

Functional Characterization of Nucleoid Associated Proteins Acting as Global Transcription Factors in *Mycobacterium tuberculosis*

THÈSE N° 8176 (2017)

PRÉSENTÉE LE 6 DÉCEMBRE 2017
À LA FACULTÉ DES SCIENCES DE LA VIE
UNITÉ DU PROF. COLE

PROGRAMME DOCTORAL EN APPROCHES MOLÉCULAIRES DU VIVANT

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

POUR L'OBTENTION DU GRADE DE DOCTEUR ÈS SCIENCES

PAR

Nina Theres ODERMATT

acceptée sur proposition du jury:

Prof. B. Lemaitre, président du jury
Prof. S. Cole, directeur de thèse
Prof. R. Brosch, rapporteur
Prof. R. Manganelli, rapporteur
Prof. M. Blokesch, rapporteuse



ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Suisse
2017

CONTENTS

Abstract		iii
Zusammenfassung		v
Abbreviations		vii
Chapter 1	Introduction	1
Chapter 2	Rv3852 (H-NS) of <i>Mycobacterium tuberculosis</i> Is Not Involved in Nucleoid Compaction and Virulence Regulation	29
Chapter 3	Characterization of mIHF	
Chapter 3.1	Essentiality of <i>mihF</i> and Impact of Gene Silencing on <i>Mycobacterium tuberculosis</i> Physiology and Transcriptional Landscape	49
Chapter 3.2	Structure and DNA-binding Mechanism of <i>Mycobacterium tuberculosis</i> mIHF	89
Chapter 4	Activity and Mode of Action of Chrysomycins in <i>Mycobacterium tuberculosis</i>	113
Chapter 5	Conclusions and Perspectives	125
Curriculum Vitae		131
Acknowledgements		135

ABSTRACT

The fatal lung disease tuberculosis is caused by the airborne *Mycobacterium tuberculosis*, a versatile pathogen adapted to rapidly changing environments. Instead of being eradicated by phagocytic cells of its human host, bacilli tune macrophages to support their own growth and even mask their presence from the immune system for several decades. Rapid adjustment of gene expression is critical for bacterial survival and heavily relies on nucleoid-associated proteins (NAPs). NAPs contribute to active DNA management by altering the chromosomal topology through bending, bridging and looping the DNA. These conformational changes can bring distant genetic loci into close spatial proximity or influence DNA supercoiling and therefore accessibility of the transcription machinery. Apart from their architectural role, NAPs moreover act as global transcription factors by direct regulation of numerous genes.

In *M. tuberculosis*, five proteins were assigned a role as NAP. Among these, EspR, HupB and Lsr2 are crucial not only for virulence but also for cellular metabolism. This thesis focuses on the NAPs mIHF and H-NS with the objective of determining their function and target regulon. I investigated both proteins by means of genetic manipulation, phenotype assessment and structural studies, and additionally assessed the efficiency of a potential new anti-tuberculosis drug acting on Lsr2.

Chrysomycin, described as specific inhibitor of Lsr2-DNA complex formation, was found to intercalate into the DNA. The resulting toxic effect on both its target *M. tuberculosis* as well as on eukaryotic cells rendered further development of the compound as an anti-tuberculosis drug futile.

We demonstrated that Rv3852, formerly annotated as H-NS, does not act as a NAP. Deletion of the *rv3852* gene had no effect on the *in vitro* phenotype of *M. tuberculosis*, did not alter nucleoid spread nor position and had no influence on virulence in mice.

The mIHF protein on the other hand is not only essential for active bacterial growth, but also indispensable for survival. Generation of a conditional knockdown mutant showed that depletion of mIHF led to elongated cells devoid of septa with abnormal DNA localization and finally to cell death. The target regulon of mIHF was thoroughly studied by mapping its binding sites on the bacterial genome and by identifying genes that were differentially expressed upon depletion of the protein. We found that mIHF has a strong effect on virulence gene expression and, similar to EspR, possesses a major binding site upstream of one of the main virulence factor operons *espACD*. Analysis of the transcriptional response revealed that mIHF is further involved in the bacterial response to the host's immune system, including control of nutrient pathways as well as global protein and nucleic acid synthesis. To define how mIHF interacts with DNA and influences its three-dimensional organisation, the protein structure of mIHF was determined by nuclear magnetic resonance spectroscopy. Binding of mIHF introduced left-hand loops into linear as well as supercoiled DNA substrates, therefore unwinding condensed DNA. We identified two DNA binding domains in mIHF and showed that its stability increased substantially upon DNA binding.

All together, the findings of this thesis contribute to a better understanding of the complex gene regulatory network of *M. tuberculosis*, advancing the knowledge necessary to eventually defeat tuberculosis, a disease that has plagued humanity for millennia.

Keywords: *Mycobacterium tuberculosis*, global transcription factor, nucleoid associated protein, molecular genetics, phenotype characterization, RNA-seq, ChIP-seq, mIHF (*rv1388*), NMR

ZUSAMMENFASSUNG

Die tödliche Lungenerkrankung Tuberkulose, auch genannt Schwindsucht, wird verursacht durch *Mycobacterium tuberculosis*, welches durch die Luft von Mensch zu Mensch übertragen werden kann. Das vielseitige Bakterium passt sich den schnell ändernden Umweltbedingungen an und kann sich nicht nur vor dem Immunsystem über Jahrzehnte verbergen, sondern sogar Makrophagen, deren einzige Bestimmung es ist, eindringende Krankheitserreger zu eliminieren, für ihren eigenen Zweck zu manipulieren, um das bakterielle Wachstum in ihrem Inneren zu fördern. Steuerung der Genexpression ist kritisch für das Überleben des Tuberkelbazillus in diesem feindlichen Umfeld und basiert zu einem grossen Teil auf den Nukleoid-assoziierten Proteinen (NAP), welche die Topologie der DNS aktiv ändern, indem sie das bakterielle Chromosom biegen, Brücken oder Schleifen bilden. Diese dreidimensionale Anordnung kann entfernt liegende Gene in räumliche Nähe bringen und die globale Kondensation der DNS beeinflussen. Neben dieser architektonischen Funktion, können NAP durch Aktivierung oder Repression ihre Zielgene direkt regulieren und wirken deshalb auch als globale Transkriptionsfaktoren.

In *M. tuberculosis* wurden bisher fünf Proteine als NAP identifiziert; EspR, HupB und Lsr2 sind nicht nur für die Virulenz, sondern auch für den zellulären Stoffwechsel von entscheidender Bedeutung. Diese Doktorarbeit konzentriert sich auf die verbleibenden zwei NAP in *M. tuberculosis*, mIHF und H-NS, mit dem Ziel der genauen Definition ihrer Funktionen. Ich habe beide Proteine mittels Genmanipulation, Phänotyp-Analyse und Strukturstudien erforscht und zusätzlich die Wirksamkeit eines potenziellen Vorläufers eines Tuberkulose-Medikaments, das auf Lsr2 wirkt, untersucht.

Chrysomycin wurde als spezifischer Inhibitor der Lsr2-DNS-Komplexbildung beschrieben, jedoch konnte ich beweisen, dass sich dieses natürliche Präparat unspezifisch in die DNS einlagert. Die daraus resultierende toxische Wirkung sowohl auf das eigentliche Ziel *M. tuberculosis*, als auch auf den eukaryotischen Wirt hatte zur Folge, dass die Weiterentwicklung des Wirkstoffs zu einem Anti-Tuberkulose-Medikament als zwecklos eingestuft wurde.

Wir konnten nachweisen, dass Rv3852, zuvor als H-NS gehandelt, nicht als NAP fungiert, da die Entfernung des Gens *rv3852* vom *M. tuberculosis* Chromosom keinen Einfluss auf den *in vitro* Phänotyp hatte, weder eine Veränderung in der DNS-Ausbreitung noch in der Position gefunden wurde, und keinen Einfluss auf die Virulenz bei Mäusen zeigte.

Das mIHF-Protein hingegen hatte einen weitaus grösseren Effekt auf *M. tuberculosis*, da es nicht nur für aktives Wachstum unentbehrlich ist, sondern auch notwendig für das Überleben der Bakterien. Ein gentechnisch veränderter *M. tuberculosis* Stamm mit einer konditionell herunterfahrbaren Variante des *mihF* Genes bestätigte, dass *mihF* essentiell ist für die Bakterien, und ein niedriger mIHF Spiegel führte zu verlängerten Zellen ohne Septum mit nur einem Nukleoid. Die durch mIHF kontrollierten Gene wurden ermittelt durch Analyse der Kontaktstellen des Proteins mit der DNS und verglichen mit der Identifizierung derjenigen Gene, die bei tiefem mIHF Spiegel differenziell exprimiert waren. mIHF hatte einen starken Einfluss auf Gene, die, ähnlich wie EspR, mit Virulenzfunktionen assoziiert sind. Eine Hauptbindestelle von mIHF befand sich oberhalb dem wichtigen Virulenzoperon *espACD*. Insgesamt war mIHF involviert in die bakterielle Reaktion auf das Wirt-Immunsystem, einschließlich Nährstoff-Stoffwechselwegen und Nukleinsäuren- sowie Protein-Synthese.

