2DLC TX114	hypo protein Rv3614c	Rv3614c	19784	100.00%	4	5	11	0.06%	31.50%	KTVGETWGLPSPEEAAAAE AEVFATR	95.00%	2.68	0.355	0	0	1	0	2	2,718.86	144	169
2DLC TX114	hypo protein Rv3614c	Rv3614c	19784	100.00%	4	5	11	0.06%	31.50%	TVSVSTLMDGR	95.00%	2.44	0.535	0	3	0	0	2	1,182.27	73	83
2DLC TX114	hypo protein Rv3193c	Rv3193c	107455	100.00%	3	3	4	0.02%	4.44%	LLDPTIVSPAFTQFQQGK	95.00%	5.06	0.68	0	0	1	0	2	1,991.34	370	387
2DLC TX114	hypo protein Rv3193c	Rv3193c	107455	100.00%	3	3	4	0.02%	4.44%	VAALQEIQAAIGAAR	95.00%	3.09	0.482	0	1	0	0	2	1,482.72	949	963
2DLC TX114	hypo protein Rv3193c	Rv3193c	107455	100.00%	3	3	4	0.02%	4.44%	WGNLLTLPVAR	95.00%	3.04	0.451	0	2	0	0	2	1,240.54	833	843
2DLC TX114	rpsC (rpsC)	Rv0707	29970	100.00%	3	3	5	0.03%	9.85%	AGIADVEIER	95.00%	3.14	0.344	0	1	0	0	2	1,073.20	49	58
2DLC TX114	rpsC (rpsC)	Rv0707	29970	100.00%	3	3	5	0.03%	9.85%	PGIVIGR	95.00%	2.5	0.263	0	3	0	0	2	711.9084	72	78
2DLC TX114	rpsC (rpsC)	Rv0707	29970	100.00%	3	3	5	0.03%	9.85%	QVQLNILEVK	95.00%	3.49	0.433	0	1	0	0	2	1,184.36	97	106
2DLC TX114	rpsJ (rpsJ)	Rv0700	11381	99.80%	2	2	4	0.02%	24.80%	LIDIIDPTPK	95.00%	2.34	0.467	0	2	0	0	2	1,125.56	73	82
2DLC TX114	rpsJ (rpsJ)	Rv0700	11381	99.80%	2	2	4	0.02%	24.80%	TGASVVGPVPLPTEK	95.00%	2.32	0.402	0	2	0	0	2	1,452.55	32	46
2DLC TX114	hypo protein Rv2536	Rv2536	24608	99.80%	2	3	3	0.02%	10.00%	AADTDVFSAVR	95.00%	1.15	0.414	1	1	0	0	2	1,152.18	157	167
2DLC TX114	hypo protein Rv2536	Rv2536	24608	99.80%	2	3	3	0.02%	10.00%	TTESDTPTEVIR	95.00%	2.63	0.401	0	1	0	0	2	1,349.39	202	213
2DLC TX114	hypo protein Rv0036c	Rv0036c	27528	99.80%	2	2	3	0.02%	10.10%	LHEELLAVPDGR	95.00%	2.9	0.417	0	1	0	0	2	1,349.54	106	117
2DLC TX114	hypo protein Rv0036c	Rv0036c	27528	99.80%	2	2	3	0.02%	10.10%	WLTIAQAFAGPPGR	95.00%	2.72	0.383	0	2	0	0	2	1,485.80	242	255
2DLC TX114	hypo protein Rv0079	Rv0079	29457	100.00%	4	4	8	0.05%	17.90%	EHGLPFLFFTDQATGR	95.00%	3.87	0.549	0	0	2	0	2	1,837.12	231	246
2DLC TX114	hypo protein Rv0079	Rv0079	29457	100.00%	4	4	8	0.05%	17.90%	GQLLYSR	95.00%	1.88	0.383	0	2	0	0	2	837.0484	247	253
2DLC TX114	hypo protein Rv0079	Rv0079	29457	100.00%	4	4	8	0.05%	17.90%	LTTPAEALVTR	95.00%	1.57	0.376	0	3	0	0	2	1,172.37	127	137
2DLC TX114	hypo protein Rv0079	Rv0079	29457	100.00%	4	4	8	0.05%	17.90%	SRPTPVLTEAAAVDR	95.00%	3.48	0.508	0	0	1	0	2	1,583.72	214	228
2DLC TX114	hypo protein Rv2115c	Rv2115c	67385	99.80%	2	2	2	0.01%	4.27%	AIDTESNLGQYL	95.00%	2.33	0.389	0	1	0	0	2	1,324.53	598	609
2DLC TX114	hypo protein Rv2115c	Rv2115c	67385	99.80%	2	2	2	0.01%	4.27%	ENAVGSHAPTRSAR	95.00%	1.53	0.398	0	1	0	0	2	1,453.46	41	54
2DLC TX114	hypo protein Rv1996	Rv1996	33860	99.80%	2	2	6	0.03%	6.94%	GGFPGMHLGSVSR	95.00%	2.65	0.509	0	4	0	0	2	1,302.47	282	294
2DLC TX114	hypo protein Rv1996	Rv1996	33860	99.80%	2	2	6	0.03%	6.94%	LNWAPIEWR	95.00%	2.89	0.455	0	2	0	0	2	1,185.42	217	225
2DLC TX114	hypo protein Rv2509	Rv2509	27996	100.00%	4	4	8	0.05%	17.50%	AIVAPIVGAFYK	95.00%	2.67	0.557	0	2	0	0	2	1,249.57	252	263
2DLC TX114	hypo protein Rv2509	Rv2509	27996	100.00%	4	4	8	0.05%	17.50%	REDVLTELAAR	95.00%	3.46	0.421	0	1	0	0	2	1,273.45	42	52
2DLC TX114	hypo protein Rv2509	Rv2509	27996	100.00%	4	4	8	0.05%	17.50%	VSLNALER	95.00%	2.32	0.287	0	2	0	0	2	902.0484	221	228
2DLC TX114	hypo protein Rv2509	Rv2509	27996	100.00%	4	4	8	0.05%	17.50%	VTVDVRPADLADPQER	95.00%	3.71	0.512	0	0	3	0	2	1,781.83	59	74
2DLC TX114	rplI (rplI)	Rv0056	16152	100.00%	4	4	6	0.03%	28.90%	AAIEALGPIALPVK	95.00%	3.71	0.476	0	3	0	0	2	1,363.80	70	83
2DLC TX114	rplI (rplI)	Rv0056	16152	100.00%	4	4	6	0.03%	28.90%	GLAIVASR	95.00%	1.74	0.139	1	0	0	0	2	786.9784	34	41
2DLC TX114	rplI (rplI)	Rv0056	16152	100.00%	4	4	6	0.03%	28.90%	LFGSVTAADVVAAIKK	95.00%	2.94	0.286	0	1	0	0	2	1,590.88	91	106
2DLC TX114	rplI (rplI)	Rv0056	16152	100.00%	4	4	6	0.03%	28.90%	NFLLPR	84.40%	1.49	0.235	0	1	0	0	2	760.0384	28	33
2DLC TX114	rpsF (rpsF)	Rv0053	10917	99.80%	2	2	2	0.01%	26.00%	AAPATVSELDR	95.00%	2.96	0.54	0	1	0	0	2	1,130.23	67	77
2DLC TX114	rpsF (rpsF)	Rv0053	10917	99.80%	2	2	2	0.01%	26.00%	TVAPSLETFLNVVR	95.00%	2.64	0.383	0	1	0	0	2	1,546.71	18	31
2DLC TX114	(mmaA2) mmaA2	Rv0644c	32707	100.00%	3	3	7	0.04%	9.41%	TGFTLTR	95.00%	1.64	0.383	0	3	0	0	2	795.9684	218	224
2DLC TX114	(mmaA2) mmaA2	Rv0644c	32707	100.00%	3	3	7	0.04%	9.41%	VGYIDVNQFTLAK	95.00%	3.3	0.577	0	2	0	0	2	1,468.69	275	287
2DLC TX114	(mmaA2) mmaA2	Rv0644c	32707	100.00%	3	3	7	0.04%	9.41%	YLTGCAK	95.00%	0.874	0.264	2	0	0	0	2	755.9484	265	271
2DLC TX114	hypo protein Rv1957	Rv1957	20089	99.80%	2	2	3	0.02%	16.60%	EILSRPMPVSPGAQWPATR GTP	95.00%	2.11	0.236	0	0	1	0	2	2,349.71	160	181
2DLC TX114	hypo protein Rv1957	Rv1957	20089	99.80%	2	2	3	0.02%	16.60%	FALYPYIR	95.00%	2.22	0.343	0	2	0	0	2	1,043.42	135	142
2DLC TX114	hypo protein Rv3780	Rv3780	19466	100.00%	3	4	10	0.06%	11.20%	AQLQQMR	95.00%	1.9	0.394	0	3	0	0	2	874.9984	150	156
2DLC TX114	hypo protein Rv3780	Rv3780	19466	100.00%	3	4	10	0.06%	11.20%	IGTMIK	95.00%	1.19	0.153	2	1	0	0	2	662.9684	54	59
2DLC TX114	hypo protein Rv3780	Rv3780	19466	100.00%	3	4	10	0.06%	11.20%	QLLEEVR	94.90%	1.81	0.206	0	4	0	0	2	886.9884	60	66
2DLC TX114	hypo protein Rv2721c	Rv2721c	72322	100.00%	10	12	24	0.14%	20.30%	AAGGAAGPLGAK	95.00%	3.36	0.488	0	2	0	0	2	941.1384	232	243
2DLC TX114	hypo protein Rv2721c	Rv2721c	72322	100.00%	10	12	24	0.14%	20.30%	AGGDTSTLGVR	95.00%	3.26	0.485	0	2	0	0	2	1,034.10	52	62
2DLC TX114	hypo protein Rv2721c	Rv2721c	72322	100.00%	10	12	24	0.14%	20.30%	FTGGMISWNR	95.00%	3.13	0.469	0	2	0	0	2	1,169.40	370	379
2DLC TX114	hypo protein Rv2721c	Rv2721c	72322	100.00%	10	12	24	0.14%	20.30%	GALNAAWDK	95.00%	2.7	0.421	0	2	0	0	2	946.1084	154	162
2DLC TX114	hypo protein Rv2721c	Rv2721c	72322	100.00%	10	12	24	0.14%	20.30%	GAMVAAWDK	95.00%	2.3	0.421	0	1	0	0	2	949.0784	336	344
2DLC TX114	hypo protein Rv2721c	Rv2721c	72322	100.00%	10	12	24	0.14%	20.30%	KGDVYPIGDGFALDFAGGK	95.00%	5.24	0.699	0	1	2	0	2	1,928.32	63	81

hypo protein									KGGQYPIGGDGIAQDFVG										
2DLC TX114	Rv2721c	Rv2721c	72322	100.00%	10	12	24	0.14%	20.30%	GK	95.00%	4.25	0.598	0	2	1	0	2	1,965.24	244	263
hypo protein																					
2DLC TX114	Rv2721c	Rv2721c	72322	100.00%	10	12	24	0.14%	20.30%	LGAPVGDQTVDGDVVSQK	95.00%	3.11	0.567	0	0	1	0	2	1,785.73	352	369
hypo protein																					
2DLC TX114	Rv2721c	Rv2721c	72322	100.00%	10	12	24	0.14%	20.30%	MFFTPATGAK	95.00%	1.76	0.401	0	3	0	0	2	1,071.35	82	91
hypo protein									VSTFSAADNPVIFWTPEHG												
2DLC TX114	Rv2721c	Rv2721c	72322	100.00%	10	12	24	0.14%	20.30%	AFVVR	95.00%	3.41	0.613	0	0	5	0	2	2,648.79	130	153
2DLC TX114	ppiA (ppiA)	Rv0009	19222	100.00%	10	14	34	0.20%	53.80%	ASGGPSGPFYDGAVFHR	95.00%	1.83	0.297	0	0	1	0	2	1,722.85	57	73
2DLC TX114	ppiA (ppiA)	Rv0009	19222	100.00%	10	14	34	0.20%	53.80%	DYSTQNASGGPSGPFYDG AVFHR	95.00%	3.62	0.536	0	0	1	0	2	2,431.53	51	73
2DLC TX114	ppiA (ppiA)	Rv0009	19222	100.00%	10	14	34	0.20%	53.80%	FGNHAPK	95.00%	2.3	0.227	0	2	0	0	2	770.8684	31	37
2DLC TX114	ppiA (ppiA)	Rv0009	19222	100.00%	10	14	34	0.20%	53.80%	HTIFGEVIDAESQR	95.00%	3.17	0.648	0	3	5	0	2	1,602.70	141	154
2DLC TX114	ppiA (ppiA)	Rv0009	19222	100.00%	10	14	34	0.20%	53.80%	IALFGN	91.50%	1.25	0.0147	1	0	0	0	2	634.8484	28	33
2DLC TX114	ppiA (ppiA)	Rv0009	19222	100.00%	10	14	34	0.20%	53.80%	IALFGNHAPK	95.00%	2.73	0.415	0	8	1	0	2	1,068.35	28	37
2DLC TX114	ppiA (ppiA)	Rv0009	19222	100.00%	10	14	34	0.20%	53.80%	SPLATATATLHTNR	95.00%	3.1	0.47	0	0	2	0	2	1,454.66	10	23
2DLC TX114	ppiA (ppiA)	Rv0009	19222	100.00%	10	14	34	0.20%	53.80%	TVANFVGLAQGTK	95.00%	3.57	0.531	0	1	0	0	2	1,306.42	38	50
2DLC TX114	ppiA (ppiA)	Rv0009	19222	100.00%	10	14	34	0.20%	53.80%	VIQGFMIQGGDPTGTGR	95.00%	4.43	0.495	0	2	3	0	2	1,750.96	74	90
2DLC TX114	ppiA (ppiA)	Rv0009	19222	100.00%	10	14	34	0.20%	53.80%	VVEAISK	95.00%	1.74	0.272	2	2	0	0	2	745.8184	155	161
hypo protein																					
2DLC TX114	Rv0007	Rv0007	31024	99.80%	2	3	3	0.02%	9.54%	FISGASAPVTGPAAAVR	95.00%	3.55	0.488	0	1	1	0	2	1,572.76	85	101
hypo protein																					
2DLC TX114	Rv0007	Rv0007	31024	99.80%	2	3	3	0.02%	9.54%	IPDAGDPPPWQR	95.00%	2.05	0.219	0	1	0	0	2	1,349.53	37	48
2DLC TX114	gyrA (gyrA)	Rv0006	92259	99.80%	2	2	5	0.03%	2.15%	LLSLNVVR	95.00%	2.13	0.356	0	2	0	0	2	914.1384	711	718
2DLC TX114	gyrA (gyrA)	Rv0006	92259	99.80%	2	2	5	0.03%	2.15%	LTPLAMEMLR	95.00%	1.91	0.438	0	3	0	0	2	1,175.62	134	143
2DLC TX114	gyrB (gyrB)	Rv0005	78422	100.00%	3	3	6	0.03%	3.92%	GLTIINL	95.00%	1.26	0.105	2	0	0	0	2	630.8684	242	247
2DLC TX114	gyrB (gyrB)	Rv0005	78422	100.00%	3	3	6	0.03%	3.92%	SALTSVVNK	95.00%	2.25	0.376	0	2	0	0	2	918.9984	361	369
2DLC TX114	gyrB (gyrB)	Rv0005	78422	100.00%	3	3	6	0.03%	3.92%	TFHYPGGLVDFVK	95.00%	4.62	0.553	0	0	2	0	2	1,480.68	289	301
2DLC TX114	desA1	Rv0824c	38752	100.00%	4	4	5	0.03%	12.70%	QYLYVTR	95.00%	2	0.432	0	2	0	0	2	865.9584	116	122
2DLC TX114	desA1	Rv0824c	38752	100.00%	4	4	5	0.03%	12.70%	IFEREDFTGEGAK	95.00%	3.29	0.473	0	0	1	0	2	1,499.70	272	284
2DLC TX114	desA1	Rv0824c	38752	100.00%	4	4	5	0.03%	12.70%	IHLDEVVMPVLK	95.00%	3.92	0.559	0	1	0	0	2	1,393.65	257	268
2DLC TX114	desA1	Rv0824c	38752	100.00%	4	4	5	0.03%	12.70%	SVDPVELEKLR	95.00%	1.7	0.287	0	1	0	0	2	1,285.46	123	133
2DLC TX114	dnaN (dnaN)	Rv0002	42096	100.00%	4	4	5	0.03%	8.71%	AGIGGSDVR	95.00%	1.96	0.249	0	1	0	0	2	831.8984	217	225
2DLC TX114	dnaN (dnaN)	Rv0002	42096	100.00%	4	4	5	0.03%	8.71%	LVALVADR	95.00%	2.64	0.423	0	1	0	0	2	857.0184	286	293
2DLC TX114	dnaN (dnaN)	Rv0002	42096	100.00%	4	4	5	0.03%	8.71%	MEFADGSVR	95.00%	1.15	0.228	1	0	0	0	2	1,028.10	299	307
2DLC TX114	dnaN (dnaN)	Rv0002	42096	100.00%	4	4	5	0.03%	8.71%	VGLTDLTFR	95.00%	2.95	0.451	0	2	0	0	2	1,022.24	8	16
fadD31																					
2DLC TX114	(fadD31)	Rv1925	66311	100.00%	9	11	17	0.10%	18.20%	AFAPYGLPR	95.00%	2.52	0.335	0	1	0	0	2	992.2684	340	348
fadD31																					
2DLC TX114	(fadD31)	Rv1925	66311	100.00%	9	11	17	0.10%	18.20%	HGLSVADLSFLPAGAIPR	95.00%	6.49	0.719	0	3	1	0	2	1,822.23	580	597
fadD31																					
2DLC TX114	(fadD31)	Rv1925	66311	100.00%	9	11	17	0.10%	18.20%	IADLLTIDGR	95.00%	2.95	0.473	2	2	0	0	2	1,087.45	501	510
fadD31										LQSPLAEGSHADGSAIDDT											
2DLC TX114	(fadD31)	Rv1925	66311	100.00%	9	11	17	0.10%	18.20%	WLR	95.00%	2.49	0.508	0	0	2	0	2	2,340.62	462	483
fadD31																					
2DLC TX114	(fadD31)	Rv1925	66311	100.00%	9	11	17	0.10%	18.20%	LVIIAER	93.70%	1.82	0.192	0	1	0	0	2	814.0684	550	556
fadD31																					
2DLC TX114	(fadD31)	Rv1925	66311	100.00%	9	11	17	0.10%	18.20%	NAVEGFLNNVPR	95.00%	2.67	0.395	0	1	0	0	2	1,330.41	154	165
fadD31																					
2DLC TX114	(fadD31)	Rv1925	66311	100.00%	9	11	17	0.10%	18.20%	VVTAAPNFAYEWAAQR	95.00%	2.67	0.566	0	1	0	0	2	1,794.88	291	306
fadD31																					
2DLC TX114	(fadD31)	Rv1925	66311	100.00%	9	11	17	0.10%	18.20%	WIQALSEGSR	95.00%	2.56	0.48	0	2	0	0	2	1,147.31	278	287
fadD31																					
2DLC TX114	(fadD31)	Rv1925	66311	100.00%	9	11	17	0.10%	18.20%	YVGDLVAYR	95.00%	2.83	0.439	0	1	0	0	2	1,056.19	46	54
2DLC TX114	lpqH (lpqH)	Rv3763	15096	100.00%	5	7	18	0.10%	25.80%	ITGTATGVDMAN	95.00%	2.13	0.546	0	1	0	0	2	1,151.25	132	143
2DLC TX114	lpqH (lpqH)	Rv3763	15096	100.00%	5	7	18	0.10%	25.80%	ITGTATGVDMANPMSPVNK	95.00%	4.49	0.508	0	4	6	0	2	1,905.14	132	150
2DLC TX114	lpqH (lpqH)	Rv3763	15096	100.00%	5	7	18	0.10%	25.80%	PMSPVNK	85.60%	2	0.14	0	1	0	0	2	788.9078	144	150
2DLC TX114	lpqH (lpqH)	Rv3763	15096	100.00%	5	7	18	0.10%	25.80%	SVGLGNVNGVTL	95.00%	2.24	0.388	0	2	0	0	2	1,130.16	99	110
2DLC TX114	lpqH (lpqH)	Rv3763	15096	100.00%	5	7	18	0.10%	25.80%	SVGLGNVNGVTLGYTSGTG QGN	95.00%	5.41	0.652	0	3	1	0	2	2,053.04	99	120
2DLC TX114	fadA2	Rv0243	46058	100.00%	5	5	8	0.05%	11.60%	LDMVVGGAVLK	95.00%	3.5	0.468	0	2	0	0	2	1,118.31	59	69
2DLC TX114	fadA2	Rv0243	46058	100.00%	5	5	8	0.05%	11.60%	LGLDAALGSIDR	95.00%	4.32	0.44	0	2	0	0	2	1,201.56	365	376
2DLC TX114	fadA2	Rv0243	46058	100.00%	5	5	8	0.05%	11.60%	PVFGVK	95.00%	1.67	0.314	0	1	0	0	2	646.7284	251	256
2DLC TX114	fadA2	Rv0243	46058	100.00%	5	5	8	0.05%	11.60%	TGLSMGGHAAVTAK	89.90%	1.67	0.224	0	0	1	0	2	1,373.51	179	192
2DLC TX114	fadA2	Rv0243	46058	100.00%	5	5	8	0.05%	11.60%	VAVLGGNR	95.00%	1.99	0.295	0	2	0	0	2	785.8384	15	22
2DLC TX114	glcB (glcB)	Rv1837c	80386	100.00%	3	3	5	0.03%	4.72%	FALNAANAR	95.00%	2.68	0.378	0	1	0	0	2	948.1384	126	134
2DLC TX114	glcB (glcB)	Rv1837c	80386	100.00%	3	3	5	0.03%	4.72%	NYTAPGGGQFTLPGR	95.00%	2.85	0.593	0	1	0	0	2	1,536.74	319	333
2DLC TX114	glcB (glcB)	Rv1837c	80386	100.00%	3	3	5	0.03%	4.72%	VIEPIDMDAYR	95.00%	2.16	0.459	0	3	0	0	2	1,322.59	72	82
hypo protein																					
2DLC TX114	Rv1836c	Rv1836c	69659	100.00%	3	3	4	0.02%	7.39%	LSNVVAALENR	95.00%	3.73	0.564	0	1	0	0	2	1,186.30	507	517
hypo protein																					
2DLC TX114	Rv1836c	Rv1836c	69659	100.00%	3	3	4	0.02%	7.39%	PAALTAALGK	95.00%	2.67	0.564	0	2	0	0	2	913.2084	556	565
hypo protein									VSDVKPPSSPVTSFPALPST												
2DLC TX114	Rv1836c	Rv1836c	69659	100.00%	3	3	4	0.02%	7.39%	LSVGDDSMR	95.00%	4.4	0.591	0	0	1	0	2	2,975.29	445	473
2DLC TX114	rho (rho)	Rv1297	65084	100.00%	5	5	8	0.05%	11.80%	KNELIAAIEEIR	95.00%	4.42	0.549	0	1	0	0	2	1,399.78	62	73
2DLC TX114	rho (rho)	Rv1297	65084	100.00%	5	5	8	0.05%	11.80%	LVEQGKDVVVLLDSITR TSGYLPGPHDVYYVSMNMV	95.00%	3.84	0.581	0	1	0	0	2	1,885.13	428	444
2DLC TX114	rho (rho)	Rv1297	65084	100.00%	5	5	8	0.05%	11.80%	R	95.00%	1.75	0.299	0	0	1	0	2	2,156.33	241	259
2DLC TX114	rho (rho)	Rv1297	65084	100.00%	5	5	8	0.05%	11.80%	VFPAVDVNPSGTR	95.00%	2.08	0.32	0	2	0	0	2	1,359.38	529	541
2DLC TX114	rho (rho)	Rv1297	65084	100.00%	5	5	8	0.05%	11.80%	VIDLIMPIGK	95.00%	3.23	0.446	0	3	0	0	2	1,099.56	336	345
2DLC TX114	sigA (sigA)	Rv2703	57752	99.80%	2	2	4	0.02%	4.17%	SQVLRDYLD	95.00%	2.47	0.456	0	2	0	0	2	1,109.27	520	528
2DLC TX114	sigA (sigA)	Rv2703	57752	99.80%	2	2	4	0.02%	4.17%	TLDEIGQVYGVTR	95.00%	2.42	0.393	0	2	0	0	2	1,451.56	488	500
2DLC TX114	lpqN (lpqN)	Rv0583c	23665	100.00%	4	7	29	0.17%	11.40%	LLPESSR	95.00%	1.53	0.3	2	0	0	0	2	801.9984	95	101
2DLC TX114	lpqN (lpqN)	Rv0583c	23665	100.00%	4	7	29	0.17%	11.40%	LTGDIDPAK	95.00%	1.71	0.428	1	2	0	0	2	930.1684	128	136
2DLC TX114	lpqN (lpqN)	Rv0583c	23665	100.00%	4	7	29	0.17%	11.40%	LTGDIDPAKVLQFAPGELK	95.00%	4.94	0.502	0	2	8	0	2	2,013.49	128	146
2DLC TX114	lpqN (lpqN)	Rv0583c	23665	100.00%	4	7	29	0.17%	11.40%	VLQFAPGELK	95.00%	2.81	0.502	2	12	0	0	2	1,102.34	137	146
hypo protein																					
2DLC TX114	Rv1231c	Rv1231c	20396	99.40%	2	2	3	0.02%	10.60%	LAIGEVPTR	95.00%	2.5	0.347	0	2	0	0	2	956.1384	131	139
hypo protein																					
2DLC TX114	Rv1231c	Rv1231c	20396	99.40%	2	2	3	0.02%	10.60%	YLLVQTLLVL	87.70%	1.59	0.117	1	0	0	0	2	1,175.56	44	53
2DLC TX114	lprC (lprC)	Rv1275	18853	100.00%	4	4	11	0.06%	26.10%	AIQGVDSIVMR	95.00%	3.01	0.516	0	3	0	0	2	1,189.37	123	133
2DLC TX114	lprC (lprC)	Rv1275	18853	100.00%	4	4	11	0.06%	26.10%	FWFEQGSLSNER	95.00%	4.48	0.567	0	3	0	0	2	1,500.63	98	109
2DLC TX114	lprC (lprC)	Rv1275	18853	100.00%	4	4	11	0.06%	26.10%	LMELTLATNA	95.00%	2.06	0.396	0	1	0	0	2	1,093.38	171	180

