DISSERTATION

COMPREHENSIVE CHARACTERIZATION OF *Mycobacterium tuberculosis* STRAINS

AFTER ACQUISITION OF ISONIAZID RESISTANCE

Submitted by

Luisa Maria Nieto Ramirez

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

Fall 2016

Doctoral Committee:

    Advisor: Karen M. Dobos

    Anne Lenaerts
    Jessica Prenni
    Richard Slayden

ABSTRACT

COMPREHENSIVE CHARACTERIZATION OF *Mycobacterium tuberculosis* STRAINS

AFTER ACQUISITION OF ISONIAZID RESISTANCE

Despite the global efforts to reduce tuberculosis (TB) rates, the emergence of drug

resistant TB has not allowed effective control of this disease. In the last decade, there were

roughly 10 million new TB cases per year and isoniazid resistant (INHr) TB accounted for 9.5%

of these cases around the world. In 2012, United States had an interruption in the supply of

isoniazid (INH), which increased the likelihood of INH resistance rates. Although INH resistance

in *Mycobacterium tuberculosis (Mtb)* is multigenic, mutations in the catalase-peroxidase (*katG)*

gene predominate amongst INHr *Mtb* strains. The characterization of the *Mtb* proteome before

and after acquiring INH resistance remains understudied. Additionally, the effect of these drug-

resistance-conferring mutations on *Mtb* fitness and virulence is variable. The purpose of this

work is to describe a complete biochemical and immunological characterization of the INHr

acquisition in *Mtb*. In this way, a global exploration of the protein and mycolic acids differences

in *Mtb* cultures, as well as differences in the immune response and bacterial virulence in the

mouse model comparing clonal susceptible and INHr pairs of *Mtb* were evaluated. After this,

common trends were analyzed and the findings were interpreted in the context of bacterial

metabolism and host-interaction.

For this work, two clonal clinical *Mtb* strains and one laboratory clonal pair of the H37Rv

strain with different susceptibility profiles to INH were studied. The H37Rv INHr strain was

isolated from a mouse that was exposed to INH in the lab and developed the same *katG* mutation

that one of the clinical INHr strain has (V1A). In all cases, the first strain was susceptible to all

tested drugs (mostly known as the INHs strain in this dissertation) while the second strain was resistant only to INH (named INHr throughout this work). The clinical pairs were confirmed as clonal pairs of the Beijing and T genotype respectively by spoligotyping and restriction fragment polymorphism analysis that uses the patterns given by the distribution of the insertion sequence (*IS*)-6110. Previous whole genome sequencing analysis of the clinical clonal pairs showed a *katG* mutation and the presence of some additional non-synonymous polymorphisms in the INHr strains. After the proteomic analysis, a *katG* PCR sequencing confirmed two mutations in *katG* for the T INHr pair (V1A and E3V) while the L101R mutation previously identified for the Beijing INHr was not confirmed. This mutation was highly unstable and the Beijing INHr might have reversed its phenotype after the absence of INH during *in vitro* growth. Therefore, the analysis with the Beijing clonal pair is only presented in chapter II.

Protein comparison of secreted and cellular fractions (membrane, cytosol and cell wall) between clinical and lab clonal pairs of *Mtb* before and after acquisition of INH resistance revealed at least 25 commonly altered proteins looking at the same cellular fractions. These proteins were involved in ATP synthase machinery, lipid metabolism, regulatory events, virulence, detoxification and adaptation processes. Western blot analysis supported some of our findings, particularly the lower level of bacterial enzyme KatG in the INHr strains. Mycolic acid (MA) analysis in these clonal pairs did not reveal a common trend in these molecules for INHr strains but generated supporting information about an alternative fatty acid biosynthetic pathway in the clinical INHr strain. These analyses are further described in chapter III. Additionally, differences in bacterial growth, immune response and pathology induced by *Mtb* strains harboring mutations at the N-terminus of KatG were evaluated in the C57BL/6 mouse model. The results in the mouse study support the idea of the individual effect of specific located

mutations in the *katG* gene together with the associated changes in the bacterial proteome induce differences in the *Mtb* virulence and pathogenicity. In addition, the *in vivo* results also suggest the contribution of innate immune response via TLR-2 in the clearance of the INHr-attenuated *Mtb* strains. Further details of this work are described in chapter IV.

This work provides a better understanding of new compensatory mechanisms in *Mtb* after INH resistance acquisition providing novel information that could be used to address alternative combined therapies as well as the identification of new drug targets in INHr strains. The results presented here also contribute to the generation of new hypothesis regarding RNA decay in *Mtb* and the need to evaluate if the observed biochemical differences are also associated with the bacterial exposure to the first line drug therapy that occurred in the patient. After the results obtained in this study, a subsequent biochemical analysis of *Mtb* strains obtained from patients before and after drug treatment is proposed to improve the description of the evolution of the acquired drug resistant phenomena observed in TB cases that limit the global disease control and hence its eradication (chapter V).

ACKNOWLEDGEMENTS

Wilusz. Also, thanks to members of the Proteomics and Metabolomics Facilities at Colorado State University and my multidisciplinary writing group for their valuable contributions during my PhD work.

I want to give special thanks to my husband, love and laboratory partner Gustavo Diaz who has not only been the best partner in the happiest and most difficult moments of my life, but also who has shared his scientific point of view and comments that have contributed a lot in my pathway to become a PhD. Thanks to my family and friends that offered me their support and incredible moments in this exciting journey. Finally, thanks to the financial aid offered by the Colombian government through the "Francisco Jose de Caldas scholarship" to complete my doctoral degree overseas.

DEDICATION

To my family:

My daughter Violeta and my husband Gustavo, who make my life colorful, lovely and every day

better.

My parents Jorge and Ruby to whom I will love and be forever grateful. All of you are in every

step that I take.

TABLE OF CONTENTS

CHAPTER I - LITERATURE REVIEW: TUBERCULOSIS HISTORY AND BIOLOGICAL
ASPECTS RELATED TO DRUG RESISTANCE

1.1 Introduction

The history of tuberculosis (TB), the disease caused by *Mycobacterium tuberculosis (Mtb)*,

has a remarkable involvement in human history; particularly in the evolution of human society

and in the development of many scientific disciplines. As more research tools have beeen

developed, TB is no longer considered a curse or a bad air. However, the first appearance of *Mtb*

in the world is still a controversy. TB has a negative role in many pages of human history, taking

lives of many renowned artists, politicians, as well as poor, wealthy, young or adult individuals.

After its apparent "resurgence" at the end of the last century with the concomitant arises of

HIV/AIDS cases, TB has been mainly associated with poverty and immunosuppression. *Mtb* was

the model microorganism that inspired Koch to develop his postulates and also sciences such as

Microbiology and Immunology. What Robert Koch never probably imagined was that over a

century later, this disease was going to continue as the leading cause of death. Regarding TB

treatment, it is discouraging to see how this lethal disease only started to be cured and controlled

during 1950's with the discovery of the first chemotherapy. It is difficult to imagine the labor of

a medical doctor taking care of TB patients before 1950's without an available treatment option.

However, this could be the same scenario that many health providers have to face nowadays with

the current spreading of multidrug resistant (MDR) and extensively drug resistant (XDR) TB

cases. In this chapter, major milestones in the TB/human history are described, followed by the

enumeration of important characteristics of *Mtb* physiology. *Mtb* is distinct from other

microorganisms not only due to its particular slow growth, biology and difficulties to study it in

the laboratory, but also, due to the complexity to eliminate it from infected people. These

1

important aspects will give a context about the chemotherapy for TB control and the current

global challenges towards it. This work explores the interaction between the bacteria and

isoniazid, which is one of the most effective drugs against *Mtb*, and particularly focuses on the

events that follow after the bacteria acquire resistance to this drug from a biochemical and

immunological perspective. In this regard, a review focused on mycobacterial cell envelope,

major physiological pathways and *Mtb*-host interaction are also included in this chapter to

provide a better background and illustration of the experimental findings.

1.2 Tuberculosis in the human history

TB is one of the most ancient human disease and has killed more people than any other

infectious disease in history (more than a billion in the last 200 years) (1). Learning TB history is

learning human history. Signs of *Mtb*, the bacteria agent causing the disease, were found in

Neolithic remains and in the spines of Egyptian mummies of more than 5.000 years ago (2) . In

Egyptian mummies, a particular form of the disease affecting vertebrae was observed, that was

called Pott's disease which is tuberculous spondylitis (3, 4). Despite these findings, references of

TB in Egyptian papyri have not been found (5). The 'white plague" –as it was commonly called

during the epidemic phase in Europe that last over two centuries— is a remarkably enduring

disease and once it arrives in a community, it stays. *Mtb* interaction with humanity has received

different names: white plague, "consumption" (which was derived from the Latin *consūmptiōn*,

that means "wasting"), *phthisis* (greek), *schachepheth* (ancient Hebrew for wasting disease),

scrofula (the king's evil, a colorful description for swollen and discharging tuberculous lymph

glands), *tabey diq* (urdu), *xoy* (waning of the moon, hindu) among others. The term

"tuberculosis" was actually given by the German medical professor Johann Lukas Schönlein in

1839 (6-8). *Mtb* was also found in bones of pre-Columbian America, Huron Indians of Ontario,

native Indians of New York State, and Peruvian mummies. The Peruvian remains suggested that sea mammals (specifically pinnipeds) had a role in transmitting the disease to humans between AD 700 and AD 1000 (9). This bacterium was also found in the skeleton of *Ainu* (from protohistoric period in Japan) and elephant remains in India before 2000 BC (8).

The TB history has being closely associated with important events in human history and directly related with the progress in biological sciences and influencing human culture. Around the 19th century, the death rate due to TB was estimated at 7 million people worldwide. As TB was taking lives of many, the major event in the medical history of this century was revealed when Robert Koch published for the first time the discovery of the tubercle bacillus: *Mycobacterium tuberculosis (Mtb).* Previous attempts to demonstrate that TB was contagious by the aerosol route were already done in 1877 by Tappeiner, but Koch was the first one able to demonstrate that TB was caused by a microorganism and to identify the bacteria (6). This milestone initiated the modern era of TB, a disease that was previously thought to be hereditary, since it "ran" in families. During his talk in 1882, Koch performed a demonstrative description of the steps that he followed to reach his discovery. These collective steps received the name of the Koch's postulates and allowed the discovery of other bacterial-causing diseases such as cholera. Koch also described the involvement of the host in TB pathology, setting then the basis for two beautiful sciences Microbiology and Immunology. However, Koch attempts to find a cure for TB were not fruitful (8). In modern times, TB was represented in some artistic pieces of music, opera, novels and paintings. During mid 1800's and early 1900's, the opera *La Traviata (Violetta)* and the novel *The magic mountain* described how this disease affected the society of those times and the first attempts to control it by sending patients to sanatoria.

With sanatoria started the first measures to treat TB patients that later evolved to the chemotherapy treatment; both interventions had a significant impact in the society. Sanatoria were a non-active medical intervention that looked for providing rest and better air. This idea was proposed by the English doctor, George Bodington who recommended "a pure atmosphere" to treat TB patients. This was probably in line with Koch thinking that the answer for TB cure was improving the natural defences in the patient's bodies. This could be reached by resting to mobilize the immune response more effectively to the infection site. Sanatoria generated a different way to perceive this disease, separating families and segregating TB patients as well and health care providers indirectly, most of whom ended up infected with the bacteria. This strategy could at least reduce the spreading of the disease in the increasingly more populated cities. The idea of a better air was even the motivation for some of the first migrants to the western cities of United States (7, 8). However, the high altitude and the less contaminated air was not enough for many TB patients who inevitably encounter death after their migration. By that time, the future of most TB patients was hopeless since they did not have a therapy to control the disease. However, TB patients and their families started to feel hopeful after the discovery of the first focused anti-TB drug therapy around 1940's. With the new antibiotic era, the number of TB cases and deaths were drastically reduced and with it also most of the sanatoria. TB was no longer considered a major health concern as it was primarily restricted to areas of poverty and crowding. Unfortunately, after the use of new drugs, drug-resistant TB cases emerged too, which created additional challenges as alternative chemotherapies were needed for these cases. However, from 1963 to 2012 the efforts and the creation of new specific anti-TB chemotherapy were strongly reduced. During this time, drugs developed against other clinically relevant bacteria were repurposed to also treat resistant TB cases.

The increasing spread of HIV/AIDS and other immunosuppressive conditions brought back new TB cases to countries where this disease was thought to be eradicated. Additionally, intrinsic conditions of the microorganism as well as increasing complex social dynamics make this situation worse. On one hand, the explosion of drug resistant and hypervirulent *Mtb* strains and on the other hand the globalization plus economic and social changes, including the expanding human migrations are likely causes to the increase of the TB global prevalence. All these events generated many concerns worldwide and motivated the creation of effective strategies to mitigate TB effects and to eventually eradicate TB as a human health problem (1).

1.3 Tuberculosis disease

Around 9 million new TB cases have been estimated per year in the last decade. More than half of the TB cases are from South-East Asia and Western Pacific regions, however the Africa region has had the most severe burden relative to population with 281cases for every 100,000 people. This number is more than double of the global average reported for the same year 2014 (133 in 100,000 population) (10, 11).

TB is caused by *Mtb*, a bacterium mainly transmitted by aerosol route, with an infective dose as low as 1 to 5 bacilli (12). Regarding mycobacterial diseases in general, Daffe and Draper in an early paper proposed four descriptive points (13):

1. *The Mycobacteria is encountered intracellularly during most of the course of the disease*
2. *Host immune response is mainly cellular, with a very decent or not very clear role of antibodies for the disease progression and prognosis.*
3. *Mycobacteria representatives can induce some immunomodulatory properties.*

4. *Finally, some of the major pathological findings are due to the inappropriate immunological response within the host, which is also related with the persistent forms of the disease.*

Tuberculosis is a good example of a mycobacterial disease that fits all the previously described points. *Mtb* has the capacity of infecting every organ, the pulmonary form being the most commonly diagnosed. Extra-pulmonary forms can also be found in organs such as bones, brain, glands, among others (14). There is also milliary TB, where *Mtb* is spread by the blood stream and scatter millet seeds can be seen in an X-ray of the chest. Lastly, there is also infection without clinical symptoms or latent TB, which can remain dormant or reactivate in patients during their life time. This reactivation will depend on an immunosuppressive event that no longer contains the bacteria in the granuloma structures in the lung, allowing the transition from latent to active TB; where the patient becomes symptomatic and infective. Due to the chronicity of the disease, a TB patient could be diagnosed after he/she has exposed many contacts to the bacteria, creating great concerns for public health. TB is also a very debilitating disease that in many cases reduces the patient's ability to work, generating also an important socio-economic impact.

When an individual enters in contact with the bacterium, three different outcomes can occur: the bacteria could be eradicated by the immune system, develop into an active disease (less than 10% of cases) or contained in a latent infection. Once *Mtb* reach the lungs, it is phagocytosed by alveolar macrophages that attract more immune cells to build up the hallmark structure of *Mtb* infection: the granuloma. The granuloma is a dynamic structure that varies from solid to necrotic to caseous and all three forms can coexist in a patient (15). Symptoms of pulmonary TB usually include chronic and productive coughing sometimes with hemoptysis, fever, night sweats, weight

loss and in some cases unusual lassitude, nausea or persistent indigestion, among others. By the time of diagnosis (that is usually done by detecting the bacilli in the sputum), signs of a late ulceration of a pulmonary lesion into a bronchus with at least one third of the lung being affected can already be occuring.

The severity of the disease depends on bacterial as well as host factors. From the bacteria side there are different genotypes of *Mtb* that have been associated with higher virulence and higher probability to develop resistant forms of the disease. Regarding the host, an immunocompromised status increases the probability to develop an active disease with higher risk of bacteremia and spread to other organs. Moreover, many researchers in the field have concluded that most of the pathology and symptoms of the disease are generated by the host immune responses against the bacteria.

1.4 Tuberculosis treatment

In the last century, the number of TB-related deaths has declined due to the discovery of anti-TB therapy. The first antibacterial therapy was developed in 1933 when Dr. Gerald Domagk synthesized Prontosil, a sulfonamide with bactericidal activity against deadly *Streptococcus* strains. Prontosil was effective in containing secondary TB infections in guinea pigs, but only at extremely high doses. Even at such doses, Prontosil did not eliminate *Mtb* (7). After sulfonamide discovery, thiosemicarbazones were introduced as a more effective synthetic anti-*Mtb* compound in 1950 (16).

A more focused anti-TB therapy was developed in 1944 with the discovery of streptomycin, a compound that was purified from *Streptomyces griseus.* This compound was tested against one of the most virulent bacteria of that time, H37Rv, which continues to be a lab

reference strain to this day. With streptomycin treatment, clinicians realized that long-term treatment (more than three months) was required to see some improvement in patient health (17). But the initial euphoria due to the beneficial properties of streptomycin did not last long as *Mtb* strains resistant to this drug were detected in as many as 80% of total TB cases in two years after its discovery. In 1943, the para-amino salicylic acid (PAS or 4-aminosalicylic acid), a derivative of salicylic acid (aspirin) was discovered by the Swedish chemist Jörgen Lehmann. PAS bacteriostatic activity was tested against *Mtb* in 1948. The synthesis of PAS was an innovative approach that took into account the previous work with sulfonamides, where the amino group in the *para* position provided bacteriostatic activity (8, 18). With PAS discovery, the first combined therapy that included streptomycin started in USA in 1948 (16, 19).

In 1951, the anti-TB properties of a new drug, isoniazid (INH), were reported. This was a critical event in TB history that was optimistically described as the "new treatment for the white scourge" in magazines and newspapers of that time. In general, there was excitement about the possible end of TB (7, 8). INH was first synthesized by Hans Meyer and Josef Mally at Charles University of Prague in 1912, as part of their doctoral degree (8, 20). Decades later in 1951, the anti-*Mtb* properties of INH were independently discovered in United States and in Germany (21). Prior to the discovery of INH, *in vitro* and *in vivo* bactericidal activity of the nicotinamide (or vitamin B3) against *Mtb* and *M. leprae* were revealed. During the search for derivative synthetic compounds that would potentiate nicotinamide activity, researchers found an intermediate compound, INH, with a powerful activity against *Mtb*. This new compound was more potent than any other existing drug against *Mtb* with relatively minor side effects (21). In 1952, the options for TB treatment continued to be rest, anti-TB chemotherapy and, in some cases, pulmonary resection. At that time, INH started to be tested against untreated TB patients from

the Navajo Reservation at Many Farms, Arizona (22) and, after the observed success, it became

part of the anti-TB regimes throughout the United States. Pictures of recovering patients were

published revealing the big hopes that this discovery gave to TB control.

Followed the discovery of INH, pyrazinamide (1954), cycloserine (1955), kanamycin

(1957), ethionamide (1961), ethambutol (1961), capreomycin (1963), rifampicin (1963),

amikacin (1972) and ofloxacin (1980) were discovered. Some of these drugs were replaced by

newly discovered drugs due to their harmful side effects. This period of time was considered the

golden age for anti-TB drugs. The inclusion of rifampicin (RIF) reduced the anti-TB treatment to

9 months and then after the inclusion of pyrazinamide, the treatment was reduced to 6 months.

By the end of 1990s, levofloxacin (1996) and moxifloxacin (1999) were approved by the Food

and Drugs Administration (FDA) for the use in the United States. These two quinolones were

originally designed for the treatment of other bacterial infections before discovering their anti-

TB properties. In the 1990s, a new class of drugs were discovered (oxazolidinones), but only

approved for use in humans until 2000. Linezolid is a good representative of these drugs and it

was initially intended to treat drug resistant Gram positive bacteria with harmful side effects. A

complete description of the most relevant anti-TB drugs with the target, effect and genes

encoding for activation is provided in Table 1.1.

After almost 40 years, a new drug originally designed for TB treatment was approved by

FDA, opening the door for a new anti-mycobacterial era. This drug was bedaquiline and started

to be prescribed in 2012 with the restraint to be used only in multi-drug resistant TB cases

(MDR-TB) (23). More recently, delamanid (2014) was conditionally approved as a last resource

care option for the most resistant forms of the disease (24, 25) (Figure 1.1). Antibiotic such as

thiolactomycin (produced by *Nocardia spp*) and its analogues have been proven to be effective

against *Mtb,* although they have not been used in the clinical setting yet (26-28). Teixobactin is another new antibiotic produced by *Eleftheria terrae* with potential activity against drug resistant *Mtb* (29). Drug candidates such as Q203, TBI-166, and TBA-7371, among others including new combinations and higher doses of currently approved drugs, are currently in various stages of clinical trials (Figure 1). Among other treatment strategies, many research groups are evaluating alternative routes for drug delivery (via aerosol) and new compounds that include liposomes, nanoparticles, and immunomodulatory therapy such as treatment with IFN-γ (17, 30). Currently, shorter combined therapy with moxifloxacin as a possible first line treatment option or higher doses of rifampicin or isoniazid are being evaluated (17). The rationale behind the increase dose of rifampicin is that the currently used dose of rifampicin was proposed in 1971 with the basis of generating a cost-effective treatment that was non-toxic for TB patients, albeit a study of the maximum dose of the drug tolerated in human has never been performed (31). Recent studies animal models have shown that higher doses of this drug could be effective even in shorter regimes, reducing also the probability to generate resistant microorganisms to the drug (31-33).
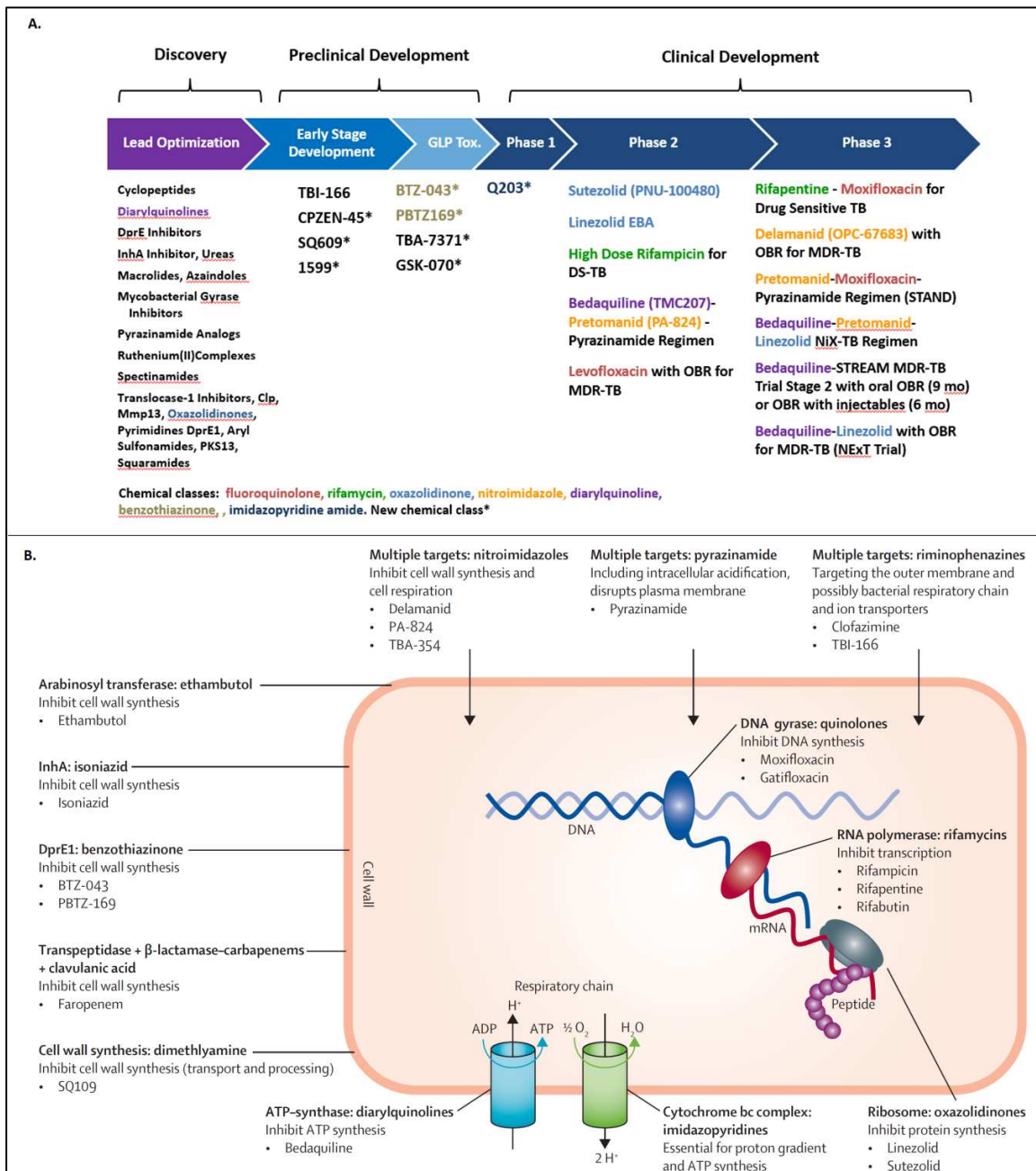
Figure 1.1. A. Current stages of the recent chemotherapy against *Mtb* adapted from the Working group on new TB drugs http://www.newtbdrugs.org/pipeline.php updated May 2016. B. Representation of *Mtb* cell and the targets of the already used therapy and some drug candidates. OBR: Optimized background regimen. DS-TB: Drug susceptible TB. MDR-TB: Multidrug resistant TB. Adapted from Wallis R et al, 2016 (34) and Zumla AI et al, 2014 (35), by permission of Elsevier.

Table 1.1. Description of first and second line drugs for TB treatment and their proposed targets and effects in *Mtb* (36, 37)

Antibiotic	Target	Effect	References
First line drugs			
Isoniazid	Enoyl-acyl-carrier-protein reductase (InhA)	Inhibits mycolic acid synthesis	(38)
Rifampicin	RNA polymerase, beta subunit (RpoB)	Inhibits transcription	(39)
Pyrazinamide	a. S1 component of 30S ribosomal subunit (RpsA). b. Aspartate 1-decarboxylase precursor PanD	a. Inhibits translation and trans-translation b. Inhibition of pantothenate and β-alanine biosynthesis	(40) (41)
Ethambutol	Arabinosyl transferases (EmbABC)	Inhibits arabinogalactan biosynthesis	(42)
Second line drugs			
Injectable agents			
Streptomycin	S12 and 16S rRNA components of 30S ribosomal subunit	Inhibits protein synthesis	(43)
Capreomycin	Remain to be determined (probably Interbridge between 30S and 50S ribosomal subunits or RplJ-RplL interaction)	Inhibits protein synthesis	(44)
Kanamycin	30S ribosomal subunit	Inhibits protein synthesis	(45)
Amikacin	30S ribosomal subunit	Inhibits protein synthesis	(45)
Fluoroquinoles			(43)
Ofloxacin Moxifloxacin Gatifloxacin	DNA gyrase (gyrA, gyrB)	Inhibit DNA synthesis	
Other oral agents			
Para-amino salicylic acid	Dihydropteroate synthase	Inhibits folate biosynthesis	(46)
Cycloserine	D-alanine racemase and ligase (Arl, Ddl)	Inhibits peptidoglycan synthesis	(47)
Ethionamide (thioamide)	Enoyl-acyl-carrier-protein reductase InhA	Inhibits mycolic acid biosynthesis	(38)
Agents with unclear efficacy			
Linezolid	50S subunit of the ribosome	Inhibits protein synthesis	(48)

The current standard treatment for TB is a combined therapy during 6 months that is divided in two phases: two months of four drugs (INH, RIF, pyrazinamide and ethambutol) in the intensive phase followed by 4 months of INH and RIF of continuation phase (17). This therapy is normally effective. The previously described drugs are grouped in the first line treatment option category and has been in place to treat TB patients since 1984 (16). The cure rates can be drastically reduced, however, after infection with a resistant *Mtb* strain. Resistant *Mtb* strains can cause complicated TB cases such as MDR TB, which are caused by *Mtb* strains resistant to INH and RIF. For this new form of TB, the treatment options extend to the second line drugs which include fluoroquinolones such as ofloxacin and moxifloxacin and injectable drugs such as amikacin, kanamycin, and capreomycin. Extensively drug resistant (XDR) and Totally Drug resistant (TDR) TB are forms of the disease very difficult to treat and cure. XDR describes the infection caused by those *Mtb* strains that in addition to being MDR, are resistant to one fluoroquinolone and at least one of three injectable second-line drugs. TDR is the most recently described and refers to MDR *Mtb* strains that are resistant to all second-line drug classes (i.e., aminoglycosides, cyclic polypeptides, fluoroquinolones, thioamides, serine analogues, and salicylic acid derivatives) (49).

1.4.1   Isoniazid and its relation with Mtb

After reviewing most of the anti-TB therapy developed up to date, it is really interesting to see that INH, one of the initial drugs used against *Mtb,* is still part of the first line drug therapy and is also used to prevent reactivation of latent TB cases (50, 51). That is the reason why after so many decades, the study of this old antimicrobial continues to be very relevant in the TB field. Additionally, INH is also effective against other Mycobacterial species such as *M. xenopi*, and *M. kansasii* but not effective against *M. avium or M. leprae* (52). INH is a synthetic compound

with the chemical formula $C_6H_7N_3O$ (137.14 g/mol) (Figure 1.2) that is freely soluble in water

and slightly soluble in alcohol and chloroform. Because of the urgency of an effective anti-TB

therapy and the way this drug was discovered in the 1950's, INH was used in patients without

knowledge of the mechanism of action.
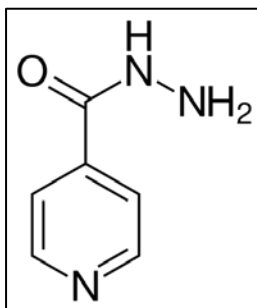
Figure 1.2. Structure of isoniazid (obtained from
https://commons.wikimedia.org/w/index.php?curid
=2781328)

Evaluating the pharmacodynamics, INH is a prodrug (similar to pyrazinamide, isoxyl-

thioacetazone, and ethambutol) that after entering the cell by passive diffusion interacts with the

bacterial enzyme KatG (catalase-peroxidase) to produce the isonicotinic acyl- nicotinamide

adenine dinucleotide complex or INH-NAD+ adduct (among other products) (16, 53-55). The

elucidation of INH mechanism of action was the subject of scientific research for almost 40

years (21). INH is a bactericidal drug only in dividing cells and under aerobic conditions (21).

Several studies have suggested that INH has multiple targets, but the enzyme InhA is the most

highly recognized (38). InhA is an enoyl acyl carrier reductase that participates in the synthesis

of mycolic acids which are key structures in the mycobacterial cell wall. Specifically, this

enzyme participates in the fatty acid synthase (FAS) II pathway that is responsible for the

elongation steps to generate these complex structures. Additional targets include other enzymes

in the FAS II route (such as KasA, AcpM and MabA) (56-58), depletion of nucleic acids (particularly nicotine adenine dinucleotide-NAD), and cell division, among others (16). INH is also associated with a greater generation of reactive oxygen intermediates inside the bacterium which are detrimental for its integrity (59-61). The internal structure and external morphology of *Mtb* are also altered after exposure to INH (21, 51).

During infection, *Mtb* has a variety of populations that have different responses to INH. For instance, bacteria in high growth rate are the most sensitive to INH, having a complete loss of their ability to synthesize mycolic acids as soon as 60 minutes after exposure to 0.5 µg/mL of the drug. The slower the bacteria growth rate, the less sensitive bacteria become to the drug to the point that dormant *Mtb* strains (metabolically inactive) are insensitive to INH. These particular populations are considered metabolically resistant to INH and are called "drug tolerant strains". In fact slow growth to non-growing state is a known strategy for the phenotypic bacterial persistence that is a fraction of the cell population able to tolerate the antibiotic. A recent work developed by John McKinney et al, demonstrated the presence of persistent *M. smegmatis* strains after exposure to INH, due to the intermittent release of KatG. These strains were tolerant to INH without having a genetic mutation and their progeny recover the susceptibility to INH after a second exposure to the drug (62).

Regarding INH pharmacokinetics, INH is administered to TB patients orally or parenterally, reaching its peak in plasma in one or two hours after an oral dose (with minimal binding to albumin) and is extensively metabolized in the liver by acetylation. INH can penetrate all tissues in the body and even the macrophage where it can exert its activity against intracellular bacteria. INH must be administered in combination with other anti-TB drugs as it has been demonstrated that 70% of *Mtb* isolates develop resistance in three months with INH

monotherapy (52). INH is excreted primarily in the urine as the intact drug or its biologically inactive metabolites acetylisoniazid, mono and diacetylhydrazine (63). In 2011, Mahapatra and colleagues demonstrated that a metabolite derived from INH-NAD adduct was also present in the urine of uninfected mice treated with INH, highlighting the possibility that this antimicrobial could also be activated by the host (54). Despite this, a close proximity of the active form of INH to the bacteria is required since, despite the activation of INH by eukaryotic peroxidases, the drug is still ineffective against *katG* mutant strains of *Mtb* during an in vivo infection (54). Because INH metabolism is primarily in the liver, hepatotoxicity is the major side effect but neurotoxicity and rheumatic complications have been also reported (52, 64).

1.5 *Mtb* drug resistance: the specific example of INH resistance

As is the case in other microorganisms, drug resistance in *Mtb* has either an intrinsic or acquired nature. *Mtb* cell wall structure and its low permeability are the major factors accounting for the high degree of intrinsic or natural resistance to many antibiotics and other chemotherapy. Highly abundant mycolic acids in the cell wall reduce the cell permeability and create a crystalline-like structure after the cytosolic membrane. As seen in other mycobacterial species (especially in saprophytic species such as *M. chelonae)*, the more impermeable the cell wall, the more antimicrobial agents the mycobacteria can resist. Drugs such as sulfonamides, penicillin, tetracycline, and vancomycin are ineffective against *Mtb* (65)*.* Additionally, the reduced number of porins in the *Mtb* pseudo-outer membrane possibly contributes to the intrinsic *Mtb* resistance against hydrophilic compounds. Among other intrinsic factors, *Mtb* possess β-lactamase enzymes (encoded by *blaC* and *blaS*) that make this bacterium naturally resistant to β-lactams (66, 67). For acquired drug resistance, spontaneous mutations in chromosomal genes during a sub-optimal drug therapy are the most common cause for drug resistance in *Mtb*. Efflux

mechanisms are less common but also present in these bacteria (66). These intrinsic and acquired mechanisms have synergistic effects and make TB treatment particularly cumbersome.

Although a combined therapy for TB is normally effective for most cases, TB cases resistant to a subsection or all of these drugs have been reported in clinical settings. INH resistance is one of the most common forms of drug resistant TB. The appearance of INH resistant cases started the same year this drug was introduced in medical practice (21). For the purpose of the present work, isoniazid resistance will be used as a model to describe the major acquired resistant events that have been identified in *Mtb*.

INH resistance is multigenic and can be divided in three categories: prevention of drug activation, alteration of the target, and differential expression of the target. In the first group, mutations in *katG* that prevent the activation of INH are present in the majority of resistant cases to this drug (16, 68). KatG function was first correlated to INH resistance in 1953, when Middlebrook and colleagues discovered that INH resistant *Mtb* strains lack catalase-peroxidase activity and were less virulent in guinea pigs (69). The molecular validation of this observation was completed later by Zhang and colleagues, restoring the sensitivity to INH in some *Mtb* resistant strains after the introduction of the *katG* gene from *Escherichia coli* (70). Currently, there are more than 300 known mutations in the *katG* gene alone with a wide range of minimum inhibitory concentrations (0.2 to 256 mg/L)(16). These mutations include missense mutations, insertions, deletions, truncations and full gene deletion. Depending on the position and nature of the mutation, *katG* mutants have different degrees of catalase-peroxidase activity (16, 66). The outcome of these mutations is the lack of the active drug, because of the absence of its activator. The mutation rate for the generation of INH-resistant strains is around 3.2×10^{-7} mutations/cell

division (after exposure to 1 mg/L INH) *in vitro* (71-73) and presumably one in 10^{8-9} organisms *in vivo* (51).

In the category of alteration of the target and differential larger expression of the target, mutations in the *inhA* gene or its promoter are accounted. InhA is the most commonly validated target for INH (21, 74). Currently, around 15 mutations in the *inhA* gene have been identified in *Mtb* strains with low-level resistance to INH. Mutations on *inhA* also drive resistance to ethionamide, since INH and ethionamide share this enzyme as target (16).The most studied mutation is the S94A that results in the reduction of the enzyme affinity for NADH and a reduced ability of INH-NAD adduct to inhibit the enzyme. Additionally, mutations in the *inhA* promoter that increase the InhA levels have been also identified. Therefore, both the reduction in enzymatic activity and the overexpression of the target serve as resistance mechanisms against INH. Other mechanisms of INH resistance include the accumulation of NADH (by redox alteration) that binds InhA and protects it from the inhibitory effect of the INH-NAD adduct. An additional resistance mechanism includes acetylation of the drug by the *nat* encoded arylamine N-acetyl-transferase which prevents INH activation by KatG. Finally, the drug efflux mechanisms include the participation of the protein EfpA which is induced upon INH treatment (21). It is important to describe also that INH resistance in *Mtb* can be either low or high-level when there is >1% of bacterial growth in the presence of 0.2 ug/mL or 1 ug/mL of INH, respectively. Regularly, mutations in the *inhA* promoter are linked to low-level of INH resistance while mutations in *kaG* are associated with high-level of INH resistance in *Mtb* (75).

1.6 A brief review of *Mtb* physiology

1.6.1 Generalities of the microorganism and basic detection methods in the lab

Mtb belongs to the Actinomicetales order and the Mycobacteriaceae family. *Mtb* is a non-motile bacillus that is able to synthesize all of its metabolites, including aminoacids, vitamins and cofactors from inorganic sources (67); this ability classifies *Mtb* as a prototrophic organism. The size of the bacteria varies from 2-4 µm in length and 0.3-0.5µm in width, with a generation time of 15-20 h at 37°C. *Mtb* is a facultative intracellular and non-spore-forming bacterium, although it can have a persistent state with low metabolic profile activity (15, 76). Because of the structure of its cell wall (which will be discussed later in this Chapter), *Mtb* is similar to a Gram-negative bacteria with a pseudo-outer membrane, but is phylogenetically linked with Gram-positive bacteria based on its 16S ribosomal RNA sequence.

Because of the fatty acid composition of the *Mtb* cell wall, specific staining is needed to visualize this acid-alcohol resistant bacterium. Here, options include Ziehl-Neelsen (ZN) or Kinyoun stain to observe the bacilli under a light microscope and auramine rhodamine (AR) stain using fluorescent microscopy. Smear microscopy is the most widely used strategy to diagnose active pulmonary TB cases, particularly in low income countries. This technique detects the bacteria in the sputum with 50-60 % sensitivity (via ZN staining). Although AR sensitivity is 10% higher compared to ZN, AR is only used as a screening tool because of its lower specificity (77). Auramine O and carbol fuchsine (the basic components of ZN staining) have the ability to bind to the cell wall, generating the fluorescence or the red-fucsia color, respectively.

Currently, the most sensitive method to detect *Mtb* is culturing the infected sample (including sputum, broncho-alveolar lavage, among others), which can take up to 12 weeks of incubation in paucibacillary specimens (samples with low bacterial burden) (78, 79). Culturing is also known as the reference method for TB detection and it can be done either on solid (Löwenstein-Jensen, Middlebrook 7H11 or Ogawa media) or liquid culture media (mycobacterial growth indicator tube BACTEC™ MGIT™, BD or BacT/ALERT® System, bioMérieux). Molecular technologies are being implemented in recent years including those that are RNA, or more commonly DNA, based. These include commercially available kits such as Xpert MTB/RIF (Cepheid), INNO-LiPA Rif.TB (Innogenetics), GenoType® MTBDRplus assay (Hain LifeScience) or in-house developed PCR approaches (77, 80-84). Some commercial kits can also provide information about drug susceptibility status of the *Mtb* strain. These methodologies, however, have shown controversial results particularly in the accuracy of drug susceptibility profile of the bacteria, such as the Xpert MTB/RIF which is highly used in clinical settings. This is because of uncommon mutations (not present in the detection kits) that do generate drug resistant *Mtb* strains but are erroneously classified as drug susceptible *Mtb* strains by the commercial kit. False resistant reports can be also generated due to mutations that had an intermediate MIC value or susceptible phenotype. This last one creates more confusion in the clinical management of these TB cases and increases the length of treatment and toxicity unnecessarily. Additionally, one of the most highly used DNA based-tools, Xpert MTB/RIF, is not able to detect monoresistant cases to INH, as only detect rifampicin resistance-associated mutations.

1.6.2. Mtb Genome

The genome of the *Mtb* reference strain H37Rv was fully sequenced by Cole and colleagues in 1998 (67, 85). The genome size is about 4.4 Mbp with around 4,000 predicted proteins which are highly devoted to metabolism, particularly lipid metabolism (15, 67). The *Mtb* genome has a high G+C content (62-70%) with little or no incorporation from other bacterial genera. After complete sequencing, the protein coding sequences received a number of four digits that start with the prefix Rv. Only about 15% of the genes could not be related to any known gene or protein function. Nearly 27% showed similarity to putative proteins and were referred to as conserved hypotheticals.

The *Mtb* genome has a number of repetitive genetic elements that have been used for typing methodologies with epidemiological and genetic evolution studies purposes. These are insertion elements (IS) which have been used in molecular techniques such as RFLP to classify *Mb* strains. Other methodologies include spoligotyping (based on the presence absence of Directly Repeated (DR) elements), Mycobacterial Interspersed Repetitive Units (MIRU) of 12, 15 or 24 loci, up to Whole Genome Sequencing (WGS) (86-88). These techniques have allowed the investigation of outbreaks, relationship between cases, evolutionary trends of *Mtb* strains and human populations, and differentiate *Mtb* strains from *M. bovis* among other applications (88, 89).

Mtb DNA repair and mutation rate are constantly evolving and therefore interesting topics for study. *Mtb* lacks the conserved proteins involved in mismatch repair (MutS and MutL normally found in *E. coli*) (90), but has defined strategies to overcome different DNA induced damage (either by hydrolysis, alkylation and oxidation) encoded by the genes *ung, mutY, nth,*

fpg, nei, tag1, xthA and *nfo* (90). Mismatch repair is strand specific and corrects mutations that

arise during DNA replication and facilitates intra-chromosomal exchange for instance between

tandem repeats or insertion elements generating genetic variation among strains (90). According

to this, it could be hypothesized that *Mtb* have a higher probability of heritable defects, but

instead, the deficiency of this repair system has been associated with a low degree of

microsatellites polymorphism and frameshift mutations (91). Despite the absence of DNA

mismatch repair proteins, *Mtb* has also been recognized by its low *in vitro* mutation rate

(probability of a mutation occurring per cell division) compared to other human pathogens (such

Plasmodium or HIV) (71, 92). On the other hand, there is a relatively high genetic variation

among clinical strains (different genetic lineages) as well as a significant number of drug induced

resistant strains (with their associated SNPs). As it has been previously discussed, the mutation

rate can increase as a result of external stresses which includes antimicrobial pressure, oxidative

stress, pH variation, among others that can occur in the host (71). Therefore, this field is

expected to continue evolving looking deeper at *in vivo* mutation rate that, as it was shown in the

macaque model, contradicts previous expectations about mutation rate in *Mtb*. This model

demonstrated a similar *Mtb* mutation rate during active or latent infection (92).

1.6.3 Mtb envelope

Moving to another important aspect of *Mtb* physiology, the cell envelope of this

bacterium has been the focus of research for many decades because of its distinct features,

importance in bacterial pathogenicity and the generation of immune response in the host. The

mycobacterial cell envelope is complex such that nutrients penetrate 10,000 times slower than

they can do in the *E. coli* outer membrane (65). Components of the cell envelope, particularly the

enzymes that participate in their synthesis, have been recognized as possible drug targets. The

understanding of the cell envelope is also required to design drugs that will be able to cross this impermeable barrier efficiently (13).

The *Mtb* envelope forms the interface between host and the bacteria. From the outside to the inside, the *Mtb* cell envelope is composed of a layer of non-covalently linked glycolipids, proteins, carbohydrates and some lipids (the capsule), a covalently linked peptidoglycan layer that contains carbohydrates and lipids (the cell wall), and a plasmatic membrane (phospholipid bilayer). In 1991, Minnikin proposed visualizing the lipid material in the *Mtb* envelope as two distinct membranes, analogous to a Gram-negative bacterium (93).

The most external layer of *Mtb* has been called the "capsule" by some authors. This layer contains hydrophilic carbohydrates and 2-3% lipids. The carbohydrates present in this external layer are arabinogalactan and arabinomannan which can be also secreted to the extracellular space (13). Glycolipids such as trehalose monomycolate (TMM) and trehalose dimicolate (TDM); phenolic lipids and glycopeptidolipids can be found in the outer part of the capsule and some of them are also cell wall-associated. TDM is also known as cord factor since it causes grow in "cords" in vitro. This particular glycolipid has been associated with the pathogenesis and immunogenicity of *Mtb* strains (94). Lipoproteins such as LpqH (Rv3763), proteins such as Psts1 (Rv0934) and the Ag85 complex (FbpA, Rv3804; FbpB, Rv1886; and FbpC, Rv0129) are also commonly found in the capsular material (13).

The *Mtb* cell wall has a covalently linked backbone with a collection of cell wall-associated lipids and polypeptides. The covalently linked molecules include peptidoglycan, arabinogalactan and mycolic acids. In addition to the presence of the last two biomolecules, there are two important hallmarks of the *Mtb* cell wall. First, the muramic acid is N-acylated,

instead of N-acetylated as regularly observed in most eubacteria. Second, there are unusual cross-links between two chains of peptidoglycan that include bounds of two residues of diaminopimelic acid in addition to the usual D-alanyl-diaminopimelate linkage. Furthermore, mycolic acids represent about 40% of the cell wall (95, 96).

Mtb has a great variety of lipids that can be grouped in categories that were deeply explored in the reference strain H37Rv during *in vitro* growth by Sartain and colleagues. They were able to identify 6 lipid categories with around 2,512 lipid groups (97). However, for the purpose of this dissertation, only mycolic acids, which are categorized in the fatty acyl group, will be further explored.

Mycolic acids (MA) are the major constituent of the cell envelope. They were first named by Stodola and collegues in 1938, who also depicted basic groups of their chemical structure. MA structure was further defined by Asselineau in 1950 (98). These are α-alkyl, β-hydroxy, long chain fatty acids that can be primarily covalently attached as esters of arabinogalactan in the cell wall or as "free lipids" in the capsule associated to trehalose in the TMM or TDM structures (99). Specifically, MA form an ester bound to the 5-position of the arabinose residue of the arabinogalactan (100). MA can also bind to glucose (101). The covalently attached MA can be obtained by saponification or methanolysis of the cell wall of the delipidated *Mtb* cells. Because MA are not soluble in methanol, they can be separated from moderately long chain fatty acids with ether or chloroform solutions. MA have one carbon chain bound to the hydroxyl group called the meromycolic chain and another (shorter) carbon chain that is bound to the α-carbon. These chains are synthesized in the cell separately and then condensed by a Claisen-type reaction (13).

MA are not unique structures of *Mycobacterium* genera, they can be present in

Corynebacterium, Nocardia, and *Rhodococcus*. MA from *Mycobacterium* are longer in carbon

units (C_{70} to C_{90}) and have the largest meromycolic chain (96). Additional modifications such as

the introduction of cyclopropane rings in the meromycolate chain, unsaturations, ethylenic

groups, and methyl branches are also observed. Both *cis* and *trans* double bounds as well as

cyclopropane rings can be found in the same type of mycolate. Some MA have additional

oxygen functionality that is one feature used to classify them. These functionalities are keto,

methoxy, carboxy and epoxy. Other types of MA lack of these oxygen groups, they are called α-

MA and α'-MA. α and α'-MA differ in their chain length, α'-MA are shorter (usually of 60

carbon units) while α-mycolic acids contains more than 70 carbon units. α-MA represent more

than 70% of the total MA found in *Mtb*, followed by keto and methoxy MA (15% and 10%)

(102). The cyclopropane structures in MA has been shown that contributes not only to its cell

wall structure, but to protect the bacteria from oxidizing agents such as H_2O_2 (Takayama).

Finally, the plasmatic membrane includes different types of phospholipids such as

phosphatidylglycerol, phosphatidylethanolamine, phosphatidylinositol and phosphatidylinositol

mannosides (PIMs). PIMs are mainly located in the outer leaflet. Other important components

are lipoarabinomannan (LAM) and lipomannan (96). Due to the high abundance of this

glycolopid in the *Mtb* envelope, LAM has been tested as a biomarker for a point of care test with

a wide range of sensitivity and specificity results in HIV-positive patients (103).

1.6.4 Major central metabolic pathways in Mtb

Similar to the variable composition of the *Mtb* envelope, this bacterium also has the

ability to use very variable carbon sources *in vitro*, such as carbohydrates, alcohols, and lipids

(including cholesterol and fatty acids) (104). Similar to other representative species of the

Actinomycetales order, *Mtb* possesses a predominant aerobic metabolism, with the genes

encoding for enzymes of the main energetic metabolic pathways such as glycolysis,

Tricarboxilyc Acid (TCA) cycle and pentose phosphate pathway. Additionally, this

microorganism also has the glyoxylate shunt which allows the bacteria to bypass some enzymes

of the regular TCA cycle under specific metabolic conditions (hypoxic conditions or growing

mainly in fatty acids) (67). Isocitrate lyase (Icl) plays an important role in this alternative to the

TCA pathway and this enzyme has been associated with virulence.

Despite genetic evidence of a complete TCA cycle in *Mtb* (67), there is no sufficient

biochemical evidence to show the presence of all the enzymatic reactions of the TCA cycle in

Mtb. In fact, some researchers propose that TCA cycle in *Mtb* is not complete because this

organism lacks alpha ketoglutarate dehydrogenase (α-KDH) (105-107). Therefore, an alternative

route or enzyme catalyzing the step from alpha ketoglutarate (α-KG) to succinate was explored.

Alpha ketoglutarate decarboxylase (α-KGD, Rv1248c) and succinic semialdehyde

dehydrogenase (GabD1/ 2, Rv0234 and Rv1731 respectively) were proposed as the enzymes that

complete this reaction in the *Mtb* TCA cycle under normoxic conditions. A-KGD catalyzes the

production of succinic semialdehyde, which can then be converted to succinate by GabD1/2

(105). Under anaerobic conditions, Icl activity together with α-KG ferredoxin oxidoreductase are

believed to complete this reaction (108). Also, experimental evidence suggests that *Mtb* operates

a reversed TCA cycle with the reduction of fumarate to succinate to maintain the membrane

potential in the absence of oxygen.

Mtb has the ability to use proteins in multiple metabolic pathways to prolong its survival,

a feature that is known as metabolic plasticity. For instance, Icl participates in the glyoxylate

shunt and the methyl citrate cycle, plus also protects *Mtb* from the oxidative stress generated by the treatment with isoniazid, rifampicin, and streptomycin (109). Another example is the dihydrolipoamide dehydrogenase (Lpd) that can act as the E3 component of the pyruvate deshydrogenase or can provide electrons to the dihydrolipoamide succinyltransferase (DlaT, previously known as SucB) or be part of the branched-chain keto acid dehydrogenase complex to metabolize branched-chain amino acids (109). On the other hand, *Mtb* has pathways with redundant enzymes (including isozymes) that can catalyze the same reaction, which guarantee that vital processes occur despite possible external or internal stresses. A good example of this is the fatty acid degradation or β-oxidation pathway, which suggest that *Mtb* not only has a high lipid catabolism activity, but also that this is crucial part of its metabolism (104).

1.6.5 Lipid metabolism: β oxidation and fatty acid synthesis

Lipid metabolism is a higly relevant physiologic process in *Mtb* which is demonstrated by the amount of the genome devoted to these reactions. Specifically, lipid metabolism is an important part of this dissertation as some enzymes of the biosynthetic metabolic pathway are the target of INH. Compared to *E. coli, Mtb* possess five times more enzymes dedicated to lipid metabolism. In fact, *Mtb* lipid metabolism is more lipolytic than lipogenic, probably as a result of the wide variety and amount of lipid sources in the human host as well as in the bacterial envelope (67). For this reason, fatty acid degradation is going to be described first and a subsequent description of the major and more recent findings regarding the fatty acid synthesis will be explored. Fatty acid degradation is a key process in *Mtb* metabolism and can explain more about its metabolic plasticity, while fatty acid synthesis is crucial in the understanding mechanism of action of INH and possible the related events in the INH resistance acquisition in *Mtb*.

Fatty acid degradation: This is a process of successive oxidations where the β carbon of the fatty acid is oxidized to a carbonyl group (figure 1.3). In this process, the main goal is the synthesis of acetyl-CoA and reduced cofactors (such as NADH, $FADH_2$) that can fulfill energy requirements in the cell and also intermediates that can serve as substrate for anabolic processes (67). Specifically, odd-chain fatty acids produce acetyl CoA while even-chain fatty acids produce acetyl-CoA and propionyl-CoA in addition to acyl-CoA derivatives with two less carbon units that are produced in both cases (67, 110). By studying the *Mtb* genome, Cole and colleagues found 36 genes encoding for enzymes that catalyze the first step of fatty acid degradation only. As shown in figure 1.3, most reactions in this pathway can be carried out by several isozymes. However, there are two enzymes, EchA5 and FadB3, that are essential for *Mtb* growth and considered possible drug targets (111).

Fatty acid synthesis: The complexity of *Mtb* lipids could be partially explained by the fact that *Mtb* has both fatty acid synthase (FAS) type I and II. Cole and colleges described the main enzymes of FAS I and II at the genetic level and recent reviews have compiled previous biochemical work, all of which have generated a better understanding of the complex pathways responsible for the MA synthesis in *Mtb* (99, 102, 112). FAS I is found mainly in eukaryotes and all the reactions are performed by a single multidomain homodimeric enzyme Fas (Rv2524) (67, 102). This enzyme has seven catalytic domains: acyltransferase, enoyl reductase, dehydratase, molonyl/palmitoyl transferase, acyl carrier protein, β ketoacyl reductase, and β ketoacyl synthase (99). Fas (Rv2524) uses acetyl-CoA and malonyl-CoA as substrate for the synthesis of acyl-CoA derivatives of 16 and 18 carbon units that are used for the synthesis of membrane phospholipids. FAS I route also produces an acyl-CoA derivative with 26 carbon units that becomes the short α-alkyl chain or α-branch of the mycolic acids. FAS I and II are connected by the synthesis of acyl-

CoA derivatives with 20 carbon atoms that are used in the FAS II pathway as the starting

molecule for the elongation of MA (102) (figure 1.4).
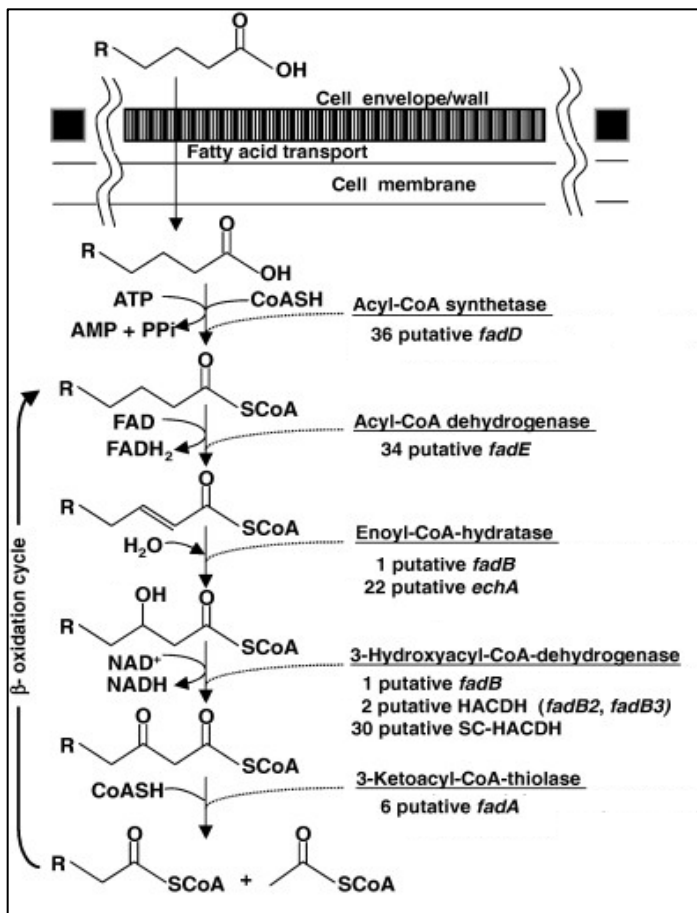
Figure 1.3. ß oxidation of fatty acids and the enzymes in *Mtb*. Adapted from Muñoz-Elias E and
McKinney J, 2006 (110) by permission of John Wiley and Sons.
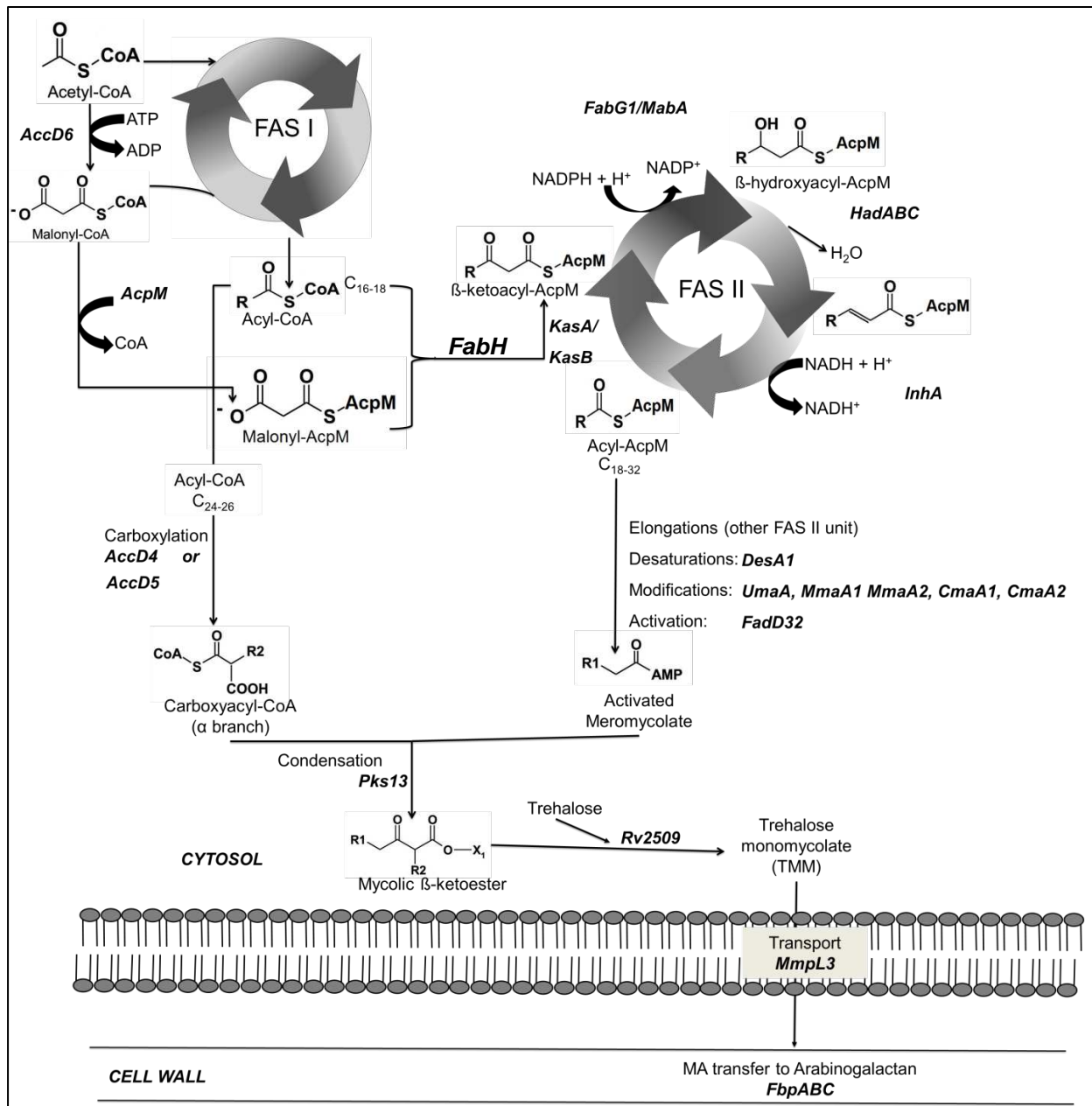
Figure 1.4. Reactions of the FAS I and FAS II describing the enzymes (or complexes) that participate in each reaction with the intermediate products. Adapted from (112), by permission of Elsevier.

There are important aspects to highlight regarding FAS I and II in *Mtb*. First, proteins FabD, AcpM, and FabH act in the transition between FAS I and FAS II, generating ACP derivatives (the substrate required for the FAS II pathway). Second, there are two known

30

Claisen-type reactions occurring: one before the FAS II starts (responsible for the condensation

of malonyl-ACP with acyl-CoA and catalyzed by FabH) and one shared with the polyketide

synthase system (catalyzed by Pks13). The latter reaction generates a carbon-carbon bond

between two activated fatty acids at the end of the MA synthesis (Figure 1.4, Table 1.2). This

second condensation takes the α-branch (produced through FAS I) and the longer meromycolate

chain (produced through FAS I and II) to form a "pre-mature MA".

Regarding FAS II specifically, this pathway is involved in fatty acid elongation instead of

de novo synthesis (contrary to what occurs in most bacteria, where FAS II has *de novo* synthesis

capacity) (99). *Mtb* needs to use both FAS I and II to generate its characteristic mycolic acids

(102, 112). Therefore, the study of mycolic acids synthesis is in fact a study of both FAS

pathways in *Mtb*. In FAS II, there is one different enzyme for each specific step, allowing for

various levels of regulation. Most of the core enzymes of FAS II are NADPH or NADH

dependent and organized in different clusters distributed through the genome (Figure 1.5, Table

1.2). FAS II can be further divided in type I and type II elongation (E1 FAS II and E2 FAS II

respectively). Here, both types are catalyzed by the core proteins InhA, MabA, HadABC and

FabD, and elongation can be done by either KasA (E1) or KasB (E2) (Figure 1.4). Despite the

sequence homology between the condensases KasA and KasB, they are predicted to participate

in two different stages during the FAS II pathway: KasA may catalyze the first elongation steps

(E1-FAS II) while KasB might be involved in the later steps (E2-FAS II), ultimately producing

full-length mycolates with more than 40 carbon units (99, 112). A representation of matured α-

MA is depicted in Figure 1.5B.

The meromycolate chain resulting from FAS II cycle can be "decorated" with chemical

modifications such cyclopropanations and methylations that are introduced before the second

Claisen-type reaction occurs. These modifications can be at distal or proximal positions and are carried out by S-adenosyl-methionine (SAM)-dependent methyl transferases (Table 1.2, Figure 1.4 and 6). Unsaturations on the other hand, are proposed to occur differently under aerobic or anaerobic conditions. The method of double bond introduction in MA in *Mtb*, however, remains unclear. Under aerobic conditions, desaturases encoded by *desA1, 2* and *3* and other candidates such as Rv1371 are believed to complete the double bond introductions at the distal position, before the Clainsen-type condensation take place. Under anaerobic conditions, unsaturations are believed to take place during the FAS II cycle in the transition of the *trans* 2-enoyl intermediate to its 3-*cis* isomer under in the distal position, resembling what FabM does in *Steptococcus pneumoniae* (112). By sequence homology, this enzymatic reaction could be mediated by EchA10 and EchA11 in *Mtb*, however, there is not enough experimental evidence to support this hypothesis (Table 1.2). Finally, the oxygenated MA (keto and methoxy MA) have a common precursor (hydroxymycolate) that is synthesized by the action of the SAM-dependent methoxy mycolic acid synthase (MmaA) 4. The synthesis of methoxy MA is additionally driven by the MmaA3 enzyme (Figure 1.4, Table 1.2) (102, 112).
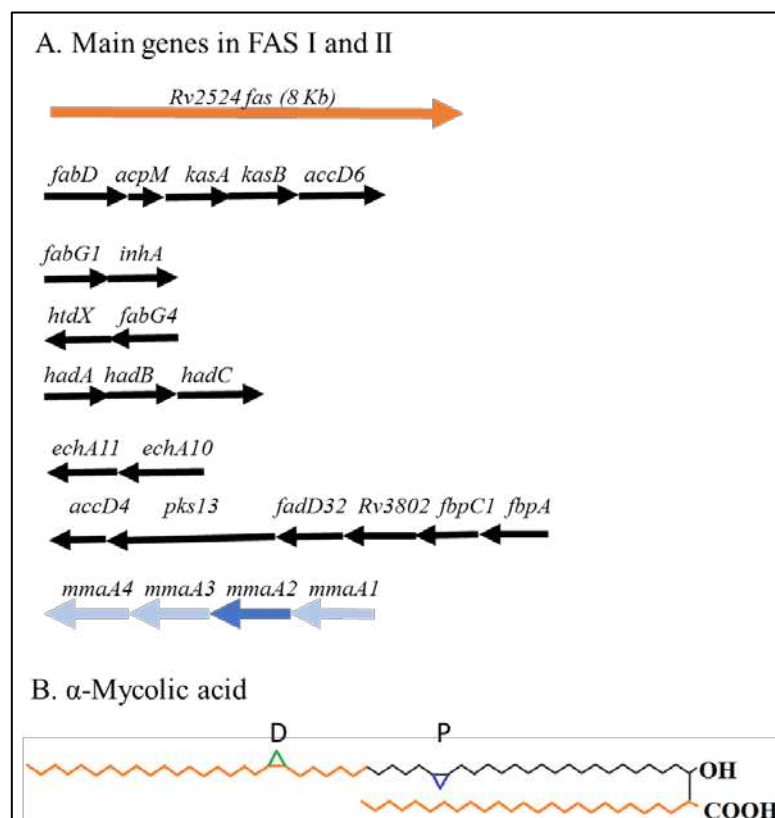
Figure 1.5. A. Major operons involved in mycolic acid synthesis in Mtb. B. Structure of an alpha-MA, the color represent the source of the carbon chain by either FAS I (orange) or FAS II (black). P: Proximal, D: Distal.

Table 1.2. Enzymes that partipate in the FAS I and II pathways in *Mtb*

Description	Gene	Rv number	Enzyme
FAS I	*fas*	2524	Fatty acid synthetase
Transition FAS I to FAS II	*fabD*	2243	Malonyl-CoA ACP transacylase
	accD6	2247	Acetyl/propionyl-CoA carboxylase (beta subunit)
	acpM	2244	Acyl carrier protein
	fabH	0533	β-ketoacyl-ACP synthase III
FAS II	*kasA/B*	2245/2246	β-ketoacyl-ACP synthase
	fab1 or MabA	1483	β-ketoacyl-ACP reductase
	hadA/B/C	0635/0636/0637	(3)-hydroxyacyl-ACP dehydratase subunit A/B/C
	htdX	0241	3-hydroxyacyl-thioester dehydratase

	echA10/11	1142/1141	Currently annotated as a enoyl-CoA hydratase*, but proposed to be 2-*trans*-enoyl-ACP isomerase[a]
	inhA	1484	2-*trans*-enoyl-ACP reductase
Modifications			
Desaturases	*desA1/2/3*	0824/1094/3229	Acyl-carrier protein desaturase
Methyltransferases	*mmaA1*	0645c	Methoxy mycolic acid synthase 1
(cyclopropanation,	*mmaA2*	0644c	Methoxy mycolic acid synthase 2 (distal cyclopropane in α-MA, proximal *cis*-cyclopropane in keto-MA)
methylation and	*mmaA3*	0643c	Methoxy mycolic acid synthase 3 (oxygenated MA)
oxygen function	*mmaA4*	0642c	Methoxy mycolic acid synthase 4 (oxygenated MA)
introduction)	*cmaA1*	3392c	Cyclopropane-fatty-acyl-phospholipid synthase 1 (distal position)
	cmaA2	0503c	Cyclopropane-fatty-acyl-phospholipid synthase 2 (proximal position-specific in methoxy-MA)
	pcaA (umaA2)	0470c	Mycolic acid synthase (proximal cyclopropanation function α-MA)
	umaA	0469	Mycolic acid synthase
Clainsen-type condensation	*accD4*	3799c	Acyl-CoA carboxylase
	accD5	3280	Acyl-CoA carboxylase
	fadD32	3801	Fatty-acid-AMP ligase
	pks13	3800	Polyketide synthase-13
Mycolic acid processing	*mmpL3*	0206	Transmembrane transport protein-3
	Rv3802	3802	Proposed to be a Mycolyltransferase I[a], recently shown to have phospholipase and thioesterase activity[b]
	cmrA	2509	Reductase
	fbpA/fbpB/ fbpC2	3804c/1886c/ 0129c	Fibronectin-binding protein ABC or antigen 85 complex

After the modification in the meromycolate chain and the last condensation reaction

occur, a MA (either α, keto or metoxy-MA) molecule is formed and can be attached to a

trehalose molecule by the action of the reductase encoded by Rv2509 also known CmrA (113).

Another fate of MA is to be covalently linked to arabinogalactan in the cell wall, a reaction that is catalyzed by the Fibronectin-binding proteins (Fbp) ABC (114) (Figure 1.4). Much of the understanding of the FAS I and II routes has been based on sequence homology with reference bacterial strains and mutation analysis using model organism such as *M. smegmatis and M. phlei(112)*. Despite the vast knowledge about the MA synthesis pathway, many unanswered questions remain regarding components of the FAS II pathway that are under current research (112).

1.6.6 Redox metabolism

In general, reduction–oxidation (i.e., redox) reactions are highly relevant for *Mtb,* since they not only comprise the necessary of defense mechanisms developed by the host during the bacterial infection, but they are also part of its own metabolism. Redox reactions could generate endogenous or exogenous stress for the bacteria. The endogenous redox stress is generated during aerobic or anaerobic respiration, where *Mtb* is exposed to reactive oxygen (ROI) and nitrogen reactive intermediates (RNI), generated when the bacterium uses oxygen and nitrogen as the final electron acceptor in the electron transport chain. RNI can be also generated when *Mtb* relies on glutamate metabolism for survival. During host-infection, *Mtb* can experience a wide range of oxygen levels that can drastically alter its metabolism going from hyperoxic stress (when is in aerosol droplets) to low oxygen tension (during the intracellular phase in alveolar macrophages) to finally hypoxic to anoxic stress (in granulomas). Additionally, inside the macrophage, *Mtb* is exposed to both ROI and RNI. H_2O_2 and the superoxide radical (O^{-2}) are the two most common ROI forms that are produced by macrophages and neutrophils to eliminate *Mtb*. During hypoxic conditions, the alteration in redox homeostasis leads to a higher NADH/NAD+ ratio which generates superoxide radicals that disrupt the redox balance in the

cell. Consequently, enzymes with heme and sulfur complexes (i.e., cytochrome C, aconitase) can be severely affected. Therefore, the ability of *Mtb* to survive the redox stress from the host would determine its success during the infection process. This stress has an impact on the bacterial metabolic pathways as well as on the expression of virulence factors (115).

Intracellular or exogenously originated ROS and RNI have the potential to damage lipids, DNA and proteins by oxidation, peroxidation and nitration reactions (116), that can result in protein inactivation, and alteration in both cell organization and signal transduction. Therefore it is crucial to successfully maintain redox homeostasis to keep the integrity of the cell. Intracellularly, the changes in the redox and nutrient levels are sensed by WhiB proteins (WhiB1-7) while extracellularly different molecules such as nitric oxide (NO), carbon monoxide (CO), hydrogen peroxide (H_2O_2) and the reduced and oxidized forms of the nicotinamide adenine dinucleotide (NADH/NAD+) can work as sensors that induce a direct transcriptional response or a two-component regulatory system DosRS-DosRT. (67, 115). Moreover, different bacterial enzymes participate in the neutralization of the host induced ROI and NOI such as superoxide dismutase (SodA and B), catalase-peroxidase (KatG), and the antioxidant complex formed by alkyl-hydroperoxidases (AhpC and AhpD), dihydrolipoamide acyltransferase (DlaT) and dehydrogenase (LpdC). Other enzymes such as peroxiredoxins (AhpE, TPx, Bcp and BcpB) and thioredoxins (TrxA, B and C) are also accounted.

Since KatG function is playing a central role in this study, a further description of this enzyme is required. *Mtb* has only one single copy of *katG* with a coding sequence of 2,223 base pairs (bp) that generates a protein with 704 amino acid residues that has a molecular weight of 80.6 KDa approximately. KatG is present as a dimeric hemoprotein that belongs to class I peroxidase superfamily, because of its high homology with yeast cytochrome c peroxidase (117).