Um zu definieren, wie mIHF mit der DNS interagiert und deren dreidimensionale Organisation beeinflusst, wurde die Proteinstruktur von mIHF durch nukleare magnet-resonanz Spektroskopie bestimmt. mIHF zeigte eine Bindung an lineare und superspiralisierte DNS-Substrate, führte zu links-gewinkelten Schleifen in kondensierter DNS und entwindete sie dadurch. Im globularen Teil des mIHF Proteins konnten zwei verschiedene DNS-Bindungsdomänen identifiziert werden, und die Stabilität von mIHF wurde wesentlich erhöht, wenn es in Kontakt mit DNS stand.

Zusammenfassend tragen die Ergebnisse dieser Arbeit zu einem besseren Verständnis der komplexen Genregulation in *M. tuberculosis* bei, und erweitern das notwendige Wissen, um hoffentlich eines Tages die Tuberkulose, eine Krankheit, die die Menschheit schon seit Jahrtausenden plagt, endgültig zu besiegen.

Stichworte: *Mycobacterium tuberculosis*, Globaler Transkriptionsfaktor, Nucleoid-assoziiertes Protein, molekulare Genetik, Phänotyp-Charakterisierung, RNA-seq, ChIP-seq, mIHF (*rv1388*), NMR

ABBREVIATIONS

AFM	atomic force microscopy
ATc	anhydrotetracycline
BCG	Bacille de Calmette et Guérin
CD	circular dichroism
ChIP-seq	chromatin-immuno precipitation followed by DNA-fragment sequencing
DBD	DNA binding domain
EMSA	electrophoretic mobility shift assay
Esp	ESX secretion associated protein
EspR	ESX-1 secreted protein regulator, a NAP
GI	genomic island
H-NS	histone-like nucleoid structuring protein, a NAP
HupB	heat-unstable protein, a NAP
Hyg	hygromycin
ITC	isothermal titration calorimetry
Kan	kanamycin
Lsr2	iron-regulated H-NS-like protein, a NAP
MIC	minimal inhibitory concentration
mIHF	mycobacterial integration host factor, a NAP
Mtb	<i>Mycobacterium tuberculosis</i>
NAP	nucleoid associated protein
NMR	nuclear magnetic resonance
NOE	Nuclear Overhauser effect
OD	optical density
PCR	polymerase chain reaction
PDIM	phthiocerol dimycocerosate
RNA-seq	whole transcriptome sequencing
Str	streptomycin
T7SS	type-VII secretion system
TCS	two-component system
Tm	melting temperature
TSS	transcription start site
TTS	transcription termination site
WHO	world health organization

CHAPTER 1

Introduction

Tuberculosis	3
History and Epidemiology	3
Host-Pathogen Interaction.....	4
Diagnostics, Vaccines, Treatment and Drug Discovery	6
Microbiology of <i>M. tuberculosis</i>	9
Cell Structure	9
Secretion Systems.....	10
Genomics and Genetics	11
Physiology and Metabolism	12
Gene Regulation	12
Transcription Regulatory Proteins	13
Global Transcription Factors	14
HupB	14
H-NS	15
EspR	15
mIHF	16
Lsr2	16
Other NAPs	17
Thesis Context and Rationale.....	17
References	20

TUBERCULOSIS

Tuberculosis claims 1.8 million lives each year, every third person in the world is estimated to be latently infected and this contagious airborne disease ranks among the top 10 causes of death worldwide. Not only is the infection itself highly dangerous, but also the existing vaccine is not effective enough to grant immunity, diagnosis is difficult and treatment is lengthy with harmful side effects (WHO, 2016)(WHO, 2014).

Mycobacterium tuberculosis (Mtb) is the causative agent of human tuberculosis. When the bacterium enters the lungs of its host via the airways, it often establishes a latent infection without inducing any symptoms and can rest inside macrophages for decades. Instead of being eliminated by these phagocytes, Mtb modulates the cellular physiology to create an environment that supports its propagation before it is contained by the immune system. The bacilli can regain growth later in the individual's life and cause active tuberculosis, which leads to death, when untreated, in 50% of the cases. How these bacteria can adapt their gene regulation to a very rapidly changing environment from entering the host over phagocytosis by macrophages to remaining for years seemingly inactive remains elusive. Sophisticated regulatory mechanisms are necessary to orchestrate global gene expression via highly specific activation or repression of transcription. This PhD thesis investigated and characterized global transcription factors of Mtb.

HISTORY AND EPIDEMIOLOGY

The first encounter between man and tuberculosis dates very far back. *Homo erectus*, first emerging nearly 2 million years ago, might have suffered from tuberculosis, as it was speculated based on a 500,000 years old find (Kappelman, 2008). Though questioned that tuberculosis occurred so early in human history (Roberts, 2009), it is accepted that tuberculosis accompanied early *Homo sapiens* at the out-of-Africa migration 70,000 years ago (Hershberg, 2008). The devastating disease was mentioned

in a written document for the first time in the Old Testament about 3000 - 2500 years ago (Daniel, 1999), soon described by Hippocrates (460 – 370 B.C.) as phthisis and later referred to as “consumption”, due to the emaciated look of tuberculosis patients. The lengthy interaction between the pathogen and its human host allowed genomic adaptations on both sides, resulting in differences in tuberculosis susceptibility and various genotypes of the pathogen associated with its virulence. The reciprocal effect on genomes of both sides led to co-evolution of this particular host-pathogen pair (Gagneux, 2012).

Tuberculosis reached its peak of incidence in Europe during the industrial revolution around 1760, when increased population size and hence crowded living, poor hygienic standards and malnutrition especially in the lower class, led to a major epidemic. In the beginning of the 19th century, tuberculosis was the reason for every seventh death (Murray, 2015). The cause of this so called “white plague” was not known then, and it was erroneously thought to be hereditary as stated by Hippocrates (Cambau, 2014), because several tuberculosis cases often appeared in the same family. Only when Robert Koch stained bacilli from human and animal lesions in 1882, was the etiological agent of tuberculosis discovered. Koch infected guinea pigs with samples extracted from the lungs of tuberculosis patients and confirmed that these bacteria were responsible for causing tuberculosis. Furthermore, he isolated the slow growing Mtb in pure culture, hence fulfilling all four criteria of his postulates to establish a causative relationship between a disease and its causing microbe (Koch, 1882). One year later, the bacterium was officially named *Mycobacterium tuberculosis*, and with the knowledge that Mtb was transmitted from human to human, and therefore contagious rather than hereditary, prevention of active transmission slowly began. Together with the rise of the middle class which could afford better food and overall improved living standards, tuberculosis incidence started to decrease (Murray, 2015).

Tuberculosis was thought to be incurable for a long time. When it was discovered that people living in higher altitudes seldom fell ill with

tuberculosis, it was assumed that “tuberculosis-free” places existed. Subsequently, sanatoria were built in elevated regions, the first one in Switzerland was located in Leysin, 1854, and a little later the famous Schatzalp in Davos was opened. Patients were prescribed healthy food, cold-water showers and exposing their body to the sun in the clean alpine air. Nowadays, the extensive southward facing balconies of the sanatoria testify to the ancient treatment method of sunbathing. Although proof of cure from tuberculosis was absent until the end of that era, several other luxurious hotel-like sanatoria as well as publicly accessible ones for the less wealthy were built until the 1950s (Stiftung Historisches Lexikon der Schweiz, 2001). At present, the number of tuberculosis cases in Switzerland is slowly declining since 1996 from about 780 to 478 cases in 2007. Since then, a slight increase occurred, most probably caused by higher immigration numbers. 21 deaths per year were attributed to tuberculosis in Switzerland in 2015, 130 tuberculosis cases were hospitalized and 50 thereof were resistant to at least one drug (WHO, 2016).

The World Health Organization (WHO) surveys tuberculosis worldwide, estimates the number of patients and treatment success, recommends a treatment regimen and publishes yearly a global tuberculosis report. Although the number of tuberculosis deaths declined in the past 15 years by 22%, tuberculosis ranks ninth in the causes of deaths worldwide. An estimated 1.4 million tuberculosis deaths occurred in 2015 and an additional 400,000 people co-infected with HIV died of tuberculosis. Over 10 million new tuberculosis cases were estimated in 2015, among them 580,000 cases of drug-resistant tuberculosis. Additionally, 2 – 3 billion people have latent tuberculosis, which means that one third of the people living in the world carry Mtb, but have not fallen ill yet with the disease. Six countries accounted for 60% of new tuberculosis cases in 2016 (India, Indonesia, China, Nigeria, Pakistan and South Africa), which shows that the disease is more prevalent in certain high burden regions. Europe has a relatively small proportion with only 3% of all tuberculosis cases (WHO, 2016).