2DLC TX114	lprC (lprC)	Rv1275	18853	100.00%	4	4	11	0.06%	26.10%	WQAANPAGLIDITR	95.00%	4.21	0.647	0	4	0	0	2	1,526.83	84	97
2DLC TX114	hypo protein Rv1220c	Rv1220c	22090	99.80%	2	2	3	0.02%	6.98%	AQEVLTR	95.00%	2.09	0.305	0	2	0	0	2	816.8684	120	126
2DLC TX114	hypo protein Rv1220c	Rv1220c	22090	99.80%	2	2	3	0.02%	6.98%	DAEVIAVR	95.00%	2.35	0.144	0	1	0	0	2	872.9184	179	186
2DLC TX114	hypo protein Rv1265	Rv1265	25211	100.00%	4	4	12	0.07%	21.20%	ILFTHIAVPVMTR	95.00%	2.11	0.363	0	0	3	0	2	1,514.86	152	164
2DLC TX114	hypo protein Rv1265	Rv1265	25211	100.00%	4	4	12	0.07%	21.20%	LVGEHTEEWLAK	95.00%	2.79	0.484	0	1	0	0	2	1,412.52	195	206
2DLC TX114	hypo protein Rv1265	Rv1265	25211	100.00%	4	4	12	0.07%	21.20%	QVLDTLVDAGVAR	95.00%	3.69	0.572	0	6	0	0	2	1,357.44	171	183
2DLC TX114	hypo protein Rv1265	Rv1265	25211	100.00%	4	4	12	0.07%	21.20%	TAMSAVDDLR	95.00%	2.71	0.498	0	2	0	0	2	1,079.23	209	218
2DLC TX114	hypo protein Rv0580c	Rv0580c	18017	100.00%	5	8	23	0.13%	25.20%	AVVSELFLR	95.00%	1.62	0.164	1	4	0	0	2	1,034.22	106	114
2DLC TX114	hypo protein Rv0580c	Rv0580c	18017	100.00%	5	8	23	0.13%	25.20%	IPTLEEFAEAVDR	95.00%	3.58	0.66	0	5	1	0	2	1,490.67	137	149
2DLC TX114	hypo protein Rv0580c	Rv0580c	18017	100.00%	5	8	23	0.13%	25.20%	MLGLSFR	95.00%	1.9	0.379	2	5	0	0	2	824.1484	127	133
2DLC TX114	hypo protein Rv0580c	Rv0580c	18017	100.00%	5	8	23	0.13%	25.20%	RIPTLEEFAEAVDR	95.00%	3.33	0.497	0	4	0	0	2	1,646.87	136	149
2DLC TX114	hypo protein Rv0580c	Rv0580c	18017	100.00%	5	8	23	0.13%	25.20%	TQLMVVSFTGR	95.00%	2.73	0.519	0	1	0	0	2	1,239.39	39	49
2DLC TX114	lpqZ (lpqZ)	Rv1244	29552	100.00%	4	4	5	0.03%	26.20%	AENIVPLYR	95.00%	2.55	0.337	0	1	0	0	2	1,075.29	226	234
2DLC TX114	lpqZ (lpqZ)	Rv1244	29552	100.00%	4	4	5	0.03%	26.20%	AIVSALPEGIAAGDYTTAAE DKPALVVTQSTAK	95.00%	2.4	0.573	0	0	2	0	2	3,260.67	108	140
2DLC TX114	lpqZ (lpqZ)	Rv1244	29552	100.00%	4	4	5	0.03%	26.20%	NDATMFAALR	95.00%	2.58	0.339	0	1	0	0	2	1,110.36	184	193
2DLC TX114	lpqZ (lpqZ)	Rv1244	29552	100.00%	4	4	5	0.03%	26.20%	QVAAGADPAAVAAGWLAE HPLGR	95.00%	1.79	0.299	0	0	1	0	2	2,229.39	264	286
2DLC TX114	mdh (mdh)	Rv1240	34273	100.00%	5	6	8	0.05%	20.40%	ALNAVAADDVR	95.00%	3.25	0.542	0	2	0	0	2	1,115.18	116	126
2DLC TX114	mdh (mdh)	Rv1240	34273	100.00%	5	6	8	0.05%	20.40%	ASGSLLGPDRPIELR	95.00%	2.63	0.475	0	1	1	0	2	1,582.00	26	40
2DLC TX114	mdh (mdh)	Rv1240	34273	100.00%	5	6	8	0.05%	20.40%	GAAIIDAR	95.00%	2.35	0.244	0	1	0	0	2	787.0084	228	235
2DLC TX114	mdh (mdh)	Rv1240	34273	100.00%	5	6	8	0.05%	20.40%	GASSAASAASATIDAAR	95.00%	2.56	0.391	0	1	0	0	2	1,478.63	236	252
2DLC TX114	mdh (mdh)	Rv1240	34273	100.00%	5	6	8	0.05%	20.40%	SDLLEANGAIFTAQGK	95.00%	4.34	0.65	0	2	0	0	2	1,635.94	100	115
2DLC TX114	htrA (htrA)	Rv1223	56527	100.00%	15	22	60	0.35%	28.60%	GGILENDVIVK	95.00%	1.76	0.204	1	1	0	0	2	1,157.36	492	502
2DLC TX114	htrA (htrA)	Rv1223	56527	100.00%	15	22	60	0.35%	28.60%	IVHPTLGISTR	95.00%	3.41	0.411	0	2	6	0	2	1,194.44	458	468
2DLC TX114	htrA (htrA)	Rv1223	56527	100.00%	15	22	60	0.35%	28.60%	KTAEVVDAFTTSK	95.00%	3.07	0.483	0	2	0	0	2	1,397.50	222	234
2DLC TX114	htrA (htrA)	Rv1223	56527	100.00%	15	22	60	0.35%	28.60%	LVANSLIK	95.00%	2.29	0.191	0	1	0	0	2	858.1484	447	454
2DLC TX114	htrA (htrA)	Rv1223	56527	100.00%	15	22	60	0.35%	28.60%	PVFRPPVDPASR	95.00%	2.71	0.433	0	0	4	0	2	1,338.50	66	77
2DLC TX114	htrA (htrA)	Rv1223	56527	100.00%	15	22	60	0.35%	28.60%	QLAIGQDAPIEVVR	95.00%	3.81	0.633	0	4	1	0	2	1,509.69	520	533
2DLC TX114	htrA (htrA)	Rv1223	56527	100.00%	15	22	60	0.35%	28.60%	SLSDSASGLGFAIPVNEMK	95.00%	4.72	0.5	0	5	2	0	2	1,924.29	428	446
2DLC TX114	htrA (htrA)	Rv1223	56527	100.00%	15	22	60	0.35%	28.60%	SVSNAIASGAQVANVK	95.00%	4.32	0.518	0	6	4	0	2	1,516.54	469	484
2DLC TX114	htrA (htrA)	Rv1223	56527	100.00%	15	22	60	0.35%	28.60%	TAEVVDAFTTSK	95.00%	3.64	0.588	0	3	0	0	2	1,269.30	223	234
2DLC TX114	htrA (htrA)	Rv1223	56527	100.00%	15	22	60	0.35%	28.60%	TDLAVLK	95.00%	1.96	0.348	0	3	0	0	2	759.9684	326	332
2DLC TX114	htrA (htrA)	Rv1223	56527	100.00%	15	22	60	0.35%	28.60%	TTVVFNDGK	95.00%	1.66	0.256	0	2	0	0	2	981.0084	305	313
2DLC TX114	htrA (htrA)	Rv1223	56527	100.00%	15	22	60	0.35%	28.60%	TTVVFNDGKEVPANLVGR	95.00%	2.49	0.268	0	0	1	0	2	1,917.00	305	322
2DLC TX114	htrA (htrA)	Rv1223	56527	100.00%	15	22	60	0.35%	28.60%	VDNVDNLTVAR	95.00%	3.13	0.342	0	2	0	0	2	1,216.21	333	343
2DLC TX114	htrA (htrA)	Rv1223	56527	100.00%	15	22	60	0.35%	28.60%	VGDEVLAVGAPLGLR	95.00%	3.37	0.581	0	4	2	0	2	1,466.66	352	366
2DLC TX114	htrA (htrA)	Rv1223	56527	100.00%	15	22	60	0.35%	28.60%	VRVGDEVLAVGAPLGLR	95.00%	4.78	0.596	0	2	2	0	2	1,721.93	350	366
2DLC TX114	nusA (nusA)	Rv2841c	37623	99.80%	2	2	4	0.02%	5.76%	GISVNELLETIK	95.00%	2.6	0.452	0	3	0	0	2	1,316.62	16	27
2DLC TX114	nusA (nusA)	Rv2841c	37623	99.80%	2	2	4	0.02%	5.76%	SALLTAYR	95.00%	1.89	0.212	0	1	0	0	2	895.1584	28	35
2DLC TX114	hypo protein Rv1209	Rv1209	13071	100.00%	4	4	8	0.05%	27.00%	ATTATTLPAFGVTR	95.00%	1.91	0.408	0	1	0	0	2	1,407.60	38	51
2DLC TX114	hypo protein Rv1209	Rv1209	13071	100.00%	4	4	8	0.05%	27.00%	ELEALR	95.00%	1.13	0.132	1	0	0	0	2	730.8984	81	86
2DLC TX114	hypo protein Rv1209	Rv1209	13071	100.00%	4	4	8	0.05%	27.00%	GYKTSEVDWVLER	95.00%	2.89	0.408	0	0	2	0	2	1,582.69	65	77
2DLC TX114	hypo protein Rv1209	Rv1209	13071	100.00%	4	4	8	0.05%	27.00%	TSEVDWVLER	95.00%	3.49	0.524	0	4	0	0	2	1,234.24	68	77
2DLC TX114	hypo protein Rv3671c	Rv3671c	40703	100.00%	3	3	4	0.02%	14.10%	LSGPDIYGDPEPVTR	95.00%	2.5	0.458	0	1	0	0	2	1,616.83	315	329
2DLC TX114	hypo protein Rv3671c	Rv3671c	40703	100.00%	3	3	4	0.02%	14.10%	TPVIPVASPDPALVNNPVVA ATEPSVVK	95.00%	4	0.574	0	0	1	0	2	2,782.91	179	206
2DLC TX114	hypo protein Rv3671c	Rv3671c	40703	100.00%	3	3	4	0.02%	14.10%	VLEGTGFVISPDR	95.00%	1.99	0.365	0	2	0	0	2	1,390.56	217	229
2DLC TX114	hypo protein Rv0686	Rv0686	28526	99.80%	2	2	4	0.02%	6.04%	INDQPLVK	95.00%	2.21	0.344	0	1	0	0	2	927.1084	92	99
2DLC TX114	hypo protein Rv0686	Rv0686	28526	99.80%	2	2	4	0.02%	6.04%	LAELESLR	95.00%	2.27	0.368	0	3	0	0	2	931.1784	237	244
2DLC TX114	tuf (tuf)	Rv0685	43543	100.00%	13	19	61	0.35%	37.10%	AFDQIDNAPEER	95.00%	2.29	0.396	0	1	1	0	2	1,405.50	48	59
2DLC TX114	tuf (tuf)	Rv0685	43543	100.00%	13	19	61	0.35%	37.10%	EHVLLAR	95.00%	2.03	0.287	0	3	0	0	2	837.9684	120	126
2DLC TX114	tuf (tuf)	Rv0685	43543	100.00%	13	19	61	0.35%	37.10%	ELLAAQEFDEDAPVVR	95.00%	2.42	0.444	0	2	0	0	2	1,802.92	158	173
2DLC TX114	tuf (tuf)	Rv0685	43543	100.00%	13	19	61	0.35%	37.10%	ETDKPFLMPVEDVFTITGR	95.00%	2.32	0.342	0	2	0	0	2	2,196.55	207	225
2DLC TX114	tuf (tuf)	Rv0685	43543	100.00%	13	19	61	0.35%	37.10%	GTVTYGR	95.00%	1.17	0.254	2	0	0	0	2	689.6584	226	232
2DLC TX114	tuf (tuf)	Rv0685	43543	100.00%	13	19	61	0.35%	37.10%	GTVVTGRVER	95.00%	1.81	0.295	0	1	0	0	2	1,074.03	226	235
2DLC TX114	tuf (tuf)	Rv0685	43543	100.00%	13	19	61	0.35%	37.10%	GVINVNEEVEIVGIRPSTTK	95.00%	1.97	0.269	0	0	2	0	2	2,155.30	236	255
2DLC TX114	tuf (tuf)	Rv0685	43543	100.00%	13	19	61	0.35%	37.10%	KLLDQGQAGDNVGL	95.00%	3.73	0.53	0	3	0	0	2	1,428.62	266	279
2DLC TX114	tuf (tuf)	Rv0685	43543	100.00%	13	19	61	0.35%	37.10%	LIQPVAMDEGLR	95.00%	4.38	0.547	3	15	1	0	2	1,358.64	365	376
2DLC TX114	tuf (tuf)	Rv0685	43543	100.00%	13	19	61	0.35%	37.10%	PQFYFR	95.00%	1.58	0.254	2	10	0	0	2	858.0384	331	336
2DLC TX114	tuf (tuf)	Rv0685	43543	100.00%	13	19	61	0.35%	37.10%	TTLTAAITK	92.70%	1.47	0.286	0	1	0	0	2	920.1784	26	34
2DLC TX114	tuf (tuf)	Rv0685	43543	100.00%	13	19	61	0.35%	37.10%	TTVTGVEMFR	95.00%	3.06	0.468	2	8	0	0	2	1,141.21	256	265
2DLC TX114	tuf (tuf)	Rv0685	43543	100.00%	13	19	61	0.35%	37.10%	VSALKALEGDAK	95.00%	3.02	0.474	0	1	1	0	2	1,202.46	174	185
2DLC TX114	fusA (fusA)	Rv0684	77153	100.00%	7	7	14	0.08%	13.10%	AHVPLSEMFGYVGDLR	95.00%	5.25	0.65	0	3	0	0	2	1,792.04	652	667
2DLC TX114	fusA (fusA)	Rv0684	77153	100.00%	7	7	14	0.08%	13.10%	ATTDEPFAALAFK	95.00%	3.4	0.611	0	3	0	0	2	1,382.66	308	320
2DLC TX114	fusA (fusA)	Rv0684	77153	100.00%	7	7	14	0.08%	13.10%	IAGSQVLK	95.00%	2.07	0.338	0	2	0	0	2	815.9984	591	598
2DLC TX114	fusA (fusA)	Rv0684	77153	100.00%	7	7	14	0.08%	13.10%	IATHPFFGK LGETYDTVEIPADLAEQAEE	95.00%	1.85	0.293	0	0	3	0	2	1,018.27	321	329
2DLC TX114	fusA (fusA)	Rv0684	77153	100.00%	7	7	14	0.08%	13.10%	YRTK	95.00%	2.02	0.307	0	0	1	0	2	2,743.00	195	218
2DLC TX114	fusA (fusA)	Rv0684	77153	100.00%	7	7	14	0.08%	13.10%	LTYIR	90.80%	1.62	0.239	0	1	0	0	2	665.9184	330	334
2DLC TX114	fusA (fusA)	Rv0684	77153	100.00%	7	7	14	0.08%	13.10%	VYSGTVESGSQVINATK	95.00%	4.32	0.59	0	0	1	0	2	1,740.75	335	351
2DLC TX114	rpsG (rpsG)	Rv0683	17583	99.90%	2	2	3	0.02%	12.20%	ALDNVKPALEVR	95.00%	2.14	0.345	0	0	2	0	2	1,325.54	65	76
2DLC TX114	rpsG (rpsG)	Rv0683	17583	99.90%	2	2	3	0.02%	12.20%	WLVGYSR	95.00%	1.69	0.295	0	1	0	0	2	881.0184	103	109
2DLC TX114	rpoC (rpoC)	Rv0668	146721	100.00%	4	4	6	0.03%	4.56%	ELDRLEDIWSTFTK	95.00%	2.48	0.172	0	0	1	0	2	1,754.10	215	228

2DLC TX114	rpoC (rpoC)	Rv0668	146721	100.00%	4	4	6	0.03%	4.56%	IGLATAEDIR	95.00%	2.84	0.377	0	3	0	0	2	1,059.33	12	21
2DLC TX114	rpoC (rpoC)	Rv0668	146721	100.00%	4	4	6	0.03%	4.56%	LGIQAFEPMLVEGKAIQLH	95.00%	1.79	0.23	0	0	1	0	2	2,111.57	507	525
2DLC TX114	rpoC (rpoC)	Rv0668	146721	100.00%	4	4	6	0.03%	4.56%	TFHQGGVGEDITGGLPR	95.00%	4.18	0.33	0	0	1	0	2	1,741.86	1014	1030
2DLC TX114	rpoB (rpoB)	Rv0667	129201	100.00%	8	8	16	0.09%	6.91%	ALMGANMQR	95.00%	2.26	0.449	0	1	0	0	2	992.2284	599	607
2DLC TX114	rpoB (rpoB)	Rv0667	129201	100.00%	8	8	16	0.09%	6.91%	FGFSEIMR	95.00%	2.88	0.526	0	1	0	0	2	987.2484	246	253
2DLC TX114	rpoB (rpoB)	Rv0667	129201	100.00%	8	8	16	0.09%	6.91%	LSALGPGGLSR	95.00%	2.68	0.389	0	2	0	0	2	1,028.33	449	459
2DLC TX114	rpoB (rpoB)	Rv0667	129201	100.00%	8	8	16	0.09%	6.91%	RQPVTVLLK	95.00%	2.62	0.412	0	1	0	0	2	1,054.28	225	233
2DLC TX114	rpoB (rpoB)	Rv0667	129201	100.00%	8	8	16	0.09%	6.91%	SEAPLVGTGMELR	95.00%	2.31	0.539	0	2	0	0	2	1,360.57	615	627
2DLC TX114	rpoB (rpoB)	Rv0667	129201	100.00%	8	8	16	0.09%	6.91%	TVGELIQNQIR	95.00%	2.62	0.318	0	3	0	0	2	1,271.44	374	384
2DLC TX114	rpoB (rpoB)	Rv0667	129201	100.00%	8	8	16	0.09%	6.91%	VNPFGFIETPYR	95.00%	3.1	0.6	0	2	0	0	2	1,440.68	500	511
2DLC TX114	rpoB (rpoB)	Rv0667	129201	100.00%	8	8	16	0.09%	6.91%	VVVSQLVR	95.00%	2.75	0.387	0	4	0	0	2	899.8784	168	175
2DLC TX114	hypo protein Rv0093c	Rv0093c	29580	99.90%	2	2	2	0.01%	13.10%	DGERAPVPSARVDEHLGEC SACR	95.00%	2.12	0.241	0	0	1	0	2	2,512.55	61	83
2DLC TX114	hypo protein Rv0093c	Rv0093c	29580	99.90%	2	2	2	0.01%	13.10%	HCATIDGMDCEVAR	95.00%	1.78	0.29	0	0	1	0	2	1,578.65	40	53
2DLC TX114	rplL (rplL)	Rv0652	13423	100.00%	10	15	52	0.30%	43.10%	AKLEAAGATVTVK	95.00%	3.75	0.53	1	5	2	0	2	1,259.43	118	130
2DLC TX114	rplL (rplL)	Rv0652	13423	100.00%	10	15	52	0.30%	43.10%	DLVDGAPKPLL	95.00%	2.65	0.339	0	1	0	0	2	1,138.46	95	105
2DLC TX114	rplL (rplL)	Rv0652	13423	100.00%	10	15	52	0.30%	43.10%	DLVDGAPKPLLEK	95.00%	1.56	0.377	3	7	0	0	2	1,395.76	95	107
2DLC TX114	rplL (rplL)	Rv0652	13423	100.00%	10	15	52	0.30%	43.10%	EAADEAKAKLEAAGATVTV K	95.00%	3.23	0.466	0	0	1	0	2	1,974.17	111	130
2DLC TX114	rplL (rplL)	Rv0652	13423	100.00%	10	15	52	0.30%	43.10%	EIVSGL	95.00%	1.32	0.301	1	0	0	0	2	617.7184	83	88
2DLC TX114	rplL (rplL)	Rv0652	13423	100.00%	10	15	52	0.30%	43.10%	EIVSGLGLK	95.00%	1.86	0.38	7	4	0	0	2	916.1684	83	91
2DLC TX114	rplL (rplL)	Rv0652	13423	100.00%	10	15	52	0.30%	43.10%	EIVSGLGLKEAK	95.00%	3.26	0.473	0	6	0	0	2	1,244.55	83	94
2DLC TX114	rplL (rplL)	Rv0652	13423	100.00%	10	15	52	0.30%	43.10%	LEAAGATVTVK	95.00%	3.02	0.563	0	8	0	0	2	1,060.15	120	130
2DLC TX114	rplL (rplL)	Rv0652	13423	100.00%	10	15	52	0.30%	43.10%	LSTDELLDAFK	95.00%	3.32	0.598	2	2	0	0	2	1,252.58	4	14
2DLC TX114	rplL (rplL)	Rv0652	13423	100.00%	10	15	52	0.30%	43.10%	VDGAPKPLLEK	95.00%	2.14	0.448	0	2	0	0	2	1,167.46	97	107
2DLC TX114	rplJ (rplJ)	Rv0651	18461	100.00%	3	3	8	0.05%	23.00%	AAGLFNAPASQLAR	95.00%	2.9	0.385	0	1	0	0	2	1,387.67	147	160
2DLC TX114	rplJ (rplJ)	Rv0651	18461	100.00%	3	3	8	0.05%	23.00%	GGYMDGHPLTVAEVER	95.00%	4.04	0.393	0	0	2	0	2	1,731.81	107	122
2DLC TX114	rplJ (rplJ)	Rv0651	18461	100.00%	3	3	8	0.05%	23.00%	GLTVANLAELR	95.00%	2.87	0.411	0	5	0	0	2	1,157.40	30	40
2DLC TX114	lppZ (lppZ)	Rv3006	38734	100.00%	4	6	12	0.07%	18.50%	LAPSTGAVTGEPDVVR	95.00%	2.02	0.451	0	3	0	0	2	1,569.61	303	318
2DLC TX114	lppZ (lppZ)	Rv3006	38734	100.00%	4	6	12	0.07%	18.50%	LMYAYISTPTDNR	95.00%	2.35	0.467	0	1	0	0	2	1,545.90	138	150
2DLC TX114	lppZ (lppZ)	Rv3006	38734	100.00%	4	6	12	0.07%	18.50%	TAGDAEKLDDVVFPLFPQG GGFPR	95.00%	3.49	0.329	0	1	1	0	2	2,534.88	344	367
2DLC TX114	lppZ (lppZ)	Rv3006	38734	100.00%	4	6	12	0.07%	18.50%	VADGDIPKDILTGIPK	95.00%	3.22	0.469	0	2	4	0	2	1,653.11	154	169
2DLC TX114	hypo protein Rv2969c	Rv2969c	26811	100.00%	3	3	5	0.03%	17.30%	IKEIVGDVPGIDSAAATATS	95.00%	2.69	0.531	0	3	0	0	2	1,916.16	236	255
2DLC TX114	hypo protein Rv2969c	Rv2969c	26811	100.00%	3	3	5	0.03%	17.30%	LVTQPGTSNPK	95.00%	1.88	0.345	0	1	0	0	2	1,142.26	68	78
2DLC TX114	hypo protein Rv2969c	Rv2969c	26811	100.00%	3	3	5	0.03%	17.30%	YIEKVDGLAAAVN	95.00%	1.81	0.391	0	1	0	0	2	1,363.55	198	210
2DLC TX114	rplA (rplA)	Rv0641	24739	99.80%	2	2	3	0.02%	11.10%	LLAENYGAAIDEVLR	95.00%	3.81	0.483	0	2	0	0	2	1,647.98	184	198
2DLC TX114	rplA (rplA)	Rv0641	24739	99.80%	2	2	3	0.02%	11.10%	TNVTPLQAAK	95.00%	1.72	0.438	0	1	0	0	2	1,220.50	18	28
2DLC TX114	nusG (nusG)	Rv0639	25396	100.00%	7	9	25	0.14%	29.00%	ETPVELTFGQVSKI	95.00%	1.98	0.286	0	1	0	0	2	1,548.71	225	238
2DLC TX114	nusG (nusG)	Rv0639	25396	100.00%	7	9	25	0.14%	29.00%	KVLPGYILVR	95.00%	2.27	0.312	0	2	0	0	2	1,158.53	103	112
2DLC TX114	nusG (nusG)	Rv0639	25396	100.00%	7	9	25	0.14%	29.00%	MDLTDDSWAAVR	95.00%	2.43	0.482	0	2	0	0	2	1,380.53	113	124
2DLC TX114	nusG (nusG)	Rv0639	25396	100.00%	7	9	25	0.14%	29.00%	NTPGVTGFVGATSR	95.00%	2.23	0.524	0	2	0	0	2	1,364.39	125	138
2DLC TX114	nusG (nusG)	Rv0639	25396	100.00%	7	9	25	0.14%	29.00%	NTPGVTGFVGATSRPSALA LDDVVK	95.00%	2.1	0.522	0	1	2	0	2	2,473.69	125	149
2DLC TX114	nusG (nusG)	Rv0639	25396	100.00%	7	9	25	0.14%	29.00%	PSALALDDVVK	95.00%	4.09	0.449	0	2	0	0	2	1,128.32	139	149
2DLC TX114	nusG (nusG)	Rv0639	25396	100.00%	7	9	25	0.14%	29.00%	VLVSIFGR	95.00%	1.56	0.202	2	11	0	0	2	891.0884	217	224
2DLC TX114	secE (secE)	Rv0638	16906	99.80%	2	3	3	0.02%	6.21%	TALVTK	95.00%	1.11	0	1	0	0	0	2	632.7684	28	33
2DLC TX114	secE (secE)	Rv0638	16906	99.80%	2	3	3	0.02%	6.21%	TALVTKPVVR	95.00%	2.52	0.369	0	1	1	0	2	1,084.23	28	37
2DLC TX114	hypo protein Rv0636	Rv0636	14916	100.00%	3	3	4	0.02%	14.80%	FTAVVPVPNDGK	95.00%	1.85	0.312	0	2	0	0	2	1,244.30	87	98
2DLC TX114	hypo protein Rv0636	Rv0636	14916	100.00%	3	3	4	0.02%	14.80%	FTAVVPVPNDGKGAELVFN GR	95.00%	2.66	0.518	0	0	1	0	2	2,188.35	87	107
2DLC TX114	hypo protein Rv0636	Rv0636	14916	100.00%	3	3	4	0.02%	14.80%	GAELVFNGR	95.00%	2.29	0.334	0	1	0	0	2	963.0684	99	107
2DLC TX114	gpsI (gpsI)	Rv2783c	79718	100.00%	3	4	6	0.03%	4.52%	DGLVHISK	95.00%	2.65	0.359	0	2	0	0	2	869.0184	690	697
2DLC TX114	gpsI (gpsI)	Rv2783c	79718	100.00%	3	4	6	0.03%	4.52%	IAKVEDVVNVGDK	95.00%	2.79	0.552	0	1	2	0	2	1,386.43	704	716
2DLC TX114	gpsI (gpsI)	Rv2783c	79718	100.00%	3	4	6	0.03%	4.52%	VPVDKIGEVIGPK	95.00%	1.73	0.382	0	0	1	0	2	1,351.57	602	614
2DLC TX114	htpG (htpG)	Rv2299c	72945	100.00%	4	4	6	0.03%	5.87%	ELISNASDALDK	95.00%	2.97	0.56	0	2	0	0	2	1,276.54	35	46
2DLC TX114	htpG (htpG)	Rv2299c	72945	100.00%	4	4	6	0.03%	5.87%	FAELLAER	95.00%	2.41	0.26	0	1	0	0	2	949.1784	635	642
2DLC TX114	htpG (htpG)	Rv2299c	72945	100.00%	4	4	6	0.03%	5.87%	SPHLEAFK	84.90%	2.12	0.138	0	0	1	0	2	929.0984	459	466
2DLC TX114	htpG (htpG)	Rv2299c	72945	100.00%	4	4	6	0.03%	5.87%	YSDFIAWPIR	95.00%	2.86	0.454	0	2	0	0	2	1,268.60	206	215
2DLC TX114	ahpC (ahpC)	Rv2428	21549	100.00%	4	4	4	0.02%	19.50%	KGDPTLDAGELLK	95.00%	3.38	0.452	0	1	0	0	2	1,357.72	180	192
2DLC TX114	ahpC (ahpC)	Rv2428	21549	100.00%	4	4	4	0.02%	19.50%	NVDEVLR	95.00%	2.03	0.268	0	1	0	0	2	844.8584	157	163
2DLC TX114	ahpC (ahpC)	Rv2428	21549	100.00%	4	4	4	0.02%	19.50%	TLPFPMLSDIK	95.00%	1.8	0.338	0	1	0	0	2	1,262.74	105	115
2DLC TX114	ahpC (ahpC)	Rv2428	21549	100.00%	4	4	4	0.02%	19.50%	VVFFWPK	95.00%	2.06	0.439	0	1	0	0	2	923.0784	49	55
2DLC TX114	hypo protein Rv1703c	Rv1703c	21520	100.00%	3	3	5	0.03%	23.00%	AYLPDLQSILDR	95.00%	3.35	0.596	0	3	0	0	2	1,404.80	125	136
2DLC TX114	hypo protein Rv1703c	Rv1703c	21520	100.00%	3	3	5	0.03%	23.00%	FAYEKSMLINVGDEKGTL	95.00%	2.42	0.259	0	0	1	0	2	2,016.45	8	25
2DLC TX114	hypo protein Rv1703c	Rv1703c	21520	100.00%	3	3	5	0.03%	23.00%	VYSVELAEANASNAR	95.00%	4.31	0.436	0	0	1	0	2	1,594.64	61	75
2DLC TX114	oppD (oppD)	Rv1281c	65275	99.90%	2	2	2	0.01%	3.27%	RRAVELLDL	95.00%	2.35	0.29	1	0	0	0	2	1,085.37	130	138
2DLC TX114	oppD (oppD)	Rv1281c	65275	99.90%	2	2	2	0.01%	3.27%	VAELLDIVGLR	95.00%	3.52	0.425	0	1	0	0	2	1,198.49	479	489
2DLC TX114	mmpL3 (mmpL3)	Rv0206c	100857	99.80%	2	2	4	0.02%	2.75%	IGLGEIHLPDER	95.00%	2.9	0.465	0	0	2	0	2	1,349.64	744	755
2DLC TX114	mmpL3 (mmpL3)	Rv0206c	100857	99.80%	2	2	4	0.02%	2.75%	PARPPVTAGLVAAR	95.00%	3.85	0.369	0	0	2	0	2	1,376.59	764	777
2DLC TX114	hypo protein Rv1875	Rv1875	15961	100.00%	4	4	4	0.02%	31.30%	AQLVGPDDPRPWLVDGER	95.00%	2.29	0.299	0	0	1	0	2	2,021.20	85	102
2DLC TX114	hypo protein Rv1875	Rv1875	15961	100.00%	4	4	4	0.02%	31.30%	ARPQLAVTFR	95.00%	1.85	0.27	0	1	0	0	2	1,159.37	64	73
2DLC TX114	hypo protein Rv1875	Rv1875	15961	100.00%	4	4	4	0.02%	31.30%	AVVLITPTR	95.00%	2.52	0.435	0	1	0	0	2	970.1584	134	142
2DLC TX114	hypo protein Rv1875	Rv1875	15961	100.00%	4	4	4	0.02%	31.30%	GLAVVSTVR	95.00%	2.59	0.47	0	1	0	0	2	901.9384	16	24
2DLC TX114	hypo protein Rv0020c	Rv0020c	56862	100.00%	7	7	11	0.06%	18.20%	GGYPPETGGYPPQPGYPRP R	95.00%	1.53	0.309	0	0	1	0	2	2,144.41	200	219
2DLC TX114	hypo protein Rv0020c	Rv0020c	56862	100.00%	7	7	11	0.06%	18.20%	GQDAQFR	95.00%	2.18	0.295	0	2	0	0	2	821.8484	460	466