The role of KatG in *Mtb* is beyond the activation of the prodrug INH. This enzyme is in fact one of the most important catalase-peroxidases that help the bacterium overcome external and internal redox stress. KatG possesses a monofunctional catalase, broad-spectrum peroxidase and peroxynitritase activity (118, 119). The catalase-peroxidase activity is in the N-terminal domain of the protein that contains a heme-binding motif, however, the C-terminus is also required for its catalytic function (55, 117). KatG activity has been associated with virulent *Mtb* strains, which are able to infect for longer periods and create more pathology in the host compared with mutant strains lacking of the enzyme activity (120-122).

As discussed above, redox reactions play an important role in bacterial respiration. In the next section, details about the cellular respiration process in *Mtb* will be covered. The relation with this topic and this dissertation arises because mutations in important genes involved in redox homeostasis such as *katG* gene (also INH activator) possibly generate an alteration in the respiration complexes in *Mtb*.

1.6.7 Respiration in Mtb

Due to the *Mtb* lifestyle, respiration in the bacterium should be highly adaptable. Specifically, during respiration, *Mtb* uses oxygen and other compounds (such as fumarate or nitrate) as the final electron acceptor depending on the specific bacterial metabolic status and the surrounding environment (123, 124). The respiratory apparatus is responsible for generating ATP and regenerating reducing equivalents. Respiration is made possible by select membrane-associated asymmetric complexes that allow for generation of proton motive force (PMF) and ATP, which are the major sources of energy in the cell. Different from other model organisms (such as *E.coli* or *Bacillus subtillis), Mtb* obtains the majority of its ATP by the electron transport

chain and the F_1F_0-ATP synthase machinery, with very little contributions from substrate level phosphorylation (125). In fact, the ATP synthase is a recently exploited target for developing anti-TB drugs of the drug class diarylquinolines (such as the clinical drug bedaquiline) (123, 125). Specifically, diarylquinolines interact with the transmembrane subunit C of the ATP synthase machinery (126). This again emphasizes the importance of ATP synthase machinery in the respiration process in *Mtb*.

Most of the *Mtb* enzymes/complexes involved in aerobic respiration have been identified and are composed primarily of two NADH dehydrogenases (NDH-1 and NDH-2) and two terminal cytochrome oxidases (aa_3-type cytochrome c oxidase and bd-type cytochrome oxidase). These participate in oxygen reduction and are coupled to generate the PMF that is used by the ATP synthase for the production of ATP. NDH-1 is encoded by the *nuo* operon (*nuoA-N)* and NDH-2 is present in two copies encoded by *ndh* and *ndhA*. Previous studies demonstrated that NDH-2 does not have a proton-translocating-function and is the main dehydrogenase in *Mtb*. NDH-2 reduces menaquinone to menaquinol that in turn can be oxidized by one of the terminal aa_3-type cytochrome c oxidase and bd-type cytochrome oxidase complexes. Because the bd-type cytochrome oxidase (CytA-B) is not coupled to proton pumping, the direct oxidation of menaquinol by this oxidase is less energetically efficient compared to the aa_3-type (CtaC-F). Instead, the oxidation of menaquinol can happen in a two-step process with the participation of the cytochrome bc1 complex (QcrA-C) and the terminal aa3-type cytochrome oxidase (CtaC-F) with a higher energy yield (123, 125).

Contrary to aerobic respiration, mediators in *Mtb* anaerobic respiration are poorly defined. However, in vitro hypoxic studies have allowed the identification of some important enzymes involved. In a reduced-oxygen environment, the nitrate reductase (NarG-I), the nitrate

transporter (NarK-2) as well as the NDH-2 dehydrogenase are upregulated. On the other hand, the ATP synthase subunits and the aa_3-type cytochrome oxidase are downregulated. During a low oxygen tension, the bd-type cytochrome oxidase is believed to be more utilized since it has a higher affinity for oxygen. ATP synthase is still active even though at a lower membrane potential which is not commonly seen in other organisms, underlining the importance of PMF in keeping the bacterium alive during this metabolic state. This could be a regular scenario for *Mtb* inside the granuloma that can drive low metabolic activity with low or no *Mtb* growth (dormancy) (15). Also, in absence of oxygen, *Mtb* uses a set of reductases (such as succinate/fumarate reductase and nitrate reductase), hydrogenases (coupling H_2 oxidation to respiration, encoded by Rv0082 and Rv0087) and ferredoxins (such as the encoded by *fdxA*) that preserve the PMF for bacterial survival (123, 124). Other changes have been detected in anaerobic adaptation, for instance, the E1 subunit of the pyruvate dehydrogenase is upregulated. Under anaerobic condition, *Mtb* can stay alive but its growth is strongly reduced (123). This theme is relevant because as it was previously described, there is a wide variety of oxygen tension in the *Mtb* interaction with the host.

1.7 *Mtb* and the host: Immune response

Mtb interaction with the human host has been an evolving topic looking at different facets of immune response generated in the host. For many years, the production of T helper (Th) type 1cytokines such Interferon (IFN)-γ and tumor necrosis factor (TNF)-α together with macrophage activation have been considered the central players for the TB protective immunity in the murine model and humans (127, 128). IFN-γ is essential for macrophage activation which is responsible for killing of intracellular bacteria, demonstrating the importance of adaptive cellular immunity against *Mtb* (129). On the other hand, the participation of the humoral (antibody)

responses has been thought to have a minor role during an infection process with *Mtb*. The role

of the humoral response has been controversial, but it has been proposed that it could participate

in bacterial opsonization with antibodies to enhance the phagocytosis and elimination of *Mtb*

(130, 131). Further controversy exists about the fact that adaptive cellular immunity rarely

eradicates infection completely and does not protect from TB reactivation or reinfection (128).

Additionally, a recent vaccination effort that triggered cellular immunity failed to protect against

TB infection (132). Similar results were obtained in other studies where no correlation of IFN-γ

with protection after BCG vaccination was observed, creating more controversy about Th1

immunity stimulation and protection against TB. After this, new research efforts have been

stablished not only towards alternate vaccination strategies but also in deciphering the role of

other components of host immunity such as the innate immunity for an effective anti-TB

response (133).

The first encounter of the bacteria with the host at the cellular level occurs with the alveolar

macrophages (favorite cell type) and resident dendritic cells after the exposure of the host with

an aerosol droplet that contains *Mtb*. According to initial studies in rabbits, 1 to 3 bacilli of *M.

bovis* or 50-2000 units of *Mtb* bacilli are enough to reach the alveolar space and cause a

pulmonary lesion (127). Additional studies in guinea pigs, revealed that one bacilli is enough to

cause infection. Due to many similarities in pathology between this animal model and human

disease, one can extrapolate this findings to the human infection (134). As part of the initial

innate immune response, different pathways induce the production of vitamin D-dependent

peptides with mycobactericidal activity. These pathways include the activation of Toll Like

Receptor (TLR) 2, 4 and 9. The infected immune cells recruit more inflammatory cells to the

site of infection in the lung, carry the *Mtb* antigens, and travel to the draining mediastinal lymph

nodes to prime naïve T cells and initiate adaptive T-cell immune responses. Bacteria can also

disseminate to other parts of the body stimulating adaptive immune responses at these locations.

Lymphocytes and more macrophages migrate to the site of infection to form the granuloma. This

complex structure contains macrophages, T CD4+, CD8+ and B lymphocytes, fibroblasts,

neutrophils and multinucleated giant cells (which are fused macrophages attacking an infected

macrophage). The purpose of this structure is to contain the bacteria and prevent dissemination.

Inside the granuloma CD4+ T cells activate macrophages to restrict the replication and dispersal

of *Mtb*. However, this response does not always lead to complete clearance of the bacteria. In

fact, *Mtb* uses the granuloma to be masked from the host immune response (130, 135).

Macrophages are multifunctional cells of both the innate and adaptive anti-TB responses.

These cells have the ability to phagocytize the bacteria, present antigen, stimulate other cells (for

instance T cells) and finally eliminate the bacteria in the phagolysosome. After phagocytosis is

complete, the pH can go from 6.2 to 4.5 due to the presence of ATPases that pump protons inside

the phagosomal compartment. The phagosomal compartment also has NADPH oxidase activity,

which is an enzyme present in the plasma membrane. This oxidase leads to superoxide (O_2^-)

production that can combine with nitric oxide (NO) and iron to produce RNI and hypervalent

iron. Additionally, O_2^- can be converted to H_2O_2 by the eukaryotic superoxide dismutase. The

phagosome also fuses with a secretory lysosome to form the phagolysosome, which have

hydrolytic activity from lysosomal-enzymes (136-138). The degradation that happens in this

ultrastructure induces bacterial killing, but does not induce more immune responses. However, if

the macrophage is activated by IFN-γ, the cell maximizes lysosomal proteolysis for epitope

production and presents the *Mtb* antigens through the Major Histocompatibility complex (MHC)

II, stimulating T cells (136).

Many different types of CD4 T cells are stimulated in a host response against *Mtb* that together with CD8 T cell induced responses are critical to overcome the infection. CD4 T cells can be additionally classified into Th1, Th2, Th17, and regulatory T (T regs) cells, based on their cytokine production. Proinflammatory Th1 cytokines include IFN-γ, TNF-α, IL-6, IL-12 and IL-1β. Mutations in IFN-γ and TNF-α have been demonstrated as genetic defects that lead to a reduced levels or activity of those cytokines in mice and the human host, thereby increasing the susceptibility to *Mtb* infection and reduce the ability to control the pathogen (128). IFN-γ induces autophagy to reduce the bacterial load in humans and nitric oxide synthase 2 (NOS2) expression in mice. TNF-α is critical for macrophage activation, increases the expression of the Major Histocompatibility Complex (MHC) and leukocyte adhesion. IL-6 is a pleiotropic cytokine with pro and anti-inflamatory properties (139). Lastly, IL-6 has an inhibitory effect in IFN-γ-dependent macrophage activation (140). However, IL-6 is required for a protective response in mice infected with high doses of *Mtb* (141). IL-12 promotes Th1 cell differentiation and production of IFN-γ (142). Th2 cytokines IL-4 and IL-10 are associated with humoral and anti-inflammatory responses respectively. IL-10 has a regulatory function to avoid exacerbated host tissue damage in the host crated by the excessive proinflammatory Th1 response. IL-10 is also produced by another group of CD4 T cells, T regs, that also produce transforming growth factor (TGF)-β (130, 135). Th17 produce IL-17, IL22, among others cytokines, which connect innate with adaptive immunity. The last group of CD4 T cells, Th17 are important in the granuloma formation by the stimulation of Th1 cells and recruitment of neutrophils, in turn producing denfensins, cathelicidins, lipocalin, superoxides, among other molecules. Finally, CD8 T cells have an important cytotoxic activity and also produce perforin and granzymes that

result in the apoptosis of the infected cell as well as by the production of granulysin that can stimulate the production of more IFN-γ (135).

The anti-*Mtb* innate immune response continues to be a subject of research, especially to define which events control the progression from initial infection to disease progression in the host. In this initial immune response, cells such as alveolar macrophages, lung epithelial cells, dendritic cells, neutrophils, γδT cells, and invariant natural killers; as well as Toll like Receptors (TLR 2, 4, 9) are accounted as some of the most important factors (133, 143, 144). Because of their location alveolar macrophages and lung epithelial cells are the most important initial responders after exposure to *Mtb*. These cells have a rapid response against *Mtb* antigens and restrain early spread of the pathogen. Some roles of these cells have been previously described in this chapter, mainly phagocytosis, antigen presentation and production of cytokines, chemokines and antimicrobial peptides (142-144). The TLR pathway and the internal signaling through the myeloid differentiation factor 88 (MYD88) are known to have important roles in the response against *Mtb*. TLR 2 is a major receptor for lipoproteins, the mycolic-arabinogalactan-peptidoglycan complex and lipoarabinomannan (142, 143).

1.8 Animal models of TB

For *Mtb* infection, mice, guinea pigs, rabbits, macaques, zebrafish and others have been tested to recreate the human infection and identify relevant responses in the bacteria-host interaction (145, 146). One important characteristic of animal models is that the majority develop a pulmonary infection, which is the main manifestation of *Mtb* infection in humans. Notably, small animal models overlap only with respect to some features of the natural host for this bacterium. Therefore, all the findings revealed in those models should be interpreted with

43

caution (145). Especially when studying animal models such as zebrafish that can be infected with the *Mtb* closely related *M. marinum* that causes a skin infection with macrophage aggregates and granulomas in this animal.

Of the smaller animal models, the guinea pig and rabbit models most closely resemble the human TB pathology, specifically the granuloma formation with caseous lesions. Most initial insights of the immune response against *Mtb* were obtained from studies in the rabbit that also develop liquefaction and cavity lesions (127). However, this animal model is more resistant to *Mtb* infection, requiring higher doses of the bacteria. Guinea pigs can rapidly develop a high bacterial load in the lung, resulting in animal death. Non-human primates, such as the *Macaca fasicularis,* infected with *Mtb* develop very similar pathology to human TB, it is the only animal model forming cavities to date. Moreover, the non-human primate model has been used to study latent TB (124, 147). In addition to this larger model, mice represent an excellent model for *Mtb* infection and have been extensively used in the lab because of their size, cost-effectivity, the presence of genetically defined mouse linages, the availability of reagents to test their immune response, and certain similarities found with the human induced-cytokine response. As with humans, the aerosol route of infection in mice generates granuloma formation in the lung. The limitations of this model include that the *Mtb* generated granulomas are structurally different and do not have caseous lesions or liquefaction and generally fail to develop necrotic lesions and cavities, all of which are key features of human pulmonary TB (124).

1.9 Overview of the dissertation

Mtb is a fascinating microorganism because of its historical symbiont-interaction with the human host and for its distinctive physiology. The purpose of this dissertation includes

evaluating the bacterial cost and subsequent events after acquisition of INH resistance in *Mtb* strains, which requires a review of relevant aspects of the bacterial physiology and the immune response generated in the host. Additionally, we have gained important insights about *Mtb* physiology through the understanding of the mechanism of action and resistance to anti-TB drugs, specifically INH. Although different metabolic pathways in *Mtb* have been studied independently using genomic, transcriptomic, and proteomic approaches, each of these pathways and reactions should be studied comprehensively as they normally occur in the cell. For instance, the TCA cycle connects with not only other central carbon metabolic pathways, but also with anabolic pathways and redox metabolism processes.

    Mtb physiology expands the limits of what microbiologists have learned from model organisms such as *E. coli* and *B. subtilis*, creating its own identity and gradually informing researchers of its successful mechanisms to persist within the host. The complex infection dynamic of *Mtb* in the host is further intensified by specific characteristics of the bacteria dictated by its genetic background that sometimes can be translated in a readable phenotype. In the next sections of this dissertation, the study of clonal strains of *Mtb* before and after acquisition of INH resistance will explore biochemical and immunological aspects specific to this resistant, clinically-originated phenotype.

REFERENCES

1.   Paulson T. Epidemiology: A mortal foe. Nature. 2013 Oct 10;502(7470):S2-3.
2.   Hershkovitz I, Donoghue HD, Minnikin DE, Besra GS, Lee OY, Gernaey AM, et al. Detection and molecular characterization of 9,000-year-old Mycobacterium tuberculosis from a Neolithic settlement in the Eastern Mediterranean. PLoS One. 2008;3(10):e3426.
3.   Cave AJE. The evidence for the incidence of tuberculosis in ancient Egypt. Br J Tuberc. 1939;33:142-52.
4.   Zink A, Haas CJ, Reischl U, Szeimies U, Nerlich AG. Molecular analysis of skeletal tuberculosis in an ancient Egyptian population. J Med Microbiol. 2001 Apr;50(4):355-66.
5.   Daniel VS, Daniel TM. Old Testament biblical references to tuberculosis. Clin Infect Dis. 1999 Dec;29(6):1557-8.
6.   Sakula A. Robert Koch: centenary of the discovery of the tubercle bacillus, 1882. Thorax. 1982 Apr;37(4):246-51.
7.   Houston M. The white death: A history of tuberculosis. BMJ. 1999 Jun 19;318(7199):1705.
8.   Ryan F. The forgotten plague : how the battle against tuberculosis was won--and lost. 1st American ed. Boston: Little, Brown; 1993.
9.   Bos KI, Harkins KM, Herbig A, Coscolla M, Weber N, Comas I, et al. Pre-Columbian mycobacterial genomes reveal seals as a source of New World human tuberculosis. Nature. 2014 Oct 23;514(7523):494-7.
10.   Organization WH. Global Tuberculosis Report2015.
11.   Organization WH. Global tuberculosis report 2014.
12.   Balasubramanian V, Wiegeshaus EH, Taylor BT, Smith DW. Pathogenesis of tuberculosis: pathway to apical localization. Tuber Lung Dis. 1994 Jun;75(3):168-78.
13.   Daffé M, Draper P. The envelope layers of mycobacteria with reference to their pathogenicity. Adv Microb Physiol. 1998;39:131-203.
14.   North RJ, Jung YJ. Immunity to tuberculosis. Annu Rev Immunol. 2004;22:599-623.
15.   Gengenbacher M, Kaufmann SH. Mycobacterium tuberculosis: success through dormancy. FEMS Microbiol Rev. 2012 May;36(3):514-32.
16.   Vilchèze C, Jacobs WR. Resistance to Isoniazid and Ethionamide in Mycobacterium tuberculosis: Genes, Mutations, and Causalities. Microbiol Spectr. 2014 Aug;2(4):MGM2-0014-2013.
17.   Zumla A, Nahid P, Cole ST. Advances in the development of new tuberculosis drugs and treatment regimens. Nat Rev Drug Discov. 2013 May;12(5):388-404.
18.   Lehmann J. Para-Aminosalicylic Acid in the Treatment of Tuberculosis. The Lancet. 1946;247(6384):15-6.
19.   TREATMENT of pulmonary tuberculosis with streptomycin and para-aminosalicylic acid; a Medical Research Council investigation. Br Med J. 1950 Nov;2(4688):1073-85.
20.   Riva MA. From milk to rifampicin and back again: history of failures and successes in the treatment for tuberculosis. J Antibiot (Tokyo). 2014 Sep;67(9):661-5.
21.   Vilchèze C, Jacobs WR. The mechanism of isoniazid killing: clarity through the scope of genetics. Annu Rev Microbiol. 2007;61:35-50.
22.   Jones DS. The health care experiments at Many Farms: the Navajo, tuberculosis, and the limits of modern medicine, 1952-1962. Bull Hist Med. 2002;76(4):749-90.

23.   Cohen J. Infectious disease. Approval of novel TB drug celebrated--with restraint. Science. 2013 Jan;339(6116):130.
24.   Hoagland DT, Liu J, Lee RB, Lee RE. New agents for the treatment of drug-resistant Mycobacterium tuberculosis. Adv Drug Deliv Rev. 2016 May.
25.   Gler MT, Skripconoka V, Sanchez-Garavito E, Xiao H, Cabrera-Rivero JL, Vargas-Vasquez DE, et al. Delamanid for multidrug-resistant pulmonary tuberculosis. N Engl J Med. 2012 Jun;366(23):2151-60.
26.   Kremer L, Douglas JD, Baulard AR, Morehouse C, Guy MR, Alland D, et al. Thiolactomycin and related analogues as novel anti-mycobacterial agents targeting KasA and KasB condensing enzymes in Mycobacterium tuberculosis. J Biol Chem. 2000 Jun;275(22):16857-64.
27.   Slayden RA, Lee RE, Armour JW, Cooper AM, Orme IM, Brennan PJ, et al. Antimycobacterial action of thiolactomycin: an inhibitor of fatty acid and mycolic acid synthesis. Antimicrob Agents Chemother. 1996 Dec;40(12):2813-9.
28.   Douglas JD, Senior SJ, Morehouse C, Phetsukiri B, Campbell IB, Besra GS, et al. Analogues of thiolactomycin: potential drugs with enhanced anti-mycobacterial activity. Microbiology. 2002 Oct;148(Pt 10):3101-9.
29.   Ling LL, Schneider T, Peoples AJ, Spoering AL, Engels I, Conlon BP, et al. A new antibiotic kills pathogens without detectable resistance. Nature. 2015 Jan;517(7535):455-9.
30.   Condos R, Rom WN, Schluger NW. Treatment of multidrug-resistant pulmonary tuberculosis with interferon-gamma via aerosol. Lancet. 1997 May;349(9064):1513-5.
31.   Boeree MJ, Diacon AH, Dawson R, Narunsky K, du Bois J, Venter A, et al. A dose-ranging trial to optimize the dose of rifampin in the treatment of tuberculosis. Am J Respir Crit Care Med. 2015 May 1;191(9):1058-65.
32.   Jayaram R, Gaonkar S, Kaur P, Suresh BL, Mahesh BN, Jayashree R, et al. Pharmacokinetics-pharmacodynamics of rifampin in an aerosol infection model of tuberculosis. Antimicrob Agents Chemother. 2003 Jul;47(7):2118-24.
33.   Rosenthal IM, Tasneen R, Peloquin CA, Zhang M, Almeida D, Mdluli KE, et al. Dose-ranging comparison of rifampin and rifapentine in two pathologically distinct murine models of tuberculosis. Antimicrob Agents Chemother. 2012 Aug;56(8):4331-40.
34.   Wallis RS, Maeurer M, Mwaba P, Chakaya J, Rustomjee R, Migliori GB, et al. Tuberculosis-advances in development of new drugs, treatment regimens, host-directed therapies, and biomarkers. Lancet Infect Dis. 2016 Apr;16(4):e34-46.
35.   Zumla AI, Gillespie SH, Hoelscher M, Philips PP, Cole ST, Abubakar I, et al. New antituberculosis drugs, regimens, and adjunct therapies: needs, advances, and future prospects. Lancet Infect Dis. 2014 Apr;14(4):327-40.
36.   Lienhardt C, Vernon A, Raviglione MC. New drugs and new regimens for the treatment of tuberculosis: review of the drug development pipeline and implications for national programmes. Curr Opin Pulm Med. 2010 May;16(3):186-93.
37.   Zumla A, Nahid P, Cole ST. Advances in the development of new tuberculosis drugs and treatment regimens. Nature Reviews Drug Discovery. 2013;12(5):388-404.
38.   Banerjee A, Dubnau E, Quemard A, Balasubramanian V, Um KS, Wilson T, et al. inhA, a gene encoding a target for isoniazid and ethionamide in Mycobacterium tuberculosis. Science. 1994 Jan;263(5144):227-30.

39.   Campbell EA, Korzheva N, Mustaev A, Murakami K, Nair S, Goldfarb A, et al. Structural mechanism for rifampicin inhibition of bacterial rna polymerase. Cell. 2001 Mar;104(6):901-12.

40.   Cole ST. Microbiology. Pyrazinamide--old TB drug finds new target. Science. 2011 Sep 16;333(6049):1583-4.

41.   Shi W, Chen J, Feng J, Cui P, Zhang S, Weng X, et al. Aspartate decarboxylase (PanD) as a new target of pyrazinamide in Mycobacterium tuberculosis. Emerging Microbes & Infections. 2014;3(8):e58.

42.   Belanger AE, Besra GS, Ford ME, Mikusová K, Belisle JT, Brennan PJ, et al. The embAB genes of Mycobacterium avium encode an arabinosyl transferase involved in cell wall arabinan biosynthesis that is the target for the antimycobacterial drug ethambutol. Proc Natl Acad Sci U S A. 1996 Oct;93(21):11919-24.

43.   Kohanski MA, Dwyer DJ, Collins JJ. How antibiotics kill bacteria: from targets to networks. Nat Rev Microbiol. 2010 Jun;8(6):423-35.

44.   Lin Y, Li Y, Zhu N, Han Y, Jiang W, Wang Y, et al. The Antituberculosis Antibiotic Capreomycin Inhibits Protein Synthesis by Disrupting Interaction between Ribosomal Proteins L12 and L10. Antimicrobial Agents and Chemotherapy. 2014;58(4):2038-44.

45.   Recht MI, Douthwaite S, Puglisi JD. Basis for prokaryotic specificity of action of aminoglycoside antibiotics. EMBO J. 1999 Jun;18(11):3133-8.

46.   Zheng J, Rubin EJ, Bifani P, Mathys V, Lim V, Au M, et al. para-Aminosalicylic acid is a prodrug targeting dihydrofolate reductase in Mycobacterium tuberculosis. J Biol Chem. 2013 Aug;288(32):23447-56.

47.   Prosser GA, de Carvalho LP. Kinetic mechanism and inhibition of Mycobacterium tuberculosis D-alanine:D-alanine ligase by the antibiotic D-cycloserine. FEBS J. 2013 Feb;280(4):1150-66.

48.   Livermore DM. Linezolid in vitro: mechanism and antibacterial spectrum. Journal of Antimicrobial Chemotherapy. 2003;51(90002):9ii-16.

49.   Velayati AA, Masjedi MR, Farnia P, Tabarsi P, Ghanavi J, Ziazarifi AH, et al. Emergence of new forms of totally drug-resistant tuberculosis bacilli: super extensively drug-resistant tuberculosis or totally drug-resistant strains in iran. Chest. 2009 Aug;136(2):420-5.

50.   Snider DE, Farer LS. Preventive therapy with isoniazid for "inactive" tuberculosis. Chest. 1978 Jan;73(1):4-6.

51.   Slayden RA, Barry CE. The genetics and biochemistry of isoniazid resistance in mycobacterium tuberculosis. Microbes Infect. 2000 May;2(6):659-69.

52.   Antimicrobial therapy and vaccines. Yu VL, Merigan TC-, Barriere SL, editors. Baltimore: Williams & Wilkins; 1999.

53.   Lei B, Wei CJ, Tu SC. Action mechanism of antitubercular isoniazid. Activation by Mycobacterium tuberculosis KatG, isolation, and characterization of inha inhibitor. J Biol Chem. 2000 Jan;275(4):2520-6.

54.   Mahapatra S, Woolhiser LK, Lenaerts AJ, Johnson JL, Eisenach KD, Joloba ML, et al. A novel metabolite of antituberculosis therapy demonstrates host activation of isoniazid and formation of the isoniazid-NAD+ adduct. Antimicrob Agents Chemother. 2012 Jan;56(1):28-35.

55.   Zhang Y, Heym B, Allen B, Young D, Cole S. The catalase-peroxidase gene and isoniazid resistance of Mycobacterium tuberculosis. Nature. 1992 Aug;358(6387):591-3.

56.   Slayden RA, Lee RE, Barry CE. Isoniazid affects multiple components of the type II fatty acid synthase system of Mycobacterium tuberculosis. Mol Microbiol. 2000 Nov;38(3):514-25.

57.   Ducasse-Cabanot S, Cohen-Gonsaud M, Marrakchi H, Nguyen M, Zerbib D, Bernadou J, et al. In vitro inhibition of the Mycobacterium tuberculosis beta-ketoacyl-acyl carrier protein reductase MabA by isoniazid. Antimicrob Agents Chemother. 2004 Jan;48(1):242-9.

58.   Mdluli K, Slayden RA, Zhu Y, Ramaswamy S, Pan X, Mead D, et al. Inhibition of a Mycobacterium tuberculosis beta-ketoacyl ACP synthase by isoniazid. Science. 1998 Jun;280(5369):1607-10.

59.   Timmins GS, Deretic V. Mechanisms of action of isoniazid. Mol Microbiol. 2006 Dec;62(5):1220-7.

60.   Zhang Y, Dhandayuthapani S, Deretic V. Molecular basis for the exquisite sensitivity of Mycobacterium tuberculosis to isoniazid. Proc Natl Acad Sci U S A. 1996 Nov;93(23):13212-6.

61.   Zhang Y, Young DB. Molecular mechanisms of isoniazid: a drug at the front line of tuberculosis control. Trends Microbiol. 1993 Jun;1(3):109-13.

62.   Wakamoto Y, Dhar N, Chait R, Schneider K, Signorino-Gelo F, Leibler S, et al. Dynamic persistence of antibiotic-stressed mycobacteria. Science. 2013 Jan 4;339(6115):91-5.

63.   Ellard GA, Gammon PT, Wallace SM. The determination of isoniazid and its metabolites acetylisoniazid, monoacetylhydrazine, diacetylhydrazine, isonicotinic acid and isonicotinylglycine in serum and urine. Biochem J. 1972 Feb;126(3):449-58.

64.   Goldman AL, Braman SS. Isoniazid: a review with emphasis on adverse effects. Chest. 1972 Jul;62(1):71-7.

65.   Jarlier V, Nikaido H. Mycobacterial cell wall: structure and role in natural resistance to antibiotics. FEMS Microbiol Lett. 1994 Oct;123(1-2):11-8.

66.   Almeida Da Silva PE, Palomino JC. Molecular basis and mechanisms of drug resistance in Mycobacterium tuberculosis: classical and new drugs. J Antimicrob Chemother. 2011 Jul;66(7):1417-30.

67.   Cole ST, Brosch R, Parkhill J, Garnier T, Churcher C, Harris D, et al. Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence. Nature. 1998;393(6685):537-44.

68.   Seifert M, Catanzaro D, Catanzaro A, Rodwell TC. Genetic mutations associated with isoniazid resistance in Mycobacterium tuberculosis: a systematic review. PLoS One. 2015;10(3):e0119628.

69.   MIDDLEBROOK G, COHN ML. Some observations on the pathogenicity of isoniazid-resistant variants of tubercle bacilli. Science. 1953 Sep;118(3063):297-9.

70.   Zhang Y, Garbe T, Young D. Transformation with katG restores isoniazid-sensitivity in Mycobacterium tuberculosis isolates resistant to a range of drug concentrations. Mol Microbiol. 1993 May;8(3):521-4.

71.   McGrath M, Gey van Pittius NC, van Helden PD, Warren RM, Warner DF. Mutation rate and the emergence of drug resistance in Mycobacterium tuberculosis. Journal of Antimicrobial Chemotherapy. 2013;69(2):292-302.

72.   Mdluli K, Swanson J, Fischer E, Lee RE, Barry CE, 3rd. Mechanisms involved in the intrinsic isoniazid resistance of Mycobacterium avium. Mol Microbiol. 1998 Mar;27(6):1223-33.

73.   Bergval IL, Schuitema ARJ, Klatser PR, Anthony RM. Resistant mutants of Mycobacterium tuberculosis selected in vitro do not reflect the in vivo mechanism of isoniazid resistance. Journal of Antimicrobial Chemotherapy. 2009;64(3):515-23.

74.   Larsen MH, Vilchèze C, Kremer L, Besra GS, Parsons L, Salfinger M, et al. Overexpression of inhA, but not kasA, confers resistance to isoniazid and ethionamide in

Mycobacterium smegmatis, M. bovis BCG and M. tuberculosis. Mol Microbiol. 2002 Oct;46(2):453-66.

75. Mokrousov I, Dantes R, Metcalfe J, Kim E, Kato-Maeda M, Hopewell PC, et al. Impact of Isoniazid Resistance-Conferring Mutations on the Clinical Presentation of Isoniazid Monoresistant Tuberculosis. PLoS One. 2012;7(5):e37956.

76. Cook GM, Berney M, Gebhard S, Heinemann M, Cox RA, Danilchanka O, et al. Physiology of mycobacteria. Adv Microb Physiol. 2009;55:81-182, 318-9.

77. Steingart KR, Schiller I, Horne DJ, Pai M, Boehme CC, Dendukuri N. Xpert® MTB/RIF assay for pulmonary tuberculosis and rifampicin resistance in adults. Cochrane Database Syst Rev. 2014;1:CD009593.

78. Kent PT, Kubica GP. Public Health Mycobacteriology. A guide for the level III laboratory. U.S Department of Health and Human Services PHS, Centers for Disease Control CDC, editor. Atlanta, GA1985.

79. Noussair L, Bert F, Leflon-Guibout V, Gayet N, Nicolas-Chanoine MH. Early diagnosis of extrapulmonary tuberculosis by a new procedure combining broth culture and PCR. J Clin Microbiol. 2009 May;47(5):1452-7.

80. Parsons LM, Somoskövi A, Gutierrez C, Lee E, Paramasivan CN, Abimiku A, et al. Laboratory diagnosis of tuberculosis in resource-poor countries: challenges and opportunities. Clin Microbiol Rev. 2011 Apr;24(2):314-50.

81. Garc©*a-Elorriaga G, Rey-Pineda Gd. Practical and laboratory diagnosis of tuberculosis : from sputum smear to molecular biology.

82. Boehme CC, Nicol MP, Nabeta P, Michael JS, Gotuzzo E, Tahirli R, et al. Feasibility, diagnostic accuracy, and effectiveness of decentralised use of the Xpert MTB/RIF test for diagnosis of tuberculosis and multidrug resistance: a multicentre implementation study. Lancet. 2011 Apr;377(9776):1495-505.

83. Rossau R, Traore H, De Beenhouwer H, Mijs W, Jannes G, De Rijk P, et al. Evaluation of the INNO-LiPA Rif. TB assay, a reverse hybridization assay for the simultaneous detection of Mycobacterium tuberculosis complex and its resistance to rifampin. Antimicrob Agents Chemother. 1997 Oct;41(10):2093-8.

84. Ling DI, Zwerling AA, Pai M. Rapid diagnosis of drug-resistant TB using line probe assays: from evidence to policy. Expert Rev Respir Med. 2008 Oct;2(5):583-8.

85. Cole ST, Barrell BG. Analysis of the genome of Mycobacterium tuberculosis H37Rv. Novartis Found Symp. 1998;217:160-72; discussion 72-7.

86. Kamerbeek J, Schouls L, Kolk A, van Agterveld M, van Soolingen D, Kuijper S, et al. Simultaneous detection and strain differentiation of Mycobacterium tuberculosis for diagnosis and epidemiology. J Clin Microbiol. 1997 Apr;35(4):907-14.

87. Bifani P, Kurepina N, Mathema B, Wang XM, Kreiswirth B. Genotyping of Mycobacterium tuberculosis clinical isolates using IS6110-based restriction fragment length polymorphism analysis. Methods Mol Biol. 2009;551:173-88.

88. Lee RS, Behr MA. The implications of whole-genome sequencing in the control of tuberculosis. Ther Adv Infect Dis. 2016 Apr;3(2):47-62.

89. Cannas A, Mazzarelli A, Di Caro A, Delogu G, Girardi E. Molecular Typing of Mycobacterium Tuberculosis Strains: A Fundamental Tool for Tuberculosis Control and Elimination. Infect Dis Rep. 2016 Jun;8(2):6567.

90. Mizrahi V, Andersen SJ. DNA repair in Mycobacterium tuberculosis. What have we learnt from the genome sequence? Mol Microbiol. 1998 Sep;29(6):1331-9.

91.   Wanner RM, Guthlein C, Springer B, Bottger EC, Ackermann M. Stabilization of the genome of the mismatch repair deficient Mycobacterium tuberculosis by context-dependent codon choice. BMC Genomics. 2008;9:249.

92.   Sherman DR, Gagneux S. Estimating the mutation rate of Mycobacterium tuberculosis during infection. Nat Genet. 2011 May;43(5):400-1.

93.   Minnikin DE. Chemical principles in the organization of lipid components in the mycobacterial cell envelope. Res Microbiol. 1991 May;142(4):423-7.

94.   Hunter RL, Olsen MR, Jagannath C, Actor JK. Multiple roles of cord factor in the pathogenesis of primary, secondary, and cavitary tuberculosis, including a revised description of the pathology of secondary disease. Ann Clin Lab Sci. 2006 Autumn;36(4):371-86.

95.   Crick DC, Mahapatra S, Brennan PJ. Biosynthesis of the arabinogalactan-peptidoglycan complex of Mycobacterium tuberculosis. Glycobiology. 2001;11(9):107R-18R.

96.   Brennan PJ, Nikaido H. The envelope of mycobacteria. Annu Rev Biochem. 1995;64:29-63.

97.   Sartain MJ, Dick DL, Rithner CD, Crick DC, Belisle JT. Lipidomic analyses of Mycobacterium tuberculosis based on accurate mass measurements and the novel "Mtb LipidDB". J Lipid Res. 2011 May;52(5):861-72.

98.   ASSELINEAU J, LEDERER E. Structure of the mycolic acids of Mycobacteria. Nature. 1950 Nov;166(4227):782-3.

99.   Bhatt A, Molle V, Besra GS, Jacobs WR, Kremer L. The Mycobacterium tuberculosis FAS-II condensing enzymes: their role in mycolic acid biosynthesis, acid-fastness, pathogenesis and in future drug development. Mol Microbiol. 2007 Jun;64(6):1442-54.

100.   McNeil M, Daffe M, Brennan PJ. Location of the mycolyl ester substituents in the cell walls of mycobacteria. J Biol Chem. 1991 Jul;266(20):13217-23.

101.   Besra GS, Sievert T, Lee RE, Slayden RA, Brennan PJ, Takayama K. Identification of the apparent carrier in mycolic acid synthesis. Proc Natl Acad Sci U S A. 1994 Dec;91(26):12735-9.

102.   Takayama K, Wang C, Besra GS. Pathway to synthesis and processing of mycolic acids in Mycobacterium tuberculosis. Clin Microbiol Rev. 2005 Jan;18(1):81-101.

103.   Shah M, Hanrahan C, Wang ZY, Dendukuri N, Lawn SD, Denkinger CM, et al. Lateral flow urine lipoarabinomannan assay for detecting active tuberculosis in HIV-positive adults. 2016.

104.   Gouzy A, Poquet Y, Neyrolles O. Nitrogen metabolism in Mycobacterium tuberculosis physiology and virulence. Nat Rev Microbiol. 2014 Nov;12(11):729-37.

105.   Tian J, Bryk R, Itoh M, Suematsu M, Nathan C. Variant tricarboxylic acid cycle in Mycobacterium tuberculosis: identification of alpha-ketoglutarate decarboxylase. Proc Natl Acad Sci U S A. 2005 Jul;102(30):10670-5.

106.   Maksymiuk C, Balakrishnan A, Bryk R, Rhee KY, Nathan CF. E1 of alpha-ketoglutarate dehydrogenase defends Mycobacterium tuberculosis against glutamate anaplerosis and nitroxidative stress. Proc Natl Acad Sci U S A. 2015 Oct 27;112(43):E5834-43.

107.   Rhee KY, Carvalho LPSd, Bryk R, Ehrt S, Marrero J, Park SW, et al. Central carbon metabolism in Mycobacterium tuberculosis: an unexpected frontier. Trends in Microbiology. 2011;19(7):307-14.

108.   Baughn AD, Garforth SJ, Vilchèze C, Jacobs WR. An anaerobic-type alpha-ketoglutarate ferredoxin oxidoreductase completes the oxidative tricarboxylic acid cycle of Mycobacterium tuberculosis. PLoS Pathog. 2009 Nov;5(11):e1000662.

109.   Cumming BM, Steyn AJ. Metabolic plasticity of central carbon metabolism protects mycobacteria. Proc Natl Acad Sci U S A. 2015 Oct;112(43):13135-6.
110.   Munoz-Elias EJ, McKinney JD. Carbon metabolism of intracellular bacteria. Cellular Microbiology. 2006;8(1):10-22.
111.   Williams KJ, Boshoff HI, Krishnan N, Gonzales J, Schnappinger D, Robertson BD. The Mycobacterium tuberculosis β-oxidation genes echA5 and fadB3 are dispensable for growth in vitro and in vivo. Tuberculosis (Edinb). 2011 Nov;91(6):549-55.
112.   Marrakchi H, Lanéelle MA, Daffé M. Mycolic acids: structures, biosynthesis, and beyond. Chem Biol. 2014 Jan;21(1):67-85.
113.   Lea-Smith DJ, Pyke JS, Tull D, McConville MJ, Coppel RL, Crellin PK. The reductase that catalyzes mycolic motif synthesis is required for efficient attachment of mycolic acids to arabinogalactan. J Biol Chem. 2007 Apr;282(15):11000-8.
114.   Belisle JT, Vissa VD, Sievert T, Takayama K, Brennan PJ, Besra GS. Role of the major antigen of Mycobacterium tuberculosis in cell wall biogenesis. Science. 1997 May;276(5317):1420-2.
115.   Trivedi A, Singh N, Bhat SA, Gupta P, Kumar A. Redox biology of tuberculosis pathogenesis. Adv Microb Physiol. 2012;60:263-324.
116.   Leichert LI, Scharf C, Hecker M. Global characterization of disulfide stress in Bacillus subtilis. J Bacteriol. 2003 Mar;185(6):1967-75.
117.   Bertrand T, Eady NA, Jones JN, Jesmin, Nagy JM, Jamart-Grégoire B, et al. Crystal structure of Mycobacterium tuberculosis catalase-peroxidase. J Biol Chem. 2004 Sep;279(37):38991-9.
118.   Ghiladi RA, Medzihradszky KF, Rusnak FM, Ortiz de Montellano PR. Correlation between isoniazid resistance and superoxide reactivity in mycobacterium tuberculosis KatG. J Am Chem Soc. 2005 Sep;127(38):13428-42.
119.   Wengenack NL, Jensen MP, Rusnak F, Stern MK. Mycobacterium tuberculosis KatG is a peroxynitritase. Biochem Biophys Res Commun. 1999 Mar;256(3):485-7.
120.   Ng VH, Cox JS, Sousa AO, MacMicking JD, McKinney JD. Role of KatG catalase-peroxidase in mycobacterial pathogenesis: countering the phagocyte oxidative burst. Mol Microbiol. 2004 Jun;52(5):1291-302.
121.   Pym AS, Saint-Joanis B, Cole ST. Effect of katG mutations on the virulence of Mycobacterium tuberculosis and the implication for transmission in humans. Infect Immun. 2002 Sep;70(9):4955-60.
122.   Manca C, Paul S, Barry CE, Freedman VH, Kaplan G. Mycobacterium tuberculosis catalase and peroxidase activities and resistance to oxidative killing in human monocytes in vitro. Infect Immun. 1999 Jan;67(1):74-9.
123.   Cook GM, Hards K, Vilchèze C, Hartman T, Berney M. Energetics of Respiration and Oxidative Phosphorylation in Mycobacteria. Microbiol Spectr. 2014 Jun;2(3).
124.   Boshoff HI, Barry CE. Tuberculosis - metabolism and respiration in the absence of growth. Nat Rev Microbiol. 2005 Jan;3(1):70-80.
125.   Bald D, Koul A. Respiratory ATP synthesis: the new generation of mycobacterial drug targets? FEMS Microbiol Lett. 2010 Jul;308(1):1-7.
126.   Koul A, Dendouga N, Vergauwen K, Molenberghs B, Vranckx L, Willebrords R, et al. Diarylquinolines target subunit c of mycobacterial ATP synthase. Nat Chem Biol. 2007 Jun;3(6):323-4.

127.   Dannenberg AM, Jr. Pathogenesis of pulmonary Mycobacterium bovis infection: basic principles established by the rabbit model. Tuberculosis (Edinb). 2001;81(1-2):87-96.

128.   Andersen P, Woodworth JS. Tuberculosis vaccines – rethinking the current paradigm. Trends in Immunology. 2014;35(8):387-95.

129.   Fletcher HA, Schrager L. TB vaccine development and the End TB Strategy: importance and current status. Transactions of The Royal Society of Tropical Medicine and Hygiene. 2016;110(4):212-8.

130.   Nunes-Alves C, Booty MG, Carpenter SM, Jayaraman P, Rothchild AC, Behar SM. In search of a new paradigm for protective immunity to TB. Nat Rev Microbiol. 2014 Apr;12(4):289-99.

131.   Cooper AM. Cell-Mediated Immune Responses in Tuberculosis. Annual Review of Immunology. 2009;27(1):393-422.

132.   Tameris MD, Hatherill M, Landry BS, Scriba TJ, Snowden MA, Lockhart S, et al. Safety and efficacy of MVA85A, a new tuberculosis vaccine, in infants previously vaccinated with BCG: a randomised, placebo-controlled phase 2b trial. The Lancet. 2013;381(9871):1021-8.

133.   Divangahi M, King IL, Pernet E. Alveolar macrophages and type I IFN in airway homeostasis and immunity. Trends in Immunology. 2015;36(5):307-14.

134.   Turner RD, Bothamley GH. Cough and the Transmission of Tuberculosis. J Infect Dis. 2014 Nov.

135.   Lin PL, Flynn JL. Understanding latent tuberculosis: a moving target. J Immunol. 2010 Jul;185(1):15-22.

136.   Russell DG, Vanderven BC, Glennie S, Mwandumba H, Heyderman RS. The macrophage marches on its phagosome: dynamic assays of phagosome function. Nat Rev Immunol. 2009 Aug;9(8):594-600.

137.   Vergne I, Chua J, Lee HH, Lucas M, Belisle J, Deretic V. Mechanism of phagolysosome biogenesis block by viable Mycobacterium tuberculosis. Proc Natl Acad Sci U S A. 2005 Mar;102(11):4033-8.

138.   Ehrt S, Schnappinger D. Mycobacterial survival strategies in the phagosome: defence against host stresses. Cell Microbiol. 2009 Aug;11(8):1170-8.

139.   Scheller J, Chalaris A, Schmidt-Arras D, Rose-John S. The pro- and anti-inflammatory properties of the cytokine interleukin-6. Biochimica et Biophysica Acta (BBA) - Molecular Cell Research. 2011;1813(5):878-88.

140.   Nagabhushanam V, Solache A, Ting LM, Escaron CJ, Zhang JY, Ernst JD. Innate inhibition of adaptive immunity: Mycobacterium tuberculosis-induced IL-6 inhibits macrophage responses to IFN-gamma. J Immunol. 2003 Nov 1;171(9):4750-7.

141.   Ladel CH, Blum C, Dreher A, Reifenberg K, Kopf M, Kaufmann SH. Lethal tuberculosis in interleukin-6-deficient mutant mice. Infect Immun. 1997 Nov;65(11):4843-9.

142.   van Crevel R, Ottenhoff THM, van der Meer JWM. Innate Immunity to Mycobacterium tuberculosis. Clinical Microbiology Reviews. 2002;15(2):294-309.

143.   Flynn JL, Chan J. Immunology of tuberculosis. Annu Rev Immunol. 2001;19:93-129.

144.   Orme IM, Robinson RT, Cooper AM. The balance between protective and pathogenic immune responses in the TB-infected lung. Nature Immunology. 2014;16(1):57-63.

145.   Gupta UD, Katoch VM. Animal models of tuberculosis. Tuberculosis. 2005;85(5-6):277-93.

146.   Myllymäki H, Bäuerlein CA, Rämet M. The Zebrafish Breathes New Life into the Study of Tuberculosis. Frontiers in Immunology. 2016;7.

147.   Walsh GP, Tan EV, dela Cruz EC, Abalos RM, Villahermosa LG, Young LJ, et al. The Philippine cynomolgus monkey (Macaca fasicularis) provides a new nonhuman primate model of tuberculosis that resembles human disease. Nat Med. 1996 Apr;2(4):430-6.

CHAPTER II - PROTEIN VARIATION OF *Mycobacterium tuberculosis* CLONAL STRAINS PRE- AND POST-ISONIAZID RESISTANCE ACQUISITION IN THE CLINICAL SETTING

2.1 Introduction

Globally the tuberculosis (TB) prevalence and mortality were 42% and 47% lower respectively in 2015 compared to 1990, however, in the African region there were still 281 cases for every 100,000 people contributing to a global incidence of 9.6 million TB cases in 2014. Similarly, in some American countries, such as Haiti and Peru, the TB incidence rate was as high as 200 and 120 cases per 100,000 population, respectively, for the same year (1). More worrisome is the fact that drug resistant TB is a growing phenomenon especially in densely populated regions of the world, where the resistant bacteria have more opportunities to propagate and are associated with other lethal pathogens such as HIV. Although the TB rate in the United States (U.S.) has been decreasing since 1992, it continues to be a major public health concern with still a rate of 3 cases per 100,000 population that has been unchanged since 2013 (2). This number seems low compared to Haiti and Peru and even lower compared to South Africa (that has a TB rate of 834 cases per 100,000 people), but because of the chronicity of the disease, its high dissemination, the economic impact, and the drug resistant TB cases that are still not efficiently diagnosed or treated, this number is still a burden on society. The magnitude of this problem is clearer when the TB mortality rates are measured, as this disease is the current leading cause of death due to an infectious agent worldwide. In 2012, the U.S. began to experience a severe interruption in the supply of isoniazid (INH), a first-line antitubercular drug, which increased the threat of incomplete treatment regimens and the potential for higher INH resistance rates (3). Even worse, the shortage in the supply of INH was not exclusive to the U.S.

This event has been experienced in many areas of the world with even higher TB and resistant TB rates due to irregular drug supply and inferior quality of anti-TB drugs (4, 5).

As it was mentioned in chapter 1, INH is one of the most effective drugs within the multi-drug regimen to treat patients with active TB (6). INH is also often used as monotherapy for latent TB infection (LTBI), and in some countries as a preventive therapy for people living with HIV (1). According to the World Health Organization, around 9.5% of the total global TB cases were resistant to INH in the last decade (7). In clinical settings, TB resistance to INH only (or TB INH-monoresistant) is the most frequent form of drug resistant TB (8, 9). INH monoresistant *Mtb* strains predominantly correlates with *katG* mutations and is reported to result in higher mutation rates, increasing the probability to acquire resistance to other drugs such as rifampicin (9, 10). Consequently, these resistant strains have an increased likelihood to cause multidrug-resistant (MDR) TB. Simultaneous to INH resistance development, *Mtb* can have further adaptations including changes in protein levels which in turn may compensate for the potential fitness loss due to the new drug resistant phenotype. Therefore, it is crucial to understand not only how the INH resistant mechanisms occur in *Mtb* but also how this new phenotype is able to support bacteria growth.

The study of drug resistant TB has been an ongoing process, mainly because the understanding of the mechanism of action of several first line drugs (such as INH and pyrazinamide) has been subject of intensive research and controversies (11). This is the case for INH resistant (INHr) TB, where many bacterial factors are considered. The cumulative exposure of *Mtb* strains to suboptimal concentrations of INH in an intermittent manner creates most of the acquired INHr TB cases. In this way, many TB patients lose one of the best options to get cleared from a disease that was initially curable. Some of the proposed mechanisms of action of

INH were described in chapter I, with InhA (protein of the fatty acid synthase route II) being the most recognized target (12-14). More than 60 years of chemotherapy with INH in TB cases has allowed the development of different *Mtb* genotypes of INH resistant (INHr) profile and their associated phenotypes. Similarly discussed in the first chapter, INHr mechanisms include mutations in multiple genes most often in the drug activator, *katG* gene followed by mutation in the drug target *inhA* (11). The INHr phenotypes include low and high level of INH resistance as well as increased or decreased virulence (15).

A comparative shotgun proteomic approach looking at different cellular fractions of *Mtb* will help to describe the effects of INHr conferring mutations in the new INHr phenotype. After the elucidation of the genome of many organisms, proteomics emerged as a powerful methodology that not only describes the sequence, structure and function of the proteins, but also extends to the analysis of complex mixture of proteins using high throughput techniques (16, 17). Proteomics analyze mature proteins considering all the complex post-translational events that occur in the cell and that finally represent the bacterial phenotype. As it was stated by Joshua LaBaer in 2011, "proteins provide the verbs to biology" (18), and proteomics allow for naming different biological events (19). As the proteome of the cell variate parallel to internal metabolic variation and external cues, proteomics is considered the most direct scaffold to measure cell activity (20).

Mass spectrometry (MS) based technologies are central components of the protein analysis. These methods include shotgun and targeted proteomics that have different modes for acquiring mass spectra. Shotgun proteomics a term coined by John Yates III and his lab, offers an indirect measurement of proteins through peptides derived from their enzymatic digestion (16). Shotgun proteomics, also known as discovery proteomics, uses liquid chromatography (LC)

connected to tandem MS (MS/MS) for the identification of the protein components in the sample. The protein identification is based on the determination of the aminoacid sequence that is done by the comparison of the experimental tandem mass spectra and the theoretical tandem mass spectra generated from an *in silico* digestion of a protein database. Quantitative analysis of proteins became an integral and significant part in the interpretation of a biological event. Through shotgun proteomics, relative changes in the concentration of peptides (and indirectly proteins) can be detected across the samples using labeling elements and label-free approaches. A description of different labeling methods defined by the acronyms ICAT, iTRAQ®, TMT®, SILAC, among others, has been reviewed by other researchers (21-28). This dissertation is specifically focused on label-free shotgun proteomic analysis that uses spectral counting (SpC) or ion intensity for the relative quantification of proteins. In SpC the frequency of the peptides that correspond to a particular protein is used to infer the abundance of the protein in the sample. In the ion intensity strategy, the protein abundance is reported as the intensity of the MS chromatographic peak of the peptides assigned to the protein. In this work, the normalized spectral abundance factor (NSAF) which is based on SpC was used to do the relative quantification of the proteins which has demonstrated a linear correlation and good reproducibility (16, 29), with a limitation in the analysis of low abundance proteins (30). The complete definition of NSAF is described in the methods section.

The general workflow of a proteomics study starts with the selection of the experimental biological conditions to evaluate, protein enzymatic digestion, a separation of the peptide mixture through a reversed-phase LC, an MS analysis and the subsequent data/bioinformatics analysis. For biological samples, a nano-LC is typically used as it offers higher sensitivity and requires lower volumes of samples (31). In shotgun proteomics, product ion scanning is the most

58

commonly used mode for acquisition of mass spectra where the ionized peptide is fragmented

and the products are analyzed in a second MS, repeating this process for the different precursor

ions (19). In other words, peptides are subjected to two sequential MS. In a sequential process,

peptides are ionized and transferred to a gas phase using an electrospray technique (ESI) and

transferred to the first MS analyzer resulting in the measurement of mass over charge ratio (m/z)

of the generated peptide ions. Additionally, during the first MS specific precursor ions are

selected for fragmentation in a second MS round. The precursor in the first MS ion is used to

determine the quantity of the peptide, while the second MS is used to infer the peptide sequence.

The selection of the precursor ion is automatically done in a survey scan, considering all the ions

detected in a process called data-dependent analysis (DDA) (32, 33). In the second MS, the

precursor peptide ions are accelerated to increase their kinetic energy and then subjected to

collision with an inert gas (such as argon or helium) that leads to a fragmentation in the peptide

bound, determining the aminoacid sequence (17, 19, 33, 34). In summary, a general workflow of

a proteomic experiment includes sample selection (which includes the setup of the biological

conditions to analyze, in this case INHs versus INHr *Mtb* strains), sample simplification, protein

extraction, concentration and digestion, separation of the peptide mixture, and a final analytical

process in a tandem MS (Figure 2.1).

In *Mtb* proteomics studies, MS-based methodologies are used for deciphering the *Mtb*

proteome during *in vivo* and *in vitro* growth (31, 35-37). The protein identification is based on

the genome derived prediction of the protein sequence of the bacterial reference strains (mainly

H37Rv). Since the whole sequencing of *Mtb* genome, all the coding sequences were grouped in

ten functional categories that include a big portion of (a) conserved hypotheticals, (b) regulatory

proteins,(c) intermediate metabolism and respiration, (d) virulence, detoxification and

adaptation, (e) information pathways, (f) insertion sequences and phages, (g) lipid metabolism, (h) cell wall and cell wall processes, (j) PE/PPE and (i) proteins with unknown function (20). From *Mtb* proteomic studies, different proteins are identified as markers for virulence, playing important roles in host interaction and also specific metabolic stages such as low oxygen concentration and starvation (37, 38). This allowed the recognition of important protein markers in the bacteria proposed as biomarkers of the disease or as stimulators of an effective immune response with potential applications in a vaccination strategy (39-41) . Regarding INH resistance, a previous proteomic analysis focused on INH susceptible (INHs) and resistant (INHr) strains of *Mtb* and found five proteins overexpressed in the INHr strains (42). These proteins were found through matrix assisted laser desorption ionization time of flight-MS (MALDI-TOF) and include Rv1446c, Rv3028c, Rv0491, Rv2971, and Rv2145. However, these proteins are not related to any of the known INH resistance mechanisms. Most of these proteins were involved in cellular metabolism, including redox metabolism (such as Rv1446, Rv2971, and Rv3028c) and there was one transcriptional regulatory protein (Rv0491). To a large extent, this could be because the strains grouped as INHr in that study were from different genetic lineages and respective protein expression profiles with different mutations conferring INH resistance and were compared to other very heterogeneous non-related group of INHs *Mtb* strains.
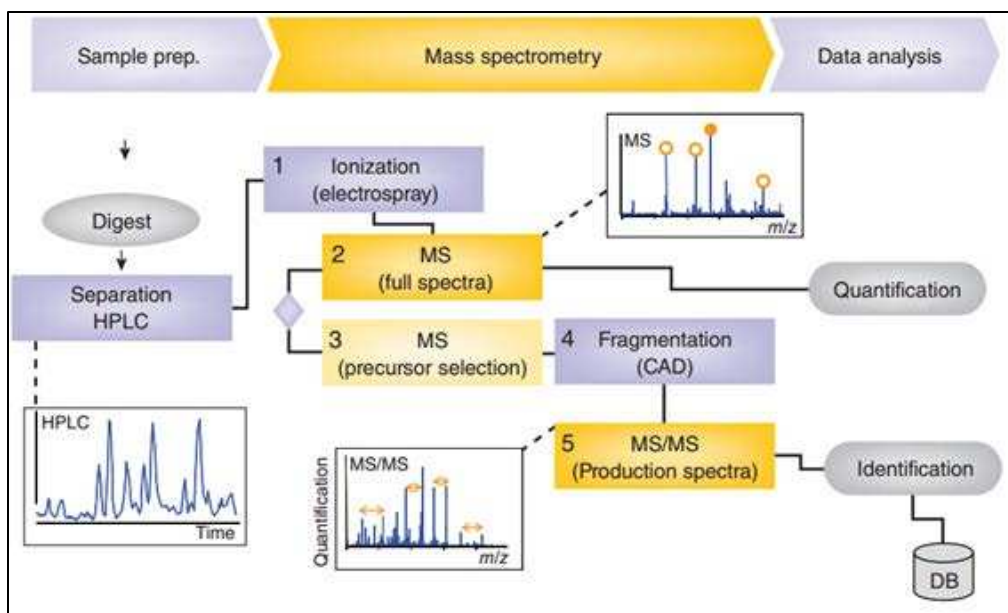
Figure 2.1. Adapted from Domon and Aebersold (43), with the permission of Nature Publishing Group.

In this initial experimental section of this dissertation, a study of clonal pairs of clinical *Mtb* strains, derived from the same patient before and after INHr acquisition, is presented with the goal of evaluating the variation in protein levels after development of the INHr phenotype. Clonal strains differentiate from each other because after some replication events, one of them developed one or more mutations within-host that in many cases is associated with phenotypic drug resistance (44). Clonal clinical strains provide a unique setting to reduce variability in the comparative analysis since they are derived from a common ancestor and variation in proteins can be attributed mostly to the mutation-conferring drug resistant phenotype. Additionally, these mutations, derived in a clinical setting, already contain the known variables that *Mtb* is normally exposed within the host, such as immune responses and a complete anti-TB drug regime. A shotgun proteomics approach is particularly helpful to address the hypothesis that after acquisition of INHr, clinical *Mtb* strains can experience variations in their protein array that are associated with the new INHr phenotype.

2.2 Materials and methods

2.2.1. Bacterial strains and preparation of Culture Filtrate Proteins (CFPs) and cellular fractions

Two clonal pairs of *Mtb* were isolated from two patients that had pulmonary TB that were enrolled in a previous study of tuberculosis transmission in San Francisco following approval from Stanford University and the University of California San Francisco's Institutional Review Boards (15, 45). Both strains were provided to Colorado State University (to the laboratory of Dr. Ian Orme, with subsequent provision to the laboratory of Dr. Karen Dobos). *Mtb* isolates were sent to the National Jewish Hospital, Denver, CO to confirm the drug susceptibility profile initially obtained by the clinical labs where the patients were being treated. An *Mtb* clinical strain susceptible to all first line drugs tested was obtained from a patient in San Francisco CA., prior to initiating standard anti-tubercular therapy (INHs). After several weeks on therapy, another isolate was obtained from the same patient and was confirmed to be resistant to INH by agar proportion test (tested as above) but susceptible to the other 3 first-line agents (INHr). These isolates were also confirmed as clonal pairs by spoligotyping and IS*6110* RFLP, and were determined to belong to the T family (46, 47). After drug treatment failure, the second strain isolated from each patient became resistant to both concentrations of INH tested (0.2 and 1.0 µg/ml). These patients were enrolled in a former study and had no disease recurrence in a two-year follow up (15, 45).

A second *Mtb* strain susceptible to all first line drugs tested (INH, RIF, pyrazinamide, and ethambutol) was obtained from a patient living in Spain prior to initiating anti-TB therapy (INHs). Another isolate was later obtained from the same patient after several weeks on treatment and was confirmed resistant to INH (tested at 0.2 and 1.0 µg/ml) by the agar

proportion test (48), but susceptible to the other 3 first line drugs (INHr). These isolates were two parent-mutant pairs defined by spoligotyping and insertion sequence (IS)*6110* restriction fragment length polymorphism (RFLP) pattern that belong to the Beijing genotype (46, 47).

For the proteomic analysis, three replicates of each bacterial strain were cultured in one liter of Glycerol-Alanine-salts (GAS) media. The preparation of CFP and cell fractions required an initial filtration step of the liquid media (using a 0.2µm filter) and γ irradiation of the washed cells (dose of 2.4 Mrad) respectively. CFPs contain all the proteins secreted by the microorganism and also those released during bacteria lysis. Bacterial death of the irradiated cells was confirmed using alamarBlue assay (Invitrogen) and subsequent work was performed at the biosafety level II laboratory (BSL2). In the BSL2, cells were resuspended in breaking buffer (0.6 µg/mL DNAase and 0.6 µg/mL RNAase and protease inhibitor (cOmplete, EDTA-free, Roche, at the manufacturer's suggested ratio) in PBS-EDTA buffer) and lysed using 10 cycles of probe-sonication (90 s on and 30 s off) maintaining cells in ice. After this, whole cell lysates were generated and the *Mtb* cellular fractions including membrane (MEM), cytosol (CYT) and cell wall (CW) were obtained by continuous cycles of ultracentrifugation, following the steps described by Lucas *et al* (49). The CFPs were concentrated from 1 L to approximately 20 mL using a Millipore™ Amicon™ Bioseparation Stirred Cell with a 3-KDa mass cutoff membrane (Millipore). From this lower volume, the remaining culture media still present in the CFP samples was replaced for 10mM ammonium bicarbonate, using Amicon Ultra-15 centrifugal filter units with a 3-kDa molecular mass cutoff. CYT and MEM pellets were resuspended and further concentrated using the previous described Amicon centrifugal filter units that also allowed the buffer exchange for 10mM ammonium bicarbonate. Before the digestion of proteins, the CW fraction was delipidated as described elsewhere (50). This protocol uses lysozyme,

incubation with the organic solvents chloroform/methanol (2:1, v/v), chloroform/methanol/water (10:10:3) and a final precipitation of the proteins with acetone at -20°C. The resulting pelleted proteins were resuspended in 10mM ammonium bicarbonate.

2.2.2 In-solution digestion of proteins

After the separation of CFP and mycobacterial cell fractions, total protein was quantified by the bicinchoninic acid method (Thermo Scientific Pierce). Cellular and secreted fractions were subjected to in solution protein digestion that include the following steps: acetone precipitation at -80°C, solubilization, reduction with dithiothreitol, alkylation with iodoacetamide, and trypsin digestion at 37°C during 3 hours (using a mix of 1% ProteaseMaxTM Surfactant (Promega) and 0.1 µg/µl trypsin (Roche)) for 30µg of total protein in a 1:17 ratio approximately. Following digestion, samples were desalted with Pierce® PepClean C18 columns (ThermoScientific) following the manufacturer instructions and resuspended in buffer. Further description of this protocol is available at https://vprnet.research.colostate.edu/PMF/wp-content/uploads/2015/06/In_Solution_diestion-protocol_040512.pdf.

2.2.3 Liquid chromatography coupled with tandem mass spectrometry (LC-MS/MS)

One microgram of digested cellular fractions and CFP of all the three biological replicates were randomly analyzed in triplicate using LC-MS/MS as described previously at the Proteomics and Metabolomics Facilites (Colorado State University) (50). Briefly, the peptides of each fraction were separated on a nanospray column (Zorbax C18, 5 mm, 75 mm ID 6 150 mm column). Samples were eluted into a 2D linear quadruple ion trap mass spectrometer-LTQ (Thermo-Finningan). Resulting raw data were converted into mzXML files using ProteoWizard (51). LTQ is a robust and inexpensive technology made of four rods with high trapping volume efficiency and sensitivity (19). LC-MS/MS spectra were then compared against *Mtb* genomic

database (MtbReverse041712) using SORCERER (Sage-N Research, version 5.0.1). The parameters used for the analysis were: trypsin digestion, a maximum of 2 missed cleavages, a precursor mass range of 400 to 4500 amu, peptide mass tolerance of 1.5 amu, reduction and alkylation of cysteine residues (resulting in the addition of a carbamidomethyl group, 57.02 amu) and the oxidation of methionine (15.99 amu).

2.2.4 Tandem mass spectrometry (MS/MS) data analysis

For each cellular fraction, peptide identifications from the MS/MS spectra previously searched were combined in the proteomic software Scaffold (version Scaffold version 4.5.3, Proteome Software Inc., Portland, OR) summing all the technical replicates results for each biological sample. Scaffold also was used to validate MS/MS based peptide and protein identifications. Peptide identifications were accepted if they could be established at greater than 90.0% probability by the Peptide Prophet algorithm (52). Protein identifications were accepted if they could be established at greater than 99.0% probability and contained at least 2 identified peptides. Protein probabilities were assigned by the Protein Prophet algorithm (53). Proteins that contained similar peptides and could not be differentiated based on MS/MS analysis alone were grouped to satisfy the principles of parsimony. NSAF algorithm was used for the comparative analysis between INH susceptible and resistant strains of each genotype (54). NSAF is an algorithm defined in 2010 that allows the comparison of abundance of individual proteins in multiple independent samples. NSAF is calculated as the ratio of the spectral counts from peptides uniquely mapping to a protein (uSpC) over the sum of the protein amino acid lengths mapping to unique (uL) and shared peptides (sL) (see the formula in Figure 2.2). In this way, NSAF only counts peptides that are unique to a particular protein and takes into account the

protein length and then normalized this value over the total sum of spectral counts/length of all the proteins in the experiment (54).

2.2.5 Statistical analysis

Protein abundance differences in each clonal pair comparing susceptible and resistant strains to INH were tested by two-tailed Student's t test with a significance level of 95%. These tests were done in the Scaffold software previous assignation of the NSAF algorithm to do the comparison. Additionally, a fold change analysis was also performed and supported the significantly different results.

2.2.6 Western Blot (WB)

Some of the proteomic findings were further validated by WB. Primary antibodies were obtained though the Biodefense and Emerging Infections Research Resources Repository-BEI, https://www.beiresources.org/). Except for AcpM, all WB were revealed with a chromogenic system, using a secondary antibody conjugated with alkaline phosphatase and BCIP/NBT (5-bromo-4-chloro-3'-indolyphosphate/ nitro-blue tetrazolium) as substrate. For AcpM, a chemiluminescent strategy was followed using the secondary antibodies Goat Anti-Rabbit F(ab)2 fragment–HRP (horseradish peroxidase) conjugated (Thermo Scientific) with the SuperSignal West Pico Stable peroxide solution (Thermo Scientific) for a higher sensitivity assay. A summary of the proteomic experimental approach is given in figure 2.2.

66
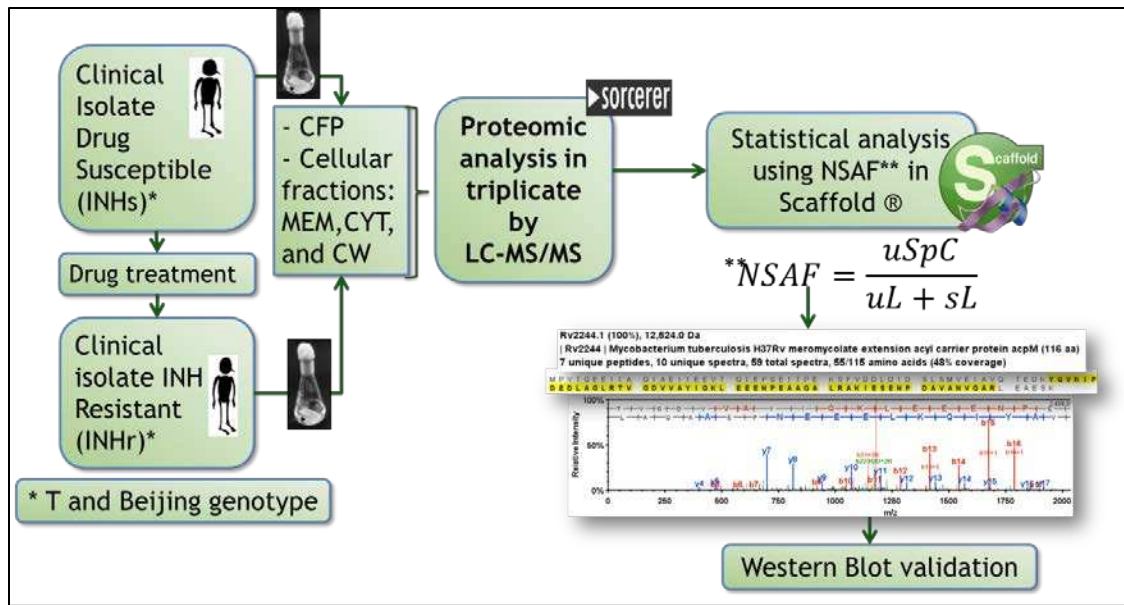
Figure 2.2. Proteomics experimental approach for the clonal strain comparison. INH: Isoniazid, CFP: culture filtrate proteins, MEM: Membrane, CYT: Cytosol, CW: Cell wall. LC-MS/MS: Liquid chromatography coupled with tandem mass spectrometry. uSpC: spectral counts of peptides uniquely mapping to a protein, uL: protein amino acid lengths mapping to unique peptide, sL: protein amino acid lengths mapping to shared peptides.

2.2.7 Bacterial growth curve

As an additional in vitro characterization of these clonal strains, five replicates of each strain were resuspended in Middlebrook 7H9 media supplemented with 0.05% tween to evaluate their growth. Ten mL of bacteria suspension were cultured in 16 x 125 mm tubes starting with an optical density (OD) close to 0.1 measured at 600nm for each replicate. These tubes were incubated at $37^{o}C$ in constant agitation, using a stir bar. Subsequent OD readings at 600 nm were performed daily until day 24. H37Rv reference strain was also included as a control in the experiment.

2.2.8 katG PCR sequencing

PCR amplification for Sanger sequencing was performed to detect the mutations in those strains after the proteomic analysis. *katG* amplification was done using a reaction mix that

included Takara Ex Taq® Hot Start DNA polymerase (1.25 U) (TAKARA BIO INC), a mixture

of deoxynucleoside triphosphates (dNTP) (200 µM each), betaine (1 M), primers for katG

(forward 5′ CGA TTT TCG GTG CCG TGC GTT TT 3′ and reverse 5′ CGC CCA GCC ATG

CAT GAG CAT TAT 3′) and *Mtb* gDNA (250 ng) quantified by Nanodrop.

An initial denaturation step at 94 °C for 5 min followed by a total of 40 cycles of

amplification were performed, with template denaturation at 94 °C for 10 s, primer annealing at

59.5 °C for 30 s and extension at 72 °C for 3 min. Then a final extension at 72 °C was done for 6

min. After visualization of the PCR product on a 1.5% agarose gel, post-PCR cleanup and

Sanger sequencing occurred at the genomic core of the Proteomics and Metabolomics facilities

at Colorado State University.

2.3 Results and Discussion

This section presents the findings from each clonal pair comparison separately, in order

to evaluate specific changes after the drug resistant event and avoid possible skew results due to

very strong differences in one of the genotypes compared. After this, the common differences

will be analyzed together to detect common trends.