Although global tuberculosis numbers decrease slowly, co-morbidity with HIV and type 2 diabetes, as well as multidrug-resistant-tuberculosis, make it difficult to reach WHO’s “end-tuberculosis strategy”, which aims to decrease tuberculosis incidence by 95% in 2035. Unchanged since the 18th century is the fact that tuberculosis remains associated with socio-economic factors like overcrowding and poverty (Lawn, 2011).

Host-Pathogen Interaction

Inhaling an aerosol droplet containing minute numbers of bacilli is sufficient for an infection (Kaufmann, 2001). Mtb then travels down the airways into the lungs and causes pulmonary tuberculosis, which is the main outcome (85% of the cases). Only 15% of all infections involve non-pulmonary sites (Farer, 1979). Although more than 2 billion people are infected with Mtb, only 2 – 23% will develop active tuberculosis (Parrish, 1998). The first and most common symptom of active pulmonary tuberculosis is coughing. With disease progression, the increased inflammation and tissue necrosis shows in blood stained sputum, accompanied by fever, sweating and weight loss (Lawn, 2011). Untreated, tuberculosis has a high case fatality of over 50% (Tiemersma, 2011).

Once the bacilli have arrived in the lungs, the innate immune system tries to eliminate the intruders. Alveolar macrophages and recruited dendritic cells recognize Mtb with pattern recognition receptors for example complement receptors, mannose receptors or Toll-like receptors (TLRs). The TLRs, first shown to be important for host immunity against fungal infections in *Drosophila* (Lemaitre, 1996) were later found to recognize bacteria extra- and intracellularly (Schlesinger, 1996; Torrelles, 2017). The phagocytes take up the bacteria (Schlesinger, 1996; Torrelles, 2017), and are then engulfed themselves by neutrophils (Eum, 2010). The first compartment Mtb encounters inside the macrophage is the phagosome, purely designed to inactivate any particle it contains by acidification, production of reactive oxygen species and finally, by fusion with the lysosome, which releases hydrolytic enzymes

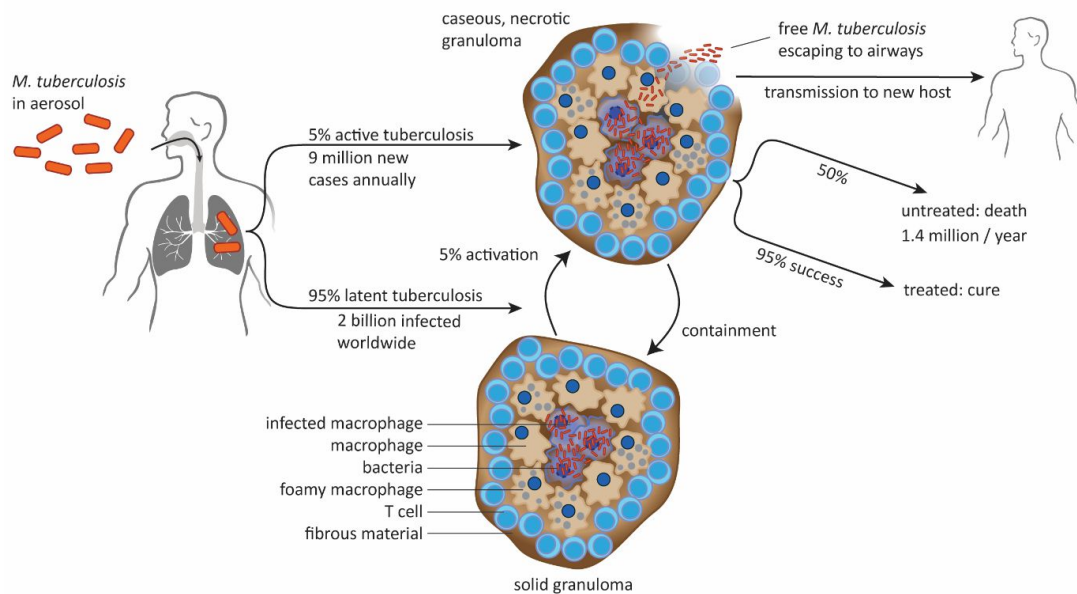


Figure 1: Infection cycle of *M. tuberculosis*. Transmission of Mtb occurs via aerosols containing infectious bacteria. After reaching the lungs, only few cases develop active tuberculosis directly, while the majority can contain the bacteria in solid granulomas, where macrophages are enclosed by T cells and fibrous material. Healthy individuals can control the infection for a long time, but stay latently infected with the risk of lifelong activation, maintaining a reservoir of two billion latently infected individuals. Upon activation, necrosis occurs from the inside and the granuloma becomes caseous. Mtb grows to high numbers, reaches the blood stream and airways and is finally transmitted to a new host by coughing contagious aerosols, leading to an estimated 50 million infections per year. If untreated, active tuberculosis has a high fatality rate of 50%. Drug-sensitive tuberculosis is cured with a set of four different antibiotics and shows a good success rate of 95%. (After (Gengenbacher, 2012; Zumla, 2013)).

and antimicrobial peptides. However, instead of being degraded, Mtb prevents phagosome maturation and the subsequent fusion with the lysosome. The mycobacterial cell wall component lipoarabinomannan inhibits the signalling cascade that induces fusion and fission necessary for phagosome maturation (Johansson, 2015). The bacteria then are able to escape from the phagosome to the less hostile cytosol (van der Wel, 2007), where TLRs sense bacterial components and induce production of the proinflammatory cytokine interleukin-1 β (IL-1 β). Mtb can modulate macrophage activity through the cyclic dinucleotide (CDN) sensor STING to produce more anti-inflammatory type I interferons (IFNs, e.g. IFN- α , IFN- β), which inhibit antibacterial pathways and antagonize the activity of IL-1 β . cGAS, the synthase of cyclic guanosine monophosphate-adenosine monophosphate (cGAMP), acts as a DNA receptor and promotes the production of type I

IFNs (Wassermann, 2015; Watson, 2015). These anti-inflammatory reactions allow the bacteria to actively divide and spread through the bloodstream (Pieters, 2013) until the onset of adaptive immunity. T cells are recruited to the site of infection, produce interferon- γ (IFN- γ) and build up a granuloma to contain the infection (Russell, 2010). Formation of the dense cellular granuloma, also called “tubercle”, is the hallmark of tuberculosis infection. In the centre of the granuloma are the Mtb-infected macrophages, surrounded by lipid-rich foamy macrophages. These again are enclosed by T cells and sealed towards the outside by fibrous material (see Figure 1 for an illustration of tuberculosis transmission and pathology). Healthy individuals can control Mtb for many years in such granulomas without showing any symptoms. Immunocompromised patients are unable to control and contain Mtb inside a tubercle. Necrosis then starts from

inside the granuloma and continues to form cavities in lung tissue until Mtb can escape, regain active growth and disseminate, finally being released into the airways to infect a new host (Gengenbacher, 2012; Fitzgerald, 2014; Subbian, 2015).

A peculiarity of Mtb is its ability to establish a latent infection where bacilli rest in a dormant phase, characterized by a non-replicating state with low metabolic rate. In latent tuberculosis, no clinical symptoms are visible and diagnosis is therefore difficult. During this latent phase, so-called persister cells are phenotypically drug tolerant. Although these cells did not undergo any mutations that can lead to true resistance, they are metabolically inactive at a level where most drugs lose their effect (Gomez, 2004; Keren, 2011). To regain growth when the host's immune system begins to weaken, some Mtb cells sporadically resuscitate and re-establish metabolic and replicative activity. If conditions are disadvantageous, these active cells die without affecting the rest of the Mtb community and the majority of the pathogens remain dormant. If the environment is favourable on the other hand, the active cells produce activation signals, consequently the other bacteria revive to establish active tuberculosis (Hett, 2008; Epstein, 2009).