2DLC TX114	hypo protein Rv0020c	Rv0020c	56862	100.00%	7	7	11	0.06%	18.20%	HPDQGDYPEQIGYPDQGG YPEQR	95.00%	4.25	0.55	0	0	1	0	2	2,647.70	220	242
2DLC TX114	hypo protein Rv0020c	Rv0020c	56862	100.00%	7	7	11	0.06%	18.20%	KLEQTVGDAFAR	95.00%	2.57	0.372	0	0	1	0	2	1,335.50	14	25
2DLC TX114	hypo protein Rv0020c	Rv0020c	56862	100.00%	7	7	11	0.06%	18.20%	LGHSEIIVR	95.00%	2.54	0.261	0	0	2	0	2	1,024.22	517	525
2DLC TX114	hypo protein Rv0020c	Rv0020c	56862	100.00%	7	7	11	0.06%	18.20%	LPDTGVSR	95.00%	1.99	0.299	0	3	0	0	2	844.9384	467	474
2DLC TX114	hypo protein Rv0020c	Rv0020c	56862	100.00%	7	7	11	0.06%	18.20%	VPGYAPQGGGYAEPAGR	95.00%	2.55	0.396	0	1	0	0	2	1,647.74	385	401
2DLC TX114	hypo protein Rv3699	Rv3699	24980	99.80%	2	2	3	0.02%	12.00%	FPAYLLTAHK	95.00%	2.5	0.454	0	0	2	0	2	1,161.50	222	231
2DLC TX114	hypo protein Rv3699	Rv3699	24980	99.80%	2	2	3	0.02%	12.00%	FSTVIDSTLFHSLPVDSR	95.00%	3.66	0.563	0	1	0	0	2	2,022.30	113	130
2DLC TX114	hypo protein Rv2477c	Rv2477c	61876	100.00%	3	3	6	0.03%	5.56%	DLSFSLPR	95.00%	2.04	0.352	0	3	0	0	2	935.1984	339	346
2DLC TX114	hypo protein Rv2477c	Rv2477c	61876	100.00%	3	3	6	0.03%	5.56%	LTEELAWVR	95.00%	3.14	0.167	0	1	0	0	2	1,117.27	266	274
2DLC TX114	hypo protein Rv2477c	Rv2477c	61876	100.00%	3	3	6	0.03%	5.56%	TIVGLETPDSGSVK	95.00%	2.15	0.461	0	2	0	0	2	1,403.54	366	379
2DLC TX114	rpsA (rpsA)	Rv1630	53184	100.00%	5	5	14	0.08%	11.60%	AWLEQTQSEVR	95.00%	3.04	0.409	0	2	0	0	2	1,347.35	189	199
2DLC TX114	rpsA (rpsA)	Rv1630	53184	100.00%	5	5	14	0.08%	11.60%	FAAAEAAGR	95.00%	2.9	0.352	0	2	0	0	2	863.9684	435	443
2DLC TX114	rpsA (rpsA)	Rv1630	53184	100.00%	5	5	14	0.08%	11.60%	GFLPASLVEMR	95.00%	2.75	0.543	0	5	0	0	2	1,236.52	146	156
2DLC TX114	rpsA (rpsA)	Rv1630	53184	100.00%	5	5	14	0.08%	11.60%	SEFLNNLQK	95.00%	2.09	0.397	0	2	0	0	2	1,093.30	200	208
2DLC TX114	rpsA (rpsA)	Rv1630	53184	100.00%	5	5	14	0.08%	11.60%	TAGGSLASDAQLAALR	95.00%	3.58	0.456	0	3	0	0	2	1,502.78	459	474
2DLC TX114	hypo protein Rv3690	Rv3690	23702	99.80%	2	2	3	0.02%	14.70%	ELEETGMLNPAAGR	95.00%	2.8	0.396	0	2	0	0	2	1,488.70	128	141
2DLC TX114	hypo protein Rv3690	Rv3690	23702	99.80%	2	2	3	0.02%	14.70%	GGDYQLFDAGQLTAAQHR	95.00%	3.18	0.51	0	0	1	0	2	1,949.11	85	102
2DLC TX114	hypo protein Rv3676	Rv3676	24742	99.80%	2	2	2	0.01%	8.93%	SVLISDSER	95.00%	1.84	0.34	0	1	0	0	2	1,006.13	210	218
2DLC TX114	hypo protein Rv3676	Rv3676	24742	99.80%	2	2	2	0.01%	8.93%	TSSATTITEVR	95.00%	2.81	0.416	0	1	0	0	2	1,166.23	90	100
2DLC TX114	hypo protein Rv1829	Rv1829	18097	100.00%	4	5	13	0.08%	20.70%	DLIAALGHSLK	95.00%	3.45	0.36	0	2	0	0	2	1,138.51	63	73
2DLC TX114	hypo protein Rv1829	Rv1829	18097	100.00%	4	5	13	0.08%	20.70%	IAALGHSLK	95.00%	2.5	0.406	0	2	0	0	2	910.2084	65	73
2DLC TX114	hypo protein Rv1829	Rv1829	18097	100.00%	4	5	13	0.08%	20.70%	PSDSVAIALR	95.00%	3.02	0.336	0	3	0	0	2	1,029.23	101	110
2DLC TX114	hypo protein Rv1829	Rv1829	18097	100.00%	4	5	13	0.08%	20.70%	VEQPQNQPVLLLR	95.00%	3.2	0.43	0	4	2	0	2	1,534.70	11	23
2DLC TX114	hypo protein Rv1827	Rv1827	17201	100.00%	3	5	12	0.07%	19.10%	FLLDQAITSAGR	95.00%	4.11	0.504	0	3	0	0	2	1,292.61	70	81
2DLC TX114	hypo protein Rv1827	Rv1827	17201	100.00%	3	5	12	0.07%	19.10%	LVFLTGPK	95.00%	1.61	0.336	2	4	0	0	2	875.1584	144	151
2DLC TX114	hypo protein Rv1827	Rv1827	17201	100.00%	3	5	12	0.07%	19.10%	LVFLTGPKQGEDDGSTGGP	95.00%	4.5	0.627	0	2	1	0	2	1,876.06	144	162
2DLC TX114	secA2	Rv1821	88902	99.80%	2	3	6	0.03%	3.84%	MLAGDVIEMATGEGK	95.00%	5.09	0.535	0	3	0	0	2	1,522.80	131	145
2DLC TX114	secA2	Rv1821	88902	99.80%	2	3	6	0.03%	3.84%	TLAGAIAAAGYALAGR	95.00%	3.04	0.476	0	2	1	0	2	1,447.83	146	161
2DLC TX114	dnaK (dnaK)	Rv0350	66813	100.00%	16	23	73	0.42%	27.00%	AALGGSDISAIK	95.00%	3.02	0.591	2	10	0	0	2	1,103.42	553	564
2DLC TX114	dnaK (dnaK)	Rv0350	66813	100.00%	16	23	73	0.42%	27.00%	DAGQIAGLNVLR	95.00%	3.78	0.425	4	5	2	0	2	1,227.45	132	143
2DLC TX114	dnaK (dnaK)	Rv0350	66813	100.00%	16	23	73	0.42%	27.00%	EAEGGSKVPEDTLNKVDAA VAEAK	95.00%	5.21	0.544	0	0	3	0	2	2,429.53	529	552
2DLC TX114	dnaK (dnaK)	Rv0350	66813	100.00%	16	23	73	0.42%	27.00%	GTSGIDLTK	95.00%	2.03	0.362	0	1	0	0	2	892.0984	218	226
2DLC TX114	dnaK (dnaK)	Rv0350	66813	100.00%	16	23	73	0.42%	27.00%	ITQDLLDR	95.00%	1.96	0.191	1	3	0	0	2	974.2184	281	288
2DLC TX114	dnaK (dnaK)	Rv0350	66813	100.00%	16	23	73	0.42%	27.00%	IVNEPTAAALAYGLDKGEK	95.00%	1.6	0.31	0	0	1	0	2	1,961.33	144	162
2DLC TX114	dnaK (dnaK)	Rv0350	66813	100.00%	16	23	73	0.42%	27.00%	LLGSFELTGIPPAPR	95.00%	2.46	0.417	0	2	0	0	2	1,569.04	427	441
2DLC TX114	dnaK (dnaK)	Rv0350	66813	100.00%	16	23	73	0.42%	27.00%	MPAVTDLVK	95.00%	1.69	0.299	1	9	0	0	2	974.1584	319	327
2DLC TX114	dnaK (dnaK)	Rv0350	66813	100.00%	16	23	73	0.42%	27.00%	NGEVLVGQPAK	95.00%	2.55	0.387	0	1	0	0	2	1,112.16	45	55
2DLC TX114	dnaK (dnaK)	Rv0350	66813	100.00%	16	23	73	0.42%	27.00%	NQAETLVYQTEK	95.00%	3.3	0.497	0	3	3	0	2	1,424.47	511	522
2DLC TX114	dnaK (dnaK)	Rv0350	66813	100.00%	16	23	73	0.42%	27.00%	NQAPNVDR	95.00%	2.47	0.298	0	2	0	0	2	1,016.94	56	64
2DLC TX114	dnaK (dnaK)	Rv0350	66813	100.00%	16	23	73	0.42%	27.00%	TTPSIVAFAR	95.00%	2.07	0.511	0	6	0	0	2	1,063.25	35	44
2DLC TX114	dnaK (dnaK)	Rv0350	66813	100.00%	16	23	73	0.42%	27.00%	TTPSIVAFARNGEVLVGQPA K	95.00%	1.98	0.495	0	0	1	0	2	2,156.39	35	55
2DLC TX114	dnaK (dnaK)	Rv0350	66813	100.00%	16	23	73	0.42%	27.00%	VVDWLVDK	95.00%	2.32	0.46	0	4	0	0	2	974.0284	208	215
2DLC TX114	dnaK (dnaK)	Rv0350	66813	100.00%	16	23	73	0.42%	27.00%	VVDWLVDKFK	95.00%	4.26	0.554	0	2	3	0	2	1,249.43	208	217
2DLC TX114	dnaK (dnaK)	Rv0350	66813	100.00%	16	23	73	0.42%	27.00%	YTAPEISAR	95.00%	2.1	0.45	0	4	0	0	2	1,008.18	86	94
2DLC TX114	grpE (grpE)	Rv0351	24482	99.80%	2	2	3	0.02%	8.51%	QGYQLGEQVLR	95.00%	2.11	0.364	0	1	0	0	2	1,291.39	170	180
2DLC TX114	grpE (grpE)	Rv0351	24482	99.80%	2	2	3	0.02%	8.51%	VQADFANYR	95.00%	2.21	0.488	0	2	0	0	2	1,084.15	65	73
2DLC TX114	hypo protein Rv1198	Rv1198	9910	99.80%	2	2	3	0.02%	31.90%	AQAGLLEAEHQAIIR	95.00%	1.92	0.502	0	0	2	0	2	1,620.89	19	33
2DLC TX114	hypo protein Rv1198	Rv1198	9910	99.80%	2	2	3	0.02%	31.90%	NFQVIYEQANAHGQK	95.00%	2.08	0.256	0	0	1	0	2	1,747.80	60	74
2DLC TX114	PPE (PPE)	Rv1196	39140	100.00%	4	5	11	0.06%	15.30%	AGGGLSGVLR	95.00%	3.02	0.252	0	5	0	0	2	887.0484	367	376
2DLC TX114	PPE (PPE)	Rv1196	39140	100.00%	4	5	11	0.06%	15.30%	ALPLTSLTSAAER	95.00%	2.83	0.613	0	2	0	0	2	1,330.64	337	349
2DLC TX114	PPE (PPE)	Rv1196	39140	100.00%	4	5	11	0.06%	15.30%	GFAPAAAAQAVQTAAQNGV R	95.00%	2.66	0.447	0	0	1	0	2	1,899.92	268	287
2DLC TX114	PPE (PPE)	Rv1196	39140	100.00%	4	5	11	0.06%	15.30%	GPGQMLGGLPVGQMGAR	95.00%	3.58	0.531	0	2	1	0	2	1,626.91	350	366
2DLC TX114	pks7 (pks7)	Rv1661	221021	100.00%	3	3	3	0.02%	2.07%	LDRAAL	91.10%	1.16	0.0246	1	0	0	0	2	658.8784	1922	1927
2DLC TX114	pks7 (pks7)	Rv1661	221021	100.00%	3	3	3	0.02%	2.07%	MTQAGLAPLTTEQALGFLDT AL	84.80%	2.33	0.153	0	1	0	0	2	2,279.76	1889	1910
2DLC TX114	pks7 (pks7)	Rv1661	221021	100.00%	3	3	3	0.02%	2.07%	SGEQADRAAEVAALLR	95.00%	2.28	0.23	0	0	1	0	2	1,657.82	1728	1743
2DLC TX114	infB (infB)	Rv2839c	93992	100.00%	4	4	6	0.03%	5.44%	GGGVGAAPGTGFR	95.00%	2.27	0.373	0	1	0	0	2	1,104.17	234	246
2DLC TX114	infB (infB)	Rv2839c	93992	100.00%	4	4	6	0.03%	5.44%	GPVATVLVQR	95.00%	3.3	0.5	0	1	0	0	2	1,040.08	591	600
2DLC TX114	infB (infB)	Rv2839c	93992	100.00%	4	4	6	0.03%	5.44%	LITFIDTPGHEAFTAMR	95.00%	1.62	0.292	0	0	1	0	2	1,937.35	450	466
2DLC TX114	infB (infB)	Rv2839c	93992	100.00%	4	4	6	0.03%	5.44%	LSEQGEFVK	95.00%	2.61	0.438	0	3	0	0	2	1,037.12	26	34
2DLC TX114	pks4 (pks4)	Rv1181	168064	100.00%	6	6	9	0.05%	5.75%	AGIELLAFGGR	95.00%	3.89	0.527	0	1	0	0	2	1,104.43	1119	1129
2DLC TX114	pks4 (pks4)	Rv1181	168064	100.00%	6	6	9	0.05%	5.75%	ANLEQAGLR	95.00%	2.29	0.338	0	1	0	0	2	972.1284	813	821
2DLC TX114	pks4 (pks4)	Rv1181	168064	100.00%	6	6	9	0.05%	5.75%	ATAAQPL	81.30%	0.809	0	1	0	0	0	2	671.7784	1362	1368
2DLC TX114	pks4 (pks4)	Rv1181	168064	100.00%	6	6	9	0.05%	5.75%	ELAELPGEPPR	95.00%	2.03	0.505	0	1	0	0	2	1,208.41	786	796
2DLC TX114	pks4 (pks4)	Rv1181	168064	100.00%	6	6	9	0.05%	5.75%	FLAELGSLPR	95.00%	3.27	0.374	0	2	0	0	2	1,103.45	1492	1501
2DLC TX114	pks4 (pks4)	Rv1181	168064	100.00%	6	6	9	0.05%	5.75%	FLAELGSLPREEWPR	95.00%	2.18	0.197	0	0	2	0	2	1,801.17	1492	1506
2DLC TX114	pks4 (pks4)	Rv1181	168064	100.00%	6	6	9	0.05%	5.75%	IHNVAAFPGAAYCEMALAA AR	95.00%	2.68	0.327	0	1	0	0	2	2,148.50	430	450
2DLC TX114	pks4 (pks4)	Rv1181	168064	100.00%	6	6	9	0.05%	5.75%	LGLFPFR	95.00%	1.87	0.319	0	2	0	0	2	850.1884	1144	1150

2DLC TX114	pks4 (pks4)	Rv1181	168064	100.00%	6	6	9	0.05%	5.75%	TTLGELSEVR	95.00%	2.73	0.479	0	2	0	0	2	1,105.22	451	460
2DLC TX114	hypo protein Rv1586c	Rv1586c	51693	99.50%	2	2	2	0.01%	6.40%	QAYADILAGASL	95.00%	1.83	0.249	0	0	1	0	2	1,193.47	183	194
2DLC TX114	hypo protein Rv1586c	Rv1586c	51693	99.50%	2	2	2	0.01%	6.40%	VDEATFWAAQAVLDAPGR	90.40%	1.64	0.227	0	0	1	0	2	1,918.03	257	274
2DLC TX114	hypo protein Rv3237c	Rv3237c	17216	99.80%	2	2	2	0.01%	10.00%	EVLLPGVGLR	95.00%	2.32	0.448	0	1	0	0	2	1,053.28	5	14
2DLC TX114	hypo protein Rv3237c	Rv3237c	17216	99.80%	2	2	2	0.01%	10.00%	FTELTR	95.00%	1.61	0.354	0	1	0	0	2	766.9184	75	80
2DLC TX114	hypo protein Rv1144	Rv1144	25770	99.80%	2	2	4	0.02%	7.20%	DGVFPL	95.00%	1.46	0.0613	1	0	0	0	2	647.7584	95	100
2DLC TX114	hypo protein Rv1144	Rv1144	25770	99.80%	2	2	4	0.02%	7.20%	GGVVGMTLPIAR	95.00%	3.47	0.539	0	3	0	0	2	1,171.41	161	172
2DLC TX114	groES	Rv3418c	10754	100.00%	7	9	37	0.21%	67.00%	DVLAVVSK	95.00%	2.03	0.44	7	6	0	0	2	830.8884	93	100
2DLC TX114	groES	Rv3418c	10754	100.00%	7	9	37	0.21%	67.00%	EKPQEGTVVAVGPGR	95.00%	2.29	0.273	0	0	2	0	2	1,524.50	36	50
2DLC TX114	groES	Rv3418c	10754	100.00%	7	9	37	0.21%	67.00%	IPLDVAEGDTVIYSK	95.00%	3.84	0.568	0	6	0	0	2	1,620.89	59	73
2DLC TX114	groES	Rv3418c	10754	100.00%	7	9	37	0.21%	67.00%	RIPLDVAEGDTVIYSK	95.00%	3.44	0.586	0	3	5	0	2	1,777.09	58	73
2DLC TX114	groES	Rv3418c	10754	100.00%	7	9	37	0.21%	67.00%	VNIKPLEDK	95.00%	1.93	0.304	0	0	1	0	2	1,056.31	4	12
2DLC TX114	groES	Rv3418c	10754	100.00%	7	9	37	0.21%	67.00%	YGGTEIK	95.00%	2.13	0.389	0	2	0	0	2	767.9184	74	80
2DLC TX114	groES	Rv3418c	10754	100.00%	7	9	37	0.21%	67.00%	YNGEEYLILSAR	95.00%	4.07	0.573	0	5	0	0	2	1,428.73	81	92
2DLC TX114	groEL1 (groEL1)	Rv3417c	55859	100.00%	22	29	107	0.62%	47.30%	AADAVSEALLASATPVSGK	95.00%	3.71	0.591	0	2	5	0	2	1,758.95	122	140
2DLC TX114	groEL1 (groEL1)	Rv3417c	55859	100.00%	22	29	107	0.62%	47.30%	AAVEEGIVPGGGASLIHQA R	95.00%	4.31	0.656	0	7	11	0	2	1,933.08	403	422
2DLC TX114	groEL1 (groEL1)	Rv3417c	55859	100.00%	22	29	107	0.62%	47.30%	AFGGPTVTNDGVTVAR	95.00%	2.55	0.38	0	2	0	0	2	1,562.56	42	57
2DLC TX114	groEL1 (groEL1)	Rv3417c	55859	100.00%	22	29	107	0.62%	47.30%	AGGVAVIK	95.00%	1.59	0.304	2	1	0	0	2	714.8184	371	378
2DLC TX114	groEL1 (groEL1)	Rv3417c	55859	100.00%	22	29	107	0.62%	47.30%	AMEVGMDKLADTVR	95.00%	3.85	0.529	0	3	6	0	2	1,552.77	14	27
2DLC TX114	groEL1 (groEL1)	Rv3417c	55859	100.00%	22	29	107	0.62%	47.30%	ESVEDAVAAAK	95.00%	1.96	0.571	0	3	0	0	2	1,090.06	392	402
2DLC TX114	groEL1 (groEL1)	Rv3417c	55859	100.00%	22	29	107	0.62%	47.30%	EVGLEVLGSAR	95.00%	3.16	0.566	3	13	0	0	2	1,130.22	309	319
2DLC TX114	groEL1 (groEL1)	Rv3417c	55859	100.00%	22	29	107	0.62%	47.30%	GPYFGDR	95.00%	1.99	0.439	0	7	0	0	2	811.9384	276	282
2DLC TX114	groEL1 (groEL1)	Rv3417c	55859	100.00%	22	29	107	0.62%	47.30%	KALTELR	95.00%	2.03	0.243	0	4	0	0	2	831.0984	423	429
2DLC TX114	groEL1 (groEL1)	Rv3417c	55859	100.00%	22	29	107	0.62%	47.30%	KESVEDAVAAAK	95.00%	4.23	0.442	0	1	3	0	2	1,218.26	391	402
2DLC TX114	groEL1 (groEL1)	Rv3417c	55859	100.00%	22	29	107	0.62%	47.30%	LAGGVAVIK	95.00%	1.86	0.247	5	15	0	0	2	828.0184	370	378
2DLC TX114	groEL1 (groEL1)	Rv3417c	55859	100.00%	22	29	107	0.62%	47.30%	LIEYDETAR	95.00%	2.51	0.429	0	1	0	0	2	1,110.30	4	12
2DLC TX114	groEL1 (groEL1)	Rv3417c	55859	100.00%	22	29	107	0.62%	47.30%	LPLLEK	95.00%	1.77	0.368	3	1	0	0	2	713.0384	232	237
2DLC TX114	groEL1 (groEL1)	Rv3417c	55859	100.00%	22	29	107	0.62%	47.30%	LVAAGVNPIAL	95.00%	2.04	0.32	2	0	0	0	2	1,038.27	105	115
2DLC TX114	groEL1 (groEL1)	Rv3417c	55859	100.00%	22	29	107	0.62%	47.30%	LVAAGVNPIALGVGIGK	95.00%	4.79	0.619	0	3	4	0	2	1,549.89	105	121
2DLC TX114	groEL1 (groEL1)	Rv3417c	55859	100.00%	22	29	107	0.62%	47.30%	MVLTTETVVVDKPAK	95.00%	1.99	0.392	0	0	1	0	2	1,647.80	513	527
2DLC TX114	groEL1 (groEL1)	Rv3417c	55859	100.00%	22	29	107	0.62%	47.30%	PDAGMVLR	95.00%	2.64	0.449	0	2	0	0	2	859.0384	301	308
2DLC TX114	groEL1 (groEL1)	Rv3417c	55859	100.00%	22	29	107	0.62%	47.30%	SAVLNASSVAR	95.00%	3.76	0.425	0	7	0	0	2	1,075.14	502	512
2DLC TX114	groEL1 (groEL1)	Rv3417c	55859	100.00%	22	29	107	0.62%	47.30%	TGIAQVATVSSR TGIAQVATVSSRDEQIGDL	95.00%	3.06	0.553	0	1	0	0	2	1,190.23	141	152
2DLC TX114	groEL1 (groEL1)	Rv3417c	55859	100.00%	22	29	107	0.62%	47.30%	VGEAMSK	95.00%	5.29	0.73	0	0	6	0	2	2,663.86	141	166
2DLC TX114	groEL1 (groEL1)	Rv3417c	55859	100.00%	22	29	107	0.62%	47.30%	VGAATETALK	95.00%	2.49	0.296	0	2	0	0	2	961.0784	379	388
2DLC TX114	groEL1 (groEL1)	Rv3417c	55859	100.00%	22	29	107	0.62%	47.30%	VGAATETALKER	95.00%	2.68	0.492	0	1	3	0	2	1,246.38	379	390
2DLC TX114	groEL1 (groEL1)	Rv3417c	55859	100.00%	22	29	107	0.62%	47.30%	VGEAMSK	95.00%	0.983	-0.0794	1	0	0	0	2	737.7978	160	166
2DLC TX114	groEL1 (groEL1)	Rv3417c	55859	100.00%	22	29	107	0.62%	47.30%	VSELPAGHGLNVN VVVSKDDTVIVDGGGTAEA	95.00%	2.54	0.325	0	2	0	0	2	1,307.34	469	481
2DLC TX114	groEL1 (groEL1)	Rv3417c	55859	100.00%	22	29	107	0.62%	47.30%	VANR	95.00%	1.79	0.26	0	0	1	0	2	2,273.21	321	343
2DLC TX114	hypo protein Rv1100	Rv1100	24545	100.00%	4	6	11	0.06%	22.30%	LPQLPGGWTPNSGGR	95.00%	2.91	0.503	0	2	0	0	2	1,537.76	97	111
2DLC TX114	hypo protein Rv1100	Rv1100	24545	100.00%	4	6	11	0.06%	22.30%	NAATSIVGFISPTGR	95.00%	3.22	0.347	0	4	1	0	2	1,491.73	128	142
2DLC TX114	hypo protein Rv1100	Rv1100	24545	100.00%	4	6	11	0.06%	22.30%	TLASATQSQPPLPAR	95.00%	3.24	0.446	0	1	2	0	2	1,538.78	219	233
2DLC TX114	hypo protein Rv1100	Rv1100	24545	100.00%	4	6	11	0.06%	22.30%	TLGFPIR	88.70%	1.93	0.213	0	1	0	0	2	804.0884	90	96
2DLC TX114	hypo protein Rv0504c	Rv0504c	18342	99.80%	2	2	4	0.02%	10.80%	APDHFLVGR	95.00%	2.66	0.41	0	0	1	0	2	1,012.14	17	25
2DLC TX114	hypo protein Rv0504c	Rv0504c	18342	99.80%	2	2	4	0.02%	10.80%	FNIPINIAR	95.00%	2.79	0.264	0	3	0	0	2	1,058.42	78	86
2DLC TX114	nrdZ (nrdZ)	Rv0570	74472	100.00%	5	5	6	0.03%	10.40%	AAFPTAAEIAPQWHLR FLDDVIDVSRYPFPELGEAA R	81.10%	1.55	0.204	0	0	1	0	2	1,780.06	594	609
2DLC TX114	nrdZ (nrdZ)	Rv0570	74472	100.00%	5	5	6	0.03%	10.40%	R	95.00%	5	0.595	0	0	1	0	2	2,410.79	419	439
2DLC TX114	nrdZ (nrdZ)	Rv0570	74472	100.00%	5	5	6	0.03%	10.40%	GAFPAFTDSR	95.00%	2.24	0.357	0	2	0	0	2	1,069.23	494	503
2DLC TX114	nrdZ (nrdZ)	Rv0570	74472	100.00%	5	5	6	0.03%	10.40%	LGEAGLDDVAR	95.00%	3.03	0.407	0	1	0	0	2	1,116.25	72	82
2DLC TX114	nrdZ (nrdZ)	Rv0570	74472	100.00%	5	5	6	0.03%	10.40%	LYDSAAGVVSMGGR	95.00%	1.47	0.341	0	1	0	0	2	1,383.53	252	265
2DLC TX114	fum (fum)	Rv1098c	50124	100.00%	4	4	7	0.04%	5.70%	AVENFPISGR	95.00%	3.25	0.408	0	2	0	0	2	1,090.22	35	44
2DLC TX114	fum (fum)	Rv1098c	50124	100.00%	4	4	7	0.04%	5.70%	DRRLDVLAMAK	95.00%	1.47	0.329	0	0	1	0	2	1,288.65	455	465
2DLC TX114	fum (fum)	Rv1098c	50124	100.00%	4	4	7	0.04%	5.70%	LDVLAMAK	95.00%	2.34	0.455	0	1	0	0	2	861.1484	458	465
2DLC TX114	fum (fum)	Rv1098c	50124	100.00%	4	4	7	0.04%	5.70%	LSIEDLDR	95.00%	2.17	0.213	0	3	0	0	2	961.1984	449	456
2DLC TX114	grcC1	Rv0562	35510	100.00%	4	4	6	0.03%	14.60%	AKDVLAQYAAQAR	95.00%	3.05	0.441	0	0	1	0	2	1,405.59	296	308
2DLC TX114	grcC1	Rv0562	35510	100.00%	4	4	6	0.03%	14.60%	ALAALVDYTVSR	95.00%	4.01	0.435	0	1	0	0	2	1,279.48	322	333
2DLC TX114	grcC1	Rv0562	35510	100.00%	4	4	6	0.03%	14.60%	GTSENVDSIEQYLK	95.00%	3.53	0.474	0	2	0	0	2	1,583.70	164	177
2DLC TX114	grcC1	Rv0562	35510	100.00%	4	4	6	0.03%	14.60%	TGSLIGAAGR	95.00%	2.89	0.476	0	2	0	0	2	903.1084	183	192
2DLC TX114	pknD (pknD)	Rv0931c	69494	100.00%	5	7	22	0.13%	9.19%	LISPQYSDNAVFR	95.00%	3.86	0.556	0	5	2	0	2	1,510.75	45	57