2.3.1 Protein abundance comparison in the T clonal pair

The NSAF analysis revealed a total of 99 proteins being significantly differential when

comparing INH susceptible versus the resistant strains in the T genotype pair (p <0.05). The

major protein functional categories affected were intermediate metabolism and respiration (37%)

(Figure 2.3). The cellular fraction with the highest contribution to protein differences was the

membrane, with a total of 40 proteins with either higher or lower representation in the INHr

strain (Table 2.1). These 99 proteins were distributed in the different cellular fractions and 17 of

68

them were present in more than one cellular fraction. Additionally, some of the significantly

differential proteins were encoded by a group of four identified operons (Figure 2.4).
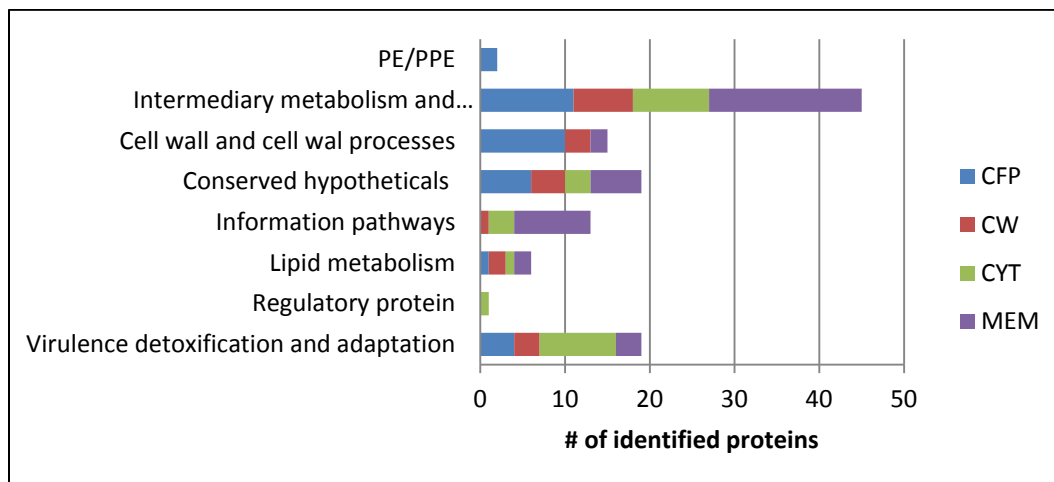
Figure 2.3. Protein categories affected in T clonal strains after acquisition of INHr differentiated
by cellular fraction. All categories are listed according to Tuberculist (version 2.6, Release 27 -
March 2013, http://tuberculist.epfl.ch/)

KatG and GroES are two examples of proteins significantly differential in almost all

cellular fractions of the T INHr strain (Table 2.1). KatG, as expected, was significantly reduced

in the INHr strain. This difference was one of the strongest after looking at the fold change

values (Table 2.1). This confirms the role of KatG as the activator of the prodrug INH and its

relation with the INHr phenotype of this strain. In the same group of virulence, detoxification

and adaptation, GroES was reduced in the INHr strain in all fractions except in cell wall

(similarly to KatG). Together with them the chaperone HtpG, the peroxidase AhpC, the

thioredoxin TrxB2 and bactoferritin BfrB were also reduced in the T INHr strain, while other

proteins in the same category (SodA, Tpx, GroEL and HbhA) had higher levels possibly

compensating the lack of those important proteins in the detoxification and virulence function.

The chaperone DnaK showed variable results depending on the cellular fraction studied (Table

2.1), suggesting that there is a possible compartmentalization of this protein in the cell wall that induced the low levels observed in the membrane of the T INHr strain.
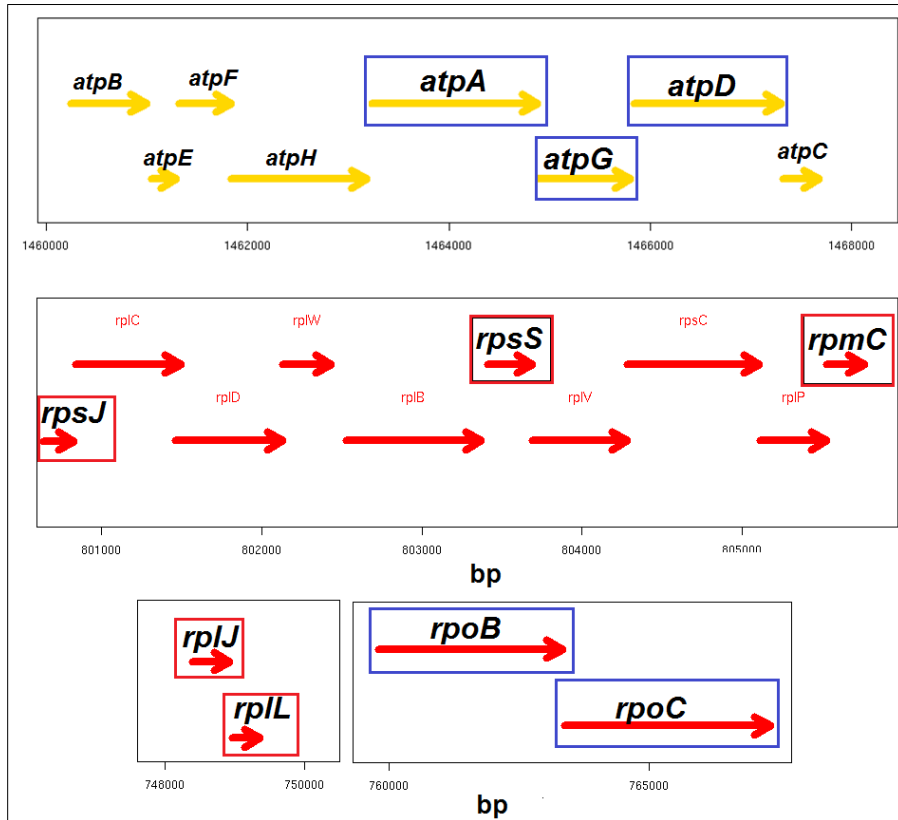
Figure 2.4. Organization of the coding genes for the proteins that were found altered in the T clonal pair comparative analysis. Blue boxes indicate that the proteins encoded by those genes were higher, while red boxes indicate they were lower in the INH resistant strain of the T genotype. Adapted from http://tuberculist.epfl.ch/.

There was a group of proteins with different abundances between T INHs and INHr strain, but after looking at other cellular fractions, this difference was more likely due to the protein localization. This is the case for pterin-4-alpha-carbinolamine dehydratase MoaB2 and SseC2, which is annotated to be involved in sulphur metabolism; both proteins were significantly more concentrated in the membrane and reduced in the CFP of the T INHr strain. Finally, the serine hydroxymethyltransferase GlyA1 was more secreted (higher in the CFP) than in the CW.

70

Interestingly, one group of altered proteins is part of the ATP synthase machinery which is crucial for *Mtb* survival (Figure 2.4 and Table 2.1). In the T INHr strain, the proteins AtpA, AtpG and AtpD were higher compared to the INHs strain. As it was reviewed in Chapter 1, ATP synthase is required for optimal growth in both active and dormant metabolic states. Also, the respiratory driven ATP is in fact the major source for energy in *Mtb* (55). The ATP synthase was recently identified as a target to treat bacterial infections. In line with this, bedaquiline—recently approved by the Food and Drug Administration to treat MDR-TB infection—inhibits ATP synthase, binding to the subunit C (AtpE, Rv1305) of the complex (56). This drug is effective against both active and dormant *Mtb*. The higher levels of these ATP synthase subunits in the INHr strain does not alter directly the drug target AtpE, instead, the increase of the other ATP synthase subunits emphasizes the importance of this metabolic process in *Mtb* even in drug resistant strains. In a previous proteomic analysis, *M. smegmatis* (a mycobacterium that is less sensitive to INH) had reduced levels of ATP synthase proteins after exposure to INH (57). Therefore, the alteration of the proteins in this important machinery for *Mtb* could have started just with the exposure of *Mtb* strain to INH within the patient. After becoming INH resistant, this trend in ATP synthase proteins had to change to cope with the new redox metabolic state of the bacteria.

Another interesting finding was the protein RpoB—the target for rifampicin (RIF)—and RpoC were higher in the T INH resistant strain (Table 2.1, Figure 2.4). This finding suggests that after the alteration of KatG levels, *Mtb* could also potentially acquire RIF resistance, by increasing the amount of the target of the drug. The increased RpoB levels could be the result of the combined anti-TB therapy (that contains RIF) that the patient received in the first place, which exerted a selective pressure for those strains that were more able to tolerate this drug

(probably those with higher levels of the target). Additionally, because of the absence of

catalase-peroxidase (KatG), compensatory metabolic mechanisms that allow the regulation of the

redox stress were induced in the bacteria. In this way, an increase of the transcription rate of

genes encoding for proteins with this compensatory function can occur with the need of higher

levels of the β and β' subunits of the RNA polymerase (RpoB and RpoC respectively) (Table 2.1

and figure 2.4). As the increased levels of RpoB and C suggested an increase in the transcription

rate, the reduced levels of the proteins RpsJ RpmC, RpsS, RplJ and RplL in the INHr strain

(Table 2.1 and Figure 2.4), indicates the opposite about translation in the bacterium. In the

previously described proteomics study, levels of RpsJ and RpsS were strongly reduced in *M.

smegmatis* after treatment with not only INH but also with ethambutol and pyrazinamide (57).

Again, the only exposure to the anti-TB drugs could have started the reduced expression of these

proteins that seem to be permanent in the later acquired INHr phenotype.

Table 2.1. Description of significantly different proteins in the INHr vs INHs T clonal
strain comparison (t-test, p<0.05)

Proteins significantly different (*t*-test, p<0.05)	Gene name	Functional category	Rv number	Fold Change (INHs/ INHr)[a]
CFP (n=32)				
iron-regulated peptidyl-prolyl-cis-trans-isomerase A ppiA	*ppiA*	IMR	Rv0009.1	0.1
isocitrate dehydrogenase icd2	*icd2*	IMR	Rv0066c.1	0.2
succinate-semialdehyde dehydrogenase dependent gabD1	*gabD1*	IMR	Rv0234c.1	6.1
PE family protein	*Rv0285*	PE	Rv0285.1	0.3
Mg2+ transport transmembrane protein mgtE	*mgtE*	CW	Rv0362.1	0.1
protein transport protein secE2	*secE2*	CW	Rv0379.1	0.0
dihydrolipoamide dehydrogenase lpd	*lpd*	IMR	Rv0462.1	6.5
hypothetical protein sseC2	*sseC2*	IMR	Rv0814c.1	0.1
enolase eno	*eno*	IMR	Rv1023.1	0.4
conserved hypothetical protein	*Rv1056*	C	Rv1056.1	0.0
serine hydroxymethyltransferase 1 glyA1	*glyA1*	IMR	Rv1093.1	0.0
esat-6 like protein esxK	*esxK*	CW	Rv1197.1	0.2

			(+1)	
conserved hypothetical protein	*Rv1906*	C	Rv1906c.1	0.2
catalase-peroxidase-peroxynitritase T katG	*katG*	V	Rv1908c.1	INF
immunogenic protein mpt63	*mpt63*	CW	Rv1926c.1	0.0
thiol peroxidase tpx	*tpx*	V	Rv1932.1	0.8
cutinase precursor cfp21	*cfp21*	CW	Rv1984c.1	0.0
proteasome beta subunit prcB	*prcB*	IMR	Rv2110c.1	3.5
conserved hypothetical protein	*Rv2204*	C	Rv2204c.1	1.9
glutamine synthetase glnA1	*glnA1*	IMR	Rv2220.1	2.5
meromycolate extension acyl carrier protein acpM	*acpM*	L	Rv2244.1	3.8
low molecular weight protein antigen cfp2	*cfp2*	CW	Rv2376c.1	0.3
PE-PGRS family protein	*Rv2396*	PE	Rv2396.1	0.4
conserved alanine and glycine rich membrane protein	*Rv2721*	CW	Rv2721c.1	0.1
soluble secreted antigen mpt53 precursor	*mpt53*	CW	Rv2878c.1	0.1
ketol-acid reductoisomerase ilvC	*ilvC*	IMR	Rv3001c.1	0.0
conserved hypothetical protein	*Rv3267*	C	Rv3267.1	0.2
10 kda chaperonin groES	*groES*	V	Rv3418c.1	1.8
bifunctional membrane-associated penicillin-binding protein 1A/1B	*ponA2*	CW	Rv3682.1	0.2
conserved hypothetical protein	*Rv3705*	C	Rv3705c.1	0.0
conserved hypothetical protein	*Rv3716*	C	Rv3716c.1	8.4
10 kda culture filtrate antigen esxB	*esxB*	CW	Rv3874.1	0.2
CW (n=20)				
conserved membrane protein	*Rv0283*	C	Rv0283.1	0.4
chaperone protein dnaK	*dnaK*	V	Rv0350.1	0.8
iron-regulated heparin binding hemagglutinin hbhA	*hbhA*	V	Rv0475.1	0.8
citrate synthase I gltA2	*gltA2*	IMR	Rv0896.1	1.3
pterin-4-alpha-carbinolamine dehydratase moaB2	*moaB2*	IMR	Rv0984.1	0.8
enolase eno	*eno*	IMR	Rv1023.1	0.6
serine hydroxymethyltransferase 1 glyA1	*glyA1*	IMR	Rv1093.1	3.8
serine protease htrA	*htrA*	IMR	Rv1223.1	0.3
lipoprotein lprA	*lprA*	CW	Rv1270c.1	0.3
lipoprotein lprF	*lprF*	CW	Rv1368.1	0.1
30S ribosomal protein S1 rpsA	*rpsA*	IP	Rv1630.1	INF
alanine and proline rich secreted protein apa	*apa*	CW	Rv1860.1	0.1
conserved hypothetical protein	*Rv2159*	C	Rv2159c.1	0.3
long-chain fatty-acid-CoA ligase fadD15	*fadD15*	L	Rv2187.1	0.3
D-3-phosphoglycerate dehydrogenase serA1	*serA1*	IMR	Rv2996c.1	0.6
ketol-acid reductoisomerase ilvC	*ilvC*	IMR	Rv3001c.1	0.6
acyl-CoA dehydrogenase fadE22	*fadE22*	L	Rv3061c.1	INF
conserved hypothetical protein	*Rv3716*	C	Rv3716c.1	12.3
bacterioferritin bfrB	*bfrB*	V	Rv3841.1	1.2

conserved hypothetical protein	*Rv3867*	C	Rv3867.1	0.2
	CYT (n=26)			
iron-regulated peptidyl-prolyl-cis-trans-isomerase A ppiA	*ppiA*	IMR	Rv0009.1	0.7
conserved hypothetical protein	*Rv0036*	C	Rv0036c.1	0.1
60 kda chaperonin 2 groEL2	*groEL2*	V	Rv0440.1	0.7
iron-regulated heparin binding hemagglutinin hbhA	*hbhA*	V	Rv0475.1	0.3
iron-regulated elongation factor tu tuf	*tuf*	IP	Rv0685.1	0.3
pterin-4-alpha-carbinolamine dehydratase moaB2	*moaB2*	IMR	Rv0984.1	1.9
5-methyltetrahydropteroyltriglutamate-homocysteine methyltransferase metE	*metE*	IMR	Rv1133c.1	0.2
ATP synthase beta chain atpD	*atpD*	IMR	Rv1310.1	0.1
thioredoxin	*Rv1324*	V	Rv1324.1	0.6
oxpp cycle protein opcA	*opcA*	IMR	Rv1446c.1	0.1
catalase-peroxidase-peroxynitritase T katG	*katG*	V	Rv1908c.1	139.3
conserved hypothetical protein	*Rv2185*	C	Rv2185c.1	0.5
long-chain fatty-acid-CoA ligase fadD15	*fadD15*	L	Rv2187.1	0.2
chaperone protein htpG	*htpG*	V	Rv2299c.1	INF
alkyl hydroperoxide reductase C protein ahpC	*ahpC*	V	Rv2428.1	INF
haloalkane dehalogenase dhaA	*dhA*	IMR	Rv2579.1	0.5
iron-dependent repressor and activator ideR	*ideR*	RP	Rv2711.1	0.1
40 kda secreted L-alanine dehydrogenase ald	*ald*	IMR	Rv2780.1	0.0
histone-like DNA-binding protein hupB	*hupB*	IP	Rv2986c.1	0.2
electron transfer flavoprotein alpha subunit fixB	*fixB*	IMR	Rv3028c.1	0.7
conserved hypothetical protein	*Rv3040*	C	Rv3040c.1	0.3
10 kda chaperonin groES	*groES*	V	Rv3418c.1	1.3
30S ribosomal protein S4 rpsD	*rpsD*	IP	Rv3458c.1	0.0
acetohydroxyacid synthase ilvX	*ilvX*	IMR	Rv3509c.1	0.4
superoxide dismutase sodA	*sodA*	V	Rv3846.1	0.6
thioredoxin reductase trxB2	*trxB2*	V	Rv3913.1	4.5
	MEM (n=37)			
dihydroxy-acid dehydratase ilvD	*ilvD*	IMR	Rv0189c.1	0.1
fatty-acid-CoA ligase fadD2	*fadD2*	L	Rv0270.1	0.1
chaperone protein dnaK	*dnaK*	V	Rv0350.1	1.7
adenylosuccinate synthetase purA	*purA*	IMR	Rv0357c.1	0.0
aldehyde dehydrogenase	*Rv0458*	IMR	Rv0458.1	0.4
dihydrolipoamide dehydrogenase lpd	*lpd*	IMR	Rv0462.1	INF
conserved hypothetical protein	*Rv0580*	C	Rv0580c.1	0.1
50S ribosomal protein L10 rplJ	*rplJ*	IP	Rv0651.1	5.6
50S ribosomal protein L7/L12 rplL	*rplL*	IP	Rv0652.1	4.4

Protein	Gene	Category	Rv number	NSAF
DNA-directed RNA polymerase beta chain rpoB	*rpoB*	IP	Rv0667.1	0.4
DNA-directed RNA polymerase beta chain rpoC	*rpoC*	IP	Rv0668.1	0.2
30S ribosomal protein S10 rpsJ	*rpsJ*	IP	Rv0700.1	INF
30S ribosomal protein S19 rpsS	*rpsS*	IP	Rv0705.1	INF
50S ribosomal protein L29 rpmC	*rpmC*	IP	Rv0709.1	INF
protease IV sppA	*sspA*	CW	Rv0724.1	0.0
phosphoribosylformylglycinamidine cyclo-ligase purM	*purM*	IMR	Rv0809.1	0.0
hypothetical protein sseC2	*sseC2*	IMR	Rv0814c.1	INF
thiosulfate sulfurtransferase cysA2	*cysA2*	IMR	Rv0815c.1	INF
conserved hypothetical protein	*Rv0831*	C	Rv0831c.1	0.3
enoyl-CoA hydratase echA6	*echA6*	L	Rv0905.1	0.4
periplasmic phosphate-binding lipoprotein pstS1	*psts1*	IMR	Rv0934.1	0.4
lipoprotein lprA	*lprA*	CW	Rv1270c.1	0.0
transcription termination factor rho	*rho*	IP	Rv1297.1	0.1
ATP synthase alpha chain atpA	*atpA*	IMR	Rv1308.1	0.1
ATP synthase gamma chain atpG	*atpG*	IMR	Rv1309.1	0.2
ATP synthase beta chain atpD	*atpD*	IMR	Rv1310.1	0.5
inosine-5-monophosphate dehydrogenase guaB1	*guaB1*	IMR	Rv1843c.1	INF
conserved hypothetical protein	*Rv1871*	C	Rv1871c.1	0.0
L-lactate dehydrogenase lldD2	*lldD2*	IMR	Rv1872c.1	0.1
catalase-peroxidase-peroxynitritase T katG	*katG*	V	Rv1908c.1	39.2
conserved hypothetical protein	*Rv2054*	C	Rv2054.1	0.0
conserved hypothetical protein	*Rv2159*	C	Rv2159c.1	0.1
glutamine synthetase glnA2	*glnA2*	IMR	Rv2222c.1	0.1
haloalkane dehalogenase		IMR	Rv2296.1	0.0
conserved hypothetical protein	*Rv2298*	C	Rv2298.1	0.1
bifunctional polyribonucleotide nucleotidyltransferase gpsI	*gpsI*	IMR	Rv2783c.1	3.4
10 kda chaperonin groES	*groES*	V	Rv3418c.1	4.0
DNA-directed RNA polymerase alpha chain rpoA	*rpoA*	IP	Rv3457c.1	2.7
ATP-dependent protease ATP-binding subunit clpC1	*clpC1*	IMR	Rv3596c.1	0.2

a. The quantitative method chosen for the statistical analysis and p value calculation was NSAF.

b. INF: NSAF in INHr strain was zero. IMR: Intermediary metabolism and respiration, V: Virulence, detoxification, adaptation, IP: Information pathways, L: Lipid metabolism, R: Regulatory protein, CW: Cell wall and cell wall processes, C: Conserved Hypothetical.

Respiration and metabolism including lipid metabolism related proteins were found with different abundances after acquisition of INH resistance (Table 2.1). Most of the proteins in the metabolism and respiration category participate in different metabolic pathways including glycolysis (Eno), TCA cycle (GabD1)(58) among others. Lipid metabolism proteins including the acyl carrier protein, AcpM (Rv2244), which is an important enzyme involved in fatty acids elongation in the type II fatty acid synthase system, was found in a greater abundance in a previous study *Mtb* after exposure to INH (59). From this publication, AcpM, together with KasA (another protein in FAS II pathway) were proposed as possible INH targets (59). In this study, the low levels of AcpM, in addition to the already known effect of the low levels of KatG, could contribute to the INHr phenotype and additionally suggests a different lipid profile after acquisition of INH resistance.

2.3.2 Protein abundance analysis in the Beijing clonal pair

In the second clinical clonal pair, 46 proteins were either more or less abundant after acquisition of INHr (with p <0.05) (Table 2.2). These were grouped in seven different categories (Figure 2.5) and the protein differences were mostly observed in the CFP (39.6%) and MEM (35.4%) fractions (Figure 2.5, Table 2.2). As it was observed in the T clonal pair, KatG levels were significantly reduced in the INHr pair in all the cell fractions except the CW. Some of these proteins were encoded by clustered genes (Figure 2.6). Although the proximity of Rv0241c, Rv0242c, and Rv0243, they are present in different DNA strands; as denoted by letter c at the end in the first two proteins, indicating the coding sequences are in the complementary strand (60). These three proteins are increased in the Beijing INHr strain (Table 2.2 and figure 2.6). The raw data of the Beijing pair was deposited to the ProteomeXchange Consortium (61)via the PRIDE partner repository with the dataset identifier PXD002986 and 10.6019/PXD002986.
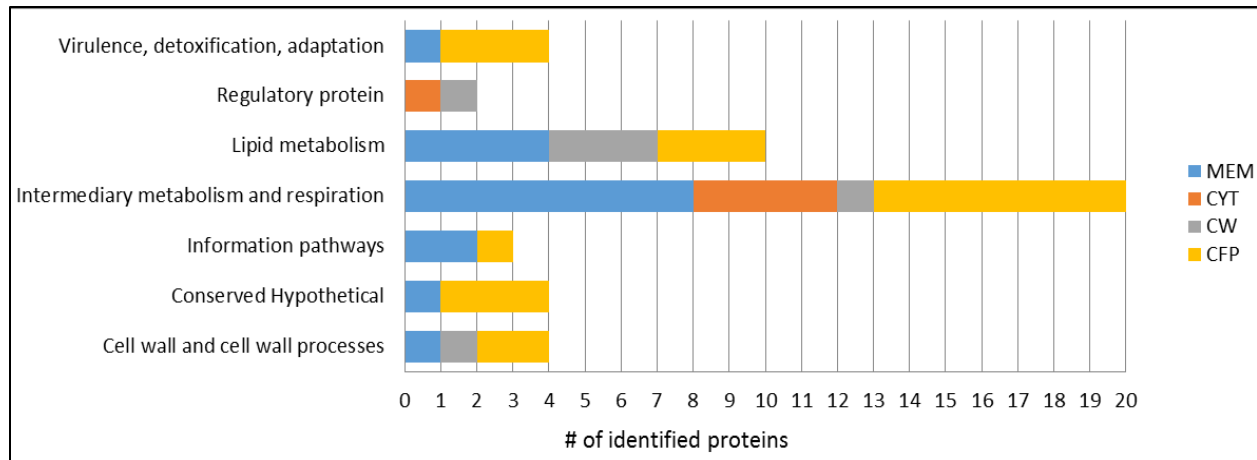
Figure 2.5. Functional categories of the *Mtb* proteins showing different levels among the INHs and INHr clonal Beijing strains (p value <0.05). All categories are listed according to Tuberculist (version 2.6, Release 27 - March 2013, http://tuberculist.epfl.ch/)

Table 2.2. Description of significantly different proteins in the INHr vs INHs Beijing strain comparison (t-test, p<0.05)

Proteins significantly different (*t*-test, p<0.05)	Gene name	Rv number	Functional category	Fold Change (INHs/ INHr)[a]
CFP (n=19)				
Iron-regulated peptidyl-prolyl-cis-trans-isomerase A	*ppiA*	Rv0009	IMR	*1.6*
Chaperone protein DnaK	*dnaK*	Rv0350	V	0.5
Succinyl-CoA synthetase beta chain	*sucC*	Rv0951	IMR	*4.1*
Succinyl-CoA synthetase alpha chain	*sucD*	Rv0952	IMR	*3.9*
Enoyl-CoA hydratase EchA9	*echA9*	Rv1071c	L	*1.8*
6-phosphogluconate dehydrogenase, decarboxylating Gnd2	*gnd2*	Rv1122	IMR	*2.9*
Integration host factor MihF	*mihF*	Rv1388	IP	*1.7*
Transaldolase	*tal*	Rv1448c	IMR	0.5
Catalase-peroxidase-peroxynitritase T KatG	*katG*	Rv1908c	V	*14*
Conserved protein	*Rv2204c*	Rv2204c	C	0.5
Trigger factor protein	*tig*	Rv2462c	CW	*3.5*
Conserved protein	*Rv2699c*	Rv2699c	C	*3.9*
Adenosylhomocysteinase	*sahH*	Rv3248c	IMR	0.4
Thiosulfate sulfurtransferase	*sseA*	Rv3283	IMR	*3.6*
3-hydroxyacyl-thioester dehydratase HtdY	*htdY*	Rv3389c	L	*3.9*
10 kDa chaperonin (protein CPN10), MPT57	*groES*	Rv3418c	V	0.8
Conserved protein	*Rv3433c*	Rv3433c	C	0.2
Conserved membrane protein	*Rv3587c*	Rv3587c	CW	0.6
Secreted fibronectin-binding protein antigen	*fbpD*	Rv3803c	L	0.4

protein

CW (n=6)				
3-oxoacyl-acyl-carrier protein] reductase FabG4	*fabG4*	Rv0242c	L	0.4
Acetyl-CoA acyltransferase FadA2	*fadA2*	Rv0243	L	0.5
Immunogenic protein MPT63	*mpt63*	Rv1926c	CW	*2.9*
ATP-dependent clp protease proteolytic subunit 2	*clpP2*	Rv2460c	IMR	0.5
Fatty-acid synthase (FAS)	*fas*	Rv2524c	L	0.3
Transcriptional regulator, crp/fnr-family	*crp*	Rv3676	R	0.1
CYT (n=6)				
Two component system transcriptional regulator PrrA	*prrA*	Rv0903c	R	0.3
5-methyltetrahydropteroyltriglutamate-homocysteine methyltransferase MetE	*metE*	Rv1133c	IMR	*5.2*
Malate dehydrogenase	*mdh*	Rv1240	IMR	*1.3*
Phosphoglycerate kinase	*pgk*	Rv1437	IMR	*1.8*
Catalase-peroxidase-peroxynitritase T KatG	*katG*	Rv1908c	V	*61*
Aminomethyltransferase	*gcvT*	Rv2211c	IMR	0.2
MEM (n=17)				
3-hydroxyacyl-thioester dehydratase HdtX	*htdX*	Rv0241c	L	0.09
Transport protein SecE2	*secE2*	Rv0379	CW	*1.3*
Polyprenyl-diphosphate synthase	*grcC1*	Rv0562	IMR	*INF[b]*
50S ribosomal protein L23, RplW	*rplW*	Rv0703	IP	0.5
Phosphoribosylformylglycinamidine synthase II	*purL*	Rv0803	IMR	*1.7*
Transcription termination factor Rho	*rho*	Rv1297	IP	*1.9*
Thioredoxin	*Rv1324*	Rv1324	IMR	*3.2*
Iron-regulated aconitate hydratase	*acn*	Rv1475c	IMR	*1.3*
Glycine dehydrogenase	*gcvB*	Rv1832	IMR	*3.6*
Catalase-peroxidase-peroxynitritase T KatG	*katG*	Rv1908c	V	*2.6*
Monophosphatase	*cysQ*	Rv2131c	IMR	*7.5*
Pyruvate dehydrogenase E1 component	*aceE*	Rv2241	IMR	*1.2*
Conserved protein	*Rv2402*	Rv2402	C	*1.6*
Chorismate synthase	*aroF*	Rv2540c	IMR	0.4
Acyl-CoA dehydrogenase FadE22	*fadE22*	Rv3061c	L	0.5
Acyl-CoA dehydrogenase FadE32	*fadE32*	Rv3563	L	0.1
Enoyl-CoA hydratase EchA21	*echA21*	Rv3774	L	*2.7*

a. The quantitative method chosen for the statistical analysis and p value calculation was NSAF. [b.] INF: NSAF in INHr strain was zero. IMR: Intermediary metabolism and respiration, V: Virulence, detoxification, adaptation, IP: Information pathways, L: Lipid metabolism, R: Regulatory protein, CW: Cell wall and cell wall processes, C: Conserved Hypothetical.

The category intermediary metabolism and respiration presented the highest number of proteins (n=20) with variable abundance among the Beijing clonal pair. In this category the enzymes from the tricarboxylic acid (TCA) cycle SucC, SucD (located in the same operon, Figure 2.6), Mdh, Acn and AceE were all decreased in the INHr strain (Figure 2.7). AceE belongs to the aerobic oxidative TCA cycle. Additionally, two enzymes of the pentose phosphate pathway Gnd2 and Tal were also significantly different in this analysis but with a different tendency (higher and lower in the INHr strain respectively) (Table 2.2).
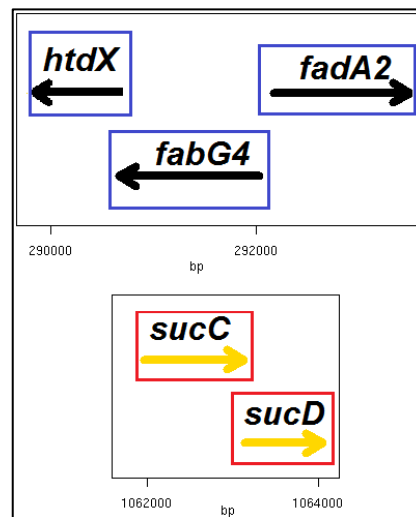
Figure 2.6. Organization of the coding genes for the proteins that were found altered in the Beijing pair comparative analysis. Blue boxes indicate that the proteins encoded by those genes were higher, while red boxes indicate they were lower in the INH resistant strain of Beijing genotype. Adapted from http://tuberculist.epfl.ch/.
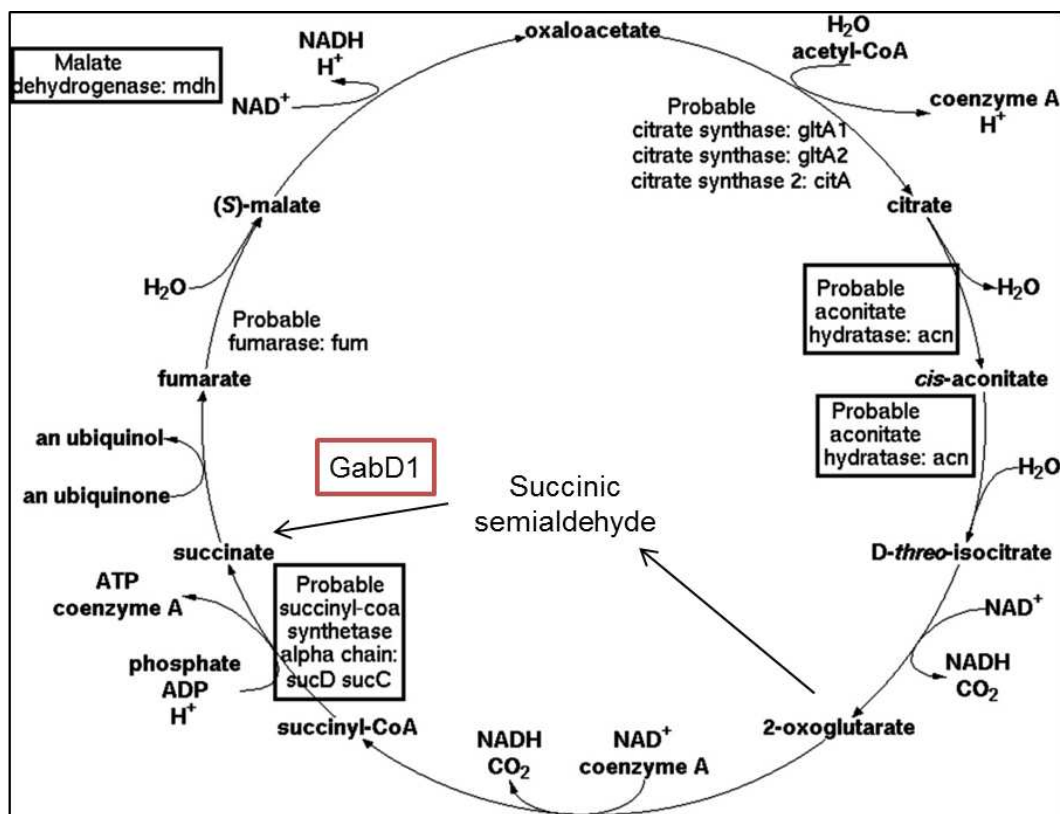
Figure 2.7. TCA cycle in *Mtb*. The enzymes in the black boxes are reduced in the Beijing INHr strain, while the red indicate the enzyme reduced in the T INHr strain. Adapted from http://biocyc.org/MTBRV and from Nieto, *et al* (62).

In the lipid metabolism category, we detected differences in proteins involved in lipid biosynthesis and β oxidation pathways. From the former, Fas and other FAS II enzymes were increased in the INHr strain (Table 2.2). Fas is a structurally integrated type I fatty acid synthase (FAS-I) with multiple domains as described in chapter I. Increased levels of enzymes from FAS II included FabG4, that participates in the elongation of fatty acids and mycolic acid synthesis and possibly has the same function of FabG1, and the 3-hydroxyacyl-thioester dehydratase HtdX. Due to the similar sequence and structure with HtdX, the enzyme HtdY was initially proposed to be part of FAS II, as another 3-hydroxyacil thioester dehydratases. However, HtdX has a higher capacity to produce lipoic acid and an increased preference for its substrate (3-hydroxyacyl-acyl carrier protein-ACP) (63). In our study, HdtX was similarly higher as the other

enzymes involved in lipid biosynthesis but HtdY was significantly decreased. The different levels observed between HtdX and HtdY may be in line with results elsewhere indicating that HdtY may not be part of the ACP-dependent FAS-II system (63). For fatty acid β oxidation, the dehydrogenases FadE22 and FadE32 and the acetyl-CoA acyltransferase FadA2 were increased while the crotonases EchA9 and EchA21 were decreased in the INHr strain (Table 2.2).

2.3.3. Common protein differences

After evaluating both clonal pair comparisons, KatG was equally altered in both INHr strains and there were some common metabolic pathways with altered proteins but with different trends at each clonal pair. KatG was reduced in both Beijing and T INHr strains compared to their respective INHs pairs. KatG is found in the three biological domains (prokaryotes, archaea and some eukaryotes) where it is mostly a bifunctional enzyme. KatG is a multifunctional enzyme with different activities, where the catalase-peroxidase functions predominate as presented in chapter 1 (64, 65). Although it is not as efficient as the catalase-peroxidase activity, KatG is also a isonicotinoyl-NAD synthase which is responsible of the INH activation (64). The implications of the lower levels of KatG have been previously discussed in section 2.3.1, such as the lack of activation INH, and the unbalanced redox metabolism.

Evaluating protein biological function, two important metabolic pathways are affected in these INHr strains. All the proteins involved in the TCA cycle found in this work were similarly reduced in the INHr strains. The Beijing INHr had five reduced proteins in this metabolic pathway (SucC, SucD, Mdh AceE, Acn) while the T INHr strain had only 1 (GabD1) (Figure 2.6, Table 2.1 and 2.2). Although there was not a protein in common or a similar trend in FAS II, this was an affected pathway in both Beijing and T INHr strains. There were proteins with

increased levels in this biosynthetic route in the Beijing INHr (FabG4 and HtdX), while

decreased levels in the T INHr strain (AcpM, Table 2.2).

Furthermore, 11 proteins were present in both comparisons albeit with different trends in

each genotype. These proteins are PpiA, DnaK, SecE2, MetE, Rho, GcvB, Mpt63, FadE22,

GroES, the probable thioredoxin Rv1324 and the conserved protein Rv2204c. (Table 2.1 and

2.2). The levels of DnaK levels were higher in the Beijing INHr strains but lower at the MEM of

the T genotype (Table 2.1 and 2.2). However, the different trends of this protein in the MEM and

the CW of the T INHr strain suggest that the difference observed of this protein could be

associated with a compartmentalization of the protein in the bacteria MEM, which cause it to be

reduced in the CW fraction. Most of these 11 proteins are grouped in the virulence,

detoxification and adaptation category (Table 2.1 and 2.2). FadE22 is a protein with different

trends in both INHr, but even in the same Beijing pair comparison this category had proteins

with different trends. The remaining proteins are classified in the intermediary metabolism and

respiration category.

2.3.4 Western Blot (WB) results

For the T clonal pair, five proteins were confirmed through WB and one protein could be

confirmed for the Beijing pair. The common protein confirmed in both clonal pairs with the same

tendency was KatG (Figure 2.8). In the case of KatG levels in the Beijing pair, they were only

significantly different at the CFP level. The protein abundances confirmed in the clinical pair

comparison includes proteins in the glutamine biosynthesis (by GlnA1), protein folding (by the

antigenic chaperone GroES) and mycolic acid synthesis (by AcpM). The enzyme PstS1, which is

involved in the acquisition of inorganic phosphate by importing inorganic phosphate in an ATP-

dependent manner, was confirmed to be higher in the membrane of the INHr T strain (Figure 2.8).
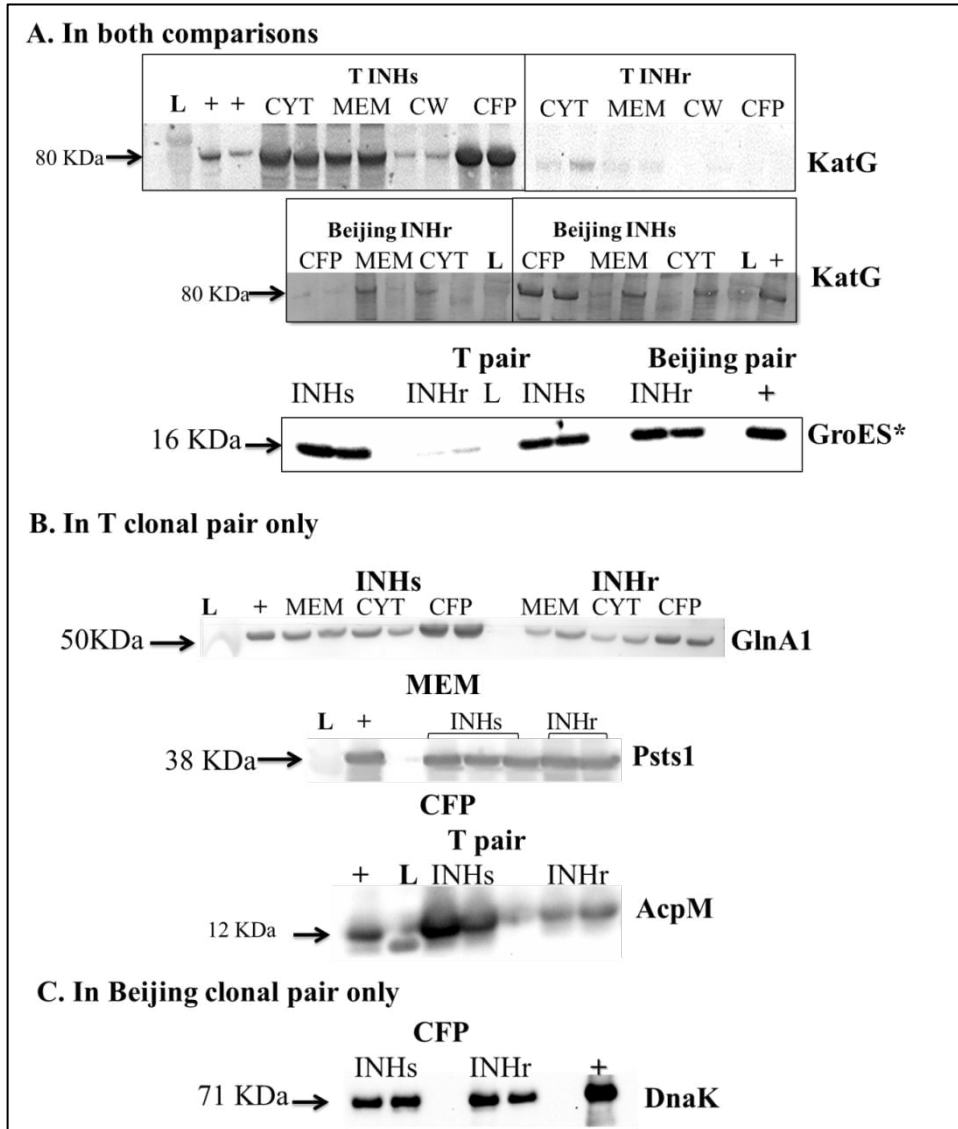
Figure 2.8. Western blot results confirming the protein differences at each cellular fraction in the INHr clonal strains of the T and Beijing genotype. L: protein standard; +, positive control.* Differences at the CFP fraction.

2.3.5 Bacterial growth

From an experimental observation, INH resistant strain of the T genotype when grown in GAS media, a minimal media without protein or non-ionic surfactant Tween 80, had a poorer

growth rate compared with its clonal strain and the Beijing pair. Moreover, when the biological

replicates were upscaled to 1 L of culture media; one of the T INHr replicates was unable to

grow. Therefore, this growth curve was intended to confirm this experimental observation, using

also the H37Rv strain as internal reference. This growth curve was done in 7H9 media

supplemented with Tween 80, in order to measure the OD reading. The results were the opposite

of what was previously observed in the GAS media. In the 7H9 supplemented with tween 80, the

T INHr strain growth was comparable to the reference strain H37Rv was even a little higher

compared to its susceptible pair (Figure 2.9).
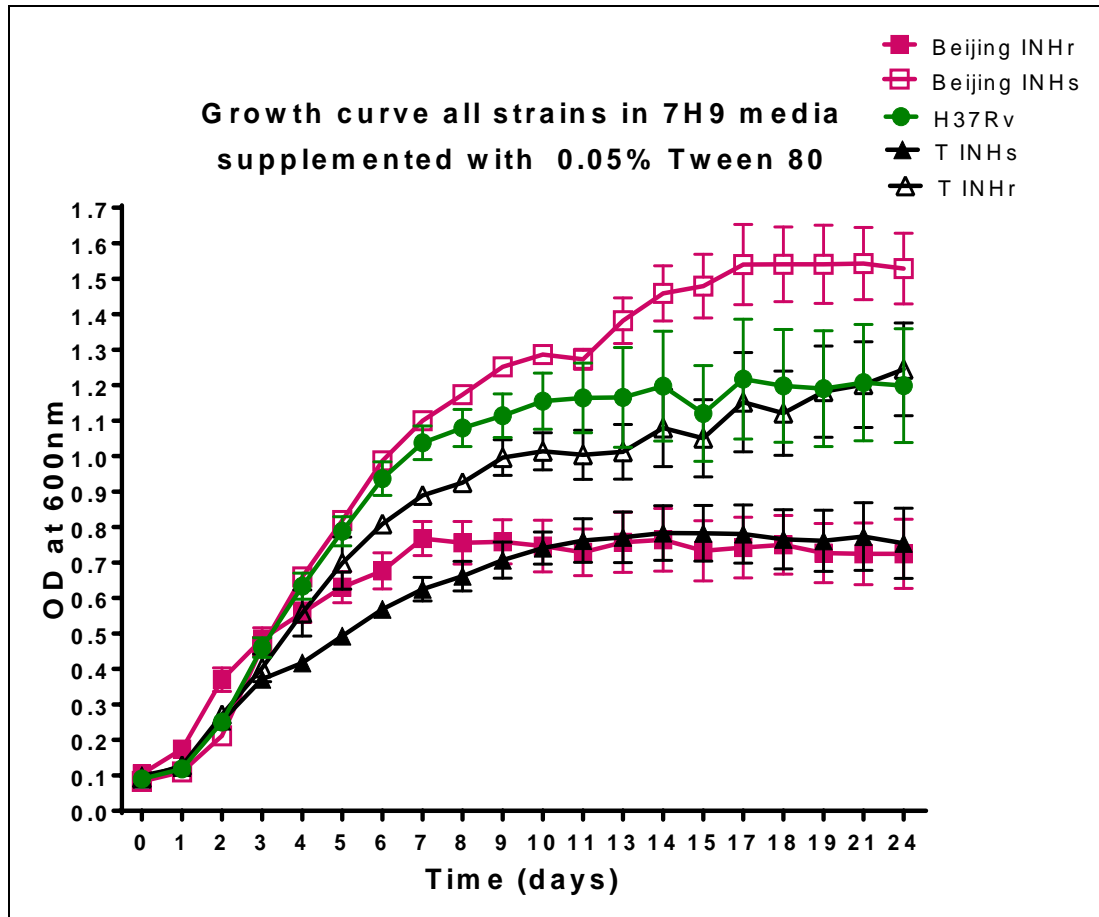
Figure 2.9. *In vitro* growth kinetics from the strains used in the study.

The higher growth rate of the T INHr opposed to the experimental observation from the same strain growing in GAS media could be explained by the fact the nonionic surfactant tween 80 is an important contributor for the T INHr strain growth (Figure 2.8). T INHr strain may have an altered metabolic profile that enhances the utilization of tween as an important substrate for energy and growth. Tween 80 benefits *Mtb* growth, not only by preventing bacteria clumps in liquid media (66), but also serving as a carbon source for the mycobacteria. The surfactant tween 80 (which is an ester of oleic acid and polyethoxylated sorbitan) can be hydrolyzed by *Mycobacterium* species into oleic acid that is metabolized via the TCA cycle for biomass synthesis (67). This may have contributed to a higher growth rate in this particular INHr strain. Correlating this with the proteomic results, we found the protein cutinase (Rv1984) higher in the CFP of the T INHr strain. This protein also known as Cfp21 has an esterase and lipase activity. *Mtb* has seven known cutinases, some of them even encoded in lipid metabolism operons (such as Clp6). Although as it is suggested by the protein name, the primary substrate for this enzyme is cutin (which is found in plants), this protein has been associated with lipid metabolism and a phospholipase activity. In fact, this protein was also shown to have the capacity to hydrolyze tween 80 (68). In that study, however, the reference strain H37Rv did not have detectable phospholipase activity at the CFP level but it did in the CW and MEM fraction, which suggests that this protein is not regularly secreted in high amounts by *Mtb* strains (68).

2.3.6 katG sequencing analysis

To confirm the driver of the INHr phenotype of those clinical strains at the genetic level, mutations in *katG* gene were further evaluated in the INHr strains compared with their parental clonal strain. In the clonal pair of the T genotype, two mutations: V1A (GTG to GCG) and E3V (GAG to GTG) at the N-terminus of the protein were identified in this work. These were the only

mutations found in the *katG* gene in two biological replicates of the T genotype INHr strain, which also had a wild type sequence in the 50 nucleotide upstream its sequence. A previous whole genome sequencing (WGS) analysis of the T clonal pair demonstrated a mutation at nucleotide position T2C (V1A) in *katG* for the INHr strain (69). As it was previously described by Wilming et al, the N-terminus of KatG is required for the protein dimerization and its interaction with the C-terminus is required for its enzymatic function (70).

In the Beijing strain, the *katG* mutation L101R was previously identified in the INHr strain in the same WGS study (69). However, the *katG* sequencing results did not show any mutation in the INHr strain. This finding was supported with a phenotypic drug susceptibility testing performed by the National Jewish laboratory that showed that both clonal Beijing pairs were drug susceptible. The previously described proteomics study was performed before the DNA analysis, therefore it is possible that after the sub-culturing of the strains for the *katG* gene analysis, the INHr went back to the wild type *katG* sequence and INHs phenotype. Therefore, this L101R *katG* mutation was not very stable for the Beijing INHr strain. A previous report of an INHr reversion in *Mtb* was already published in a *katG* mutant in the absence of the drug pressure (71).

2.4. Conclusions

This work, presents for the first time a comparative proteomic analysis of clonal clinical strains with different INH susceptibility profile, further parsing the protein differences by localization within the cell. The study of drug resistant organisms has evolved from phenotypic characterization after exposure to the drug, to the study of the genetic basis of resistance, and then the effect of the associated mutations at the protein level. The significantly reduced levels of KatG affects also all its multi-functionality (catalase, peroxidase, peroxinitritase, INH activator)

and have a direct impact in the bacterial redox homeostasis and virulence (64, 65, 72, 73). There have been many *in vitro* and *in vivo* studies that focused on the virulence and fitness cost in mutant strains of INHr *Mtb*, concluding that *katG* mutations not only generate INH resistant strains, but also bacteria with a wide range of virulence and ability to transmit; probably depending on the catalase-peroxidase activity (15, 74, 75). Accordingly, the KatG levels observed in the INHr strains of this dissertation could be associated with a less virulent phenotype.

Working with clonal pairs of *Mtb*, it was possible to compare specific protein abundance variations that were coincidental to the acquisition of INH resistance. These differences were highly dependent on the genetic background, although some common trends were observed. The Beijing pair comparison provided the less variable proteome (with only differences in 46 proteins) which can be related with the only 8 non-synonymous SNPs found in a previous whole genome sequencing (WGS) analysis (69). The Beijing INHr strain lost its resistant phenotype after serial *in vitro* cultures, therefore, it was not possible to continue studying the INHr event in this genetic background. Instead, the T pair comparison revealed a high variability (100 proteins in total) between INHs and INHr strains together with a previously reported 14 non-synonymous SNPs detected by WGS (69). The advantage of evaluating the protein differences at each cellular fraction allowed confirming if these differences were due to a confident difference of the protein levels (such as the case of KatG and GroES whose levels were affected in three out of four cellular fractions studied) or possibly a specific compartmentalization of the molecule (i.e., DnaK that was lower in the MEM but more concentrated in the CW fraction of the INHr T strain).

In summary, the findings of this chapter demonstrated that acquisition of INH resistance can result in significant changes in the mycobacteria proteome, affecting in at least four different biological pathways in *Mtb*. These pathways include the categories of metabolism and respiration, virulence, detoxification and adaptation and lipid metabolism, which may result as a compensatory mechanism after KatG reduced levels and its consequent impact on mycobacterial physiology and fitness. This work shows the mycobacterial drug resistance is a complex process that orchestrates a variety of physiological events in *Mtb* that are highly dependent on the specific genetic background. Additionally, this work also provides evidence that INH resistance is an unstable phenotype for some *Mtb* strains, such as the Beijing INHr described in this chapter that return to its wild type *katG* sequence and recovered its susceptibility to INH. This could be dependent on the specific *katG* mutation that drove to the INHr phenotype in the first place, however, the mechanism that explains this event remains unknown. This study also provides a better understanding of new compensatory mechanisms in *Mtb* after INHr that could address alternative combined therapies (such as the use of drugs targeting the ATP synthesis in susceptible instead of INHr *Mtb* strains) as well as candidates for new drug targets in INH resistant strains.

REFERENCES

1.   Organization WH. Global Tuberculosis Report2015.
2.   Salinas JL, Mindra G, Haddad MB, Pratt R, Price SF, Langer AJ. Leveling of Tuberculosis Incidence — United States, 2013–2015. MMWR Morbidity and Mortality Weekly Report. 2016;65(11):273-8.
3.   (CDC) CfDCaP. Impact of a shortage of first-line antituberculosis medication on tuberculosis control - United States, 2012-2013. MMWR Morb Mortal Wkly Rep. 2013 May;62(20):398-400.
4.   Laserson KF, Kenyon AS, Kenyon TA, Layloff T, Binkin NJ. Substandard tuberculosis drugs on the global market and their simple detection. Int J Tuberc Lung Dis. 2001 May;5(5):448-54.
5.   Bate R, Jensen P, Hess K, Mooney L, Milligan J. Substandard and falsified anti-tuberculosis drugs: a preliminary field analysis. Int J Tuberc Lung Dis. 2013 Mar;17(3):308-11.
6.   Taylor Z, Nolan CM, Blumberg HM. Controlling tuberculosis in the United States. Recommendations from the American Thoracic Society, CDC, and the Infectious Diseases Society of America. MMWR Recomm Rep. 2005 Nov 4;54(RR-12):1-81.
7.   Organization WH. Global tuberculosis report 2014.
8.   Hu Y, Hoffner S, Jiang W, Wang W, Xu B. Extensive transmission of isoniazid resistant M. tuberculosis and its association with increased multidrug-resistant TB in two rural counties of eastern China: a molecular epidemiological study. BMC Infect Dis. 2010;10:43.
9.   O'Sullivan DM, McHugh TD, Gillespie SH. The effect of oxidative stress on the mutation rate of Mycobacterium tuberculosis with impaired catalase/peroxidase function. J Antimicrob Chemother. 2008 Oct;62(4):709-12.
10.   McGrath M, Gey van Pittius NC, van Helden PD, Warren RM, Warner DF. Mutation rate and the emergence of drug resistance in Mycobacterium tuberculosis. J Antimicrob Chemother. 2014 Feb;69(2):292-302.
11.   Vilchèze C, Jacobs Jr WR. Resistance to Isoniazid and Ethionamide in Mycobacterium tuberculosis: Genes, Mutations, and Causalities. Microbiology Spectrum. 2014;2(4).
12.   Seifert M, Catanzaro D, Catanzaro A, Rodwell TC. Genetic mutations associated with isoniazid resistance in Mycobacterium tuberculosis: a systematic review. PLoS One. 2015;10(3):e0119628.
13.   Vilchèze C, Jacobs WR. Resistance to Isoniazid and Ethionamide in Mycobacterium tuberculosis: Genes, Mutations, and Causalities. Microbiol Spectr. 2014 Aug;2(4):MGM2-0014-2013.
14.   Banerjee A, Dubnau E, Quemard A, Balasubramanian V, Um KS, Wilson T, et al. inhA, a gene encoding a target for isoniazid and ethionamide in Mycobacterium tuberculosis. Science. 1994 Jan;263(5144):227-30.
15.   Gagneux S, Burgos MV, DeRiemer K, Encisco A, Munoz S, Hopewell PC, et al. Impact of bacterial genetics on the transmission of isoniazid-resistant Mycobacterium tuberculosis. PLoS Pathog. 2006 Jun;2(6):e61.
16.   Zhang Y, Fonslow BR, Shan B, Baek M-C, Yates JR. Protein Analysis by Shotgun/Bottom-up Proteomics. Chemical Reviews. 2013;113(4):2343-94.
17.   Aebersold R, Goodlett DR. Mass Spectrometry in Proteomics. Chemical Reviews. 2001;101(2):269-96.

18.   Hancock W, LaBaer J, Marko-Varga GrA. Journal of Proteome Research- 10th Anniversary. Journal of Proteome Research. 2011;10(1):1-2.

19.   Nogueira FCS, Domont GB. Survey of Shotgun Proteomics. 2014;1156:3-23.

20.   Gengenbacher M, Mouritsen J, Schubert OT, Aebersold R, Kaufmann SH. Mycobacterium tuberculosis in the Proteomics Era. Microbiol Spectr. 2014 Apr;2(2).

21.   Rauniyar N, Yates JR. Isobaric Labeling-Based Relative Quantification in Shotgun Proteomics. Journal of Proteome Research. 2014;13(12):5293-309.

22.   Zhang F, Altorki NK, Wu YC, Soslow RA, Subbaramaiah K, Dannenberg AJ. Duodenal reflux induces cyclooxygenase-2 in the esophageal mucosa of rats: evidence for involvement of bile acids. Gastroenterology. 2001 Dec;121(6):1391-9.

23.   Thompson A, Schafer J, Kuhn K, Kienle S, Schwarz J, Schmidt G, et al. Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. Anal Chem. 2003 Apr 15;75(8):1895-904.

24.   Ong SE, Blagoev B, Kratchmarova I, Kristensen DB, Steen H, Pandey A, et al. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. Mol Cell Proteomics. 2002 May;1(5):376-86.

25.   Gygi SP, Rist B, Gerber SA, Turecek F, Gelb MH, Aebersold R. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. Nat Biotechnol. 1999 Oct;17(10):994-9.

26.   Ross PL, Huang YN, Marchese JN, Williamson B, Parker K, Hattan S, et al. Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. Mol Cell Proteomics. 2004 Dec;3(12):1154-69.

27.   Mann M. Functional and quantitative proteomics using SILAC. Nature Reviews Molecular Cell Biology. 2006;7(12):952-8.

28.   Dunkley TPJ, Dupree P, Watson RB, Lilley KS. The use of isotope-coded affinity tags (ICAT) to study organelle proteomes in Arabidopsis thaliana. Biochemical Society Transactions. 2004;32(3):520-3.

29.   Liu H, Sadygov RG, Yates JR, 3rd. A model for random sampling and estimation of relative protein abundance in shotgun proteomics. Anal Chem. 2004 Jul 15;76(14):4193-201.

30.   Old WM, Meyer-Arendt K, Aveline-Wolf L, Pierce KG, Mendoza A, Sevinsky JR, et al. Comparison of label-free methods for quantifying human proteins by shotgun proteomics. Mol Cell Proteomics. 2005 Oct;4(10):1487-502.

31.   Mehaffy MC, Kruh-Garcia NA, Dobos KM. Prospective on Mycobacterium tuberculosis proteomics. J Proteome Res. 2012 Jan;11(1):17-25.

32.   Aebersold R, Mann M. Mass spectrometry-based proteomics. Nature. 2003;422(6928):198-207.

33.   Domon B, Aebersold R. Options and considerations when selecting a quantitative proteomics strategy. Nature Biotechnology. 2010;28(7):710-21.

34.   Vachet RW, Glish GL. Effects of heavy gases on the tandem mass spectra of peptide ions in the quadrupole ion trap. Journal of the American Society for Mass Spectrometry. 1996;7(12):1194-202.

35.   Kruh NA, Troudt J, Izzo A, Prenni J, Dobos KM. Portrait of a pathogen: the Mycobacterium tuberculosis proteome in vivo. PLoS One. 2010;5(11):e13938.

36.   Mattow J, Siejak F, Hagens K, Becher D, Albrecht D, Krah A, et al. Proteins unique to intraphagosomally grown Mycobacterium tuberculosis. Proteomics. 2006 Apr;6(8):2485-94.

37.   Rosenkrands I, Slayden RA, Crawford J, Aagaard C, Barry CE, 3rd, Andersen P. Hypoxic response of Mycobacterium tuberculosis studied by metabolic labeling and proteome analysis of cellular and extracellular proteins. J Bacteriol. 2002 Jul;184(13):3485-91.

38.   Albrethsen J, Agner J, Piersma SR, Højrup P, Pham TV, Weldingh K, et al. Proteomic profiling of Mycobacterium tuberculosis identifies nutrient-starvation-responsive toxin-antitoxin systems. Mol Cell Proteomics. 2013 May;12(5):1180-91.

39.   Zhang C, Song X, Zhao Y, Zhang H, Zhao S, Mao F, et al. Mycobacterium tuberculosisSecreted Proteins As Potential Biomarkers for the Diagnosis of Active Tuberculosis and Latent Tuberculosis Infection. Journal of Clinical Laboratory Analysis. 2015;29(5):375-82.

40.   Tucci P, GonzÃ¡lez-Sapienza G, Marin M. Pathogen-derived biomarkers for active tuberculosis diagnosis. Frontiers in Microbiology. 2014;5.

41.   Olsen AW, van Pinxteren LAH, Okkels LM, Rasmussen PB, Andersen P. Protection of Mice with a Tuberculosis Subunit Vaccine Based on a Fusion Protein of Antigen 85B and ESAT-6. Infection and Immunity. 2001;69(5):2773-8.

42.   Jiang X, Zhang W, Gao F, Huang Y, Lv C, Wang H. Comparison of the proteome of isoniazid-resistant and -susceptible strains of Mycobacterium tuberculosis. Microb Drug Resist. 2006 Winter;12(4):231-8.

43.   Domon B, Aebersold R. Options and considerations when selecting a quantitative proteomics strategy. Nat Biotechnol. 2010 Jul;28(7):710-21.

44.   Chindelevitch L, Colijn C, Moodley P, Wilson D, Cohen T. ClassTR: Classifying Within-Host Heterogeneity Based on Tandem Repeats with Application to Mycobacterium tuberculosis Infections. PLoS Comput Biol. 2016 Feb;12(2):e1004475.

45.   Burgos M, DeRiemer K, Small PM, Hopewell PC, Daley CL. Effect of drug resistance on the generation of secondary cases of tuberculosis. J Infect Dis. 2003 Dec 15;188(12):1878-84.

46.   Kamerbeek J, Schouls L, Kolk A, van Agterveld M, van Soolingen D, Kuijper S, et al. Simultaneous detection and strain differentiation of Mycobacterium tuberculosis for diagnosis and epidemiology. J Clin Microbiol. 1997 Apr;35(4):907-14.

47.   van Soolingen D, de Haas PE, Kremer K. Restriction fragment length polymorphism typing of mycobacteria. Methods Mol Med. 2001;54:165-203.

48.   Kent PT, and George P Kubica. Public Health Mycobacteriology: A Guide for the Level III Laboratory. US Dept of Health and Human Services, Public Health Service, Centers for Disease Control. 1985.

49.   Lucas MC, Wolfe LM, Hazenfield RM, Kurihara J, Kruh-Garcia NA, Belisle J, et al. Fractionation and analysis of mycobacterial proteins. Methods Mol Biol. 2015;1285:47-75.

50.   Bisson GP, Mehaffy C, Broeckling C, Prenni J, Rifat D, Lun DS, et al. Upregulation of the phthiocerol dimycocerosate biosynthetic pathway by rifampin-resistant, rpoB mutant Mycobacterium tuberculosis. J Bacteriol. 2012 Dec;194(23):6441-52.

51.   Chambers MC, Maclean B, Burke R, Amodei D, Ruderman DL, Neumann S, et al. A cross-platform toolkit for mass spectrometry and proteomics. Nat Biotechnol. 2012 Oct;30(10):918-20.

52.   Keller A, Nesvizhskii AI, Kolker E, Aebersold R. Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. Anal Chem. 2002 Oct 15;74(20):5383-92.

53.   Nesvizhskii AI, Keller A, Kolker E, Aebersold R. A statistical model for identifying proteins by tandem mass spectrometry. Anal Chem. 2003 Sep 1;75(17):4646-58.

54.   Zhang Y, Wen Z, Washburn MP, Florens L. Refinements to label free proteome quantitation: how to deal with peptides shared by multiple proteins. Anal Chem. 2010 Mar 15;82(6):2272-81.

55.   Bald D, Koul A. Respiratory ATP synthesis: the new generation of mycobacterial drug targets? FEMS Microbiol Lett. 2010 Jul;308(1):1-7.

56.   Andries K. A Diarylquinoline Drug Active on the ATP Synthase of Mycobacterium tuberculosis. Science. 2005;307(5707):223-7.

57.   Wang R, Marcotte EM. The Proteomic Response ofMycobacterium smegmatisto Anti-Tuberculosis Drugs Suggests Targeted Pathways. Journal of Proteome Research. 2008;7(3):855-65.

58.   Tian J, Bryk R, Itoh M, Suematsu M, Nathan C. Variant tricarboxylic acid cycle in Mycobacterium tuberculosis: identification of alpha-ketoglutarate decarboxylase. Proc Natl Acad Sci U S A. 2005 Jul 26;102(30):10670-5.

59.   Mdluli K. Inhibition of a Mycobacterium tuberculosis -Ketoacyl ACP Synthase by Isoniazid. Science. 1998;280(5369):1607-10.

60.   Camus JC, Pryor MJ, Médigue C, Cole ST. Re-annotation of the genome sequence of Mycobacterium tuberculosis H37Rv. Microbiology. 2002 Oct;148(Pt 10):2967-73.

61.   Vizcaino JA, Cote R, Reisinger F, Barsnes H, Foster JM, Rameseder J, et al. The Proteomics Identifications database: 2010 update. Nucleic Acids Res. 2010 Jan;38(Database issue):D736-42.

62.   Nieto RL, Mehaffy C, Dobos KM. Comparing isogenic strains of Beijing genotype Mycobacterium tuberculosis after acquisition of Isoniazid resistance: A proteomics approach. Proteomics. 2016 May;16(9):1376-80.

63.   Gurvitz A, Hiltunen JK, Kastaniotis AJ. Heterologous Expression of Mycobacterial Proteins in Saccharomyces cerevisiae Reveals Two Physiologically Functional 3-Hydroxyacyl-Thioester Dehydratases, HtdX and HtdY, in Addition to HadABC and HtdZ. Journal of Bacteriology. 2009;191(8):2683-90.

64.   Njuma OJ, Ndontsa EN, Goodwin DC. Catalase in peroxidase clothing: Interdependent cooperation of two cofactors in the catalytic versatility of KatG. Archives of Biochemistry and Biophysics. 2014;544:27-39.

65.   Nagy JM, Cass AE, Brown KA. Purification and characterization of recombinant catalase-peroxidase, which confers isoniazid sensitivity in Mycobacterium tuberculosis. J Biol Chem. 1997 Dec 12;272(50):31265-71.

66.   DUBOS RJ, DAVIS BD. Factors affecting the growth of tubercle bacilli in liquid media. J Exp Med. 1946 May;83:409-23.

67.   Tang YJ, Shui W, Myers S, Feng X, Bertozzi C, Keasling JD. Central metabolism in Mycobacterium smegmatis during the transition from O2-rich to O2-poor conditions as studied by isotopomer-assisted metabolite analysis. Biotechnol Lett. 2009 Aug;31(8):1233-40.

68.   Parker SK, Curtin KM, Vasil ML. Purification and Characterization of Mycobacterial Phospholipase A: an Activity Associated with Mycobacterial Cutinase. Journal of Bacteriology. 2007;189(11):4153-60.

69.   Datta G, Nieto LM, Davidson RM, Mehaffy C, Pederson C, Dobos KM, et al. Longitudinal whole genome analysis of pre and post drug treatment Mycobacterium tuberculosis isolates reveals progressive steps to drug resistance. Tuberculosis (Edinb). 2016 May;98:50-5.

70.   Wilming M, Johnsson K. Inter- and intramolecular domain interactions of the catalase-peroxidase KatG from M. tuberculosis. FEBS Lett. 2001 Dec 7;509(2):272-6.

71.   Richardson ET, Lin SY, Pinsky BA, Desmond E, Banaei N. First documentation of isoniazid reversion in Mycobacterium tuberculosis. Int J Tuberc Lung Dis. 2009 Nov;13(11):1347-54.

72.   Cade CE, Dlouhy AC, Medzihradszky KF, Salas-Castillo SP, Ghiladi RA. Isoniazid-resistance conferring mutations inMycobacterium tuberculosisKatG: Catalase, peroxidase, and INH-NADH adduct formation activities. Protein Science. 2010:NA-NA.

73.   Ng VH, Cox JS, Sousa AO, MacMicking JD, McKinney JD. Role of KatG catalase-peroxidase in mycobacterial pathogenesis: countering the phagocyte oxidative burst. Molecular Microbiology. 2004;52(5):1291-302.

74.   Ordway DJ, Sonnenberg MG, Donahue SA, Belisle JT, Orme IM. Drug-resistant strains of Mycobacterium tuberculosis exhibit a range of virulence for mice. Infect Immun. 1995 Feb;63(2):741-3.

75.   Pym AS, Saint-Joanis B, Cole ST. Effect of katG Mutations on the Virulence of Mycobacterium tuberculosis and the Implication for Transmission in Humans. Infection and Immunity. 2002;70(9):4955-60.

CHAPTER III – COMPREHENSIVE BIOCHEMICAL CHARACTERIZATION OF M*ycobacterium tuberculosis* STRAINS AFTER ACQUISITION OF ISONIAZID RESISTANCE

3.1 Introduction

During the first and second chapter of this dissertation, the importance of catalase-peroxidase KatG in isoniazid (INH) activation, as well as *katG* mutations and INH resistance, were addressed through reviewing of related literature (Chapter I) and an initial shotgun proteomics experimental approach (Chapter II). In chapter II, the study of two clinically-isolated clonal pairs with different genotypes, unexpectedly did not generate many common protein findings between the two INH resistant (INHr) strains, with the exception of decreased KatG levels .The lack of common protein findings could be explained by the genetic background differences that include a different mutation in the *katG* gene, as well as the fact that they belong to different *Mtb* lineages (Beijing vs T genotype). Due to the reversion of the INH resistance in the Beijing clonal pair, it was not possible to complete a biochemical characterization using this pair. Instead, a laboratory-isolated clonal pair with a *katG* mutation similar to that of the T genotype clinical INHr clonal pair is included in the present chapter. Having two different clonal pairs, one clinical and one laboratory generated, with the same *katG* mutation, offers the convenience of confirming protein changes specifically related to the INHr phenotype caused by N-terminus *katG* mutations while eliminating confounding factors including previous exposure to other anti-TB drugs. Because of the already established relationship between the synthesis of mycolic acids (MA) and INH mechanism of action (See Chapter I), this chapter will describe a broader biochemical analysis using a high-resolution shotgun proteomics method and a comparison of the MA profiles between each of the clonal pairs.

As described in the previous chapter, mass analyzers are central to protein analysis. Furthermore, new developments in this field has created instrumentation with higher accuracy, throughput capacity, and resolving power. Mass spectrometry (MS)-based proteomics has facilitated significant advancement in various biological sciences employing soft ionization of peptides; notably, this method was awarded with a Nobel Prize in 2002 (1). Soft ionization strategies such as electrospray ionization (ESI) convert peptides into ionized compounds that can be transferred to the MS analyzer where they are discriminated by their mass over charge (m/z) values. MS instruments most commonly used in proteomics analyze samples in a liquid phase using ESI. An important MS experimental variable is the resolution, which is a function of the mass analyzer's resolving power (i.e., the ability to separate two peaks or near-isobaric species to determine their mass accurately) (2, 3). Therefore, a larger mass analyzer resolution indicates a better separation of peaks (4). This concept is particularly relevant for the work presented in this chapter because a hybrid linear ion trap (LTQ)-Orbitrap is used for the proteomic analysis, an instrument with a higher resolution than the LTQ alone (discussed in Chapter II).

Both LTQ and Orbitrap are ion trap mass spectrometers, but they differ in many aspects. In ion traps, ions are captured for the MS/MS analysis and the sensitivity is related to the number of ions that can be trapped. The LTQ uses dynamic electric fields to capture the ionized peptides. Specifically, four symmetrically parallel rod-like electrodes (commonly known as a quadrupole) are used to constrain ions in a cylindrical volume. This feature increases the sensitivity, resolution and mass accuracy of the LTQ compared to the 3D ion trap analyzer (1). The resolution depends on the mass and in the quadruple it could be as high as 800 at m/z=130 m/z or as low as 250 at m/z=720 (5, 6). The four electrodes act as a mass filter that can be set to allow ions with selected m/z to pass through. The Orbitrap mass analyzer was invented by Alexander

Makarov and first introduced in 2005 (7). The Orbitrap was included in high-performance hybrid spectrometers that have higher mass accuracy and resolution than the LTQ alone. The Orbitrap is an electrostatic trapping mass spectrometer where the ions are trapped and oscillate on a central electrode in a limited space defined by two external electrodes (3, 7). This is a smaller and faster device than the LTQ, where the oscillation of the ions induces an image current into the outer electrodes which is then amplified and detected for the proteomic analysis. The hybrid system used in this work operates in an Orbitrap-LTQ mode having advantages of both methodologies. Because of its high resolution and mass accuracy (resolution of 100,000 at m/z=400 and mass accuracy $\leq$ 1 ppm with internal calibration) (8), the Orbitrap is used for the first MS for the determination of the peptide masses. Following this step, fragmentation of the peptides using a collision gas occurs in the LTQ analyzer, resulting in MS/MS spectra used to determine amino acid sequence.

In addition to the proteomic analysis, and as part of a comprehensive characterization, a comparison of MA in the clonal pairs was performed using liquid chromatography (LC)-time of flight (TOF) MS. TOF is another MS-based technique that can be also coupled with matrix-assisted laser desorption ionization (MALDI), when samples are in a solid phase. For this analysis, the TOF instrument was paired with ESI source (the same ionization strategy used for the proteomic analysis). The MA are separated using reverse-phase High-performance LC (HPLC) and then subjected to high voltage in a cone-shaped space to transform the sample from the liquid phase to aerosol. During this process, the particles in the droplet become more charged and then are separated from each other as the solvent evaporates, creating single charged ions for further analysis in the MS. The TOF analyzer consists of a tube divided in two sections: a short region with high electric field where the ions are accelerated and a longer electric field-free

section where the ions "fly" (9). The time that a given ion takes to travel across the longer region and reach the final detector depends on its m/z, therefore, based on this time the m/z is calculated for a particular ion. MS is a popular and well-suited methodology for lipidomics, a discipline that investigates changes in lipid abundance, structure and function in the cell.