Complete eradication of dormant Mtb cells from the human body is difficult because of their phenotypic drug resistance. Dormancy models are necessary to assess the activity of a drug against non-replicating Mtb, but they are scarce and do not appropriately reflect the physiological *in vivo* environment. To induce dormancy in Mtb, models based on nutrient deprivation or under oxygen depleted conditions were introduced, which showed up-regulation of key dormancy genes like the DosR regulon (Gengenbacher, 2012). In the static culture model, Mtb is grown for 100 days without agitation until self-induced nutrient and oxygen depletion occurs. Several other dormancy models exist based on a similar approach of either nutrient or oxygen depletion. Most dormancy models only allow investigation of Mtb *in vitro*, take a long time to be established and manipulations are limited to the respective environment. A different approach is used by the non-replicating model of strain *M. tuberculosis* 18b. Strain 18b is

dependent on streptomycin for growth, as it has a mutation in the 16S rRNA that needs streptomycin to be bound to the ribosome for functionality. In the absence of streptomycin, the bacteria reach a non-replicating state after ten days which resembles Mtb residing within macrophages (Benjak, 2016). The 18b model also showed that isoniazid and BTZ043 displayed negligible activity on the non-replicating strain, but rifampicin and PA-824 sterilized an 18b culture *in vitro* (Sala, 2010). Isoniazid and BTZ043 target cell wall synthesis, a process that is not susceptible to drugs in non-growing bacteria. Streptomycin dependent growth control was successfully tested in animal models, where the bacteria stopped replicating when streptomycin was withdrawn (Kashino, 2008). Mouse experiments with a latent 18b infection showed that rifampicin, bedaquiline, delamanid and pyrazinamide were active *in vivo*, while cell wall inhibitors had no effect (Zhang, 2012).

Diagnostics, Vaccines, Treatment and Drug Discovery

Rapid and accurate diagnosis of tuberculosis is difficult, especially in low-income countries with limited means. Although many different methods are available, they have limitations, as most are time-consuming, expensive or not precise. The most widely used sputum-smear microscopy test identifies Mtb by acid-fast staining of the bacilli in the sputum of patients. Technicians have to be highly trained to correctly handle and identify Mtb in sputum samples. Still, not all tuberculosis patients are sputum positive, smear-negative tuberculosis patients show clinical symptoms, but no bacteria are visible in their sputum (Siddiqi, 2003).

The gold standard in tuberculosis diagnosis is inoculation of sputum to culture and verify Mtb infection. Sputum cultures even permit the analysis of drug susceptibility, but it heavily relies on technically demanding incubation steps. The biggest disadvantage though is the very long incubation time of 4 – 6 weeks until a result is available (Gillissen, 2016).

The Mantoux-tuberculin test dates back to Koch's trial of developing a treatment against Mtb, and uses purified protein derivatives (PPD)

from Mtb. PPD is injected intradermally to provoke an immune reaction in case the patient encountered Mtb. Although the tuberculin test can even detect a latent Mtb infection, it may cross-react with non-tuberculous bacteria or with the Bacille Calmette-Guérin (BCG) vaccine in healthy individuals (Siddiqi, 2003; Laal, 2005).

About a decade ago, another test capable of detecting latent tuberculosis was developed. The IFN- γ release assay (IGRA) is based on antigen specific T cells in the blood that release the pro-inflammatory cytokine IFN- γ upon contact with Mtb antigens. When stimulated *in vitro* with immunodominant antigens like EsxA and EsxB, primed T cells subsequently release IFN- γ that can then be measured (Whitworth, 2013).

With the emergence of molecular diagnostics, nucleic acid-based technologies to identify Mtb have been developed over 25 years ago. PCR is a sensitive and fast tool, but can be affected by inhibitors present in clinical specimens (Marchetti, 1998). The standard recommended by WHO since 2010 is the PCR-based Xpert Mtuberculosis/RIF (developed by Cepheid, USA) test for detection of drug-susceptible and drug-resistant strains of Mtb (WHO, 2016). It isolates genomic DNA from the sputum and subsequently performs PCR to detect Mtb and genes with rifampicin-resistance mutations. Nonetheless, even the most advanced Xpert Mtuberculosis/RIF test only detects active tuberculosis.

Vaccination to prevent tuberculosis was tried shortly after the discovery of the causative agent of the disease. So far, a potent vaccine that protects completely from Mtb has not been found, despite much effort and several decades of research. In 1920, Albert Calmette and Camille Guérin produced the first vaccine strain, called Bacille Calmette-Guérin (BCG). They sub-cultured the tuberculosis causing strain from cattle, *M. bovis*, over 200 times for several years, until the resulting bacteria lost their virulence in guinea pigs, and could be used as a live attenuated vaccine. BCG has low side effects and was administered several million times worldwide. Unfortunately its efficiency shows high variability, partially due to the wide variety of BCG strains used (McShane, 2012). Some studies reported up to 80% protection while others claimed BCG had a very low

protective rate, or only protected against childhood tuberculosis (McShane, 2012; Tran, 2014). Based on the partial protection conferred by BCG, engineering improved live attenuated vaccine candidates might be successful (Gengenbacher, 2017). Similar approaches, but based on dead bacterial cells included killed *M. vaccae* (de Bruyn, 2003) and fragmented Mtb, showing good efficacy in treatment of latent tuberculosis in animal models by stimulating the immune system (Cardona, 2006). A promising candidate vaccine is MTBVAC, an Mtb *phoP* mutant strain that induces higher IFN- γ expression than BCG in vaccinated groups and protects better against tuberculosis in guinea pigs (Martin, 2006). Recently, MTBVAC entered the first clinical trial showing good safety and no adverse effects (Spertini, 2015). Instead of using whole cells or cell parts, a subunit vaccine includes certain components of the cell that are required to trigger an immune response. Often the potent antigens Ag85 and EsxA are used to establish adaptive immunity against Mtb (Hawn, 2014), but so far a subunit vaccine was not more successful than any of the other attempts.

While the treatment of tuberculosis patients in sanatoria until the 1950s was not very successful, a real breakthrough in defeating tuberculosis came with the advent of antibiotics. Penicillin, the first antibiotic discovered by Alexander Fleming in 1928, saved many lives and allowed to cure diseases that were previously a certain death sentence. Surprisingly, when used against tuberculosis, penicillin had no effect. Many years later it was discovered that Mtb possesses a beta-lactamase conferring resistance to any beta-lactam antibiotic, including penicillin (Iland, 1949). The first drug effective against Mtb was the aminoglycoside antibiotic streptomycin, discovered in 1943 by Waksman and Schulz.

Today, the WHO recommends directly observed treatment short course (DOTS), where treatment adherence and completion are closely monitored (WHO, 2016). Active tuberculosis is thus treated with a combination of four drugs that are administered for at least 6 months, and 9 months in high-risk groups such as immuno-compromised patients. During the first two months, isoniazid, rifampicin,

Group A Fluoroquinolones	<ul style="list-style-type: none"> – Moxifloxacin – Levofloxacin – Gatifloxacin 	Group C Other core second-line drugs	<ul style="list-style-type: none"> – Ethionamide / prothionamide – Cycloserine – Terizodone – Clofazimine
Group B Second-line injectable agents	<ul style="list-style-type: none"> – Capreomycin – Kanamycin – Amikacin – (Streptomycin) 	Group D Add-on agents	D1 <ul style="list-style-type: none"> – Pyrazinamide – Ethambutol D2 <ul style="list-style-type: none"> – High-dose isoniazid – Bedaquiline – Delamanid D3 <ul style="list-style-type: none"> – <i>p</i>-Aminosalicylic acid – Imipenem-cilastatin – Meropenem – Amoxicillin-clavulanate
First-line oral drugs Treatment of drug-sensitive tuberculosis	<ul style="list-style-type: none"> – Ethambutol – Pyrazinamide – Isoniazid – Rifampicin 	New agents	<ul style="list-style-type: none"> – Q203 – PBTZ169 – PA-824

Table 1: Current drugs used to treat tuberculosis, divided into WHO groups A – D for drug-resistant tuberculosis, first-line oral drugs and the discussed new agents.

ethambutol and pyrazinamide are administered, followed by a four-month continuation phase with isoniazid and rifampicin only. While drug-sensitive tuberculosis has a high probability of cure, drug-resistance is a considerable problem in tuberculosis treatment. Acquired drug-resistance was observed soon after the onset of streptomycin treatment, and the same happened with isoniazid monotherapy (Marais, 2016). The most common resistance in *Mtb* is to streptomycin, and multidrug-resistant tuberculosis (MDR-TB) displays resistance to isoniazid and rifampicin. Therefore, a combination of drugs, usually at least four together, is necessary to prevent *Mtb* from developing resistance. Still, resistance mechanisms are accumulating in *Mtb* and the list of effective reserve drugs is short.