2DLC TX114	pknD (pknD)	Rv0931c	69494	100.00%	5	7	22	0.13%	9.19%	MIDGTSLR	95.00%	1.89	0.265	0	1	0	0	2	893.1484	94	101
2DLC TX114	pknD (pknD)	Rv0931c	69494	100.00%	5	7	22	0.13%	9.19%	TPLAVAVDSDR	95.00%	2.49	0.452	0	2	0	0	2	1,144.22	637	647
2DLC TX114	pknD (pknD)	Rv0931c	69494	100.00%	5	7	22	0.13%	9.19%	VPPALDQVIAK	95.00%	2.46	0.548	0	3	0	0	2	1,151.36	242	252
2DLC TX114	pknD (pknD)	Rv0931c	69494	100.00%	5	7	22	0.13%	9.19%	WSPGDSATVAGPLAADSR	95.00%	4.13	0.453	0	6	3	0	2	1,758.89	327	344
2DLC TX114	hypo protein Rv1097c	Rv1097c	29855	100.00%	3	4	7	0.04%	9.56%	TSSITVDGVR	95.00%	2.76	0.493	0	3	0	0	2	1,035.07	222	231
2DLC TX114	hypo protein Rv1097c	Rv1097c	29855	100.00%	3	4	7	0.04%	9.56%	VDADITIADSSR	95.00%	3.61	0.547	0	1	0	0	2	1,263.41	235	246
2DLC TX114	hypo protein Rv1097c	Rv1097c	29855	100.00%	3	4	7	0.04%	9.56%	VIEALK	95.00%	1.63	0.198	2	1	0	0	2	672.8684	284	289
2DLC TX114	hypo protein Rv2161c	Rv2161c	30750	97.40%	2	2	2	0.01%	9.72%	AFEVIAR	84.80%	2.06	0.104	0	1	0	0	2	805.9484	185	191
2DLC TX114	hypo protein Rv2161c	Rv2161c	30750	97.40%	2	2	2	0.01%	9.72%	MQFVTDLTPPPQLVAVWAE ER	82.90%	2.41	0.13	0	0	1	0	2	2,428.65	6	26
2DLC TX114	guaB2	Rv3411c	54848	99.80%	2	2	4	0.02%	5.67%	FEVDQSKQVAEVMTK	95.00%	3.51	0.537	0	0	2	0	2	1,739.79	177	191
2DLC TX114	guaB2	Rv3411c	54848	99.80%	2	2	4	0.02%	5.67%	SAAAALVDAGADAVK	95.00%	2.75	0.496	0	2	0	0	2	1,330.45	318	332
2DLC TX114	dapB (dapB)	Rv2773c	25715	99.80%	2	2	3	0.02%	10.20%	GADVDGIPVHAVR	95.00%	2.37	0.343	0	0	1	0	2	1,306.31	174	186
2DLC TX114	dapB (dapB)	Rv2773c	25715	99.80%	2	2	3	0.02%	10.20%	TSFVPGVLLAVR	95.00%	2.61	0.554	0	2	0	0	2	1,259.46	215	226
2DLC TX114	mrr (mrr)	Rv2528c	33631	99.80%	2	2	4	0.02%	5.23%	LLEAMGYGR	95.00%	3.13	0.389	0	2	0	0	2	1,010.30	175	183
2DLC TX114	mrr (mrr)	Rv2528c	33631	99.80%	2	2	4	0.02%	5.23%	VDMAVLR	95.00%	2.08	0.377	0	2	0	0	2	803.9384	97	103
2DLC TX114	ppp (ppp)	Rv0018c	53763	99.80%	2	2	4	0.02%	2.33%	KEIVK	95.00%	1.4	0.0609	3	0	0	0	2	616.7884	278	282
2DLC TX114	ppp (ppp)	Rv0018c	53763	99.80%	2	2	4	0.02%	2.33%	LIELALR	95.00%	2.54	0.0909	0	1	0	0	2	828.1984	218	224
2DLC TX114	hypo protein Rv2955c	Rv2955c	35901	99.80%	2	2	3	0.02%	8.41%	AVIAGAAELLR	95.00%	3.61	0.527	0	2	0	0	2	1,084.36	247	257
2DLC TX114	hypo protein Rv2955c	Rv2955c	35901	99.80%	2	2	3	0.02%	8.41%	LAPVVALEPAPGTHSR	95.00%	2.04	0.383	0	0	1	0	2	1,615.79	148	163
2DLC TX114	hypo protein Rv2345	Rv2345	70012	99.80%	2	2	2	0.01%	3.94%	ALLDAVDSAATDIR	95.00%	4.79	0.652	0	1	0	0	2	1,431.69	412	425
2DLC TX114	hypo protein Rv2345	Rv2345	70012	99.80%	2	2	2	0.01%	3.94%	SLEQALFTAESR	95.00%	3.02	0.577	0	1	0	0	2	1,352.54	517	528
2DLC TX114	argH (argH)	Rv1659	49725	99.80%	2	3	5	0.03%	5.96%	FAGGPSDALAALSK	95.00%	4.57	0.561	0	2	1	0	2	1,305.62	13	26
2DLC TX114	argH (argH)	Rv1659	49725	99.80%	2	3	5	0.03%	5.96%	VAEQLNAIGEAAER	95.00%	3.34	0.496	0	2	0	0	2	1,471.56	450	463
2DLC TX114	sigK (sigK)	Rv0445c	21017	99.80%	2	2	4	0.02%	10.20%	LAANLSTIK	95.00%	2.58	0.323	0	1	0	0	2	931.2584	162	170
2DLC TX114	sigK (sigK)	Rv0445c	21017	99.80%	2	2	4	0.02%	10.20%	LSSDLDALLR	95.00%	2.67	0.403	0	3	0	0	2	1,103.46	7	16
2DLC TX114	hypo protein Rv2298	Rv2298	34969	100.00%	3	3	4	0.02%	8.67%	FLTDMVR	95.00%	2.11	0.188	0	2	0	0	2	882.0884	311	317
2DLC TX114	hypo protein Rv2298	Rv2298	34969	100.00%	3	3	4	0.02%	8.67%	IEPLLATLR	83.00%	1.69	0.286	0	1	0	0	2	1,026.42	233	241
2DLC TX114	hypo protein Rv2298	Rv2298	34969	100.00%	3	3	4	0.02%	8.67%	YLDVDGIGQVSR	95.00%	2.65	0.521	0	1	0	0	2	1,322.44	3	14
2DLC TX114	hypo protein Rv0444c	Rv0444c	23865	100.00%	5	7	14	0.08%	35.30%	ETMAVVSAATTAEPPAHLR	95.00%	1.5	0.249	0	0	1	0	2	1,969.08	53	71
2DLC TX114	hypo protein Rv0444c	Rv0444c	23865	100.00%	5	7	14	0.08%	35.30%	GAGTATVVFSR	95.00%	2.59	0.403	0	1	0	0	2	1,066.10	138	148
2DLC TX114	hypo protein Rv0444c	Rv0444c	23865	100.00%	5	7	14	0.08%	35.30%	GVLTRPSPPPTVAEQVLTAP DVR	95.00%	2.96	0.511	0	0	3	0	2	2,401.59	109	131
2DLC TX114	hypo protein Rv0444c	Rv0444c	23865	100.00%	5	7	14	0.08%	35.30%	TAILDATKPEVR	95.00%	2.9	0.359	0	2	2	0	2	1,314.57	72	83
2DLC TX114	hypo protein Rv0444c	Rv0444c	23865	100.00%	5	7	14	0.08%	35.30%	VAAAPSPVAAAFNDEVR	95.00%	4.7	0.632	0	2	3	0	2	1,685.73	33	49
2DLC TX114	fas (fas)	Rv2524c	326201	100.00%	17	17	23	0.13%	6.94%	ADAILIR	95.00%	2.43	0.194	0	1	0	0	2	772.0784	316	322
2DLC TX114	fas (fas)	Rv2524c	326201	100.00%	17	17	23	0.13%	6.94%	AVVITGTPEQLSR	95.00%	1.95	0.25	0	1	0	0	2	1,371.49	235	247
2DLC TX114	fas (fas)	Rv2524c	326201	100.00%	17	17	23	0.13%	6.94%	ELADAILIR	95.00%	2.04	0.232	0	1	0	0	2	1,014.38	314	322
2DLC TX114	fas (fas)	Rv2524c	326201	100.00%	17	17	23	0.13%	6.94%	FVEIGVK	94.10%	2.27	0.115	0	1	0	0	2	791.9084	1684	1690
2DLC TX114	fas (fas)	Rv2524c	326201	100.00%	17	17	23	0.13%	6.94%	GLGIGIVPAATR	95.00%	3.34	0.365	0	2	0	0	2	1,125.42	360	371
2DLC TX114	fas (fas)	Rv2524c	326201	100.00%	17	17	23	0.13%	6.94%	LASAIAGGAPMYQRPGDR	95.00%	1.81	0.308	0	0	1	0	2	1,832.21	3021	3038
2DLC TX114	fas (fas)	Rv2524c	326201	100.00%	17	17	23	0.13%	6.94%	LFTLDRDFIQEIR	95.00%	2.82	0.434	0	0	2	0	2	1,667.12	1606	1618
2DLC TX114	fas (fas)	Rv2524c	326201	100.00%	17	17	23	0.13%	6.94%	LLDGGATVIATTSKLDEER	95.00%	1.94	0.344	0	0	1	0	2	1,990.33	2137	2155
2DLC TX114	fas (fas)	Rv2524c	326201	100.00%	17	17	23	0.13%	6.94%	MLVDTQGTDQWISAGK	95.00%	3.59	0.405	0	2	0	0	2	1,750.95	654	669
2DLC TX114	fas (fas)	Rv2524c	326201	100.00%	17	17	23	0.13%	6.94%	NLFTVGATPEVAR	95.00%	2.43	0.415	0	2	0	0	2	1,375.49	376	388
2DLC TX114	fas (fas)	Rv2524c	326201	100.00%	17	17	23	0.13%	6.94%	PVAACATAAVSVEEGVDKI R	95.00%	1.67	0.301	0	0	1	0	2	1,987.05	2716	2735
2DLC TX114	fas (fas)	Rv2524c	326201	100.00%	17	17	23	0.13%	6.94%	SFILVPGIDVPFHSR	95.00%	2.05	0.413	0	0	1	0	2	1,685.01	1554	1568
2DLC TX114	fas (fas)	Rv2524c	326201	100.00%	17	17	23	0.13%	6.94%	SLVHLDHAAR	89.30%	1.79	0.256	0	1	0	0	2	1,119.23	1116	1125
2DLC TX114	fas (fas)	Rv2524c	326201	100.00%	17	17	23	0.13%	6.94%	TYKPYGPVLSDAINDQLR	95.00%	4.62	0.535	0	0	3	0	2	2,051.44	1848	1865
2DLC TX114	fas (fas)	Rv2524c	326201	100.00%	17	17	23	0.13%	6.94%	VGIDQGAEIVDVAAR	95.00%	3.84	0.499	0	1	0	0	2	1,513.57	1300	1314
2DLC TX114	fas (fas)	Rv2524c	326201	100.00%	17	17	23	0.13%	6.94%	VGVAEFR	95.00%	1.67	0.233	0	1	0	0	2	777.7884	1572	1578
2DLC TX114	fas (fas)	Rv2524c	326201	100.00%	17	17	23	0.13%	6.94%	VLLWAVQR	95.00%	2.37	0.224	0	1	0	0	2	985.1384	2242	2249
2DLC TX114	pknB (pknB)	Rv0014c	66491	99.80%	2	2	2	0.01%	4.15%	ADLARDPSFYLR	95.00%	2.58	0.223	0	0	1	0	2	1,424.78	44	55
2DLC TX114	pknB (pknB)	Rv0014c	66491	99.80%	2	2	2	0.01%	4.15%	GQSSADAIATLQNR	95.00%	2.98	0.41	0	1	0	0	2	1,432.57	367	380
2DLC TX114	hypo protein Rv1794	Rv1794	32381	99.80%	2	2	5	0.03%	6.00%	GPSASVLSLK	95.00%	2.8	0.485	0	3	0	0	2	959.1784	236	245
2DLC TX114	hypo protein Rv1794	Rv1794	32381	99.80%	2	2	5	0.03%	6.00%	IALYQQAR	95.00%	2.48	0.485	0	2	0	0	2	963.1784	252	259
2DLC TX114	hypo protein Rv2302	Rv2302	8572	99.80%	2	2	4	0.02%	23.70%	SADGSPPYVVR	95.00%	2.15	0.391	0	2	0	0	2	1,148.19	30	40
2DLC TX114	hypo protein Rv2302	Rv2302	8572	99.80%	2	2	4	0.02%	23.70%	VGDYLVVK	95.00%	2.98	0.247	0	2	0	0	2	892.9784	5	12
2DLC TX114	hypo protein Rv1784	Rv1784	101451	100.00%	3	3	4	0.02%	3.00%	FGVEQVR	95.00%	2.44	0.339	0	2	0	0	2	834.8084	671	677
2DLC TX114	hypo protein Rv1784	Rv1784	101451	100.00%	3	3	4	0.02%	3.00%	FLDALWGEIAR	95.00%	2.91	0.369	0	1	0	0	2	1,291.63	105	115
2DLC TX114	hypo protein Rv1784	Rv1784	101451	100.00%	3	3	4	0.02%	3.00%	VGTVIIDQLR	95.00%	2.89	0.369	0	1	0	0	2	1,114.31	333	342
2DLC TX114	hypo protein Rv1782	Rv1782	53671	99.80%	2	2	2	0.01%	7.71%	ADALVEHDTLPMDMTPAEL VVPK	95.00%	3.2	0.227	0	0	1	0	2	2,493.83	484	506
2DLC TX114	hypo protein Rv1782	Rv1782	53671	99.80%	2	2	2	0.01%	7.71%	VVSGPTIPVAATEMNK	95.00%	2.65	0.497	0	0	1	0	2	1,614.76	379	394

APPENDIX II

Supplemental Table 2. Cell Wall Protein Identifications

ORF	SHORT NAME	PROTEIN_NAME	CLASS ID TUBERCULIST	CLASS ID SANGER	BIOLOGICAL SAMPLE	REFERENCE
Rv0002	dnaN	DNA POLYMERASE III (BETA CHAIN) DNAN (DNA NUCLEOTIDYLTRANSFERASE)	2	II.A.5	D	16
Rv0005	gyrB	DNA GYRASE (SUBUNIT B) GYRB (DNA TOPOISOMERASE (ATP-HYDROLYSING))	2	II.A.5	D	22, 24
Rv0006	gyrA	DNA GYRASE (SUBUNIT A) GYRA (DNA TOPOISOMERASE (ATP-HYDROLYSING))	2	II.A.5	D	22
Rv0007	Rv0007	POSSIBLE CONSERVED MEMBRANE PROTEIN	3	V	D	
Rv0009	ppiA	PROBABLE IRON-REGULATED PEPTIDYL-PROLYL CIS-TRANS ISOMERASE A PPIA (PPIase A)	2	II.A.6	A, C, D	42, 22, 16, 24
Rv0014c	pknB	TRANSMEMBRANE SERINE/THREONINE-PROTEIN KINASE B PKNB (PROTEIN KINASE B) (STPK B)	9	I.J.3	D	
Rv0018c	pstP	POSSIBLE SERINE/THREONINE PHOSPHATASE PSTP	9	I.J.3	D	
Rv0020c	TB39.8	CONSERVED HYPOTHETICAL PROTEIN TB39.8	10	V	D	42, 24
Rv0036c	Rv0036c	CONSERVED HYPOTHETICAL PROTEIN	10	V	B, D	42, 22
Rv0048c	Rv0048c	POSSIBLE MEMBRANE PROTEIN	3	VI	D	22, 24
Rv0053	rpsF	PROBABLE 30S RIBOSOMAL PROTEIN S6 RPSF	2	II.A.1	D	42, 22
Rv0056	rplI	PROBABLE 50S RIBOSOMAL PROTEIN L9 RPLI	2	II.A.1	D	42, 22, 24
Rv0064	Rv0064	PROBABLE CONSERVED TRANSMEMBRANE PROTEIN	3	II.C.5	B	16, 24
Rv0066c	icd2	PROBABLE ISOCITRATE DEHYDROGENASE [NADP] ICD2 (OxALOSUCCINATE DECARBOxYLASE)	7	I.B.3	D	22, 24
Rv0079	Rv0079	HYPOTHETICAL PROTEIN	16	VI	D	22, 24
Rv0092	ctpA	PROBABLE CATION TRANSPORTER P-TYPE ATPASE A CTPA	3	III.A.2	B	
Rv0093c	Rv0093c	PROBABLE CONSERVED MEMBRANE PROTEIN	3	VI	D	24
Rv0101	nrp	PROBABLE PEPTIDE SYNTHETASE NRP (PEPTIDE SYNTHASE)	1	I.I	B	24
Rv0108c	Rv0108c	HYPOTHETICAL PROTEIN	16	VI	D	42
Rv0111	Rv0111	POSSIBLE TRANSMEMBRANE ACYLTRANSFERASE	7	IV.H	A, B	24
Rv0118c	o37cA	PROBABLE OxALYL-CoA DECARBOxYLASE OxCA	7	I.A.1	D	24
Rv0120c	fusA2	PROBABLE ELONGATION FACTOR G FUSA2 (EF-G)	2	II.A.6	B	42, 22, 24
Rv0125	pepA	PROBABLE SERINE PROTEASE PEPA (SERINE PROTEINASE) (MTB32A)	7	II.B.3	A	42, 16
Rv0129c	fbpC	SECRETED ANTIGEN 85-C FBPC (85C) (AG58C) (MYCOLYL TRANSFERASE 85C)	1	I.H.3	A, B, D	42, 16
Rv0139	Rv0139	POSSIBLE OxIDOREDUCTASE	7	I.B.7	A	24
Rv0142	Rv0142	CONSERVED HYPOTHETICAL PROTEIN	10	VI	A, B	
Rv0147	Rv0147	PROBABLE ALDEHYDE DEHYDROGENASE (NAD+) DEPENDENT	7	I.B.7	D	22, 24
Rv0148	Rv0148	PROBABLE SHORT-CHAIN TYPE DEHYDROGENASE/REDUCTASE	7	I.B.7	A, B, D	42, 22
Rv0169	mce1A	MCE-FAMILY PROTEIN MCE1A	0	IV.A	D	24
Rv0172	mce1D	MCE-FAMILY PROTEIN MCE1D	0	IV.A	D	16
Rv0173	lprK	POSSIBLE MCE-FAMILY LIPOPROTEIN LPRK (MCE-FAMILY LIPOPROTEIN MCE1E)	3	II.C.1	D	16
Rv0174	mce1F	MCE-FAMILY PROTEIN MCE1F	0	IV.A	D	16, 24
Rv0183	Rv0183	POSSIBLE LYSOPHOSPHOLIPASE	7	I.B.7	D	42, 22, 24
Rv0194	Rv0194	PROBABLE DRUGS-TRANSPORT TRANSMEMBRANE ATP-BINDING PROTEIN ABC TRANSPORTER	3	II.C.5	C	24
Rv0202c	mmpL11	PROBABLE CONSERVED TRANSMEMBRANE TRANSPORT PROTEIN MMPL11	3	II.C.4	A, B, C	24
Rv0206c	mmpL3	POSSIBLE CONSERVED TRANSMEMBRANE TRANSPORT PROTEIN MMPL3	3	II.C.4	D	24
Rv0211	pckA	PROBABLE IRON-REGULATED PHOSPHOENOLPYRUVATE CARBOxYKINASE [GTP]	7	I.C.2	B, C	42, 22, 16, 24
Rv0227c	Rv0227c	PROBABLE CONSERVED MEMBRANE PROTEIN	3	II.C.5	D	22, 24
Rv0237	lpqI	PROBABLE CONSERVED LIPOPROTEIN LPQI	3	II.C.1	D	16, 24
Rv0242c	fabG4	PROBABLE 3-OxOACYL-[ACYL-CARRIER PROTEIN] REDUCTASE FABG4	1	I.H.1	D	42, 22, 16, 24
Rv0243	fadA2	PROBABLE ACETYL-CoA ACYLTRANSFERASE FADA2 (3-KETOACYL-CoA THIOLASE)	1	I.A.3	D	42, 22, 24
Rv0247c	Rv0247c	PROBABLE SUCCINATE DEHYDROGENASE [IRON-SULFUR SUBUNIT] (SUCCINIC DEHYDROGENASE)	7	I.B.7	D	22
Rv0248c	Rv0248c	PROBABLE SUCCINATE DEHYDROGENASE [IRON-SULFUR SUBUNIT] (SUCCINIC DEHYDROGENASE)	7	I.B.7	B	22, 24
Rv0251c	hsp	HEAT SHOCK PROTEIN HSP (HEAT-STRESS-INDUCED RIBOSOME-BINDING PROTEIN A)	0	III.B	D	22, 24
Rv0252	nirB	PROBABLE NITRITE REDUCTASE [NAD(P)H] LARGE SUBUNIT [FAD FLAVOPROTEIN] NIRB	7	I.B.6.b	D	24
Rv0270	fadD2	PROBABLE FATTY-ACID-CoA LIGASE FADD2 (FATTY-ACID-CoA SYNTHETASE)	1	I.A.3	D	42, 22, 24
Rv0271c	fadE6	PROBABLE ACYL-CoA DEHYDROGENASE FADE6	1	I.A.3	A	22, 24
Rv0282	Rv0282	CONSERVED HYPOTHETICAL PROTEIN	10	V	B, D	22, 24
Rv0283	Rv0283	POSSIBLE CONSERVED MEMBRANE PROTEIN	3	V	D	22, 16, 24
Rv0284	Rv0284	POSSIBLE CONSERVED MEMBRANE PROTEIN	3	V	B, D	22, 24
Rv0290	Rv0290	PROBABLE CONSERVED TRANSMEMBRANE PROTEIN	3	II.C.5	D	22, 24
Rv0291	mycP3	PROBABLE MEMBRANE-ANCHORED MYCOSIN MYCP3 (SERINE PROTEASE)	7	II.C.5	D	42, 22, 16
Rv0292	Rv0292	PROBABLE CONSERVED TRANSMEMBRANE PROTEIN	3	II.C.5	D	22
Rv0309	Rv0309	POSSIBLE CONSERVED ExPORTED PROTEIN	3	VI	D	16
Rv0327c	cyp135A1	POSSIBLE CYTOCHROME P450 135A1 CYP135A1	7	IV.F	B	
Rv0334	rmlA	ALPHA-D-GLUCOSE-1-PHOSPHATE THYMIDYLYLTRANSFERASE RMLA (DTDP-GLUCOSE SYNTHASE)	7	I.C.3	A	22
Rv0338c	Rv0338c	PROBABLE IRON-SULFUR-BINDING REDUCTASE	7	V	A	22, 24
Rv0350	dnaK	PROBABLE CHAPERONE PROTEIN DNAK (HEAT SHOCK PROTEIN 70)	0	III.B	A, B, D	42, 22, 16, 24
Rv0351	grpE	PROBABLE GRPE PROTEIN (HSP-70 COFACTOR)	0	III.B	D	42
Rv0352	dnaJ1	PROBABLE CHAPERONE PROTEIN DNAJ1	0	III.B	D	24
Rv0361	Rv0361	PROBABLE CONSERVED MEMBRANE PROTEIN	3	II.C.5	D	
Rv0363c	fba	PROBABLE FRUCTOSE-BISPHOSPHATE ALDOLASE FBA	7	I.B.1	D	42, 22, 16
Rv0379	secE2	POSSIBLE PROTEIN TRANSPORT PROTEIN SECE2	3	III.D	D	42, 22
Rv0384c	clpB	PROBABLE ENDOPEPTIDASE ATP BINDING PROTEIN (CHAIN B) (HEAT SHOCK PROTEIN F84.1)	0	III.B	D	42, 22, 24
Rv0386	Rv0386	PROBABLE TRANSCRIPTIONAL REGULATORY PROTEIN (PROBABLY LUxR/UHPA-FAMILY)	9	I.J.1	A, D	24
Rv0402c	mmpL1	PROBABLE CONSERVED TRANSMEMBRANE TRANSPORT PROTEIN MMPL1	3	II.C.4	A	16, 24
Rv0407	fgd1	PROBABLE F420-DEPENDENT GLUCOSE-6-PHOSPHATE DEHYDROGENASE FGD1	7	IV.G	D	22
Rv0431	Rv0431	PUTATIVE TUBERCULIN RELATED PEPTIDE	3	II.C.2	D	
Rv0432	sodC	PROBABLE PERIPLASMIC SUPEROxIDE DISMUTASE [CU-ZN] SODC	0	III.F	C, D	22, 24
Rv0440	groEL2	60 KDA CHAPERONIN 2 GROEL2 (PROTEIN CPN60-2) (GROEL PROTEIN 2) (65 KDA ANTIGEN)	0	III.B	A,B,C, D	42, 22, 24
Rv0444c	Rv0444c	CONSERVED HYPOTHETICAL PROTEIN	10	V	D	22
Rv0445c	sigK	PROBABLE ALTERNATIVE RNA POLYMERASE SIGMA FACTOR SIGK	2	II.A.7	D	22, 24
Rv0450c	mmpL4	PROBABLE CONSERVED TRANSMEMBRANE TRANSPORT PROTEIN MMPL4	3	II.C.4	A	24
Rv0468	fadB2	PROBABLE 3-HYDROxYBUTYRYL-CoA DEHYDROGENASE FADB2	1	I.A.3	D	42, 22, 24
Rv0474	Rv0474	PROBABLE TRANSCRIPTIONAL REGULATORY PROTEIN	9	I.J.1	D	24
Rv0480c	Rv0480c	POSSIBLE AMIDOHYDROLASE	7	V	A	24
Rv0497	Rv0497	PROBABLE CONSERVED TRANSMEMBRANE PROTEIN	3	II.C.5	D	
Rv0504c	Rv0504c	CONSERVED HYPOTHETICAL PROTEIN	10	V	D	22
Rv0507	mmpL2	PROBABLE CONSERVED TRANSMEMBRANE TRANSPORT PROTEIN MMPL2	3	II.C.4	B	24
Rv0512	hemB	PROBABLE DELTA-AMINOLEVULINIC ACID DEHYDRATASE HEMB (PORPHOBILINOGEN SYNTHASE)	7	I.G.12	C	
Rv0540	Rv0540	CONSERVED HYPOTHETICAL PROTEIN	10	V	B	22
Rv0559c	Rv0559c	POSSIBLE CONSERVED SECRETED PROTEIN	3	II.C.5	D	42, 16
Rv0562	grcC1	PROBABLE POLYPRENYL-DIPHOSPHATE SYNTHASE GRCC1	7	I.G.11	D	22
Rv0569	Rv0569	CONSERVED HYPOTHETICAL PROTEIN	10	V	D	42, 24
Rv0570	nrdZ	PROBABLE RIBONUCLEOSIDE-DIPHOSPHATE REDUCTASE (LARGE SUBUNIT)	2	I.F.3	D	22, 24
Rv0580c	Rv0580c	CONSERVED HYPOTHETICAL PROTEIN	10	VI	D	42, 24
Rv0583c	lpqN	PROBABLE CONSERVED LIPOPROTEIN LPQN	3	II.C.1	D	22, 16
Rv0614	Rv0614	CONSERVED HYPOTHETICAL PROTEIN	10	V	A	
Rv0622	Rv0622	POSSIBLE MEMBRANE PROTEIN	3	II.C.5	B	
Rv0631c	recC	PROBABLE ExONUCLEASE V (GAMMA CHAIN) RECC (ExODEOxYRIBONUCLEASE V GAMMA CHAIN)	2	II.A.5	A	16, 24