Lipids are diverse molecules that are further categorized according to their chemical properties and complexity. Specifically, lipids are categorized based on chemical structure and mechanism of biosynthesis (10). Particularly for lipids of *Mtb*, a classification method proposed by Mark Sartain *et al* (available at http://www.lipidmaps.org and http://mrl.colostate.edu/mtb/) proceeds hierarchically from six categories (fatty acyls, glycerophospholipids, glycolipids, prenol lipids, saccharolipids, and polyketides), to 15 main lipid classes, 46 lipid subclasses, and 16 4th level lipid classes (11). Mycolic acids (MA) form a subclass in the fatty acyl category with six 4th level lipid classes (including the entire free and covalently attached MA) (11). Another *Mtb* lipid database was created around the same time by Branch Moody's group, called MycoMass. This database also allows the identification of lipids based on their m/z value (12) and complements Sartain's database by including the m/z values for α'-MA that are MA species with a shorter mycolate chain.

This chapter describes the complete biochemical exploration of the *Mtb* phenotype after acquisition of INH resistance, specifically targeting the global analysis of proteins and MA. The idea behind the exploration of INH resistance development and maintenance in *Mtb* is supported by the fact that INHr strains are estimated to be causing disease in 9.5% of all TB cases globally (13). These INH resistant *Mtb* strains have a higher probability to acquire further mutations that drive multidrug resistant (MDR) TB cases which have more deadly outcomes (14, 15). Therefore, the understanding of the INH resistance event is urgently needed with the final goal of

the design of more directed treatment strategies that effectively control and eliminate a

proportion of *Mtb* strains that increase the morbidity and mortality rates due to TB worldwide.

3.2 Materials and Methods

3.2.1 Bacterial strains and preparation of Culture Filtrate Proteins (CFP) and cellular fractions

One clonal clinical and one clonal lab pair of *Mtb* were separated into the different cellular

fractions as described in chapter II. The clinical pair were isolated from the same TB patient and

belong to the T lineage by restriction fragment length polymorphism RFLP -IS6110 (16) and

spoligotyping (17). The second strain, obtained after drug treatment failure, was resistant to INH

only, while the initial isolate was susceptible to all drugs tested (this pair was also described in

Chapter II). The other *Mtb* clonal pair was a laboratory pair consisting of wild type H37Rv

reference strain and a *katG* mutant generated from this strain after exposure to INH in the mouse

model of TB (18). The laboratory strain H37Rv (Trudeau Mycobacteria Culture Collection #

107) was part of our own bacterial collection and its mutant pair was provided as a kind donation

from Doctors Gyanu Lamichhane and Eric Nuermberger at the Center of Tuberculosis Research

of Johns Hopkins University (Baltimore, MD). In this section, the INH susceptible and resistant

Mtb strains will be denoted as INHs and INHr, respectively. The laboratory INHr strain has a

mutation in the first codon of the *katG* gene (V1A) (18). A PCR-sequencing analysis of the *katG*

gene confirmed the mutation for the lab-INHr strain and revealed two mutations for the clinical

INHr strain: V1A and E3V (19).

For the proteomic analysis, all *Mtb* strains were cultured in triplicate in a final volume of

1L of Glycerol Alanine Salts (GAS) media and incubated at 37$^{\text{o}}$C in constant agitation. The

preparation of CFP and cellular fractions were performed as described in chapter II. The cellular

fractions include cytosol (CYT), membrane (MEM) and cell wall (CW) and were prepared after

the confirmation of bacterial death using the Alamar Blue assay (Invitrogen).

3.2.2 In-solution trypsin digestion

Total protein concentration of all cellular fractions and CFPs was measured using the

bicinchoninic acid method (Thermo Scientific™Pierce™BCA Protein Assay BCA). The

protocol for in solution digestion of proteins was performed as described in Chapter II. Briefly,

starting with 30 µg of each secreted and cell fraction protein sample the sequential steps of

acetone precipitation, denaturation (with urea), reduction (with dithiothreitol), alkylation (with

idoacetamide), trypsin digestion (using a mix of trypsin and ProteaseMaxTM Surfactant) and a

final desalting step were followed.

3.2.3 Liquid chromatography coupled with tandem mass spectrometry assay (LC-MS/MS)

One microliter (0.5 µg) of digested peptides from cellular fractions and CFPs for all the

three biological replicates were randomly injected in duplicate using the Orbitrap Velos MS

coupled with nano-HPLC instrument (Thermo Scientific) at the Proteomics and Metabolomics

Facilities (PMF), Colorado State University. Each sample was injected using an EASY nanoLC-

II system (Thermo Scientific, San Jose, CA). Peptides were purified and concentrated using an

on-line enrichment column (EASY-Column, 100 µm ID × 2 cm ReproSil-Pur C18). Subsequent

chromatographic separation was performed on a reverse phase nanospray column (EASY-

Column, 3 µm, 75 µm ID × 100mm ReproSil-Pur C18) using a 90 minute linear gradient from

5%-45% buffer B (100% ACN, 0.1% formic acid) at a flow rate of 400 nanoliters/min. Peptides

were eluted directly into the mass spectrometer (Thermo Scientific Orbitrap Velos). The

instrument was operated in Orbitrap-LTQ mode where precursor measurements were acquired in

the Orbitrap (60,000 resolution) and the tandem MS/MS spectra (top 20) were acquired in the LTQ ion trap with a normalized collision energy of 35%.

3.2.4 Database Searching

Tandem mass spectra raw data were converted to mzXML files using ProteoWizard (20). All MS/MS samples were analyzed using Sequest (Thermo Fisher Scientific, San Jose, CA, USA; version 1.0). Sequest was set up to search the MtbReverse041712 database (7992 entries) assuming trypsin as the digestion enzyme with up to two missed cleavage sites. Sequest was searched with a fragment ion mass tolerance of 1.00 Da and a parent ion tolerance of 20 PPM. Oxidation of methionine and carbamidomethyl of cysteine were specified in Sequest as variable modifications.

3.2.5 Criteria for protein identification and statistical analysis

Scaffold (version Scaffold_4.3.2, 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 95.0% probability by the Scaffold Local FDR algorithm. Protein identifications were accepted if they could be established at greater than 99.0% probability to achieve an FDR less than 1.0% and contained at least 2 identified peptides. Protein probabilities were assigned by the Protein Prophet algorithm (21). Proteins that contained similar peptides and could not be differentiated based on MS/MS analysis alone were grouped to satisfy the principles of parsimony. A brief description of the procedure for the proteomic analysis is shown in figure 1A.

Differences between protein abundances, expressed as normalized spectra abundance factors (NSAF values) among the two different phenotypes (susceptible versus resistant to INH) were tested by two tailed Student's t test. NSAF concept and description was described in

chapter II (22). Additional fold change was included to visualize if the proteins levels were higher or lower in the INHr strain.

3.2.6 Western blot validation

For western blot analysis, 5 µg of total protein of cytosol and CFP of two biological replicates for each strain were transferred into a nitrocellulose membrane. Anti-*Mtb* KatG (clone IT-57) mouse monoclonal antibody was used as the primary antibody at a 1:1,000 dilution (available through Biodefense and Emerging Infections Research Resources Repository-BEI, https://www.beiresources.org/), the secondary antibodies Goat Anti-Rabbit F(ab)2 fragment (Thermo Scientific) and Goat anti-Mouse IgG (H+L) (Thermo Scientific)-HRP conjugated were used with the chemiluminescent detection system (SuperSignal West Pico Stable peroxide solution, Thermo Scientific Pierce). Recombinant KatG protein (also from BEI) was used as positive control.

3.2.7 Peroxidase activity testing

For the peroxidase activity assay, each sample was diluted to 1.6 µg/µl of total protein with 1X PBS buffer and catalase-peroxidase activity was evaluated at 25°C using 3,3′,5,5′-Tetramethylbenzidine liquid substrate system (TMB, Sigma-Aldrich) that already contains hydrogen peroxide (H_2O_2) in a 1:1 volume ratio. After 15 minutes incubation, the reaction was stopped with 0.5 M H_2SO_4 solution and then the optical density (OD) was read at 450 nm. Additionally, a standard curve was prepared by making five serial dilutions of recombinant KatG within a concentration range from 0.125 mg/mL to 2 mg/mL.

A linear regression was built using the OD values from different concentrations of recombinant KatG to interpolate the KatG enzymatic activity in the cytosol and culture supernatants of the laboratory and clinical pairs. Differences for peroxidase activity in each

100

clonal pair were compared using an unpaired t-test. Statistical analysis and graphs were done in Graphpad Prism® version 6.05.

3.2.8 Lipidomics

For this section the clonal pairs were cultured in 100 mL of GAS media and incubated under the same conditions of the cultures used for the proteomic analysis. After removing the excess of media by centrifugation (at 3 *x g* for 20 minutes at 4°C), the bacterial cell pellets were incubated with 6mL of chloroform/methanol/water 10:10:3 overnight at room temperature in constant agitation to inactivate the bacteria and extract the total lipids, as previously described (11). After drying the lipid solutions under the nitrogen bath, samples were diluted in chloroform/methanol 2:1. The overview of the lipidomic approach is summarized in figure 3.1.
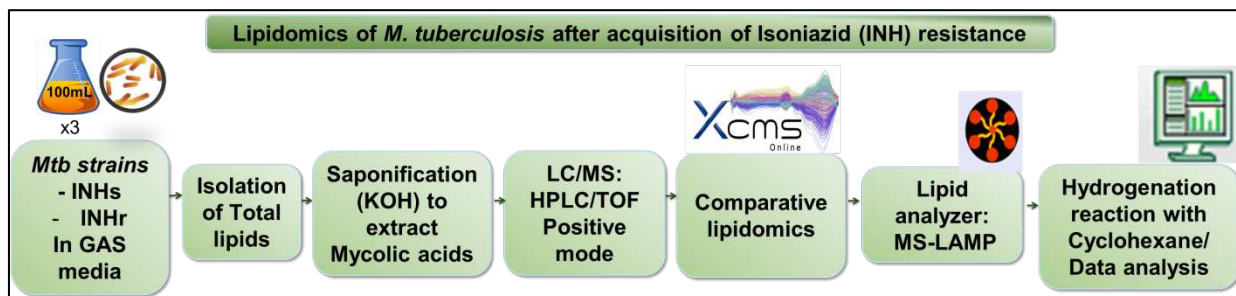
Figure 3.1. Overview of the lipidomics approach.

3.2.8.1 Mycolic acids extraction- The total lipid solutions previously diluted were dried again under the nitrogen gas. Dried lipids were mixed with 2 mL of 1M KOH (Sigma, dissolved in methanol) and incubated for 2 hours at -80°C. After setting the tubes at room temperature for 20 minutes, pH was adjusted to 5.0 and 2 mL of diethyl ether (Sigma) were added to extract the saponified material. The diethyl ether extraction step was repeated, collecting the upper layer in both cases in the same tube. After this, 2 mL of HPLC water were added and the upper (organic) layer was carefully separated. This layer was then transferred to a pre-weighted glass tube and

this resulting solution of mycolic acids was dried under nitrogen gas. The MA samples were dried and diluted to 1 µg/µl in solvent B (5 mM ammonium acetate in n-propanol-hexane-water (79:20:1; v/v/v)). This protocol was kindly provided by the laboratory of Dr. John Belisle, Colorado State University.

3.2.8.2 LC-MS analysis- Following Sartain's method (11) with minor modifications, five µl of mycolic acid solution (1µg/µl, in solvent B) were used for the LC-TOF analysis. The autosampler was kept at 20°C. A Water's XBridge (Ethylene-Bridged Hybrid, BEH) C18 column (2.1 x 100 mm, 2.5 µm) was used for chromatographic separation of lipids at 45 °C, over 35 min with a flow rate of 0.32 ml/min. The system was equilibrated with 80% solvent A (5 mM ammonium acetate in methanol-water 99:1; v/v) and 20% solvent B (5 mM ammonium acetate in n-propanol-hexane-water (79:20:1; v/v/v)). Solvent A was maintained at 80% for 2.0 min, followed by a 30.0 min linear gradient to 80% solvent B. An Agilent 6230 time-of-flight (TOF) mass spectrometer for the mass analysis of the LC eluent was used. Mass calibration was performed with an Agilent tune mix from 100 to 2,700 Da. The conditions for the TOF analyzer were established following Mark Sartain protocol, positive (+) ion data were generated by operation of the mass spectrometer in a dual ESI mode with a capillary voltage of 3500 V, nebulizer of 45 psig, drying gas of 8 l/min, gas temperature of 310°C, fragmentor of 120 V, skimmer of 60 V, and octopole radio frequency voltage of 250 V. Mass spectra were acquired at 1.2 spectra/s and 8,297 transients/spectrum.. Data were collected over a mass range of 250 to 3,200 Da with the Agilent MassHunter WorkStation Data Acquisition software, version B.05.

3.2.8.3 Hydrogenation reaction-In order to further validate the identification of MA with the same m/z value but with discrepant retention time values, a hydrogenation reaction was performed to evaluate the possible unsaturation in the meromycolate chain as the reason for the

observed discrepancy. The reduction of the double bounds is usually performed using hydrogen gas, but because of the highly flammable nature of this reaction which requires very specific glassware and laboratory setting, an alternative reaction was followed. This alternative hydrogenation reaction uses cyclohexane as a donor for the hydrogen and the palladium on carbon Pd/C works as catalyst of the reaction. Palladium is the most active catalyst for a transfer reduction. The use of cyclohexane as a hydrogen donor has been used in organic chemistry since 1975 (23). Based on this, the expected reaction would occur as in figure3.2. In this reaction, two µg of MA were resuspended in 3 ml of cyclohexane and mixed with about 1mg of 10% Pd/C (Sigma-Aldrich) in a 13x100 test tube. The tube was sealed heated for 30 min at 80°C. A centrifugation at 3,000 x g at room temperature and a filtration step was followed to remove the Pd/C. Remaining particles of Pd/C were removed by filtering the solution through a column of approximately 3 cm of celite in a Pasteur pipet. The filtered solution was collected in a new pre-weighted 13 x100 tube. The resultant translucid liquid was then dried in a nitrogen bath and resuspended in solvent B to a final concentration of 1µg/µl for LC/MS analysis. LC/MS analysis was performed in an Agilent 6220 with the same conditions previously described above for the Agilent 6230 instrument. The correspondent original non-hydrogenated samples were simultaneously analyzed to have a more accurate comparison.
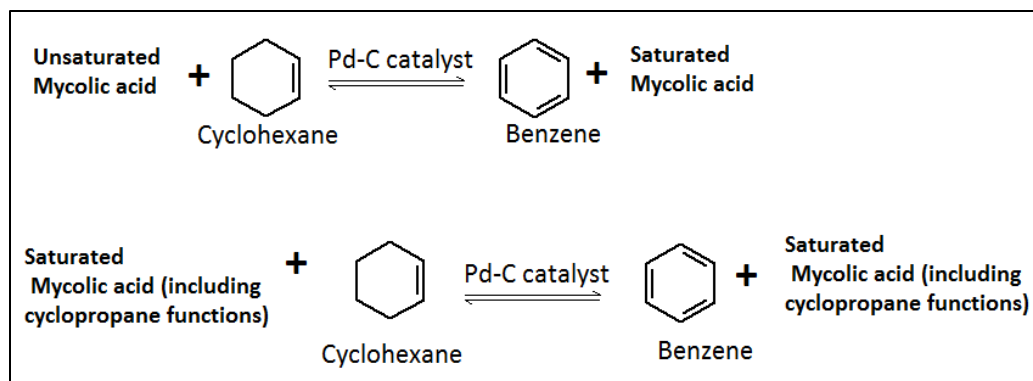
Figure 3.2. Reactions that are expected to occur with the hydrogenation strategy of mycolic acid

(MA). As it is described, only those unsaturated MA are expected to change after the hydrogenation reaction.

3.2.8.4 *Data analysis and statistical comparison-* All raw data was converted to mzXML files using proteowizard (20). Then data was uploaded to the XCMS online platform version 2.2.1 (https://xcmsonline.scripps.edu/) (24) for the mycolic acids abundance comparison in the clonal pair samples. This platform is also able to calculate statistical differences using unpaired Welch t test (assuming unequal variances). Significantly different features ($p<0.05$) with 2-fold difference or higher between INHs and INHr strains were exported and uploaded into Mass Spectrometry based Lipid(ome) Analyzer & Molecular Platform (MS-LAMP) using the "Query type 2" option (25). This platform contains the *Mtb* lipid database that is also found at http://www.mrl.colostate.edu/ and allows for the identification of specific *Mtb* lipids according to their m/z values. For the feature identification, the m/z values were interrogated against different adducts (M+H)+, (M+Na) (M+2Na-H)+, and (M+NH4)+. Those features that were not identified using MS-LAMP platform, were interrogated against MycoMass (http://www.brighamandwomens.org/research/depts/medicine/rheumatology/Labs/Moody/defaul t.aspx). The comparison of the extracted ion chromatograms for the significantly different features before and after hydrogenation was done in Agilent MassHunter Qualitative Analysis B.04.00.

3.3 Results and Discussion

3.3.1 Shared protein trends in the Mtb *INHr strains*

When comparing INHs vs INHr strains in each group (clinical and laboratory strains), numerous proteins have significantly different abundance (t-test, $p<0.05$). Every protein with significantly different abundance was classified according with its function or metabolic pathway involved. In general, more proteins with affected abundances were found in the clinical

comparison versus the laboratory comparison (Figure 3.3). A total of 45 proteins with different

abundances were shared between both clinical and laboratory comparisons, looking at each

cellular fraction (some proteins were found in more than one cellular fraction) (Table 3.1).

Metabolism and respiration as well as lipid metabolism were the most affected functional

categories of proteins (Figure 3.3B). In order to clearly present the protein findings, the next

figures in this chapter describe the commonly altered proteins that participate in particular

metabolic routes or cell functions in black boxes. Outside the boxes, significant differences in

protein levels for each clonal pair comparison are described. The commonly affected proteins

represent particular functions in the cell such as virulence and detoxification proteins (Figure

3.4), energy metabolism and respiration (Figure 3.5), lipid metabolism (Figure 3.6 and 3.7), S-

adenosyl-methionine (SAM) metabolism (Figure 3.8), β-oxidation and proteins in the lipid

degradation (Figure 3.9 and 3.10), regulatory proteins (Table 3.2) and lipoproteins (Figure 3.11).

After that, proteins affected without similar trends and specific protein differences for each

clonal pair comparison are presented.

Increased levels of EphD and EphF suggest a higher oxidation rate of fatty acids under

low KatG levels. Using a more robust proteomic strategy, a strong reduction of the INH activator

(KatG) in both INHr strains was now observed at all cellular fractions, concurrent with

compensatory action of related enzymes in the virulence and detoxification category (Figure

3.4). KatG, as shown in Table 3.1 and Figure 3.4, is significantly reduced in CYT, CW, MEM

and CFP of both INHr strains, suggesting that the single V1A mutation at the first codon is able

to cause a reduction of this important redox enzyme. On the other hand, there were four similarly

increased enzymes in both clinical and laboratory INHr strains, including the dehydrogenases

EphF and EphD and two probable oxidoreductases, possibly to compensate the lack of KatG

(Figure 3.4). Interestingly, EphF and EphD are involved in detoxification reactions following the degradation of lipids, a category that, as it will be explored later, is increased in both INHr strains.

A.

B.
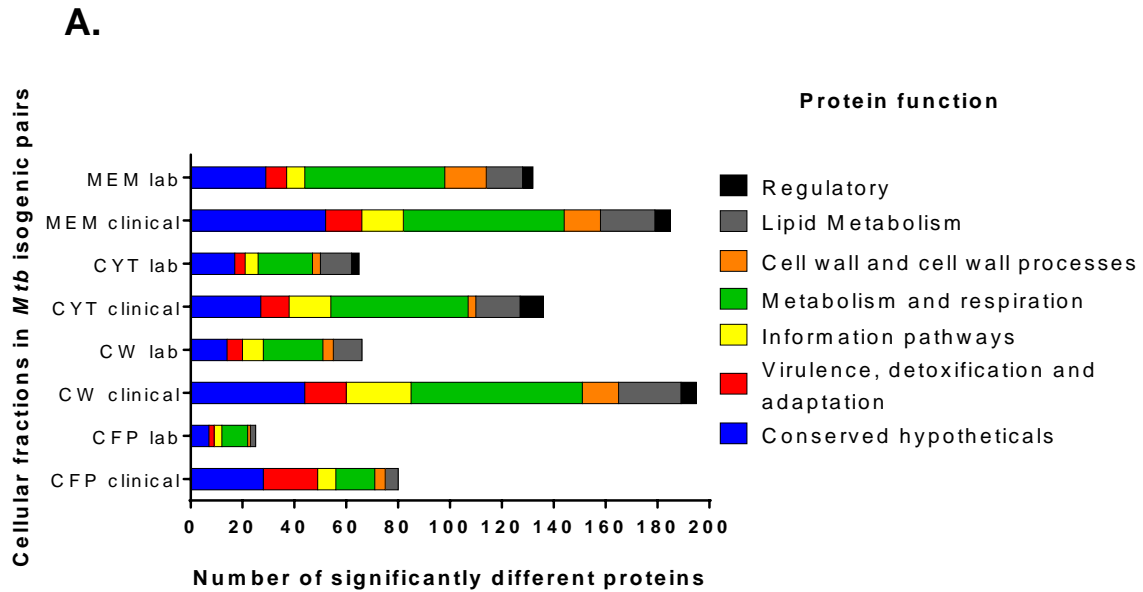
Cellular fraction	Number of proteins with significant differential abundance among pairs ($p<0.05$, t- test)		
	Clinical pair comparison	Lab pair comparison	Common proteins
CFP	80	25	4
CW	197	66	7
CYT	135	65	7
MEM	183	133	27

Figure 3.3. A. Number of proteins with significant differences in their abundance between resistance vs susceptible pairs, grouped by functional category according to Tuberculist (http://tuberculist.epfl.ch/) ($p<0.05$, t-test). B. Number of proteins with significant abundance differences among INH susceptible and resistant pairs for each cellular fraction. CFP: secreted proteins, CW: Cell wall, CYT: Cytosol, MEM: Membrane.

Table 3.1. Proteins with significant differential abundance in both clinical and lab strains after acquisition of INHr

Accession Number	Gene	Identified Proteins CFP (4)	Fold change $\frac{NSAF\ INHs}{NSAF\ INHr}$	
			Clinical pair	Lab pair
Rv0363c.1	fba	Fructose-bisphosphate aldolase*	2.3	0.9
Rv1908c.1	katG	Catalase-peroxidase-peroxynitritase T	25.1	2.7
Rv2495	pdhC	Dihydrolipoamide S-acetyltransferase E2 component*	a	0.4
Rv1910c.1	Rv1910c.1	Hypothetical exported protein	0.2	0.8
Accession Number	**Gene**	**Identified Proteins CW (7)**		
Rv1731	gabD2	Succinate-semialdehyde dehydrogenase NADPH-dependent	0.8	0.7
Rv1908c.1	katG	Catalase-peroxidase-peroxynitritase T	98.6	5
Rv2213	pepB	Aminopeptidase	0.7	0.7
Rv2214	ephD	Short-chain type dehydrogenase	0.4	0.3
Rv2346	esxO	Esat-6 like protein	1.3	1.7
Rv2567	Rv2567	Conserved alanine and leucine rich protein*	0.1	1.4
Rv2984	ppk	Polyphosphate kinase *	0.2	17.2
Accession Number	**Gene**	**Identified Proteins CYT (7)**		
Rv0757.1	phoP	Two component system transcriptional regulator	0.5	0.2
Rv1454	qor	Quinone reductase*	0.8	1.8
Rv1821.1	secA2	Preprotein translocase ATPase*	7.4	0.1
Rv1908c.1	katG	Catalase-peroxidase-peroxynitritase T	4.5	3.1
Rv2428.1	ahpC	Alkyl hydroperoxide reductase C*	2.5	0.8
Rv2703.1	sigA	RNA polymerase sigma factor*	0.5	0.5
Rv3801c.1	fadD32	Fatty-acid-coa ligase*	1.3	0.2
Accession Number	**Gene**	**Identified Proteins MEM (27)**		
Rv0134	ephF	Epoxide hydrolase	0.1	0.6
Rv0243.1	fadA2	Acetyl-coa acyltransferase*	0.5	1.3
Rv0636.1	hadB	(3R)-hydroxyacyl-ACP dehydratase catalytic subunit*	2.0	0.3
Rv0710.1	rpsQ	30S ribosomal protein S17	1.6	1.5
Rv0896.1	gltA2	Citrate synthase I*	0.8	1.2
Rv1077	cbs	Cystathionine beta-synthase	a	1.6
Rv1206	fadD6	fatty-acid-CoA ligase	0.1	0.4
Rv1326c.1	glgB	1,4-α--glucan branching enzyme	4.5	2.2
Rv1392.1	metK	S-adenosylmethionine synthetase	1.4	1.4
Rv1647	Rv1647	Adenylate cyclase (ATP pyrophosphate-lyase)	0.2	0.1
Rv1872c.1	lldD2	L-lactate dehydrogenase	0.5	0.7

Rv1908c.1	*katG*	Catalase-peroxidase-peroxynitritase T	**5.3**	**4.91**
Rv1978.1	*Rv1978.1*	Conserved hypothetical protein	**0.1**	**0.5**
Rv2037c.1	*Rv2037c.1*	Conserved membrane protein	**0.1**	**0.3**
Rv2198	*mmpS3*	Membrane protein mmpS3*	***a***	***b***
Rv2225	*panB*	3-methyl-2-oxobutanoate hydroxymethyltransferase	**9.2**	**1.7**
Rv2338c	*moeW*	Molybdopterin biosynthesis protein	**0.1**	**0.06**
Rv2357	*glyS*	glycyl-tRNA synthetase	**0.4**	**0.6**
Rv2429.1	*ahpD*	Alkyl hydroperoxide reductase D protein*	**7.9**	**0.6**
Rv2564.1	*glnQ*	Glutamine-transport ATP-binding protein ABC transporter	**0.5**	**0.6**
Rv2574.1	*Rv2574.1*	Conserved hypothetical protein	**0.1**	**0.6**
Rv2773	*dapB*	Dihydrodipicolinate reductase	**0.03**	**0.9**
Rv2951c.1	*Rv2951c.1*	Oxidoreductase	**0.4**	**0.6**
Rv2971	*Rv2971*	Oxidoreductase*	**1.9**	**0.1**
Rv2984	*ppk*	Polyphosphate kinase*	**0.3**	**1.6**
Rv3050c.1	*Rv3050c.1*	Transcriptional regulator, asnc-family	**0.1**	**0.4**
Rv3137.1	*Rv3137.1*	Monophosphatase	**9.0**	**1.6**

* Proteins with different trend between clinical and lab INHr strains. *a*: NSAF values for the INHr strain were zero. *b:* NSAF values for the INHs strain were zero.

Lab INHs	Lab INHr	Clinical INHs	Clinical INHr	Rv number	Gene	Protein name
				Rv0134§	*ephF*	**Epoxide hydrolase**
				Rv1751§	*Rv1751*	**Oxidoreductase**
				Rv2951§	*Rv2951c*	**Oxidoreductase**
				Rv1908§	*katG*	**Catalase-peroxidase-peroxynitritase T CFP**
				Rv1908§	*katG*	**Catalase-peroxidase-peroxynitritase T CW**
				Rv1908§	*katG*	**Catalase-peroxidase-peroxynitritase T CYT**
				Rv1908§	*katG*	**Catalase-peroxidase-peroxynitritase T MEM**
				Rv2214§	*ephD*	**Short-chain type Dehydrogenase**
				Rv2428§	*ahpC*	Alkyl hydroperoxide reductase C
				Rv2971§	*Rv2971*	Oxidoreductase
				Rv2429§	*ahpD*	Alkyl hydroperoxide reductaseD protein
				Rv0432**	*sodC*	Periplasmic superoxide dismutase
				Rv0475**	*hbhA*	Iron-regulated heparin binding hemagglutinin
				Rv0913c**	*Rv0913c*	Dioxygenase
				Rv1471**	*trxB1*	Thioredoxin
				Rv1855c**	*Rv1855c*	Oxidoreductase
				Rv2129c**	*Rv2129c*	Oxidoreductase
				Rv2263**	*Rv2263*	Oxidoreductase
				Rv2391**	*sirA*	Ferredoxin-dependent nitrite reductase
				Rv2454c**	*Rv2454c*	Oxidoreductase beta subunit
				Rv3790**	*Rv3790*	Oxidoreductase
				Rv0251**	*hsp*	Heat shock Protein
				Rv1860**	*apa*	Alanine and proline rich secreted protein
				Rv2416**	*eis*	Enhanced Intracellular survival protein
				Rv3418**	*groES*	10 kDa chaperonin
				Rv3913**	*trxB2*	thioredoxin reductase
				Rv1876**	*bfrA*	bacterioferritin
				Rv2238c*	*ahpE*	Peroxiredoxin
				Rv0352*	*dnaJ1*	Probable chaperone
				Rv2031*	*hspX*	Heat shock Protein
				Rv3161c*	*Rv3161c*	Dioxygenase

Figure 3.4. Proteins in the virulence and detoxification category with significant abundance differences between clonal pairs of *Mtb* after INHr acquisition. *$p < 0.05$ (t-test) in the laboratory strains comparison only, **$p < 0.05$ (t-test) in the clinical strains comparison only, §$p < 0.05$ (t-test) in both (clinical and lab) comparisons. The black box indicates the proteins that were not only shared but had the same trend in both comparisons. The cellular fraction is only indicated for KatG to emphasize the strong reduction of this protein. CFP: culture filtrate, CW: cell wall, MEM: membrane, CYT: Cytosol proteins.

katG mutant Mtb *strains had a highly active energy metabolism which could be an indirect measure of the increased intracellular redox potential.* In the next category of energy metabolism and respiration, there was an common increase in the abundance of two proteins of the ATP synthase machinery, and one of the cytochrome bc_1 complex (QcrC), while the dihydrolipoamide dehydrogenase LpdC, also involved in energy metabolism in *Mtb,* was

reduced in both INHr strains (Figure 3.5A and B). Additional proteins in this category include other members of the ATP synthase (AtpF and H) and the cytochrome bc_1 complex (QcrA and B), as well as the cytochrome aa_3-type oxidase (CtaC and D). All of these proteins that participate in the electron-transport chain in the INHr strains were increased in the INHr strains (Figure 3.5). The increase of these proteins indicates a strong redox cytosolic environment, particularly a high reductive stress in the cell with a high production of ATP.

In the same figure 3.5B, increased abundance of three enzymes of the central carbon metabolism were commonly observed for both INHr strains; these include the enzymes Icd2, GabD2 and LldD2 (Figure 3.5B). Icd2 (isocitrate dehydrogenase) is part of the Tricarboxylic acid (TCA) cycle and catalyzes the conversion of isocitrate to oxoglutarate; using $NADP^+$ as a cofactor. *Mtb* Icd2 has a higher affinity for $NADP^+$ rather than for NAD^+ (26). Also part of the TCA cycle, GabD2 is a succinate-semialdehyde dehydrogenase, with increased levels in the INHr strains and higher activity with the $NADP^+$ cofactor rather than NAD^+. This enzyme participates in the production of succinate from succinic semialdehyde, bypassing the TCA cycle enzyme α-ketoglutarate dehydrogenase (which is absent in *Mtb)* (27). On the other hand, LldD2 catalyzes the interconversion from pyruvate to lactate with the oxidation or reduction of a c-type cytochrome. The overall increased tendency of TCA cycle enzymes observed in these clinical and laboratory strains are correlated with the higher levels of energy metabolism proteins in both INHr strains. However, in chapter II other TCA enzymes were significantly reduced particularly in the INHr clinical Beijing INHr strain (Chapter II). Collectively, these findings emphasizes that more than a general metabolic trend in the carbon energy metabolism, there could be differences in particular enzymes of this pathway that participate in redox reaction using NADH and

NADPH as cofactors. Additionally, these findings suggest a possible specific trend in the TCA cycle enzymes depending on the genetic background.

A

B

Clinical INHr	Clinical INHs	Lab INHr	Lab INHs	Rv number	Gene	Protein name
				Rv1307*	*atpH*	ATP synthase delta chain
				Rv1306*	*atpF*	ATP synthase B chain
				Rv1310§	*atpD*	ATP synthase beta chain
				Rv1309§	*atpG*	ATP synthase gamma chain
				Rv3150*	*nuoF*	NADH dehydrogenase I chain G nuoF
				Rv3141*	*fadB4*	NADPH quinone oxidoreductase fadB4
				Rv2196*	*qcrB*	Ubiquinol-cytochrome C reductase QcrB (cytochrome B subunit)
				Rv1161**	*narG*	Respiratory nitrate reductase alpha chain narG
				Rv2200**	*ctaC*	Transmembrane cytochrome C oxidase (subunit II)
				Rv3043**	*ctaD*	Cytochrome C oxidase polypeptide I
				Rv3059**	*cyp136*	Cytochrome P450 136 cyp136
				Rv3147**	*nuoC*	NADH dehydrogenase I chain G nuoC
				Rv3148**	*nuoD*	NADH dehydrogenase I chain G nuoD
				Rv1279**	*gmc*	Dehydrogenase FAD flavoprotein gmc oxidoreductase
				Rv2195**	*qcrA*	Ubiquinol-cytochrome C reductase iron-sulfur subunit (cyoB)
				Rv2495c§	*bkdC*	Branched-chain keto acid dehydrogenase E2 component BkdC
				Rv3151§	*nuoG*	NADH dehydrogenase I chain G nuoG
				Rv1454c§	*qor*	Quinone reductase qor
				Rv0462§	*lpdC*	Dihydrolipoamide dehydrogenase
				Rv2194§	*qcrC*	Ubiquinol-cytochrome C reductase (cytochrome C subunit)
				Rv3010c*	*pfkA*	6-phosphofructokinase pfkA
				Rv0951*	*sucC*	Succinyl-CoA ligase [ADP-forming] β subunit
				Rv2215*	*sucB*	Pyruvate dehydrogenase E2 component sucB
				Rv1449c**	*tkt*	Transketolase tkt
				Rv3214**	*gpm2*	Phosphoglycerate mutase
				Rv1438**	*tpi*	Triosephosphate isomerase
				Rv1437**	*pgk*	Phosphoglycerate kinase
				Rv1617**	*pykA*	Pyruvate kinase-CYT
				Rv1617**	*pykA*	Pyruvate kinase-CFP
				Rv1023**	*eno*	Enolase
				Rv3318**	*sdhA*	Succinate dehydrogenase (flavoprotein subunit)
				Rv0467**	*icl1*	Isocitrate lyase
				Rv0248**	*Rv0248*	Succinate dehydrogenase (iron-sulfur subunit)
				Rv3339c**	*icd1*	Isocitrate dehydrogenase icd1
				Rv224§**	*aceE*	Pyruvate dehydrogenase E1 component
				Rv0363§	*fba*	Fructose-bisphosphate aldolase
				Rv0896§	*gltA2*	Citrate synthase I
				Rv0066§	*icd2*	Isocitrate dehydrogenase
				Rv1731§	*gabD2*	Succinate-semialdehyde dehydrogenase [NADP+]-dependent
				Rv1872§	*lldD2*	L-lactate dehydrogenase

Figure 3.5. A. Components of the electron-transport chain and ATP synthase proteins altered

in the INHr strains. MK: Menaquinone, MKH$_2$: Menaquinol. B. Proteins in the energetic metabolism and respiration category significantly different between clonal pairs of *Mtb* after INHr aquisition. p< 0.05 (t-test) *in the laboratory strains comparison only, **in the clinical strains comparison only, §in both (clinical and lab) comparisons. The black box indicates the proteins that were not only shared but had the same trend in both comparisons.

A similar reduction in the abundance of proteins involved in the MA synthesis in both INHr strains. There were three commonly reduced enzymes in the fatty acid biosynthetic pathway: UmaA, FbpA, and InhA in both INHr strains (Figure 3.6 and 3.7). UmaA is a probable methyl-transferase (MT) whose function in *Mtb* remains unknown. As a MT, UmaA can participate in the modification of mycolates, since it has a cyclopropane synthase domain according to the protein families database Pfam (http://pfam.xfam.org/family/CMAS) (28). However, the naturally occurring UmaA mutant CDC1551 did not exhibit a different mycolic acid profile compared to H37Rv by MALDI-TOF MS (29). A possible explanation for the lack of affectation in MA composition in *umaA* mutants is that although UmaA could serve as a modifying enzyme of MA, there is some level of redundancy of its function, potentially catalyzed by other protein(s) (29). The next enzyme, FbpA is part of the Antigen 85 complex which has a mycolyltransferase activity that is part of the later stages of *Mtb* cell wall assembly. Specifically, FbpA participates in the synthesis of trehalose monomycolate (TMM) and dimicolate (TDM) as well as in the transfer of mycolic acid to arabinogalactan to form cell-wall mycolyl-arabinogalactan-peptidoglycan complex (30, 31). The lower levels of FbpA have been associated with a reduction of particularly TDM without alteration of mycolic acid methyl esters (MAME, covalently bound to the cell wall) (32). Additionally, *fbpA* mutants of H37Rv strain had a reduced growth compared with the wild type strain in human and mouse macrophage cell lines, suggesting an additional role of FbpA in *Mtb* virulence (33). Finally, InhA is an NADH-dependent meromycolic acid reductase enzyme that participates in the mycolate biosynthesis and

is the most accepted target of INH, which was reduced in both INHr strains. Inhibition of the InhA protein could affect the last elongation steps for the formation of the meromycolate chain, with the possible accumulation of the *cis*-unsaturated short chain precursors (34). Additional differences of proteins in this category specific to each of the clonal pairs are presented in Figure 3.6 and 3.7 and discussed in later sections. In summary, the reduced enzymes in the synthesis of MA can result in the alteration of the MA profile including the non-covalently bound TDM and *Mtb* survival (35).

Lab INHs	Lab INHr	Clinical INHs	Clinical INHr	Rv number	Gene	Protein name
				Rv0469§	***umaA***	**Mycolic acid synthase**
				Rv1484§	***inhA***	**NADH-dependent enoyl-[acyl-carrier-protein] reductase**
				Rv1683§	***Rv1683***	**Long-chain acyl-CoA synthase**
				Rv3804c§	***fbpA***	**Secreted fibronectin-binding protein antigen fbpA**
				Rv0636§	*hadB*	(3R)-hydroxyacyl-ACP dehydratase subunit HadB
				Rv3089§	*fadD13*	fatty-acid-CoA ligase
				Rv3801c§	*fadD32*	Fatty-acid-CoA ligase
				Rv0242c**	*fabG4*	3-oxoacyl-[acyl-carrier protein] reductase
				Rv0468**	*fadB2*	3-hydroxybutyryl-CoA dehydrogenase
				Rv0470c**	*pcaA*	cyclopropane synthase
				Rv1521**	*fadD25*	fatty-acid-CoA ligase
				Rv2244**	*acpM*	meromycolate extension acyl carrier protein
				Rv2246**	*kasB*	3-oxoacyl-[acyl-carrier protein] synthase 2
				Rv2503c**	*scoB*	succinyl-CoA:3-ketoacid-coenzyme A transferase beta subunit
				Rv2504c**	*scoA*	succinyl-CoA:3-ketoacid-coenzyme A transferase alpha subunit
				Rv2524c**	*fas*	fatty-acid synthase fas
				Rv2930**	*fadD26*	fatty-acid-CoA ligase
				Rv2935**	*ppsE*	phenolpthiocerol synthesis type-I polyketide synthase
				Rv2940c**	*mas*	multifunctional mycocerosic acid synthase membrane-associated
				Rv2959c**	*Rv2959c*	methyltransferase/methylase
				Rv3265c **	*wbbL1*	a-3-L-rhamnosyl transferase
				Rv3280**	*accD5*	propionyl-CoA carboxylase beta chain 5
				Rv3285**	*accA3*	bifunctional acetyl-/propionyl-coenzyme A carboxylase alpha chain
				Rv3522**	*ltp4*	lipid transfer protein or keto acyl-CoA thiolase
				Rv3667**	*acs*	acetyl-CoA synthetase acs
				Rv0635**	*hadA*	(3R)-hydroxyacyl-ACP dehydratase subunit HadA
				Rv0637**	*hadC*	(3R)-hydroxyacyl-ACP dehydratase subunit HadC
				Rv0241c**	*htdX*	Probable 3-hydroxyacyl-thioester dehydratase HtdX
				Rv0503c*	*cmaA2*	Cyclopropane-fatty-acyl-phospholipid synthase 2
				Rv0644c*	*mmaA2*	methoxy mycolic acid synthase 2
				Rv0645c*	*mmaA1*	methoxy mycolic acid synthase 1
				Rv0855*	*far*	fatty-acid-CoA racemase far
				Rv1181*	*pks4*	polyketide beta-ketoacyl synthase
				Rv1323*	*fadA4*	acetyl-CoA acetyltransferase
				Rv1483*	*fabG1*	3-oxoacyl-[acyl-carrier protein] reductase
				Rv1543*	*Rv1543*	fatty-acyl-CoA reductase
				Rv2006*	*otsB1*	trehalose-6-phosphate phosphatase otsB1
				Rv2941*	*fadD28*	fatty-acid-CoA ligase
				Rv3392c*	*cmaA1*	Cyclopropane-fatty-acyl-phospholipid synthase 1
				Rv3720*	*Rv3720*	fatty-acid synthase
				Rv3800c*	*pks13*	polyketide synthase
				Rv3825c*	*pks2*	polyketide synthase

Scale: 1.5 — 0.05 — -1.4

Figure 3.6. Proteins involved in the lipid biosynthetic pathway with significant differences between clonal pairs of *Mtb* after INHr acquisition. p< 0.05 (t-test) *in the laboratory strains comparison only, **in the clinical strains comparison only, or §in both (clinical and lab)

comparisons. The black box indicates the proteins that were not only shared but had the same trend in both comparisons.

Reduction of proteins of the SAM metabolism correlates with the alteration in the levels of multiple MT in INHr Mtb strains. In line with the findings of altered MTs modifying enzymes in the MA biosynthesis (Figure 3.6 and 3.7), levels of proteins MetK and SahH, involved in the S-adenosylmethionine (SAM) metabolism, were significantly reduced in the INHr strains. SAM is an intermediate that participates in the formation of fatty acid modifications such as the introduction of *trans* methyl branches and cyclopropane rings in the meromycolic acid chain of mycolic acids (36). MetK participates in the final conversion of methionine to SAM and was reduced in both INHr strains (Table 3.1 and Figure 3.8). Methionine can be obtained by excision from the N-terminus of a peptide by the enzyme MapB, which was also reduced in the clinical INHr strain (Figure 3.8B). SahA was also reduced in both INHr strains; however, the differences were present in different cellular fractions (Figure 3.8B). The trend of these SAM metabolism proteins indicate a possible reduced synthesis of SAM in both strains which probably is related with the variation of the MT of the FAS II. The reduction of these enzymes suggests a diminished availability of SAM (Figure 3.8A), therefore, MT with higher affinity for SAM are needed.
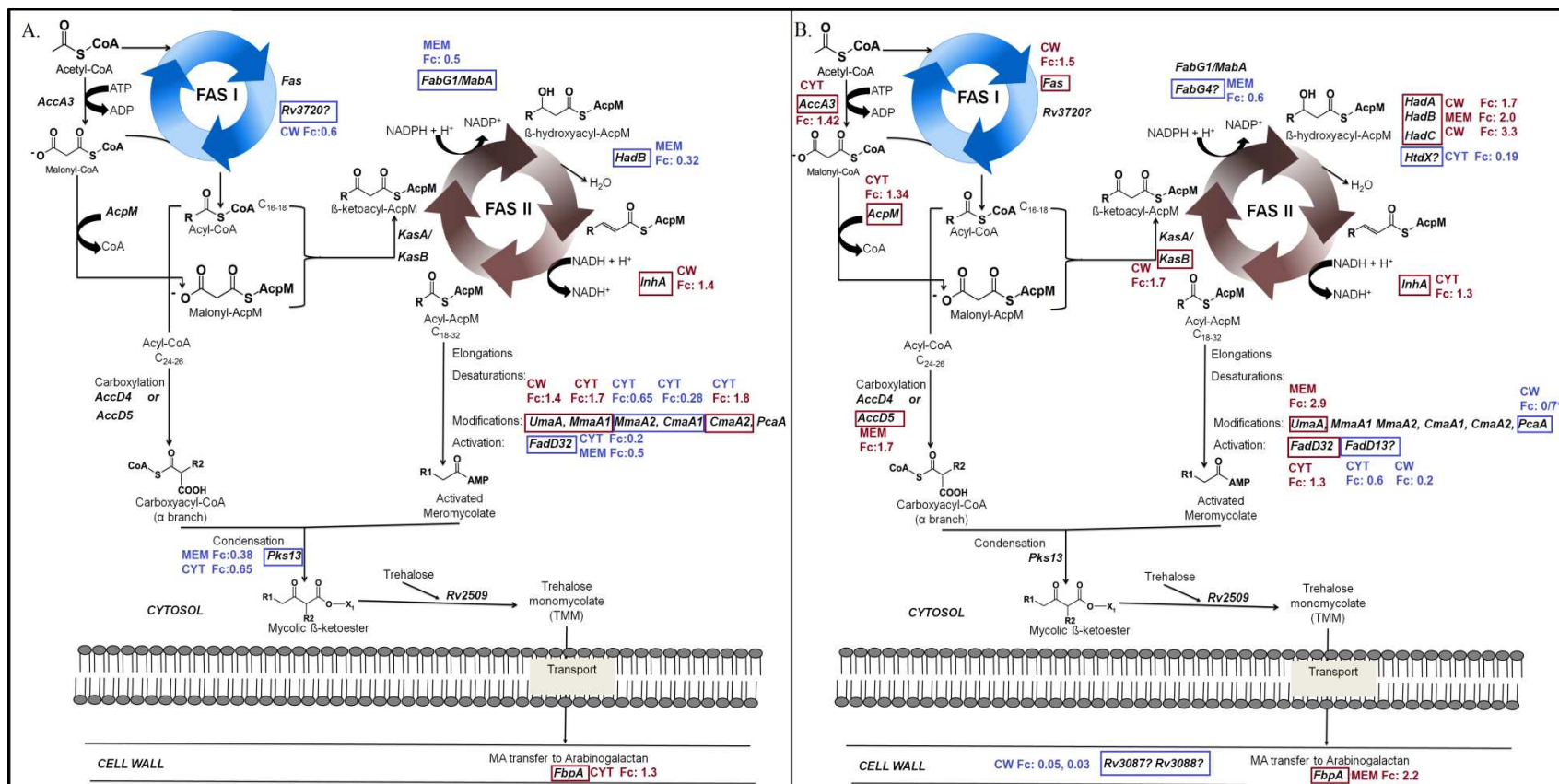
Figure 3.7. Altered enzymes in the fatty acid synthesis I and II pathway of *M. tuberculosis* (p<0.05). A. Lab strains comparison. B. Clinical pair comparison. Blue and red squares represent higher or lower levels of enzymes in the INHr strain respectively. Fold change (Fc)= NSAF INHs/INHr. *PcaA NSAF values for the fold change (numerator and denominator) were multiplied by 1000 to facilitate the comparison. Alternative enzymes marked with quotation marks (?) for the clinical pair comparison are added; see main text to find more details. Adapted from Marrakchi, et al., 2014 (37) with permission from of Elsevier.

A.

O-acetyl-L-homoserine + Hydrogen sulfide

O-acetyl-L-homoserine + Methanethiol

MetC

MetC

Succinil-CoA + H$^+$ → Acetate

Acetate

Succinyl L-homoserine + L-cysteine

MetB

→ Cystationine → Homocysteine → Methionine

MetE

ATP + H$_2$O

MetK

PP + P$_i$

SahH → Adenine

MT

S-adenosyl-homocysteine (SAH)

S-adenosyl-L-methionine (SAM)

Methylated biomolecules

B.

Lab INHs	Lab INHr	Clinical INHs	Clinical INHr	Rv number	Gene	Protein name
				Rv1392§	*metK*	**S-adenosylmethionine synthetase metK**
				Rv3248c**	*sahH*	**adenosylhomocysteinase sahH-CFP**
				Rv3248c*	*sahH*	**adenosylhomocysteinase sahH-CYT**
				Rv2861c**	*mapB*	Methionine aminopeptidase mapB (286 aa)
				Rv1079**	*metB*	Cystathionine gamma-synthase metB (389 aa)
				Rv1405**	*Rv1405*	methyltransferase
				Rv3340**	*metC*	O-acetylhomoserine sulfhydrylase metC
				Rv1133c**	*metE*	5-methyltetrahydropteroyltriglutamate-homocysteine methyltransferase metE-CYT
				Rv1133c*	*metE*	5-methyltetrahydropteroyltriglutamate-homocysteine methyltransferase metE-CFP

Scale: 1.4, -0.05, -1.5

Figure 3.8. A. Altered enzymes in the methionine and SAM metabolism (blue and red squares indicate increased and decreased levels in INHr strain respectively), dashed purple square indicates a different trend between clinical and lab INHr (see details in heatmap), MT: methyltransferase. B. Heatmap showing the abundance of the protein in each clonal pair. p< 0.05 *in the laboratory strains comparison only, ** in the clinical strains comparison only and § in both clinical and lab comparisons.

Increased abundance of proteins in the fatty acid β-oxidation in katG mutants. Proteins affected that participated in the ß-oxidation of fatty acids and lipolysis are described in figure 3.8. The proteins FadD6, FadE24 and FadB are similarly increased in both INHr strains. FadD6 is a fatty acid CoA ligase that catalyzes the first reaction and FadE24is an acyl CoA dehydrogenase that catalyzes the second reaction of the ß-oxidation. FadB forms a canonical ß-oxidation complex with FadA (38) which is also increased in the clinical INHr strain (Figure 3.8). So far, five *fadB* genes have been identified in *Mtb* with no very specific functions assigned, but probably with different chain specificity length for their substrates (39). Overall

116

there was a net increase of the majority of these enzymes that participate in different reactions in the ß oxidation pathway in both INHr strain (Figure 3.9 and 3.10). It is important to emphasize that ß oxidation of fatty acids is a metabolic route with a great redundancy (38), as it was observed in this analysis, many protein representatives can participate in one single reaction of the metabolic pathway (Figure 3.10). The increase of the fatty acid oxidation enzymes suggests a switch of INHr strains to use lipids as the main source of energy, with the generation of substantial quantities of NADPH (40). The higher β-oxidation activity in *Mtb* is also related with an altered energetic metabolism that was also a commonly altered metabolic pathway. For instance, the β-oxidation of palmitate generates 106 ATP molecules, while the oxidation of glucose yields only 38 ATP molecules, demonstrating a relationship with the also increased enzymes in the ATP synthesis (40). The increase of β-oxidation is also associated with the reduced levels of the pyruvate kinase (PykA) in the INHr strains (Figure 3.5B) that is associated with the use of non-glycolytic carbon sources such as lipids (41).

Although this category only has three proteins in common, the overall trend in clinical and laboratory INHr strains was consistently higher compared to their respective INHs pairs (Figures 3.9 and 3.10). This category also includes lipolytic enzymes such as lipases/esterases, phospholipases, and enzymes involved in cholesterol oxidation which were mostly increased in the INHr strains. Among the lipases, higher levels of LipD and LipW were observed in the clinical INHr strain while only higher levels of LipN were detected in the laboratory INHr strain (Figure 3.9). Lip family genes comprise 24 genes, some of them with the function of release fatty acids-intermediates for the glyoxylate shunt, while others have an assigned role in virulence (42). LipN was recently demonstrated to have a role in the degradation of exogenous lipids (43).

LipD on the other hand, is not normally expressed in H37Rv strain but is induced under oxidative stress and is active under acidic and high temperature conditions (44).

Lab INHs	Lab INHr	Clinical INHs	Clinical INHr	Rv number	Gene	Protein name
				Rv1206 §	*fadD6*	**fatty-acid-CoA ligase**
				Rv3139 §	*fadE24*	**acyl-CoA dehydrogenase**
				Rv0860 §	*fadB*	**fatty-acid oxidation protein**
				Rv0243 §	*fadA2*	acetyl-CoA acyltransferase
				Rv2301 §	*cut2*	cutinase
				Rv3546 §	*fadA5*	acetyl-CoA acetyltransferase in cholesterol degradation
				Rv0131c**	*fadE1*	acyl-CoA dehydrogenase
				Rv0217c **	*lipW*	esterase
				Rv0231**	*fadE4*	acyl-CoA dehydrogenase
				Rv0270 **	*fadD2*	fatty-acid-CoA ligase
				Rv0271c **	*fadE6*	acyl-CoA dehydrogenase
				Rv0400c **	*fadE7*	acyl-CoA dehydrogenase
				Rv0859 **	*fadA*	acyl-CoA thiolase
				Rv0873 **	*fadE10*	acyl-CoA dehydrogenase
				Rv0905 **	*echA6*	enoyl-CoA hydratase
				Rv1070c **	*echA8*	enoyl-CoA hydratase
				Rv1071c **	*echA9*	enoyl-CoA hydratase
				Rv1923**	*lipD*	Lipase
				Rv2349c**	*plcC*	phospholipase C 3
				Rv2350c**	*plcB*	Membrane-associated phospholipase C 2
				Rv2590 **	*fadD9*	fatty-acid-CoA ligase
				Rv2831 **	*echA16*	enoyl-CoA hydratase
				Rv3409c**	*choD*	cholesterol oxidase precursor choD
				Rv3797 **	*fadE35*	acyl-CoA dehydrogenase
				Rv1427c**	*fadD12*	long-chain fatty-acid-CoA ligase
				Rv1106c**	*Rv1106c*	Cholesterol dehydrogenase
				Rv0672*	*fadE8*	acyl-CoA dehydrogenase
				Rv2500c*	*fadE19*	acyl-CoA dehydrogenase
				Rv2679*	*echA15*	enoyl-CoA hydratase
				Rv2970c*	*lipN*	lipase
				Rv3140*	*fadE23*	acyl-CoA dehydrogenase
				Rv1193*	*fadD36*	fatty-acid-CoA ligase
				Rv2187*	*fadD15*	long-chain fatty-acid-CoA ligase

Scale: 1.5 — 0.05 — -1.4

Figure 3.9. Proteins in the β-Oxidation and lipid hydrolysis with significant differences between clonal pairs of *Mtb* after acquisition of INH resistance. $p < 0.05$ (t-test) *in the laboratory strains comparison only, **in the clinical strains comparison only, or §in both (clinical and lab) comparisons. The black box indicates the proteins that were not only shared but had the same trend in both comparisons.

Higher levels of transcriptional regulatory proteins in both INHr strains suggest an alteration in the metabolism of complex lipids, among other functions. PhoP is a transcriptional regulatory protein that works together with PhoQ to positively regulate some genes implied in the growth of fully virulent *Mtb* strains in the mouse model (45). In both INHr strains significantly higher abundance of PhoP was detected respective to their susceptible pairs. A relationship between PhoP and the metabolism of complex lipids in *Mtb* has been shown;

particularly an increased expression of this protein was previously associated with higher levels of sulfolipids. Jackson et al., suggested that PhoP regulates the expression of genes involved in the polyketide synthesis (46). This was true for the laboratory pair comparison as Pks2 and Pks4 were significantly increased in the INHr strain (Figure 3.6) simultaneous to an increase of the PhoP levels. Although PhoP was also increased in the clinical INHr strains, there were not significantly increased levels of proteins in the polyketide synthesis for this INHr strain. It should be noted that the fold change showed as stronger PhoP increase in the laboratory pair (Table 3.2).

Another regulatory protein with increased abundance in both INHr strains was Rv3050c.1. Little is known about this particular enzyme that is part of the AsnC-family which groups those regulators that are specifically triggered by asparagine binding. Among the possible events that members these proteins regulate are amino acid metabolism, DNA repair, central carbon metabolism and persistence (47). Particular differences for each pair comparison are further described in table 3.2.

Figure 3.10. Enzymes involved in the ß-oxidation of fatty acids. For each reaction, the name of the enzyme is described in black, close to them are the different isozymes that were found in the analysis significantly higher (blue) or lower (red) in the INHr strain for each comparison.
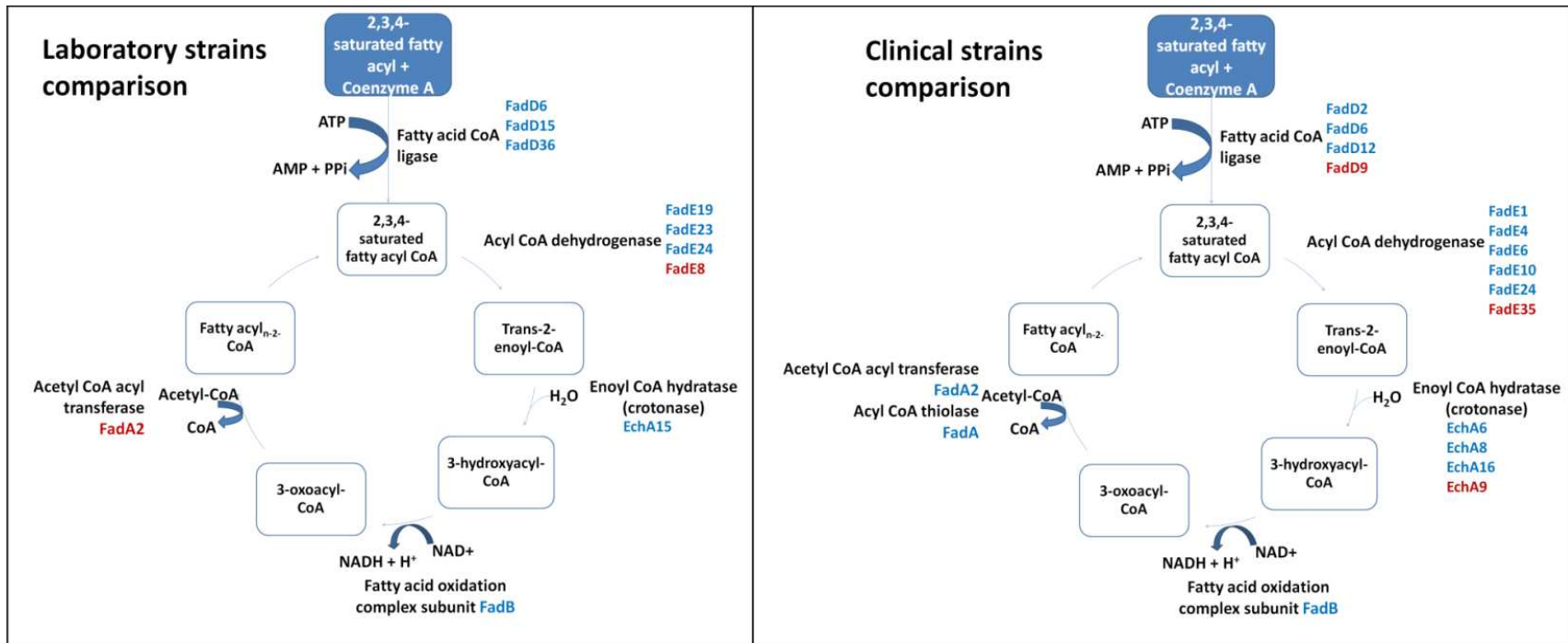
Table 3.2. Regulatory proteins altered in both clinical pair comparisons*

Rv number	Gene	Protein name	Fold change clinical INHs/INHr	Fold change laboratory INHs/INHr
Rv0757.1	*phoP*	Two component system transcriptional regulator phoP	**0.5**	**0.2**
Rv3050c.1	*Rv3050c.1*	Transcriptional regulator, asnC-family	**0.1**	**0.4**
Rv0042c.1	*Rv0042c.1*	Transcriptional regulator, marR-family	**0.2**	NS
Rv0043c.1	*Rv0043c.1*	Transcriptional regulator, gntR-family	***a***	NS
Rv0474.1	*Rv0474.0*	Transcriptional regulator	***b***	NS
Rv0844c.1	*narL*	Nitrate/nitrite response transcriptional regulator narL	***b***	NS
Rv1423.1	*whiA*	Transcriptional regulator whiA	**0.5**	NS
Rv1479.1	*moxR1*	Transcriptional regulator moxR1	**0.3**	NS
Rv1626.1	*Rv1626.1*	Two component system transcriptional regulator	**0.8**	NS
Rv2919c.1	*glnB*	Nitrogen regulatory protein P-II glnB	**0.3**	NS
Rv3143.1	*Rv3143.1*	Response regulator	**0.2**	NS
Rv3291c.1	*Rv3291c.1*	Transcriptional regulator, asnC-family	**0.7**	NS
Rv3295.1	*Rv3295.1*	Transcriptional regulator, tetR-family	**0.7**	NS
Rv3676.1	*Rv3676.1*	Transcriptional regulator, crp/fnr-family	**0.5**	NS
Rv3692.1	*moxR2*	Methanol dehydrogenase transcriptional regulator moxR2	***b***	NS
Rv0576.1	*Rv0576*	Transcriptional regulator, arsR-family	NS	**0.4**
Rv1267c.1	*embR*	Transcriptional regulator embR	NS	**0.2**
Rv1379.1	*pyrR*	Pyrimidine operon regulatory protein pyrR	NS	**0.2**
Rv1719.1	*Rv1719.1*	Transcriptional regulator	NS	**0.1**

*Fold change values provided only for significant differences, t-test, p ≤0.05). *a:* NSAF values in the INHr strain were zero, *b:* NSAF values in the INHs strain were zero. NS: not statistically significant.

Increased levels of the lipoproteins in the INHr strains. Even though there was not a single lipoprotein in common in both clinical and laboratory comparisons, there was a common trend in the abundance of proteins in this category. Nine conserved lipoproteins were increased in the INHr strains, while the protein LppX was significantly decreased in the clinical INHr

strain (Figure 3.11). LppX is a lipoprotein involved in the translocation of phthiocerol dimycocerosates (DIMs) across the plasma membrane of *Mtb (48)*. This finding is in line with reduced levels of PpsE (phenolpthiocerol synthesis type-I polyketide synthase) found in the clinical INHr strain (Figure 3.6). Lower levels of LppX also suggest a possible reduced virulence of the clinical INHr strain (48).

Lab INHs	Lab INHr	Clinical INHs	Clinical INHr	Rv number	Gene	Protein name
				Rv2518c*	*lppS*	Lipoprotein lppS
				Rv1244*	*lpqZ*	Lipoprotein lpqZ
				Rv1270c*	*lprA*	Lipoprotein lprA
				Rv1275*	*lprC*	Lipoprotein lprC
				Rv1368**	*lprF*	Conserved lipoprotein
				Rv1899*	*lppD*	Possible lipoprotein
				Rv2585c**	*Rv2585*	Conserved lipoprotein
				Rv2945**	*lppX*	Lipoprotein lppX
				Rv3007**	*lppZ*	Conserved lipoprotein
				Rv3623*	*lpqG*	Conserved lipoprotein

Figure 3.11. Lipoproteins with significant differences between clonal pairs of *Mtb* after INHr acquisition. p< 0.05 (t-test) *in the laboratory strains comparison only, **in the clinical strains comparison only.

3.3.2 Protein differences between clinical and laboratory strains without similar trends

Increased abundance of the alkyl peroxidases are not common events in all katG mutants. In the analysis of proteins that compensate the reduction of KatG, the alkyl peroxidases AhpC and AhpD were evaluated at both INHr strains. As shown in table 3.1 and Figure 3.4, these enzymes were lower in the clinical INHr but higher in the laboratory INHr strain compared to their susceptible pairs respectively (Table 3.1, Figure 3.4). AhpC has been highly recognized as an important compensatory mechanism in INHr strains with *katG* mutations to combat the redox stress. KatG and AhpC share organic peroxides and reactive nitrogen intermediates as substrates (49, 50). Previously, Shelman *et al.,* demonstrated that *katG* mutant clinical isolates had a wide range of increased levels of the AhpC compared with the reference strain H37Rv (50). This was

the trend for AhpC observed in the INHr laboratory strain compared to its clonal pair in the present study (Figure 3.4). Additionally, in the same study by Shelman *et al.*, some laboratory *katG* mutants did not have increased levels of AhpC. They attributed this result to the fact that the non-altered AhpC strains were derived in the laboratory and were not previously exposed to a second selection event, such as being exposed to the host immune response. The results obtained in this dissertation contradict Sherman's argument because the clinical INHr strain had even significantly reduced levels of AhpC and AhpD compared to its INHs clonal pair (Figure 3.4). This finding suggests that general statements regarding INHr *katG* mutants and their phenotype should be always validated in the context of the same bacterial genetic background and emphasizes the importance to compare the protein levels to a closely related parental strain (instead to a reference strain), reducing in this way possible confounding factors in the analysis.

NADH dehydrogenase complex I. Previously, it was proposed in this chapter that many of the differences in the proteome variation of the INHr strains could be related with the different metabolism of the NADH or NADPH. Interestingly, proteins the NADH dehydrogenase complex-NDH I were altered in both INHr strains but with a different trend. The protein NuoF was significantly reduced in the laboratory INHr while the proteins NuoC, D, and G were higher in the clinical INHr strain (Figure 3.5). This suggest that NDH-1 complex alteration could be a common result of the INHr phenotype but the trend to which it could be affected may depend on additional intrinsic and external factors of the bacteria such as its genetic background and the specific time to which the Mtb strain was exposed to the drug as well if it was a single or multiple external cues (such as a combined therapy).

3.3.3 Protein abundance differences specific to the laboratory INHr strain

Increased levels of FtsZ in the laboratory INHr strain. Remarkably, the essential protein

for cell division FtsZ (acronym for Filamentous temperature-sensitive protein Z) was

significantly higher in the cytosol of the laboratory INHr only and confirmed by western blot

(Appendix I Table 5, Figure 3.13). Because of the importance of this protein in cell division and

its exploration as an attractive drug target against *Mtb*, it is important to further explore the

implications of higher levels in the laboratory INHr strain (51). In other bacteria models such as

Escherichia coli, a seven-fold increase in the levels of FtsZ is associated with a the formation of

minicells at the cell poles (52). This protein seems to control the frequency of polar division but

not the medial division in this bacteria model that has another set of proteins identified in this

process such as FtsA, ZapA, ZipA and Min proteins (51-54). *Mtb*, on the other hand, does not

have completely annotated these other proteins associated with division and may lack of some of

those. The cell division process itself in *Mtb* is very different from model Gram positive and

Gram negative bacteria (51), not only because *Mtb* grow slowly but also asymmetrically (55). In

our proteomics comparison, FtsZ was 9.3 fold higher in the laboratory INHr compared to its

INHs pair, looking at the NSAF values (Appendix I Table 5). Although higher levels of FtsZ are

probably associated with the INHr acquisition event, this effect is not universal of all INHr *Mtb*

strains, since this protein levels were not increased in the other INHr strain that overall has a

more variable proteome (Figure 3.3B).

Altered metabolism of complex lipids in the laboratory INHr strain. In the laboratory

INHr strain, there were enzymes in the phthiocerol, phenolphthiocerol and polyketide

biosynthesis with increased levels after acquisition of INHr. The first one, PpsA is an NADPH

dependent enzyme that participates in the synthesis of the phthiocerol and phenolphthiocerol.

This enzyme was increased in the membrane fraction of the laboratory INHr strain. Similarly, the NADPH-dependent phthioceranic/hydroxyphthioceranic acid synthase Pks2 and the probable polyketide beta-ketoacyl synthase Pks4 were higher in the lab INHr strain (Figure 3.6). Pks2 catalyzes the synthesis of phthioceranic acids, whose hydroxylated version are major components of the mycobacterial sulfolipids (56). The specific function of Pks4 is not known yet. Most of the products of *pks* genes are associated with the synthesis of complex *Mtb* lipids, such as sulfolipids, mycocerosic acid, di- and tri-acylated trehaloses, poly-acyltrehaloses and phthiocerol dimycocerosates.

Other enzymes in the fatty acid synthesis increased in the laboratory INHr strains. In section 3.3.1, a similar reduction of three FAS II-related- proteins were identified in both INHr strains, one of them the recognized INH target: InhA. However, there were also variation core enzymes of this pathway and the later MA modifying enzymes that were unique for each clonal pair comparison. In the laboratory comparison, the NADPH dependent protein FabG1 and the catalytic subunit of the (3R)-hydroxyacyl-ACP dehydratase complex, HadB, were significantly increased in the INHr strain (Figures 3.6 and 3.7).   These are core enzymes of the FAS II route for the elongation of MA. Although it was only significant for the laboratory pair, FabG1 was also increased in the clinical INHr strain. In *Mtb* as well as in *M. bovis* BCG, *fabG1* and *inhA* are not only closely located but part of the same transcription unit, which means that they are controlled by the same promoter (57). The different trend observed between FabG1 and InhA in the laboratory INHr strain suggests that there is a strong post-translational event affecting the NADH dependent INH enzyme (InhA) that is also contributing to the INHr phenotype observed.

Additionally, the laboratory INHr strain had the highest number of altered enzymes that participate in the MA modification step. These reactions are catalyzed by a group of MT that use

125

SAM in the methyl transfer reactions (32). In the laboratory INHr strain, an increased ratio of MmaA2/MmaA1 and CmaA1/CmaA2 was observed (Figure 3.6 and 3.7). As Tanaka *et al.,* and Marrakchi *et al*., already described, a difference in this ratio may not affect the total amount of MA but causes a possible increase in the *cis/trans* unsaturation and cyclopropanation ratio in the oxygenated MA as a result of the reduced levels of the MmaA1 and CmaA2. This is because both MmaA1 and CmaA2 are required for the introduction of *trans*-cyclopropane ring and the conversion of *cis* to *trans* unsaturation in oxygenated meroacids (37, 58). Also looking at MmaA2 levels, it could be possible that the number of α-MA with distal cyclopropane rings will be higher than those with *cis* unsaturation (37, 58). The increase in the abundance of CmaA1in the laboratory INHr strain could be related with the increase of the bacterial need to resist the H_2O_2 stress and adapt to the lack of KatG, as it was previously demonstrated by Yuan et al (59).

In the activation and condensation reactions following the core enzymes of FAS II, the fatty-acid-AMP ligase FadD32 and the polyketide synthase-Pks13 are both increased in the laboratory INHr strain. FadD32 converts the meroacyl-S-ACP derived from the FAS-II system to meroacyl-AMP by using an ATP molecule. After this, Pks13 condenses the activated FAS II product with the CoA-derivative from FAS I (that was previously further carboxylated by AccD4 and AccD5) in a Cleinsen-type reaction (58). Both FadD32 and Pks13 are in the same transcriptional unit. Looking at all the FAS II enzymes altered in the laboratory INHr strain, only those in common with the clinical INHr strain (InhA, UmaA, and FbpA) had reduced levels. Consequently, only the reduction of the previous three proteins of FAS II can be closely related with the INHr phenotype. After this event, the laboratory INHr strain had a very specific strategy to keep this important metabolic pathway functional for the bacteria; resulting in the increase of the majority of the other enzymes of FAS II route (Figure 3.6 and 3.7).

3.3.4 Protein differences specific to the clinical INHr strain

Highly reduced levels of the FAS II enzymes reveal an alternative or intermediate fatty-acid synthetic pathway in the clinical INHr strain. In this dissertation, members of the core reactions of FAS II route acyl-carrier protein (ACP) AcpM, 3-oxoacyl-ACP synthase 2 KasB and the (3R)-hydroxyacyl-ACP dehydratase HadABC complex, were all reduced in the clinical strain. Additionally, the fatty-acid-AMP ligase FadD32 and the carboxylase AccD5 were also reduced in this INHr *Mtb* strain (Figure 3.6 and 3.7). The reduction of enzymes in this vital metabolic pathway seems to be compensated by the increase of alternative enzymes that catalyzes similar reactions. An alternative fatty acid biosynthetic pathway, called the "fatty acid elongation system II," was first observed in *M. smegmatis* in 1983, that was (60).The proposed alternative enzymes that can participate in this alternative pathway are FabG4, HtdX, FadD13 and other members of the *mymA* operon, that were increased in the INHr stain (Figure 3.6 and 3.7). These enzymes can participate in the same reactions but have a different affinity for the NADH and NADPH cofactors and other substrates.

The initial reduction of the AcpM in the clinical INHr strain suggests a slow entrance of the ACP derivatives into the FAS II pathway. AcpM reduced levels were in line with results obtained in the previous proteomics approach (Chapter II) and were confirmed by western blot in the culture filtrate fraction (chapter 2, Figure 2.8) and cytosol (Figure 3.13). This reduction of AcpM derivatives may induce the increased levels of FabG4. FabG4 (probable 3-oxoacyl-acyl-carrier protein reductase) share the FabG1function of converting oxo (=O) to hydroxyl (-OH) functional groups in the mycolate biosynthesis FabG4 has not only an affinity for both NADH and NADPH cofactors, but also for both CoA and ACP fatty acid derivatives; this is not seen for FabG1 which has a higher affinity for NADPH and ACP derivatives (61, 62). In fact, FabG4 has

a higher affinity for NADH and is more efficient in the presence this cofactor (61). The increase

of FabG4 also demonstrate that the variation of enzymes of MA biosynthetic pathway is

associated again with a differently affected NADH and NADPH metabolism in clinical versus

laboratory INHr strains.