Table 1 lists anti-tuberculosis drugs, divided in first- and second-line drugs and groups A – D (World Health Organization, 2016). WHO guidelines recommend to treat rifampicin-resistant or MDR-TB with at least five medicines during the intensive phase, including pyrazinamide, one drug from group A and B each and at least two drugs from group C. If these drugs are not available, they should be replaced with group D medicines to bring the total to five (World Health Organization, 2016). Group A contains fluoroquinolones as core second-line drugs with good bactericidal and

sterilizing activity and a good safety profile to treat MDR-TB. In general, second-line drugs have a lower efficacy, are more difficult to administer and are more expensive than first-line drugs. Group B includes second-line injectable drugs with bactericidal but no sterilizing activity with a worse safety profile than fluoroquinolones. Group C contains drugs under validation that could be future Group B drugs. And finally group D covers agents for which the safety profile or effectiveness in combination has not been confirmed yet, or that have a lower activity (Tiberi, 2017). Extensively drug-resistant-tuberculosis (XDR-TB) is additionally resistant to second-line drugs including a fluoroquinolone and one of the injectables (Cole, 2016) and is treated for at least 18 months and up to 24 months. This extensively long treatment including third-line drugs (e.g. clofazimine, linezolid, clarithromycin) is not only costly but also associated with high rates of toxic effects (Lawn, 2011).

The long drug-discovery process, the difficult search for new targets and the cost intensive clinical trials impede progress in finding new active compounds, and the global drug pipeline for new anti-tuberculosis medicines remains thin. Nevertheless, the combined effort of many research teams recently brought some anti-tuberculosis drugs into clinical trials. The promising compound

Q203 targets the cytochrome bc₁ complex in the respiratory chain of Mtb and is currently in phase I clinical trials (Working Group for new TB Drugs). Bedaquiline (a diarylquinoline, formerly TMC207) targets the F₁F₀-ATP synthase, responsible for establishing the proton motive force. In medical use since only about three years, clinical resistance unfortunately has already emerged, because bedaquiline is administered together with weaker third- and second-line drugs (Kalia, 2017). The benzothiazinone lead compound BTZ043 inhibits the DprE1 enzyme involved in cell wall synthesis. Further development of BTZ043 to PBTZ169 improved its potency, safety and efficacy. PBTZ has now entered clinical trials (Makarov, 2014) and is promoted by Innovative Medicines for Tuberculosis based in Lausanne, Switzerland (iM4TB). The nitroimidazoles PA-824 and delamanid are highly active against Mtb, and delamanid was recently approved by several health authorities for use in patients (Zumla, 2014).

MICROBIOLOGY OF *M. TUBERCULOSIS*

Mycobacteria, belonging to the phylum of the Gram-positive Actinobacteria, are a diverse group including more than 150 species that have a high GC content in their DNA and a lipid-rich cell envelope in common. Although most are harmless saprophytes, some mycobacteria developed into major pathogens. Tuberculosis in humans is mainly caused by Mtb or by the very closely related *Mycobacterium africanum*. Together with *M. canettii* (showing a rare smooth colony phenotype), *M. bovis* (a cattle pathogen, but likewise capable of causing tuberculosis in humans), *M. caprae* (infecting sheep and goats), *M. pinnipedii* (seals and sea lions) and *M. microti* (voles), *M. africanum* and Mtb are collectively referred to as the *Mycobacterium tuberculosis* complex (MTBC) (Rodriguez-Campos, 2014).

Alongside the MTBC, also *Mycobacterium ulcerans*, the causative agent of the severe skin lesions called Buruli ulcers, or the infamous leprosy causing *Mycobacterium leprae* are part of this genus (Hopewell, 2005). Most

pathogenic mycobacteria are slow growers, in the case of Mtb the division time is 18 – 24 hours (Gengenbacher, 2012). In contrast, most environmental mycobacteria are fast growers, like the non-pathogenic *M. smegmatis*, which is often used as a model organism for Mtb.

Mtb is a non-motile, straight or slightly curved rod-shaped bacterium with a cell length of 2 – 4 µm and a thickness of 0.2 – 0.5 µm (see Figure 2 for electron micrographs of Mtb). Colonies have a rough, wrinkled appearance of whitish colour (Todar, 2017). Mtb is predominantly aerobic but also grows as a facultative anaerobe, prefers 37°C and neutral pH and is surrounded by an almost impermeable capsule. Due to their hydrophobic cell envelope, mycobacteria form clumps in liquid culture. The tubercle bacillus does not form spores to survive for a long time in unfavourable conditions like other Gram-positive bacteria, but it can enter a dormant state, with reduced metabolism and very slow to no growth (Gengenbacher, 2012).

Cell Structure

Mtb is phylogenetically a Gram-positive bacterium, and its thick peptidoglycan layer and absence of a classical outer membrane reflect the general structure of this group of bacteria. However, Mtb not only fails to retain the Gram-stain thus mimicking Gram-negative bacteria, but its cell envelope also has unusual properties like porins in the outer lipid layer, a periplasm-like space normally only attributed to Gram-negative bacteria, and an uncommon, thick capsule (Hett, 2008).

The plasma membrane found in mycobacteria does not differ from the plasma membranes found in other bacteria and consists of phospholipids with either straight, unsaturated or mono-methyl branched fatty acids (Daffé, 2008). A potential periplasmic space is hypothesized to be located between the plasma membrane and the adjacent mycolic acid cell wall layer (Beveridge, 2008).

The inner compartment of the cell wall consists of peptidoglycan, covalently linked to arabinogalactan chains, which are attached to long-chain mycolic acids. These three macromolecules are collectively referred to as

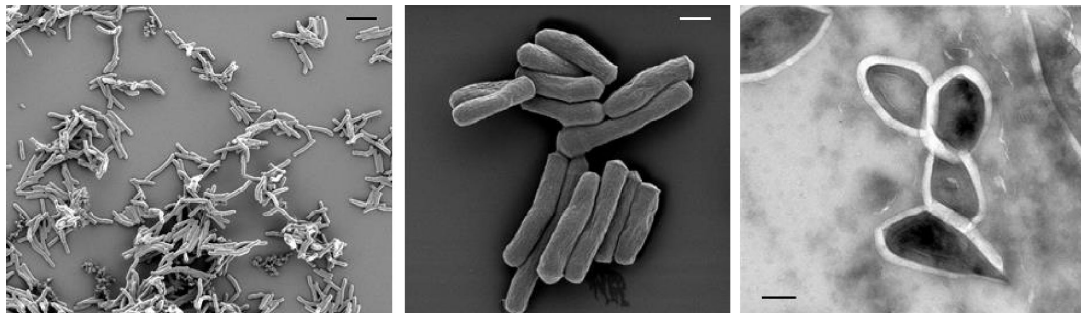


Figure 2: *Mycobacterium tuberculosis*. Surface scanning electron micrographs show *M. tuberculosis* strain H37Rv wildtype in the first two pictures, scale bars represent 3 μm and 500 nm, respectively. The third picture shows a cross section of *M. tuberculosis* H37Rv wildtype inside a macrophage. The white area is the electron-transparent capsule surrounding every bacterium. Scale bar represents 200 nm. (BIO-EM facility at EPFL, N. Odermatt)

the mycolyl-arabinogalactan-peptidoglycan (mAGP) complex, the target of several drugs (e.g. ethambutol, isoniazid, benzothiazinone). Mycolic acids can further be attached to trehalose. While the inner layer is insoluble and necessary for bacterial growth, the outer cell wall layer, sometimes referred to as the outer membrane of Mtb, consists of lipids and proteins that are intercalated into the mycolic acids. These soluble components interact with the immune system and contribute to the virulence of Mtb (Hett, 2008). Mycobacterial glycolipids are recognized by the macrophage and activate the secretion of pro-inflammatory cytokines, while other cell envelope components like the lipoarabinomannan inhibit phagosome maturation (Britton, 2008).

The outermost layer is the capsule, which was first reported as the “electron-transparent zone” to describe the space between the phagosomal membrane and the enclosed Mtb in transmission electron micrographs (Daffé, 2008) (see Fig. 2). The capsule consists mainly of glucan and mannan and also contains several proteins, but only a minor fraction of lipids (Draper, 2005). Overall, this “thick and waxy coat” of Mtb renders it resistant to many drugs and to degradation by host enzymes, as the thickly packed and impermeable cell envelope limits access of external agents.

Secretion Systems

Mtb possesses several different export machineries to secrete metabolites and proteins. The ATP-binding cassette (ABC) transporters are represented by approximately

40 complete and incomplete systems in Mtb. These complexes transport a large variety of small molecules as importers or exporters across biological membranes and can act as efflux pumps in *M. smegmatis* to export fluoroquinolones (Content, 2008). Proteins with typical signal sequences, like the T cell antigen Ag85 are secreted by the general secretion pathway (Sec), which transports unfolded proteins only. Interestingly, Mtb has a second copy, the accessory Sec pathway, which is important for virulence (Braunstein, 2003). The TAT (twin-arginine translocation) system can export folded proteins, among them the beta-lactamases BlaS and BlaC of Mtb (Pieters, 2013).