Rv0636	Rv0636	CONSERVED HYPOTHETICAL PROTEIN	10	VI	D	42, 24
Rv0638	secE1	PROBABLE PREPROTEIN TRANSLOCASE SECE1	3	III.D	D	
Rv0639	nusG	PROBABLE TRANSCRIPTION ANTITERMINATION PROTEIN NUSG	2	II.A.7	D	42, 22
Rv0641	rplA	PROBABLE 50S RIBOSOMAL PROTEIN L1 RPLA	2	II.A.1	D	42, 22, 24
Rv0644c	mmaA2	METHOxY MYCOLIC ACID SYNTHASE 2 MMAA2 (METHYL MYCOLIC ACID SYNTHASE 2)	1	I.H.2	D	22
Rv0651	rplJ	PROBABLE 50S RIBOSOMAL PROTEIN L10 RPLJ	2	II.A.1	D	42, 22, 24
Rv0652	rplL	PROBABLE 50S RIBOSOMAL PROTEIN L7/L12 RPLL (SA1)	2	II.A.1	D	42, 22, 24
Rv0667	rpoB	DNA-DIRECTED RNA POLYMERASE (BETA CHAIN) RPOB (TRANSCRIPTASE BETA CHAIN)	2	II.A.7	D	22, 24
Rv0668	rpoC	DNA-DIRECTED RNA POLYMERASE (BETA' CHAIN) RPOC (TRANSCRIPTASE BETA' CHAIN)	2	II.A.7	D	22, 24
Rv0673	echA4	POSSIBLE ENOYL-CoA HYDRATASE ECHA4 (ENOYL HYDRASE)	1	I.A.3	B	24
Rv0683	rpsG	PROBABLE 30S RIBOSOMAL PROTEIN S7 RPSG	2	II.A.1	D	22, 24
Rv0684	fusA1	PROBABLE ELONGATION FACTOR G FUSA1 (EF-G)	2	II.A.6	B, D	42
Rv0685	tuf	PROBABLE IRON-REGULATED ELONGATION FACTOR TU TUF (EF-TU)	2	II.A.6	A, B, C, D	42, 22, 24
Rv0686	Rv0686	PROBABLE MEMBRANE PROTEIN	3	II.C.5	D	22
Rv0700	rpsJ	30S RIBOSOMAL PROTEIN S10 RPSJ (TRANSCRIPTION ANTITERMINATION FACTOR NUSE)	2	II.A.1	D	42, 22, 24
Rv0707	rpsC	PROBABLE 30S RIBOSOMAL PROTEIN S3 RPSC	2	II.A.1	D	22, 24
Rv0712	Rv0712	CONSERVED HYPOTHETICAL PROTEIN	10	V	B	
Rv0724	sppA	POSSIBLE PROTEASE IV SPPA (ENDOPEPTIDASE IV) (SIGNAL PEPTIDE PEPTIDASE)	3	II.B.3	D	42, 22, 24
Rv0730	Rv0730	CONSERVED HYPOTHETICAL PROTEIN	10	V	D	42, 22
Rv0733	adk	PROBABLE ADENYLATE KINASE ADK (ATP-AMP TRANSPHOSPHORYLASE)	7	I.F.5	D	42, 22
Rv0771	Rv0771	POSSIBLE 4-CARBOxYMUCONOLACTONE DECARBOxYLASE (CMD)	7	II.B.6	D	
Rv0798c	cfp29	29 KDa ANTIGEN CFP29	0	IV.E	A	22, 24
Rv0814c	sseC2	CONSERVED HYPOTHETICAL PROTEIN SSEC2	7	I.C.5	D	42, 22
Rv0815c	cysA2	PROBABLE THIOSULFATE SULFURTRANSFERASE CYSA2 (RHODANESE-LIKE PROTEIN)	7	I.D.3	A, B, D	42, 22, 24
Rv0824c	desA1	PROBABLE ACYL-[ACYL-CARRIER PROTEIN] DESATURASE DESA1 (ACYL-[ACP] DESATURASE)	1	I.H.2	D	42, 22, 24
Rv0830	Rv0830	CONSERVED HYPOTHETICAL PROTEIN	10	V	A	22
Rv0831c	Rv0831c	CONSERVED HYPOTHETICAL PROTEIN	10	V	A, B, D	22, 24
Rv0854	Rv0854	CONSERVED HYPOTHETICAL PROTEIN	10	V	D	42
Rv0858c	Rv0858c	PROBABLE AMINOTRANSFERASE	7	IV.H	A	
Rv0859	fadA	POSSIBLE ACYL-CoA THIOLASE FADA	1	I.A.3	B, D	42, 22
Rv0860	fadB	PROBABLE FATTY OxIDATION PROTEIN FADB	1	I.A.3	D	42, 22, 24
Rv0871	cspB	PROBABLE COLD SHOCK-LIKE PROTEIN B CSPB	0	III.E	D	42, 22
Rv0873	fadE10	PROBABLE ACYL-CoA DEHYDROGENASE FADE10	1	I.A.3	D	22, 24
Rv0892	Rv0892	PROBABLE MONOOxYGENASE	7	I.B.7	B	
Rv0899	ompA	OUTER MEMBRANE PROTEIN A OMPA	3	II.C.2	D	
Rv0901	Rv0901	POSSIBLE CONSERVED ExPORTED OR MEMBRANE PROTEIN	3	V	D	22
Rv0902c	prrB	TWO COMPONENT SENSOR HISTIDINE KINASE PRRB	9	I.J.2	D	
Rv0905	echA6	POSSIBLE ENOYL-CoA HYDRATASE ECHA6 (ENOYL HYDRASE)	1	I.A.3	A, D	42, 22, 24
Rv0926c	Rv0926c	CONSERVED HYPOTHETICAL PROTEIN	10	V	A	
Rv0928	pstS3	PERIPLASMIC PHOSPHATE-BINDING LIPOPROTEIN PSTS3 (PBP-3) (PSTS3) (PHOS1)	3	III.A.4	D	16
Rv0931c	pknD	TRANSMEMBRANE SERINE/THREONINE-PROTEIN KINASE D PKND	9	I.J.3	D	22, 24
Rv0934	pstS1	PERIPLASMIC PHOSPHATE-BINDING LIPOPROTEIN PSTS1 (PBP-1) (PSTS1)	3	III.A.4	B, C, D	42, 22, 16, 24
Rv0944	Rv0944	POSSIBLE FORMAMIDOPYRIMIDINE-DNA GLYCOSYLASE (FAPY-DNA GLYCOSYLASE)	2	II.A.5	A	24
Rv0951	sucC	PROBABLE SUCCINYL-CoA SYNTHETASE (BETA CHAIN) SUCC (SCS-BETA)	7	I.B.3	D	24
Rv0952	sucD	PROBABLE SUCCINYL-CoA SYNTHETASE (ALPHA CHAIN) SUCD (SCS-ALPHA)	7	I.B.3	A, B, D	42, 22, 24
Rv0954	Rv0954	PROBABLE CONSERVED TRANSMEMBRANE PROTEIN	3	II.C.2	D	
Rv0968	Rv0968	CONSERVED HYPOTHETICAL PROTEIN	10	V	D	22
Rv0969	ctpV	PROBABLE METAL CATION TRANSPORTER P-TYPE ATPASE CTPV	3	III.A.2	A, D	22
Rv0981	mprA	MYCOBACTERIAL PERSISTENCE REGULATOR MRPA	9	I.J.2	D	22
Rv1005c	pabB	Probable para-aminobenzoate synthase component I PABD	7	I.G.2	A	24
Rv1006	Rv1006	HYPOTHETICAL PROTEIN	16	VI	B, D	22, 24
Rv1016c	lpqT	PROBABLE CONSERVED LIPOPROTEIN LPQT	3	II.C.1	D	
Rv1017c	prsA	PROBABLE RIBOSE-PHOSPHATE PYROPHOSPHOKINASE PRSA	7	I.F.1	D	42, 22, 24
Rv1023	eno	PROBABLE ENOLASE ENO	7	I.B.1	D	22, 24
Rv1029	kdpA	Probable Potassium-transporting ATPase A chain KDPA	3	III.A.2	A	
Rv1038c	es37J	ESAT-6 LIKE PROTEIN ESxJ (ESAT-6 LIKE PROTEIN 2)	3	V	A, B, D	42, 22, 16, 24
Rv1070c	echA8	PROBABLE ENOYL-CoA HYDRATASE ECHA8 (ENOYL HYDRASE)	1	I.A.3	D	22
Rv1074c	fadA3	PROBABLE BETA-KETOACYL CoA THIOLASE FADA3	1	I.A.3	D	42, 22, 16
Rv1075c	Rv1075c	CONSERVED ExPORTED PROTEIN	3	V	B	16, 24
Rv1077	cbs	Probable cystathionine beta-synthase CBS (Serine sulfhydrase) (Beta-thionase)	7	I.D.3	D	22, 16
Rv1078	pra	Probable Proline-rich antigen homolog pra	10	V	A	22
Rv1080c	greA	PROBABLE TRANSCRIPTION ELONGATION FACTOR GREA	2	II.A.6	D	42, 22, 24
Rv1093	glyA1	Probable Serine hydroxymethyltransferase 1 glyA1	7	I.D.3	D	42, 22, 24
Rv1094	desA2	POSSIBLE ACYL-[ACYL-CARRIER PROTEIN] DESATURASE DESA2	1	I.H.2	B, D	22, 24
Rv1096	Rv1096	POSSIBLE GLYCOSYL HYDROLASE	7	I.A.1	D	
Rv1097c	Rv1097c	PROBABLE MEMBRANE GLYCINE AND PROLINE RICH PROTEIN	3	II.C.5	D	24
Rv1098c	fum	PROBABLE FUMARASE FUM (Fumarate hydratase)	7	I.B.3	D	42, 22, 16, 24
Rv1100	Rv1100	CONSERVED HYPOTHETICAL PROTEIN	10	V	D	
Rv1109c	Rv1109c	CONSERVED HYPOTHETICAL PROTEIN	10	VI	D	42, 22, 24
Rv1128c	Rv1128c	CONSERVED HYPOTHETICAL PROTEIN	5	IV.B.2	B	
Rv1130	Rv1130	CONSERVED HYPOTHETICAL PROTEIN	10	V	A	24
Rv1133c	metE	HOMOCYSTEINE METHYLTRANSFERASE METE	7	I.D.2	D	22, 24
Rv1144	Rv1144	PROBABLE SHORT-CHAIN TYPE DEHYDROGENASE/REDUCTASE	7	I.B.7	D	22
Rv1151c	Rv1151c	PROBABLE TRANSCRIPTIONAL REGULATORY PROTEIN	9	I.J.1	A	
Rv1161	narG	PROBABLE RESPIRATORY NITRATE REDUCTASE (ALPHA CHAIN) NARG	7	I.B.6.b	B, C	16, 24
Rv1162	narH	PROBABLE RESPIRATORY NITRATE REDUCTASE (BETA CHAIN) NARH	7	I.B.6.b	A	24
Rv1165	typA	POSSIBLE GTP-BINDING TRANSLATION ELONGATION FACTOR	2	V	B	24
Rv1175c	fadH	PROBABLE NADPH DEPENDENT 2,4-DIENOYL-COA REDUCTASE FADH	1	I.A.3	A	24
Rv1179c	Rv1179c	HYPOTHETICAL PROTEIN	16	VI	B	22, 24
Rv1181	pks4	PROBABLE POLYKETIDE BETA-KETOACYL SYNTHASE PKS4	1	I.I	A, D	22, 24
Rv1184c	Rv1184c	POSSIBLE ExPORTED PROTEIN	3	V	D	
Rv1196	PPE18	PPE FAMILY PROTEIN	6	IV.C.2	D	
Rv1198	es37L	PUTATIVE ESAT-6 LIKE PROTEIN ESxL (ESAT-6 LIKE PROTEIN 4)	3	V	D	42, 22, 16, 24
Rv1201c	Rv1201c	PROBABLE TRANSFERASE	7	V	B	42
Rv1209	Rv1209	CONSERVED HYPOTHETICAL PROTEIN	10	V	D	
Rv1220c	Rv1220c	PROBABLE METHYLTRANSFERASE	7	IV.H	D	22
Rv1223	htrA	PROBABLE SERINE PROTEASE HTRA (DEGP PROTEIN)	7	II.B.3	D	24
Rv1230c	Rv1230c	POSSIBLE MEMBRANE PROTEIN	3	II.C.5	B	
Rv1231c	Rv1231c	PROBABLE MEMBRANE PROTEIN	3	VI	D	24
Rv1240	mdh	PROBABLE MALATE DEHYDROGENASE MDH	7	I.B.3	D	42, 22, 24
Rv1244	lpqZ	PROBABLE LIPOPROTEIN LPQZ	3	II.C.1	D	24
Rv1245c	Rv1245c	PROBABLE SHORT-CHAIN TYPE DEHYDROGENASE/REDUCTASE	7	I.B.7	A	42, 22, 24
Rv1248c	sucA	PROBABLE 2-OxOGLUTARATE DEHYDROGENASE SUCA (Alpha-ketoglutarate dehydrogenase)	7	I.B.3	B, C	24
Rv1253	deaD	PROBABLE COLD-SHOCK DEAD-BOx PROTEIN A HOMOLOG DEAD	2	II.A.7	A	24
Rv1256c	cyp130	PROBA BLE CYTOCHROME P450 130 CYP130	7	IV.F	B	

Rv1261c	Rv1261c	CONSERVED HYPOTHETICAL PROTEIN	10	V	D	22
Rv1265	Rv1265	HYPOTHETICAL PROTEIN	16	VI	D	22
Rv1266c	pknH	PROBABLE TRANSMEMBRANE SERINE/THREONINE-PROTEIN KINASE H PKNH (STPK H)	9	I.J.3	B, D	
Rv1267c	embR	PROBABLE TRANSCRIPTIONAL REGULATORY PROTEIN EMBR	9	I.J.1	D	24
Rv1269c	Rv1269c	CONSERVED PROBABLE SECRETED PROTEIN	3	V	D	42, 16
Rv1270c	lprA	POSSIBLE LIPOPROTEIN LPRA	3	II.C.1	C, D	22, 16, 24
Rv1273c	Rv1273c	PROBABLE DRUGS-TRANSPORT TRANSMEMBRANE ATP-BINDING PROTEIN ABC TRANSPORTER	3	II.C.5	B	
Rv1275	lprC	POSSIBLE LIPOPROTEIN LPRC	3	II.C.1	D	22, 24
Rv1278	Rv1278	HYPOTHETICAL PROTEIN	16	VI	B	24
Rv1279	Rv1279	PROBABLE DEHYDROGENASE FAD flavoprotein GMC oxidoreductase	7	I.B.7	B	22, 24
Rv1281c	oppD	PROBABLE OLIGOPEPTIDE-TRANSPORT ATP-BINDING PROTEIN ABC TRANSPORTER OPPD	3	III.A.1	D	24
Rv1294	thrA	PROBABLE HOMOSERINE DEHYDROGENASE THRA	7	I.D.2	A	22
Rv1297	rho	PROBABLE TRANSCRIPTION TERMINATION FACTOR RHO HOMOLOG	2	II.A.7	D	42, 22, 24
Rv1306	atpF	PROBABLE ATP SYNTHASE B CHAIN ATPF	7	I.B.8	B, C, D	22
Rv1307	atpH	PROBABLE ATP SYNTHASE DELTA CHAIN ATPH	7	I.B.8	D	22, 24
Rv1308	atpA	PROBABLE ATP SYNTHASE ALPHA CHAIN ATPA	7	I.B.8	A, B, D	42, 22, 24
Rv1309	atpG	PROBABLE ATP SYNTHASE GAMMA CHAIN ATPG	7	I.B.8	A, B, D	22, 24
Rv1310	atpD	PROBABLE ATP SYNTHASE BETA CHAIN ATPD	7	I.B.8	B, D	22, 16, 24
Rv1311	atpC	PROBABLE ATP SYNTHASE EPSILON CHAIN ATPC	7	I.B.8	D	22, 24
Rv1324	Rv1324	POSSIBLE THIOREDOxIN	7	VI	D	42, 22, 24
Rv1330c	Rv1330c	CONSERVED HYPOTHETICAL PROTEIN	10	V	B	
Rv1361c	PPE19	PPE FAMILY PROTEIN	6	IV.C.2	A	
Rv1368	lprF	PROBABLE CONSERVED LIPOPROTEIN LPRF	3	II.C.1	C, D	22, 24
Rv1379	pyrR	PROBABLE PYRIMIDINE OPERON REGULATORY PROTEIN PYRR	9	I.J.1	D	
Rv1388	mihF	PUTATIVE INTEGRATION HOST FACTOR MIHF	2	II.A.4	D	22, 24
Rv1390	rpoZ	PROBABLE DNA-DIRECTED RNA POLYMERASE (OMEGA CHAIN) RPOZ	2	VI	D	42
Rv1392	metK	PROBABLE S-ADENOSYLMETHIONINE SYNTHETASE METK (MAT) (AdoMet synthetase)	7	I.D.2	D	42, 22, 24
Rv1393c	Rv1393c	PROBABLE MONOxYGENASE	7	I.B.7	C	
Rv1407	fmu	PROBABLE FMU PROTEIN (SUN PROTEIN)	2	II.A.4	B	24
Rv1411c	lprG	PROBABLE CONSERVED LIPOPROTEIN LPRG	3	II.C.1	B, D	22, 24
Rv1423	whiA	PROBABLE TRANSCRIPTIONAL REGULATORY PROTEIN WHIA	9	I.J.1	D	22, 24
Rv1424c	Rv1424c	POSSIBLE MEMBRANE PROTEIN	3	II.C.5	A	24
Rv1436	gap	PROBABLE GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE GAP (GAPDH)	7	I.B.1	D	42, 22, 16, 24
Rv1442	bisC	PROBABLE BIOTIN SULFOxIDE REDUCTASE BISC (BDS reductase) (BSO reductase)	7	I.G.1	A, D	
Rv1449c	tkt	PROBABLE TRANSKETOLASE TKT (TK)	7	I.B.5	B	22, 16, 24
Rv1450c	PE_PGRS27	PE-PGRS FAMILY PROTEIN	6	IV.C.1.b	A	
Rv1451	ctaB	PROBABLE CYTOCHROME C OxIDASE ASSEMBLY FACTOR CTAB	7	I.B.6.a	B	
Rv1466	Rv1466	CONSERVED HYPOTHETICAL PROTEIN	10	V	D	
Rv1473	Rv1473	PROBABLE MACROLIDE-TRANSPORT ATP-BINDING PROTEIN ABC TRANSPORTER	3	II.C.5	D	24
Rv1479	mo37R1	PROBABLE TRANSCRIPTIONAL REGULATORY PROTEIN MOxR1	9	I.J.1	A, B, D	42, 22, 24
Rv1484	inhA	NADH-DEPENDENT ENOYL-[ACYL-CARRIER-PROTEIN] REDUCTASE INHA	1	I.H.1	D	42, 22
Rv1488	Rv1488	POSSIBLE ExPORTED CONSERVED PROTEIN	3	V	D	22, 16, 24
Rv1492	mutA	PROBABLE METHYLMALONYL-CoA MUTASE SMALL SUBUNIT MUTA (MCM)	1	I.A.3	B	24
Rv1520	Rv1520	probable sugar transferase	7	IV.H	C	24
Rv1522c	mmpL12	PROBABLE CONSERVED TRANSMEMBRANE TRANSPORT PROTEIN MMPL12	3	II.C.4	B	24
Rv1527c	pks5	Probable polyketide synthase pks5	1	I.I	B	24
Rv1543	Rv1543	POSSIBLE FATTY ACYL-CoA REDUCTASE	1	I.H.3	D	42, 22, 24
Rv1544	Rv1544	Possible ketoacyl reductase	1	I.I	D	42, 22
Rv1547	dnaE1	PROBABLE DNA POLYMERASE III (ALPHA CHAIN) DNAE1 (DNA NUCLEOTIDYLTRANSFERASE)	2	II.A.5	A	22
Rv1551	plsB1	Possible acyltransferase plsB1	1	I.H.3	A	24
Rv1558	Rv1558	CONSERVED HYPOTHETICAL PROTEIN	10	V	D	42, 22
Rv1569	bioF1	PROBABLE 8-AMINO-7-OxONONANOATE SYNTHASE BIOF1 (AONS)	7	I.G.1	A, B	24
Rv1586c	Rv1586c	Probable phiRv1 integrase	5	IV.B.3	D	
Rv1589	bioB	PROBABLE BIOTIN SYNTHETASE BIOB	7	I.G.1	C	24
Rv1593c	Rv1593c	CONSERVED HYPOTHETICAL PROTEIN	10	V	A	22
Rv1594	nadA	Probable quinolinate synthetase nadA	7	I.G.7	B	
Rv1596	nadC	Probable nicotinate-nucleotide pyrophosphatase nadC	7	I.G.7	A, B	42, 22
Rv1599	hisD	Probable histidinol dehydrogenase HisD (HDH)	7	I.D.5	D	
Rv1602	hisH	Probable amidotransferase hisH	7	I.D.5	B	
Rv1611	trpC	Probable indole-3-glycerol phosphate synthase trpC	7	I.D.4	D	24
Rv1623c	cydA	Probable integral membrane cytochrome D ubiquinol oxidase (subunit I) cydA	7	I.B.6.c	D	
Rv1626	Rv1626	Probable two-component system transcriptional regulator	9	I.J.2	D	42, 22, 24
Rv1627c	Rv1627c	Probable nonspecific lipid-transfer protein	1	I.H.3	A	22
Rv1629	polA	PROBABLE DNA POLYMERASE I POLA	2	II.A.5	D	22, 24
Rv1630	rpsA	PROBABLE RIBOSOMAL PROTEIN S1 RPSA	2	II.A.1	A, B, D	22, 24
Rv1636	TB15.3	IRON-REGULATED CONSERVED HYPOTHETICAL PROTEIN TB15.3	10	V	D	42, 22
Rv1640c	lys37	Possible lysyl-tRNA synthetase 2 lysx	2	II.A.3	D	24
Rv1654	argB	Probable Acetylglutamate kinase argB	7	I.D.1	D	22
Rv1659	argH	Probable Argininosuccinate lyase argH	7	I.D.1	D	
Rv1661	pks7	Probable polyketide synthase pks7	1	I.I	D	24
Rv1666c	cyp139	Probable cytochrome P450 139 CYP139	7	IV.F	A	
Rv1689	tyrS	Probable Tyrosyl-tRNA synthase tyrS (TYRRS)	2	II.A.3	B	24
Rv1703c	Rv1703c	Probable catechol-o-methyltransferase	7	IV.H	D	42, 22, 24
Rv1747	Rv1747	PROBABLE CONSERVED TRANSMEMBRANE ATP-BINDING PROTEIN ABC TRANSPORTER	3	II.C.5	D	24
Rv1748	Rv1748	HYPOTHETICAL PROTEIN	16	VI	D	
Rv1771	Rv1771	PROBABLE OxIDOREDUCTASE	7	I.B.7	D	22, 24
Rv1777	cyp144	Probable cytochrome p450 144 CYP144	7	IV.F	A	
Rv1782	Rv1782	PROBABLE CONSERVED MEMBRANE PROTEIN	3	V	D	22
Rv1784	Rv1784	CONSERVED HYPOTHETICAL PROTEIN	10	V	D	16, 24
Rv1794	Rv1794	CONSERVED HYPOTHETICAL PROTEIN	10	V	B, D	42, 22
Rv1801	PPE29	PPE FAMILY PROTEIN	6	IV.C.2	A	
Rv1815	Rv1815	CONSERVED HYPOTHETICAL PROTEIN	10	V	B	16
Rv1821	secA2	POSSIBLE PREPROTEIN TRANSLOCASE ATPase SECA2	3	III.D	D	22, 24
Rv1827	cfp17	CONSERVED HYPOTHETICAL PROTEIN CFP17	10	V	D	42, 22, 16
Rv1829	Rv1829	CONSERVED HYPOTHETICAL PROTEIN	10	V	D	42, 22
Rv1836c	Rv1836c	CONSERVED HYPOTHETICAL PROTEIN	10	V	D	
Rv1837c	glcB	PROBABLE MALATE SYNTHASE G GLCB	7	I.B.4	A, B, D	42, 22, 16, 24
Rv1866	Rv1866	CONSERVED HYPOTHETICAL PROTEIN	1	V	B	24
Rv1871c	Rv1871c	CONSERVED HYPOTHETICAL PROTEIN	10	VI	D	22
Rv1872c	lldD2	POSSIBLE L-LACTATE DEHYDROGENASE (CYTOCHROME) LLDD2	7	I.B.6.a	D	22, 24
Rv1875	Rv1875	CONSERVED HYPOTHETICAL PROTEIN	10	V	D	42, 24
Rv1886c	fbpB	SECRETED ANTIGEN 85-B FBPB (85B) (ANTIGEN 85 COMPLEx B) (MYCOLYL TRANSFERASE 85B)	1	I.H.3	A, B	42, 22, 16
Rv1908c	katG	CATALASE-PEROxIDASE-PEROxYNITRITASE T KATG	0	III.F	B, D	16, 24
Rv1911c	lppC	PROBABLE LIPOPROTEIN LPPC	3	II.C.1	A	16
Rv1919c	Rv1919c	CONSERVED HYPOTHETICAL PROTEIN	10	II.C.2	D	22