FabG4 is in the same operon of the 3-hydroxyacyl-thioester dehydratase HtdX, as shown

in figure 1.5 of chapter I (61, 62). However the specific role of both enzymes in FAS II remains

unclear (61). A potential catalytic role of the HtdX in the FAS II route has been proposed in the

dehydration step of the hydroxyacyl-ester to an enoyl intermediate that is catalyzed by the

HadABC complex (Figure 3.7) (63). HtdX was proposed to have a small contribution in the FAS

II reactions since this enzyme is strictly specific for CoA substrates rather than ACP derivatives

(64). Since the entire HadABC complex was significantly lower in the clinical INHr, it is

possible that under the absence of ACP substrates (because of the reduced AcpM levels); more

CoA derivatives are used in the synthesis of MA alternatively with the mediation of FabG4 and

HtdX which are highly abundant in this INHr strain. It should be noted that FabG4 and HtdX

were also increased clinical Beijing INHr (Table 2.2 from chapter 2), albeit without detectable

altered levels of AcpM.

Finally, the fatty-acid-AMP ligase FadD13 (which also annotated to be part of the β-

oxidation of fatty acids) is significantly increased in the clinical but decreased in the laboratory

INHr strain (Figure 3.6 and 3.7). FadD13 was proposed to participate in the biosynthetic

pathway of mycolic acids by the activation of fatty acids. This enzyme has a high affinity for

fatty acids with 24 and 26 carbons (C24 and C26) (65). This protein is part of the *mymA* operon

where the acyl transferases Rv3087 and Rv3088 that are significantly increased in the clinical

INHr strain also belong (Figure 3.6 and 3.7). The last two enzymes are proposed to transfer

FadD13 activated fatty acids to an acceptor in the mycobacterial cell wall (66). The *mymA* operon is upregulated under acidic growth conditions simultaneous to a downregulation of the FAS-II pathway (67). The accumulation of the C24 and C26 fatty acids are the perfect substrate for the enzymes of the *mymA* operon to possibly replenish the synthesis of the very vital MA for *Mtb* (66). Again, this suggests the increase of some alternative enzymes for the synthesis of mycolic acids when the canonical FAS II enzymes are reduced due to the unbalanced redox intracellular space.

Only one modifying MA enzyme was increased in the INHr, the cyclopropane synthase PcaA (Figure 3.6). This enzyme participates in the modification of MA by introducing a proximal cyclopropane ring in α-MA and is also considered a virulence factor (68).

Reduction of FAS I in the clinical INHr strain. Interestingly, not only FAS II but FAS I was also affected in the clinical pair comparison. Fas, the eukaryotic-like FAS I enzyme was significantly lower only in the CW fraction and the carboxylase AccA3 was reduced in the cytosol fraction of clinical INHr strain. The carboxylase AccA3 activates the substrates that are going to be adding carbons to the growing FAS I molecule. In contrast, the probable fatty acid synthase Rv3720 was increased in the INHr strain (Figure 3.6 and 3.7). Rv3720 is known as a cyclopropane-fatty-acyl-phospholipid synthase (69) and is an NADPH-dependent enzyme that could be working in parallel with the reduced levels of Fas enzyme (figure 3.6 and 3.7).

The role of cutinases in the enhanced in vitro growth of the clinical INHr in a media containing Tween 80. In the group of altered lipases/esterases described in figure 3.9, there is a significant increase of the enzyme cutinase 2 (Cut2, Culp2 or Cfp25) in the clinical INHr strain. *Mtb* has seven putative cutinases which are esterases with the ability to hydrolyze phospholipids and Tween 80 as was discussed in chapter II and reported by Parker, *et al* (70). Cutinases have

been experimentally associated with phospholipases which were also increased in the clinical

INHr strain (PlcC, PlcB, Figure 3.9) (70). Both phospholipases are PlcC and B are in the same

transcription unit and are virulence factors involved in the bacterial pathogenesis which are

particularly relevant in the late infectious stage in mice (71, 72). The role of these enzymes is to

hydrolyze the membrane phospholipids of the host cells (71). PlcC is differentiated from other

phospholipases as is the most active phospholipase under acidic conditions (71). Culp2 as well

as Culp1 that was found increased in the CFP of the clinical INHr strain (in chapter II) are

secreted proteins with the great potential to favor the growth of this strain under oleic derivatives

as the carbon source. The increase of these different lipases and cutinases offer more supporting

reasons to explain the higher growth of INHr observed only in Tween-containing media, in a

strain that has growth inhibition in the minimal GAS media (Chapter II, Figure 2.9).

Increased cholesterol metabolism-related enzymes in the clinical INHr strain also

suggest an altered host-pathogen interaction. Interestingly, enzymes involved in cholesterol

metabolism in *Mtb*, FadA5 (Rv3546, steroid 3-ketoacyl-CoA thiolase), Rv1106 ((3) β-

hydroxysteroid dehydrogenase) and ChoD (Rv3409, cholesterol oxidase) were higher in the

clinical INHr strain in the cell wall and membrane fractions (Fig 3.9). FadA5 is required for

utilization of cholesterol as a sole carbon source *in vitro* and for virulence at chronic stages

during *in vivo* infection in the mouse model (73, 74). In the same line, Rv1106 participates in the

first step of the cholesterol degradation pathway, using NAD^+ as a cofactor (75). ChoD is an

oxidoreductase that degrades cholesterol generating hydrogen peroxide (76). The increased

levels of this enzyme together with the reduced levels of the major catalase-peroxidase enzyme

in *Mtb* (KatG) could be detrimental for the bacterial redox balance and could affect bacterial

growth during host infection. Common to other proteins described in lipid hydrolysis in this

section, ChoD is considered a virulence factor because of its ability to modulate the macrophage responses, including its own intracellular growth and stimulating Interleukin-10 production via Toll like receptor (TLR) 2 (77).

Finally, this proteomics analysis confirms 56% of the proteins out of the 99 identified in the T pair comparison of chapter 2. These proteins have the same trend in both studies and are described in Appendix I Table 9. This table also includes six proteins (Psts1, EchA6, GpsI, Rv0283, Rv2054 and Rv3716) that had the same trend in both proteomic analysis but in the current one the difference had p values between 0.05 and 0.08

3.3.5 Western Blot validation results and evaluation of the peroxidase activity

A confirmation of the KatG levels and the peroxidase activity in the soluble fractions (cytosol and CFP) of laboratory and clinical pair as were performed to validate the proteomic findings and to evaluate the catalytic impairment of this enzyme in the INHr pairs. The confirmation of the catalytic capacity of the INHr strains is also relevant since previous experiments suggest that *katG* mutations conferring INH resistance cause different levels of peroxidase activity in *Mtb* (78). *In vitro* cultures revealed lower levels of the protein and reduced peroxidase activity in the cytosol and culture supernatant of both INHr strains in comparison with equivalent protein quantities of their respective INHs strains. These results confirm that *katG* N-terminus single or double mutations (Val1Ala and Va1Ala/Glu3Val, observed in the laboratory and clinical INHr strain respectively) resulted in a significantly reduced intracellular and secreted levels of atG and an associated lower peroxidase activity (p-value<0.05) (Figure 3.12). Under this condition, *Mtb* have an impaired ability to neutralize and overcome the highly oxidative intracellular and external environment in the phagosome and it could be rapidly cleared by macrophages and other phagocytes (79, 80). Other proteins such as

AcpM, FbpA, HspX in the cytosol and Mpt32 (Apa) in the CFP fraction of the INHr strain were

confirmed by western blot. Furthermore, higher levels of the protein FtsZ were confirmed for the

laboratory INHr strain only (NSAF data on Appendix I Table 5) (Figure 3.13).
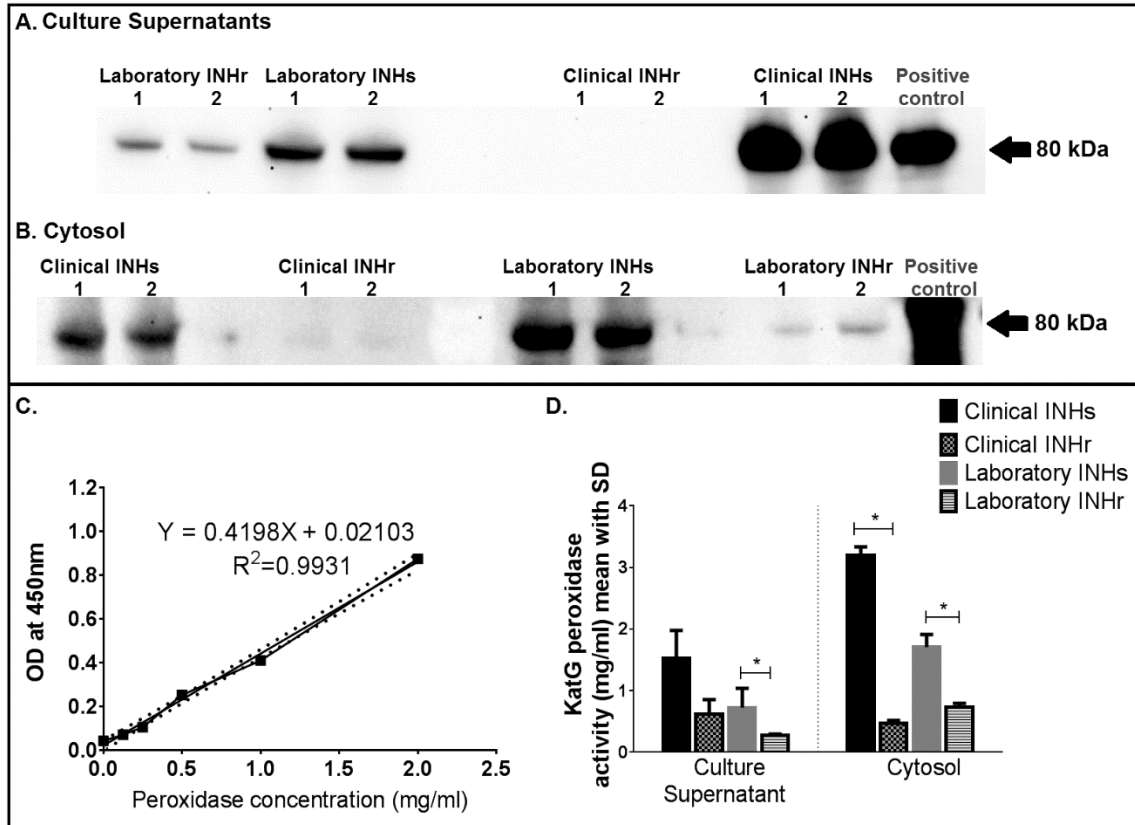
Figure 3.12. KatG levels and peroxidase activity in the clonal *Mtb* strains. KatG western blot in
A. CFP and B. Cytosol of the clonal strains. C. Standard curve for peroxidase activity using
3,3′,5,5′-Tetramethylbenzidine (TMB) as substrate and recombinant KatG. D. KatG peroxidase
activity (derived from the standard curve, from laboratory and clinical *Mtb* pairs using
recombinant KatG as reference. Non-paired *t*-test without assuming consistent standard
deviation, *$p<0.05$.

Although catalase and peroxidase are both functions of KatG that reduce H_2O_2

generating water, the reaction that both catalyzes is different as the peroxidase requires a

hydrogen donor to complete its reaction (81). The assay previously described measures

peroxidase activity mainly, as TMB functions as the hydrogen donor and is oxidized to 3,3',5,5'-

tetramethylbenzidine diimine. The catalase, on the contrary, can directly degrade H_2O_2 with the

generation of oxygen. Therefore, the assay presented here is specific to measure the peroxidase

activity of KatG. The catalase and peroxidase activity are independent functions of KatG, as is

the INH-NAD synthase activity (81). However, the N-terminus mutations present in clinical and

laboratory strains of this dissertation generated a reduced activity in both peroxidase and INH-

NAD synthase activity of KatG that was confirmed by the enzymatic level and by looking at the

INHr phenotype.
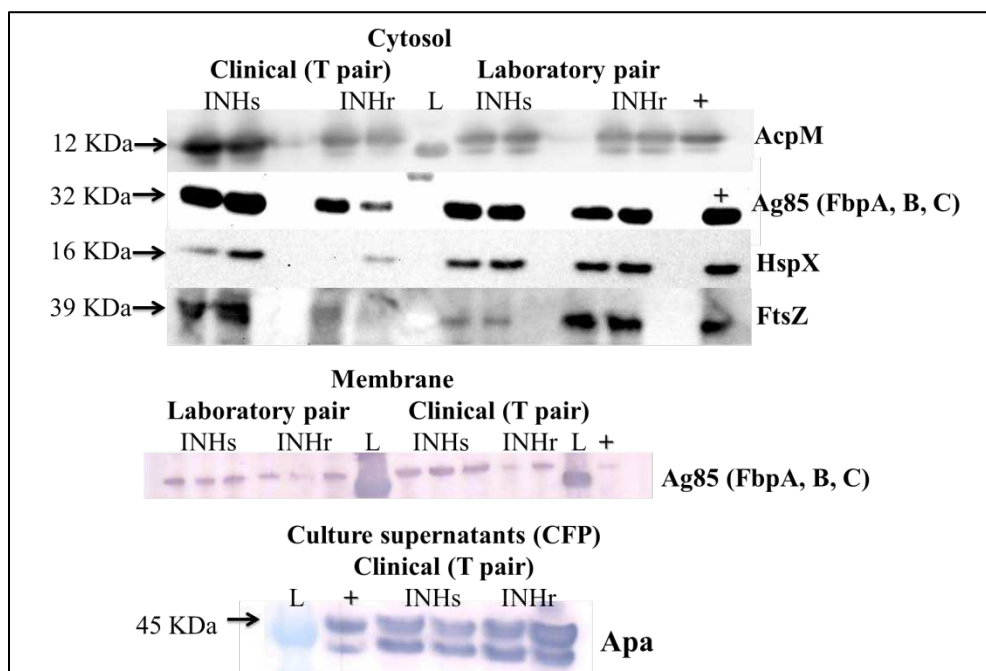
Figure 3.13. Western blot (WB) validation for some significantly different proteins the clinical and laboratory INHr strains revealed by chemiluminescent and colorimetric reaction. Two biological replicates were used for most WB. For Ag85 in the membrane fraction, three biological replicates were used, except for the clinical INHr strain (two replicates).

3.3.4 Lipidomics

After evaluating all the differences in the proteome at each *Mtb* pair comparison, it is

evident that FAS I and II routes, responsible for the generation of fatty acids including the *Mtb*

hallmark MA, are affected in each INHr strain, albeit with different trends in each one (Figure

3.6 and 3.7). Additionally, being one of the most recognized targets for INH, MA biosynthesis is

a very probable pathway to be altered in INHr strains that were previously exposed to the drug. In general, the MA analysis of both pairs showed more differences in the clinical than in the laboratory comparison, similar to the discovery of more protein alteration in the clinical INHr strain (Cloud plot, Figure 3.14). Additionally, the distribution of the samples looks more homogeneous in the clinical pair comparison, as observed in the Principal component graph (Figure 3.14). These mycolic acids were obtained from the saponification of the cell wall of the delipidated *Mtb* cells with KOH dissolved in methanol at high temperature (80°C). Since MA are not soluble in methanol, they can be separated from moderately long chain fatty acids with an ether solution (82). The resulting analysis of MA by LC-TOF is presented in this section. The XCMS online platform was used to identify the significant features by the criteria of fold change >2 or <0.5 and p<0.05 (t-test). As shown in the cloud plot of figure 3.14, there were 115 and 21 significantly different features in the clinical and laboratory pair comparison respectively. These numbers were severely reduced to 23for the clinical pair comparison and to zero to the laboratory pair comparison after evaluating the spectra of the each significant features and manually excluding that the presence of those significant features that were isotopes and selecting those with m/z values between 900 and 1,500. The selection criterion includes masses in the previous range, in order to observe α- prime mycolic acids at the lower mass range or methoxy-MA with 94 carbons in the higher mass range. After this, m/z values for these features were interrogated against Sartain database with a window range of +/- 0.05 with the positive adducts Na+, H+, and NH4+. From the total 23 significantly different features in the clinical INHr, 20 were unambiguously identified by MS-LAMP (Table 3.3). The three features that could not be identified by MS-LAMP could not be identified by MycoMass database either. Most features were increased in the INHr strain (Table 3.3). The spectra and the extracted ion

chromatogram for all the significant different features from the clinical MA comparison are

presented in Appendix II Figures 1 to 20. Some features that were significant different in the

laboratory clonal pair comparison and did not meet the quality control criteria are presented also

in Appendix II Figures 21 and 22.

Table 3.3. Significant MA identified through XCMS and MS-LAMP analysis in clinical clonal *Mtb* strains comparison

MS-LAMP Identification	Adducts identified	Fold change (ion)	P value	Level in clinical INHr strain	m/z med*	Retention time med*
α-MA C75	NH$_4$+, Na, H+	0.40 (NH$_4$+)	0.013236	DOWN	1113.156 (NH$_4$+)	13.13
α-MA *C77*	NH$_4$+, Na	*2.7* (NH$_4$+)	*0.015792*	*UP*	*1141.187* (NH$_4$+)	*10.93*
α-MA C78	Na+	2.6	9.22E-05	UP	1160.157	13.33
α-MA C78	Na+	2.0	0.046855	UP	1160.158	13.78
α-*MAC79*	*NH$_4$+, Na+*	*3.4* (NH$_4$+)	*0.000483*	*UP*	*1169.217* (NH$_4$+)	*11.37*
α -MAC80	Na+	2.1	0.007323	UP	1188.189	14.19
α-MAC80	Na+	2.8	0.012174	UP	1188.188	13.75
α -*MAC81*	*NH4+, Na+*	*2.4* (NH$_4$+)	*0.000952*	*UP*	*1197.246* (NH$_4$+)	*11.83*
α -*MA C83*	*Na+*	*2.5*	*0.002396*	*UP*	*1230.237*	*12.28*
α -*MA C85*	*Na+*	*2.5*	*0.001931*	*UP*	*1258.268*	*12.69*
Keto C74	Na+	2.6	0.033158	UP	1120.126	13.41
Keto C76	Na+	2.3	0.00463	UP	1148.157	13.85
Keto C80	NH4+	2.5	0.000624	UP	1199.219	16.33
Methoxy C79	H+	2.5	0.000813	UP	1170.187	15.98
Methoxy C81	H+	2.6	0.00029	UP	1198.217	16.33
Feature 21*****		3.0	0.000137	UP	1193.261	16.32
Feature 22*****		2.9	0.002189	UP	1165.228	15.97
Feature 23*****		2.1	0.010469	UP	1168.128	10.92

*med: Median values from three biological replicates, **could not be identified by MS-LAMP.
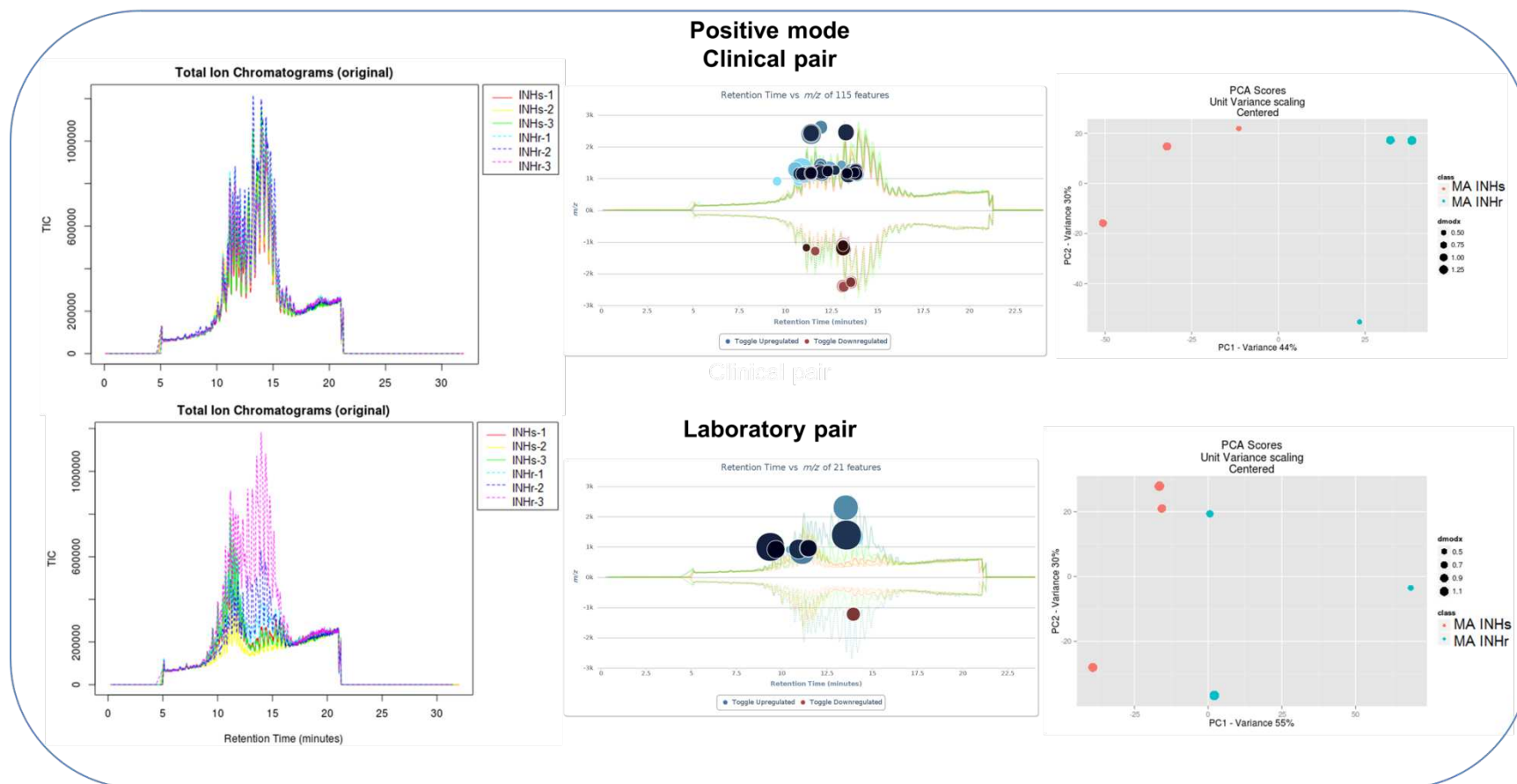
3.14. Lipidomic approach for the MA analysis of clinical and laboratory pairs. The graph showed in the lower portion of the figure is the XCMS online platform output from left to right: total ion chromatogram, cloud plot including all the significant different features (without manual validation) and the Principal component analysis (PCA) for each pair comparison.

136

Looking closely at the identified features in the clinical pair comparison, a lack of correlation between the number of carbons and retention times was observed in some representatives of the α-MA (Table 3.3). Since the HPLC column used in this experiment separates compounds based on their hydrophobicity (C18 chemistry), it is expected that there would be a sequential order in the appearance of compounds depending on the number of carbons. In this way, molecules with a lower number of carbons are eluted from the column earlier retention times compared to molecules with a longer number of carbons. Additionally, after looking the total ion chromatogram, it is evident that most of these differences occur in the earlier peaks (Figure 3.15). With this in mind, a hydrogenation reaction was performed to the MA of these strains to see if these earlier compounds had some level of unsaturation (presence of double bounds) and explain those significantly different earlier α-MA that share the m/z value according to Sartain's database but had unusual retention times compared to the other α-MA also identified in this study.
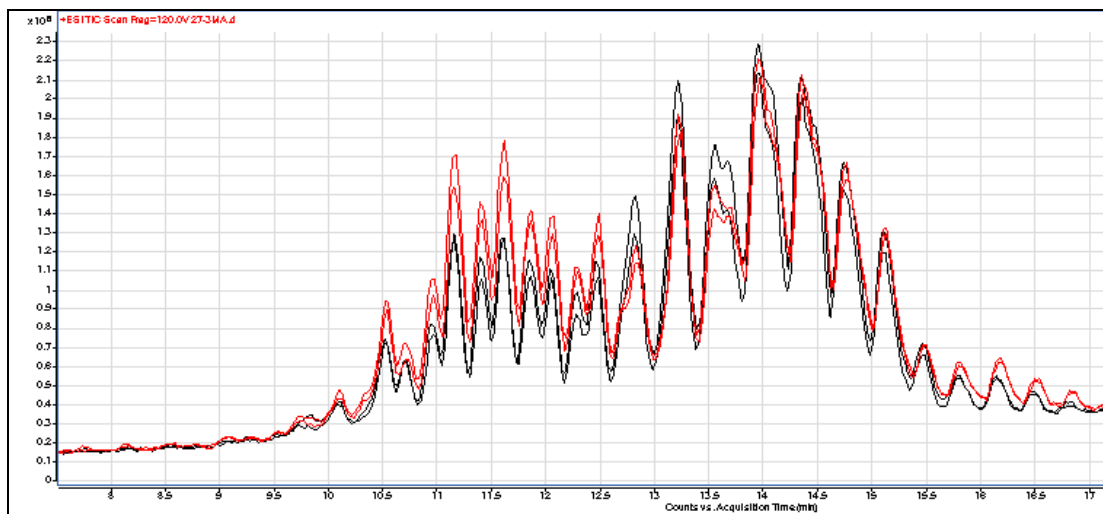
Figure 3.15. Total ion chromatogram for the MA in the clinical pair comparison. Red: replicates of the INHr, black: replicates for the INHs strain.

For the hydrogenation experiment, it should be noted that the hydrogenation experiment was analyzed in different LC-TOF instrument (Agilent 6220 vs the original analysis in Agilent

6230), therefore, both a hydrogenated and the original MA extract for each strain were analyzed simultaneously to facilitate the interpretation of the results. For this reason also, the retention time values increased about 1.1 minutes. After the hydrogenation reaction, the peaks of those features that did not have a linear correlation between the number of carbons and retention times were no longer observed in the individual EIC (Figure 3.16). This suggests that a reduction of the double bonds occurred, these significant compounds were therefore grouped as α-MA with a level of unsaturation or presence of double bonds (Figure 3.16). The remaining α-MA that did not have altered the peaks after the hydrogenation reaction occurred had a linear relationship in their retention time values with the number of carbons, which suggests that these are probably saturated α-MA (Figure 3.16). It should be noted that there are two features identified as α-MA C78 and other two features identified as C80 with almost the exact mass but different retention times, which can be seen as two different peaks in the EIC. These features could be representing structural isomers and also explains why the graph of retention time vs m/z in the saturated α-MA is more heterogeneous (Figure 3.17).
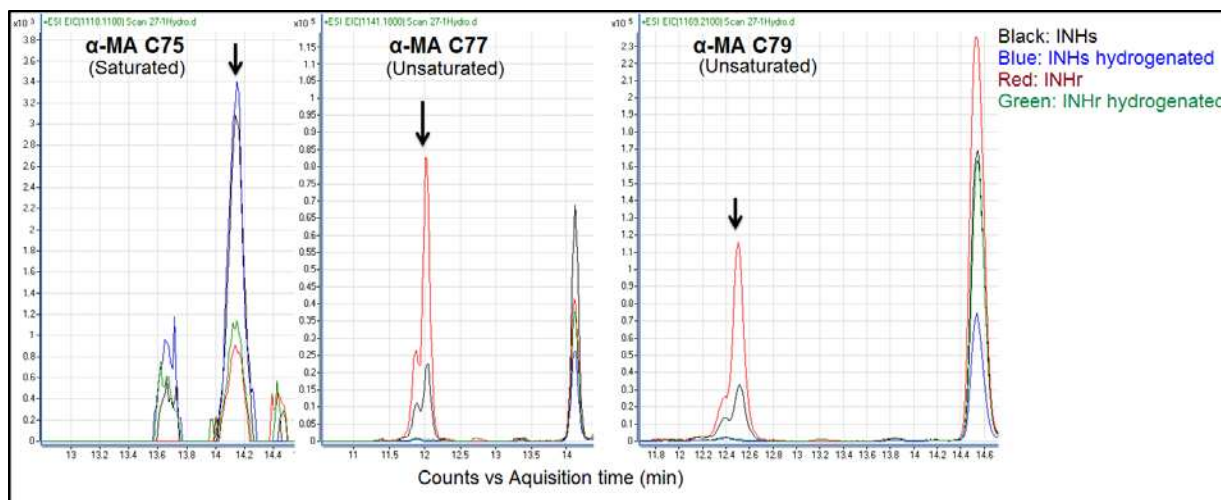
Figure 3.16. Extracted ion chromatogram from some significant different features in the MA comparison of the clinical strains before and after hydrogenation, revealing some level of unsaturation in the α-MA C77 and 79. In the case of α-MA C75 as the peaks for this feature

before and after hydrogenation stayed the same, suggesting no double bonds in this α-MA. Pictures from Agilent mass Hunter Qualitative Analysis software output, version 4.00.
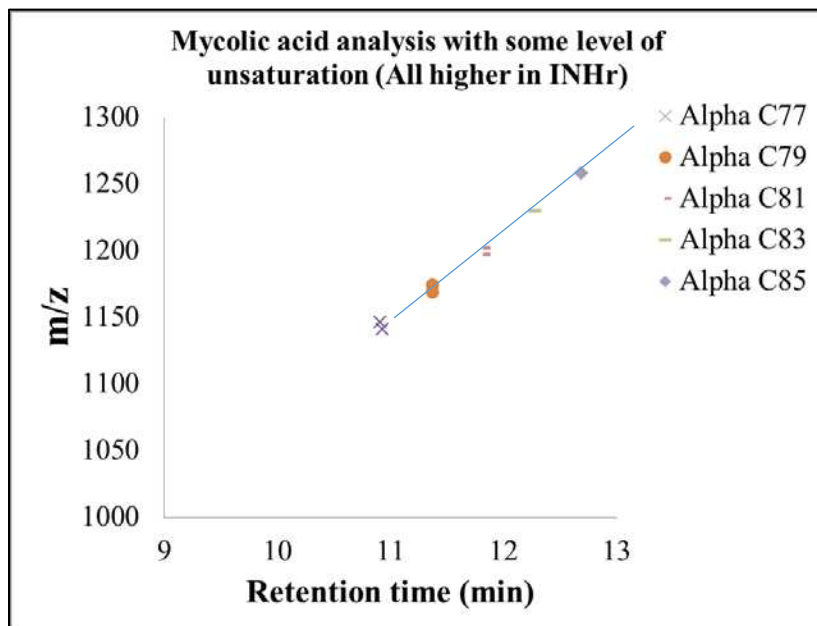
Figure 3.17. Distribution of the α-MA with a possible level of unsaturation as suggested by the hydrogenation experiment.
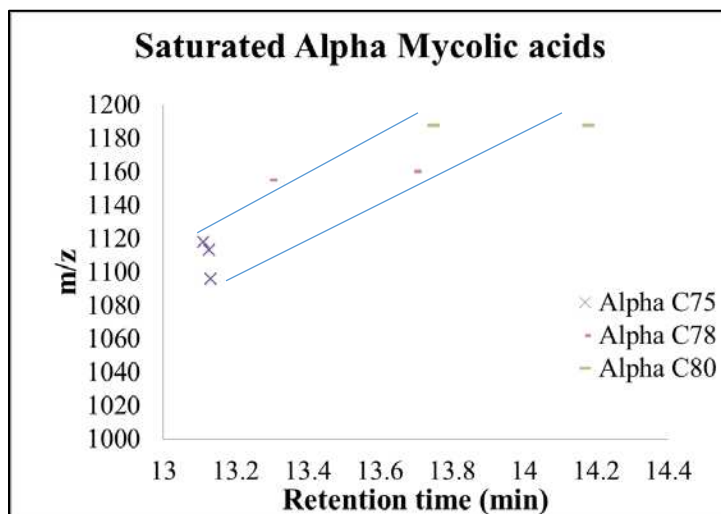
Figure 3.18. Distribution of seven significant features identified as α-MA in table 3.5, with a possibly saturated MA as suggested by the hydrogenation experiment. Two suggested trend lines indicate further variation within the group of saturated α-MA.

3.4 Conclusions

3.4.1 Protein similarities after the acquisition of INHr and relation with previous findings. This study presents a snapshot of genetically-related *Mtb* strains under the same *in vitro* conditions to detect specific changes that result in the INHr strains. At least two initial known common events occurred in the INHr *Mtb* strains analyzed here, first they were exposed to INH and then developed resistance to the drug with mutations in the N-terminus of the *katG* gene (Chapter II). These mutations generate an important reduction in KatG levels, which in addition to inhibit INH activation, also impair the bacteria's ability to handle oxidative stress, an important cellular process both *in vivo* and *in vitro* (83, 84). Additionally, a significant reduction in the peroxidase activity in both clonal INHr strains was observed in this study (Figure 3.12). Reduction in the peroxidase activity and INH resistance due to KatG mutations are not normally correlated events as previously demonstrated by Cade *et al* (83). The peroxidase activity reduction observed in these N-terminus mutants is probably associated with the overall reduction of the protein levels in general, more than a change in the protein structure that could impair its activity. Further variation in protein abundance revealed similar metabolic trends, such as an increase in the energy metabolism and ATP synthesis, increase of detoxifying enzymes that are particularly important after the also increased β-oxidation pathway, among others (Table 3.1, Figures 3.3-3.10).

Many of the altered enzymes found in this study have a high affinity for the INH-NAD as previously demonstrated by a computational study by Jena et al (85). Both clinical and laboratory pair comparisons had a significant variation in the levels of these enzymes described by Jena et al (85) that included SahH, Rv0753c, Rv2723, FadB2, InhA, Rv2971, and Rv2766c. The oxidoreductase Rv2971 was also altered in a previous proteomic study that did not use

genetically-related *Mtb* pairs, however with an opposite trend to the one described for the both

INHr of this dissertation (Table 3.1) (86). Other proteins described in the computational study

were also found with different abundances in the INHr strains, although the t-test only revealed a

tendency, such as increased levels of PntaA (Rv0199, p=0.052) that participate in the NAD

phosphorylation and trans-hydrogenation as well as decreased levels of Mtn (Rv0091, p=0.053)

that participate in the SAM metabolism. It is important to highlight that other proteins associated

with SAM and NAD metabolism with alterations were found in this dissertation. Besides being

the cofactor required for the synthesis of the active form of the INH, NADH is an important

nucleotide used in many relevant reactions vital for the cell. These include, central carbon

metabolism, energetic metabolism, and lipid metabolism, most of which were increased in the

INHr strains. Probably the tendency to be increased or decreased depends on the affinity of the

enzyme for NADH or the flexibility to use also NADPH as a cofactor, as is the case of FabG4

(that can use both cofactors but has a higher affinity for NADH).

A previous study with clinic clonal pairs of the Central-Asian (CAS) 1-Delhi lineage that

transition from drug susceptible to multidrug-resistant (INHr plus rifampicin resistant) MDR,

demonstrated the increased levels of 26 proteins, including some that were also observed in the

present study, such as FadB, FabG4 (Figure 3.6 and 3.9), LipN (Figure 3.9), MoxR (Table 3.2),

and RpoA (Appendix I Table 5) (87). In addition to increased RpoA levels in the laboratory

INHr strain, there were also higher levels of RpoB and RpoC in the clinical INHr strain

(Appendix I Table 6 and 8). Interestingly, we found increased levels of the protein Rv3327 (a

probable transposase fusion protein) in the membrane and cell wall of the clinical INHr strain

(Appendix I Table 2 and 8). In the proteomic study of the clinical *Mtb* strain CAS1 genotype that

developed a MDR phenotype, they found altered levels of five proteins in the insertion

sequences and phages category were found; however the name of the proteins was not specified (87). The role of Rv3327 is not clear yet, but may be associated with drug resistance in specific *Mtb* lineages.

3.4.2 Proteomics and lipidomics in the clinical INHr strain strongly suggest an alternative MA biosynthetic pathway in the clinically-isolated drug resistant Mtb strains.

In the particular case of the clinical pair, the alteration observed in the FAS II route is probably related to the fact that InhA, KasB and AcpM are important proposed targets for INH. Previous exposure to the drug followed by acquisition of INHr affects the levels of these proteins. These FAS-II canonical FAS II enzymes were reduced in both INHr strains. However, these enzymes are normally upregulated in *Mtb* strains after exposure to INH (88). These protein changes are not associated with mutations in those genes, as previously revealed by a whole genomics sequencing study (19). The laboratory strain on the other hand, only showed reduced levels of InhA, which could be due to a variety of reasons including the genetic background already discussed in chapter one (one versus two mutations in the *katG* gene and different genotype). According to the sequence homology, Rv3485 and other hypotheticals share up to 26% of identity with InhA (34). Rv3485 is a short chain dehydrogenase as well as EphD. EphD was similarly increased in both INHr strains. It is possible that EphD could compensate for the reduced abundance of InhA in the INHr strains. Additionally, the similarly increased epoxide hydrolase EphF shares 18.7% identity with InhA according the Multiple Sequence Alignment platform MUSCLE. Further functional analysis would confirm if these increased proteins play a role in the synthesis of fatty acids in the INHr phenotype. Additional factors should be considered to explain the higher reduction in proteins of the FAS II in the clinical INHr strains.

These include external factors, such as the number of drugs to which *Mtb* strain are exposed in the host, the duration of the drug exposure, among others.

Interestingly, all the α-MA (either saturated or not) were increased in the clinical INHr strain; except tor α-MA C75, simultaneous with the reduced levels of some canonical FAS II enzymes. It is possible that an alternative fatty acid biosynthetic pathway is the main route for the production of MA in the clinical INHr. This idea is supported by the proteomic findings that showed an increase of FabG4, HtdX and FadD13 (Figure 3.6 and 3.7) with the parallel reduction of the core FAS II enzymes AcpM, KasB, and FadD32. Importantly, FabG4 was not increased as a possible compensation of its homolog FabG1, since the last one was also increased in the CYT and MEM of the INHr strain albeit not significantly (p=0.07 and 0.09 respectively). It should be noted that the FAS I enzymes, Fas and AccA3, were reduced in the clinical INHr strain (Figure 3.7). Putting this together, it is possible that FabG1prefers substrates with longer carbon chains for elongation, therefore, the increase of FabG4 could be due to the higher availability of fatty acids with lower number of carbons as a result of reduced number of FAS I cycles (Figure 3.7). This alternative biosynthetic pathway was proposed to be required for drug resistance in *Mtb* (64, 89). Also important to note, the FabG4 protein was commonly increased in the MDR strain from the CAS1genotype (87). FabG4 and HtdX were also increased in the INHr stains from the Beijing genotype (Chapter II, table 2.2). The increase of FabG4 is probably related to the exposure to the first line drugs with the subsequent development of the INHr phenotype, events that all these *Mtb* strains had in common, that then can stimulate the entire branch of alternative fatty acid biosynthetic pathway.

Additionally, in this alternative biosynthetic pathway, the lower levels observed only for α-MA C75 could be a result of this molecule serving as a precursor for the introduction of

unsaturations and cyclopropane rings for the synthesis of MA with higher number of carbons. The presence of two features with the same m/z values and with the same adduct (Na^+) for α-MA C78 and α-MA C80, but different retention times, may indicate that there are also increased intensity peaks for variants of cyclopropanated species among the saturated α-MA, (Table 3.3). The latter can also be associated with the increased levels of the cyclopropane synthase PcaA (Figure 3.6 and 3.7). The lipidomics results suggest that the synthesis of the MA could be related with the alternative biosynthetic pathway. Furthermore, common trends in MA levels were not observed between the clinical and laboratory comparison. This was expected since the degree of alteration of the core FAS II enzymes in the clinical INHr was higher than in the laboratory INHr strain (Figure 3.8). On the other hand, the laboratory INHr strain has altered level in most of the mycolates modifying MT enzymes (Figure 3.8).

It is important to note that most of the differences in the MA presented in this study were in the less intense peaks. This was expected for the case of keto and methoxy-MA that together represent the 25% of the total MA (58), which further explains the peaks with low intensity observed in the EIC for the clinical strain comparison that represent those oxygenated species of MA (Appendix II and Table 3.3). In the case of the laboratory strain comparison, probably the differences in the MA are mostly in the oxygenated MA and according to the proteomics analysis; this difference could be particularly in the ratio of *cis* versus *trans* double bounds or cyclopropane rings that could not be detected following the experimental strategy used in this study.

Finally, it is also clear that, by looking at the metabolic rearrangements and the different levels of the proteins in the virulence and detoxification group as well as in those involved in protein secretion; it is expected that INHr would have an altered behavior during their growth *in*

144

vivo. These proteins are grouped in Figure 3.4 and Table 3.2 respectively and add more supporting information to the already expected reduced virulence phenotype due to the reduced level of KatG in both strains. This work demonstrates that mycobacterial resistance to INH is a complex process that orchestrates a series of physiological events in a mutation- and strain-dependent manner. Our finding of a new set of proteins involved in four different biological pathways in *Mtb* after acquisition of INHr related to the mechanism of action of this drug could be also responsible for an *Mtb* strain with a reduced virulent phenotype. This study also provides a better understanding of new compensatory mechanisms in *Mtb* after INHr that could address alternative combined therapies as well as candidates for new drug targets in INH resistant strains such as FabG4 and provides supporting data about possible new metabolic routes, such as the alternative MA biosynthetic pathway.

In summary, after acquisition of INHr due to *katG* N-terminus mutations, *Mtb* strains have a proteomic rearrangement that includes common changes for instance in the ß-oxidation of fatty acids and energetic metabolism, as well as individual changes that are going to depend on the genetic background of the bacterium. The genetic background includes not only the bacterial linage, but also the specific SNP that is causing the INHr phenotype. Among specific proteome changes, the clinical INHr strain provided the highest differences probably linked to the double mutation found in the *katG* gene. Interestingly, one of the altered pathways in the clinical INHr strain, FAS II pathway, was compensated by an alternative biosynthetic pathway which could be corroborated after looking at the mycolic acid profile differences (some of them even higher in the clinical INHr strain). Proteins of this pathway were also found to be increased in another proteomic study of MDR-clinical *Mtb* strains, as well as in the study described in chapter II. The later also propose the idea that some of these changes could be generated due to the bacterial

exposure to the combined drug therapy in the first place. This idea should be further explored in

a clinical context to provide a better understanding of the natural process of drug-resistance

acquisition in *Mtb*.

REFERENCES

1. Aebersold R, Mann M. Mass spectrometry-based proteomics. Nature. 2003;422(6928):198-207.
2. Domon B, Aebersold R. Options and considerations when selecting a quantitative proteomics strategy. Nature Biotechnology. 2010;28(7):710-21.
3. Nogueira FCS, Domont GB. Survey of Shotgun Proteomics. 2014;1156:3-23.
4. McNaught AD, Wilkinson A, International Union of Pure and Applied Chemistry. Compendium of chemical terminology : IUPAC recommendations. 2nd ed. Oxford England ; Malden, MA, USA: Blackwell Science; 1997.
5. Schwartz JC, Senko MW, Syka JEP. A two-dimensional quadrupole ion trap mass spectrometer. Journal of the American Society for Mass Spectrometry. 2002;13(6):659-69.
6. Welling M, Schuessler HA, Thompson RI, Walther H. Ion/molecule reactions, mass spectrometry and optical spectroscopy in a linear ion trap. International Journal of Mass Spectrometry and Ion Processes. 1998;172(1-2):95-114.
7. Hu Q, Noll RJ, Li H, Makarov A, Hardman M, Graham Cooks R. The Orbitrap: a new mass spectrometer. Journal of Mass Spectrometry. 2005;40(4):430-43.
8. Makarov A, Denisov E, Kholomeev A, Balschun W, Lange O, Strupat K, et al. Performance evaluation of a hybrid linear ion trap/orbitrap mass spectrometer. Anal Chem. 2006 Apr 1;78(7):2113-20.
9. Cole RB, Ebooks Corporation. Electrospray and MALDI mass spectrometry fundamentals, instrumentation, practicalities, and biological applications. Hoboken, N.J.: Wiley,; 2010. Available from: http://www.CSU.eblib.com/patron/FullRecord.aspx?p=573866.
10. Leray C. Introduction to lipidomics : from bacteria to man. Boca Roton: CRC Press; 2013.
11. Sartain MJ, Dick DL, Rithner CD, Crick DC, Belisle JT. Lipidomic analyses of Mycobacterium tuberculosis based on accurate mass measurements and the novel "Mtb LipidDB". J Lipid Res. 2011 May;52(5):861-72.
12. Layre E, Sweet L, Hong S, Madigan CA, Desjardins D, Young DC, et al. A comparative lipidomics platform for chemotaxonomic analysis of Mycobacterium tuberculosis. Chem Biol. 2011 Dec;18(12):1537-49.
13. Organization WH. Global Tuberculosis Report2015.
14. O'Sullivan DM, McHugh TD, Gillespie SH. The effect of oxidative stress on the mutation rate of Mycobacterium tuberculosis with impaired catalase/peroxidase function. J Antimicrob Chemother. 2008 Oct;62(4):709-12.
15. Bergval I, Kwok B, Schuitema A, Kremer K, van Soolingen D, Klatser P, et al. Pre-existing isoniazid resistance, but not the genotype of Mycobacterium tuberculosis drives rifampicin resistance codon preference in vitro. PLoS One. 2012;7(1):e29108.
16. van Soolingen D, de Haas PE, Kremer K. Restriction fragment length polymorphism typing of mycobacteria. Methods Mol Med. 2001;54:165-203.
17. Kamerbeek J, Schouls L, Kolk A, van Agterveld M, van Soolingen D, Kuijper S, et al. Simultaneous detection and strain differentiation of Mycobacterium tuberculosis for diagnosis and epidemiology. J Clin Microbiol. 1997 Apr;35(4):907-14.

18.   Almeida D, Nuermberger E, Tasneen R, Rosenthal I, Tyagi S, Williams K, et al. Paradoxical effect of isoniazid on the activity of rifampin-pyrazinamide combination in a mouse model of tuberculosis. Antimicrob Agents Chemother. 2009 Oct;53(10):4178-84.

19.   Datta G, Nieto LM, Davidson RM, Mehaffy C, Pederson C, Dobos KM, et al. Longitudinal whole genome analysis of pre and post drug treatment Mycobacterium tuberculosis isolates reveals progressive steps to drug resistance. Tuberculosis. 2016;98:50-5.

20.   Chambers MC, Maclean B, Burke R, Amodei D, Ruderman DL, Neumann S, et al. A cross-platform toolkit for mass spectrometry and proteomics. Nat Biotechnol. 2012 Oct;30(10):918-20.

21.   Nesvizhskii AI, Keller A, Kolker E, Aebersold R. A statistical model for identifying proteins by tandem mass spectrometry. Anal Chem. 2003 Sep 1;75(17):4646-58.

22.   Zhang Y, Wen Z, Washburn MP, Florens L. Refinements to label free proteome quantitation: how to deal with peptides shared by multiple proteins. Anal Chem. 2010 Mar;82(6):2272-81.

23.   Johnstone Robert A. W. WAH, Entwistle Ian D. Heterogeneous catalytic transfer hydrogenation and its relation to other methods for reduction of organic compounds. *Chemical Reviews*. 1985;85(2):41.

24.   Tautenhahn R, Patti GJ, Rinehart D, Siuzdak G. XCMS Online: a web-based platform to process untargeted metabolomic data. Anal Chem. 2012 Jun 5;84(11):5035-9.

25.   Sabareesh V, Singh G. Mass spectrometry based lipid(ome) analyzer and molecular platform: a new software to interpret and analyze electrospray and/or matrix-assisted laser desorption/ionization mass spectrometric data of lipids: a case study from Mycobacterium tuberculosis. J Mass Spectrom. 2013 Apr;48(4):465-77.

26.   Banerjee S, Nandyala A, Podili R, Katoch V, Hasnain SE. BMC Biochemistry. 2005;6(1):20.

27.   Tian J, Bryk R, Itoh M, Suematsu M, Nathan C. Variant tricarboxylic acid cycle in Mycobacterium tuberculosis: identification of alpha-ketoglutarate decarboxylase. Proc Natl Acad Sci U S A. 2005 Jul;102(30):10670-5.

28.   Finn RD, Bateman A, Clements J, Coggill P, Eberhardt RY, Eddy SR, et al. Pfam: the protein families database. Nucleic Acids Res. 2014 Jan;42(Database issue):D222-30.

29.   Laval F, Haites R, Movahedzadeh F, Lemassu A, Wong CY, Stoker N, et al. Investigating the function of the putative mycolic acid methyltransferase UmaA: divergence between the Mycobacterium smegmatis and Mycobacterium tuberculosis proteins. J Biol Chem. 2008 Jan;283(3):1419-27.

30.   Belisle JT. Role of the Major Antigen of Mycobacterium tuberculosis in Cell Wall Biogenesis. Science. 1997;276(5317):1420-2.

31.   Puech V, Guilhot C, Perez E, Tropis M, Armitige LY, Gicquel B, et al. Evidence for a partial redundancy of the fibronectin-binding proteins for the transfer of mycoloyl residues onto the cell wall arabinogalactan termini of Mycobacterium tuberculosis. Molecular Microbiology. 2002;44(4):1109-22.

32.   Nguyen L, Chinnapapagari S, Thompson CJ. FbpA-Dependent biosynthesis of trehalose dimycolate is required for the intrinsic multidrug resistance, cell wall structure, and colonial morphology of Mycobacterium smegmatis. J Bacteriol. 2005 Oct;187(19):6603-11.

33.   Armitige LY, Jagannath C, Wanger AR, Norris SJ. Disruption of the genes encoding antigen 85A and antigen 85B of Mycobacterium tuberculosis H37Rv: effect on growth in culture and in macrophages. Infect Immun. 2000 Feb;68(2):767-78.

34.   Takayama K, Wang C, Besra GS. Pathway to synthesis and processing of mycolic acids in Mycobacterium tuberculosis. Clin Microbiol Rev. 2005 Jan;18(1):81-101.

35.   Marrakchi H, Lanéelle MA, Daffé M. Mycolic acids: structures, biosynthesis, and beyond. Chem Biol. 2014 Jan;21(1):67-85.

36.   de Souza MVN, Ferreira MdL, Pinheiro AC, Saraiva MF, de Almeida MV, Valle MS. Synthesis and Biological Aspects of Mycolic Acids: An Important Target AgainstMycobacterium tuberculosis. The Scientific World JOURNAL. 2008;8:720-51.

37.   Marrakchi H, Laneelle MA, Daffe M. Mycolic acids: structures, biosynthesis, and beyond. Chem Biol. 2014 Jan 16;21(1):67-85.

38.   Cole ST, Brosch R, Parkhill J, Garnier T, Churcher C, Harris D, et al. Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence. Nature. 1998;393(6685):537-44.

39.   Taylor RC, Brown AK, Singh A, Bhatt A, Besra GS. Characterization of a -hydroxybutyryl-CoA dehydrogenase from Mycobacterium tuberculosis. Microbiology. 2010;156(7):1975-82.

40.   Kumar A, Farhana A, Guidry L, Saini V, Hondalus M, Steyn AJ. Redox homeostasis in mycobacteria: the key to tuberculosis control? Expert Rev Mol Med. 2011;13:e39.

41.   Chavadi S, Wooff E, Coldham NG, Sritharan M, Hewinson RG, Gordon SV, et al. Global effects of inactivation of the pyruvate kinase gene in the Mycobacterium tuberculosis complex. J Bacteriol. 2009 Dec;191(24):7545-53.

42.   Singh G, Singh G, Jadeja D, Kaur J. Lipid hydrolizing enzymes in virulence:Mycobacterium tuberculosisas a model system. Critical Reviews in Microbiology. 2010;36(3):259-69.

43.   Jadeja D, Dogra N, Arya S, Singh G, Singh G, Kaur J. Characterization of LipN (Rv2970c) ofMycobacterium TuberculosisH37Rv and its Probable Role in Xenobiotic Degradation. Journal of Cellular Biochemistry. 2016;117(2):390-401.

44.   Singh G, Arya S, Kumar A, Narang D, Kaur J. Molecular Characterization of Oxidative Stress-Inducible LipD of Mycobacterium tuberculosis H37Rv. Current Microbiology. 2013;68(3):387-96.

45.   Pérez E, Samper S, Bordas Y, Guilhot C, Gicquel B, Martín C. An essential role for phoP in Mycobacterium tuberculosis virulence. Molecular Microbiology. 2001;41(1):179-87.

46.   Asensio JG, Maia C, Ferrer NL, Barilone N, Laval F, Soto CY, et al. The Virulence-associated Two-component PhoP-PhoR System Controls the Biosynthesis of Polyketide-derived Lipids inMycobacterium tuberculosis. Journal of Biological Chemistry. 2006;281(3):1313-6.

47.   Deng W, Wang H, Xie J. Regulatory and pathogenesis roles of Mycobacterium Lrp/AsnC family transcriptional factors. Journal of Cellular Biochemistry. 2011;112(10):2655-62.

48.   Sulzenbacher G, Canaan S, Bordat Y, Neyrolles O, Stadthagen G, Roig-Zamboni V, et al. LppX is a lipoprotein required for the translocation of phthiocerol dimycocerosates to the surface of Mycobacterium tuberculosis. The EMBO Journal. 2006;25(7):1436-44.

49.   Chen L, Xie QW, Nathan C. Alkyl hydroperoxide reductase subunit C (AhpC) protects bacterial and human cells against reactive nitrogen intermediates. Mol Cell. 1998 May;1(6):795-805.

50.   Sherman DR, Mdluli K, Hickey MJ, Arain TM, Morris SL, Barry CE, et al. Compensatory ahpC gene expression in isoniazid-resistant Mycobacterium tuberculosis. Science. 1996 Jun;272(5268):1641-3.

51.   Kumar K, Awasthi D, Berger WT, Tonge PJ, Slayden RA, Ojima I. Discovery of anti-TB agents that target the cell-division protein FtsZ. Future Medicinal Chemistry. 2010;2(8):1305-23.
52.   Ward JE, Lutkenhaus J. Overproduction of FtsZ induces minicell formation in E. coli. Cell. 1985;42(3):941-9.
53.   Begg K, Nikolaichik Y, Crossland N, Donachie WD. Roles of FtsA and FtsZ in activation of division sites. J Bacteriol. 1998 Feb;180(4):881-4.
54.   Rowlett VW, Margolin W. The Min system and other nucleoid-independent regulators of Z ring positioning. Frontiers in Microbiology. 2015;6.
55.   Vijay S, Nagaraja M, Sebastian J, Ajitkumar P. Asymmetric cell division in Mycobacterium tuberculosis and its unique features. Archives of Microbiology. 2014;196(3):157-68.
56.   Sirakova TD, Thirumala AK, Dubey VS, Sprecher H, Kolattukudy PE. The Mycobacterium tuberculosis pks2 gene encodes the synthase for the hepta- and octamethyl-branched fatty acids required for sulfolipid synthesis. J Biol Chem. 2001 May 18;276(20):16833-9.
57.   Banerjee A, Sugantino M, Sacchettini JC, Jacobs WR. The mabA gene from the inhA operon of Mycobacterium tuberculosis encodes a 3-ketoacyl reductase that fails to confer isoniazid resistance. Microbiology. 1998 Oct;144 (Pt 10):2697-704.
58.   Takayama K, Wang C, Besra GS. Pathway to Synthesis and Processing of Mycolic Acids in Mycobacterium tuberculosis. Clinical Microbiology Reviews. 2005;18(1):81-101.
59.   Yuan Y, Lee RE, Besra GS, Belisle JT, Barry CE, 3rd. Identification of a gene involved in the biosynthesis of cyclopropanated mycolic acids in Mycobacterium tuberculosis. Proc Natl Acad Sci U S A. 1995 Jul 3;92(14):6630-4.
60.   Kikuchi S, Kusaka T. Purification and characterization of 3-oxoacyl-CoA synthase of Mycobacterium smegmatis. J Biochem. 1983 Oct;94(4):1045-51.
61.   Gurvitz A. The essential mycobacterial genes, fabG1 and fabG4, encode 3-oxoacyl-thioester reductases that are functional in yeast mitochondrial fatty acid synthase type 2. Mol Genet Genomics. 2009 Oct;282(4):407-16.
62.   Dutta D, Bhattacharyya S, Roychowdhury A, Biswas R, Das Amit K. Crystal structure of hexanoyl-CoA bound to β-ketoacyl reductase FabG4 ofMycobacterium tuberculosis. Biochemical Journal. 2013;450(1):127-39.
63.   Gurvitz A, Hiltunen JK, Kastaniotis AJ. Heterologous Expression of Mycobacterial Proteins in Saccharomyces cerevisiae Reveals Two Physiologically Functional 3-Hydroxyacyl-Thioester Dehydratases, HtdX and HtdY, in Addition to HadABC and HtdZ. Journal of Bacteriology. 2009;191(8):2683-90.
64.   Sacco E, Slama N, Backbro K, Parish T, Laval F, Daffe M, et al. Revisiting the Assignment of Rv0241c to Fatty Acid Synthase Type II of Mycobacterium tuberculosis. Journal of Bacteriology. 2010;192(15):4037-44.
65.   Cheruvu M, Plikaytis BB, Shinnick TM. The acid-induced operon Rv3083–Rv3089 is required for growth of Mycobacterium tuberculosis in macrophages. Tuberculosis. 2007;87(1):12-20.
66.   Singh A, Jain S, Gupta S, Das T, Tyagi AK. mymA operon of Mycobacterium tuberculosis: its regulation and importance in the cell envelope. FEMS Microbiol Lett. 2003 Oct 10;227(1):53-63.

67.   Fisher MA, Plikaytis BB, Shinnick TM. Microarray analysis of the Mycobacterium tuberculosis transcriptional response to the acidic conditions found in phagosomes. J Bacteriol. 2002 Jul;184(14):4025-32.

68.   Forrellad MA, Klepp LI, Gioffré A, Sabio y García J, Morbidoni HR, Santangelo MdlP, et al. Virulence factors of theMycobacterium tuberculosiscomplex. Virulence. 2014;4(1):3-66.

69.   Finn RD, Bateman A, Clements J, Coggill P, Eberhardt RY, Eddy SR, et al. Pfam: the protein families database. Nucleic Acids Research. 2014;42(D1):D222-D30.

70.   Parker SK, Curtin KM, Vasil ML. Purification and Characterization of Mycobacterial Phospholipase A: an Activity Associated with Mycobacterial Cutinase. Journal of Bacteriology. 2007;189(11):4153-60.

71.   Bakala N'Goma JC, Schué M, Carrière F, Geerlof A, Canaan S. Evidence for the cytotoxic effects of Mycobacterium tuberculosis phospholipase C towards macrophages. Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids. 2010;1801(12):1305-13.

72.   Raynaud C, Guilhot C, Rauzier J, Bordat Y, Pelicic V, Manganelli R, et al. Phospholipases C are involved in the virulence of Mycobacterium tuberculosis. Molecular Microbiology. 2002;45(1):203-17.

73.   Schaefer CM, Lu R, Nesbitt NM, Schiebel J, Sampson NS, Kisker C. FadA5 a thiolase from Mycobacterium tuberculosis: a steroid-binding pocket reveals the potential for drug development against tuberculosis. Structure. 2015 Jan 6;23(1):21-33.

74.   Nesbitt NM, Yang X, Fontan P, Kolesnikova I, Smith I, Sampson NS, et al. A thiolase of Mycobacterium tuberculosis is required for virulence and production of androstenedione and androstadienedione from cholesterol. Infect Immun. 2010 Jan;78(1):275-82.

75.   Yang X, Dubnau E, Smith I, Sampson NS. Rv1106c from Mycobacterium tuberculosis is a 3beta-hydroxysteroid dehydrogenase. Biochemistry. 2007 Aug 7;46(31):9058-67.

76.   Terao J. Cholesterol Hydroperoxides and Their Degradation Mechanism. 2014;77:83-91.

77.   Briken V, Klink M, Brzezinska M, Szulc I, Brzostek A, Kielbik M, et al. Cholesterol Oxidase Is Indispensable in the Pathogenesis of Mycobacterium tuberculosis. PLoS One. 2013;8(9):e73333.

78.   Heym B, Alzari PM, Honore N, Cole ST. Missense mutations in the catalase-peroxidase gene, katG, are associated with isoniazid resistance in Mycobacterium tuberculosis. Mol Microbiol. 1995 Jan;15(2):235-45.

79.   Pym AS, Saint-Joanis B, Cole ST. Effect of katG mutations on the virulence of Mycobacterium tuberculosis and the implication for transmission in humans. Infect Immun. 2002 Sep;70(9):4955-60.

80.   Ng VH, Cox JS, Sousa AO, MacMicking JD, McKinney JD. Role of KatG catalase-peroxidase in mycobacterial pathogenesis: countering the phagocyte oxidative burst. Mol Microbiol. 2004 Jun;52(5):1291-302.

81.   Njuma OJ, Ndontsa EN, Goodwin DC. Catalase in peroxidase clothing: Interdependent cooperation of two cofactors in the catalytic versatility of KatG. Archives of Biochemistry and Biophysics. 2014;544:27-39.

82.   Daffé M, Draper P. The envelope layers of mycobacteria with reference to their pathogenicity. Adv Microb Physiol. 1998;39:131-203.

83.   Cade CE, Dlouhy AC, Medzihradszky KF, Salas-Castillo SP, Ghiladi RA. Isoniazid-resistance conferring mutations inMycobacterium tuberculosisKatG: Catalase, peroxidase, and INH-NADH adduct formation activities. Protein Science. 2010:NA-NA.

84.   Vilchèze C, Jacobs Jr WR. Resistance to Isoniazid and Ethionamide in Mycobacterium tuberculosis: Genes, Mutations, and Causalities. Microbiology Spectrum. 2014;2(4).

85.   Jena L, Deshmukh S, Waghmare P, Kumar S, Harinath BC. Study of mechanism of interaction of truncated isoniazid–nicotinamide adenine dinucleotide adduct against multiple enzymes of Mycobacterium tuberculosis by a computational approach. International Journal of Mycobacteriology. 2015;4(4):276-83.

86.   Jiang X, Zhang W, Gao F, Huang Y, Lv C, Wang H. Comparison of the proteome of isoniazid-resistant and -susceptible strains of Mycobacterium tuberculosis. Microb Drug Resist. 2006 Winter;12(4):231-8.

87.   Singh A, Gopinath K, Sharma P, Bisht D, Singh N, Singh S. Comparative proteomic analysis of sequential isolates of Mycobacterium tuberculosis from a patient with pulmonary tuberculosis turning from drug sensitive to multidrug resistant. Indian J Med Res. 2015 Jan;141(1):27-45.

88.   Mdluli K, Slayden RA, Zhu Y, Ramaswamy S, Pan X, Mead D, et al. Inhibition of a Mycobacterium tuberculosis beta-ketoacyl ACP synthase by isoniazid. Science. 1998 Jun;280(5369):1607-10.

89.   Biswas R, Dutta D, Das AK. Cloning, overexpression, purification, crystallization and preliminary X-ray diffraction analysis of Rv0241c (HtdX) fromMycobacterium tuberculosisH37Rv. Acta Crystallographica Section F Structural Biology and Crystallization Communications. 2013;69(10):1110-3.

CHAPTER IV – DIFFERENTIAL *IN VIVO* CHARACTERISTICS OF *Mycobacterium tuberculosis* AFTER ACQUISITION OF ISONIAZID RESISTANCE

4.1 Introduction

The *Mtb* catalase-peroxidase KatG is a crucial enzyme for the balance of endogenous and exogenous oxidative stress as well as in the activation of isoniazid (INH), an effective drug used to control tuberculosis (TB) infection (1). These KatG functions have been explored in the first two experimental chapters of this dissertation, while in the present one the role of this enzyme as an important virulence factor is further evaluated. This evaluation is performed in the context of *katG* mutants and with the subsequent INHr phenotype generated after drug exposure. These *Mtb* strains also suffered an alteration in their proteome that support this new phenotype, as it is described in chapters II and III. *katG* mutations that may or may not end with the reduction of the enzyme activity are highly associated with resistance to INH in *Mtb* (2-8). The effect of these mutations on *Mtb* fitness and virulence is variable, for instance, the most frequently found in clinical multidrug resistant-MDR TB cases (Ser315Thr) is associated with minimal to no reduction in fitness (9, 10), and a functional catalase-peroxidase activity (10). Conversely, total deletion and other *katG* mutations result in severely attenuated *Mtb* strains (10-12). In the previous chapter a biochemical description of *katG* mutants suggested a new metabolic rearrangement and a possible alteration in the virulent profile of the INHr strains. This alteration is suggested by seeing the different levels of proteins in the virulence and detoxification category as well as those involved in protein export (Sec and ESX-related proteins). The next step is to evaluate the impact of this new phenotype of the INHr strains *in vivo*. In the present chapter, an exploration of the role of specific *katG*-conferring mutations in the virulence of *Mtb* strains is

further explored looking at the growth of these strains and the in the induced immune response and pathology in the mouse model.

Mouse models for *Mtb* infection has been used since 1880´s by Robert Koch in his initial experiments (13). Since then, this model has been used to understand many aspects of the host-pathogen interaction and insights in the immune response against *Mtb*. The mouse model has many advantages, such as the resemblance of the cell-mediated immunity with human disease, the availability of reagents to measure their immune response, the tropism of *Mtb* for the mouse lung, also the availability of mouse strains with different genetic backgrounds and its economic cost. These advantages allow to understand the role of specific host cells and gene products during the *Mtb* infection (14). Different mouse strains have been characterized in terms of their susceptibility to *Mtb* infection. For instance, the C57BL/6 mouse strain is recognized for being resistant to *Mtb* challenge with a longer survival after infection (compared to other mouse strains such as DBA/2 and C3H/HeJ) (14-16). In this mouse strain, the low infection dose of *Mtb* (10^2 colony forming units (CFU) deposited in the lung) was standardized and has been widely used for *Mtb* infection studies (14, 16, 17). The mouse model of *Mtb* infection using the C57BL/6 strain has been a useful model to study bacterial virulence in a homogeneous genetic background. In this way, differences in the mouse response comparing the infection with different *Mtb* strains would be mostly due to intrinsic bacteria factors and their virulent profile that specifically alters the host-bacteria interaction. For these reasons, mouse models offer clear advantages over other animal models of TB. However, it is worth to mention that the mouse model is showing us a partial picture of what occurs in the *Mtb*-human interaction, as mice do not generally develop all pathology features seen in TB patients (15). With the exception of the C3HeB/FeJ so called Kramnik model (18, 19), the differences in the pathology include the

absence of necrotic granulomas and caseous lesions in the lungs of *Mtb*–infected mice.

Additionally, there are not substantial hypoxic zones in the mouse lung around TB lesions, which

is an environment that *Mtb* regularly encounter in the human granulomas (20). Finally, the

cellular organization, including the distribution of T cells and macrophages in the TB lesions in

the mouse are not similar to the distribution of these cells in the lesions found in humans (21). In

summary, the mouse model of *Mtb* infection is one of the most successful animal models to

understand the bacterium interaction with the natural host (humans), however, the interpretation

of the pathological findings should account for the previously described differences.

   Mtb virulence is associated with both the lack of an effective immune response in the

host and ability of the bacteria to invade, survive and multiply in the host (22). These

characteristics are regularly measured in the mouse model after determining at the CFU in lungs

and other organs as well as the expressed cytokine profile. After *Mtb* interacts with macrophages

and dendritic cells from the human or murine host, cytokines of the T helper type 1 (Th1)

response, such as tumor necrosis factor (TNF)-α and interferon-γ (IFNγ), are crucial for the

control of *Mtb* infection (23). However, a "dynamic balance" between all Th cell subsets is

required to avoid an exacerbated inflammatory response in the lung tissue and to accomplish a

successful bacterial inhibitory response (24, 25). *Mtb*-host interaction and bacterial fitness are

also variable in both drug susceptible and resistant isolates depending on the *Mtb* genetic

background (26, 27). In this chapter, a study of clonal strains of *Mtb* was performed *in vivo* to

evaluate the effect of INH-resistance-conferring mutations in a similar genetic background.

Moreover, the *katG* mutations evaluated here are clinically relevant mutation causing INHr TB

cases. The aim of this chapter is to assess differences in bacterial virulence in INHr *Mtb* due to

mutations in the *katG* gene using a pair of clonal strains from a TB patient and a laboratory derived clonal pair.

4.2 Materials and Methods

4.2.1 Ethics statement

All experiments involving mice were approved by the Colorado State University Research Integrity & Compliance Review Office (RICRO) and the Institutional Animal Care and Use Committee, protocol #13-4509A. All experiments involving mice were approved by Colorado State University's Institutional Animal Care and Use Committee (IACUC protocol #13-4509A). Samples obtained from the human subject were obtained from a collection of a TB patient enrolled in a previous study of TB transmission in San Francisco following approval from Stanford University and the University of California San Francisco's Institutional Review Boards (9, 10).

4.2.2 Bacterial strains

The strains used for this study are the same clonal laboratory and clinical pairs used in the chapter III. Both clonal pairs of *Mtb* were simultaneously cultured in Proskauer-Beck (PB) media to prepare the infectivity stocks as previously described (11). The preparation of the infectivity stocks involved the culture of those cells that growth only on the surface of the liquid media, upscaling each time a pellicle formation is reached. The time for pellicle formation varied from three to eight weeks. The cultures should be kept at 37°C without any disturbance. In this way, the bacteria pellicle was up-scaled from 10mL, 25 mL, to 100 ml of PB media. The resulting bacteria pellicle were carefully taken and transferred to 20 ml PB media with 20% Glycerol. After this, bacterial clumps were disaggregated using the bath sonicator three times during 10 seconds each. Then, the bacteria were settled down for 15 minutes on ice and only the soluble

bacteria were stored in volumes of one milliliter at -80°C until its use for the mouse infection study. Interestingly, as it was observed in the GAS media growth, the clinical INHr took longer times to create pellicles that were at the same time less dense compared to its susceptible pair and any other strain of the study.

4.2.3 Mouse infection

Groups of C57BL/6 female mice (n = 5/group), 6-8 weeks old (The Jackson Laboratory, Bar Harbor, ME), were infected with the clinical and laboratory clonal pairs of *Mtb* with approximately 50-100 colony forming units (CFU) of each strain via the aerosol route using the Middlebrook Aerosol Exposure Chamber (Glas-Col, Terre Haute, IN). This route of infection was selected as it is the most relevant looking at the most common route that humans get in contact with the bacterium. At days 0, 7, 14, 21, 28, 60, and 120 post-infection (p.i.), 5 mice per group were humanely euthanized for CFU counting in the lung and spleen, and for cytokine and histopathology analysis of the lung.

4.2.4 CFU determination in lung and spleen

Lung and spleen from infected mice were excised, homogenized in sterile physiological saline (0.85 % NaCl) and 10-fold serial dilutions prepared. With the exception of day zero, dilutions were plated on 7H11 agar plates and incubated for 21 days at 37°C until the enumeration of colonies (CFU). For day zero, whole lung preparations were plated. Data were expressed as Log$_{10}$ CFU in the respective organ.

4.2.5 Cytokine analysis

The determination of the cytokines concentration in lung homogenates was performed using the Cytometric Bead Array (CBA, BD Biosciences) assay. In this assay, five bead populations coated with capture antibodies specific for Interleukin (IL) -2, IL-6, IL-10,

Interferon (IFN)-γ and Tumor Necrosis Factor (TNF)-α were mixed with the lung homogenate samples. The homogenates were previously clarified by centrifugation (1200 x g during 10 minutes) to remove cellular debris and incubated with a mix of the previously described beads. Then, samples were analyzed in the flow cytometer FACSCanto II (BD Biosciences). The final concentration of each cytokine was obtained from the indiviual standard curve using mouse recombinant cytokine proteins (from 0 to 5000 pg/mL) that define the minimum and maximum quantifiable levels. The limit of detection of each cytokine is stated at each result figure.