While the aforementioned transporters are comparable to the ones found in other bacteria, the ESX secretion systems are restricted to a small subset of Gram-positive bacteria and some of them are essential for virulence of Mtb (Simeone, 2009; Mortimer, 2017). One of the main virulence factors is the heterodimer EsxA / EsxB, originally named 6-kDa early-secreted antigenic target ESAT-6 and 10-kDa culture filtrate protein, CFP-10, respectively. These small antigens require a special secretion system, termed ESAT-6 system (ESX) or more recently renamed as the

type VII secretion system (T7SS). ESX systems export folded substrates across the mycobacterial cell envelope. In Mtb,

chromosomal loci encoding five ESX-systems (ESX-1 – ESX-5) are found, and they usually include *pe* / *ppe* genes, a pair of genes encoding

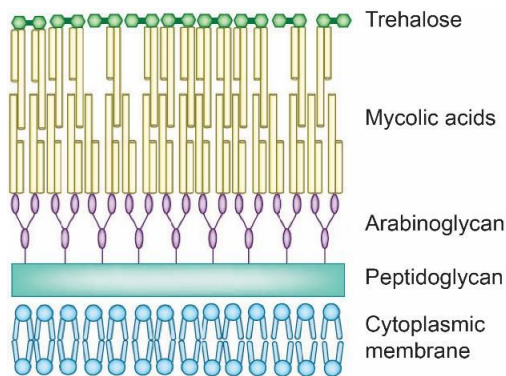


Figure 3: Schematic showing a simplified mycobacterial cell envelope (modified from (Cole, 2012)).

secreted Esx proteins and the ESX machinery proteins. The ESX-conserved components, encoded by *ecc* genes, form the core structure of the secretion system in the inner membrane and together with MycP represent the ESX machinery. MycP1 is a serine protease that cleaves EspB during export (Feltcher, 2010). Ecc proteins from the ESX-5 system (EccB₅, EccC₅, EccD₅ and EccE₅) were shown to assemble in equimolar stoichiometry with six-fold symmetry into a membrane-associated complex with a five-nanometer pore. This dimension is sufficient to secrete folded T7SS effectors (Beckham, 2017). Esp proteins are ESX secretion-associated proteins, most have homologues in all five ESX-secretion systems, others are specific to ESX-1, like EspL and EspB (Gröschel, 2016).

The T7SS was initially encoded on a plasmid (Newton-Foot, 2016), then transferred to the chromosome. Selective forces drove duplication and divergence of the ESX systems, and some mycobacterial species subsequently lost one or more copies of the ESX-secretion system, e.g. ESX-2 does not exist in *M. ulcerans* and *M. marinum* (Mortimer, 2017).

ESX-1 is important for virulence, as it is responsible for secretion of EsxA / EsxB. Deletion of the ESX-1 system results in strong attenuation of Mtb, and loss of some of the ESX-1 genes gave rise to the region of difference 1 (ΔRD1) that distinguishes the vaccine strain *M. bovis* BCG from its virulent parental strain. The ESX-1 secretion system has been the target of extensive studies since its discovery, and substantial progress was made in the structural

role of its components, especially the membrane components of the system. The heterodimer EspA / EspC is necessary for co-secretion of EsxA / EsxB (Fortune, 2005; Xu, 2007), and it was recently shown that EspC forms polymers once secreted by Mtb. It is possible that EspC is part of an outer membrane channel or acts as a needle protein itself, allowing secretion of EsxA / EsxB (Lou, 2016).

The ESX-3 system is essential in Mtb and involved in iron acquisition (Siegrist, 2009), while ESX-5 is associated with PE / PPE protein export (Abdallah, 2009). Less is known about ESX-2 and ESX-4, which are predicted to be not essential for *in vitro* growth or virulence (Feltcher, 2010; Gröschel, 2016).

Genomics and Genetics

The cornerstone of modern molecular analysis of Mtb genetics was set with the publication of the complete genome sequence of the H37Rv strain in 1998 by Stewart Cole and co-workers (Cole, 1998). The circular chromosome consists of 4,411,532 bp with a GC content of 65.6% and 4007 coding sequences, including 50 stable RNAs (Cole, 1998; Camus, 2002). The start point for numbering was chosen at the origin of replication at the *dnaA* gene with *rv0001*. The genes were clustered into 11 different functional categories, broadly describing their function (Lew, 2011).

With about 10% of the coding sequence, a large part of the genes fall into the family of glycine-rich PE / PPE proteins, only present in slow-growing pathogenic mycobacteria. While the C-terminal part is divergent, their conserved N-terminus includes a PE (Pro-Glu) or PPE (Pro-Pro-Glu) motif, after which the proteins were named. The *pe* / *ppe* genes are distributed all over the genome, often co-transcribed as a *pe-ppe* pair and the proteins are found as heterodimers (Brennan, 2017). Most of these proteins are cell surface associated or secreted (Målen, 2007; Daleke, 2011), and show immunogenic activity (Sampson, 2011).

While 13 sigma factors and more than 100 regulatory proteins have been detected, indicating tight gene regulation, only 15 sensor histidine kinases and response regulators were identified. An additional 11 eukaryotic-like serine / threonine protein kinases probably

compensate for the paucity in environmental signal transduction pathways (Cole, 1998; Boitel, 2003).

The GC content is evenly distributed throughout the entire Mtb genome, and therefore horizontally acquired genes are either rare or difficult to detect (Cole, 1998). The non-virulent *M. smegmatis* shows eukaryotic-like conjugal DNA transfer, and a low transfer rate was detected as well in *M. canettii*. In contrast, no other member of the MTBC showed horizontal gene transfer (HGT) in an extensive investigation by Boritsch et al. (Boritsch, 2016). The presence of HGT in *M. canettii* and its close relatedness to Mtb suggest that HGT was once an important mechanism for evolution in Mtb (Mortimer, 2014). Naturally occurring plasmids have not yet been described for Mtb, although they are present in other mycobacterial species (Ummels, 2014). Several insertion sequences, short transposable elements without any accessory genes, were detected in the Mtb genome, as well as two prophage-like elements (Cole, 1998). These prophages ϕ Rv1 and ϕ Rv2 have variable positions across members of the MTBC, and ϕ Rv1 was shown to be fully competent to integrate and excise (Cole, 2005).

Physiology and Metabolism

The genome sequence revealed that Mtb is equipped with diverse genes for metabolic flexibility and pathways to synthesize all macro- and micronutrients (including essential amino acids, vitamins and enzyme co-factors). Further, a variety of enzymes to metabolize carbon sources like carbohydrates, hydrocarbons, alcohols, ketones and carboxylic acids are present, as well as over 250 enzymes involved in fatty acid metabolism (Cole, 1998). Mtb can survive on a wide array of organic substrates. Glycerol is usually added to Mtb cultures *in vitro*, as it supports maximal growth rates and yields. *In vivo*, on the other hand, Mtb's preferred substrates for carbon and energy metabolism are lipids such as fatty acids and cholesterol (Baughn, 2014). During the early stage of infection in aerobic conditions, bacteria access glucose and triacylglycerides as primary carbon sources. After formation of granulomas and in an increasingly anaerobic environment,

Mtb is forced to shift to the utilization of lipids (Warner, 2014). Interestingly, Mtb is able to compartmentalize discrete metabolic processes and therefore use simultaneously multiple carbon sources to enhance its growth (De Carvalho, 2010).

Apart from carbon and nitrogen sources, metals are essential for bacterial survival. As an intracellular pathogen, Mtb faces iron limitation and therefore had to develop high-affinity iron acquisition mechanisms. To import iron into the bacterial cell, Mtb produces siderophores, iron-chelators in the form of mycobactin or the soluble carboxymycobactin, or directly utilizes heme as an iron source (Rodriguez, 2014). Zinc and copper are important co-factors for many enzymes and their level has to be carefully controlled to avoid depletion or a toxic effect. Zur (Rv2359) for example regulates zinc uptake and export (Rodriguez, 2014).

GENE REGULATION

Microbes, like all living organisms, are exposed to an environment that may change within a short time. As an intracellular pathogen that spreads from host to host, Mtb has to adapt to immunological, environmental and nutritional changes, and the repertoire of transcriptional regulators present in tubercle bacilli allows them to respond to external stimuli and ensures their success. When entering the host, Mtb encounters several different milieux, starting with contact with the aerated region of the respiratory system, continuing in harsh conditions inside the phagosome and ending in an anoxic and low nutrient environment in the granulomas, before it can escape, enter the blood stream and replicate (Torrelles, 2017). Bacteria either modify gene expression at the transcriptional level or induce posttranslational protein modifications to change their activity. Transcription is controlled by several different mechanisms in a complex interplay. DNA condensation, making a gene reachable or inaccessible to the RNA polymerase, influences global gene expression. The RNA polymerase itself is directed to the -35 region of the promoters of genes by sigma factors that can be exchanged. Promoter-centred regulation is

based on transcription factors that specifically bind and activate or repress a gene (Browning, 2016).