Rv1925	fadD31	PROBABLE ACYL-CoA LIGASE FADD31 (ACYL-COA SYNTHETASE) (ACYL-CoA SYNTHASE)	1	I.A.3	D	24
Rv1926c	mpt63	IMMUNOGENIC PROTEIN MPT63 (ANTIGEN MPT63/MPB63)	3	VI	D	42, 16
Rv1932	tp37	PROBABLE THIOL PEROxIDASE TPx	0	III.F	B	42, 16
Rv1937	Rv1937	POSSIBLE OxYGENASE	7	I.B.7	A, B	24
Rv1957	Rv1957	HYPOTHETICAL PROTEIN	16	VI	D	
Rv1969	mce3D	MCE-FAMILY PROTEIN MCE3D	0	IV.A	A	
Rv1980c	mpt64	IMMUNOGENIC PROTEIN MPT64 (ANTIGEN MPT64/MPB64)	3	II.C.2	D	42, 16
Rv1984c	cfp21	PROBABLE CUTINASE PRECURSOR CFP21	3	II.C.2	A	16
Rv1996	Rv1996	CONSERVED HYPOTHETICAL PROTEIN	10	V	D	42, 22, 24
Rv2005c	Rv2005c	CONSERVED HYPOTHETICAL PROTEIN	10	V	A, B, D	42, 22, 24
Rv2028c	Rv2028c	CONSERVED HYPOTHETICAL PROTEIN	10	V	D	22
Rv2030c	Rv2030c	CONSERVED HYPOTHETICAL PROTEIN	10	V	B, D	22, 24
Rv2031c	hspX	HEAT SHOCK PROTEIN HSPx (ALPHA-CRSTALLIN HOMOLOG) (14 kDa ANTIGEN) (HSP16.3)	0	III.B	A, B, C, D	42, 22, 24
Rv2032	acg	Conserved hypothetical protein Acg	10	V	D	22, 24
Rv2047c	Rv2047c	CONSERVED HYPOTHETICAL PROTEIN	10	V	D	
Rv2048c	pks12	Probable polyketide synthase pks12	1	I.I	B, C	24
Rv2050	Rv2050	CONSERVED HYPOTHETICAL PROTEIN	10	V	D	42
Rv2063c	Rv2063c	hypothetical protein	0	VI	B	
Rv2074	Rv2074	CONSERVED HYPOTHETICAL PROTEIN	10	VI	D	42, 16
Rv2091c	Rv2091c	Probable membrane protein	3	II.C.5	D	24
Rv2095c	Rv2095c	CONSERVED HYPOTHETICAL PROTEIN	10	V	D	
Rv2096c	Rv2096c	CONSERVED HYPOTHETICAL PROTEIN	10	V	D	24
Rv2100	Rv2100	CONSERVED HYPOTHETICAL PROTEIN	5	V	B	24
Rv2101	helZ	PROBABLE HELICASE HELZ	2	II.A.5	B	24
Rv2115c	Rv2115c	Probable ATPase	3	III.C	D	42, 22, 24
Rv2122c	hisE	Probable phosphoribosyl-AMP pyrophosphatase HisE	7	I.D.5	D	
Rv2129c	Rv2129c	Probable oxidoreductase	7	I.B.7	D	24
Rv2145c	wag31	CONSERVED HYPOTHETICAL PROTEIN WAG31	3	II.C.2	A, B, D	42, 22, 24
Rv2159c	Rv2159c	CONSERVED HYPOTHETICAL PROTEIN	10	VI	D	22, 24
Rv2171	lppM	Probable conserved lipoprotein lppM	3	II.C.1	D	
Rv2178c	aroG	Probable 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase AroG	7	I.D.4	B	42, 22, 24
Rv2185c	TB16.3	CONSERVED HYPOTHETICAL PROTEIN TB16.3	10	V	D	42, 22
Rv2187	fadD15	Probable long-chain-fatty-acid-CoA ligase fadD15 (FATTY-ACID-CoA SYNTHETASE)	1	I.A.3	D	22
Rv2191	Rv2191	CONSERVED HYPOTHETICAL PROTEIN	2	II.A.5	B	24
Rv2195	qcrA	Probable Rieske iron-sulfur protein QcrA	7	I.B.6.a	D	22, 24
Rv2198c	mmpS3	PROBABLE CONSERVED MEMBRANE PROTEIN MMPS3	3	II.C.4	D	
Rv2205c	Rv2205c	CONSERVED HYPOTHETICAL PROTEIN	10	V	B	
Rv2211c	gcvT	Probable aminomethyltransferase GcvT (Glycine cleavage system T protein)	7	I.C.1	D	
Rv2213	pepB	Probable aminopeptidase PepB	7	II.B.3	D	22
Rv2214c	ephD	Possible short-chain dehydrogenase EphD	0	III.F	D	22, 24
Rv2215	sucB	Probable pyruvate dehydrogenase (E2 component) SucB	7	I.B.3	B, D	42, 22, 24
Rv2216	Rv2216	CONSERVED HYPOTHETICAL PROTEIN	10	V	D	22
Rv2220	glnA1	GLUTAMINE SYNTHETASE GLNA1 (GLUTAMINE SYNTHASE) (GS-I)	7	I.D.1	D	42, 22, 16, 24
Rv2221c	glnE	GLUTAMATE-AMMONIA-LIGASE ADENYLYLTRANSFERASE	7	I.D.1	A	22
Rv2223c	Rv2223c	Probable exported protease	3	II.C.2	B	
Rv2241	aceE	Probable pyruvate dehydrogenase E1 component aceE (PYRUVATE DECARBOxYLASE)	7	I.B.2	D	22, 16, 24
Rv2244	acpM	MEROMYCOLATE ExTENSION ACYL CARRIER PROTEIN ACPM	1	I.H.1	D	22, 24
Rv2245	kasA	3-OxOACYL-[ACYL-CARRIER PROTEIN] SYNTHASE 1 KASA (BETA-KETOACYL-ACP SYNTHASE) (KAS I)	1	I.H.1	A, B, D	42, 22, 24
Rv2246	kasB	3-OxOACYL-[ACYL-CARRIER PROTEIN] SYNTHASE 2 KASB (BETA-KETOACYL-ACP SYNTHASE) (KAS I)	1	I.H.1	B, D	42, 22, 24
Rv2247	accD6	ACETYL/PROPIONYL-CoA CARBOxYLASE (BETA SUBUNIT) ACCD6	1	I.H.1	A, D	
Rv2258c	Rv2258c	Possible transcriptional regulatory protein	9	I.J.1	B	22
Rv2260	Rv2260	CONSERVED HYPOTHETICAL PROTEIN	10	V	D	
Rv2280	Rv2280	Probable dehydrogenase	7	I.B.7	A, D	22
Rv2296	Rv2296	Probable haloalkane dehalogenase	7	IV.I	A, D	42, 22, 24
Rv2298	Rv2298	CONSERVED HYPOTHETICAL PROTEIN	10	V	D	22, 24
Rv2299c	htpG	PROBABLE CHAPERONE PROTEIN HTPG (HEAT SHOCK PROTEIN) (HSP90 FAMILY PROTEIN)	0	III.B	D	22, 24
Rv2302	Rv2302	CONSERVED HYPOTHETICAL PROTEIN	10	V	D	42
Rv2307c	Rv2307c	CONSERVED HYPOTHETICAL PROTEIN	10	V	B	
Rv2314c	Rv2314c	CONSERVED HYPOTHETICAL PROTEIN	10	V	A	16
Rv2326c	Rv2326c	POSSIBLE TRANSMEMBRANE ATP-BINDING PROTEIN ABC TRANSORTER	3	II.C.5	A, B	22
Rv2345	Rv2345	POSSIBLE CONSERVED TRANSMEMBRANE PROTEIN	3	II.C.2	D	22, 24
Rv2373c	dnaJ2	PROBABLE CHAPERONE PROTEIN DNAJ2	0	III.B	D	19
Rv2376c	cfp2	LOW MOLECULAR WEIGHT ANTIGEN CFP2 (LOW MOLECULAR WEIGHT PROTEIN ANTIGEN 2)	3	V	A	16
Rv2383c	mbtB	PHENYLOxAZOLINE SYNTHASE MBTB (PHENYLOxAZOLINE SYNTHETASE)	1	I.I	A	24
Rv2391	nirA	PROBABLE FERREDOxIN-DEPENDENT NITRITE REDUCTASE NIRA	7	I.B.6.b	A	19
Rv2428	ahpC	ALKYL HYDROPEROxIDE REDUCTASE C PROTEIN AHPC (ALKYL HYDROPEROxIDASE C)	0	III.F	D	42, 22
Rv2431c	PE25	PE FAMILY PROTEIN	6	IV.C.1.a	D	16
Rv2441c	rpmA	PROBABLE 50S RIBOSOMAL PROTEIN L27 RPMA	2	II.A.1	D	24
Rv2444c	rne	POSSIBLE RIBONUCLEASE E RNE	2	II.B.1	D	22, 24
Rv2455c	Rv2455c	PROBABLE OxIDOREDUCTASE (ALPHA SUBUNIT)	7	I.B.7	A	22
Rv2460c	clpP2	PROBABLE ATP-DEPENDENT CLP PROTEASE PROTEOLYTIC SUBUNIT 2 CLPP2	7	II.B.3	B	22, 24
Rv2462c	tig	PROBABLE TRIGGER FACTOR (TF) PROTEIN TIG	3	III.D	D	42, 22, 24
Rv2466c	Rv2466c	CONSERVED HYPOTHETICAL PROTEIN	10	V	D	
Rv2471	aglA	PROBABLE ALPHA-GLUCOSIDASE AGLA (MALTASE) (GLUCOINVERTASE)	7	II.A.8	C	24
Rv2476c	gdh	PROBABLE NAD-DEPENDENT GLUTAMATE DEHYDROGENASE GDH (NAD-GDH)	7	VI	A	22, 24
Rv2477c	Rv2477c	PROBABLE MACROLIDE-TRANSPORT ATP-BINDING PROTEIN ABC TRANSPORTER	3	II.C.5	A, D	22, 24
Rv2488c	Rv2488c	PROBABLE TRANSCRIPTIONAL REGULATORY PROTEIN (LUxR-FAMILY)	9	I.J.1	A, B	24
Rv2495c	pdhC	PROBABLE DIHYDROLIPOAMIDE S-ACETYLTRANSFERASE E2 COMPONENT	7	I.B.2	B	24
Rv2496c	pdhB	PROBABLE PYRUVATE DEHYDROGENASE E1 COMPONENT (BETA SUBUNIT)	7	I.B.2	A	22, 24
Rv2509	Rv2509	PROBABLE SHORT-CHAIN TYPE DEHYDROGENASE/REDUCTASE	7	I.B.7	D	22, 24
Rv2518c	lppS	PROBABLE CONSERVED LIPOPROTEIN LPPS	3	II.C.1	D	
Rv2524c	fas	PROBABLE FATTY ACID SYNTHASE FAS (FATTY ACID SYNTHETASE)	1	I.H.1	B, C, D	22, 24
Rv2528c	mrr	PROBABLE RESTRICTION SYSTEM PROTEIN MRR	2	II.A.5	D	22
Rv2531c	Rv2531c	PROBABLE AMINO ACID DECARBOxYLASE	7	I.A.2	B	24
Rv2536	Rv2536	PROBABLE CONSERVED TRANSMEMBRANE PROTEIN	3	II.C.5	D	24
Rv2555c	alaS	PROBABLE ALANYL-TRNA SYNTHETASE ALAS (ALANINE--TRNA LIGASE)	2	II.A.3	A	22, 24
Rv2556c	Rv2556c	CONSERVED HYPOTHETICAL PROTEIN	10	V	D	
Rv2576c	Rv2576c	POSSIBLE CONSERVED MEMBRANE PROTEIN	3	VI	A	16
Rv2579	dhaA	POSSIBLE HALOALKANE DEHALOGENASE DHAA (1-CHLOROHExANE HALIDOHYDROLASE)	7	IV.I	A, B	42, 22
Rv2587c	secD	PROBABLE PROTEIN-ExPORT MEMBRANE PROTEIN SECD	3	III.D	D	24
Rv2622	Rv2622	POSSIBLE METHYLTRANSFERASE (METHYLASE)	7	IV.H	B	24
Rv2623	TB31.7	CONSERVED HYPOTHETICAL PROTEIN TB31.7	10	V	A, B, C, D	22, 24
Rv2625c	Rv2625c	PROBABLE CONSERVED TRANSMEMBRANE ALANINE AND LEUCINE RICH PROTEIN	3	V	B	22
Rv2626c	Rv2626c	CONSERVED HYPOTHETICAL PROTEIN	10	V	D	42, 22, 24
Rv2629	Rv2629	CONSERVED HYPOTHETICAL PROTEIN	10	VI	D	24

Rv2631	Rv2631	CONSERVED HYPOTHETICAL PROTEIN	10	V	B	16
Rv2691	ceoB	TRK SYSTEM POTASSIUM UPTAKE PROTEIN CEOB	3	III.A.2	D	22
Rv2703	sigA	RNA POLYMERASE SIGMA FACTOR SIGA (SIGMA-A)	2	II.A.7	A, D	22, 24
Rv2711	ideR	IRON-DEPENDENT REPRESSOR AND ACTIVATOR IDER	9	I.J.1	D	42, 22
Rv2712c	Rv2712c	HYPOTHETICAL PROTEIN	16	VI	A	
Rv2713	sthA	PROBABLE SOLUBLE PYRIDINE NUCLEOTIDE TRANSHYDROGENASE STHA	7	I.B.7	B	24
Rv2720	le37A	REPRESSOR LExA	9	I.J.1	D	
Rv2721c	Rv2721c	POSSIBLE CONSERVED TRANSMEMBRANE ALANINE AND GLYCINE RICH PROTEIN	3	V	A, D	16
Rv2744c	35kd_ag	CONSERVED 35 KDA ALANINE RICH PROTEIN	10	V	A, B, C, D	42, 22, 24
Rv2748c	ftsK	POSSIBLE CELL DIVISION TRANSMEMBRANE PROTEIN FTSK	3	III.C	A	24
Rv2756c	hsdM	POSSIBLE TYPE I RESTRICTION/MODIFICATION SYSTEM DNA METHYLASE	2	II.A.5	B	24
Rv2773c	dapB	DIHYDRODIPICOLINATE REDUCTASE DAPB (DHPR)	7	I.D.2	D	42, 22
Rv2780	ald	SECRETED L-ALANINE DEHYDROGENASE ALD (40 KDA ANTIGEN) (TB43)	7	I.A.2	D	42, 22, 24
Rv2783c	gpsI	BIFUNCTIONAL PROTEIN POLYRIBONUCLEOTIDE NUCLEOTIDYLTRANSFERASE	2	II.A.7	D	22, 24
Rv2789c	fadE21	PROBABLE ACYL-CoA DEHYDROGENASE FADE21	1	I.A.3	A	
Rv2839c	infB	PROBABLE TRANSLATION INITIATION FACTOR IF-2 INFB	2	II.A.6	D	22, 24
Rv2841c	nusA	PROBABLE N UTILIZATION SUBSTANCE PROTEIN A NUSA	2	II.A.7	D	22
Rv2864c	Rv2864c	POSSIBLE PENICILLIN-BINDING LIPOPROTEIN	3	II.C.3	B	
Rv2868c	gcpE	PROBABLE GCPE PROTEIN	10	V	D	22
Rv2880c	Rv2880c	CONSERVED HYPOTHETICAL PROTEIN	10	V	A	
Rv2881c	cdsA	PROBABLE INTEGRAL MEMBRANE PHOSPHATIDATE CYTIDYLYLTRANSFERASE	1	I.H.3	C	
Rv2882c	frr	RIBOSOME RECYCLING FACTOR FRR (RIBOSOME RELEASING FACTOR) (RRF)	2	II.A.6	D	42, 22
Rv2888c	amiC	PROBABLE AMIDASE AMIC (AMINOHYDROLASE)	7	IV.I	D	22
Rv2889c	tsf	PROBABLE ELONGATION FACTOR TSF (EF-TS)	2	II.A.6	A, B, D	42, 22, 24
Rv2890c	rpsB	PROBABLE 30S RIBOSOMAL PROTEIN S2 RPSB	2	II.A.1	D	22, 24
Rv2921c	ftsY	PROBABLE CELL DIVISION PROTEIN FTSY (SIGNAL RECOGNITION PARTICLE RECEPTOR)	3	III.C	D	22
Rv2931	ppsA	PHENOLPTHIOCEROL SYNTHESIS TYPE-I POLYKETIDE SYNTHASE PPSA	1	I.I	A	
Rv2935	ppsE	PHENOLPTHIOCEROL SYNTHESIS TYPE-I POLYKETIDE SYNTHASE PPSE	1	I.I	D	24
Rv2936	drrA	PROBABLE DAUNORUBICIN-DIM-TRANSPORT ATP-BINDING PROTEIN ABC TRANSPORTER DRRA	3	III.A.6	D	
Rv2938	drrC	PROBABLE DAUNORUBICIN-DIM-TRANSPORT ABC TRANSPORTER DRRC	3	III.A.6	A, B	
Rv2940c	mas	PROBABLE MULTIFUNCTIONAL MYCOCEROSIC ACID SYNTHASE MEMBRANE-ASSOCIATED MAS	1	I.I	A, D	22, 24
Rv2945c	lppX	PROBABLE CONSERVED LIPOPROTEIN LPPX	3	II.C.1	C, D	22, 16, 24
Rv2953	Rv2953	CONSERVED HYPOTHETICAL PROTEIN	10	V	D	42, 22, 24
Rv2955c	Rv2955c	CONSERVED HYPOTHETICAL PROTEIN	10	V	D	24
Rv2969c	Rv2969c	POSSIBLE CONSERVED MEMBRANE OR SECRETED PROTEIN	3	II.C.5	D	22, 24
Rv2971	Rv2971	PROBABLE OxIDOREDUCTASE	7	I.B.7	D	42, 22
Rv2986c	hupB	PROBABLE DNA-BINDING PROTEIN HU HOMOLOG HUPB (HISTONE-LIKE PROTEIN) (HLP)	2	II.A.4	D	22, 24
Rv2996c	serA1	PROBABLE D-3-PHOSPHOGLYCERATE DEHYDROGENASE SERA1 (PGDH)	7	I.D.3	D	22, 24
Rv3002c	ilvN	PROBABLE ACETOLACTATE SYNTHASE (SMALL SUBUNIT) ILVN	7	I.D.7	D	22
Rv3003c	ilvB1	PROBABLE ACETOLACTATE SYNTHASE (LARGE SUBUNIT)(ACETOHYDROxY-ACID SYNTHASE)	7	I.D.7	A	19
Rv3006	lppZ	PROBABLE CONSERVED LIPOPROTEIN LPPZ	3	II.C.1	D	22, 16
Rv3012c	gatC	PROBABLE GLUTAMYL-TRNA(GLN) AMIDOTRANSFERASE (SUBUNIT C)(Glu-ADT SUBUNIT C)	2	II.A.3	D	
Rv3028c	fl37B	PROBABLE ELECTRON TRANSFER FLAVOPROTEIN (ALPHA-SUBUNIT)	7	I.B.6.c	A, C, D	42, 22, 24
Rv3029c	fl37A	PROBABLE ELECTRON TRANSFER FLAVOPROTEIN (BETA-SUBUNIT)	7	I.B.6.c	D	42, 22, 24
Rv3038c	Rv3038c	CONSERVED HYPOTHETICAL PROTEIN	10	VI	B	22
Rv3046c	Rv3046c	CONSERVED HYPOTHETICAL PROTEIN	10	V	D	42
Rv3071	Rv3071	CONSERVED HYPOTHETICAL PROTEIN	10	VI	B	24
Rv3075c	Rv3075c	CONSERVED HYPOTHETICAL PROTEIN	10	V	D	42, 22
Rv3083	Rv3083	PROBABLE MONOOxYGENASE (HYDROxYLASE)	7	I.B.7	C	
Rv3086	adhD	PROBABLE ZINC-TYPE ALCOHOL DEHYDROGENASE ADHD (ALDEHYDE REDUCTASE)	7	I.B.7	B	
Rv3089	fadD13	PROBABLE CHAIN -FATTY-ACID-CoA LIGASE FADD13 (FATTY-ACYL-CoA SYNTHETASE)	1	I.A.3	A	24
Rv3099c	Rv3099c	CONSERVED HYPOTHETICAL PROTEIN	10	VI	B, D	22, 24
Rv3105c	prfB	PROBABLE PEPTIDE CHAIN RELEASE FACTOR 2 PRFB (RF-2)	2	II.A.6	D	42
Rv3106	fprA	NADPH:ADRENODOxIN OxIDOREDUCTASE FPRA (NADPH-FERREDOxIN REDUCTASE)	7	I.B.6.c	D	16
Rv3107c	agpS	POSSIBLE ALKYLDIHYDROxYACETONEPHOSPHATE SYNTHASE AGPS (ALKYL-DHAP SYNTHASE)	1	I.B.7	B	
Rv3127	Rv3127	CONSERVED HYPOTHETICAL PROTEIN	10	V	D	42, 22, 24
Rv3132c	devS	TWO COMPONENT SENSOR HISTIDINE KINASE DEVS	9	I.J.2	D	22
Rv3133c	devR	TWO COMPONENT TRANSCRIPTIONAL REGULATORY PROTEIN DEVR	9	I.J.2	D	42, 22, 24
Rv3134c	Rv3134c	CONSERVED HYPOTHETICAL PROTEIN	10	V	D	22, 24
Rv3136	PPE51	PPE FAMILY PROTEIN	6	IV.C.2	D	
Rv3139	fadE24	PROBABLE ACYL-CoA DEHYDROGENASE FADE24	1	I.A.3	D	22, 24
Rv3140	fadE23	PROBABLE ACYL-CoA DEHYDROGENASE FADE23	1	I.A.3	A, B, D	22
Rv3143	Rv3143	PROBABLE RESPONSE REGULATOR	9	I.J.2	D	42
Rv3151	nuoG	PROBABLE NADH DEHYDROGENASE I (CHAIN G) (NADH-UBIQUINONE OxIDOREDUCTASE CHAIN G)	7	I.B.6.a	B	22, 24
Rv3153	nuoI	PROBABLE NADH DEHYDROGENASE I (CHAIN I)(NADH-UBIQUINONE OxIDOREDUCTASE CHAIN I)	7	I.B.6.a	D	24
Rv3193c	Rv3193c	PROBABLE CONSERVED TRANSMEMBRANE PROTEIN	3	II.C.5	D	22, 16
Rv3205c	Rv3205c	CONSERVED HYPOTHETICAL PROTEIN	10	VI	D	22
Rv3206c	moeB1	PROBABLE MOLYBDENUM COFACTOR BIOSYNTHESIS PROTEIN(MPT-SYNTHASE SULFURYLASE)	7	I.G.4	D	
Rv3211	rhlE	PROBABLE ATP-DEPENDENT RNA HELICASE RHLE	2	II.A.7	D	
Rv3223c	sigH	ALTERNATIVE RNA POLYMERASE SIGMA-E FACTOR (SIGMA-24) SIGH (RPOE)	2	II.A.7	D	19
Rv3224	Rv3224	POSSIBLE IRON-REGULATED SHORT-CHAIN DEHYDROGENASE/REDUCTASE	7	I.B.7	A, D	42, 22, 24
Rv3237c	Rv3237c	CONSERVED HYPOTHETICAL PROTEIN	10	III.A.2	A, D	22
Rv3240c	secA1	PROBABLE PREPROTEIN TRANSLOCASE SECA1 1 SUBUNIT	3	III.D	D	22, 16, 24
Rv3243c	Rv3243c	HYPOTHETICAL PROTEIN	16	VI	A	
Rv3244c	lpqB	PROBABLE CONSERVED LIPOPROTEIN LPQB	3	II.C.1	B	16, 24
Rv3248c	sahH	PROBABLE ADENOSYLHOMOCYSTEINASE SAHH (S-ADENOSYL-L-HOMOCYSTEINE HYDROLASE)	7	I.C.5	D	22, 16, 24
Rv3269	Rv3269	CONSERVED HYPOTHETICAL PROTEIN	0	III.B	D	42, 22
Rv3274c	fadE25	PROBABLE ACYL-CoA DEHYDROGENASE FADE25	1	I.A.3	D	42, 22, 24
Rv3280	accD5	PROBABLE PROPIONYL-CoA CARBOxYLASE BETA CHAIN 5 ACCD5 (PCCASE)	1	I.H.1	D	42, 22
Rv3281	Rv3281	CONSERVED HYPOTHETICAL PROTEIN	10	V	D	
Rv3282	Rv3282	CONSERVED HYPOTHETICAL PROTEIN	10	V	A	
Rv3285	accA3	PROBABLE BIFUNCTIONAL PROTEIN ACETYL-/PROPIONYL-COENZYME A CARBOxYLASE	1	I.H.1	D	22, 24
Rv3371	Rv3371	CONSERVED HYPOTHETICAL PROTEIN	10	V	B	24
Rv3373	echA18	PROBABLE ENOYL-CoA HYDRATASE(ENOYL HYDRASE) (UNSATURATED ACYL-CoA HYDRATASE)	1	I.A.3	A	
Rv3389c	Rv3389c	POSSIBLE DEHYDROGENASE	7	I.B.7	B	42, 22, 24
Rv3396c	guaA	PROBABLE GMP SYNTHASE [GLUTAMINE-HYDROLYZING] GLUTAMINE AMIDOTRANSFERASE)	7	I.F.1	D	42
Rv3411c	guaB2	PROBABLE INOSINE-5'-MONOPHOSPHATE DEHYDROGENASE GUAB2 (IMP DEHYDROGENASE)	7	I.F.1	D	42, 24
Rv3416	whiB3	TRANSCRIPTIONAL REGULATORY PROTEIN WHIB-LIKE WHIB3	9	I.J.1	A	
Rv3417c	groEL1	60 KDA CHAPERONIN 1 GROEL1 (PROTEIN CPN60-1) (GROEL PROTEIN 1)	0	III.B	A, B, D	22, 24
Rv3418c	groES	10 KDA CHAPERONIN GROES (PROTEIN CPN10) (PROTEIN GROES)	0	III.B	A, B, D	42, 22, 16, 24
Rv3421c	Rv3421c	CONSERVED HYPOTHETICAL PROTEIN	10	V	C	
Rv3443c	rplM	PROBABLE 50S RIBOSOMAL PROTEIN L13 RPLM	2	II.A.1	D	22, 24
Rv3448	Rv3448	PROBABLE CONSERVED INTEGRAL MEMBRANE PROTEIN	3	II.C.5	D	
Rv3457c	rpoA	PROBABLE DNA-DIRECTED RNA POLYMERASE (ALPHA CHAIN)	2	II.A.7	D	42, 22, 24
Rv3458c	rpsD	PROBABLE 30S RIBOSOMAL PROTEIN S4 RPSD	2	II.A.1	D	22, 24