4.2.6 Lung histology

During necropsy, the inferior lung lobe was inflated with 10% formalin/PBS and placed into 10% formalin/PBS. Following routine processing for paraffin-embedding, 5μm sections were stained with hematoxylin and eosin (H&E) and assessed by a veterinary pathologist blinded to the experimental groupings for degree of lung involvement using a defined scoring system adapted from Dormans *et al.* (28). The score was based on the total number of lesions, overall extent of changes, presence of peribronchiolitis, perivasculitis, alveolitis, "granuloma" formation and degree of necrosis. The scale ranges from zero (no apparent changes) to five (severe changes) (Figure 4.1). Microphotographs were captured using a Nikon Eclipse 51E microscope equipped with a Nikon DS-Fi1 camera with a DS-U2 unit and NIS elements software. Microphotographs are reproduced without manipulation other than cropping and adjustment of light intensity.
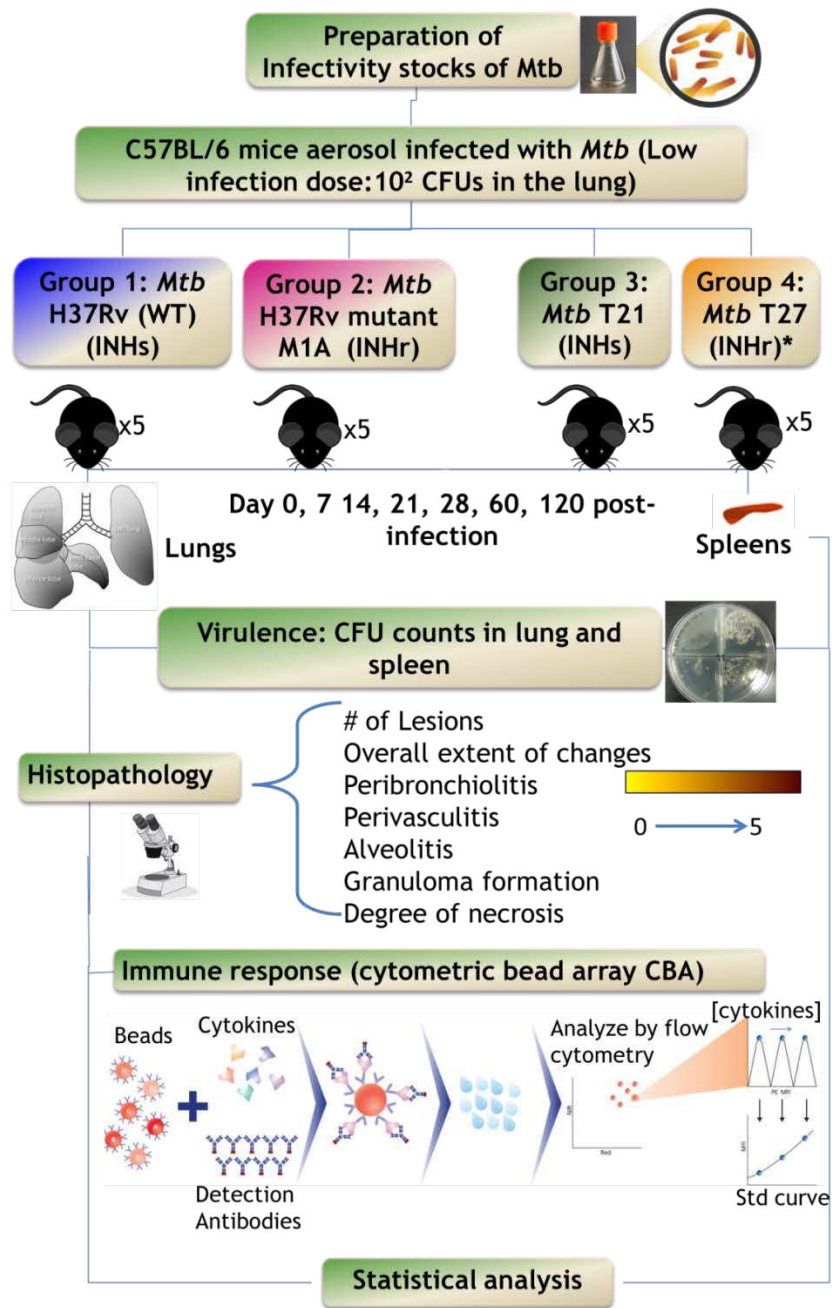
Figure 4.1. Schematic representation of the methodology followed for the virulence study of the clonal *Mtb* pairs after acquisition of INH resistance. CFUs: Colony forming units.

4.2.7 Statistical analysis

Statistical differences for Log_{10} CFU counts and cytokine level comparisons were evaluated using an unpaired t-test or Mann-Whitney test depending on whether or not the data set had a normal distribution (defined by Kolmogorov-Smirnov test). Differences in lesion scores were analyzed using Kruskal-Wallis test and Dunn's post-test. All statistical analyses were done in Graphpad Prism® version 6.05.

4.3 Results and Discussion

4.3.1 In vivo growth of Mtb *demonstrated reduced bacteria fitness in INHr strains*

The effect of the reduced KatG levels and enzymatic activity in the INHr strains was explored by determining differences in bacterial virulence compared to their correspondent INHs clonal pairs. Lower CFU counts were observed in lung and spleen mouse tissues of both clinical and laboratory INHr strains for almost all time points measured. This difference was significant in both lung and spleen tissues of the clinical pair, and it was further magnified by the minimal growth of the clinical INHr strain in the spleen (Figure 4.2). It is important to highlight that the bacterial growth of each strain was evaluated in BACTEC media using the MGIT™320 prior to the mouse infection. This confirmed the viability of all the strains and lack of significant differences in the bacterial growth and viability among them at the time of infection (Figure 4.3).

The decreased bacterial fitness observed in the INHr strains, seen as a CFU reduction, has been demonstrated in the mycobacterial complex as an effect of some *katG* mutations using isogenic strains of *Mycobacterium bovis* and *Mtb* in the guinea pig and mouse models (29-31). These studies were performed with a laboratory strain of *Mtb* with total deletion of the *katG*
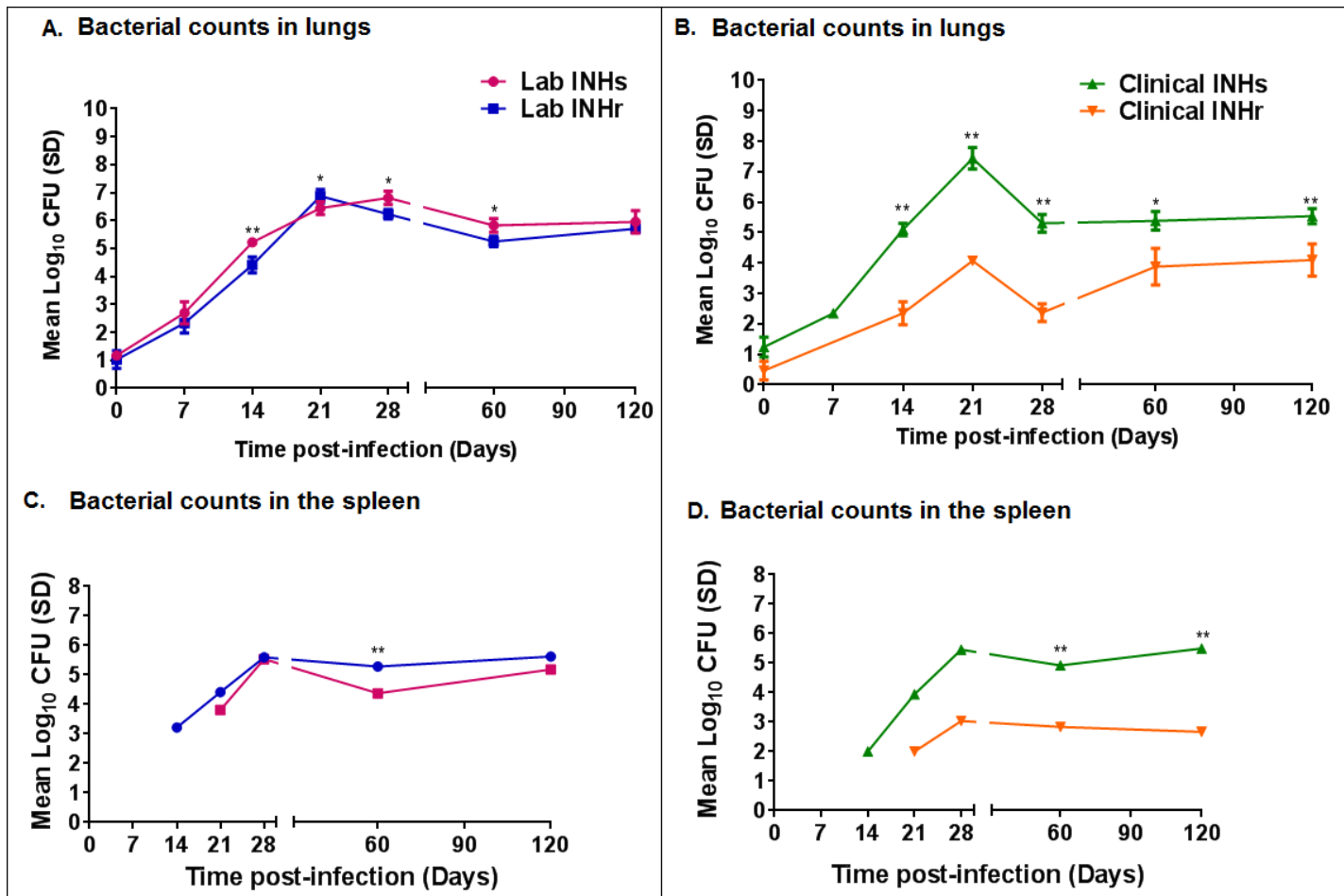
Figure 4.2. Comparison of bacterial growth in C57BL/6 mice after low dose aerosol infection. A) and C). Lung CFU count of mice infected with laboratory and clinical clonal strain pairs respectively. B) and D). Spleen CFU counts of laboratory and clinical clonal strain pairs respectively. Non-paired *t*-test without assuming consistent standard deviation (SD).*p<0.05, **p<0.001.
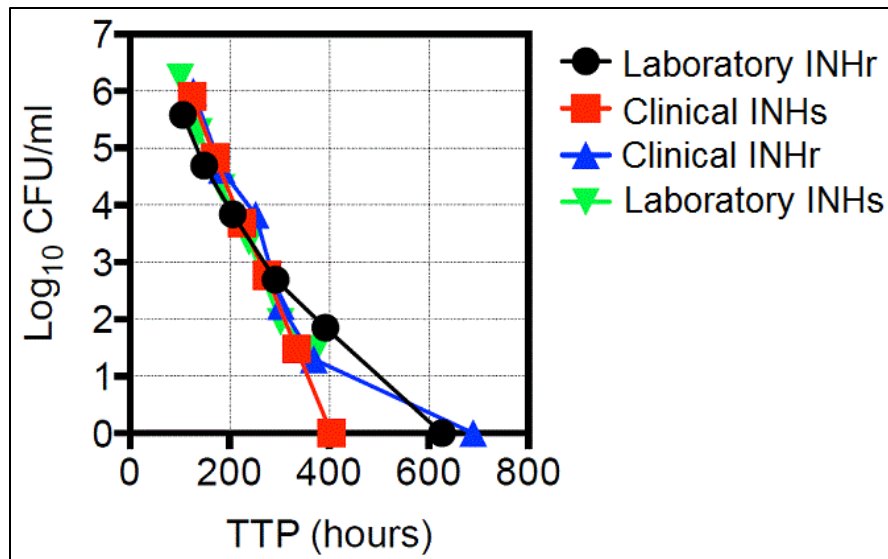
Figure 4.3. Comparison of bacterial growth rates in BACTEC media using MGIT™320. Time to positivity (TTP) of each strain was determined by seeding BACTEC tubes with10-fold serial dilutions. TTP assessed in the MGIT™320.

gene, a mutation rarely found in clinical settings (32). In addition to this, there have been supportive reports describing differences in CFU in mice infected with INHr clinical strains (12).

4.3.2 Differences in lung pathology at chronic stage of infection

To further evaluate how virulence was affected in the INHr strains, we compared the lesion in the lungs of the mice infected with the *Mtb* strains of our study. There were no statistical differences in the lung pathology scores at all the evaluated time points for the laboratory pair comparison. Pathologic findings in animals infected with both laboratory INHs and INHr strains included peribronchiolitis, perivasculitis and alveolitis from day 21 p.i. onwards, with marked granulomatous inflammation from day 60 p.i. onwards. In contrast, the lung pathology scores of the mice infected with the clinical INHr strain at days 21, 28 and 120 p.i. were significantly lower than scores for the clinical INHs strain (Figure 4.4). Mice infected with the clinical INHs strain also showed signs of inflammation from day 21 p.i., while for the clinical INHr infected animals, inflammatory lesions were rarely observed until day 60 p.i.
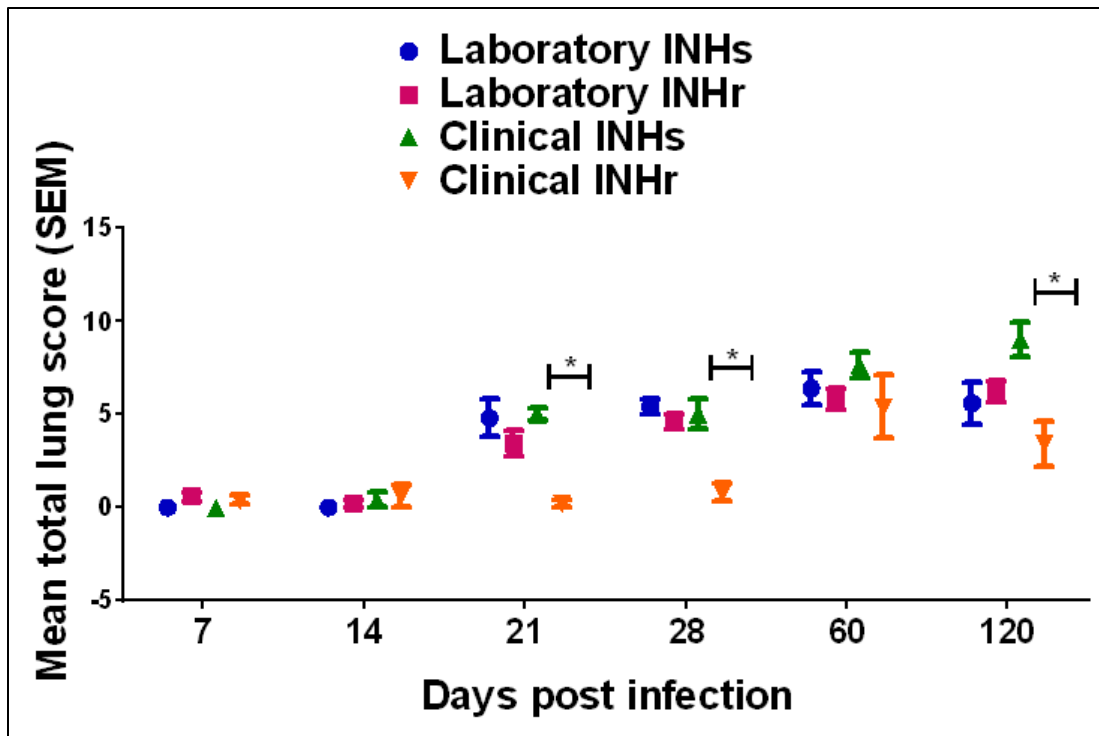
Figure 4.4. Lung pathology scores for C57BL/6 mice infected with clinical and laboratory clonal strains with different INH susceptibility profiles. Mean of the lung scores with standard error (SEM) for infected mice with each strain differentiated by color. *$p<0.05$ with Kruskal-Wallis test and Dunn's post-test.

A prior study using the clinical strains (labeled as T21 and T27) suggested no significant differences in lung pathology during the first 90 days of infection between the INHr and INHs strains (33). In our study, we were able to fully evaluate the pathology of these strains with a detailed scoring system; from this, we built upon the previous assessment and observed clear differences in the lung pathology during the course of infection, with the greatest difference observed at the latest time point (day 120 post-infection) (Figure 4.4 and 4.5).

4.3.3 Lower induction of proinflammatory cytokines after infection with INHr Mtb

To determine the immune response generated against these less virulent INHr strains, we performed an evaluation of cytokines in the mouse lung tissue. The dynamics of IFN-γ, TNF-α, and IL-6 levels showed a reduction of these proinflammatory cytokines in both INHr strains compared to their INHs pair throughout the study (Figure 4.6A). Specifically, in the clinical

INHr strain, the difference was statistically significant for the three cytokines from day 21 p.i. onward. Similar results were observed also in the laboratory INHr strain, but the difference was significantly different only at one or two time points, depending on the cytokine evaluated (Figure 4.6B). In addition, levels of IL-2 were very close for the majority of the time points in the laboratory pair. However, the levels of this cytokine in the INHr clinical strain were almost undetected On the other hand, although not significant, there was a slight increase of IL-10 in the INHr strains at the initial time points (day 7 and 14 p.i. for the laboratory and day 21 p.i for the clinical strain). The levels of this regulatory cytokine (IL-10) were very similar among both clinical and laboratory pairs after day 28 p.i. (Figure 4.6C).

The strongly reduced levels of the proinflammatory cytokines together with the slightly augmented levels of the anti-inflammatory cytokine IL-10 in mice infected with the clinical INHr strain (Figure 4.6) are in line with the deficient granulomatous reaction with lower cell infiltration observed in the lungs of these mice (Figure 4.5).

At day 21 and day 120 p.i., IL-2 levels in the laboratory clonal pair showed a contrasting trend when compared with IFN-γ levels (Figure 4.6A and C). Normally, the increased *Mtb* specific antigen levels are related to increased levels of IFN-γ while frequencies of antigen-specific IL-2-secreting T cells increased during *Mtb* clearance (34). Also, IL-2 produced by CD4+ T cells has been implied in T cell central memory and proposed as a marker for clinical monitoring to detect specific CD4+ *Mtb*-specific T-cell signatures in humans (34, 35). The cytokine levels induced by the laboratory pair from our study suggest that bacterial antigen secretion in the INHr was reduced over time together with a mild inflammatory response. However, this was true only for two time points, and only for the laboratory pair.
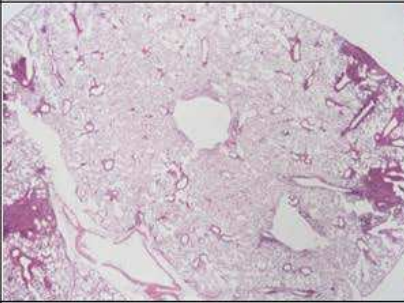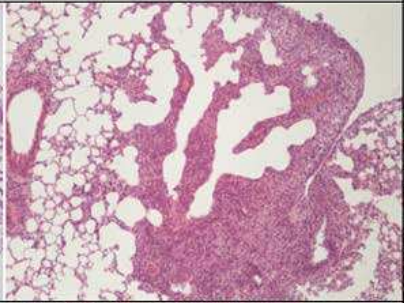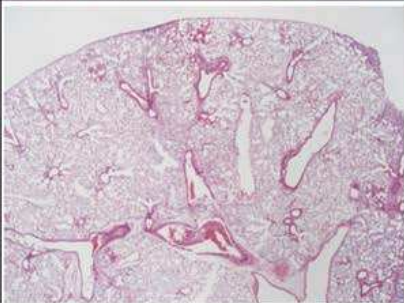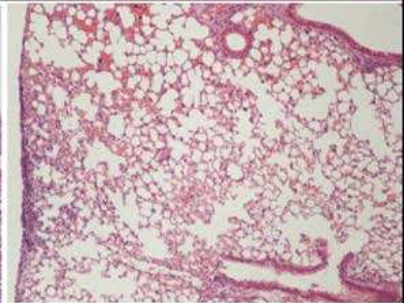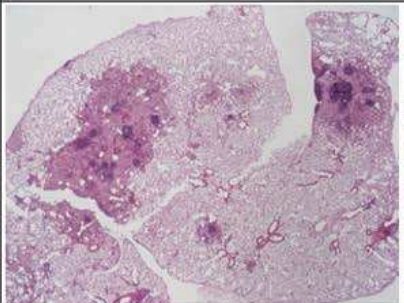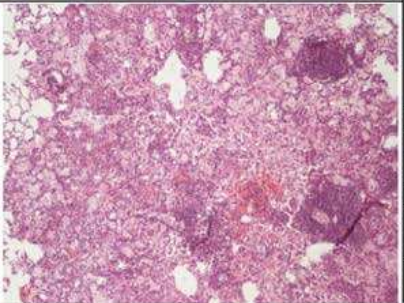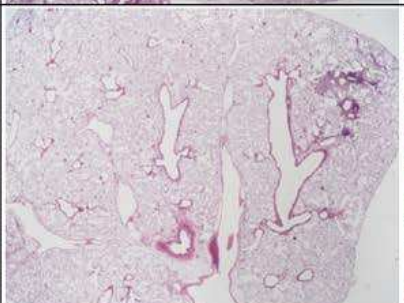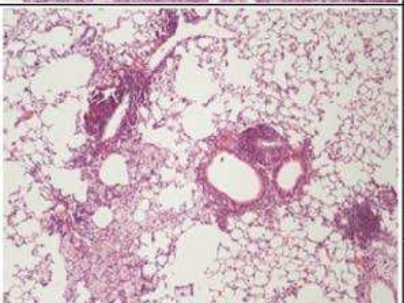
Mtb strain	Day 28 post infection		Total lung score
Clinical INHs			**7** -Number of lesions: 10 -Extent of change: ~10% -Peribronchiolitis: 2 -Perivasculitis: 2 -Alveolitis: 3 -"Granuloma"formation: 0 -Degree of necrosis: 0
Clinical INHr			**2** -Number of lesions: 2 -Extent of change: ~1 % -Peribronchiolitis: 0 -Perivasculitis: 1 -Alveolitis: 1 -"Granuloma"formation: 0 -Degree of necrosis: 0
	Day 120 post infection		
Clinical INHs			**11** -Number of lesions: ~13 -Extent of change: ~30 % -Peribronchiolitis: 1 -Perivasculitis: 3 -Alveolitis: 2 -"Granuloma"formation: 3 -Degree of necrosis: 2
Clinical INHr			**5** -Number of lesions: 2 -Extent of change: ~2 % -Peribronchiolitis: 1 -Perivasculitis: 2 -Alveolitis: 2 -"Granuloma"formation: 0 -Degree of necrosis: 0

Figure 4.5. Histopathology for C57BL/6 mouse lungs infected with clinical strains with significant differences. Four examples of the spectrum of lesions and corresponding scores are depicted. H&E stained sections at 20x and 100x original magnifications.
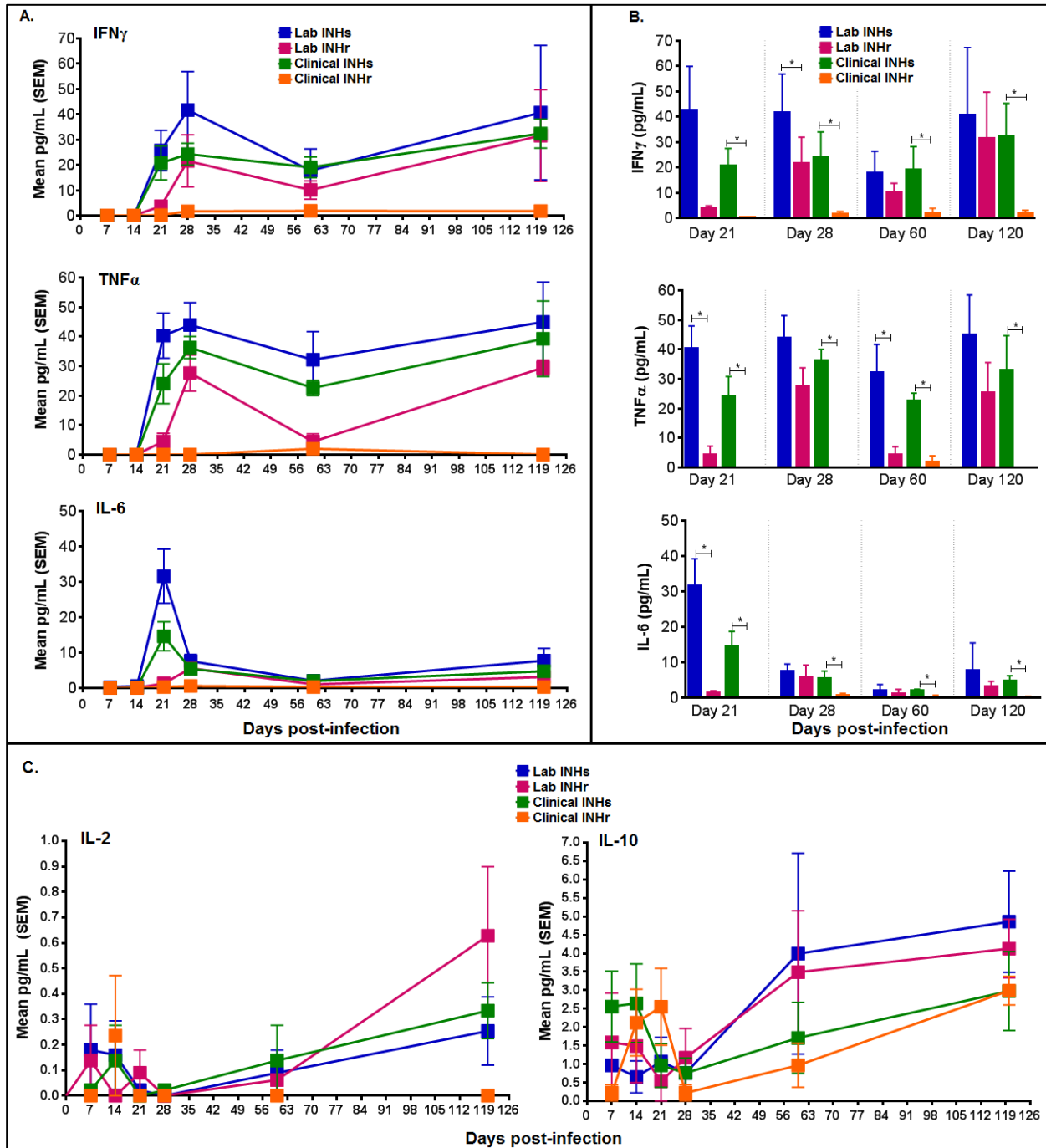
Figure 4.6. Dynamics and comparison of cytokine levels in mouse lung homogenates. A. Pro-inflammatory cytokines levels throughout the infection. Limit of detection (LD) for IFNγ: 0.5pg/mL, TNFα: 0.9 pg/mL, and IL-6: 1.4 pg/mL. B) Pair comparison between mouse infected with clinical and laboratory clonal strains using t-test, *p<0.05. C. Levels of IL-2 (LD: 0.1 pg/mL) and IL-10 (LD: 16.8 pg/mL) in mouse infected with laboratory and clinical pairs.

4.4 Conclusions

Our results concurred with previous studies that have demonstrated a reduction of bacterial virulence due to specific *katG* mutations (12, 13). This was evidenced by a significant reduction of CFU counts due to single or double clinically relevant mutations at the N-terminus of *katG* gene. Evaluation of the clinical INHr strain, we also confirmed the cumulative negative effect of the *katG* N-terminus mutations in bacterial fitness and adaptation to the host. In fact, the low frequency of these *katG* mutations in clinical settings already suggests a lower fitness for these INHr strains (11). It is important to emphasize that not every SNP combination in *katG* can cause a cumulative effect in bacterial fitness and hence virulence. For instance, the Arg463Leu mutation in *katG* is not related with either INH resistance or virulence. This particular mutation instead has been used mostly for genotyping purposes (9).

Mice infected with the clinical INHr strain still controlled the bacterial load and exhibited decreased pathology, despite the apparent lack of macrophage activation (lower induction of proinflammatory cytokines). Multiple hypothesis are possible: first, the reduced levels of the catalase-peroxidase (less bacterial virulence factor), which allowed the faster elimination of the infecting bacteria. Secondly, the very controlled (or minimal) induction of a Th1 response was enough to clear the bacteria from the body, which probably generated less tissue damage. Since macrophage activation and phagocytosis are key features to control *Mtb* growth during infection (23), other innate immune factors and cells such as invariant natural killer T cells, γδT cells, mucosa-associated invariant T (MAIT) cells (36, 37) or TLR-2 associated mechanisms could be playing an important role in the clearance of these less virulent INHr strains.

The significant reduction in virulence and pathology observed particularly in mice infected with the INHr clinical strain suggests that this INHr strain may experience additional

physiological rearrangements in addition to those caused by the double *katG* mutations (V1A and E3V). In fact, KatG levels alone were not enough to explain the extremely poor *in vivo* growth of the clinical INHr, since the laboratory INHr strain with also reduced levels of protein and peroxidase activity did not exhibit such a strong difference (Figure 4.3 and 3.12 from chapter III). Moreover, the clinical INHr strain unlike to the laboratory INHr strain was also exposed to other first line anti-TB drugs when INHr was developed within the patient. Further physiological rearrangements could occur in the clinical INHr strain that may include a variation in its metabolic profile which can be explored in a more comprehensive biochemical analysis. These changes could possibly alter the regular activation of macrophages via Toll like receptor (TLR), such as TLR-2 (23) and possibly other pathways involved in early anti-TB innate immune response that make the bacteria clearance more efficient.

The previous suggestion is supported by some of the proteomics findings in chapter III. Among the protein changes, the high levels of the protein cholesterol oxidase (ChoD) in the clinical INHr strain only, possibly alter the regular activation of macrophages via TLR-2, that normally result in the inflammatory responses of innate immune cells (23, 38). ChoD, particularly, is hypothesized to induce the production of IL-10 via TLR-2, suppressing the bactericidal activity of macrophages without altering other cytokines (38). As it was demonstrated in figure 4.6, IL-10 was the only cytokine with comparable levels between the INHs and INHr clonal pair. Contributing to that phenotype, the protein LprA was increased in the INHr strains (although only significant in the laboratory comparison, chapter III Figure 3.11). This protein is a TLR-2 agonist that is proposed to modulate also the innate immune response by stimulating dendritic cell maturation and the antigen presenting function, as well as decreasing IFN-γ-induced by the major histocompatibility complex II (39). In general most lipoproteins

increased in the INHr strains could be also contributing to the TLR-2 associated responses, giving strong support to the important role of the innate immunity in the control of these INHr strains. Additionally, the increased levels of the alkyl peroxidase AhpC in *katG* mutants previously proposed as a compensatory event in *Mtb* strains that lack of KatG activity, only occurred in the laboratory INHr strain. In fact in the clinical INHr strain, AphC was strongly reduced (Chapter III, Table 3.1). Together, this proteomic trend could explain the differences in the degree of virulence reduction observed in both INHr strains.

In summary, the evaluation of virulence and fitness in the INHr-*katG* mutant *Mtb* strains in the mouse model supports the idea of the innate response in the clearance of the INHr strains by the host including specific details in the early interaction of the bacterium with the host. Here, the results obtained in the characterization of the host response, confirm the KatG role during *in vivo* infection and confirm the variable virulent phenotypes depending on the specific *katG* mutation, even if the mutations affect the same domain of the protein (11, 12, 40). Combining this findings with the previous proteomic study findings (Chapter III), this mouse study suggest a possible role of TLR-2 in the early clearance of the INHr strains associated with the abundance of some particular *Mtb* proteins (ChoD and lipoproteins such as LprA) that are agonist for this receptor.

REFERENCES

1.   Trivedi A, Singh N, Bhat SA, Gupta P, Kumar A. Redox Biology of Tuberculosis Pathogenesis. 2012;60:263-324.

2.   Cade CE, Dlouhy AC, Medzihradszky KF, Salas-Castillo SP, Ghiladi RA. Isoniazid-resistance conferring mutations inMycobacterium tuberculosisKatG: Catalase, peroxidase, and INH-NADH adduct formation activities. Protein Science. 2010:NA-NA.

3.   Ando H, Kondo Y, Suetake T, Toyota E, Kato S, Mori T, et al. Identification of katG Mutations Associated with High-Level Isoniazid Resistance in Mycobacterium tuberculosis. Antimicrobial Agents and Chemotherapy. 2010;54(5):1793-9.

4.   Barry CE, Slayden RA, Mdluli K. Mechanisms of isoniazid resistance in Mycobacterium tuberculosis. Drug Resist Updat. 1998;1(2):128-34.

5.   Cade CE, Dlouhy AC, Medzihradszky KF, Salas-Castillo SP, Ghiladi RA. Isoniazid-resistance conferring mutations in Mycobacterium tuberculosis KatG: catalase, peroxidase, and INH-NADH adduct formation activities. Protein Sci. 2010 Mar;19(3):458-74.

6.   Zhang Y, Heym B, Allen B, Young D, Cole S. The catalase-peroxidase gene and isoniazid resistance of Mycobacterium tuberculosis. Nature. 1992 Aug;358(6387):591-3.

7.   Ramaswamy SV, Reich R, Dou SJ, Jasperse L, Pan X, Wanger A, et al. Single nucleotide polymorphisms in genes associated with isoniazid resistance in Mycobacterium tuberculosis. Antimicrob Agents Chemother. 2003 Apr;47(4):1241-50.

8.   Sandy J, Mushtaq A, Kawamura A, Sinclair J, Sim E, Noble M. The structure of arylamine N-acetyltransferase from Mycobacterium smegmatis--an enzyme which inactivates the anti-tubercular drug, isoniazid. J Mol Biol. 2002 May 10;318(4):1071-83.

9.   O'Sullivan DM, McHugh TD, Gillespie SH. Mapping the fitness of Mycobacterium tuberculosis strains: a complex picture. J Med Microbiol. 2010 Dec;59(Pt 12):1533-5.

10.   Pym AS, Saint-Joanis B, Cole ST. Effect of katG mutations on the virulence of Mycobacterium tuberculosis and the implication for transmission in humans. Infect Immun. 2002 Sep;70(9):4955-60.

11.   Gagneux S, Burgos MV, DeRiemer K, Encisco A, Munoz S, Hopewell PC, et al. Impact of bacterial genetics on the transmission of isoniazid-resistant Mycobacterium tuberculosis. PLoS Pathog. 2006 Jun;2(6):e61.

12.   Ordway DJ, Sonnenberg MG, Donahue SA, Belisle JT, Orme IM. Drug-resistant strains of Mycobacterium tuberculosis exhibit a range of virulence for mice. Infect Immun. 1995 Feb;63(2):741-3.

13.   Collins FM. The immunology of tuberculosis. Am Rev Respir Dis. 1982 Mar;125(3 Pt 2):42-9.

14.   Orme IM. The mouse as a useful model of tuberculosis. Tuberculosis. 2003;83(1-3):112-5.

15.   Gupta UD, Katoch VM. Animal models of tuberculosis. Tuberculosis. 2005;85(5-6):277-93.

16.   Medina E, North RJ. Resistance ranking of some common inbred mouse strains to Mycobacterium tuberculosis and relationship to major histocompatibility complex haplotype and Nramp1 genotype. Immunology. 1998 Feb;93(2):270-4.

17.   Ordway DJ, Orme IM. Animal models of mycobacteria infection. Curr Protoc Immunol. 2011 Aug;Chapter 19:Unit19 5.

18.   Irwin SM, Driver E, Lyon E, Schrupp C, Ryan G, Gonzalez-Juarrero M, et al. Presence of multiple lesion types with vastly different microenvironments in C3HeB/FeJ mice following aerosol infection with Mycobacterium tuberculosis. Dis Model Mech. 2015 Jun;8(6):591-602.

19.   Driver ER, Ryan GJ, Hoff DR, Irwin SM, Basaraba RJ, Kramnik I, et al. Evaluation of a mouse model of necrotic granuloma formation using C3HeB/FeJ mice for testing of drugs against Mycobacterium tuberculosis. Antimicrob Agents Chemother. 2012 Jun;56(6):3181-95.

20.   Tsai MC, Chakravarty S, Zhu G, Xu J, Tanaka K, Koch C, et al. Characterization of the tuberculous granuloma in murine and human lungs: cellular composition and relative tissue oxygen tension. Cell Microbiol. 2006 Feb;8(2):218-32.

21.   Apt A, Kramnik I. Man and mouse TB: Contradictions and solutions. Tuberculosis. 2009;89(3):195-8.

22.   Hanekom M, Gey van Pittius NC, McEvoy C, Victor TC, Van Helden PD, Warren RM. Mycobacterium tuberculosis Beijing genotype: a template for success. Tuberculosis (Edinb). 2011 Nov;91(6):510-23.

23.   Flynn JL, Chan J. Immunology of tuberculosis. Annu Rev Immunol. 2001;19:93-129.

24.   Dwivedi VP, Bhattacharya D, Chatterjee S, Prasad DV, Chattopadhyay D, Van Kaer L, et al. Mycobacterium tuberculosis directs T helper 2 cell differentiation by inducing interleukin-1beta production in dendritic cells. J Biol Chem. 2012 Sep 28;287(40):33656-63.

25.   Dheda K, Booth H, Huggett JF, Johnson MA, Zumla A, Rook GA. Lung remodeling in pulmonary tuberculosis. J Infect Dis. 2005 Oct 1;192(7):1201-9.

26.   Gagneux S, Long CD, Small PM, Van T, Schoolnik GK, Bohannan BJ. The competitive cost of antibiotic resistance in Mycobacterium tuberculosis. Science. 2006 Jun 30;312(5782):1944-6.

27.   Toungoussova OS, Caugant DA, Sandven P, Mariandyshev AO, Bjune G. Impact of drug resistance on fitness of Mycobacterium tuberculosis strains of the W-Beijing genotype. FEMS Immunol Med Microbiol. 2004 Nov 1;42(3):281-90.

28.   Dormans J, Burger M, Aguilar D, Hernandez-Pando R, Kremer K, Roholl P, et al. Correlation of virulence, lung pathology, bacterial load and delayed type hypersensitivity responses after infection with different Mycobacterium tuberculosis genotypes in a BALB/c mouse model. Clin Exp Immunol. 2004 Sep;137(3):460-8.

29.   Wilson TM, de Lisle GW, Collins DM. Effect of inhA and katG on isoniazid resistance and virulence of Mycobacterium bovis. Mol Microbiol. 1995 Mar;15(6):1009-15.

30.   Ng VH, Cox JS, Sousa AO, MacMicking JD, McKinney JD. Role of KatG catalase-peroxidase in mycobacterial pathogenesis: countering the phagocyte oxidative burst. Mol Microbiol. 2004 Jun;52(5):1291-302.

31.   Li Z, Kelley C, Collins F, Rouse D, Morris S. Expression of katG in Mycobacterium tuberculosis is associated with its growth and persistence in mice and guinea pigs. J Infect Dis. 1998 Apr;177(4):1030-5.

32.   Pym AS, Domenech P, Honoré N, Song J, Deretic V, Cole ST. Regulation of catalase-peroxidase (KatG) expression, isoniazid sensitivity and virulence by furA of Mycobacterium tuberculosis. Mol Microbiol. 2001 May;40(4):879-89.

33.   Caraway M. The evaluation of clinical isolates of Mycobacterium tuberculosis in the murine and guinea pig infection models: M.Sc. Thesis, Colorado State University; 2011. Available: http://dspace.library.colostate.edu/webclient/DeliveryManager/digitool_items/csu01_storage/2013/08/22/file_1/212343.

34.   Millington KA, Innes JA, Hackforth S, Hinks TS, Deeks JJ, Dosanjh DP, et al. Dynamic relationship between IFN-gamma and IL-2 profile of Mycobacterium tuberculosis-specific T cells and antigen load. J Immunol. 2007 Apr 15;178(8):5217-26.

35.   Suter-Riniker F, Berger A, Mayor D, Bittel P, Iseli P, Bodmer T. Clinical significance of interleukin-2/gamma interferon ratios in Mycobacterium tuberculosis-specific T-cell signatures. Clin Vaccine Immunol. 2011 Aug;18(8):1395-6.

36.   Orme IM, Robinson RT, Cooper AM. The balance between protective and pathogenic immune responses in the TB-infected lung. Nature Immunology. 2014;16(1):57-63.

37.   Sakala IG, Kjer-Nielsen L, Eickhoff CS, Wang X, Blazevic A, Liu L, et al. Functional Heterogeneity and Antimycobacterial Effects of Mouse Mucosal-Associated Invariant T Cells Specific for Riboflavin Metabolites. J Immunol. 2015 Jul 15;195(2):587-601.

38.   Briken V, Klink M, Brzezinska M, Szulc I, Brzostek A, Kielbik M, et al. Cholesterol Oxidase Is Indispensable in the Pathogenesis of Mycobacterium tuberculosis. PLoS One. 2013;8(9):e73333.

39.   Pecora ND, Gehring AJ, Canaday DH, Boom WH, Harding CV. Mycobacterium tuberculosis LprA is a lipoprotein agonist of TLR2 that regulates innate immunity and APC function. J Immunol. 2006 Jul 1;177(1):422-9.

40.   Pym AS, Saint-Joanis B, Cole ST. Effect of katG Mutations on the Virulence of Mycobacterium tuberculosis and the Implication for Transmission in Humans. Infection and Immunity. 2002;70(9):4955-60.

CHAPTER V - CONCLUDING REMARKS AND PERSPECTIVES

"When you are studying any matter, or considering any philosophy, ask yourself only what are the facts and what is the truth that the facts bear out. Never let yourself be diverted either by what you wish to believe, or by what you think would have beneficent social effects if it were believed. But look only, and solely, at what are the facts."

Bertrand Russell (1872–1970)
Philosopher, logician, mathematician, historian, and social critic

In microbiology research as well as in other research fields, presenting the data as the main characters in the story is highly relevant and allows the generation of new hypotheses and evolution of the field. The findings presented in this work include comprehensive biochemical and immunological characterization of clonal *Mycobacterium tuberculosis (Mtb)* pairs to gain deeper insight regarding the relationship between acquisition of isoniazid (INH) resistance in *Mtb* and the interaction of resistant strains with the host. Besides reviewing concepts of *Mtb* physiology and virulence factors in the first chapter and in the discussion of the results obtained in chapters II to IV, this work also highlights the individual nature of *Mtb* to adapt to new biological events such as the acquisition of drug resistance. In microbiology, there are some general statements in the field that should be always be mentioned in the context of the specific species, genetic lineage and sometimes even at the strain level. At the same time, the interpretation of different events in microbiology should always consider external factors, such as nutrient supply, oxygen tension*,* drug exposure, *in vitro* verus *in vivo* growth, etc., which are crucial variables for the analysis design and the interpretation of the results. There are universal statements in microbiology that have set up its foundation and allowed the evolution of the field, for example, differences between Gram positive from Gram negative bacteria, the absence of

mitochondria, and the expression of the life in sequential reactions from DNA to RNA to protein (1). But central paradigms in every discipline have caveats that were originated by the human interest to research and explore its surroundings to understand better these relevant life-related events.

By the use of clonal pairs of *Mtb*, this work was able to demonstrate some already known features about INH resistance in *Mtb* such as the implication of *katG* mutations and the reduced enzymatic activity (specifically peroxidase) in the attenuated phenotype during in vivo growth (2-9). It is important to clarify that the protein abundance variation in the INH resistant (INHr) strains presented here are relative to the protein levels in the parental INH susceptible (INHs) strain. In the laboratory and clinical strains comparison, similar trends were observed in protein abundance and bacterial attenuation after the INH resistance event (Chapters III and IV). These common trends reveal similar affected pathways, such as the energy metabolism, redox stress-related proteins, fatty acid oxidation and regulatory proteins. Most of these altered proteins have in common the participation of NADH or NADPH cofactors in their reactions. One of the most significantly reduced enzymes in both INHr strains, KatG, uses the NADH cofactor in its reactions (10). Additionally, a previous metabolomics study showed severely reduced levels of NADPH in *Mtb* after exposure to INH (11). Therefore, it is possible that simultaneous to the KatG alteration, metabolism and/or levels of these nucleotides (NADH and NADPH) are also altered in the INHr strains that could be related to the proteomic trend observed in the INHr strains of this dissertation research. Computational models have also suggested that the INH-NAD$^+$ adduct (the active form of INH) not only have affinity for the protein of the FAS II pathway, InhA (which had significant reduced levels both INHr strains) (12), but also for other proteins that were increased or decreased in both clinical and laboratory INHr strains of this

174

dissertation (See Chapter III). Additionally, commonly affected individual proteins with unknown function were also observed, such as the increased levels of the probably exported protein Rv1910c (Chapter III, Table 3.3); which is upstream the *furA-katG* transcription unit (13). According to this result, one could hypothesize that the increased levels of Rv1910c have a direct relation with the reduced levels of KatG or with the INHr phenotype. The exploration of this hypothesis aids to explore the role of Rv1910c in *Mtb* that has not been identified yet.

The higher differences in the bacterial proteome occurred in the clinical INHr strain that is also related with the 14 non-synonymous single nucleotide polymorphisms (SNPs) found in a previous whole genome study of this clonal pair (14). In chapter II, common differences could not be highly observed comparing both clinical INHr Beijing and T genotype pairs, besides the lower KatG levels. Further exploration of the lack of common protein findings could not be possible because of the instability of the *katG* mutation in the Beijing INHr pair that resulted in a reversion of the INHr phenotype in this clonal strain. The reversion of the INHr phenotype in *Mtb* has been previously reported in other *katG* mutant strain (15). When the laboratory clonal pair was added to the analysis, it was possible to focus on the implication of the mutation affecting the KatG N-terminus and look for common trends. The *katG* mutation V1A was present in the laboratory derived INHr strain of H37Rv. This particular mutation was found previously in a multidrug resistant (MDR) strain in Russia in 1995 (16). The clinical INHr strain had also the V1A mutation and the second mutation E3V (discussed in Chapter II). This strain was isolated from a patient that was eventually cured (personal communication). Looking at the results of the mouse study, it could be interpreted that the double *katG* mutation in the clinical INHr strain was a successful combination for bacterial attenuation that could be due to the significantly reduced levels of the catalase-peroxidase together with the protein variation

described in chapter III. The laboratory strain with one single *katG* mutation, on the other hand, had a discrete reduction in virulence with lower ability to induce proinflammatory cytokines in the infected mice and without significant differences in bacterial counts the in the lung histopathology (Figure 4.2, 4.4-4.6). Therefore, it is more likely that the single V1A mutation in *katG* evolve to a MDR phenotype, compared to the double V1A/E3V *katG* mutant. This was demonstrated in the clinical setting with the already described MDR-TB case that had the V1A mutation in *katG* gene. Conversely, the double *katG* mutation appears to be a strain with highly reduced virulence, probably easier to be cleared by the innate immune response.

Host-pathogen interaction studies in TB are moving forward to reveal more insights in the innate immune responses and to revisit the central paradigm of the effective immune response against *Mtb*. The establishment of a TB infection depends on many host factors such as the duration and frequency of *Mtb* exposure, the immune and nutrition status of the person as well as bacterial factors such the *Mtb* virulence (17, 18). TB infection can be detected by the development of a delayed immune cellular response, revealed by the purified protein derivative-PPD test, particularly useful in non-BCG (Bacille Calmette Guerin) vaccinated populations (19). Based on studies of contacts of TB cases that are expected to be in a higher risk to develop an *Mtb* infection, around 50% and sometimes more than 50% of the contacts are PPD negative, followed by latent TB cases and in a lesser extent, active TB individuals (17, 20-22). This point suggests that there is an important contribution of the early immune response in the control of *Mtb* infection for the PPD negative population that although has been in contact with the bacterium, do not generate a delayed immune cellular response (19). Currently, there are many research groups working in deciphering more aspects of the innate immune responses against *Mtb* such as the specific role of natural killer cells, dendritic cells, toll like receptors (TLRs),

antimicrobial peptides, among other players (23-25). Additionally, recent studies on BCG

vaccination suggest that there is an activation of the innate immunity part called trained

immunity (equivalent to the acquired memory response) that could be responsible of some of the

protective effect of the BCG vaccination (26). Additionally, the vaccine attempt against TB that

generated a cellular immune response against *Mtb*, failed in its protective role against the

bacterial infection (27). In this way, as it has been previously described, the acquired cellular

immune response (including the memory arm) is not sufficient to completely sterilize and protect

the host against *Mtb*. In fact, TB infection itself does not prevent of new infection episode either

with the same or with another strain (28). The lessons learned from the TB cellular immune

response during these years is probably suggesting that its activation is related with the co-

evolution of *Mtb* with the human host. A clear representation of this is the granuloma structure

created during *Mtb* infection, which is maintained by the host cellular immune response. The

granuloma is a self-contained structure that prevents *Mtb* (and the immune response towards it)

to cause symptoms to the host. In this structure, there is a mix of metabolically inactive and

active bacteria, with a predominance of the first (29). Therefore, the central paradigm of immune

response in TB, that attributes a protective role to the T helper type 1 (Th1) cellular

inflammatory responses with a balanced Th2 response, is currently enriched adding the role of

the innate immune responses and the role of other T cell populations such as Th17 against *Mtb*

infection (28, 30, 31).

Once again, the study described in this dissertation propose the possible contribution of

TLR-2 as an important player in the innate immune response for the clearance of the INHr

clinical strains in association with recently explored *Mtb* proteins ChoD and LprA (32, 33). In

addition, the increase of other lipoproteins in the INHr strain, also suggest the role of innate

immune response via TLR-2 in attenuated strains of *Mtb*. This could be further confirmed using a mouse model knock out for TLR-2 that is expected to have an impaired function of IL-6 and TNF-α. This mouse model was previously used to confirm the role of the19 KDa lipoprotein LpqH in *Mtb* infection (34). In this model, it would be possible to evaluate if the differences between INHr and INHs are maintained or attributable to an early innate response. Although, a reduction in virulence in INHr strains compared to their INHs pairs is still expected, because of the reduced catalase-peroxidase levels; it is also expected that there would be more quantifiable amounts of cytokines in those immunodeficient mouse models and less inhibited growth particularly for the clinical INHr strain (compared to what it was observed in the C57BL/6 mouse, chapter IV).

Here, the importance developing more sensitive and reliable instruments to detect changes in the levels of biomolecules (specifically proteins and mycolic acids); together with the analysis of the resulting data in the context of *Mtb* physiology is crucial to understand the biological significance of the proteome and lipid changes. This analysis is mostly relying on the previously annotated reference gene of H37Rv strain and also the available MA database from the same reference strain. The proteomics findings together with the mycolic acids (MA) analysis of the clinical clonal pair support an alternative fatty acid biosynthetic pathway. In the clinical INHr strain, there was an important reduction in the core enzymes of the fatty acid synthase II pathway (KasB, AcpM, HadABC complex, InhA) that included the MA biosynthesis and transfer to the cell wall (reduced FbpA). MA are vital molecules *Mtb* (35, 36). In this way, lipidomics and proteomics findings allow the generation of the hypothesis that the exposure of *Mtb* to a first line drug regime stimulates an alternative fatty acid biosynthetic pathway lead by FabG4 and HtdX with the possible contribution of other proteins in the *mymA* operon such as

FadD13. This hypothesis also includes the possible participation of FabG4 and HtdX in the fatty acid biosynthesis in the no matter the AcpM levels, since FabG4 and HtdX also have affinity for CoA intermediates that could be used instead of acyl carrier protein-intermediates (37-39). The upregulation of the *mymA* operon encode for proteins in the later activation, acylation and transfer to the cell wall of the MA that is upregulated when *Mtb* is in an acidic environment (40). The increased levels of the proteins of this *mymA* operon complement the pathway started with the FabG4 and HtdX enzymes as proposed in chapter III (Figure 3.7) and also suggest the unregulated redox environment that leads to an acidic intracellular space (40).

It would be very interesting to experimentally confirm if the exposure to a complete first line drug combined therapy, as normally happens in the host, is able to induce this alternative fatty acid biosynthetic pathway. This could be either done *in vitro* by exposing the bacteria to sub-inhibitory concentration of the drugs and evaluate the proteome differences before and after the drug exposure. Alternatively and probably in a more clinically relevant manner, this analysis can be performed in *Mtb* clinical isolates pre-drug treatment and post-drug treatment from TB cases, isolating clonal *Mtb* strains every month after receiving the first line drug therapy during four months. In this way, it will be possible to detect if this alternative fatty acid biosynthetic pathway is stimulated after first line drug exposure and/or because of a drug resistant *Mtb* phenotype including MDR. This is supported by the fact that the laboratory INHr strain that acquired its resistant phenotype after being initially exposed to INH only, did not have alteration in those previously mentioned proteins. In addition, a previous proteomic comparison using sequential strains of *Mtb* showed also increased levels of FabG4 in the MDR strain and the Beijing pair studied in this dissertation also had increased levels of FabG4 and HtdX in the INHr strain (41, 42). Previous suggestions of an alternative biosynthetic fatty acid pathway in *Mtb*

179

drug resistant strains was proposed since 1983 from studies in *M. smegmatis* and later supported

by functional studies of FabG4 and HtdX proteins (37, 43). This alternative pathway could

potentially serve as a useful drug target in MDR *Mtb* strains. Additionally, it is plausible that

being extremely relevant for *Mtb* physiology, fatty acid biosynthesis (including MA) should be a

pathway with some level of redundancy it happens with the fatty acid β-oxidation in *Mtb* (44).

Importantly, it is still uncertain if many of the protein differences and the trends observed

in both INHr strains could be associated with the initial exposure to INH as well. Remarkably,

previous studies have demonstrated that protein members of the FAS II pathway, polyketide

synthesis, the cytochrome bc$_1$ complex (QcrABC operon), and ATP synthase are distinctly

altered after the exposure to INH in mycobacteria (45, 46). In this dissertation, these proteins

were also altered in INHr strains (particularly the clinical strain); however their trends were

opposite to those previously described. For instance, AcpM and KasB were reduced in the INHr

clinical strain, while these and other members of the FAS II complex were previously found to

increase after exposure of INHs strains to INH (45-47). Preliminary attempts during this PhD

work which looked at the proteome of these clonal pairs before and after exposure to INH

revealed other proteins with the same pattern. According to the proteomics data presented in

chapter III, most lipoproteins were increased after the acquisition of INHr. One of them was

LpqH, although it was not statistically significant (p=0.07, fold change=0.6 in cytosol of clinical

pair comparison). Western blot evaluation of LpqH in both clinical and laboratory pairs

following exposure to sub-inhibitory concentrations of INH, demonstrated that this protein was

reduced in the membrane fraction, regardless of their susceptibility profile to INH (Figure 5.1).

Additional evaluation of protein changes after drug exposure *in vitro* as well as using samples

derived from clinical patients will help develop a complete picture regarding evolution of INH

resistant phenotypes in *Mtb*.
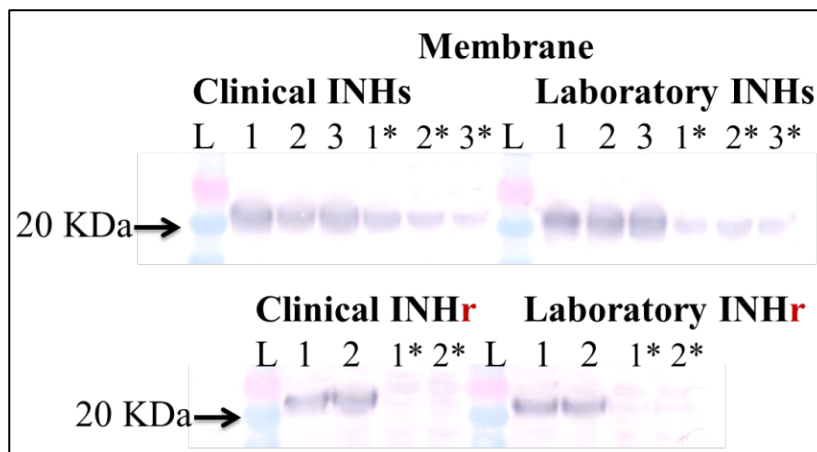
Figure 5.1. Western blot anti-LpqH in INHs and INHr clonal pairs of *Mtb* exposed to INH.
*Biological replicates of *Mtb* strains exposed to INH. INHs strains were exposed to 0.05µg/mL
and INHr strains were exposed to 0.2µg/mL of INH.

The proteomic findings also showed alteration of proteins that are currently used as drug

targets including those proteins in the ATP synthase machinery and RpoB. The recent FDA

approved drug for MDR-TB cases, bedaquiline, targets the ATP synthase in *Mtb* which is a

relatively novel target (48). The increased level of proteins of this complex commonly found in

both INHr generates concern because a common bacteria drug resistance mechanism is the

overproduction of the target. The levels of these proteins also generate some questions regarding

the minimum inhibitory concentration (MIC) of bedaquiline and rifampicin for these INHr

strains. Although previous phenotypic studies have demonstrated that the INHr strains were

susceptible to rifampicin by the agar proportion method, it would be interesting to observe if the

increased RpoB and ATP synthase subunits levels create higher MIC values of rifampicin and

bedaquiline for the INHr strains compared with their susceptible progenitor. A possible solution

could be to explore the use of this ATP-synthase inhibitor in the first line drug therapy

combination. Similarly RpoB had a higher abundance but only in the clinical INHr strain, suggesting that the increase of this protein was mostly associated with previous exposure to combined drug therapy in the patient. This combined therapy includes rifampicin that targets this increased protein (RpoB) (49, 50).

As part of this comprehensive characterization, levels of *katG* mRNA were evaluated to complement the description of *katG* mutations in clinically-isolated INHr strains, however the inconsistency of these results make it difficult to derive confident conclusions. In the central paradigm of molecular biology, genetic information goes from DNA to mRNA to proteins. However, it should be noted that changes in the proteome do not always relate with changes in the transcriptome. In fact, the correlation between protein and mRNA levels is not very strong which can occur due to a variety of post-transcriptional events such as phosphorylation and acetylation, as well as mRNA half-life and protein localization (51). For this reason it is necessary to look at both transcripts and protein levels together to achieve an accurate interpretation of the phenotypic event that is being studied, in this case, INH resistance in *Mtb*. Diverse strategies have been used to analyze RNA such as Northern blotting, microarrays, reverse transcription (RT)-quantitative PCR (qPCR), the more recently introduced in the digital droplet (dd)-PCR (52) and high throughput RNA seq whole transcriptome analysis (53). Advantages of the ddPCR strategy over RT-qPCR include its higher precision and reproducibility, lower limits of mRNA quantification, and the elimination of standard curves and the inter-well technical replicates (54).

During this dissertation, many attempts were made to study *katG* transcripts of the clinical INHr strain. The initial RT-PCR showed increased *katG* transcript levels which opposed the decreased protein abundance found in every cellular fraction of the clinical INHr stain

(Chapter II and III). Notably, there was a lot of variability in the levels of *katG* transcripts among biological replicates of the clinical INHr strain which did not occur with the biological replicates of the clinical INHs strain. After this, the second proteomics approach described in Chapter III, revealed significant differences in the reference gene chosen: *sigA* (Chapter III, table 3.3). Therefore, the RT-PCR analysis with just two biological replicates for the clinical INHr strain and without an accurate reference gene could not be validated. The later ddPCR strategy with *rrs* as the reference gene showed variable results as well for the clinical INHr strain. Consequently, an interpretation of the *katG* mRNA transcripts in the INHr strain could not be made.

In general, transcriptomics studies should be carefully designed starting with the pre-analytical phase of the experiment. This includes culture conditions, growth phases at which *Mtb* cells are collected for RNA isolation, cell density, and the number of samples processed simultaneously among other factors. Transcript levels are highly dependent on the transcription and degradation rate of the mRNA (55). Possible RNA decay events in mycobacteria as well as the possible formation of secondary mRNA structures are important factors that should also be taken into account during analysis of *katG* transcripts in the mutant clinical strain. Less is known about RNA decay in gram positive bacteria and *Mtb* compared with Gram negative bacteria (13). From studies performed in *Escherichia coli*, possible candidates for *Mtb* proteins in the RNA degradosome complex have been annotated which include the homologs for RNAse E (Rv2444c), enolase (Rv1023), helicase (Rv3211) and the polynucleotide phosphorylase (Rv2783) which participate in the degradation of mRNA and small RNAs as well in the maturation of rRNA and tRNA (56). Interestingly, two proteins of the *Mtb* RNA degradosome Rv2444c (RNAse E) and Rv1023 (enolase) were increased in the clinical INHr, with p values of 0.057 and 0.0028 respectively. Enolase (Rv1023) is also classified as a catalytic protein of

glycolysis (Chapter III, Figure3.5B). However, significant reduced levels of the phosphorylase

Rv2783, also predicted to be part of the degradosome, were observed in the clinical INHr strain

(Appendix I Table 6). Additionally, in a previous experiment that was looking for enzymes that

were part of the RNA degradosome complex in *Mtb,* the BCG protein GroEL co-precipitated

with the RNAse E protein, implying a possible role of GroEL in the RNA degradasome (56). In

this dissertation, GroEL1 was significantly higher the INHr strains (Appendix I Tables 5 and 6).

With these findings, one could hypothesize that the difficulty in obtaining a valid interpretation

of the *katG* transcript levels could be due to an altered RNA decay process occurring in the

clinical INHr strain that should be further explored.

A possible strategy is to measure global mRNA decay is by exposing the logarithmic

phase *Mtb* cultures to a defined concentration of rifampicin (>250 times the minimum inhibitory

concentration value) at different time points. After this, the total bacterial RNA will be carefully

extracted and then the amount of different gene transcripts will be determined after the

conversion to cDNA and analyzed by ddPCR. The role of the rifampicin is to inhibit

transcription by the binding to the ß-subunit of the RNA polymerase, therefore, the

determination of the amount of transcripts at each time point will be confidently comparable.

The genes to evaluate could be those encoding for the proteins with the most uniform levels

among the clinical INHs and INHr strains in the proteomic analysis and also looking at the gene

of interest *katG* and the regularly used reference gene 16S ribosomal RNA. This strategy has

been used in previous studies to determine the *Mtb* mRNA half-life which seems to be higher to

the mRNA half-life of other prokaryotes (13, 55).

Furthermore, the proteomics analysis suggested altered levels in the phthiocerol,

phthiocerol dimycocerosates (DIMs), polyketides and sulfolipids. Some enzymes involved in

their biosynthesis are increased in the laboratory INHr strain (Pks2, Pks4, PpsA) and some reduced in the clinical strain (Mas, PpsE). In the clinical INHr, the protein involved in the transport of DIMs across the membrane (LppX) was reduced. Additionally, the protein PhoP that regulates the genes involved in polyketide synthesis was commonly increased in both INHr strains. Therefore, in order to complete the lipid characterization changes after acquisition of INHr, it is possible to look for those lipids by following the lipidomics approach described by Mark Sartain *et al*. This global lipidomics strategy will focus on the comparative analysis of the total lipids in each clonal pair. Further analysis of the particular population of MA in the laboratory clonal pair by methodologies such as nuclear magnetic resonance (NMR) as previously performed by Glickman, Watanabe and others, will help to investigate if this ratio between the methyl transferases CmaA1 and 2 and MmaA1 and 2, affected the *cis/trans* ratio and the cyclopropane rings of the α- and oxygenated MA (57-59). However, since this trend was only observed for the laboratory strain, the interpretation of these results cannot serve as a general description of biochemical features in clinical INHr strains.

Finally, more than 60 years of *Mtb* exposure to INH has driven a selection of INHr strains by treating TB patients with this drug. The understanding of the phenomena happening inside the bacteria after acquiring this phenotype is giving us a better description of the circulating *Mtb* strains that are able to grow in the presence of this effective anti-TB drug and cause disease. This work has provided an opportunity to explore the integration of many omics disciplines together and advance the understanding of *Mtb* physiology. The results presented in this dissertation contribute to the research of a pathogen that has developed many strategies to co-evolve with the host in order to maintain a symbiotic relationship in the form of latent infection (29). This ancient pathogen also generates active TB cases and death, which is

associated with the host immune status as well as intrinsic bacterial characteristics such as drug resistance. TB research has experienced different phases in human history and re-emerged in 1993 after the World Health Organization (WHO) proclaimed a "global TB emergency." The work in this dissertation is in line with one of the current pillars of the WHO End TB strategy that aims to end the global TB epidemic by 2035. Specifically, the third pillar is to "intensify research and innovation' globally from basic sciences to operational research (60). The study of INH resistance event in this dissertation provides new insights of alternative *Mtb* metabolic routes, supportive information of previously discussed proteins involved in the bacteria-host interaction via TLR-2, and reveals some specific aspects of currently used drug targets for drug resistant *Mtb*. In this way, a better understanding of the INH resistant phenotype of *Mtb* strains was possible, revealing individual and common findings that will contribute to find more directed drug targets to combat these highly prevalent resistant strains.

REFERENCES

1.   Tortora GJ, Funke BR, Case CL. Microbiology : an introduction media update. Rev. 7th ed. San Francisco: Benjamin Cummings; 2002.

2.   Vilchèze C, Jacobs WR. Resistance to Isoniazid and Ethionamide in Mycobacterium tuberculosis: Genes, Mutations, and Causalities. Microbiol Spectr. 2014 Aug;2(4):MGM2-0014-2013.

3.   Cade CE, Dlouhy AC, Medzihradszky KF, Salas-Castillo SP, Ghiladi RA. Isoniazid-resistance conferring mutations in Mycobacterium tuberculosis KatG: catalase, peroxidase, and INH-NADH adduct formation activities. Protein Sci. 2010 Mar;19(3):458-74.

4.   Gagneux S, Burgos MV, DeRiemer K, Encisco A, Munoz S, Hopewell PC, et al. Impact of bacterial genetics on the transmission of isoniazid-resistant Mycobacterium tuberculosis. PLoS Pathog. 2006 Jun;2(6):e61.

5.   Pym AS, Saint-Joanis B, Cole ST. Effect of katG mutations on the virulence of Mycobacterium tuberculosis and the implication for transmission in humans. Infect Immun. 2002 Sep;70(9):4955-60.

6.   Miesel L, Rozwarski DA, Sacchettini JC, Jacobs WR. Mechanisms for isoniazid action and resistance. Novartis Found Symp. 1998;217:209-20; discussion 20-1.

7.   Zhang Y, Garbe T, Young D. Transformation with katG restores isoniazid-sensitivity in Mycobacterium tuberculosis isolates resistant to a range of drug concentrations. Mol Microbiol. 1993 May;8(3):521-4.

8.   Zhang Y, Heym B, Allen B, Young D, Cole S. The catalase-peroxidase gene and isoniazid resistance of Mycobacterium tuberculosis. Nature. 1992 Aug;358(6387):591-3.

9.   Ordway DJ, Sonnenberg MG, Donahue SA, Belisle JT, Orme IM. Drug-resistant strains of Mycobacterium tuberculosis exhibit a range of virulence for mice. Infect Immun. 1995 Feb;63(2):741-3.

10.   Mahapatra S, Woolhiser LK, Lenaerts AJ, Johnson JL, Eisenach KD, Joloba ML, et al. A novel metabolite of antituberculosis therapy demonstrates host activation of isoniazid and formation of the isoniazid-NAD+ adduct. Antimicrob Agents Chemother. 2012 Jan;56(1):28-35.

11.   Bhat AG, Vashisht R, Chandra N. Modeling metabolic adjustment in Mycobacterium tuberculosis upon treatment with isoniazid. Syst Synth Biol. 2010 Dec;4(4):299-309.

12.   Jena L, Deshmukh S, Waghmare P, Kumar S, Harinath BC. Study of mechanism of interaction of truncated isoniazid–nicotinamide adenine dinucleotide adduct against multiple enzymes of Mycobacterium tuberculosis by a computational approach. International Journal of Mycobacteriology. 2015;4(4):276-83.

13.   Sala C, Forti F, Magnoni F, Ghisotti D. The katG mRNA of Mycobacterium tuberculosis and Mycobacterium smegmatis is processed at its 5' end and is stabilized by both a polypurine sequence and translation initiation. BMC Molecular Biology. 2008;9(1):33.

14.   Datta G, Nieto LM, Davidson RM, Mehaffy C, Pederson C, Dobos KM, et al. Longitudinal whole genome analysis of pre and post drug treatment Mycobacterium tuberculosis isolates reveals progressive steps to drug resistance. Tuberculosis. 2016;98:50-5.

15.   Richardson ET, Lin SY, Pinsky BA, Desmond E, Banaei N. First documentation of isoniazid reversion in Mycobacterium tuberculosis. Int J Tuberc Lung Dis. 2009 Nov;13(11):1347-54.

16.   Heym B, Alzari PM, Honore N, Cole ST. Missense mutations in the catalase-peroxidase gene, katG, are associated with isoniazid resistance in Mycobacterium tuberculosis. Mol Microbiol. 1995 Jan;15(2):235-45.

17.   Morrison J, Pai M, Hopewell PC. Tuberculosis and latent tuberculosis infection in close contacts of people with pulmonary tuberculosis in low-income and middle-income countries: a systematic review and meta-analysis. Lancet Infect Dis. 2008 Jun;8(6):359-68.

18.   Schaaf HS, Zumla A, Grange JM. Tuberculosis : a comprehensive clinical reference. Edinburgh: Saunders/Elsevier; 2009.

19.   Black CA. Delayed type hypersensitivity: current theories with an historic perspective. Dermatol Online J. 1999 May;5(1):7.

20.   Turner RD, Bothamley GH. Cough and the Transmission of Tuberculosis. J Infect Dis. 2014 Nov.

21.   Moyo N, Tay EL, Denholm JT. Evaluation of tuberculin skin testing in tuberculosis contacts in Victoria, Australia, 2005-2013. Public Health Action. 2015 Sep;5(3):188-93.

22.   Azarkar Z, Arbabi A. Comparison of PPD test in household contacts of smear-positive and -negative tuberculosis (TB). International Journal of Mycobacteriology. 2015;4:95.

23.   Berrington WR, Hawn TR. Mycobacterium tuberculosis, macrophages, and the innate immune response: does common variation matter? Immunol Rev. 2007 Oct;219:167-86.

24.   Sia JK, Georgieva M, Rengarajan J. Innate Immune Defenses in Human Tuberculosis: An Overview of the Interactions between Mycobacterium tuberculosis and Innate Immune Cells. J Immunol Res. 2015;2015:747543.

25.   van Crevel R, Ottenhoff TH, van der Meer JW. Innate immunity to Mycobacterium tuberculosis. Clin Microbiol Rev. 2002 Apr;15(2):294-309.

26.   Netea MG, van Crevel R. BCG-induced protection: effects on innate immune memory. Semin Immunol. 2014 Dec;26(6):512-7.

27.   Tameris MD, Hatherill M, Landry BS, Scriba TJ, Snowden MA, Lockhart S, et al. Safety and efficacy of MVA85A, a new tuberculosis vaccine, in infants previously vaccinated with BCG: a randomised, placebo-controlled phase 2b trial. Lancet. 2013 Mar;381(9871):1021-8.

28.   Andersen P, Woodworth JS. Tuberculosis vaccines--rethinking the current paradigm. Trends Immunol. 2014 Aug;35(8):387-95.

29.   Gengenbacher M, Kaufmann SH. Mycobacterium tuberculosis: success through dormancy. FEMS Microbiol Rev. 2012 May;36(3):514-32.

30.   Nunes-Alves C, Booty MG, Carpenter SM, Jayaraman P, Rothchild AC, Behar SM. In search of a new paradigm for protective immunity to TB. Nat Rev Microbiol. 2014 Apr;12(4):289-99.

31.   Orme IM, Robinson RT, Cooper AM. The balance between protective and pathogenic immune responses in the TB-infected lung. Nat Immunol. 2015 Jan;16(1):57-63.

32.   Briken V, Klink M, Brzezinska M, Szulc I, Brzostek A, Kielbik M, et al. Cholesterol Oxidase Is Indispensable in the Pathogenesis of Mycobacterium tuberculosis. PLoS One. 2013;8(9):e73333.

33.   Pecora ND, Gehring AJ, Canaday DH, Boom WH, Harding CV. Mycobacterium tuberculosis LprA is a lipoprotein agonist of TLR2 that regulates innate immunity and APC function. J Immunol. 2006 Jul 1;177(1):422-9.

34.   Takeuchi O, Sato S, Horiuchi T, Hoshino K, Takeda K, Dong Z, et al. Cutting edge: role of Toll-like receptor 1 in mediating immune response to microbial lipoproteins. J Immunol. 2002 Jul 1;169(1):10-4.

35.	Marrakchi H, Lanéelle MA, Daffé M. Mycolic acids: structures, biosynthesis, and beyond. Chem Biol. 2014 Jan;21(1):67-85.

36.	Takayama K, Wang C, Besra GS. Pathway to synthesis and processing of mycolic acids in Mycobacterium tuberculosis. Clin Microbiol Rev. 2005 Jan;18(1):81-101.

37.	Dutta D, Bhattacharyya S, Roychowdhury A, Biswas R, Das Amit K. Crystal structure of hexanoyl-CoA bound to β-ketoacyl reductase FabG4 ofMycobacterium tuberculosis. Biochemical Journal. 2013;450(1):127-39.

38.	Gurvitz A. The essential mycobacterial genes, fabG1 and fabG4, encode 3-oxoacyl-thioester reductases that are functional in yeast mitochondrial fatty acid synthase type 2. Molecular Genetics and Genomics. 2009;282(4):407-16.

39.	Sacco E, Slama N, Backbro K, Parish T, Laval F, Daffe M, et al. Revisiting the Assignment of Rv0241c to Fatty Acid Synthase Type II of Mycobacterium tuberculosis. Journal of Bacteriology. 2010;192(15):4037-44.

40.	Singh A, Jain S, Gupta S, Das T, Tyagi AK. mymAoperon ofMycobacterium tuberculosis: its regulation and importance in the cell envelope. FEMS Microbiology Letters. 2003;227(1):53-63.

41.	Nieto R LM, Mehaffy C, Dobos KM. Comparing isogenic strains of Beijing genotypeMycobacterium tuberculosisafter acquisition of Isoniazid resistance: A proteomics approach. Proteomics. 2016;16(9):1376-80.

42.	Singh A, Gopinath K, Sharma P, Bisht D, Singh N, Singh S. Comparative proteomic analysis of sequential isolates of Mycobacterium tuberculosis from a patient with pulmonary tuberculosis turning from drug sensitive to multidrug resistant. Indian J Med Res. 2015 Jan;141(1):27-45.