Transcription Regulatory Proteins

The RNA polymerase of Mtb, like in other bacteria, is composed of a core enzyme consisting of two α subunits, one β and one β' subunit, and ω . Together with the σ factor, which recognizes and binds to specific promoter sequences and directs the transcription of a certain gene set, the holo-enzyme is formed and ready for transcription. Mtb, subject to a very variable environment, expresses 13 different sigma factors, in contrast to *Escherichia coli*, which only has seven. Every σ factor confers specificity to transcription initiation, but due to the high GC content in Mtb, it is often difficult to find a consensus sequence for the -10 and -35 regions upstream of the transcription start site (Smith, 2005; Newton-Foot, 2013).

σ^A , the essential housekeeping sigma factor encoded by the *sigA* gene *rv2703*, recognizes promoters similar to the canonical *E. coli* one and is relatively stably expressed in exponential growth *in vitro*, in macrophages as well as under certain stress conditions (Manganelli, 1999). Expression profiles different from the σ^A mediated one can be induced by alternative σ factors with different consensus sequences, anti- σ factors that inhibit binding of the σ factor to the RNA polymerase or to the promoter sequence, and also anti-anti- σ factors that inhibit binding of anti- σ factors to their targets. The primary-like σ^B , sharing homology with σ^A , is induced under various stress conditions like poor nutrient availability, low aeration or sodium dodecyl sulfate (SDS) exposure (Lee, 2008). σ^F , also classified as a stress response sigma factor, is induced after cold-shock or in stationary phase and upregulates a diverse set of genes, including virulence factors or several antisense transcripts (Hartkoorn, 2012).

The other sigma factors (σ^C , σ^D , σ^E , σ^G , σ^H , σ^I , σ^J , σ^K , σ^L , σ^M), are extracytoplasmic sigma factors (ECF) involved in cell envelope synthesis, secretion, protein repair or degradation functions and have a very broad range of action (Lee, 2008). While σ^C is only expressed in exponential growth, σ^E for example is induced

under stress conditions like elevated temperature, exposure to detergents or oxidative stress (Manganelli, 2014).

In addition to sigma factors, two-component regulatory systems (TCS) often influence a wide array of genes. They integrate environmental and intracellular stimuli to bring about changes in very specific or global gene expression and are composed of a sensor histidine kinase and a response regulator. The histidine kinase senses a signal and transmits it by phosphorylation to a cognate response regulator acting as a transcription factor. Mtb has relatively few TCS compared to other bacteria (Parish, 2014). One of the best-studied TCS is DosRST (Rv3133, Rv3132, Rv2027), the regulator of the dormancy regulon that controls the shift to hypoxia-induced *in vitro* dormancy. The DosR regulon includes 48 genes, which enable the pathogen to persist under hypoxic conditions (Chen, 2012). Another TCS pair is the histidine kinase PhoR (Rv0758) and its response regulator PhoP (Rv0757). Its regulon is induced upon Mg^{2+} starvation. PhoP negatively regulates its own gene and is involved in virulence regulation, as PhoP mutants in Mtb are severely attenuated (Smith, 2005; Gonzalo-Asensio, 2008). The TCS MprAB (Rv0981, Rv0982) is involved in virulence regulation by direct binding of MprA to the *espA* promoter and is required for ESX-1 function (Pang, 2013). MprA shows enhanced binding to its target genes after phosphorylation by MprB and regulates the stress-induced σ^E , which is involved in the stringent-response (Sureka, 2007).

Apart from TCS, other regulators directly sense a substrate and bind to the DNA. TetR family transcription regulators often act as repressors and are present in a wide range of bacterial families. The prototype of this family is the regulator of the expression of a tetracycline efflux pump in Gram-negative bacteria, where TetR specifically binds to the *tetO* operators and blocks access to the *tet* promoter by RNAP. In Mtb, over 50 TetR-type regulators are encoded in the genome (Balhana, 2015). For example, EthR regulates expression of *ethA*, a monooxygenase that activates the antibiotic ethionamide. KstR controls the expression of a number of genes involved in lipid metabolism,

including the *mce4* operon coding for a cholesterol importer (Kendall, 2010).

CRP, the cyclic AMP (cAMP)-receptor protein acts as a global transcription factor in *E. coli* by binding the second messenger cAMP in response to nutrient starvation (Grainger, 2005). In Mtb, CRP (Rv3676) regulates, among others, *rpfA* encoding a resuscitation promoting factor, the transcriptional regulator *whiB1* and the fumarate reductase *frdA*, playing an important role under starvation conditions. Furthermore, CRP binds upstream of *espACD* and activates its transcription (Rickman, 2005). CRP has been recently proposed to behave as a global regulator, as it contacts more than 200 loci distributed all over the Mtb chromosome and is often associated with the transcription start sites (Kahramanoglou, 2014).

Metal ions are important signals and contribute to gene expression regulation. For example, IdeR, the iron dependent repressor, binds to divalent metal cations including iron, zinc or manganese and activates iron-storage genes (Newton-Foot, 2013). IdeR also regulates *hupB* expression, and *ideR*-deficient Mtb exhibits unrestricted iron uptake and lacks iron storage (Pandey, 2014). Another important regulatory mechanism is the control of the SOS response upon DNA damage. After induction of the SOS response, DNA repair and mutagenesis are induced while the cell is resting. The transcription factor LexA represses its target genes by binding to the SOS box in the promoter region. When DNA damage occurs, RecA stimulates the cleavage of LexA, which can no longer bind to DNA and therefore activates the SOS-response (Newton-Foot, 2013).

Global Transcription Factors

Apart from the abovementioned rather specific transcription factors and response regulators that influence a relatively small subset of genes, global transcription factors have a major impact on the expression of a large number of genes. One class of global regulators are the nucleoid associated proteins (NAPs), small and low-molecular-weight, alkaline proteins, that often form homodimers. NAPs can have up to several hundred binding sites on a bacterial chromosome and are able to bend, bridge or loop the DNA. These

conformational changes in the DNA topology also lead to supercoiling and DNA condensation in bacteria. In addition to playing a purely architectural role, NAPs often act as transcription factors by directly silencing or activating genes (Dillon, 2010). In *E. coli*, topologically isolated regions of the chromosome were identified as macrodomains of about 1 Mbp size, which are further segmented into microdomains with an average size of 10 kbp and it is believed that NAPs are responsible for this organization thanks to their ability to bridge DNA (Song, 2015). Formation of loops in the DNA bring loci, which are distal on the primary sequence of the chromosome, into close spatial proximity. mRNA and protein products are then close in space upon transcription and translation (Badrinarayanan, 2015), which facilitates interaction between them, like a regulatory gene and its target operons (Dorman, 2013).

Twelve NAPs have been identified so far in *E. coli* (Dillon, 2010) but, contrary to the high number of sigma factors, NAPs are underrepresented in Mtb. Only four (HU, mlHF, Lsr2, EspR) have a confirmed function as global transcriptional regulators today.

HupB

HU, the Histone-like U93 protein has two subunits in *E. coli*, HU α and HU β . HU exists in homo- or heterodimeric forms, depending on the growth stage of the bacterium. HU does bind DNA in a sequence unspecific manner, but prefers distorted regions like Holliday junctions, mismatches or nicks. It can introduce negative supercoils and condense the DNA by interacting with topoisomerase I or by multimerization (Rouvière-Yaniv, 1979; Kamashev, 2008; Dillon, 2010; Hammel, 2016).

It was shown that HU controls 8% of the entire *E. coli* transcriptome and is involved in the SOS response and in the cellular response during aerobic and acid stress (Oberto, 2009).

The *hupB* gene (*rv2986*) codes for the Mtb homologue of *E. coli* HU and its N-terminal amino acid sequence is similar to *E. coli* HupB. HupB is the only conserved NAP in Mtb with sequence similarity to NAPs outside the domain of the Actinobacteria. Contrary to the two

subunits forming the HU heterodimer in *E. coli*, Mtb only has one copy of the *hupB* gene.

In *M. smegmatis* and *M. bovis*, HupB is named MDP1 for mycobacterial DNA-binding protein 1. The level of MDP1 is higher in *M. bovis* in stationary cultures and in dormant *M. smegmatis*. Interestingly, downregulation of MDP1 in *M. bovis* BCG led to faster growth in macrophage-like cells and a higher cell mass in broth cultures (Lewin, 2008). The protein is also required for adaptation to acidic pH and induction of cytokine secretion in macrophages (Kunisch, 2012).

Mtb HupB binds preferentially to GC-rich sequences and negatively regulates gene expression. Of interest was the control of *katG* expression (Niki, 2012), as the catalase-peroxidase KatG activates the prodrug isoniazid and downregulation of *katG* leads to phenotypic resistance to this drug (Niki, 2012). It was also shown that HupB is regulated by iron levels, with a strong increase in protein levels upon iron limiting conditions (Yeruva, 2006). Later on it was demonstrated that it positively regulates the *mbt* operon, responsible for siderophore production (Pandey, 2014). HupB regulates a diverse set of genes beside those required for siderophore expression. For instance, it is involved in cell wall synthesis or adhesion (Pandey, 2014). Unusual for a NAP, HupB was shown to be associated with the mycobacterial cell wall (Pandey, 2014) and to trigger the host's immune response (Prabhakar, 1998; Sivakolundu, 2013).