Rv3462c	infA	PROBABLE TRANSLATION INITIATION FACTOR IF-1 INFA	2	II.A.6	D	42, 24
Rv3490	otsA	PROBABLE ALPHA, ALPHA-TREHALOSE-PHOSPHATE SYNTHASE [UDP-FORMING]	0	III.E	D	22, 24
Rv3496c	mce4D	MCE-FAMILY PROTEIN MCE4D	0	IV.A	C	
Rv3509c	ilv37	PROBABLE ACETOHYDROxYACID SYNTHASE ILVx (ACETOLACTATE SYNTHASE)	7	I.D.7	A, D	42, 22, 16
Rv3515c	fadD19	PROBABLE FATTY-ACID-CoA LIGASE FADD19 (FATTY-ACID-CoA SYNTHETASE)	1	I.A.3	B	
Rv3520c	Rv3520c	POSSIBLE COENZYME F420-DEPENDENT OxIDOREDUCTASE	7	IV.G	D	22
Rv3534c	Rv3534c	PROBABLE 4-HYDROxY-2-OxOVALERATE ALDOLASE (HOA)	7	II.B.6	C	
Rv3547	Rv3547	CONSERVED HYPOTHETICAL PROTEIN	10	V	D	22, 24
Rv3565	aspB	POSSIBLE ASPARTATE AMINOTRANSFERASE ASPB (TRANSAMINASE A) (ASPAT)	7	I.D.2	B	
Rv3584	lpqE	POSSIBLE CONSERVED LIPOPROTEIN LPQE	3	II.C.1	D	22, 16
Rv3592	TB11.2	CONSERVED HYPOTHETICAL PROTEIN TB11.2	10	V	A	42, 22
Rv3596c	clpC1	PROBABLE ATP-DEPENDENT PROTEASE ATP-BINDING SUBUNIT CLPC1	7	II.B.3	A, D	22, 24
Rv3597c	lsr2	PROBABLE IRON-REGULATED LSR2 PROTEIN PRECURSOR	10	V	D	22, 24
Rv3614c	Rv3614c	CONSERVED HYPOTHETICAL PROTEIN	10	V	D	22
Rv3615c	Rv3615c	CONSERVED HYPOTHETICAL PROTEIN	10	V	D	
Rv3623	lpqG	PROBABLE CONSERVED LIPOPROTEIN LPQG	3	II.C.1	C	22, 24
Rv3646c	topA	DNA TOPOISOMERASE I TOPA (OMEGA-PROTEIN) (RELAxING ENZYME)	2	II.A.5	D	22, 24
Rv3651	Rv3651	CONSERVED HYPOTHETICAL PROTEIN	10	VI	D	
Rv3666c	dppA	PROBABLE PERIPLASMIC DIPEPTIDE-BINDING LIPOPROTEIN DPPA	3	III.A.1	C	
Rv3670	ephE	POSSIBLE EPOxIDE HYDROLASE EPHE (EPOxIDE HYDRATASE) (ARENE-OxIDE HYDRATASE)	0	III.F	A	24
Rv3671c	Rv3671c	POSSIBLE MEMBRANE-ASSOCIATED SERINE PROTEASE	7	II.B.3	D	16
Rv3676	Rv3676	PROBABLE TRANSCRIPTIONAL REGULATORY PROTEIN (PROBABLY CRP/FNR-FAMILY)	9	I.J.1	D	42, 22
Rv3682	ponA2	PROBABLE BIFUNCTIONAL MEMBRANE-ASSOCIATED PENICILLIN-BINDING PROTEIN	3	II.C.3	A	22, 16, 24
Rv3688c	Rv3688c	CONSERVED HYPOTHETICAL PROTEIN	10	V	D	
Rv3690	Rv3690	PROBABLE CONSERVED MEMBRANE PROTEIN	3	VI	D	
Rv3699	Rv3699	CONSERVED HYPOTHETICAL PROTEIN	10	IV.H	D	42
Rv3708c	asd	ASPARTATE-SEMIALDEHYDE DEHYDROGENASE ASD (ASA DEHYDROGENASE)	7	I.D.2	D	22
Rv3709c	ask	ASPARTOKINASE ASK (ASPARTATE KINASE)	7	I.D.2	D	22
Rv3718c	Rv3718c	CONSERVED HYPOTHETICAL PROTEIN	10	VI	D	42, 22
Rv3720	Rv3720	POSSIBLE FATTY ACID SYNTHASE	1	I.H.3	D	22
Rv3725	Rv3725	POSSIBLE OxIDOREDUCTASE	7	I.B.7	B	16
Rv3763	lpqH	19 KDA LIPOPROTEIN ANTIGEN PRECURSOR LPQH	3	II.C.1	C, D	22, 24
Rv3778c	Rv3778c	POSSIBLE AMINOTRANSFERASE	7	V	B	22
Rv3780	Rv3780	CONSERVED HYPOTHETICAL PROTEIN	10	VI	D	42, 22
Rv3786c	Rv3786c	HYPOTHETICAL PROTEIN	16	VI	A, B, C	24
Rv3792	Rv3792	PROBABLE CONSERVED TRANSMEMBRANE PROTEIN	3	II.C.5	C	24
Rv3799c	accD4	PROBABLE PROPIONYL-CoA CARBOxYLASE BETA CHAIN 4 ACCD4 (PCCASE)	1	I.H.1	B	22
Rv3800c	pks13	POLYKETIDE SYNTHASE PKS13	1	I.I	A, D	22, 24
Rv3801c	fadD32	PROBABLE FATTY-ACID-CoA LIGASE FADD32 (FATTY-ACID-CoA SYNTHETASE)	1	I.A.3	D	22, 24
Rv3804c	fbpA	SECRETED ANTIGEN 85-A FBPA (MYCOLYL TRANSFERASE 85A)(FIBRONECTIN-BINDING PROTEIN A)	1	I.H.3	B, D	42, 22, 16, 24
Rv3809c	glf	UDP-GALACTOPYRANOSE MUTASE GLF (UDP-GALP MUTASE)	3	II.C.3	A	24
Rv3816c	Rv3816c	POSSIBLE ACYLTRANSFERASE	7	IV.H	D	22
Rv3823c	mmpL8	PROBABLE CONSERVED INTEGRAL MEMBRANE TRANSPORT PROTEIN MMPL8	3	II.C.4	C	24
Rv3825c	pks2	PROBABLE POLYKETIDE SYNTHASE PKS2	1	I.I	A, C, D	22, 24
Rv3835	Rv3835	PROBABLE CONSERVED MEMBRANE PROTEIN	3	VI	A	16, 24
Rv3841	bfrB	POSSIBLE BACTERIOFERRITIN BFRB	7	I.G.14	D	42, 22, 16, 24
Rv3849	Rv3849	CONSERVED HYPOTHETICAL PROTEIN	16	VI	D	42, 22, 16, 24
Rv3852	hns	POSSIBLE HISTONE-LIKE PROTEIN HNS	2	II.A.4	D	
Rv3865	Rv3865	CONSERVED HYPOTHETICAL PROTEIN	10	V	D	42
Rv3869	Rv3869	POSSIBLE CONSERVED MEMBRANE PROTEIN	3	V	D	24
Rv3870	Rv3870	POSSIBLE CONSERVED TRANSMEMBRANE PROTEIN	3	V	D	24
Rv3871	Rv3871	CONSERVED HYPOTHETICAL PROTEIN	10	V	D	24
Rv3874	es37B	10 KDA CULTURE FILTRATE ANTIGEN ESxB (LHP) (CFP10)	3	V	A, B, D	42, 22, 16, 24
Rv3877	Rv3877	PROBABLE CONSERVED TRANSMEMBRANE PROTEIN	3	V	B	
Rv3880c	Rv3880c	CONSERVED HYPOTHETICAL PROTEIN	10	V	D	22
Rv3894c	Rv3894c	POSSIBLE CONSERVED MEMBRANE PROTEIN	3	II.C.5	A, B, C	24
Rv3903c	Rv3903c	HYPOTHETICAL ALANINE AND PROLINE RICH PROTEIN	16	VI	B	
Rv3914	tr37C	THIOREDOxIN TRxC (TRx) (MPT46)	7	I.G.10	D	42, 22, 16, 24

Biological Sample: A - 2% SDS; B - 6M GuHCl; C - 4% TX-114; D - 2DLC TX-114

References:

X. Proteome Database System - Max Planck Institute (http://web.mpiib-berlin.mpg.de/cgi-bin/pdbs/2d-page/e37tern/n

Y. Tuberculosis Protein Database - Stratens Serum Institute (http://www.ssi.dk/sw14644.asp)

Z. Mawenyuega et al.

L. Gu et al

M. Malen et al

APPENDIX III

Supplementary Table 1 - List of Labeled and Unlabeled Protein Identifications of the Mtb ATPome along with corresponding protein functions and functional categories.

Gene Annotation	Name	Essential	Functional Role	Tuberculist	Class Name	SSI	Class Name
Rv0014c	**pknB**	E	TRANSMEMBRANE SERINE/THREONINE-PROTEIN KINASE B PKNB (PROTEIN KINASE B) (STPK B)	9	regulatory proteins	I.J.3	Serine-Threonine protein kinases and phosphoprotein phosphatases
Rv0015c	**pknA**	E	TRANSMEMBRANE SERINE/THREONINE-PROTEIN KINASE A PKNA (PROTEIN KINASE A) (STPK A)	9	regulatory proteins	I.J.3	Serine-Threonine protein kinases and phosphoprotein phosphatases
Rv0046c	**ino1**	NE	MYO-INOSITOL-1-PHOSPHATE SYNTHASE INO1 (Inositol 1-phosphate synthetase) (D-glucose 6-phosphate cycl	7	intermediary metabolism and respiration	V	Conserved hypotheticals
Rv0088	**Rv0088**	NE	HYPOTHETICAL PROTEIN	16	conserved hypotheticals with an orthologue in M. bovis	VI	Unknowns
Rv0211	**pckA**	-	PROBABLE IRON-REGULATED PHOSPHOENOLPYRUVATE CARBOXYKINASE [GTP] PCKA (PHOSPHOENOLPYRUVATE CARBOXYLAS	7	intermediary metabolism and respiration	I.C.2	Gluconeogenesis
Rv0231	**fadE4**	NE	ACYL-COA DEHYDROGENASE FADE4	1	lipid metabolism	I.A.3	Fatty Acids
Rv0350	**dnaK**	E	PROBABLE CHAPERONE PROTEIN DNAK (HEAT SHOCK PROTEIN 70) (HEAT SHOCK 70 KDA PROTEIN) (HSP70)	0	virulence, detoxification, adaptation	III.B	Chaperones/Heat shock
Rv0384c	**clpB**	E	PROBABLE ENDOPEPTIDASE ATP BINDING PROTEIN (CHAIN B) CLPB (CLPB PROTEIN) (HEAT SHOCK PROTEIN F84.1)	0	virulence, detoxification, adaptation	III.B	Chaperones/Heat shock
Rv0440	**groEL2**	E	60 KDA CHAPERONIN 2 GROEL2 (PROTEIN CPN60-2) (GROEL PROTEIN 2) (65 KDA ANTIGEN) (HEAT SHOCK PROTEIN	0	virulence, detoxification, adaptation	III.B	Chaperones/Heat shock
Rv0458	**Rv0458**	NE	ALDEHYDE DEHYDROGENASE	7	intermediary metabolism and respiration	I.B.7	Misc. Oxyidreductases
Rv0467	**icl**	NE	ISOCITRATE LYASE ICL (ISOCITRASE) (ISOCITRATASE)	7	intermediary metabolism and respiration	I.B.4	Glyoxylate bypass
Rv0468	**fadB2**	NE	PROBABLE 3-HYDROXYBUTYRYL-CoA DEHYDROGENASE FADB2 (BETA-HYDROXYBUTYRYL-CoA DEHYDROGENASE) (BHBD)	1	lipid metabolism	I.A.3	Fatty acids
Rv0475	**hbhA**	NE	IRON-REGULATED HEPARIN BINDING HEMAGGLUTININ HBHA (ADHESIN)	3	cell wall and cell processes	II.C.5	Other membrane proteins
Rv0636	**Rv0636**	E	CONSERVED HYPOTHETICAL PROTEIN	10	conserved hypotheticals	VI	Unknowns
Rv0652	**rplL**	E	PROBABLE 50S RIBOSOMAL PROTEIN L7/L12 RPLL (SA1)	2	information pathways	II.A.1	Ribosomal protein synthesis and modification
Rv0667	**rpoB**	E	DNA-DIRECTED RNA POLYMERASE (BETA CHAIN) RPOB (TRANSCRIPTASE BETA CHAIN) (RNA POLYMERASE BETA SUBUNI	2	information pathways	II.A.7	RNA synthesis, RNA modification and DNA transcription
Rv0685	**tuf**	E	PROBABLE IRON-REGULATED ELONGATION FACTOR TU TUF (EF-TU)	2	information pathways	II.A.6	Protein translation and modification
Rv0733	**adk**	-	PROBABLE ADENYLATE KINASE ADK (ATP-AMP TRANSPHOSPHORYLASE)	7	intermediary metabolism and respiration	I.F.5	Miscellaneous nucleoside/nucleotide reactions
Rv0772	**purD**	E	PROBABLE PHOSPHORIBOSYLAMINE--GLYCINE LIGASE PURD (GARS) (GLYCINAMIDE RIBONUCLEOTIDE SYNTHETASE) (PH	7	intermediary metabolism and respiration	I.F.1	Purine ribonucleotide biosynthesis
Rv0780	**purC**	E	PHOSPHORIBOSYLAMINOIMIDAZOLE- SUCCINOCARBOXAMIDE SYNTHASE PURC (SAICAR SYNTHETASE)	7	intermediary metabolism and respiration	I.F.1	Purine ribonucleotide biosynthesis
Rv0824c	**desA1**	E	PROBABLE ACYL-[ACYL-CARRIER PROTEIN] DESATURASE DESA1 (ACYL-[ACP] DESATURASE) (STEAROYL-ACP DESATURA	1	lipid metabolism	I.H.2	Modification of fatty and mycolic acids
Rv0860	**fadB**	NE	PROBABLE FATTY OXIDATION PROTEIN FADB	1	lipid metabolism	I.A.3	Fatty acids
Rv0902c	**prrB**	E	TWO COMPONENT SENSOR HISTIDINE KINASE PRRB	9	regulatory proteins	I.J.2	Two component systems
Rv0931c	**pknD**	NE	TRANSMEMBRANE SERINE/THREONINE-PROTEIN KINASE D PKND (PROTEIN KINASE D) (STPK D)	9	regulatory proteins	I.J.3	Serine-Threonine protein kinases and phosphoprotein phosphatases
Rv0951	**sucC**	E	PROBABLE SUCCINYL-CoA SYNTHETASE (BETA CHAIN) SUCC (SCS-BETA)	7	intermediary metabolism and respiration	I.B.3	TCA cycle
Rv0952	**sucD**	E	PROBABLE SUCCINYL-CoA SYNTHETASE (ALPHA CHAIN) SUCD (SCS-ALPHA)	7	intermediary metabolism and respiration	I.B.3	TCA cycle
Rv1080c	**greA**	NE	PROBABLE TRANSCRIPTION ELONGATION FACTOR GREA (Transcript cleavage factor greA)	2	information pathways	II.A.6	Protein translation and modification
Rv1094	**desA2**	E	POSSIBLE ACYL-[ACYL-CARRIER PROTEIN] DESATURASE DESA2 (ACYL-[ACP] DESATURASE) (STEAROYL-ACP DESATURA	1	lipid metabolism	I.H.2	Modification of fatty and mycolic acids
Rv1112	**Rv1112**	NE	Probable GTP binding protein	10	conserved hypotheticals	V	Conserved hypotheticals
Rv1185c	**fadD21**	NE	PROBABLE FATTY-ACID--CoA LIGASE FADD21 (FATTY-ACID-CoA SYNTHETASE) (FATTY-ACID-CoA SYNTHASE)	1	lipid metabolism	I.A.3	Fatty acids
Rv1201c	**dapD**	E	PROBABLE TETRAHYDRODIPICOLINATE N-SUCCINYLTRANSFERASE DAPD	7	intermediary metabolism and respiration	V	Conserved hypotheticals
Rv1266c	**pknH**	NE	PROBABLE TRANSMEMBRANE SERINE/THREONINE-PROTEIN KINASE H PKNH (PROTEIN KINASE H) (STPK H)	9	regulatory proteins	I.J.3	Serine-Threonine protein kinases and phosphoprotein phosphatases
Rv1297	**rho**	E	PROBABLE TRANSCRIPTION TERMINATION FACTOR RHO HOMOLOG	2	information pathways	II.A.7	RNA synthesis, RNA modification and DNA transcription
Rv1308	**atpA**	E	PROBABLE ATP SYNTHASE ALPHA CHAIN ATPA	7	intermediary metabolism and respiration	I.B.8	ATP-proton motive force
Rv1310	**atpD**	E	PROBABLE ATP SYNTHASE BETA CHAIN ATPD	7	intermediary metabolism and respiration	I.B.8	ATP-proton motive force
Rv1388	**mihF**	E	PUTATIVE INTEGRATION HOST FACTOR MIHF	2	information pathways	II.A.4	Nucleoproteins
Rv1437	**pgk**	E	PROBABLE PHOSPHOGLYCERATE KINASE PGK	7	intermediary metabolism and respiration	I.B.1	Glycolysis
Rv1475	**acn**	E	PROBABLE IRON-REGULATED ACONITATE HYDRATASE ACN (CITRATE HYDRO-LYASE)	7	intermediary metabolism and respiration	-	-
Rv1521	**fadD25**	NE	PROBABLE FATTY-ACID-CoA LIGASE FADD25 (FATTY-ACID-CoA SYNTHETASE) (FATTY-ACID-CoA SYNTHASE)	1	lipid metabolism	I.A.3	Fatty acids

Table Key: Bold text indicates proteins identified with High Confidence labeled peptides (see text for ranking criteria). SSI functional categories can be accessed here: http://web.mpiib-berlin.mpg.de/bioinformatik/FUNC_CLASS/overview.php?db_name=CLASS_Mtb

Gene Annotation	Name	Essential	Functional Role	Tuberculist	Class Name	SSI	Class Name
Rv1617	**pykA**	E	PROBABLE PYRUVATE KINASE PYKA	7	intermediary metabolism and respiration	I.B.1	Glycolysis
Rv1629	**polA**	E	DNA POLYMERASE I POLA	2	information pathways	II.A.5	DNA replication, repair, recombination and restriction/modification
Rv1636	**TB15.3**	NE	IRON-REGULATED CONSERVED HYPOTHETICAL PROTEIN TB15.3	10	conserved hypotheticals	V	Conserved hypotheticals
Rv1738	**Rv1738**	E	CONSERVED HYPOTHETICAL PROTEIN	10	conserved hypotheticals	V	Conserved hypotheticals
Rv1743	**pknE**	NE	TRANSMEMBRANE SERINE/THREONINE PROTEIN KINASE E				
Rv1746	**pknF**	NE	ANCHORED-MEMBRANE SERINE/THREONINE-PROTEIN KINASE PKNF (PROTEIN KINASE F) (STPK F)	9	regulatory proteins	I.J.3	Serine-Threonine protein kinases and phosphoprotein phosphatases
Rv2004c	**Rv2004c**	NE	CONSERVED HYPOTHETICAL PROTEIN	10	conserved hypotheticals	VI	Unknowns
Rv2005c	**Rv2005c**	NE	CONSERVED HYPOTHETICAL PROTEIN	10	conserved hypotheticals	V	Conserved hypotheticals
Rv2028c	**Rv2028c**	NE	CONSERVED HYPOTHETICAL PROTEIN	10	conserved hypotheticals	V	Conserved hypotheticals
Rv2031c	**hspX**	NE	HEAT SHOCK PROTEIN HSPX (ALPHA-CRSTALLIN HOMOLOG) (14 kDa ANTIGEN) (HSP16.3)	0	virulence, detoxification, adaptation	III.B	Chaperones/Heat shock
Rv2032	**acg**	NE	HYPOTHETICAL PROTEIN ACG	3	cell wall and processes		
Rv2074	**Rv2074**	-	CONSERVED HYPOTHETICAL PROTEIN	10	conserved hypotheticals	VI	Unknowns
Rv2159c	**Rv2159c**	NE	CONSERVED HYPOTHETICAL PROTEIN	10	conserved hypotheticals	VI	Unknowns
Rv2215	**dlaT**	E	Probable pyruvate dehydrogenase (E2 component) SucB	7	intermediary metabolism and respiration	I.B.3	TCA cycle
Rv2220	**glnA1**	E	GLUTAMINE SYNTHETASE GLNA1 (GLUTAMINE SYNTHASE) (GS-I)	7	intermediary metabolism and respiration	I.D.1	Glutamate family
Rv2241	**aceE**	NE	Probable pyruvate dehydrogenase E1 component aceE (PYRUVATE DECARBOXYLASE) (PYRUVATE DEHYDROGENASE)	7	intermediary metabolism and respiration	I.B.2	Pyruvate dehydrogenase
Rv2244	**acpM**	E	MEROMYCOLATE EXTENSION ACYL CARRIER PROTEIN ACPM	1	lipid metabolism	I.H.1	Synthesis of fatty and mycolic acids
Rv2296	**Rv2296**	NE	Probable haloalkane dehalogenase	7	intermediary metabolism and respiration	IV.I	Miscellaneous phosphatases, lyases, and hydrolases
Rv2299c	**htpG**	NE	PROBABLE CHAPERONE PROTEIN HTPG (HEAT SHOCK PROTEIN) (HSP90 FAMILY PROTEIN) (HIGH TEMPERATURE PROTE	0	virulence, detoxification, adaptation	III.B	Chaperones/Heat shock
Rv2319c	**Rv2319c**	E	UNIVERSAL STRESS PROTEIN	0	virulence, detoxification, adaptation	VI	Unknowns
Rv2376c	**cfp2**	NE	LOW MOLECULAR WEIGHT PROTEIN ANTIGEN	3	cell wall and processes	V	Conserved hypotheticals
Rv2428	**ahpC**	E	ALKYL HYDROPEROXIDE REDUCTASE C PROTEIN AHPC (ALKYL HYDROPEROXIDASE C)	0	virulence, detoxification, adaptation	III.F	Detoxification
Rv2477c	**Rv2477c**	E	PROBABLE MACROLIDE-TRANSPORT ATP-BINDING PROTEIN ABC TRANSPORTER	3	cell wall and cell processes	II.C.5	Other membrane proteins
Rv2501c	**accA1**	E	PROBABLE ACETYL-/PROPIONYL-COENZYME A CARBOXYLASE ALPHA CHAIN (ALPHA SUBUNIT) ACCA1: BIOTIN CARBOXYL	1	lipid metabolism	I.A.3	Fatty acids
Rv2510c	**Rv2510c**	E	CONSERVED HYPOTHETICAL PROTEIN	10	conserved hypotheticals	V	Conserved hypotheticals
Rv2607	**pdxH**	NE	PYRIDOXAMINE 5-PHOSPHATE OXIDASE PDXH	7	intermediary metabolism and respiration	I.G.6	Pyridoxine
Rv2623	**TB31.7**	E	CONSERVED HYPOTHETICAL PROTEIN TB31.7; UNIVERSAL STRESS PROTEIN	0	virulence, detoxification, adaptation	V	Conserved hypotheticals
Rv2624c	**Rv2624c**	NE	CONSERVED HYPOTHETICAL PROTEIN;UNIVERSAL STRESS PROTEIN	0	virulence, detoxification, adaptation	V	Conserved hypotheticals
Rv2773c	**dapB**	E	DIHYDRODIPICOLINATE REDUCTASE DAPB (DHPR)	7	intermediary metabolism and respiration	I.D.2	Aspartate family
Rv2780	**ald**	NE	SECRETED L-ALANINE DEHYDROGENASE ALD (40 KDA ANTIGEN) (TB43)	7	intermediary metabolism and respiration	I.A.2	Amino acids and amines
Rv2783c	**gpsI**	-	BIFUNCTIONAL PROTEIN POLYRIBONUCLEOTIDE NUCLEOTIDYLTRANSFERASE GPSI: GUANOSINE PENTAPHOSPHATE SYNTHE	2	information pathways	II.A.7	RNA synthesis, RNA modification and DNA transcription
Rv2858c	**aldC**	NE	PROBABLE ALDEHYDE DEHYDROGENASE ALDC	7	intermediary metabolism and respiration	I.B.7	Miscellaneous oxidoreductases and oxygenases
Rv2933	**ppsC**	NE	PHENOLPTHIOCEROL SYNTHESIS TYPE-I POLYKETIDE SYNTHASE PPSC	1	lipid metabolism	I.I	Polyketide and non-ribosomal peptide synthesis
Rv2950c	**fadD29**	NE	PROBABLE FATTY-ACID-CoA LIGASE FADD29 (FATTY-ACID-CoA SYNTHETASE) (FATTY-ACID-CoA SYNTHASE)	1	lipid metabolism	I.A.3	Fatty acids
Rv2984	**ppk1**	E	POLYPHOSPHATE KINASE PPK (POLYPHOSPHORIC ACID KINASE)	7	intermediary metabolism and respiration	I.A.4	Phosphorous Compounds
Rv3001c	**ilvC**	E	PROBABLE KETOL-ACID REDUCTOISOMERASE ILVC (Acetohydroxy-acid isomeroreductase) (Alpha-keto-beta-hydr	7	intermediary metabolism and respiration	I.D.7	Branched amino acid family
Rv3009c	**gatB**	E	PROBABLE GLUTAMYL-TRNA(GLN) AMIDOTRANSFERASE (SUBUNIT B) GATB (Glu-ADT SUBUNIT B)	2	information pathways	II.A.3	Aminoacyl tRNA synthases and their modification
Rv3010c	**pfkA**	NE	PROBABLE 6-PHOSPHOFRUCTOKINASE PFKA (PHOSPHOHEXOKINASE) (PHOSPHOFRUCTOKINASE)	7	intermediary metabolism and respiration	I.B.1	Glycolysis
Rv3028c	**fixB**	NE	PROBABLE ELECTRON TRANSFER FLAVOPROTEIN (ALPHA-SUBUNIT) FIXB (ALPHA-ETF) (ELECTRON TRANSFER FLAVOPRO	7	intermediary metabolism and respiration	I.B.6.c	Electron transport