43.	Biswas R, Dutta D, Das AK. Cloning, overexpression, purification, crystallization and preliminary X-ray diffraction analysis of Rv0241c (HtdX) fromMycobacterium tuberculosisH37Rv. Acta Crystallographica Section F Structural Biology and Crystallization Communications. 2013;69(10):1110-3.

44.	Cole ST, Brosch R, Parkhill J, Garnier T, Churcher C, Harris D, et al. Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence. Nature. 1998;393(6685):537-44.

45.	Boshoff HIM, Myers TG, Copp BR, McNeil MR, Wilson MA, Barry CE. The Transcriptional Responses of Mycobacterium tuberculosis to Inhibitors of Metabolism: NOVEL INSIGHTS INTO DRUG MECHANISMS OF ACTION. Journal of Biological Chemistry. 2004;279(38):40174-84.

46.	Wang R, Marcotte EM. The Proteomic Response ofMycobacterium smegmatisto Anti-Tuberculosis Drugs Suggests Targeted Pathways. Journal of Proteome Research. 2008;7(3):855-65.

47.	Mdluli K. Inhibition of a Mycobacterium tuberculosis -Ketoacyl ACP Synthase by Isoniazid. Science. 1998;280(5369):1607-10.

48.	Haagsma AC, Abdillahi-Ibrahim R, Wagner MJ, Krab K, Vergauwen K, Guillemont J, et al. Selectivity of TMC207 towards mycobacterial ATP synthase compared with that towards the eukaryotic homologue. Antimicrob Agents Chemother. 2009 Mar;53(3):1290-2.

49.	Zumla A, Nahid P, Cole ST. Advances in the development of new tuberculosis drugs and treatment regimens. Nat Rev Drug Discov. 2013 May;12(5):388-404.

50.	Campbell EA, Korzheva N, Mustaev A, Murakami K, Nair S, Goldfarb A, et al. Structural mechanism for rifampicin inhibition of bacterial rna polymerase. Cell. 2001 Mar;104(6):901-12.

51.	Maier T, Güell M, Serrano L. Correlation of mRNA and protein in complex biological samples. FEBS Letters. 2009;583(24):3966-73.

52.	Hindson BJ, Ness KD, Masquelier DA, Belgrader P, Heredia NJ, Makarewicz AJ, et al. High-Throughput Droplet Digital PCR System for Absolute Quantitation of DNA Copy Number. Analytical Chemistry. 2011;83(22):8604-10.

53.	Tang F, Barbacioru C, Wang Y, Nordman E, Lee C, Xu N, et al. mRNA-Seq whole-transcriptome analysis of a single cell. Nature Methods. 2009;6(5):377-82.

54.	Taylor SC, Carbonneau J, Shelton DN, Boivin G. Optimization of Droplet Digital PCR from RNA and DNA extracts with direct comparison to RT-qPCR: Clinical implications for quantification of Oseltamivir-resistant subpopulations. Journal of Virological Methods. 2015;224:58-66.

55.	Rustad TR, Minch KJ, Brabant W, Winkler JK, Reiss DJ, Baliga NS, et al. Global analysis of mRNA stability in Mycobacterium tuberculosis. Nucleic Acids Research. 2012;41(1):509-17.

56.	Kovacs L, Csanadi A, Megyeri K, Kaberdin VR, Miczak A. Mycobacterial RNase E-associated proteins. Microbiol Immunol. 2005;49(11):1003-7.

57.	Watanabe M, Aoyagi Y, Ridell M, Minnikin DE. Separation and characterization of individual mycolic acids in representative mycobacteria. Microbiology. 2001 Jul;147(Pt 7):1825-37.

58.	Glickman MS. The mmaA2 gene of Mycobacterium tuberculosis encodes the distal cyclopropane synthase of the alpha-mycolic acid. J Biol Chem. 2003 Mar 7;278(10):7844-9.

59.	Mederos L, Valdivia JA, Valero-Guillen PL. Analysis of the structure of mycolic acids of Mycobacterium simiae reveals a particular composition of alpha-mycolates in strain 'habana' TMC 5135, considered as immunogenic in tuberculosis and leprosy. Microbiology. 2007 Dec;153(Pt 12):4159-65.

60.	Organization WH. A Global action framework for TB research. Geneva, Switzerland: WHO; 2015. p. 76.

Appendix Table 1. Significantly different proteins in the cell wall of the clonal laboratory pair

Identified Proteins p<0.05 (t-test)	Accession Number	Fold Change Laboratory (INHs/ INHr)	Molecular Weight	NSAF	
				Average Lab INHr x 1000	Average Lab INHs x 1000
catalase-peroxidase-peroxynitritase T katG (741 aa)	Rv1908c.1	5.0	81 kDa	52.8	265.6
5-methyltetrahydropteroyltriglutamate-homocysteine methyltransferase metE (760 aa)	Rv1133c.1	1.5	82 kDa	45.4	67.8
conserved hypothetical protein (272 aa)	Rv0831c.1	1.2	30 kDa	181.6	225.2
iron-regulated peptidyl-prolyl-cis-trans-isomerase A ppiA (183 aa)	Rv0009.1	2.1	19 kDa	75.6	160.1
NADH-dependent enoyl-[acyl-carrier-protein] reductase inhA (270 aa)	Rv1484.1	1.4	29 kDa	78.6	109.7
mycolic acid synthase umaA (287 aa)	Rv0469.1	1.4	33 kDa	85.0	117.4
bifunctional methylenetetrahydrofolate dehydrogenase folD (282 aa)	Rv3356c.1	2.6	29 kDa	44.5	115.7
valyl-tRNA synthase protein valS (877 aa)	Rv2448c.1	2.0	98 kDa	10.1	19.8
sulfatase (466 aa)	Rv0296c.1	2.3	52 kDa	22.4	51.8
30S ribosomal protein S11 rpsK (140 aa)	Rv3459c.1	1.5	15 kDa	141.6	212.5
Esat-6 like protein esxO (95 aa)	Rv2346c.1	1.7	10 kDa	166.4	285.5
phosphoribosylformylglycinamidine synthase I purG (225 aa)	Rv0788.1	4.3	24 kDa	9.8	42.5
short-chain type dehydrogenase/reductase (315 aa)	Rv3485c.1	6.3	33 kDa	5.1	31.9
elongation factor P efp (188 aa)	Rv2534c.1	3.6	20 kDa	25.1	90.2
conserved hypothetical protein (152 aa)	Rv3678c.1	1.9	15 kDa	102.2	192.5
quinone reductase qor (329 aa)	Rv1454c.1	2.7	34 kDa	10.3	27.8
leucyl-tRNA synthetase leuS (970 aa)	Rv0041.1	4.6	108 kDa	4.6	21.1
3-deoxy-D-arabino-heptulosonate 7-phosphate synthase aroG (463 aa)	Rv2178c.1	2.5	51 kDa	11.8	29.6
conserved alanine and leucine rich protein (885 aa)	Rv2567.1	1.4	95 kDa	6.0	8.3
polyphosphate kinase ppk (743 aa)	Rv2984.1	17.2	83 kDa	0.5	9.3
acyl-CoA dehydrogenase fadE8 (543 aa)	Rv0672.1	6.7	59 kDa	2.1	14.2
aspartate carbamoyltransferase pyrB (320 aa)	Rv1380.1	1.4	34 kDa	42.2	60.0
50S ribosomal protein L4 rplD (224 aa)	Rv0702.1	3.4	24 kDa	16.0	55.2
fatty-acid-CoA racemase far (360 aa)	Rv0855.1	14.9	38 kDa	1.1	16.7
conserved hypothetical protein (288 aa)	Rv1061.1	4.0	32 kDa	17.1	68.3
6-phosphofructokinase pfkA (344 aa)	Rv3010c.1	2.0	37 kDa	13.3	26.3
tryptophanyl-tRNA synthetase trpS (337 aa)	Rv3336c.1	5.7	36 kDa	1.2	6.7
chorismate synthase aroF (402 aa)	Rv2540c.1	8.0	42 kDa	2.5	20.4
acetyl-CoA acetyltransferase fadA5 (392 aa)	Rv3546.1	4.3	41 kDa	3.5	15.2
aminotransferase (398 aa)	Rv0858c.1	3.7	42 kDa	4.6	17.1
dioxygenase (383 aa)	Rv3161c.1	#DIV/0!	43 kDa	0.0	7.1
dTDP-glucose 4,6-dehydratase (327 aa)	Rv3784.1	3.5	36 kDa	6.8	23.6

conserved membrane protein (539 aa)	Rv0283.1	0.7	56 kDa	108.2	80.1
polyketide beta-ketoacyl synthase pks4 (1583 aa)	Rv1181.1	0.8	168 kDa	108.4	91.2
ATP synthase gamma chain atpG (306 aa)	Rv1309.1	0.6	34 kDa	178.8	110.2
protease IV sppA (624 aa)	Rv0724.1	0.5	66 kDa	68.9	32.7
succinyl-CoA synthetase beta chain sucC (388 aa)	Rv0951.1	0.7	41 kDa	107.2	74.9
aminopeptidase pepB (516 aa)	Rv2213.1	0.7	53 kDa	81.4	56.1
succinate-semialdehyde dehydrogenase [NADP+]-dependent gabD2 (519 aa)	Rv1731.1	0.7	55 kDa	97.8	67.6
acyl-CoA dehydrogenase fadE24 (469 aa)	Rv3139.1	0.8	50 kDa	113.3	85.5
ATP synthase delta chain atpH (447 aa)	Rv1307.1	0.7	49 kDa	103.6	70.7
fatty-acyl-CoA reductase (342 aa)	Rv1543.1	0.5	37 kDa	89.2	46.3
conserved hypothetical protein (399 aa)	Rv0106.1	0.6	44 kDa	220.0	127.9
fatty-acid synthase (421 aa)	Rv3720.1	0.6	47 kDa	107.5	65.9
short-chain type dehydrogenase/reductase (277 aa)	Rv1245c.1	0.5	29 kDa	57.9	28.3
PPE family protein (392 aa)	Rv1196.1	0.4	39 kDa	64.3	23.1
long-chain acyl-CoA synthase (1000 aa)	Rv1683.1	0.4	107 kDoa	20.9	8.7
29 kda antigen cfp29 (266 aa)	Rv0798c.1	0.1	29 kDa	117.8	17.5
conserved hypothetical protein (406 aa)	Rv0530.1	0.7	43 kDa	103.6	68.5
conserved hypothetical protein (420 aa)	Rv2449c.1	0.2	44 kDa	9.0	2.2
conserved membrane protein (325 aa)	Rv2037c.1	0.3	35 kDa	71.6	24.0
conserved hypothetical protein (322 aa)	Rv2955c.1	0.2	36 kDa	23.3	4.3
conserved hypothetical protein (933 aa)	Rv1784.1	0.8	101 kDa	24.7	18.6
RNA polymerase sigma factor sigA (529 aa)	Rv2703.1	0.4	58 kDa	34.4	15.5
short-chain type dehydrogenase ephD (593 aa)	Rv2214c.1	0.3	64 kDa	21.7	6.2
conserved hypothetical protein (678 aa)	Rv1836c.1	0.5	70 kDa	18.1	9.4
alkyl hydroperoxide reductase D protein ahpD (178 aa)	Rv2429.1	0.6	19 kDa	198.2	115.3
bacterioferritin bfrA (160 aa)	Rv1876.1	0.7	18 kDa	57.0	37.3
daunorubicin-dim-transport ATP-binding protein ABC transporter drrA (332 aa)	Rv2936.1	0.5	36 kDa	79.4	40.1
membrane protein (146 aa)	Rv0008c.1	0.6	16 kDa	83.3	47.0
hypothetical protein (856 aa)	Rv0613c.1	0.7	93 kDa	16.5	12.2
transmembrane ATP-binding protein ABC transorter (698 aa)	Rv2326c.1	0.0	73 kDa	7.8	0.0
membrane-associated serine protease (398 aa)	Rv3671c.1	0.3	41 kDa	34.4	11.4
transmembrane transport protein mmpL4 (968 aa)	Rv0450c.1	0.1	105 kDa	6.9	1.0
50S ribosomal protein L1 rplA (236 aa)	Rv0641.1	0.5	25 kDa	34.7	17.3
lipoprotein lprC (181 aa)	Rv1275.1	0.4	19 kDa	57.3	20.1

Appendix Table 2. Significantly different proteins in the cell wall of the clonal clinical pair

Identified Proteins p<0.05 (t-test)	Accession Number	Molecular Weight	Fold Change (INHs/ INHr)	NSAF	
				Average Clinical INHr x 1000	Average Clinical INHS x 1000
iron-regulated elongation factor tu tuf (397 aa)	Rv0685.1	44 kDa	1.2	377.5	461.7
fatty-acid synthase fas (3070 aa)	Rv2524c.1	326 kDa	1.5	14.8	21.6
multifunctional mycocerosic acid synthase membrane-associated mas (2112 aa)	Rv2940c.1	224 kDa	2.5	15.0	38.0
10 kda chaperonin groES (101 aa)	Rv3418c.1	11 kDa	3.1	545.6	1668.9
catalase-peroxidase-peroxynitritase T katG (741 aa)	Rv1908c.1	81 kDa	98.6	0.9	88.3
trigger factor protein tig (467 aa)	Rv2462c.1	51 kDa	1.3	129.6	173.1
3-oxoacyl-[acyl-carrier protein] synthase 2 kasB (439 aa)	Rv2246.1	46 kDa	1.7	57.7	95.6
elongation factor G fusA1 (702 aa)	Rv0684.1	77 kDa	1.4	38.4	54.2
endopeptidase ATP binding protein chain B clpB (849 aa)	Rv0384c.1	93 kDa	1.6	53.8	88.2
pyruvate dehydrogenase E2 component DlaT (554 aa)	Rv2215.1	57 kDa	1.2	84.1	99.5
phenolpthiocerol synthesis type-I polyketide synthase ppsE (1489 aa)	Rv2935.1	159 kDa	#DIV/0!	0.0	8.5
aspartate-semialdehyde dehydrogenase asd (346 aa)	Rv3708c.1	36 kDa	2.1	56.0	116.4
thiosulfate sulfurtransferase cysA2 (278 aa)	Rv0815c.1 (+1)	31 kDa	1.5	82.7	126.8
translation initiation factor IF-2 infB (901 aa)	Rv2839c.1	94 kDa	1.6	22.9	37.3
immunogenic protein mpt64 (229 aa)	Rv1980c.1	25 kDa	1.3	202.5	268.4
acyl-CoA dehydrogenase fadE25 (390 aa)	Rv3274c.1	42 kDa	1.4	69.2	95.2
alkyl hydroperoxide reductase C protein ahpC (196 aa)	Rv2428.1	22 kDa	8.9	20.5	181.4
alanyl-tRNA synthetase alaS (905 aa)	Rv2555c.1	97 kDa	1.9	10.3	19.4
conserved hypothetical protein (514 aa)	Rv2226.1	56 kDa	2.1	38.8	82.8
conserved hypothetical protein (159 aa)	Rv0635.1	17 kDa	1.7	92.6	161.5
thioredoxin reductase trxB2 (336 aa)	Rv3913.1	36 kDa	2.4	17.9	42.2
threonyl-tRNA synthetase thrS (693 aa)	Rv2614c.1	77 kDa	1.7	12.5	21.8
enhanced intracellular survival protein eis (403 aa)	Rv2416c.1	44 kDa	2.3	9.0	20.7
conserved hypothetical protein (134 aa)	Rv3716c.1	13 kDa	1.4	120.0	164.2
mannose-6-phosphate isomerase manA (409 aa)	Rv3255c.1	43 kDa	1.3	19.6	26.3
hypothetical protein sseC2 (101 aa)	Rv0814c.1 (+1)	10 kDa	3.0	103.0	311.4
Esat-6 like protein esxO (95 aa)	Rv2346c.1	10 kDa	1.3	116.8	149.7
pantoate-beta-alanine ligase panC (310 aa)	Rv3602c.1	33 kDa	2.9	11.8	33.9
conserved hypothetical protein (167 aa)	Rv0637.1	19 kDa	3.3	12.1	39.8

fatty-acid-CoA ligase fadD26 (584 aa)	Rv2930.1	63 kDa	11.3	0.6	6.3
monophosphatase cysQ (268 aa)	Rv2131c.1	28 kDa	1.8	33.5	60.2
30S ribosomal protein S10 rpsJ (102 aa)	Rv0700.1	11 kDa	1.1	118.5	136.0
D-alanine-D-alanine ligase ddlA (374 aa)	Rv2981c.1	40 kDa	3.0	5.4	16.4
polyprenol-monophosphomannose synthase ppm1 (875 aa)	Rv2051c.1	94 kDa	5.0	0.7	3.7
membrane protein secretion factor yajC (116 aa)	Rv2588c.1	13 kDa	6.1	5.8	35.1
phosphomannomutase pmmA (466 aa)	Rv3257c.1	49 kDa	4.8	4.3	20.6
conserved hypothetical protein (65 aa)	Rv2111c.1	7 kDa	1.5	140.8	217.9
cysteine desulfurase csd (418 aa)	Rv1464.1	45 kDa	1.8	5.6	10.2
ribonuclease rphA (260 aa)	Rv1340.1	27 kDa	1.5	9.0	13.5
dehydrogenase FAD flavoprotein gmc oxidoreductase (529 aa)	Rv1279.1	57 kDa	1.5	10.7	15.5
50S ribosomal protein L22 rplV (198 aa)	Rv0706.1	20 kDa	2.2	18.5	40.2
conserved hypothetical protein (155 aa)	Rv3688c.1	17 kDa	2.2	28.0	61.9
50S ribosomal protein L10 rplJ (179 aa)	Rv0651.1	18 kDa	3.0	14.7	43.8
conserved hypothetical protein (127 aa)	Rv0140.1	14 kDa	2.8	31.7	88.8
50S ribosomal protein L11 rplK (143 aa)	Rv0640.1	15 kDa	9.3	4.7	43.9
hypothetical protein (188 aa)	Rv3920c.1	21 kDa	2.9	16.0	45.9
methyltransferase/methylase (246 aa)	Rv2959c.1	28 kDa	#DIV/0!	0.0	16.9
methyltransferase (275 aa)	Rv1405c.1	29 kDa	#DIV/0!	0.0	21.4
conserved hypothetical protein (345 aa)	Rv2159c.1	36 kDa	0.6	614.6	338.1
L-lactate dehydrogenase lldD2 (415 aa)	Rv1872c.1	45 kDa	0.8	239.1	202.6
3-oxoacyl-[acyl-carrier protein] reductase fabG4 (455 aa)	Rv0242c.1	47 kDa	0.9	243.2	213.4
fatty-acid oxidation protein fadB (721 aa)	Rv0860.1	76 kDa	0.6	77.1	49.0
conserved hypothetical protein (455 aa)	Rv3734c.1	49 kDa	0.6	111.3	66.9
conserved membrane protein (332 aa)	Rv0292.1	36 kDa	0.8	108.6	82.6
aminopeptidase pepB (516 aa)	Rv2213.1	53 kDa	0.7	51.0	34.5
succinate-semialdehyde dehydrogenase [NADP+]-dependent gabD2 (519 aa)	Rv1731.1	55 kDa	0.8	62.9	49.8
40 kda secreted L-alanine dehydrogenase ald (372 aa)	Rv2780.1	39 kDa	0.2	57.4	8.7
membrane-anchored mycosin mycP3 (462 aa)	Rv0291.1	46 kDa	0.7	80.0	53.9
conserved hypothetical protein (168 aa)	Rv2468c.1	17 kDa	0.6	276.9	177.9
aldehyde dehydrogenase NAD-dependent (507 aa)	Rv0147.1	55 kDa	0.5	77.5	42.6
DNA gyrase subunit A gyrA (839 aa)	Rv0006.1	92 kDa	0.4	22.6	9.9
acyl-CoA dehydrogenase fadE6 (732 aa)	Rv0271c.1	78 kDa	0.5	38.1	17.3
UDP-N-acetylmuramoylalanyl-D-glutamate-2,6-diaminopimelat E ligase murE (536 aa)	Rv2158c.1	55 kDa	0.4	63.4	22.7
conserved hypothetical protein (116 aa)	Rv3880c.1	12 kDa	0.7	378.2	253.7

acyl-CoA dehydrogenase fadE4 (569 aa)	Rv0231.1	63 kDa	0.2	21.1	3.5
dehydrogenase (460 aa)	Rv2280.1	48 kDa	0.5	46.2	22.3
prolyl-tRNA synthetase proS (583 aa)	Rv2845c.1	63 kDa	0.6	25.7	14.2
integration host factor mihF (191 aa)	Rv1388.1	21 kDa	0.6	80.4	48.1
respiratory nitrate reductase alpha chain narG (1233 aa)	Rv1161.1	137 kDa	0.4	24.8	9.2
pyruvate dehydrogenase E1 component aceE (902 aa)	Rv2241.1	100 kDa	0.4	20.0	8.2
oxidoreductase (294 aa)	Rv2129c.1	31 kDa	0.5	81.8	41.3
conserved hypothetical protein (184 aa)	Rv3867.1	20 kDa	0.6	208.7	123.3
oxidoreductase beta subunit (374 aa)	Rv2454c.1	40 kDa	0.6	51.7	29.7
conserved hypothetical protein (289 aa)	Rv2161c.1	31 kDa	0.6	101.6	64.4
oxidoreductase (308 aa)	Rv1855c.1	33 kDa	0.7	42.1	27.9
transketolase tkt (701 aa)	Rv1449c.1	76 kDa	0.2	30.3	6.4
NADH dehydrogenase I chain G nuoG (807 aa)	Rv3151.1	85 kDa	0.1	22.7	3.0
enoyl-CoA hydratase echA21 (275 aa)	Rv3774.1	29 kDa	0.6	52.1	29.3
succinate dehydrogenase flavoprotein subunit sdhA (591 aa)	Rv3318.1	65 kDa	0.4	24.8	9.5
superoxide dismutase sodA (208 aa)	Rv3846.1	23 kDa	0.8	93.2	72.7
elongation factor G fusA2 (715 aa)	Rv0120c.1	76 kDa	0.4	33.5	12.7
short-chain type dehydrogenase ephD (593 aa)	Rv2214c.1	64 kDa	0.4	27.5	11.7
acetyl-CoA synthetase acs (652 aa)	Rv3667.1	71 kDa	0.5	20.4	10.2
aldehyde dehydrogenase (508 aa)	Rv0458.1	55 kDa	0.5	47.3	25.0
conserved hypothetical protein (152 aa)	Rv3547.1	17 kDa	0.6	136.6	82.0
fatty-acid-CoA ligase fadD6 (598 aa)	Rv1206.1	64 kDa	0.2	40.5	8.7
conserved hypothetical protein (436 aa)	Rv3722c.1	47 kDa	0.0	2.3	0.0
transcription antitermination protein nusG (239 aa)	Rv0639.1	25 kDa	0.8	86.5	69.6
50S ribosomal protein L16 rplP (139 aa)	Rv0708.1	16 kDa	0.7	60.3	39.3
cation-transporter ATPase I ctpI (1633 aa)	Rv0107c.1	170 kDa	0.2	2.8	0.5
coenzyme F420-dependent oxidoreductase (348 aa)	Rv3520c.1	38 kDa	0.4	12.5	4.8
anion transporter ATPase (387 aa)	Rv3680.1	41 kDa	0.6	26.7	15.3
lipoprotein lppZ (374 aa)	Rv3006.1	39 kDa	0.5	48.9	24.3
ATP-dependent DNA helicase helY (907 aa)	Rv2092c.1	100 kDa	0.1	4.0	0.3
GDP-D-mannose dehydratase gmdA (341 aa)	Rv1511.1	38 kDa	0.0	9.7	0.0
NADP-dependent alcohol dehydrogenase adhC (347 aa)	Rv3045.1	37 kDa	0.2	33.6	8.2
conserved membrane protein (473 aa)	Rv0290.1	48 kDa	0.7	29.5	20.8
dihydrodipicolinate reductase dapB (246 aa)	Rv2773c.1	26 kDa	0.6	75.9	42.9
conserved hypothetical protein (215 aa)	Rv3241c.1	25 kDa	0.5	43.5	21.6

long-chain fatty-acid-CoA ligase fadD12 (536 aa)	Rv1427c.1	58 kDa	0.2	29.9	6.8
periplasmic superoxide dismutase sodC (241 aa)	Rv0432.1	24 kDa	0.5	60.9	30.6
DNA/pantothenate metabolism flavoprotein dfp (419 aa)	Rv1391.1	44 kDa	0.4	18.3	7.7
conserved hypothetical protein (436 aa)	Rv1232c.1	47 kDa	0.3	35.1	8.9
transmembrane carbonic anhydrase (765 aa)	Rv3273.1	81 kDa	0.3	17.0	4.3
iron-sulfur-binding reductase (883 aa)	Rv0338c.1	95 kDa	0.1	4.9	0.4
hypothetical alanine and valine rich protein (296 aa)	Rv3090.1	32 kDa	0.4	58.7	26.2
acyl-CoA dehydrogenase fadE1 (448 aa)	Rv0131c.1	50 kDa	0.2	6.0	1.3
proline rich membrane-anchored mycosin mycP5 (586 aa)	Rv1796.1	60 kDa	0.2	19.9	3.5
conserved membrane protein (661 aa)	Rv2345.1	70 kDa	0.4	17.6	7.0
conserved hypothetical protein (132 aa)	Rv0523c.1	15 kDa	0.3	89.0	31.0
oxidoreductase (461 aa)	Rv1751.1	50 kDa	0.4	10.1	4.1
acetolactate synthase small subunit ilvN (169 aa)	Rv3002c.1	18 kDa	0.3	29.7	8.1
enoyl-CoA hydratase echA16 (250 aa)	Rv2831.1	27 kDa	0.3	46.6	11.7
conserved alanine and leucine rich protein (885 aa)	Rv2567.1	95 kDa	0.1	7.9	1.0
amino acid decarboxylase (948 aa)	Rv2531c.1	106 kDa	0.2	2.8	0.6
polyphosphate kinase ppk (743 aa)	Rv2984.1	83 kDa	0.2	8.6	1.5
fatty-acid-CoA ligase fadD13 (504 aa)	Rv3089.1	54 kDa	0.2	16.5	2.9
conserved hypothetical protein (326 aa)	Rv2410c.1	36 kDa	0.6	27.6	15.7
transmembrane cytochrome C oxidase subunit II ctaC (364 aa)	Rv2200c.1	40 kDa	0.3	22.9	6.1
dehydrogenase (398 aa)	Rv3726.1	42 kDa	0.4	27.6	10.3
30S ribosomal protein S9 rpsI (152 aa)	Rv3442c.1	16 kDa	0.1	65.9	3.4
peroxidoxin bcpB (155 aa)	Rv1608c.1	17 kDa	0.5	36.6	19.7
conserved hypothetical protein (158 aa)	Rv0857.1	18 kDa	0.3	52.7	15.5
30S ribosomal protein S15 rpsO (90 aa)	Rv2785c.1	10 kDa	0.2	149.4	36.4
restriction system protein mrr (307 aa)	Rv2528c.1	34 kDa	0.2	46.8	11.5
inorganic pyrophosphatase ppa (163 aa)	Rv3628.1	18 kDa	0.5	100.6	48.9
lipoprotein lprF (262 aa)	Rv1368.1	27 kDa	0.5	72.6	36.5
ubiquinol-cytochrome C reductase qcrC cytochrome C subunit (281 aa)	Rv2194.1	29 kDa	0.5	29.7	15.1
conserved hypothetical protein (473 aa)	Rv3087.1	53 kDa	0.0	23.3	1.1
cholesterol oxidase precursor choD (579 aa)	Rv3409c.1	63 kDa	0.4	14.4	6.5
NADH dehydrogenase I chain D nuoD (441 aa)	Rv3148.1	48 kDa	0.1	18.1	1.2
transmembrane serine/threonine-protein kinase B pknB (627 aa)	Rv0014c.1	67 kDa	0.1	12.2	1.7
30S ribosomal protein S13 rpsM (125 aa)	Rv3460c.1	14 kDa	0.0	48.1	2.3

conserved hypothetical protein (225 aa)	Rv2557.1	24 kDa	0.1	66.4	9.8
transposase (571 aa)	Rv3327.1	63 kDa	0.1	17.5	2.3
phosphoglucomutase pgmA (548 aa)	Rv3068c.1	58 kDa	0.0	2.4	0.0
conserved hypothetical protein (326 aa)	Rv1021.1	35 kDa	0.7	22.5	15.6
NADH dehydrogenase I chain C nuoC (237 aa)	Rv3147.1	27 kDa	0.5	45.2	22.7
phosphoribosylformylglycinamidine cyclo-ligase purM (365 aa)	Rv0809.1	38 kDa	0.1	14.6	1.6
esterase lipM (432 aa)	Rv2284.1	47 kDa	0.2	11.6	2.0
arylsulfatase atsA (788 aa)	Rv0711.1	86 kDa	0.0	7.2	0.0
riboflavin biosynthesis protein ribG (340 aa)	Rv1409.1	35 kDa	0.1	13.8	0.8
oxpp cycle protein opcA (304 aa)	Rv1446c.1	33 kDa	0.1	26.5	2.8
transmembrane serine/threonine-protein kinase E pknE (567 aa)	Rv1743.1	61 kDa	0.0	5.3	0.0
phosphoserine phosphatase serB2 (410 aa)	Rv3042c.1	43 kDa	0.0	8.2	0.0
tyrosyl-tRNA synthase tyrS (425 aa)	Rv1689.1	46 kDa	0.6	6.3	3.7
hypothetical protein (430 aa)	Rv2305.1	46 kDa	0.5	14.7	8.1
conserved membrane protein (329 aa)	Rv0347.1	37 kDa	0.1	9.1	0.7
membrane-associated phospholipase C 2 plcB (513 aa)	Rv2350c.1	56 kDa	0.3	24.0	6.5
conserved hypothetical protein (373 aa)	Rv0906.1	41 kDa	0.4	13.4	4.7
acyl-CoA dehydrogenase fadE7 (396 aa)	Rv0400c.1	42 kDa	0.0	5.8	0.0
epoxide hydrolase ephA (323 aa)	Rv3617.1	35 kDa	0.5	26.8	14.4
transcriptional regulator, tetR-family (222 aa)	Rv3295.1	25 kDa	0.1	31.8	2.4
MCE-family protein mce1F (516 aa)	Rv0174.1	54 kDa	0.3	25.9	6.5
conserved hypothetical protein (403 aa)	Rv3864.1	42 kDa	0.4	24.0	9.5
oxidoreductase (318 aa)	Rv2263.1	34 kDa	0.1	10.5	0.8
MCE-family protein mce1C (516 aa)	Rv0171.1	54 kDa	0.2	16.9	3.1
conserved hypothetical protein (234 aa)	Rv3702c.1	25 kDa	0.3	29.8	7.6
quinone oxidoreductase (323 aa)	Rv0149.1	33 kDa	0.0	8.2	0.0
transcriptional regulator (141 aa)	Rv0474.1	15 kDa	0.6	68.9	41.0
phosphate acetyltransferase pta (691 aa)	Rv0408.1	73 kDa	0.4	7.2	3.2
mycolic acid synthase pcaA (288 aa)	Rv0470c.1	33 kDa	0.0	6.9	0.0
conserved hypothetical protein (475 aa)	Rv3088.1	51 kDa	0.0	18.9	0.5
penicillin-binding protein pbpA (492 aa)	Rv0016c.1	52 kDa	0.0	6.1	0.0
conserved hypothetical protein (470 aa)	Rv0221.1	52 kDa	0.0	9.2	0.0
mycothiol conjugate amidase mca (289 aa)	Rv1082.1	33 kDa	0.0	3.5	0.0
lipase lipD (447 aa)	Rv1923.1	47 kDa	0.4	16.4	6.6
nitrogen regulatory protein P-II glnB (113 aa)	Rv2919c.1	12 kDa	0.3	80.2	28.0
transcriptional regulator, marR-family (209 aa)	Rv0042c.1	22 kDa	0.2	24.0	3.6
oxidoreductase (462 aa)	Rv3790.1	50 kDa	0.2	12.3	2.3

cutinase cut2 (231 aa)	Rv2301.1	24 kDa	0.6	26.0	15.3
moaD-moaE fusion protein moaX (222 aa)	Rv3323c.1	24 kDa	0.4	39.1	14.7
conserved hypothetical protein (329 aa)	Rv0496.1	35 kDa	0.4	12.1	4.8
zinc-type alcohol dehydrogenase NAD dependent adhB (376 aa)	Rv0761c.1	40 kDa	0.0	15.1	0.0
30S ribosomal protein S19 rpsS (94 aa)	Rv0705.1	11 kDa	0.1	74.8	5.1
invasion-associated protein (231 aa)	Rv1566c.1	24 kDa	0.0	31.7	1.1
isocitrate dehydrogenase icd1 (410 aa)	Rv3339c.1	46 kDa	0.0	3.2	0.0
conserved lipoprotein (558 aa)	Rv2585c.1	59 kDa	0.2	7.1	1.6
phospholipase C 3 plcC (509 aa)	Rv2349c.1	55 kDa	0.0	7.9	0.0
dTDP-rha:a-D-glcnac-diphosphoryl polyprenol, a-3-L-rhamnosyl transferase wbbL1 (302 aa)	Rv3265c.1	33 kDa	0.0	5.5	0.0
PE family protein (99 aa)	Rv3477.1	10 kDa	0.4	67.7	29.1
adenylyl cyclase (398 aa)	Rv1264.1	42 kDa	0.1	5.9	0.7
PPE family protein (540 aa)	Rv1387.1	55 kDa	0.1	8.0	0.9
conserved secreted protein (113 aa)	Rv0559c.1	12 kDa	0.2	50.3	12.2
O-succinylhomoserine sulfhydrylase metZ (407 aa)	Rv0391.1	43 kDa	0.1	4.9	0.7
alternative RNA polymerase sigma factor sigJ (313 aa)	Rv3328c.1	34 kDa	0.2	11.8	2.5
conserved hypothetical protein (308 aa)	Rv3075c.1	33 kDa	0.0	6.5	0.0
conserved hypothetical protein (167 aa)	Rv2616.1	18 kDa	0.1	28.0	3.0
conserved hypothetical protein (211 aa)	Rv2680.1	23 kDa	0.0	20.6	0.0
lipid transfer protein or keto acyl-CoA thiolase ltp4 (355 aa)	Rv3522.1	38 kDa	0.0	12.2	0.0
ferric uptake regulation protein furA (151 aa)	Rv1909c.1	16 kDa	0.0	15.6	0.0
outer membrane protein A ompA (327 aa)	Rv0899.1	34 kDa	0.2	7.2	1.5
conserved hypothetical protein (94 aa)	Rv3269.1	10 kDa	0.1	42.7	2.8
hydrolase (345 aa)	Rv1333.1	34 kDa	0.1	10.6	0.9
methylated-DNA-protein-cysteine methyltransferase ogt (166 aa)	Rv1316c.1	18 kDa	0.0	20.3	0.0
hypothetical exported protein (137 aa)	Rv0203.1	14 kDa	0.1	31.7	2.1
conserved hypothetical protein (130 aa)	Rv3541c.1	14 kDa	0.2	12.9	2.0

Appendix Table 3. Significantly different proteins in the CFP of the clonal laboratory pair

Identified Proteins p<0.05 (t-test)	Accession Number	Molecular Weight	Fold Change (INHs/ INHr)	NSAF Average lab INHr x 1000	Average lab INHs x 1000
catalase-peroxidase-peroxynitritase T katG (741 aa)	Rv1908c.1	81 kDa	2.69	68.30	183.78
5-methyltetrahydropteroyltriglutamate-homocysteine methyltransferase metE (760 aa)	Rv1133c.1	82 kDa	1.68	26.55	44.71
ATP-dependent clp protease proteolytic subunit 2 clpP2 (215 aa)	Rv2460c.1	24 kDa	1.88	23.60	44.34
iron-dependent repressor and activator ideR (231 aa)	Rv2711.1	25 kDa	1.17	75.09	87.61
peroxiredoxin ahpE (154 aa)	Rv2238c.1	17 kDa	3.34	8.76	29.22
conserved hypothetical protein (98 aa)	Rv0250c.1	11 kDa	1.26	117.88	148.50
cutinase cut2 (231 aa)	Rv2301.1	24 kDa	1.31	40.76	53.41
diaminopimelate decarboxylase lysA (448 aa)	Rv1293.1	47 kDa	1.67	10.18	16.99
hypothetical glycine rich protein (144 aa)	Rv2307B.1	15 kDa	2.13	16.78	35.76
conserved alanine and arginine rich protein (451 aa)	Rv2731.1	50 kDa	2.05	8.77	18.01
conserved hypothetical protein (133 aa)	Rv3849.1	15 kDa	1.66	29.85	49.47
dehydrogenase/reductase (389 aa)	Rv0331.1	41 kDa	1.26	15.81	19.95
fructose-bisphosphate aldolase fba (345 aa)	Rv0363c.1	37 kDa	0.90	285.12	255.67
lipoprotein lprA (245 aa)	Rv1270c.1	25 kDa	0.86	131.15	113.26
conserved hypothetical protein (179 aa)	Rv3780.1	20 kDa	0.90	182.55	163.66
50S ribosomal protein L9 rplI (153 aa)	Rv0056.1	16 kDa	0.72	102.52	73.71
ATP-dependent DNA helicase helY (907 aa)	Rv2092c.1	100 kDa	0.17	6.44	1.12
enoyl-CoA hydratase echA15 (277 aa)	Rv2679.1	30 kDa	0.83	79.65	65.72
conserved hypothetical protein (148 aa)	Rv0854.1	16 kDa	0.63	150.52	95.34
ATP-dependent protease ATP-binding subunit clpC1 (849 aa)	Rv3596c.1	94 kDa	0.10	3.32	0.32
acetyl/propionyl-CoA carboxylase beta subunit accD6 (474 aa)	Rv2247.1	50 kDa	0.00	2.72	0.00
hypothetical exported protein (198 aa)	Rv1910c.1	20 kDa	0.75	47.53	35.82
dihydrolipoamide S-acetyltransferase E2 component pdhC (394 aa)	Rv2495c.1	41 kDa	0.39	12.03	4.63
glutamyl-tRNA(gln) amidotransferase subunit C gatC (100 aa)	Rv3012c.1	11 kDa	0.59	58.89	34.57
short-chain type dehydrogenase/reductase (259 aa)	Rv2857c.1	27 kDa	0.25	20.21	4.97

Appendix Table 4. Significantly different proteins in the CFP of the clonal clinical pair

Identified Proteins p<0.05 (t-test)	Accession Number	Molecular Weight	Fold Change (INHs/ INHr)	NSAF Average clinical INHr x 1000	Average clinical INHs x 1000
catalase-peroxidase-peroxynitritase T katG (741 aa)	Rv1908c.1	81 kDa	25.1	14.7	370.4
adenosylhomocysteinase sahH (496 aa)	Rv3248c.1	54 kDa	1.2	136.6	162.0
bacterioferritin bfrB (182 aa)	Rv3841.1	20 kDa	1.8	487.8	869.4
fructose-bisphosphate aldolase fba (345 aa)	Rv0363c.1	37 kDa	2.3	40.2	92.5
inorganic pyrophosphatase ppa (163 aa)	Rv3628.1	18 kDa	3.6	136.2	485.3
conserved alanine rich protein (271 aa)	Rv2744c.1	29 kDa	7.3	28.6	208.5
conserved hypothetical protein (787 aa)	Rv3401.1	87 kDa	4.4	27.4	121.3
bacterioferritin bfrA (160 aa)	Rv1876.1	18 kDa	2.2	350.3	753.5
50S ribosomal protein L7/L12 rplL (131 aa)	Rv0652.1	13 kDa	1.4	285.4	398.9
ferredoxin-dependent nitrite reductase SirA (564 aa)	Rv2391.1	63 kDa	5.1	17.7	91.2
ATP phosphoribosyltransferase hisG (285 aa)	Rv2121c.1	30 kDa	3.4	48.2	166.1
enoyl-CoA hydratase echA9 (346 aa)	Rv1071c.1	36 kDa	5.3	31.3	164.7
thioredoxin trxB1 (124 aa)	Rv1471.1	13 kDa	6.1	56.3	341.2
oxidoreductase (283 aa)	Rv2971.1	30 kDa	6.1	9.4	57.2
pyruvate kinase pykA (473 aa)	Rv1617.1	51 kDa	5.0	11.2	56.5
succinyl-CoA:3-ketoacid-coenzyme A transferase alpha subunit scoA (249 aa)	Rv2504c.1	26 kDa	#DIV/0!	0.0	38.7
50S ribosomal protein L4 rplD (224 aa)	Rv0702.1	24 kDa	9.3	8.7	80.8
conserved hypothetical protein (308 aa)	Rv3075c.1	33 kDa	9.4	6.0	56.4
50S ribosomal protein L10 rplJ (179 aa)	Rv0651.1	18 kDa	6.3	31.1	195.6
conserved hypothetical protein (149 aa)	Rv1558.1	16 kDa	22.2	6.0	132.1
phosphoglycerate mutase gpm2 (204 aa)	Rv3214.1	22 kDa	2.9	31.8	92.2
conserved hypothetical protein (185 aa)	Rv3614c.1	20 kDa	#DIV/0!	0.0	28.1
glycogen phosphorylase glgP (864 aa)	Rv1328.1	96 kDa	6.6	3.1	20.3
alkyl hydroperoxide reductase D protein ahpD (178 aa)	Rv2429.1	19 kDa	14.7	10.0	146.8
acyl-CoA dehydrogenase fadE7 (396 aa)	Rv0400c.1	42 kDa	#DIV/0!	0.0	66.5
conserved hypothetical protein (286 aa)	Rv3463.1	31 kDa	#DIV/0!	0.0	33.5
dihydroxy-acid dehydratase ilvD (576 aa)	Rv0189c.1	59 kDa	21.1	2.9	60.6
hypothetical protein (84 aa)	Rv0634A.1	9 kDa	12.0	10.6	127.1
30S ribosomal protein S7 rpsG (157 aa)	Rv0683.1	18 kDa	#DIV/0!	0.0	142.5
acyl-CoA dehydrogenase fadE25 (390 aa)	Rv3274c.1	42 kDa	18.1	2.3	41.0
hypothetical protein (55 aa)	Rv3489.1	6 kDa	6.4	30.7	195.4
conserved hypothetical protein (147 aa)	Rv2175c.1	16 kDa	#DIV/0!	0.0	42.0
F420-dependent glucose-6-phosphate dehydrogenase fgd1 (337 aa)	Rv0407.1	37 kDa	4.6	18.6	85.1
thioredoxin reductase trxB2 (336 aa)	Rv3913.1	36 kDa	5.3	7.9	42.2
conserved hypothetical protein (449 aa)	Rv2141c.1	48 kDa	#DIV/0!	0.0	3.3
50S ribosomal protein L30 rpmD (66 aa)	Rv0722.1	7 kDa	#DIV/0!	0.0	65.2
heat shock protein hsp (160 aa)	Rv0251c.1	18 kDa	14.2	6.6	93.2
glutamate-1-semialdehyde 2,1-aminomutase hemL (463 aa)	Rv0524.1	48 kDa	#DIV/0!	0.0	18.8
30S ribosomal protein S17 rpsQ (137 aa)	Rv0710.1	15 kDa	#DIV/0!	0.0	53.5
ribonuclease rphA (260 aa)	Rv1340.1	27 kDa	7.7	3.7	28.8
dihydrolipoamide S-acetyltransferase E2 component BdkC(394 aa)	Rv2495c.1	41 kDa	#DIV/0!	0.0	13.4

conserved hypothetical protein (208 aa)	Rv1700.1	23 kDa	#DIV/0!	0.0	61.3
conserved hypothetical protein (284 aa)	Rv3099c.1	31 kDa	#DIV/0!	0.0	28.6
preprotein translocase secE1 (162 aa)	Rv0638.1	17 kDa	#DIV/0!	0.0	59.1
heat shock protein hspX (145 aa)	Rv2031c.1	16 kDa	0.4	1309.5	466.0
immunogenic protein mpt64 (229 aa)	Rv1980c.1	25 kDa	0.6	593.6	363.0
low molecular weight protein antigen cfp2 (169 aa)	Rv2376c.1	17 kDa	0.2	1973.5	421.9
alanine and proline rich secreted protein apa (326 aa)	Rv1860.1	33 kDa	0.4	566.8	239.5
enolase eno (430 aa)	Rv1023.1	45 kDa	0.7	245.1	165.1
conserved hypothetical protein (528 aa)	Rv0020c.1	57 kDa	0.3	112.1	35.5
10 kda culture filtrate antigen esxB (101 aa)	Rv3874.1	11 kDa	0.4	1805.9	656.6
Esat-6 like protein esxP (99 aa)	Rv2347c.1	11 kDa	0.2	1243.7	288.3
6 kda early secretory antigenic target esxA (96 aa)	Rv3875.1	10 kDa	0.5	703.7	342.8
conserved alanine and glycine rich protein (393 aa)	Rv3616c.1	40 kDa	0.4	276.2	122.3
conserved hypothetical protein (91 aa)	Rv3208A.1	9 kDa	0.2	568.5	91.0
soluble secreted antigen mpt53 precursor (174 aa)	Rv2878c.1	18 kDa	0.3	310.8	83.0
protein transport protein secE2 (72 aa)	Rv0379.1	8 kDa	0.2	731.6	119.3
conserved hypothetical protein (157 aa)	Rv1906c.1	16 kDa	0.4	341.0	134.1
bifunctional membrane-associated penicillin-binding protein 1A/1B ponA2 (811 aa)	Rv3682.1	85 kDa	0.4	47.3	17.0
3-isopropylmalate dehydrogenase leuB (337 aa)	Rv2995c.1	35 kDa	0.2	165.5	29.0
conserved hypothetical protein (184 aa)	Rv3867.1	20 kDa	0.1	55.1	6.0
conserved hypothetical protein (215 aa)	Rv3705c.1	22 kDa	0.2	147.1	30.7
acetyl-CoA acetyltransferase fadA6 (387 aa)	Rv3556c.1	41 kDa	0.2	22.4	5.4
CpsA1 (499 aa)	Rv3267.1	51 kDa	0.2	74.3	18.0
conserved hypothetical protein (527 aa)	Rv1130.1	58 kDa	0.1	81.9	6.2
hypothetical protein (381 aa)	Rv1887.1	40 kDa	0.2	104.5	20.5
invasion-associated protein (473 aa)	Rv1477.1	50 kDa	0.2	65.2	9.9
conserved hypothetical protein (80 aa)	Rv0787A.1	9 kDa	0.4	154.2	55.0
conserved hypothetical protein (86 aa)	Rv1638A.1	10 kDa	0.7	179.6	119.8
hypothetical exported protein (198 aa)	Rv1910c.1	20 kDa	0.2	130.2	26.7
hypothetical protein (183 aa)	Rv3033.1	19 kDa	0.4	46.6	20.4
dihydrodipicolinate reductase dapB (246 aa)	Rv2773c.1	26 kDa	0.1	166.7	19.0
MCE-family protein mce1F (516 aa)	Rv0174.1	54 kDa	0.2	54.9	9.2
conserved hypothetical protein (685 aa)	Rv0822c.1	73 kDa	0.2	15.0	3.3
MCE-associated membrane protein (245 aa)	Rv0178.1	26 kDa	0.3	60.3	16.2
conserved hypothetical protein (403 aa)	Rv3864.1	42 kDa	0.4	41.8	17.2
conserved hypothetical protein (386 aa)	Rv2190c.1	40 kDa	0.1	32.3	2.1
acetyl-CoA acetyltransferase fadA5 (392 aa)	Rv3546.1	41 kDa	0.0	37.7	0.0
conserved hypothetical protein (242 aa)	Rv3717.1	25 kDa	0.0	47.1	0.0
conserved secreted protein (114 aa)	Rv1271c.1	12 kDa	0.0	220.1	0.0

Appendix Table 5. Significantly different proteins in the Cytosol of the clonal laboratory pair

Identified Proteins p<0.05 (t-test)	Accession Number	Molecular Weight	Fold Change (INHs/ INHr)	NSAF Average lab INHr x 1000	Average lab INHs x 1000
heat shock protein hspX (145 aa)	Rv2031c.1	16 kDa	1.27	1511.70	1927.17
catalase-peroxidase-peroxynitritase T katG (741 aa)	Rv1908c.1	81 kDa	3.08	92.36	284.09
adenosylhomocysteinase sahH (496 aa)	Rv3248c.1	54 kDa	1.13	197.02	222.23
acetohydroxyacid synthase ilvX (516 aa)	Rv3509c.1	52 kDa	1.87	52.50	98.37
secreted fibronectin-binding protein antigen fbpA (339 aa)	Rv3804c.1	36 kDa	1.30	143.69	186.89
conserved hypothetical protein (298 aa)	Rv2623.1	32 kDa	1.57	167.82	263.52
quinone reductase qor (329 aa)	Rv1454c.1	34 kDa	1.76	74.39	130.87
cyclopropane-fatty-acyl-phospholipid synthase 2 cmaA2 (303 aa)	Rv0503c.1	35 kDa	1.78	46.41	82.75
nicotinate-nucleotide pyrophosphatase nadC (286 aa)	Rv1596.1	30 kDa	1.59	54.40	86.76
glutamine synthetase glnA2 (447 aa)	Rv2222c.1	50 kDa	1.20	53.76	64.70
nucleoside diphosphate kinase ndkA (137 aa)	Rv2445c.1	15 kDa	1.34	137.36	183.89
50S ribosomal protein L3 rplC (218 aa)	Rv0701.1	23 kDa	1.58	80.97	127.56
conserved hypothetical protein (95 aa)	Rv1738.1	11 kDa	1.17	217.08	253.11
phosphoribosylformimino-5-aminoimidazole carboxamide ribotide isomerase hisA (246 aa)	Rv1603.1	26 kDa	1.67	38.13	63.51
conserved hypothetical protein (73 aa)	Rv1893.1	7 kDa	3.20	78.89	252.30
extragenic suppressor protein suhB (291 aa)	Rv2701c.1	30 kDa	3.30	8.68	28.61
conserved alanine, glycine and valine rich protein (260 aa)	Rv2696c.1	27 kDa	1.38	17.90	24.62
methoxy mycolic acid synthase 1 mmaA1 (287 aa)	Rv0645c.1	33 kDa	1.69	11.71	19.81
deoxycytidine triphosphate deaminase dcd (191 aa)	Rv0321.1	21 kDa	1.58	26.35	41.55
conserved hypothetical protein (148 aa)	Rv1155.1	16 kDa	3.56	9.33	33.20
fatty-acid oxidation protein fadB (721 aa)	Rv0860.1	76 kDa	0.88	198.99	175.16
60 kda chaperonin 1 groEL1 (540 aa)	Rv3417c.1	56 kDa	0.89	307.75	272.81
pyruvate dehydrogenase E2 component DlaT (554 aa)	Rv2215.1	57 kDa	0.84	265.98	222.24
DNA-directed RNA polymerase alpha chain rpoA (348 aa)	Rv3457c.1	38 kDa	0.93	309.72	288.62
conserved hypothetical protein (632 aa)	Rv0282.1	68 kDa	0.42	177.52	73.87
trigger factor protein tig (467 aa)	Rv2462c.1	51 kDa	0.72	291.49	210.10
chaperone protein htpG (648 aa)	Rv2299c.1	73 kDa	0.80	183.83	147.74
polyketide synthase pks13 (1734 aa)	Rv3800c.1	186 kDa	0.65	61.47	39.89
fatty-acid-CoA ligase fadD32 (638 aa)	Rv3801c.1	69 kDa	0.19	122.26	23.74
alkyl hydroperoxide reductase C protein ahpC (196 aa)	Rv2428.1	22 kDa	0.75	644.69	482.24
conserved hypothetical protein (574 aa)	Rv3868.1	62 kDa	0.47	47.92	22.55
conserved alanine and arginine rich protein (451 aa)	Rv2731.1	50 kDa	0.68	68.92	46.56
30S ribosomal protein S4 rpsD (202 aa)	Rv3458c.1	23 kDa	0.78	151.42	118.75

methoxy mycolic acid synthase 2 mmaA2 (288 aa)	Rv0644c.1	33 kDa	0.65	37.64	24.51
long-chain fatty-acid-CoA ligase fadD15 (601 aa)	Rv2187.1	64 kDa	0.50	30.34	15.16
transcriptional regulator embR (389 aa)	Rv1267c.1	42 kDa	0.18	38.15	7.02
conserved hypothetical protein (243 aa)	Rv0730.1	26 kDa	0.57	98.90	56.05
conserved hypothetical protein (246 aa)	Rv2927c.1	27 kDa	0.73	112.01	81.75
conserved alanine rich protein (271 aa)	Rv2744c.1	29 kDa	0.32	93.70	30.08
fatty-acid-CoA ligase fadD36 (474 aa)	Rv1193.1	50 kDa	0.56	48.87	27.53
RNA polymerase sigma factor sigA (529 aa)	Rv2703.1	58 kDa	0.53	46.86	24.86
2-hydroxy-6-oxo-6-phenylhexa-2,4-dienoate hydrolase bphD (292 aa)	Rv3569c.1	32 kDa	0.51	34.95	17.77
preprotein translocase ATPase secA2 (809 aa)	Rv1821.1	89 kDa	0.14	12.56	1.72
pantoate-beta-alanine ligase panC (310 aa)	Rv3602c.1	33 kDa	0.00	60.59	0.00
conserved hypothetical protein (141 aa)	Rv0390.1	15 kDa	0.79	114.27	90.84
conserved transmembrane ATP-binding protein ABC transporter (866 aa)	Rv1747.1	92 kDa	0.11	3.68	0.40
zinc metalloprotease (664 aa)	Rv0198c.1	74 kDa	0.54	13.91	7.49
cell division protein ftsZ (380 aa)	Rv2150c.1	39 kDa	0.08	35.49	2.73
peptide methionine sulfoxide reductase msrA (183 aa)	Rv0137c.1	20 kDa	0.17	44.84	7.47
two component system transcriptional regulator phoP (248 aa)	Rv0757.1	28 kDa	0.23	30.39	7.07
lipoprotein lppD (344 aa)	Rv1899c.1	35 kDa	0.22	24.65	5.39
3-hydroxyisobutyrate dehydrogenase mmsB (295 aa)	Rv0751c.1	30 kDa	0.41	17.61	7.25
oxidoreductase (373 aa)	Rv1260.1	41 kDa	0.04	26.18	0.92
transcriptional regulator (260 aa)	Rv1719.1	28 kDa	0.08	34.91	2.64
conserved hypothetical protein (378 aa)	Rv1367c.1	41 kDa	0.22	4.21	0.91
conserved hypothetical protein (304 aa)	Rv1896c.1	33 kDa	0.41	25.69	10.61
cyclopropane-fatty-acyl-phospholipid synthase 1 cmaA1 (288 aa)	Rv3392c.1	32 kDa	0.28	13.03	3.60
conserved hypothetical protein (316 aa)	Rv0819.1	34 kDa	0.04	28.63	1.11
conserved hypothetical protein (268 aa)	Rv0295c.1	30 kDa	0.00	5.26	0.00
precorrin-8x methylmutase cobH (209 aa)	Rv2065.1	22 kDa	0.20	16.40	3.29
thiosulfate sulfurtransferase sseB (285 aa)	Rv2291.1	29 kDa	0.00	41.60	0.00
histidinol-phosphate aminotransferase hisC1 (381 aa)	Rv1600.1	41 kDa	0.28	13.42	3.81
acyl-CoA dehydrogenase fadE19 (395 aa)	Rv2500c.1	42 kDa	0.11	15.00	1.72
phosphoribosylamine-glycine ligase purD (423 aa)	Rv0772.1	44 kDa	0.00	3.42	0.00
protease (233 aa)	Rv3668c.1	23 kDa	0.12	12.05	1.50

Appendix Table 6. Significantly different proteins in the Cytosol of the clonal clinical pair

Identified Proteins p<0.05 (t-test)	Accession Number	Molecular Weight	Fold Change (INHs/ INHr)	NSAF Average clinical INHr x 1000	Average clinical INHs x 1000
endopeptidase ATP binding protein chain B clpB (849 aa)	Rv0384c.1	93 kDa	1.41	171.90	241.56
catalase-peroxidase-peroxynitritase T katG (741 aa)	Rv1908c.1	81 kDa	4.54	79.67	362.00
bifunctional polyribonucleotide nucleotidyltransferase gpsI (753 aa)	Rv2783c.1	80 kDa	1.34	162.80	217.55
S-adenosylmethionine synthetase metK (404 aa)	Rv1392.1	43 kDa	1.17	381.12	447.01
meromycolate extension acyl carrier protein acpM (116 aa)	Rv2244.1	13 kDa	1.35	872.75	1174.20
thiosulfate sulfurtransferase cysA2 (278 aa)	Rv0815c.1 (+1)	31 kDa	1.71	217.69	372.78
aminopeptidase N pepN (862 aa)	Rv2467.1	94 kDa	1.90	71.33	135.73
pyruvate kinase pykA (473 aa)	Rv1617.1	51 kDa	1.14	143.98	164.08
bifunctional acetyl-/propionyl-coenzyme A carboxylase alpha chain accA3 (601 aa)	Rv3285.1	64 kDa	1.42	137.82	196.27
iron-regulated phosphoenolpyruvate carboxykinase pckA (607 aa)	Rv0211.1	67 kDa	1.40	78.59	110.00
fatty-acid-CoA ligase fadD32 (638 aa)	Rv3801c.1	69 kDa	1.33	141.19	187.56
alkyl hydroperoxide reductase C protein ahpC (196 aa)	Rv2428.1	22 kDa	2.46	77.96	192.12
conserved hypothetical protein (208 aa)	Rv2466c.1	23 kDa	1.71	173.48	296.26
2-isopropylmalate synthase leuA (645 aa)	Rv3710.1	70 kDa	1.75	43.65	76.37
thioredoxin reductase trxB2 (336 aa)	Rv3913.1	36 kDa	2.16	78.21	169.27
methylmalonate-semialdehyde dehydrogenase mmsA (511 aa)	Rv0753c.1	54 kDa	2.15	47.62	102.25
conserved hypothetical protein (436 aa)	Rv3722c.1	47 kDa	1.52	66.22	100.86
O-acetylhomoserine sulfhydrylase metC (450 aa)	Rv3340.1	47 kDa	2.28	63.93	146.00
conserved hypothetical protein (302 aa)	Rv2172c.1	33 kDa	2.89	49.30	142.68
succinate dehydrogenase iron-sulfur subunit (647 aa)	Rv0248c.1	71 kDa	1.68	20.11	33.69
multifunctional mycocerosic acid synthase membrane-associated mas (2112 aa)	Rv2940c.1	224 kDa	#DIV/0!	0.00	2.84
conserved hypothetical protein (134 aa)	Rv3716c.1	13 kDa	1.30	161.63	210.25
NADH-dependent enoyl-[acyl-carrier-protein] reductase inhA (270 aa)	Rv1484.1	29 kDa	1.27	81.10	103.39
30S ribosomal protein S5 rpsE (221 aa)	Rv0721.1	23 kDa	1.64	77.12	126.67
conserved hypothetical protein (98 aa)	Rv0250c.1	11 kDa	1.63	115.86	188.58
thioredoxin trxB1 (124 aa)	Rv1471.1	13 kDa	1.81	140.21	254.09
aminopeptidase pepB (516 aa)	Rv2213.1	53 kDa	2.40	18.84	45.23
aminotransferase (391 aa)	Rv0075.1	43 kDa	1.66	53.25	88.63
isocitrate lyase icl (429 aa)	Rv0467.1	47 kDa	1.89	34.21	64.68
beta-1,3-glucanase precursor (295 aa)	Rv0315.1	32 kDa	1.94	36.76	71.21
glucanase glgE (702 aa)	Rv1327c.1	79 kDa	2.28	21.42	48.86
alkyl hydroperoxide reductase D protein ahpD (178 aa)	Rv2429.1	19 kDa	15.43	5.67	87.53
alanine rich hydrolase (246 aa)	Rv2765.1	26 kDa	1.35	42.17	57.00
serine protease pepD (465 aa)	Rv0983.1	46 kDa	1.52	26.39	40.09

Protein	Rv	kDa	Ratio	Value 1	Value 2
acyl-CoA dehydrogenase fadE35 (594 aa)	Rv3797.1	65 kDa	1.46	33.14	48.41
acetyl-CoA acetyltransferase fadA6 (387 aa)	Rv3556c.1	41 kDa	1.48	46.17	68.13
fatty-acid-CoA ligase fadD25 (584 aa)	Rv1521.1	63 kDa	2.01	20.15	40.42
methionine aminopeptidase mapB (286 aa)	Rv2861c.1	31 kDa	2.42	18.26	44.16
lipoprotein lppX (234 aa)	Rv2945c.1	24 kDa	1.87	33.50	62.69
preprotein translocase ATPase secA2 (809 aa)	Rv1821.1	89 kDa	7.40	1.24	9.19
cystathionine beta-synthase cbs (465 aa)	Rv1077.1	49 kDa	2.82	7.04	19.83
glycerol kinase glpK (518 aa)	Rv3696c.1	56 kDa	3.29	8.36	27.50
thymidylate synthase thyX (251 aa)	Rv2754c.1	28 kDa	1.83	31.64	57.78
1,4-alpha-glucan branching enzyme glgB (732 aa)	Rv1326c.1	82 kDa	3.65	5.27	19.24
succinyl-CoA:3-ketoacid-coenzyme A transferase beta subunit scoB (219 aa)	Rv2503c.1	23 kDa	14.63	2.30	33.68
methyltransferase (275 aa)	Rv1405c.1	29 kDa	31.40	1.75	54.86
glutamine-dependent NAD(+) synthetase nadE (680 aa)	Rv2438c.1	75 kDa	11.12	0.71	7.84
conserved hypothetical protein (164 aa)	Rv2327.1	17 kDa	1.60	60.41	96.94
conserved hypothetical protein (129 aa)	Rv0546c.1	14 kDa	2.95	25.19	74.28
translation initiation factor IF-1 infA (74 aa)	Rv3462c.1	8 kDa	1.99	51.37	102.31
conserved hypothetical protein (237 aa)	Rv2821c.1	26 kDa	#DIV/0!	0.00	70.20
succinyl-diaminopimelate desuccinylase dapE (355 aa)	Rv1202.1	37 kDa	4.78	5.28	25.21
conserved hypothetical protein (159 aa)	Rv1276c.1	16 kDa	1.66	29.35	48.70
adenosine deaminase add (366 aa)	Rv3313c.1	40 kDa	20.69	1.22	25.16
conserved hypothetical protein (169 aa)	Rv3472.1	19 kDa	1.89	22.19	42.02
beta-glucosidase bglS (692 aa)	Rv0186.1	74 kDa	5.61	1.45	8.14
NAD-dependent malate oxidoreductase mez (549 aa)	Rv2332.1	59 kDa	#DIV/0!	0.00	4.29
heat shock protein hsp (160 aa)	Rv0251c.1	18 kDa	11.16	2.83	31.63
conserved hypothetical protein (116 aa)	Rv1466.1	12 kDa	#DIV/0!	0.00	31.07
conserved hypothetical protein (244 aa)	Rv1513.1	27 kDa	#DIV/0!	0.00	24.84
conserved hypothetical protein (208 aa)	Rv1700.1	23 kDa	3.43	11.35	38.92
5-methyltetrahydropteroyltriglutamate-homocysteine methyltransferase metE (760 aa)	Rv1133c.1	82 kDa	0.62	273.80	168.97
iron-regulated elongation factor tu tuf (397 aa)	Rv0685.1	44 kDa	0.50	359.13	180.56
iron-regulated heparin binding hemagglutinin hbhA (200 aa)	Rv0475.1	22 kDa	0.50	747.03	375.13
short-chain type dehydrogenase/reductase (287 aa)	Rv0148.1	30 kDa	0.86	437.36	376.49
DNA-directed RNA polymerase beta chain rpoB (1173 aa)	Rv0667.1	129 kDa	0.51	93.77	47.76
ATP synthase beta chain atpD (487 aa)	Rv1310.1	53 kDa	0.39	177.81	70.02
40 kda secreted L-alanine dehydrogenase ald (372 aa)	Rv2780.1	39 kDa	0.44	209.51	91.94
3-oxoacyl-[acyl-carrier protein] reductase fabG4 (455 aa)	Rv0242c.1	47 kDa	0.58	234.38	136.74
fatty-acid-CoA ligase fadD2 (561 aa)	Rv0270.1	60 kDa	0.62	167.11	104.15
conserved hypothetical protein (272 aa)	Rv0831c.1	30 kDa	0.89	303.87	271.03

DNA-directed RNA polymerase beta chain rpoC (1317 aa)	Rv0668.1	147 kDa	0.44	94.27	41.44
transcriptional regulator, crp/fnr-family (225 aa)	Rv3676.1	25 kDa	0.66	337.57	222.85
DNA topoisomerase I topA (935 aa)	Rv3646c.1	102 kDa	0.60	53.83	32.17
two component system transcriptional regulator (206 aa)	Rv1626.1	23 kDa	0.80	307.76	246.73
elongation factor G fusA2 (715 aa)	Rv0120c.1	76 kDa	0.78	136.78	106.96
iron-regulated peptidyl-prolyl-cis-trans-isomerase A ppiA (183 aa)	Rv0009.1	19 kDa	0.78	325.68	252.85
acyl-CoA thiolase fadA (404 aa)	Rv0859.1	42 kDa	0.87	151.81	132.31
acetyl-CoA acyltransferase fadA2 (441 aa)	Rv0243.1	46 kDa	0.70	126.41	88.22
quinone reductase qor (329 aa)	Rv1454c.1	34 kDa	0.80	163.86	131.56
acyl-CoA dehydrogenase fadE25 (390 aa)	Rv3274c.1	42 kDa	0.72	101.33	72.59
enoyl-CoA hydratase echA6 (244 aa)	Rv0905.1	26 kDa	0.46	108.72	50.36
ferredoxin-dependent glutamate synthase large subunit gltB (1528 aa)	Rv3859c.1	166 kDa	0.50	32.53	16.21
transferase (318 aa)	Rv1201c.1	33 kDa	0.64	90.38	58.23
reductase (412 aa)	Rv1869c.1	44 kDa	0.54	112.82	61.15
transcriptional regulator, tetR-family (222 aa)	Rv3295.1	25 kDa	0.65	183.67	119.60
enoyl-CoA hydratase echA8 (258 aa)	Rv1070c.1	27 kDa	0.52	62.13	32.29
conserved hypothetical protein (116 aa)	Rv0801.1	13 kDa	0.83	187.13	154.92
conserved hypothetical protein (345 aa)	Rv2159c.1	36 kDa	0.31	132.01	40.70
NADH dehydrogenase I chain G nuoG (807 aa)	Rv3151.1	85 kDa	0.57	33.81	19.29
bacterioferritin bfrA (160 aa)	Rv1876.1	18 kDa	0.61	290.33	177.20
conserved hypothetical protein (514 aa)	Rv2226.1	56 kDa	0.13	49.53	6.23
RNA polymerase sigma factor sigA (529 aa)	Rv2703.1	58 kDa	0.51	18.69	9.45
conserved alanine and glycine rich protein (393 aa)	Rv3616c.1	40 kDa	0.82	64.79	52.93
immunogenic protein mpt63 (160 aa)	Rv1926c.1	17 kDa	0.71	96.95	68.77
Esat-6 like protein esxO (95 aa)	Rv2346c.1	10 kDa	0.52	109.55	56.82
50S ribosomal protein L2 rplB (281 aa)	Rv0704.1	31 kDa	0.33	35.11	11.66
fatty-acid-CoA ligase fadD13 (504 aa)	Rv3089.1	54 kDa	0.61	41.66	25.50
transcriptional regulator whiA (326 aa)	Rv1423.1	35 kDa	0.52	52.92	27.70
serine hydroxymethyltransferase 1 glyA1 (427 aa)	Rv1093.1	45 kDa	0.12	65.59	7.99
conserved hypothetical protein (88 aa)	Rv0898c.1	10 kDa	0.56	133.75	74.55
NADH-dependent glutamate synthase small subunit gltD (489 aa)	Rv3858c.1	53 kDa	0.58	30.73	17.73
50S ribosomal protein L24 rplX (106 aa)	Rv0715.1	11 kDa	0.14	61.16	8.67
histidyl-tRNA synthetase hisS (424 aa)	Rv2580c.1	45 kDa	0.00	3.42	0.00
50S ribosomal protein L13 rplM (148 aa)	Rv3443c.1	16 kDa	0.03	53.10	1.83
50S ribosomal protein L15 rplO (147 aa)	Rv0723.1	16 kDa	0.17	60.73	10.59
carboxylase (495 aa)	Rv1722.1	55 kDa	0.60	15.06	9.06
glycyl-tRNA synthetase glyS (464 aa)	Rv2357c.1	53 kDa	0.21	22.32	4.78
transcriptional regulator, asnC-family (151 aa)	Rv3291c.1	17 kDa	0.71	124.24	88.59
ATPase (729 aa)	Rv0435c.1	75 kDa	0.16	20.14	3.28
two component system transcriptional regulator phoP (248 aa)	Rv0757.1	28 kDa	0.51	32.28	16.61
moaD-moaE fusion protein moaX (222 aa)	Rv3323c.1	24 kDa	0.65	67.83	44.38
hypothetical protein (225 aa)	Rv0088.1	25 kDa	0.59	29.30	17.34
cystathionine gamma-synthase metB (389	Rv1079.1	41 kDa	0.16	37.47	5.98

aa)

conserved hypothetical protein (281 aa)	Rv0241c.1	30 kDa	0.19	65.70	12.50
conserved hypothetical protein (421 aa)	Rv3311.1	46 kDa	0.32	17.65	5.57
acyl-CoA dehydrogenase fadE24 (469 aa)	Rv3139.1	50 kDa	0.26	14.94	3.95
alternative RNA polymerase sigma factor sigH (217 aa)	Rv3223c.1	24 kDa	0.29	34.43	9.83
peroxidoxin bcpB (155 aa)	Rv1608c.1	17 kDa	0.48	48.20	23.20
50S ribosomal protein L14 rplN (123 aa)	Rv0714.1	13 kDa	0.14	38.07	5.26
molybdopterin biosynthesis protein moeA1 (427 aa)	Rv0994.1	44 kDa	0.19	22.10	4.19
conserved hypothetical protein (195 aa)	Rv3773c.1	21 kDa	0.00	28.84	0.00
endonuclease III nth (246 aa)	Rv3674c.1	27 kDa	0.05	26.53	1.21
conserved hypothetical protein (419 aa)	Rv2953.1	45 kDa	0.00	8.94	0.00
NADH dehydrogenase I chain C nuoC (237 aa)	Rv3147.1	27 kDa	0.05	29.65	1.36
L-lactate dehydrogenase lldD2 (415 aa)	Rv1872c.1	45 kDa	0.22	6.76	1.51
nitrate/nitrite response transcriptional regulator narL (217 aa)	Rv0844c.1	23 kDa	0.00	19.74	0.00
aminotransferase (399 aa)	Rv3778c.1	42 kDa	0.00	8.13	0.00
methanol dehydrogenase transcriptional regulator moxR2 (359 aa)	Rv3692.1	38 kDa	0.00	14.13	0.00
conserved hypothetical protein (79 aa)	Rv0500A.1	9 kDa	0.27	71.73	19.62
dehydrogenase/reductase (249 aa)	Rv0769.1	26 kDa	0.55	13.10	7.15
conserved hypothetical protein (144 aa)	Rv1546.1	15 kDa	0.00	33.13	0.00
hypothetical protein (143 aa)	Rv0441c.1	15 kDa	0.24	36.03	8.73
conserved hypothetical protein (169 aa)	Rv2179c.1	20 kDa	0.07	24.68	1.76
transcriptional regulator (141 aa)	Rv0474.1	15 kDa	0.00	23.20	0.00
zinc-type alcohol dehydrogenase adhE2 (362 aa)	Rv2259.1	38 kDa	0.12	14.13	1.67

Appendix Table 7. Significantly different proteins in the Membrane of the clonal laboratory pair