Table Key: Bold text indicates proteins identified with High Confidence labeled peptides (see text for ranking criteria). SSI functional categories can be accessed here: http://web.mpiib-berlin.mpg.de/bioinformatik/FUNC_CLASS/overview.php?db_name=CLASS_Mtb

Gene Annotation	Name	Essential	Functional Role	Tuberculist	Class Name	SSI	Class Name
Rv3029c	**fixA**	E	PROBABLE ELECTRON TRANSFER FLAVOPROTEIN (BETA-SUBUNIT) FIXA (BETA-ETF) (ELECTRON TRANSFER FLAVOPROTE	7	intermediary metabolism and respiration	I.B.6.c	Electron transport
Rv3133c	**devR**	NE	TWO COMPONENT TRANSCRIPTIONAL REGULATORY PROTEIN DEVR (PROBABLY LUXR/UHPA-FAMILY)	9	regulatory proteins	I.J.2	Two component systems
Rv3140	**fadE23**	E	PROBABLE ACYL-CoA DEHYDROGENASE FADE23	1	lipid metabolism	I.A.3	Fatty acids
Rv3248c	**sahH**	E	PROBABLE ADENOSYLHOMOCYSTEINASE SAHH (S-ADENOSYL-L-HOMOCYSTEINE HYDROLASE) (ADOHCYASE)	7	intermediary metabolism and respiration	I.C.5	Sulphur metabolism
Rv3269	**Rv3269**	NE	CONSERVED HYPOTHETICAL PROTEIN	0	virulence, detoxification, adaptation	III.B	Chaperones/Heat shock
Rv3285	**accA3**	E	PROBABLE BIFUNCTIONAL PROTEIN ACETYL-/PROPIONYL-COENZYME A CARBOXYLASE (ALPHA CHAIN) ACCA3: BIOTIN C	1	lipid metabolism	I.H.1	Synthesis of fatty and mycolic acids
Rv3336c	**trpS**	E	PROBABLE TRYPTOPHANYL-TRNA SYNTHETASE TRPS (TRYPTOPHAN--TRNA LIGASE) (TRPRS) (TRYPTOPHAN TRANSLASE)	2	information pathways	II.A.3	Aminoacyl tRNA synthases and their modification
Rv3389c	**hdtY**	NE	PROBABLE 3-HYDROXYACYL-THIOESTER DEHYDRATASE HTDY	7	intermediary metabolism and respiration	I.B.7	Misc. Oxyidreductases
Rv3417c	**groEL1**	E	60 KDA CHAPERONIN 1 GROEL1 (PROTEIN CPN60-1) (GROEL PROTEIN 1)	0	virulence, detoxification, adaptation	III.B	Chaperones/Heat shock
Rv3418c	**groES**	E	10 KDA CHAPERONIN GROES (PROTEIN CPN10) (PROTEIN GROES) (BCG-A HEAT SHOCK PROTEIN) (10 KDA ANTIGEN)	0	virulence, detoxification, adaptation	III.B	Chaperones/Heat shock
Rv3580c	**cysS1**	E	CYSTEINYL-TRNA SYNTHETASE 1 CYSS1 (CYSTEINE--TRNA LIGASE 1)	2	information pathways	II.A.3	Aminoacyl tRNA synthases and their modification
Rv3596c	**clpC1**	E	PROBABLE ATP-DEPENDENT PROTEASE ATP-BINDING SUBUNIT CLPC1	7	intermediary metabolism and respiration	II.B.3	Proteins, peptides and glycopeptides
Rv3610c	**ftsH**	E	MEMBRANE-BOUND PROTEASE FTSH (CELL DIVISION PROTEIN)	3	cell wall and cell processes	III.C	Cell division
Rv3648c	**cspA**	E	PROBABLE COLD SHOCK PROTEIN A CSPA	0	virulence, detoxification, adaptation	III.E	Adaptations and atypical conditions
Rv3763	**lpqH**	NE	19 KDA LIPOPROTEIN ANTIGEN PRECURSOR LPQH	3	cell wall and processes	II.C.1	Lipoproteins(lppA-lpr0)
Rv3801c	**fadD32**	E	PROBABLE FATTY-ACID-CoA LIGASE FADD32 (FATTY-ACID-CoA SYNTHETASE) (FATTY-ACID-CoA SYNTHASE	1	lipid metabolism	I.A.3	Fatty acids
Rv3818	**Rv3818**	NE	HYPOTHETICAL PROTEIN	16	conserved hypotheticals with an orthologue in M. bovis	VI	Unknowns
Rv3846	**sodA**	E	SUPEROXIDE DISMUTASE [FE] SODA	0	virulence, detoxification, adaptation	III.F	Detoxification
Rv3868	**Rv3868**	NE	CONSERVED HYPOTHETICAL PROTEIN	10	conserved hypotheticals	V	Conserved hypotheticals
Rv3874	**esxB**	NE	10 KDA CULTURE FILTRATE ANTIGEN ESXB (LHP) (CFP10)	3	cell wall and cell processes	V	Conserved hypotheticals
Rv0005	***gyrB***	E	DNA GYRASE (SUBUNIT B) GYRB (DNA TOPOISOMERASE (ATP-HYDROLYSING)) (DNA TOPOISOMERASE II) (TYPE II DN	2	information pathways	II.A.5	DNA replication, repair, recombination and restriction/modification
Rv0407	***fgd1***	NE	PROBABLE F420-DEPENDENT GLUCOSE-6-PHOSPHATE DEHYDROGENASE FGD1	7	intermediary metabolism and respiration	IV.G	Coenzyme F420-dependent enzymes
Rv0702	***rplD***	E	PROBABLE 50S RIBOSOMAL PROTEIN L4 RPLD	2	information pathways	II.A.1	Ribosomal protein synthesis and modification
Rv0709	***rpmC***	E	PROBABLE 50S RIBOSOMAL PROTEIN L29 RPMC	2	information pathways	II.A.1	Ribosomal protein synthesis and modification
Rv0721	***rpsE***	E	PROBABLE 30S RIBOSOMAL PROTEIN S5 RPSE	2	information pathways	II.A.1	Ribosomal protein synthesis and modification
Rv0896	***gltA2***	-	PROBABLE CITRATE SYNTHASE I GLTA2	7	intermediary metabolism and respiration	I.B.3	TCA cycle
Rv0973c	***accA2***	E	PROBABLE ACETYL-/PROPIONYL-COENZYME A CARBOXYLASE ALPHA CHAIN (ALPHA SUBUNIT) ACCA2: BIOTIN CARBOXYL	1	lipid metabolism	I.A.3	Fatty acids
Rv1630	***rpsA***	E	PROBABLE RIBOSOMAL PROTEIN S1 RPSA	2	information pathways	II.A.1	Ribosomal protein synthesis and modification
Rv1819c	***bacA***	NE	PROBABLE DRUGS-TRANSPORT TRANSMEMBRANE ATP-BINDING PROTEIN ABC TRANSPORTER	3	cell wall and cell processes	III.A.6	Efflux proteins
Rv1837c	***glcB***	NE	PROBABLE MALATE SYNTHASE G GLCB	7	intermediary metabolism and respiration	I.B.4	Glyoxylate bypass
Rv2158c	***murE***	NE	ProbableUDP-N-acetylmuramoylalanyl-D-glutamate-2,6-diaminopimelate ligase MurE	3	cell wall and cell processes	II.C.3	Murein sacculus and peptidoglycan
Rv2211c	***gcvT***	E	Probable aminomethyltransferase GcvT (Glycine cleavage system T protein)	7	intermediary metabolism and respiration	I.C.1	General
Rv2445c	***ndkA***	NE	PROBABLE NUCLEOSIDE DIPHOSPHATE KINASE NDKA (NDK) (NDP KINASE) (NUCLEOSIDE-2-P KINASE)	7	intermediary metabolism and respiration	I.F.5	Miscellaneous nucleoside/nucleotide reactions
Rv2511	***orn***	E	OLIGORIBONUCLEASE ORN	7	intermediary metabolism and respiration	V	Conserved hypotheticals
Rv3224	***Rv3224***	NE	POSSIBLE IRON-REGULATED SHORT-CHAIN DEHYDROGENASE/REDUCTASE	7	intermediary metabolism and respiration	I.B.7	Miscellaneous oxidoreductases and oxygenases
Rv3275c	***purE***	E	PROBABLE PHOSPHORIBOSYLAMINOIMIDAZOLE CARBOXYLASE CATALYTIC SUBUNIT PURE (AIR CARBOXYLASE) (AIRC)	7	intermediary metabolism and respiration	I.F.1	Purine ribonucleotide biosynthesis
Rv3410c	***guaB3***	E	PROBABLE INOSINE-5'-MONOPHOSPHATE DEHYDROGENASE GUAB3 (IMP DEHYDROGENASE) (INOSINIC ACID DEHYDROGENA	7	intermediary metabolism and respiration	I.F.1	Purine ribonucleotide biosynthesis
Rv3443c	***rplM***	E	PROBABLE 50S RIBOSOMAL PROTEIN L13 RPLM	2	information pathways	II.A.1	Ribosomal protein synthesis and modification

Table Key: Bold text indicates proteins identified with High Confidence labeled peptides. *Bold Italic* - Medium Confidence (see text for ranking criteria). SSI functional categories can be accessed here: http://web.mpiib-berlin.mpg.de/bioinformatik/FUNC_CLASS/overview.php?db_name=CLASS_Mtb

Gene Annotation	Name	Essential	Functional Role	Tuberculist	Class Name	SSI	Class Name
Rv3456c	***rplQ***	NE	PROBABLE 50S RIBOSOMAL PROTEIN L17 RPLQ	2	information pathways	II.A.1	Ribosomal protein synthesis and modification
Rv3457c	***rpoA***	E	PROBABLE DNA-DIRECTED RNA POLYMERASE (ALPHA CHAIN) RPOA (TRANSCRIPTASE ALPHA CHAIN) (RNA POLYMERASE	2	information pathways	II.A.7	RNA synthesis, RNA modification and DNA transcription
Rv2941	*fadD28*	NE	FATTY-ACID-CoA LIGASE FADD28 (FATTY-ACID-CoA SYNTHETASE) (FATTY-ACID-CoA SYNTHASE)	1	lipid metabolism	I.A.3	Fatty acids
Rv2967c	*pca*	E	PROBABLE PYRUVATE CARBOXYLASE PCA (PYRUVIC CARBOXYLASE)	7	intermediary metabolism and respiration	I.B.3	TCA cycle
Rv3139	*fadE24*	E	PROBABLE ACYL-CoA DEHYDROGENASE FADE24	1	lipid metabolism	I.A.3	Fatty acids
Rv3157	*nuoM*	NE	PROBABLE NADH DEHYDROGENASE I (CHAIN M) NUOK (NADH-UBIQUINONE OXIDOREDUCTASE CHAIN M)	7	intermediary metabolism and respiration	I.B.6.a	aerobic
Rv3280	*accD5*	-	PROBABLE PROPIONYL-CoA CARBOXYLASE BETA CHAIN 5 ACCD5 (PCCASE) (PROPANOYL-COA:CARBON DIOXIDE LIGASE)	1	lipid metabolism	I.H.1	Synthesis of fatty and mycolic acids
Rv3411c	*guaB2*	E	PROBABLE INOSINE-5'-MONOPHOSPHATE DEHYDROGENASE GUAB2 (IMP DEHYDROGENASE) (INOSINIC ACID DEHYDROGENA	7	intermediary metabolism and respiration	I.F.1	Purine ribonucleotide biosynthesis
Rv3774	*echA21*	NE	POSSIBLE ENOYL-CoA HYDRATASE ECHA21 (ENOYL HYDRASE) (UNSATURATED ACYL-CoA HYDRATASE) (CROTONASE)	1	lipid metabolism	I.A.3	Fatty acids
Rv2524c	*fas*	E	PROBABLE FATTY ACID SYNTHASE FAS (FATTY ACID SYNTHETASE)	1	lipid metabolism	I.H.1	Synthesis of fatty and mycolic acids
Rv2590	*fadD9*	NE	PROBABLE FATTY-ACID-CoA LIGASE FADD9 (FATTY-ACID-CoA SYNTHETASE) (FATTY-ACID-CoA SYNTHASE)	1	lipid metabolism	I.A.3	Fatty acids
Rv2029c	*pfkB*	NE	Probable phosphofructokinase PfkB (PHOSPHOHEXOKINASE)	7	intermediary metabolism and respiration	I.B.1	Glycolysis
Rv1980c	*mpt64*	NE	IMMUNOGENIC PROTEIN MPT64 (ANTIGEN MPT64/MPB64)	3	cell wall and cell processes	II.C.2	Surface polysaccharides, lipopolysaccharides, proteins and antigens
Rv1925	*fadD31*	NE	PROBABLE ACYL-CoA LIGASE FADD31 (ACYL-COA SYNTHETASE) (ACYL-CoA SYNTHASE)	1	lipid metabolism	I.A.3	Fatty acids
Rv1699	*pyrG*	E	Probable CTP synthase PyrG	7	intermediary metabolism and respiration	I.F.2	Pyrimidine ribonucleotide biosynthesis
Rv1712	*cmk*	E	Probable Cytidylate kinase cmk (CMP kinase) (Cytidine monophosphate kinase) (CK)	7	intermediary metabolism and respiration	I.F.5	Miscellaneous nucleoside/nucleotide reactions
Rv1193	*fadD36*	E	PROBABLE FATTY-ACID-CoA LIGASE FADD36 (FATTY-ACID-CoA SYNTHETASE) (FATTY-ACID-CoA SYNTHASE)	1	lipid metabolism	I.A.3	Fatty acids
Rv1133c	*metE*	E	PROBABLE5-METHYLTETRAHYDROPTEROYLTRIGLUTAMATE--HOMOCYSTEINE METHYLTRANSFERASE METE (methionine synth	7	intermediary metabolism and respiration	I.D.2	Aspartate family
Rv0641	*rplA*	NE	PROBABLE 50S RIBOSOMAL PROTEIN L1 RPLA	2	information pathways	II.A.1	Ribosomal protein synthesis and modification
Rv0684	*fusA1*	E	PROBABLE ELONGATION FACTOR G FUSA1 (EF-G)	2	information pathways	II.A.6	Protein translation and modification
Rv0705	*rpsS*	E	PROBABLE 30S RIBOSOMAL PROTEIN S19 RPSS	2	information pathways	II.A.1	Ribosomal protein synthesis and modification
Rv0379	*secE2*	NE	POSSIBLE PROTEIN TRANSPORT PROTEIN SECE2	3	cell wall and cell processes	III.D	Protein and peptide secretion
Rv0041	*leuS*	E	PROBABLE LEUCYL-tRNA SYNTHETASE LEUS (LEUCINE--tRNA LIGASE) (LEURS)	2	information pathways	II.A.3	Aminoacyl tRNA synthases and their modification
Rv0009	ppiA	NE	PROBABLE IRON-REGULATED PEPTIDYL-PROLYL CIS-TRANS ISOMERASE A PPIA (PPIase A) (ROTAMASE A)	2	information pathways	II.A.6	Protein translation and modification
Rv0066c	icd2	NE	PROBABLE ISOCITRATE DEHYDROGENASE [NADP] ICD2 (OXALOSUCCINATE DECARBOXYLASE) (IDH) (NADP+-SPECIFIC I	7	intermediary metabolism and respiration	I.B.3	TCA cycle
Rv0101	nrp	NE	PROBABLE PEPTIDE SYNTHETASE NRP (PEPTIDE SYNTHASE)	1	lipid metabolism	I.I	Polyketide and non-ribosomal peptide synthesis
Rv0119	fadD7	-	PROBABLE FATTY-ACID-CoA LIGASE FADD7 (FATTY-ACID-CoA SYNTHETASE) (FATTY-ACID-CoA SYNTHASE)	1	lipid metabolism	I.A.3	Fatty acids
Rv0174	mce1F	NE	MCE-FAMILY PROTEIN MCE1F	0	virulence, detoxification, adaptation	IV.A	Virulence
Rv0189c	ilvD	E	PROBABLE DIHYDROXY-ACID DEHYDRATASE ILVD (DAD)	7	intermediary metabolism and respiration	I.D.7	Branched amino acid family
Rv0204c	Rv0204c	NE	PROBABLE CONSERVED TRANSMEMBRANE PROTEIN	3	cell wall and cell processes	II.C.5	Other membrane proteins
Rv0214	fadD4	NE	PROBABLE FATTY-ACID-CoA LIGASE FADD4 (FATTY-ACID-CoA SYNTHETASE) (FATTY-ACID-CoA SYNTHASE)	1	lipid metabolism	I.A.3	Fatty acids
Rv0405	pks6	NE	PROBABLE MEMBRANE BOUND POLYKETIDE SYNTHASE PKS6	1	lipid metabolism	I.I	Polyketide and non-ribosomal peptide synthesis
Rv0578c	PE_PGRS7	NE	PE-PGRS FAMILY PROTEIN	6	PE/PPE	IV.C.1.b	PE_PGRS subfamily
Rv0761c	adhB	NE	POSSIBLE ZINC-CONTAINING ALCOHOL DEHYDROGENASE NAD DEPENDENT ADHB	7	intermediary metabolism and respiration	I.B.7	Miscellaneous oxidoreductases and oxygenases
Rv0845	Rv0845	-	POSSIBLE TWO COMPONENT SENSOR KINASE	9	regulatory proteins	I.J.2	Two component systems
Rv0886	fprB	NE	PROBABLE NADPH:ADRENODOXIN OXIDOREDUCTASE FPRB (ADRENODOXIN REDUCTASE) (AR) (FERREDOXIN-NADP(+) REDU	7	intermediary metabolism and respiration	I.B.6.c	Electron transport
Rv1091	PE_PGRS22	-	PE-PGRS FAMILY PROTEIN	6	PE/PPE	IV.C.1.b	PE_PGRS subfamily
Rv1127c	ppdK	NE	PROBABLE PYRUVATE, PHOSPHATE DIKINASE PPDK	7	intermediary metabolism and respiration	I.C.1	General
Rv1161	narG	NE	PROBABLE RESPIRATORY NITRATE REDUCTASE (ALPHA CHAIN) NARG	7	intermediary metabolism and respiration	I.B.6.b	anaerobic

Table Key: ***Bold Italic*** **- Medium Confidence;** *italic - low confidence.* Plain - no ATP label. **(see text for ranking criteria). SSI functional categories can be accessed here: http://web.mpiib-berlin.mpg.de/bioinformatik/FUNC_CLASS/overview.php?db_name=CLASS_Mtb**

Gene Annotation	Name	Essential	Functional Role	Tuberculist	Class Name	SSI	Class Name
Rv1177	fdxC	E	PROBABLE FERREDOXIN FDXC	7	intermediary metabolism and respiration	I.B.6.c	Electron transport
Rv1183	mmpL10	NE	PROBABLE CONSERVED TRANSMEMBRANE TRANSPORT PROTEIN MMPL10	3	cell wall and cell processes	II.C.4	Conserved membrane proteins
Rv1358	Rv1358	NE	PROBABLE TRANSCRIPTIONAL REGULATORY PROTEIN	9	regulatory proteins	I.J.1	Repressors/activators
Rv1450c	PE_PGRS27	-	PE-PGRS FAMILY PROTEIN	6	PE/PPE	IV.C.1.b	PE_PGRS subfamily
Rv1522c	mmpL12	NE	PROBABLE CONSERVED TRANSMEMBRANE TRANSPORT PROTEIN MMPL12	3	cell wall and cell processes	II.C.4	Conserved membrane proteins
Rv1527c	pks5	NE	Probable polyketide synthase pks5	1	lipid metabolism	I.I	Polyketide and non-ribosomal peptide synthesis
Rv1661	pks7	E	Probable polyketide synthase pks7	1	lipid metabolism	I.I	Polyketide and non-ribosomal peptide synthesis
Rv1768	PE_PGRS31	E	PE-PGRS FAMILY PROTEIN	6	PE/PPE	IV.C.1.b	PE_PGRS subfamily
Rv1796	mycP5	-	PROBABLE PROLINE RICH MEMBRANE-ANCHORED MYCOSIN MYCP5 (SERINE PROTEASE) (SUBTILISIN-LIKE PROTEASE) (	7	intermediary metabolism and respiration	V	Conserved hypotheticals
Rv1941	Rv1941	-	PROBABLE SHORT-CHAIN TYPE DEHYDROGENASE/REDUCTASE	7	intermediary metabolism and respiration	I.B.7	Miscellaneous oxidoreductases and oxygenases
Rv1992c	ctpG	NE	PROBABLE METAL CATION TRANSPORTER P-TYPE ATPASE G CTPG	3	cell wall and cell processes	III.A.2	Cations
Rv2007c	fdxA	-	PROBABLE FERREDOXIN FDXA	7	intermediary metabolism and respiration	I.B.6.c	Electron transport
Rv2048c	pks12	NE	Probable polyketide synthase pks12	1	lipid metabolism	I.I	Polyketide and non-ribosomal peptide synthesis
Rv2089c	pepE	NE	Probable dipeptidase PepE	7	intermediary metabolism and respiration	II.B.3	Proteins, peptides and glycopeptides
Rv2155c	murD	E	Probable UDP-N-acetylmuramoylalanine-D-glutamate ligase MurD	3	cell wall and cell processes	II.C.3	Murein sacculus and peptidoglycan
Rv2157c	murF	E	ProbableUDP-N-acetylmuramoylalanyl-D-glutamyl-2,6-diaminopimelate-D-alanyl-D-alanyl ligase MurF	3	cell wall and cell processes	II.C.3	Murein sacculus and peptidoglycan
Rv2222c	glnA2	NE	PROBABLE GLUTAMINE SYNTHETASE GLNA2 (GLUTAMINE SYNTHASE) (GS-II)	7	intermediary metabolism and respiration	I.D.1	Glutamate family
Rv2245	kasA	E	3-OXOACYL-[ACYL-CARRIER PROTEIN] SYNTHASE 1 KASA (BETA-KETOACYL-ACP SYNTHASE) (KAS I)	1	lipid metabolism	I.H.1	Synthesis of fatty and mycolic acids
Rv2384	mbtA	NE	BIFUNCTIONAL ENZYME MBTA: SALICYL-AMP LIGASE (SAL-AMP LIGASE) + SALICYL-S-ArCP SYNTHETASE	1	lipid metabolism	I.I	Polyketide and non-ribosomal peptide synthesis
Rv2744c	35kd_ag	NE	CONSERVED 35 KDA ALANINE RICH PROTEIN	10	conserved hypotheticals	V	Conserved hypotheticals
Rv2931	ppsA	NE	PHENOLPTHIOCEROL SYNTHESIS TYPE-I POLYKETIDE SYNTHASE PPSA	1	lipid metabolism	I.I	Polyketide and non-ribosomal peptide synthesis
Rv2934	ppsD	NE	PHENOLPTHIOCEROL SYNTHESIS TYPE-I POLYKETIDE SYNTHASE PPSD	1	lipid metabolism	I.I	Polyketide and non-ribosomal peptide synthesis
Rv2935	ppsE	NE	PHENOLPTHIOCEROL SYNTHESIS TYPE-I POLYKETIDE SYNTHASE PPSE	1	lipid metabolism	I.I	Polyketide and non-ribosomal peptide synthesis
Rv2940c	mas	NE	PROBABLE MULTIFUNCTIONAL MYCOCEROSIC ACID SYNTHASE MEMBRANE-ASSOCIATED MAS	1	lipid metabolism	I.I	Polyketide and non-ribosomal peptide synthesis
Rv2996c	serA1	E	PROBABLE D-3-PHOSPHOGLYCERATE DEHYDROGENASE SERA1 (PGDH)	7	intermediary metabolism and respiration	I.D.3	Serine family
Rv3006	lppZ	-	PROBABLE CONSERVED LIPOPROTEIN LPPZ	3	cell wall and cell processes	II.C.1	Lipoproteins(lppA-lpr0)
Rv3116	moeB2	NE	PROBABLE MOLYBDENUM COFACTOR BIOSYNTHESIS PROTEIN MOEB2 (MPT-SYNTHASE SULFURYLASE) (MOLYBDOPTERIN SY	7	intermediary metabolism and respiration	I.G.4	Molybdopterin
Rv3132c	devS	E	TWO COMPONENT SENSOR HISTIDINE KINASE DEVS	9	regulatory proteins	I.J.2	Two component systems
Rv3147	nuoC	-	PROBABLE NADH DEHYDROGENASE I (CHAIN C) NUOC (NADH-UBIQUINONE OXIDOREDUCTASE CHAIN C)	7	intermediary metabolism and respiration	I.B.6.a	aerobic
Rv3158	nuoN	NE	PROBABLE NADH DEHYDROGENASE I (CHAIN N) NUON (NADH-UBIQUINONE OXIDOREDUCTASE CHAIN N)	7	intermediary metabolism and respiration	I.B.6.a	aerobic
Rv3343c	PPE54	E	PPE FAMILY PROTEIN	6	PE/PPE	IV.C.2	PPE family
Rv3345c	PE_PGRS50	-	PE-PGRS FAMILY PROTEIN	6	PE/PPE	IV.C.1.b	PE_PGRS subfamily
Rv3507	PE_PGRS53	NE	PE-PGRS FAMILY PROTEIN	6	PE/PPE	IV.C.1.b	PE_PGRS subfamily
Rv3508	PE_PGRS54	-	PE-PGRS FAMILY PROTEIN	6	PE/PPE	IV.C.1.b	PE_PGRS subfamily
Rv3536c	hsaE	-	PROBABLE HYDRATASE	7	intermediary metabolism and respiration	II.B.6	Aromatic hydrocarbons
Rv3628	ppa	-	INORGANIC PYROPHOSPHATASE PPA (PYROPHOSPHATE PHOSPHO-HYDROLASE) (PPASE) (INORGANIC DIPHOSPHATASE) (D	7	intermediary metabolism and respiration	I.A.4	Phosphorous compounds
Rv3800c	pks13	E	POLYKETIDE SYNTHASE PKS13	1	lipid metabolism	I.I	Polyketide and non-ribosomal peptide synthesis
Rv3826	fadD23	-	PROBABLE FATTY-ACID-CoA LIGASE FADD23 (FATTY-ACID-CoA SYNTHETASE) (FATTY-ACID-CoA SYNTHASE)	1	lipid metabolism	I.A.3	Fatty acids

Table Key: *Bold Italic* - Medium Confidence; *italic - low confidence.* Plain - no ATP label. **(see text for ranking criteria). SSI functional categories can be accessed here: http://web.mpiib-berlin.mpg.de/bioinformatik/FUNC_CLASS/overview.php?db_name=CLASS_Mtb**