Identified Proteins p<0.05 (t-test)	Accession Number	Molecular Weight	Fold Change (INHs/ INHr)	NSAF Average lab INHr x 1000	Average lab INHs x 1000
glutamine synthetase glnA1 (479 aa)	Rv2220.1	54 kDa	1.32	680.99	901.23
S-adenosylmethionine synthetase metK (404 aa)	Rv1392.1	43 kDa	1.42	475.19	674.46
ATP-dependent clp protease proteolytic subunit 2 clpP2 (215 aa)	Rv2460c.1	24 kDa	1.20	961.88	1156.90
citrate synthase I gltA2 (432 aa)	Rv0896.1	48 kDa	1.18	323.53	381.42
catalase-peroxidase-peroxynitritase T katG (741 aa)	Rv1908c.1	81 kDa	4.91	34.39	168.91
phosphoserine aminotransferase serC (377 aa)	Rv0884c.1	40 kDa	1.36	384.73	522.70
conserved hypothetical protein (298 aa)	Rv2623.1	32 kDa	1.49	386.59	576.76
glycine dehydrogenase gcvB (942 aa)	Rv1832.1	100 kDa	1.25	91.40	114.53
dihydrolipoamide dehydrogenase lpd (465 aa)	Rv0462.1	49 kDa	1.27	148.38	187.90
arginine deiminase arcA (403 aa)	Rv1001.1	43 kDa	1.32	170.93	224.98
aldehyde dehydrogenase (508 aa)	Rv0458.1	55 kDa	1.97	39.72	78.38
acetyl-CoA acyltransferase fadA2 (441 aa)	Rv0243.1	46 kDa	1.31	127.82	166.84
30S ribosomal protein S13 rpsM (125 aa)	Rv3460c.1	14 kDa	1.42	324.96	462.69
bifunctional methylenetetrahydrofolate dehydrogenase folD (282 aa)	Rv3356c.1	29 kDa	1.25	182.89	228.47
enoyl-CoA hydratase echA21 (275 aa)	Rv3774.1	29 kDa	1.59	82.87	132.06
conserved hypothetical protein (436 aa)	Rv3722c.1	47 kDa	1.43	82.25	117.85
cysteinyl-tRNA synthetase 1 cysS1 (470 aa)	Rv3580c.1	52 kDa	1.66	58.90	97.78
protein transport protein secE2 (72 aa)	Rv0379.1	8 kDa	2.21	208.93	462.44
acetyl-CoA acetyltransferase fadA4 (390 aa)	Rv1323.1	40 kDa	1.51	106.77	161.18
NADH dehydrogenase I chain G nuoG (807 aa)	Rv3151.1	85 kDa	2.02	13.59	27.49
polyphosphate kinase ppk (743 aa)	Rv2984.1	83 kDa	1.57	29.09	45.65
threonine dehydratase ilvA (430 aa)	Rv1559.1	45 kDa	1.27	60.94	77.27
cystathionine beta-synthase cbs (465 aa)	Rv1077.1	49 kDa	1.56	66.47	103.87
phosphoglucomutase pgmA (548 aa)	Rv3068c.1	58 kDa	1.74	39.53	68.70
acyl-CoA dehydrogenase fadE8 (543 aa)	Rv0672.1	59 kDa	1.92	41.20	79.07
30S ribosomal protein S17 rpsQ (137 aa)	Rv0710.1	15 kDa	1.54	98.34	151.18
trehalose-6-phosphate phosphatase otsB1 (1328 aa)	Rv2006.1	146 kDa	2.11	9.91	20.90
fatty-acid-CoA ligase fadD13 (504 aa)	Rv3089.1	54 kDa	1.57	33.26	52.22
6-phosphogluconate dehydrogenase, decarboxylating gnd2 (341 aa)	Rv1122.1	36 kDa	2.85	6.33	18.02
membrane-associated serine protease (398 aa)	Rv3671c.1	41 kDa	1.53	35.73	54.54
conserved hypothetical protein (453 aa)	Rv2097c.1	51 kDa	1.54	31.03	47.74
1,4-alpha-glucan branching enzyme glgB (732 aa)	Rv1326c.1	82 kDa	2.24	9.62	21.58
3-methyl-2-oxobutanoate hydroxymethyltransferase panB (282 aa)	Rv2225.1	29 kDa	1.72	27.27	46.86
riboflavin biosynthesis protein ribA2 (426 aa)	Rv1415.1	46 kDa	2.06	9.94	20.50
glucose-6-phosphate 1-dehydrogenase zwf1	Rv1121.1	52 kDa	2.14	11.32	24.17

(467 aa)					
monophosphatase (261 aa)	Rv3137.1	28 kDa	1.63	34.27	55.73
NADH dehydrogenase I chain F nuoF (446 aa)	Rv3150.1	48 kDa	3.15	6.22	19.58
4-hydroxy-2-oxovalerate aldolase (347 aa)	Rv3534c.1	36 kDa	1.75	11.29	19.80
aminotransferase (398 aa)	Rv0858c.1	42 kDa	25.01	0.83	20.86
ribonucleoside-diphosphate reductase beta chain nrdF2 (325 aa)	Rv3048c.1	37 kDa	11.79	1.06	12.52
fatty-acid-CoA ligase fadD32 (638 aa)	Rv3801c.1	69 kDa	0.52	314.09	163.28
polyketide synthase pks13 (1734 aa)	Rv3800c.1	186 kDa	0.38	121.51	45.84
L-lactate dehydrogenase lldD2 (415 aa)	Rv1872c.1	45 kDa	0.73	303.07	222.04
conserved hypothetical protein (632 aa)	Rv0282.1	68 kDa	0.55	195.86	106.94
conserved membrane protein (1331 aa)	Rv0284.1	145 kDa	0.43	89.96	38.38
preprotein translocase secA1 (950 aa)	Rv3240c.1	106 kDa	0.13	137.91	18.01
conserved hypothetical protein (382 aa)	Rv1488.1	41 kDa	0.79	264.27	208.36
ATP-dependent protease ATP-binding subunit clpC1 (849 aa)	Rv3596c.1	94 kDa	0.57	83.91	48.05
oxidoreductase (382 aa)	Rv2951c.1	41 kDa	0.61	232.87	141.46
transmembrane serine/threonine-protein kinase D pknD (665 aa)	Rv0931c.1	70 kDa	0.78	98.08	76.71
lipase/esterase lipN (377 aa)	Rv2970c.1	40 kDa	0.73	137.32	99.89
fatty-acid-CoA ligase fadD28 (581 aa)	Rv2941.1	63 kDa	0.83	67.95	56.45
ATP synthase delta chain atpH (447 aa)	Rv1307.1	49 kDa	0.62	121.17	75.41
transcription termination factor rho (603 aa)	Rv1297.1	65 kDa	0.55	84.42	46.19
short-chain type dehydrogenase/reductase (251 aa)	Rv1144.1	26 kDa	0.61	238.74	145.00
hypothetical protein (225 aa)	Rv0088.1	25 kDa	0.77	151.13	116.83
glutamine-transport ATP-binding protein ABC transporter glnQ (331 aa)	Rv2564.1	35 kDa	0.56	123.26	69.05
polyketide synthase pks2 (2127 aa)	Rv3825c.1	226 kDa	0.10	21.23	2.20
short-chain type dehydrogenase/reductase (261 aa)	Rv2766c.1	27 kDa	0.74	90.25	66.79
alkyl hydroperoxide reductase D protein ahpD (178 aa)	Rv2429.1	19 kDa	0.61	218.42	134.06
conserved membrane protein (661 aa)	Rv2345.1	70 kDa	0.34	62.36	21.24
anion transporter ATPase (387 aa)	Rv3680.1	41 kDa	0.69	89.80	62.40
conserved membrane protein (256 aa)	Rv2969c.1	27 kDa	0.64	154.01	97.96
fatty-acid-CoA ligase fadD6 (598 aa)	Rv1206.1	64 kDa	0.38	19.51	7.43
conserved hypothetical protein (286 aa)	Rv3463.1	31 kDa	0.55	106.40	58.11
acyl-CoA dehydrogenase fadE23 (402 aa)	Rv3140.1	43 kDa	0.51	59.03	30.16
conserved membrane protein (422 aa)	Rv0227c.1	46 kDa	0.66	44.96	29.76
fatty-acid-CoA ligase fadD36 (474 aa)	Rv1193.1	50 kDa	0.67	46.71	31.39
dehydrogenase FAD flavoprotein gmc oxidoreductase (529 aa)	Rv1279.1	57 kDa	0.57	36.81	20.80
heat shock protein hsp (160 aa)	Rv0251c.1	18 kDa	0.14	128.80	18.20
glycyl-tRNA synthetase glyS (464 aa)	Rv2357c.1	53 kDa	0.60	43.40	26.13
peptidyl-prolyl-cis-trans-isomerase B ppiB (309 aa)	Rv2582.1	32 kDa	0.69	90.27	61.89
3-oxoacyl-[acyl-carrier protein] reductase fabG1 (248 aa)	Rv1483.1	26 kDa	0.54	94.61	51.47
chaperone protein dnaJ1 (396 aa)	Rv0352.1	41 kDa	0.79	45.32	35.88
conserved hypothetical protein (143 aa)	Rv0636.1	15 kDa	0.33	76.95	25.06
cell division ATP-binding protein ftsE (230 aa)	Rv3102c.1	26 kDa	0.53	113.52	60.31
lipoprotein lpqG (241 aa)	Rv3623.1	25 kDa	0.32	112.62	36.28

conserved hypothetical protein (421 aa)	Rv3311.1	46 kDa	0.31	53.16	16.60
metal cation transporting P-type ATPase ctpC (719 aa)	Rv3270.1	76 kDa	0.21	21.56	4.46
epoxide hydrolase ephF (301 aa)	Rv0134.1	34 kDa	0.63	56.15	35.22
conserved hypothetical protein (283 aa)	Rv1978.1	31 kDa	0.47	74.73	34.97
oxidoreductase (283 aa)	Rv2971.1	30 kDa	0.08	14.86	1.26
conserved membrane protein (325 aa)	Rv2037c.1	35 kDa	0.32	65.97	20.81
preprotein translocase ATPase secA2 (809 aa)	Rv1821.1	89 kDa	0.28	15.92	4.45
short-chain type dehydrogenase ephD (593 aa)	Rv2214c.1	64 kDa	0.17	23.90	4.02
dehydrogenase/reductase (312 aa)	Rv0439c.1	34 kDa	0.33	58.55	19.46
ubiquinol-cytochrome C reductase qcrB cytochrome B subunit (550 aa)	Rv2196.1	61 kDa	0.49	26.77	13.12
conserved hypothetical protein (471 aa)	Rv3719.1	53 kDa	0.41	30.46	12.36
transcription antitermination protein nusG (239 aa)	Rv0639.1	25 kDa	0.35	8.68	3.02
glycerol-3-phosphate dehydrogenase glpD2 (586 aa)	Rv3302c.1	63 kDa	0.40	24.49	9.83
transcriptional regulator, arsR-family (435 aa)	Rv0576.1	46 kDa	0.38	10.57	4.03
oxidoreductase (461 aa)	Rv1751.1	50 kDa	0.22	8.54	1.88
conserved hypothetical protein (94 aa)	Rv3269.1	10 kDa	0.52	241.16	126.16
dihydrodipicolinate reductase dapB (246 aa)	Rv2773c.1	26 kDa	0.86	69.33	59.71
membrane protein (290 aa)	Rv0048c.1	31 kDa	0.38	55.80	21.07
hypothetical protein (238 aa)	Rv3528c.1	27 kDa	0.30	59.33	17.72
conserved hypothetical protein (301 aa)	Rv1794.1	32 kDa	0.67	62.48	41.68
molybdopterin biosynthesis protein moeW (319 aa)	Rv2338c.1	35 kDa	0.06	20.69	1.17
GTP-binding protein lepA (654 aa)	Rv2404c.1	72 kDa	0.75	14.06	10.56
transcriptional regulator, asnC-family (247 aa)	Rv3050c.1	28 kDa	0.44	69.15	30.23
ubiquinol-cytochrome C reductase qcrC cytochrome C subunit (281 aa)	Rv2194.1	29 kDa	0.72	40.38	28.99
conserved hypothetical protein (195 aa)	Rv3773c.1	21 kDa	0.71	61.38	43.76
short-chain type dehydrogenase/reductase (255 aa)	Rv3791.1	27 kDa	0.28	20.50	5.73
conserved membrane protein (1185 aa)	Rv3910.1	124 kDa	0.21	7.74	1.59
conserved hypothetical protein (168 aa)	Rv2574.1	19 kDa	0.62	56.95	35.24
glycerol-3-phosphate dehydrogenase gpdA2 (335 aa)	Rv2982c.1	34 kDa	0.30	17.61	5.25
oligopeptide-transport ATP-binding protein ABC transporter oppD (613 aa)	Rv1281c.1	65 kDa	0.36	6.85	2.46
conserved hypothetical protein (373 aa)	Rv0906.1	41 kDa	0.34	31.96	10.85
Esat-6 like protein esxW (99 aa)	Rv3620c.1	11 kDa	0.40	107.83	43.33
conserved hypothetical protein (302 aa)	Rv0830.1	33 kDa	0.18	19.80	3.64
molybdenum cofactor biosynthesis protein moeB1 (393 aa)	Rv3206c.1	42 kDa	0.22	24.39	5.44
conserved hypothetical protein (534 aa)	Rv2510c.1	57 kDa	0.12	10.90	1.35
conserved hypothetical protein (588 aa)	Rv1069c.1	64 kDa	0.00	10.23	0.00
lipoprotein lppS (409 aa)	Rv2518c.1	43 kDa	0.22	16.06	3.49
lipoprotein lpqZ (287 aa)	Rv1244.1	30 kDa	0.07	17.39	1.21
conserved membrane protein (305 aa)	Rv0007.1	31 kDa	0.11	21.00	2.29
NADPH quinone oxidoreductase fadB4 (324 aa)	Rv3141.1	34 kDa	0.36	12.31	4.48

PPE family protein (392 aa)	Rv1196.1	39 kDa	0.09	10.74	0.95
proline-rich antigen pra (241 aa)	Rv1078.1	25 kDa	0.32	32.08	10.27
serine/threonine phosphatase ppp (515 aa)	Rv0018c.1	54 kDa	0.00	6.80	0.00
conserved hypothetical protein (377 aa)	Rv0365c.1	42 kDa	0.07	13.12	0.94
amidase amiB2 (463 aa)	Rv1263.1	49 kDa	0.16	14.22	2.30
conserved hypothetical protein (228 aa)	Rv2794c.1	25 kDa	0.00	14.93	0.00
drugs-transport transmembrane ATP-binding protein ABC transporter (640 aa)	Rv1819c.1	71 kDa	0.00	5.58	0.00
ATP synthase B chain atpF (172 aa)	Rv1306.1	18 kDa	0.11	55.93	6.23
conserved hypothetical protein (317 aa)	Rv1647.1	34 kDa	0.10	12.41	1.18
short-chain type dehydrogenase/reductase (269 aa)	Rv2509.1	28 kDa	0.09	18.21	1.61
membrane protein mmpS3 (300 aa)	Rv2198c.1	31 kDa	0.00	16.73	0.00
conserved alanine and leucine rich membrane protein (394 aa)	Rv2625c.1	41 kDa	0.20	13.75	2.78
conserved hypothetical protein (148 aa)	Rv1875.1	16 kDa	0.35	13.92	4.82
pyrimidine operon regulatory protein pyrR (194 aa)	Rv1379.1	21 kDa	0.21	17.63	3.73
glycine and proline rich membrane protein (294 aa)	Rv1097c.1	30 kDa	0.16	7.40	1.21

Appendix Table 8. Significantly different proteins in the Membrane of the clonal clinical pair

Identified Proteins (1173/1181)	Accession Number	Molecular Weight	Fold Change (INHs/ INHr)	NSAF Average clinical INHr	Average clinical INHs
pyridoxine biosynthesis protein (300 aa)	Rv2606c.1	31 kDa	0.3	105.5	34.7
50S ribosomal protein L30 rpmD (66 aa)	Rv0722.1	7 kDa	1.8	207.8	371.6
MCE-family protein mce1B (347 aa)	Rv0170.1	38 kDa	0.0	11.9	0.0
membrane-bound ell division protein ftsH (761 aa)	Rv3610c.1	82 kDa	0.0	1.3	0.0
dehydrogenase/reductase (303 aa)	Rv0303.1	33 kDa	0.0	6.6	0.0
acetyl-CoA acyltransferase fadA2 (441 aa)	Rv0243.1	46 kDa	0.5	133.0	72.5
lipoate biosynthesis protein A lipA (312 aa)	Rv2218.1	35 kDa	0.0	3.2	0.0
conserved hypothetical protein (272 aa)	Rv0831c.1	30 kDa	0.7	251.1	180.5
40 kda secreted L-alanine dehydrogenase ald (372 aa)	Rv2780.1	39 kDa	0.4	249.3	93.1
conserved hypothetical protein (611 aa)	Rv1798.1	68 kDa	0.0	10.0	0.0
conserved membrane protein (342 aa)	Rv3888c.1	38 kDa	0.0	7.3	0.0
isocitrate dehydrogenase icd2 (746 aa)	Rv0066c.1	83 kDa	0.7	124.9	89.4
S-adenosylmethionine synthetase metK (404 aa)	Rv1392.1	43 kDa	1.4	312.4	425.3
conserved hypothetical protein (283 aa)	Rv1978.1	31 kDa	0.1	49.8	4.9
cystathionine gamma-synthase metB (389 aa)	Rv1079.1	41 kDa	0.5	83.6	40.2
anthranilate phosphoribosyltransferase trpD (371 aa)	Rv2192c.1	38 kDa	0.0	49.3	2.2
haloalkane dehalogenase (301 aa)	Rv2296.1	33 kDa	0.5	324.0	171.6
enolase eno (430 aa)	Rv1023.1	45 kDa	0.8	304.5	240.0
catalase-peroxidase-peroxynitritase T katG (741 aa)	Rv1908c.1	81 kDa	5.3	48.8	259.4
epoxide hydrolase ephF (301 aa)	Rv0134.1	34 kDa	0.1	53.8	3.6
transcriptional regulator, gntR-family (245 aa)	Rv0043c.1	27 kDa	#DIV/0!	0.0	18.6
glutamine synthetase glnA4 (458 aa)	Rv2860c.1	50 kDa	1.3	141.2	189.7
arginine/ornithine transport system ATPase (335 aa)	Rv1496.1	36 kDa	5.2	8.2	42.9
phosphoglycerate kinase pgk (413 aa)	Rv1437.1	43 kDa	1.5	68.7	101.0
conserved hypothetical protein (287 aa)	Rv1118c.1	32 kDa	0.3	31.2	10.9
endopeptidase ATP binding protein chain B clpB (849 aa)	Rv0384c.1	93 kDa	2.3	44.8	102.0
conserved hypothetical protein (322 aa)	Rv3218.1	34 kDa	0.2	22.7	3.5
monophosphatase (261 aa)	Rv3137.1	28 kDa	9.0	3.9	35.0
DNA/pantothenate metabolism flavoprotein dfp (419 aa)	Rv1391.1	44 kDa	11.1	3.8	41.5
thioredoxin trxB1 (124 aa)	Rv1471.1	13 kDa	2.5	60.8	155.1
thiamine biosynthesis protein thiC (548 aa)	Rv0423c.1	60 kDa	8.4	1.0	8.2
ketoacyl reductase (268 aa)	Rv1544.1	28 kDa	0.3	125.1	33.1
3-oxoacyl-[acyl-carrier protein] reductase fabG4 (455 aa)	Rv0242c.1	47 kDa	0.6	399.2	237.5
conserved hypothetical protein (302 aa)	Rv2172c.1	33 kDa	3.2	24.5	78.4
threonine synthase thrC (361 aa)	Rv1295.1	37 kDa	2.1	33.5	71.6
response regulator (134 aa)	Rv3143.1	14 kDa	0.2	65.1	15.0
cholesterol oxidase precursor choD (579 aa)	Rv3409c.1	63 kDa	0.2	21.1	4.0
dihydrodipicolinate reductase dapB (246 aa)	Rv2773c.1	26 kDa	0.03	33.7	1.2
long-chain (C50)-Z-isoprenyl diphosphate	Rv2361c.1	34 kDa	0.0	8.9	0.0

synthase (297 aa)					
3-methyl-2-oxobutanoate hydroxymethyltransferase panB (282 aa)	Rv2225.1	29 kDa	9.2	3.8	35.2
conserved hypothetical protein (320 aa)	Rv3683.1	34 kDa	0.1	21.1	2.5
DNA-directed RNA polymerase beta chain rpoB (1173 aa)	Rv0667.1	129 kDa	0.6	175.8	106.4
conserved hypothetical protein (94 aa)	Rv2632c.1	10 kDa	0.3	285.9	93.8
bacterioferritin bfrA (160 aa)	Rv1876.1	18 kDa	0.7	1242.0	905.9
lipoprotein lppZ (374 aa)	Rv3006.1	39 kDa	0.3	61.3	19.4
alkyl hydroperoxide reductase D protein ahpD (178 aa)	Rv2429.1	19 kDa	7.9	18.6	146.4
conserved hypothetical protein (152 aa)	Rv3678c.1	15 kDa	1.7	147.0	250.9
cation-transporter ATPase I ctpI (1633 aa)	Rv0107c.1	170 kDa	0.0	3.4	0.0
conserved hypothetical protein (144 aa)	Rv1546.1	15 kDa	0.1	40.6	3.7
formyltetrahydrofolate deformylase purU (311 aa)	Rv2964.1	34 kDa	0.2	12.7	2.7
dioxygenase (503 aa)	Rv0913c.1	57 kDa	0.1	18.3	1.6
polyphosphate glucokinase ppgK (266 aa)	Rv2702.1	27 kDa	0.4	84.3	34.1
glucose-6-phosphate 1-dehydrogenase zwf2 (515 aa)	Rv1447c.1	57 kDa	0.4	102.2	39.5
conserved hypothetical protein (208 aa)	Rv2466c.1	23 kDa	1.7	104.3	173.7
chorismate synthase aroF (402 aa)	Rv2540c.1	42 kDa	1.8	22.3	40.4
L-lactate dehydrogenase lldD2 (415 aa)	Rv1872c.1	45 kDa	0.5	382.0	183.4
ATP synthase gamma chain atpG (306 aa)	Rv1309.1	34 kDa	0.5	192.3	88.7
conserved hypothetical protein (191 aa)	Rv0464c.1	21 kDa	0.3	90.1	23.8
inosine-5-monophosphate dehydrogenase guaB3 (376 aa)	Rv3410c.1	39 kDa	0.8	285.5	216.3
glutamyl-tRNA(gln) amidotransferase subunit B gatB (510 aa)	Rv3009c.1	55 kDa	1.5	39.9	60.5
valyl-tRNA synthase protein valS (877 aa)	Rv2448c.1	98 kDa	2.5	9.5	24.3
secreted fibronectin-binding protein antigen fbpA (339 aa)	Rv3804c.1	36 kDa	2.2	40.2	90.0
isoleucyl-tRNA synthetase ileS (1042 aa)	Rv1536.1	117 kDa	2.4	2.5	6.0
conserved hypothetical protein (73 aa)	Rv1893.1	7 kDa	0.1	75.3	7.7
conserved hypothetical protein (201 aa)	Rv0966c.1	22 kDa	0.0	13.3	0.0
alkyl hydroperoxide reductase C protein ahpC (196 aa)	Rv2428.1	22 kDa	30.7	5.6	171.2
conserved hypothetical protein (143 aa)	Rv0636.1	15 kDa	2.0	107.0	212.0
succinyl-CoA:3-ketoacid-coenzyme A transferase alpha subunit scoA (249 aa)	Rv2504c.1	26 kDa	2.6	10.2	26.6
membrane-associated phospholipase C 2 plcB (513 aa)	Rv2350c.1	56 kDa	0.1	10.5	0.5
conserved hypothetical protein (376 aa)	Rv3796.1	39 kDa	0.0	16.5	0.7
dehydrogenase (460 aa)	Rv2280.1	48 kDa	0.1	68.9	6.3
conserved hypothetical protein (361 aa)	Rv1063c.1	38 kDa	0.1	6.1	0.8
conserved hypothetical protein (564 aa)	Rv3091.1	62 kDa	0.1	44.6	5.4
esterase lipW (303 aa)	Rv0217c.1	32 kDa	0.0	24.2	0.0
Rieske iron-sulfur protein qcrA (430 aa)	Rv2195.1	47 kDa	0.3	47.1	12.4
dehydrogenase (398 aa)	Rv3726.1	42 kDa	0.2	64.0	12.9
propionyl-CoA carboxylase beta chain 5 accD5 (549 aa)	Rv3280.1	59 kDa	1.7	84.2	144.8
cysteinyl-tRNA synthetase 2 cysS2 (415 aa)	Rv2130c.1	46 kDa	0.1	22.9	3.4
tyrosyl-tRNA synthase tyrS (425 aa)	Rv1689.1	46 kDa	1.5	64.7	94.3
cholesterol dehydrogenase (371 aa)	Rv1106c.1	41 kDa	0.2	104.7	25.7
acetaldehyde dehydrogenase (304 aa)	Rv3535c.1	32 kDa	5.3	1.8	9.7
protease IV sppA (624 aa)	Rv0724.1	66 kDa	0.2	76.6	19.1

conserved hypothetical protein (86 aa)	Rv1398c.1	9 kDa	3.2	131.4	421.6
monooxygenase (525 aa)	Rv3049c.1	59 kDa	0.0	11.5	0.0
cytochrome C oxidase polypeptide I ctaD (574 aa)	Rv3043c.1	64 kDa	0.1	6.2	0.5
glycogen phosphorylase glgP (864 aa)	Rv1328.1	96 kDa	1.4	98.1	138.4
fatty-acid-CoA ligase fadD2 (561 aa)	Rv0270.1	60 kDa	0.5	244.0	119.6
triosephosphate isomerase tpi (262 aa)	Rv1438.1	27 kDa	2.0	42.7	85.3
conserved hypothetical protein (562 aa)	Rv1215c.1	63 kDa	0.2	17.5	3.4
mycolic acid synthase umaA (287 aa)	Rv0469.1	33 kDa	2.9	25.3	73.3
conserved hypothetical protein (150 aa)	Rv2740.1	17 kDa	0.1	67.8	9.5
iron-regulated heparin binding hemagglutinin hbhA (200 aa)	Rv0475.1	22 kDa	0.8	334.8	266.3
O-acetylhomoserine sulfhydrylase metC (450 aa)	Rv3340.1	47 kDa	2.2	68.7	150.6
long-chain acyl-CoA synthase (1000 aa)	Rv1683.1	107 kDa	0.5	9.9	4.6
1,4-alpha-glucan branching enzyme glgB (732 aa)	Rv1326c.1	82 kDa	4.5	2.8	12.6
enhanced intracellular survival protein eis (403 aa)	Rv2416c.1	44 kDa	2.4	73.7	175.3
transcriptional regulator, crp/fnr-family (225 aa)	Rv3676.1	25 kDa	0.5	274.3	149.9
fatty-acid-CoA ligase fadD6 (598 aa)	Rv1206.1	64 kDa	0.1	59.4	7.7
hydrolase (263 aa)	Rv3400.1	28 kDa	2.4	24.1	58.4
conserved hypothetical protein (584 aa)	Rv2565.1	62 kDa	0.0	10.0	0.0
conserved hypothetical protein (119 aa)	Rv1489.1	12 kDa	0.2	286.0	70.9
fatty-acid-CoA ligase fadD9 (1169 aa)	Rv2590.1	127 kDa	#DIV/0!	0.0	4.7
conserved hypothetical protein (185 aa)	Rv3614c.1	20 kDa	5.2	11.5	60.0
conserved hypothetical protein (317 aa)	Rv1647.1	34 kDa	0.2	11.3	1.7
conserved hypothetical protein (130 aa)	Rv1871c.1	15 kDa	0.5	131.1	62.8
membrane protein mmpS3 (300 aa)	Rv2198c.1	31 kDa	#DIV/0!	0.0	15.4
glutamine-transport ATP-binding protein ABC transporter glnQ (331 aa)	Rv2564.1	35 kDa	0.5	83.7	39.7
amidotransferase hisH (207 aa)	Rv1602.1	21 kDa	0.0	19.7	0.0
30S ribosomal protein S11 rpsK (140 aa)	Rv3459c.1	15 kDa	4.8	51.1	244.7
29 kda antigen cfp29 (266 aa)	Rv0798c.1	29 kDa	2.0	428.9	846.9
conserved hypothetical protein (155 aa)	Rv1919c.1	17 kDa	0.4	307.4	123.2
aminotransferase (408 aa)	Rv2294.1	44 kDa	0.0	4.0	0.0
conserved hypothetical protein (514 aa)	Rv2226.1	56 kDa	0.3	54.4	15.3
conserved hypothetical protein (104 aa)	Rv3865.1	11 kDa	0.7	274.1	186.5
conserved hypothetical protein (474 aa)	Rv3433c.1	47 kDa	1.9	21.0	39.6
conserved hypothetical protein (318 aa)	Rv1480.1	34 kDa	0.2	21.8	4.6
60 kda chaperonin 1 groEL1 (540 aa)	Rv3417c.1	56 kDa	0.7	175.7	122.7
conserved hypothetical protein (322 aa)	Rv2955c.1	36 kDa	0.4	81.3	34.9
conserved hypothetical protein (302 aa)	Rv2216.1	32 kDa	0.1	47.7	5.6
conserved hypothetical protein (359 aa)	Rv0926c.1	38 kDa	0.4	21.4	8.0
conserved hypothetical protein (98 aa)	Rv2117.1	11 kDa	4.9	25.7	126.0
dihydrolipoamide dehydrogenase lpd (465 aa)	Rv0462.1	49 kDa	1.3	93.4	120.3
lipoprotein lpqD (237 aa)	Rv3390.1	25 kDa	0.7	77.3	53.0
cytochrome P450 136 cyp136 (493 aa)	Rv3059.1	56 kDa	0.0	3.3	0.0
lipoprotein lpqN (229 aa)	Rv0583c.1	24 kDa	0.3	73.9	24.0
glycyl-tRNA synthetase glyS (464 aa)	Rv2357c.1	53 kDa	0.4	40.9	16.9
DNA polymerase III beta chain dnaN (403 aa)	Rv0002.1	42 kDa	0.6	58.0	35.3
conserved hypothetical protein (237 aa)	Rv2821c.1	26 kDa	7.8	9.2	71.6
oxidoreductase (382 aa)	Rv2951c.1	41 kDa	0.4	128.2	46.8
conserved hypothetical protein (164 aa)	Rv0580c.1	18 kDa	0.4	440.6	188.5
hypothetical protein (430 aa)	Rv2305.1	46 kDa	0.2	19.9	3.9

conserved hypothetical protein (237 aa)	Rv2135c.1	25 kDa	0.6	24.7	15.8
conserved hypothetical protein (148 aa)	Rv3718c.1	16 kDa	0.5	36.1	19.4
50S ribosomal protein L13 rplM (148 aa)	Rv3443c.1	16 kDa	1.8	267.1	478.0
fatty-acid-CoA racemase far (360 aa)	Rv0855.1	38 kDa	0.2	14.5	3.0
citrate synthase I gltA2 (432 aa)	Rv0896.1	48 kDa	0.8	370.6	284.4
transcriptional regulator moxR1 (378 aa)	Rv1479.1	41 kDa	0.3	275.8	85.9
polyphosphate kinase ppk (743 aa)	Rv2984.1	83 kDa	0.3	33.9	10.7
conserved alanine and arginine rich protein (451 aa)	Rv2731.1	50 kDa	1.9	13.6	25.7
transposase (571 aa)	Rv3327.1	63 kDa	0.3	28.8	8.4
bifunctional UDP-galactofuranosyl transferase glfT (638 aa)	Rv3808c.1	72 kDa	0.0	7.4	0.0
protein-export membrane protein secF (443 aa)	Rv2586c.1	47 kDa	#DIV/0!	0.0	2.6
ATPase (610 aa)	Rv2115c.1	67 kDa	2.0	30.9	63.4
acetyl-CoA synthetase acs (652 aa)	Rv3667.1	71 kDa	0.1	3.9	0.4
transcriptional regulator, asnC-family (247 aa)	Rv3050c.1	28 kDa	0.1	36.2	2.3
conserved hypothetical protein (168 aa)	Rv2574.1	19 kDa	0.1	101.2	7.0
ferredoxin-dependent glutamate synthase large subunit gltB (1528 aa)	Rv3859c.1	166 kDa	0.6	75.4	44.5
tRNA/rRNA methyltransferase (323 aa)	Rv3579c.1	34 kDa	4.4	4.6	20.0
conserved hypothetical protein (127 aa)	Rv2420c.1	14 kDa	6.7	8.8	59.2
lipoprotein lpqE (183 aa)	Rv3584.1	19 kDa	0.2	183.8	42.7
acyl-CoA dehydrogenase fadE10 (651 aa)	Rv0873.1	71 kDa	0.5	210.3	94.7
30S ribosomal protein S3 rpsC (275 aa)	Rv0707.1	30 kDa	1.4	193.9	279.6
conserved hypothetical protein (225 aa)	Rv2557.1	24 kDa	0.3	143.8	36.0
conserved hypothetical protein (326 aa)	Rv2410c.1	36 kDa	0.2	11.3	2.6
bacterioferritin bfrB (182 aa)	Rv3841.1	20 kDa	1.5	686.8	1051.4
branched-chain amino acid transaminase ilvE (369 aa)	Rv2210c.1	40 kDa	1.9	31.3	59.2
repressor lexA (218 aa)	Rv2720.1	23 kDa	0.2	27.7	6.3
DNA-directed RNA polymerase beta chain rpoC (1317 aa)	Rv0668.1	147 kDa	0.7	276.6	184.1
lyase (347 aa)	Rv3684.1	38 kDa	0.2	28.7	6.3
lipoprotein lprF (262 aa)	Rv1368.1	27 kDa	0.1	139.5	17.3
30S ribosomal protein S14 rpsN2 (102 aa)	Rv2056c.1	11 kDa	4.8	4.6	22.2
hypothetical protein (198 aa)	Rv0078A.1	22 kDa	2.8	7.8	21.9
conserved hypothetical protein (152 aa)	Rv3547.1	17 kDa	0.3	290.6	99.9
conserved membrane protein (329 aa)	Rv0347.1	37 kDa	0.4	27.9	11.3
soluble secreted antigen mpt53 precursor (174 aa)	Rv2878c.1	18 kDa	3.0	12.4	37.4
30S ribosomal protein S17 rpsQ (137 aa)	Rv0710.1	15 kDa	1.6	132.2	205.8
adenylate kinase adk (182 aa)	Rv0733.1	20 kDa	2.2	81.3	179.9
hypothetical protein (158 aa)	Rv1060.1	17 kDa	0.2	83.0	16.3
membrane protein (221 aa)	Rv1362c.1	24 kDa	0.0	28.4	1.2
cystathionine beta-synthase cbs (465 aa)	Rv1077.1	49 kDa	#DIV/0!	0.0	13.1
molybdopterin biosynthesis protein moeW (319 aa)	Rv2338c.1	35 kDa	0.1	63.0	6.0
short-chain type alcohol dehydrogenase/reductase (288 aa)	Rv3057c.1	31 kDa	0.1	17.8	1.0
oxidoreductase (283 aa)	Rv2971.1	30 kDa	1.9	48.4	93.0
3-hydroxybutyryl-CoA dehydrogenase fadB2 (287 aa)	Rv0468.1	31 kDa	0.7	220.5	149.7
translation initiation factor IF-2 infB (901 aa)	Rv2839c.1	94 kDa	2.4	24.5	58.6
lipoprotein lpqI (389 aa)	Rv0237.1	40 kDa	0.1	10.7	1.4

bacterioferritin comigratory protein bcp (158 aa)	Rv2521.1	17 kDa	2.1	27.3	56.0
ATP synthase beta chain atpD (487 aa)	Rv1310.1	53 kDa	0.7	485.3	335.6
medium chain fatty-acid-CoA ligase fadD14 (544 aa)	Rv1058.1	60 kDa	0.5	46.1	25.0
conserved hypothetical protein (119 aa)	Rv2204c.1	13 kDa	1.5	93.2	140.3
conserved membrane protein (325 aa)	Rv2037c.1	35 kDa	0.1	42.4	6.0
conserved hypothetical protein (330 aa)	Rv0493c.1	35 kDa	0.0	46.8	1.7

Appendix Table 9. Common proteins with significant different proteins in the clinical T pair using the linear ion trap (chapter II) and the hybrid orbitrap-LTQ MS/MS (chapter III)

Proteins significantly different (*t*-test, p<0.05)	Gene name	Functional category	Rv number	Trend in T INHr
iron-regulated peptidyl-prolyl-cis-trans-isomerase A ppiA	*ppiA*	IMR	Rv0009.1	Higher
isocitrate dehydrogenase icd2	*icd2*	IMR	Rv0066c.1	Higher
dihydroxy-acid dehydratase ilvD	*ilvD*	IMR	Rv0189c.1	Higher
fatty-acid-CoA ligase fadD2	*fadD2*	L	Rv0270.1	Higher
conserved membrane protein	*Rv0283*	C	Rv0283.1	Higher
protein transport protein secE2	*secE2*	CW	Rv0379.1	Higher
dihydrolipoamide dehydrogenase lpd	*lpd*	IMR	Rv0462.1	Lower
conserved hypothetical protein	*Rv0580*	C	Rv0580c.1	Higher
50S ribosomal protein L10 rplJ	*rplJ*	IP	Rv0651.1	Lower
DNA-directed RNA polymerase beta chain rpoB	*rpoB*	IP	Rv0667.1	Higher
DNA-directed RNA polymerase beta chain rpoC	*rpoC*	IP	Rv0668.1	Higher
iron-regulated elongation factor tu tuf	*tuf*	IP	Rv0685.1	Higher
30S ribosomal protein S10 rpsJ	*rpsJ*	IP	Rv0700.1	Lower
50S ribosomal protein L29 rpmC	*rpmC*	IP	Rv0709.1	Lower
protease IV sppA	*sspA*	CW	Rv0724.1	Higher
phosphoribosylformylglycinamidine cyclo-ligase purM	*purM*	IMR	Rv0809.1	Higher
hypothetical protein sseC2	*sseC2*	IMR	Rv0814c.1	Higher
thiosulfate sulfurtransferase cysA2	*cysA2*	IMR	Rv0815c.1	Lower
enoyl-CoA hydratase echA6	*echA6*	L	Rv0905.1	Higher
periplasmic phosphate-binding lipoprotein pstS1	*psts1*	IMR	Rv0934.1	Higher
enolase eno	*eno*	IMR	Rv1023.1	Higher
5-methyltetrahydropteroyltriglutamate-homocysteine methyltransferase metE	*metE*	IMR	Rv1133c.1	Higher
lipoprotein lprA	*lprA*	CW	Rv1270c.1	Higher
transcription termination factor rho	*rho*	IP	Rv1297.1	Higher
ATP synthase gamma chain atpG	*atpG*	IMR	Rv1309.1	Higher
ATP synthase beta chain atpD	*atpD*	IMR	Rv1310.1	Higher
ATP synthase beta chain atpD	*atpD*	IMR	Rv1311	Higher
lipoprotein lprF	*lprF*	CW	Rv1368.1	Higher
oxpp cycle protein opcA	*opcA*	IMR	Rv1446c.1	Higher
alanine and proline rich secreted protein apa	*apa*	CW	Rv1860.1	Higher
conserved hypothetical protein	*Rv1871*	C	Rv1871c.1	Higher
L-lactate dehydrogenase lldD2	*lldD2*	IMR	Rv1872c.1	Higher
conserved hypothetical protein	*Rv1906*	C	Rv1906c.1	Higher
catalase-peroxidase-peroxynitritase T katG	*katG*	V	Rv1908c.1	Lower
immunogenic protein mpt63	*mpt63*	CW	Rv1926c.1	
conserved hypothetical protein	*Rv2054*	C	Rv2054.1	Higher
conserved hypothetical protein	*Rv2159*	C	Rv2159c.1	Higher
long-chain fatty-acid-CoA ligase fadD15	*fadD15*	L	Rv2187.1	Higher
glutamine synthetase glnA1	*glnA1*	IMR	Rv2220.1	Lower
meromycolate extension acyl carrier protein	*acpM*	L	Rv2244.1	Lower
haloalkane dehalogenase	R*v2296*	IMR	Rv2296.1	Higher
conserved hypothetical protein	*Rv2298*	C	Rv2298.1	Higher

low molecular weight protein antigen	*cfp2*	CW	Rv2376c.1	Higher
alkyl hydroperoxide reductase C protein	*ahpC*	V	Rv2428.1	Lower
40 kda secreted L-alanine dehydrogenase	*ald*	IMR	Rv2780.1	Higher
bifunctional polyribonucleotide nucleotidyltransferase	*gpsI*	IMR	Rv2783c.1	Lower
conserved hypothetical protein	*Rv3267*	C	Rv3267.1	Higher
10 kda chaperonin	*groES*	V	Rv3418c.1	Lower
30S ribosomal protein S4	*rpsD*	IP	Rv3458c.1	Higher
bifunctional membrane-associated penicillin-binding protein 1A/1B	*ponA2*	CW	Rv3682.1	Higher
conserved hypothetical protein	*Rv3705*	C	Rv3705c.1	Higher
conserved hypothetical protein	*Rv3716*	C	Rv3716c.1	Lower
bacterioferritin bfrB	*bfrB*	V	Rv3841.1	Lower
conserved hypothetical protein	*Rv3867*	C	Rv3867.1	Higher
10 kda culture filtrate antigen	*esxB*	CW	Rv3874.1	Higher
thioredoxin reductase	*trxB2*	V	Rv3913.1	Lower

Extracted Ion chromatogram and Spectra of significantly different features in the mycolic acid

comparison of the clinical pair obtained through XCMS online platform.
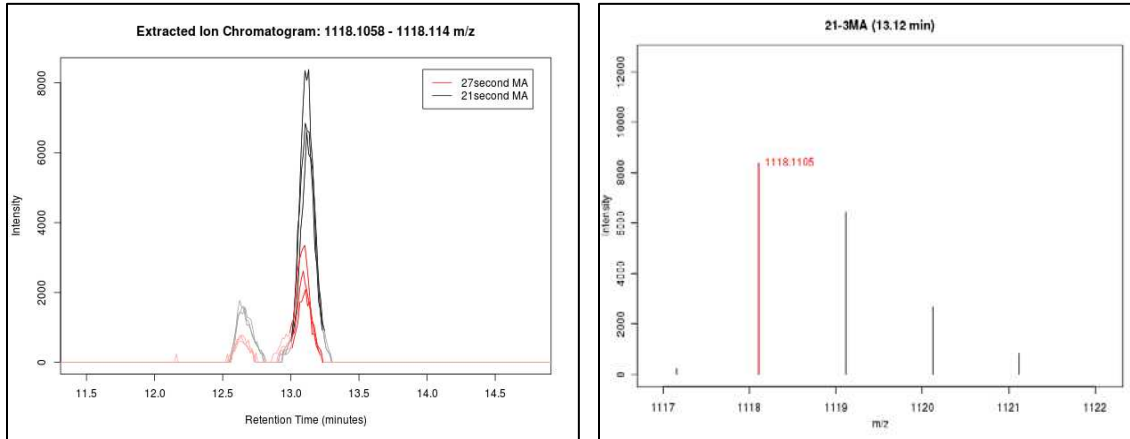
Appendix Figure 1. Alpha-MA (C75) identified with Na+ adduct - Extracted Ion Chromatogram
and Spectra
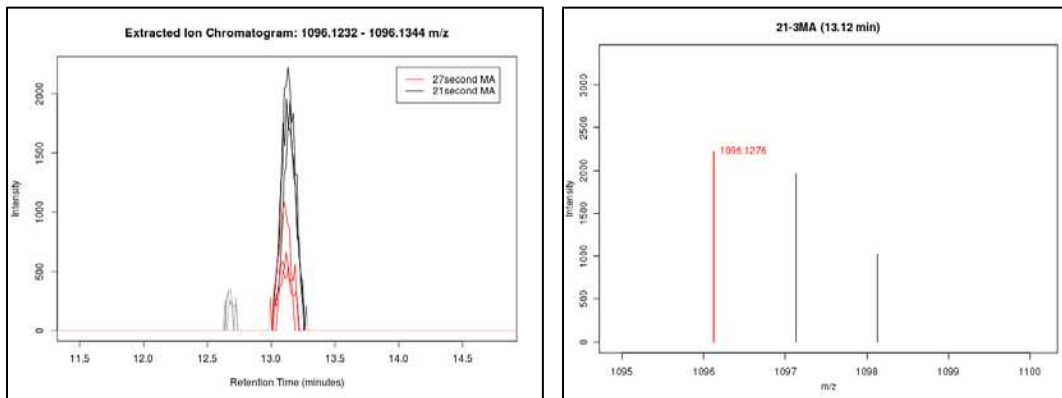
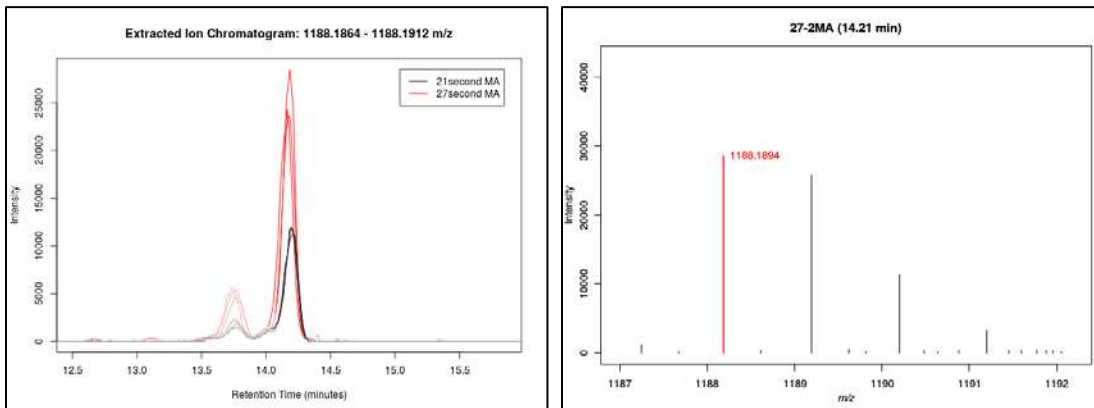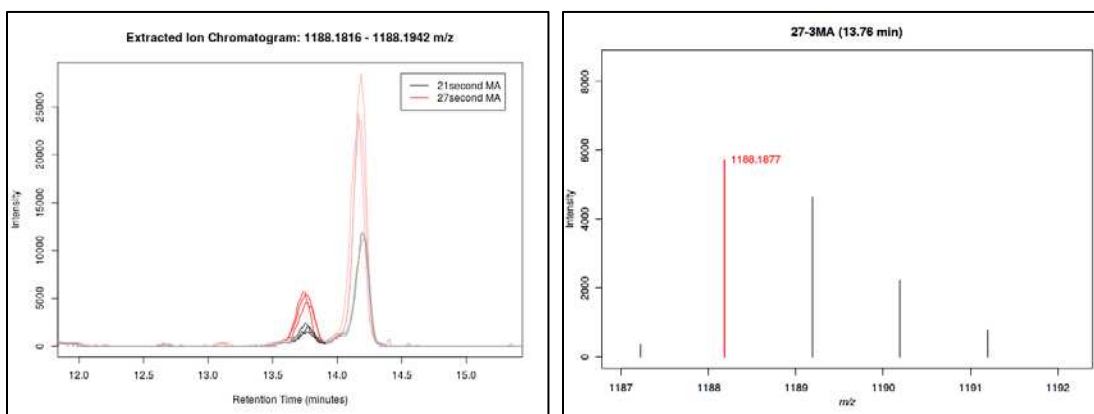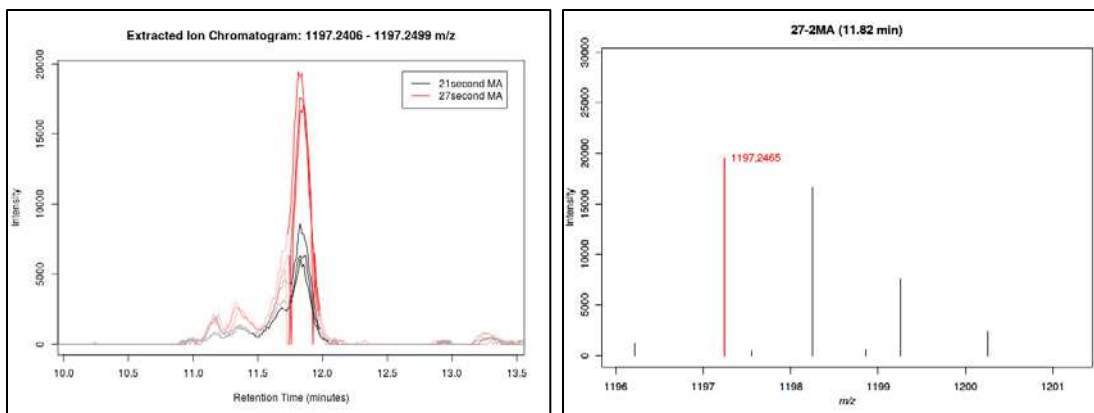*Appendix Figure 2. Alpha-MA (C75) identified with H+ adduct - Extracted Ion Chromatogram
and Spectra.*
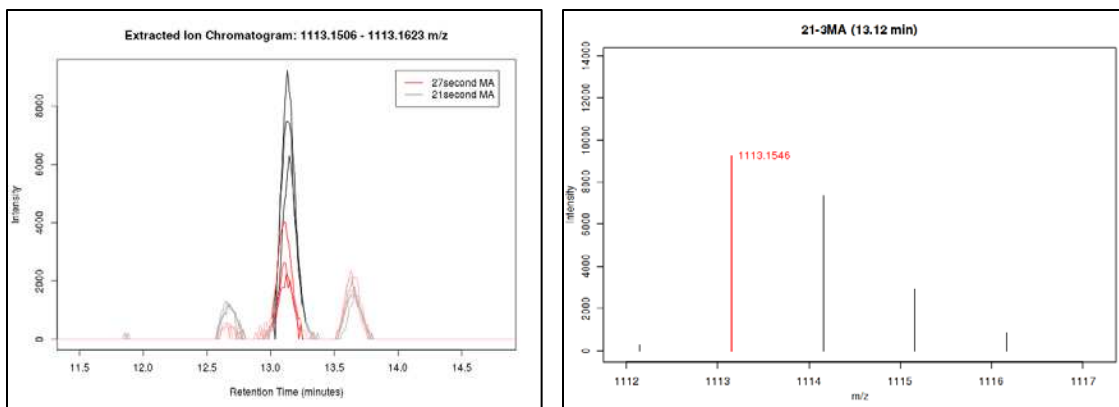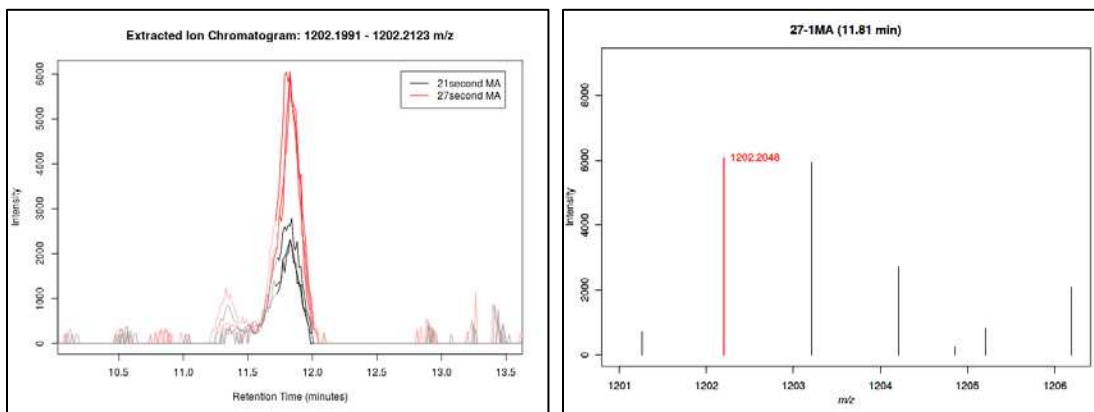
Appendix Figure 3. Alpha-MA (C75) identified with Na+ adduct - Extracted Ion Chromatogram and Spectra.
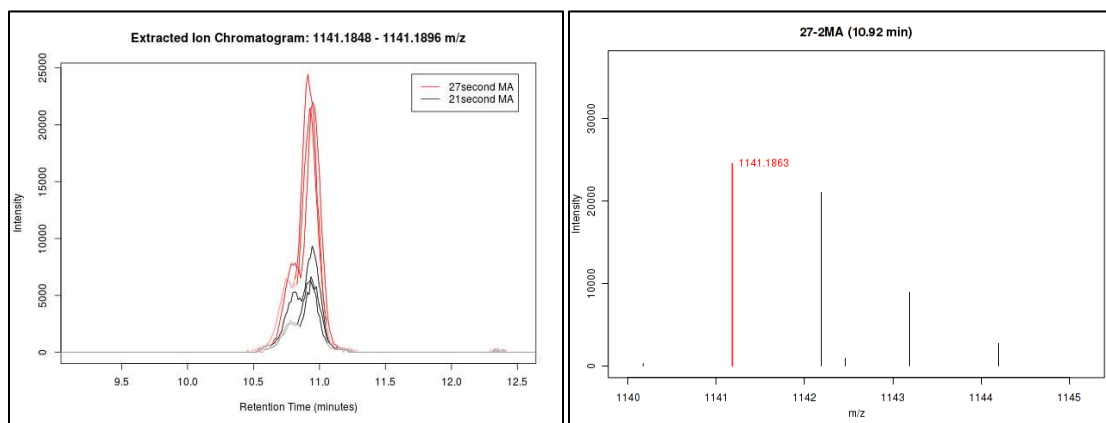
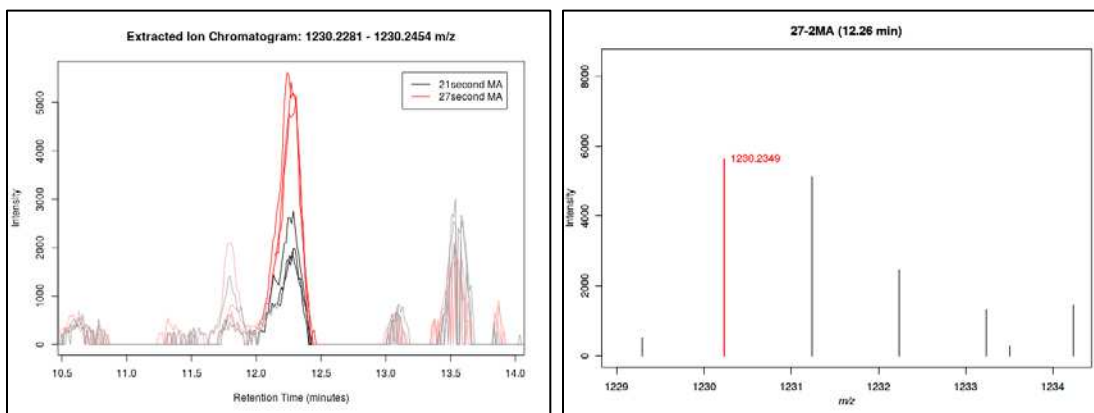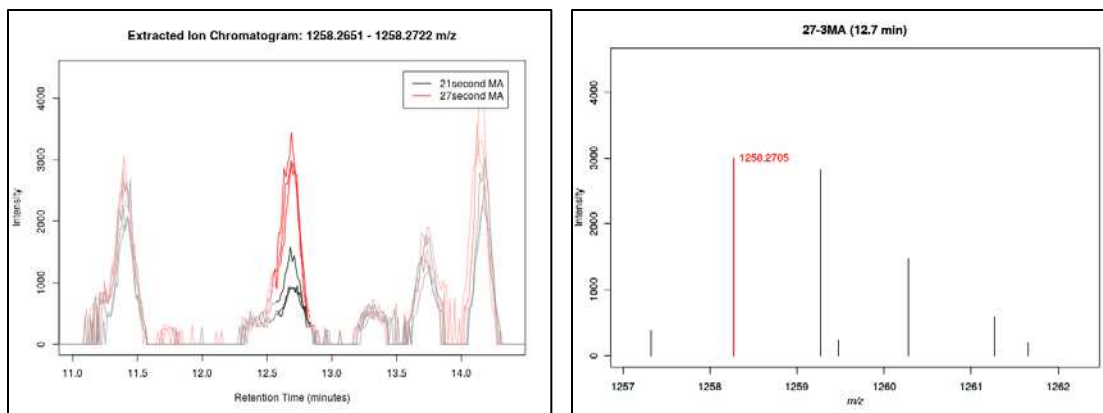Appendix Figure 4. Alpha-MA (C77) identified with Na+ adduct - Extracted Ion Chromatogram and Spectra.
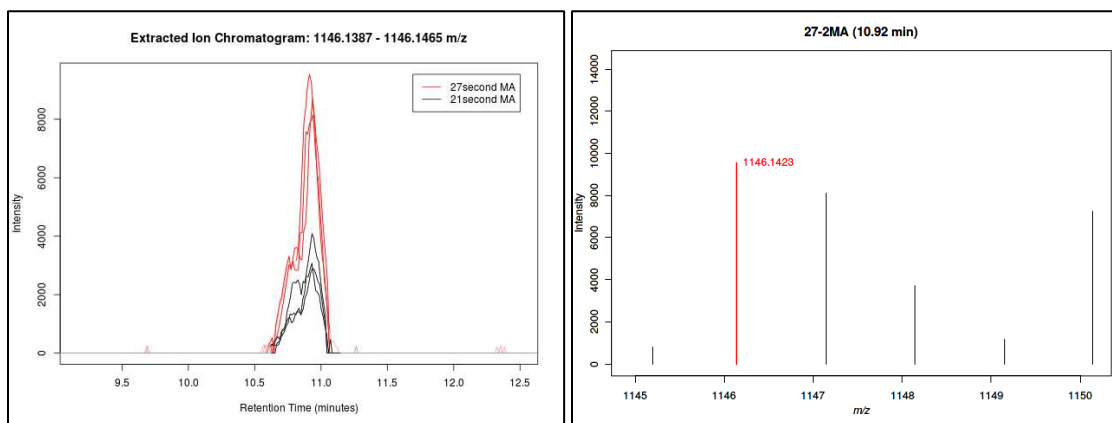
Appendix Figure 5. Alpha-MA (C77) identified with NH_4+ adduct - Extracted Ion Chromatogram and Spectra.
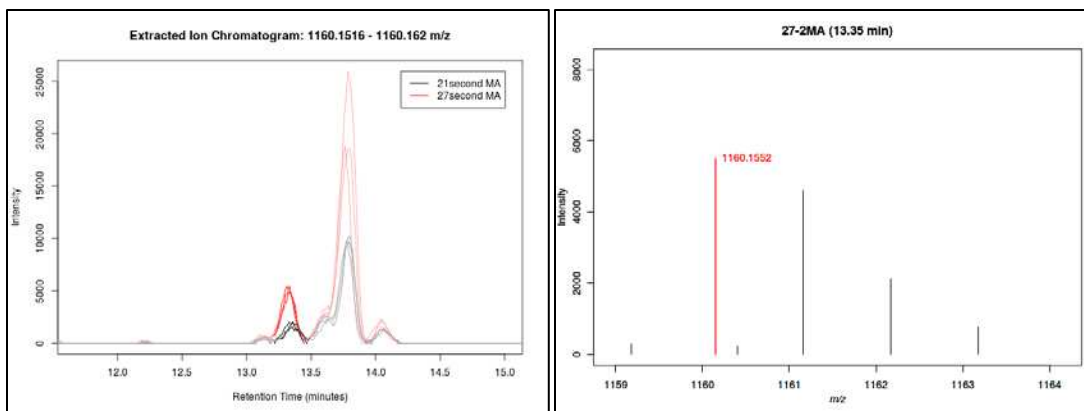
Appendix Figure 6. Alpha-MA (C78) identified with Na+ adduct - Extracted Ion Chromatogram and Spectra.
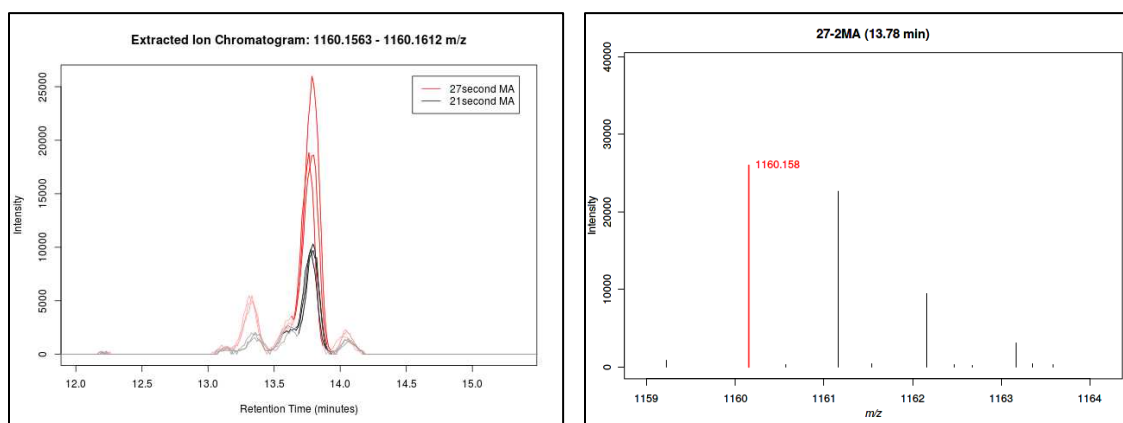
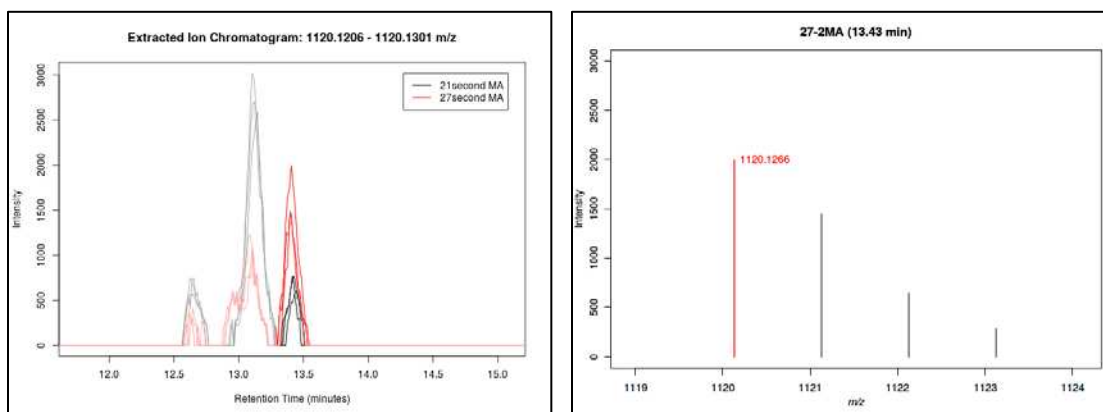Appendix Figure 7. Alpha-MA (C78) identified with Na+ adduct - Extracted Ion Chromatogram and Spectra.
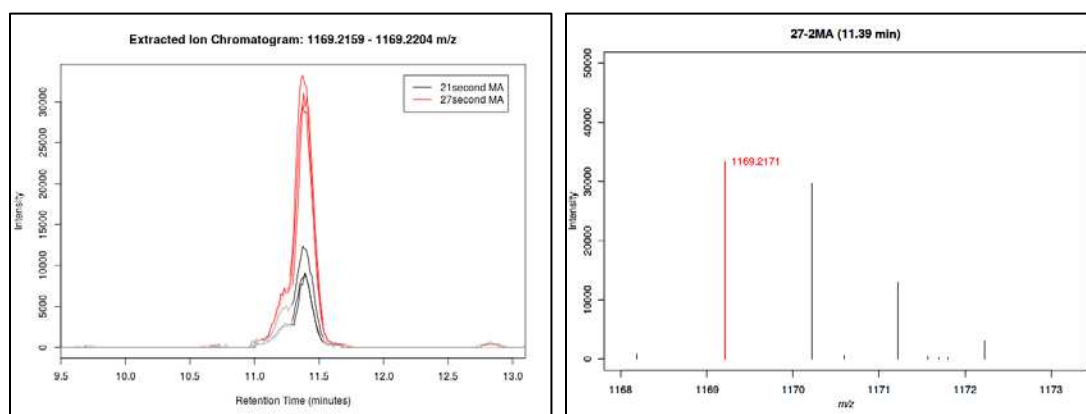
Appendix Figure 8. Alpha-MA (C79) identified with Na+ adduct - Extracted Ion Chromatogram and Spectra.

221
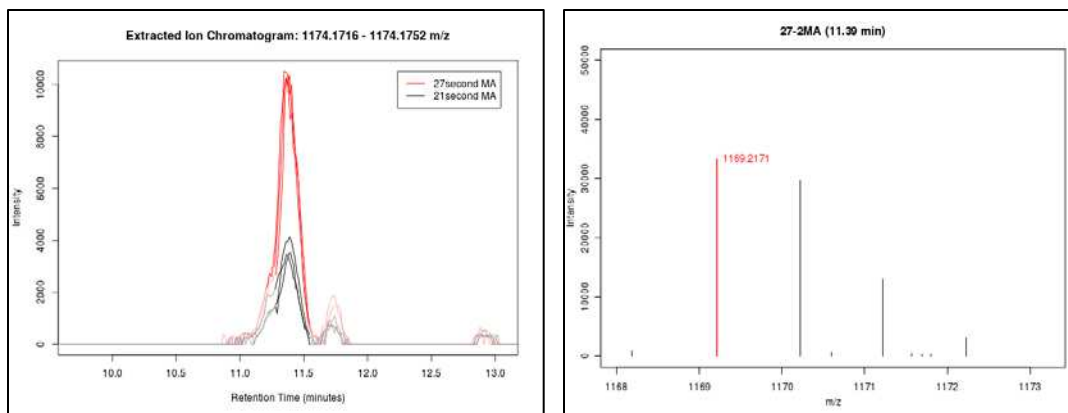
Appendix Figure 9. Alpha-MA (C79) identified with NH$_4$+ adduct - Extracted Ion Chromatogram and Spectra.

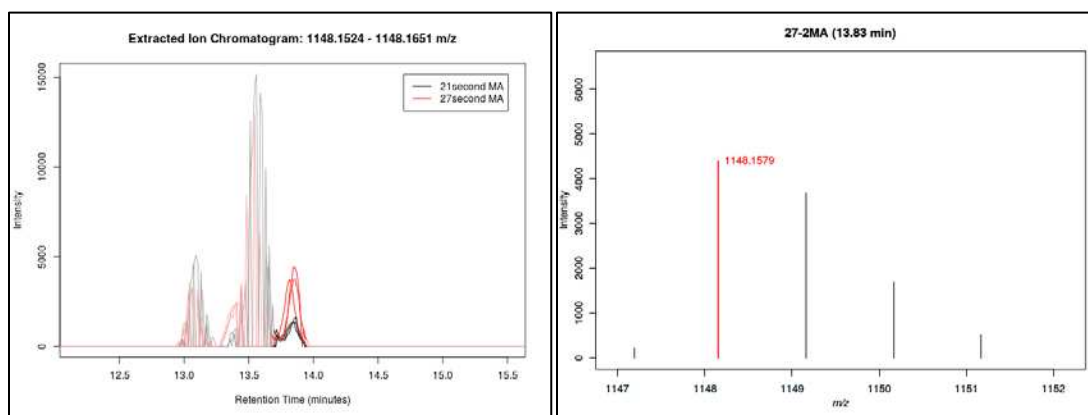Appendix Figure 10. Alpha-MA (C80) identified with Na+ adduct - Extracted Ion Chromatogram and Spectra.

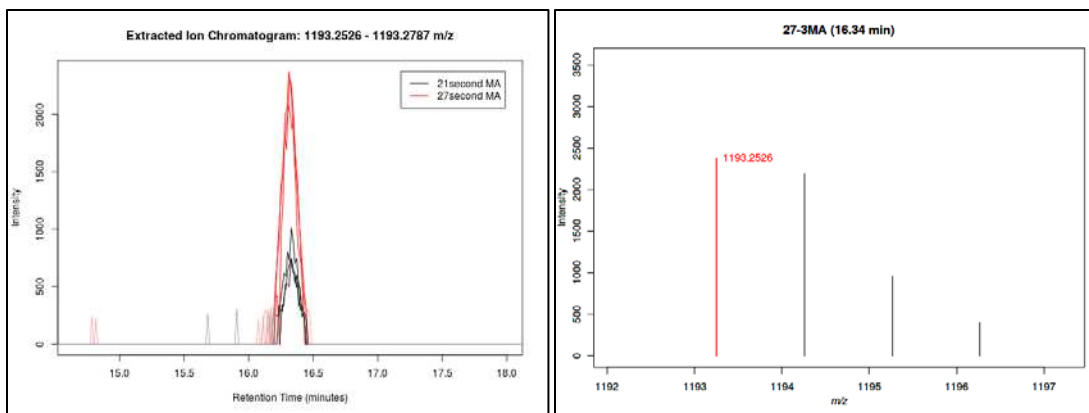Appendix Figure 11. Alpha-MA (C80) identified with Na+ adduct - Extracted Ion Chromatogram and Spectra.

Appendix Figure 12. Alpha-MA (C81) identified with NH_4+ adduct - Extracted Ion Chromatogram and Spectra.

Appendix Figure 13. Alpha-MA (C81) identified with Na+ adduct - Extracted Ion Chromatogram and Spectra.

Appendix Figure 14. Alpha-MA (C83) identified with Na+ adduct - Extracted Ion Chromatogram and Spectra.

Appendix Figure 15. Alpha-MA (C85) identified with Na+ adduct - Extracted Ion Chromatogram and Spectra.
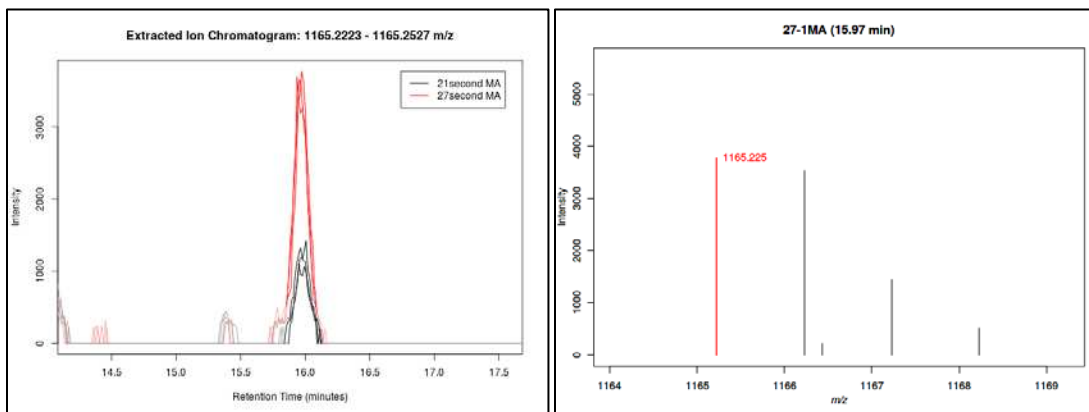

Appendix Figure 16. Keto-MA (C74) identified with Na+ adduct - Extracted Ion Chromatogram and Spectra.
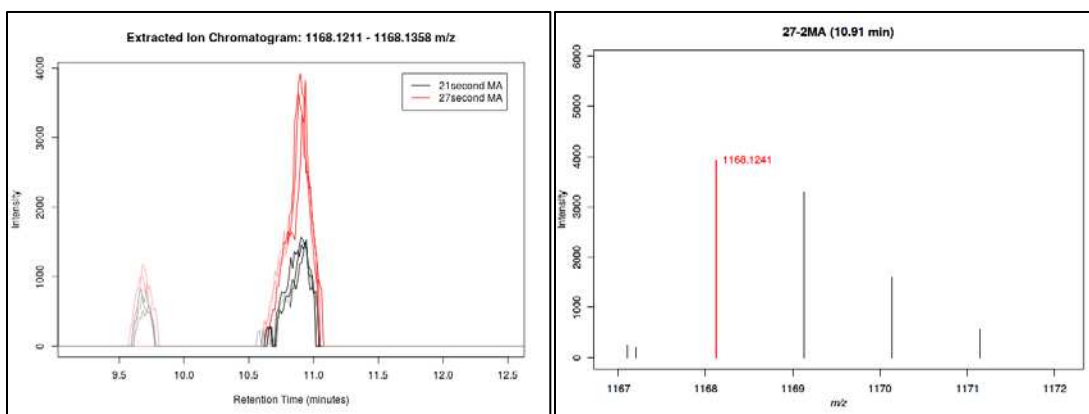

Appendix Figure 17. Keto-MA (C80) identified with Na+ adduct - Extracted Ion Chromatogram and Spectra.

Appendix Figure 18. Unidentified feature 21- Extracted Ion Chromatogram and Spectra.

Appendix Figure 19.Unidentified feature 22- Extracted Ion Chromatogram and Spectra.

Appendix Figure 20.Unidentified feature 23- Extracted Ion Chromatogram and Spectra.

Extracted Ion chromatogram and Spectra of two significantly different features in the mycolic

acid comparison of the laboratory pair obtained through XCMS online platform. (Note: The

quality of the spectra were acceptable, however, the quality of the extracted ion chromatogram

did not allow the validation of these features.

Appendix Figure 21. Unidentified feature 1- Extracted Ion Chromatogram and Spectra.

Appendix Figure 22.Unidentified feature 2- Extracted Ion Chromatogram and Spectra.