The *hupB* gene was predicted to be essential by transposon mutagenesis in *Mtb* (Griffin, 2011) and this was confirmed by Bhowmick et al. who constructed a conditional knockdown mutant and validated the essentiality of *hupB* (Bhowmick, 2014). However, recently Pandey et al. were successful in removing the *hupB* gene from Mtb. Their mutant strain showed no growth defect *in vitro* but *ex vivo* infection and multiplication were compromised (Pandey, 2014).

H-NS

The histone-like nucleoid-structuring protein, H-NS, represses horizontally acquired genes in many Gram-negative bacteria like *E. coli* or *Salmonella enterica* (Ali, 2013). This

xenogeneic silencing is thought to protect the bacteria from expression of potentially harmful foreign genes, acquired for example by plasmid uptake or viral infection. The dimeric H-NS binds non-specifically to AT-rich, curved DNA at hundreds of loci in the bacterial genome. It represses the pathogenicity islands of *Salmonella*, while another NAP, Fis (factor for inversion stimulation), activates these genes in an overall antagonistic effect on H-NS (Dorman, 2004, 2009). H-NS does not exclusively target horizontally acquired genetic elements, it also controls several other genes and represses all of the rRNA operons in *E. coli*.

In Mtb, the Rv3852 protein was identified as H-NS due to its N-terminal similarity to human histone H1 (Cole, 1998), but lacks any sequence similarity to *E. coli* H-NS. Werlang et al. showed that purified Rv3852 indeed did bind DNA with a slight preference for curved DNA. The protein formed a dimer in solution and most probably higher order oligomers upon DNA binding. Interestingly, the Mtb *rv3852* gene did not complement the *hns* mutant phenotype in *E. coli*, suggesting that these two proteins have a different function (Werlang, 2009). Expression of *rv3852* in *M. smegmatis*, whose genome lacks a homologue of *rv3852*, led to reduced biofilm formation, membrane invagination and altered colony morphology as well as a change in the lipid profile. Further, the C-terminal domain proved to be a membrane-anchor (Ghosh, 2013), a very atypical trait for a nucleoid-associated protein. It was also predicted by transposon mutagenesis that *rv3852* was not essential for *in vitro* growth (Griffin, 2011).

EspR

EspR (Rv3849), the ESX-1 secreted protein regulator, was first described as an essential factor for ESX-1 gene expression (Raghavan, 2008). Surprisingly, it was discovered that EspR not only specifically controls expression of the virulence factors EspA-EspC-EspD, but binds to over 150 loci upon dimerization, including genes in the ESX-2 and ESX-5 systems, cell wall functions, and several *pe/ppe* genes (Blasco, 2012). Despite previous evidence, EspR was shown to be not secreted and only appears in the cytosol (Blasco, 2012). The *espR* gene in

Mtb is not essential, but deletion leads to slow growth (Sassetti, 2003; Griffin, 2011) and reduced virulence in macrophages (Raghavan, 2008). Protein abundance increases from early to late exponential growth phase *in vitro* (Blasco, 2012). PhoP and MprA directly regulate EspR by binding to its promoter region (Cao, 2015).

EspR carries a helix-turn-helix DNA binding motif at its N-terminus and a dimerization domain at the C-terminus. Dimerization of the protein is necessary for DNA binding, but not sufficient, as structurally important amino acids have to be conserved for full function. EspR introduces DNA loops that are consistent in size with the distance between the binding sites throughout the chromosome (Blasco, 2014).

mIHF

Bacterial integration host factor (IHF) has many binding sites on the *E. coli* chromosome, bends the DNA sharply and often interacts with other transcription factors (Browning, 2010). IHF belongs to the HU family of DNA binding proteins due to its amino acid sequence and structural similarity, and also consists of two subunits encoded by *ihfA* and *ihfB* in *E. coli* (Oberto, 1994). It was originally identified as a host factor required for bacteriophage λ integration and, unlike HU, sequence-specifically binds to DNA. IHF regulates about 100 genes in *E. coli* and is responsible for survival during starvation (Johnson, 2005).

Comparable to phage λ , the temperate bacteriophage L5 integrates at *attB* (bacterial attachment site) into the genomes of Mtb, *M. smegmatis*, and *M. bovis* (Pedulla, 1996). Similar to the discovery of IHF in *E. coli*, also mIHF (mycobacterial integration host factor, Rv1388) was determined to be essential for L5 integration in the *M. smegmatis* chromosome (Lee, 1993). Cell extracts from *E. coli* were not able to stimulate L5 integration, indicating that mIHF and the HU-family proteins from *E. coli* not only lack any sequence or structural similarity, but they are fundamentally different in function and cannot be exchanged. mIHF has sequence homologues only among Actinobacteria, and although IHF and HU belong to the same class of DNA-binding

proteins in Gram-negative bacteria, mIHF does not resemble HupB in Mtb.

mIHF was identified among the top ten most abundant proteins in Mtb in exponential phase in a quantitative mass spectrometry experiment, and its abundance did not vary in hypoxia-induced dormancy, nor in re-aeration several days later (Schubert, 2015). By contrast, in *M. smegmatis* and in *M. bovis* BCG, mIHF reaches the highest level before entering the stationary phase with relatively low levels at the beginning of the exponential phase. Construction of a knockout mutant in *M. smegmatis* was not successful, which points to the essentiality of mIHF in this bacterium (Pedulla, 1998). Transposon mutagenesis suggested that mIHF is essential in Mtb too (Sassetti, 2003).

Lsr2

The iron-regulated H-NS-like protein Lsr2 (Rv3597) is restricted to Actinobacteria and although lacking sequence or structural similarity to *E. coli* H-NS, it is considered to be the only H-NS-like protein in Gram-positive bacteria (Gordon, 2008). Unlike Rv3852, which was wrongly proposed as the H-NS equivalent of Mtb, Lsr2 can complement *E. coli hns* knockout phenotypes, and vice versa *E. coli hns* can complement an *lsr2* knockout mutant in *M. smegmatis* (Gordon, 2008), showing the functional equivalence of these two proteins. In *M. smegmatis*, *lsr2* knockouts have a smooth colony morphology and altered lipid composition (Chen, 2006), while in Mtb the gene seems to be essential (Sassetti, 2003; Colangeli, 2009). Purified Lsr2, like H-NS in *E. coli*, bridges DNA, binds preferentially to AT-rich sequences and forms oligomers (Chen, 2008). The Lsr2-DNA complexes are strong, resistant to pH and salt changes and lead to stiffening and folding of the DNA (Qu, 2013).

In a ChIP-chip experiment, Lsr2 co-precipitated with 840 genes, among them were those for virulence factors like the *espACD* operon and over half of the total PE and PPE genes. Overall, Lsr2 has a negative effect on gene expression, similar to *E. coli* H-NS (Gordon, 2010). The expression of *lsr2* is upregulated at high temperatures and upon nutrient starvation (Betts, 2002; Stewart, 2002).

Other NAPs

It has been suggested that GroEL1 (Rv3417) shows NAP-like characteristics. GroEL1 is similar in sequence to GroEL2, an essential chaperone that prevents misfolding of proteins. Surprisingly, it was shown that GroEL1 recognizes DNA in a sequence-unspecific manner and causes DNA condensation (Basu, 2009). The DNA binding ability of GroEL1 remains doubtful, because the crystal structure of GroEL1 is remarkably similar to GroEL2 of Mtb and to GroEL of *E. coli*, and it was further shown that GroEL1 recognizes specific peptide sequences (Sielaff, 2011). NapM in *M. smegmatis*, annotated as PadR, was suggested to be a NAP by Liu et al. They showed that NapM co-localized with the bacterial nucleoid, introduced DNA-bridges and regulated a set of 156 genes in *M. smegmatis* (Liu, 2016). Rv0047 is the Mtb homologue of NapM, but so far no evidence for Rv0047 to act as a NAP exists. Rv0430 was identified as a member of the Fis (factor for inversion stimulation) protein family in an amino acid sequence homology search (Lew, 2011), but was never confirmed nor disproved to act as a NAP.

Gene regulation is indisputably the key step in adaptation to environmental factors. The immense complexity of the interplay between different response regulators, specific as well as global transcription factors and DNA topology make it a difficult task to draw a complete map of all regulatory cascades. Every addition of knowledge about the regulon of a transcription factor adds value to this goal. The relatively low number of NAPs in Mtb, their proven impact on specific virulence genes together with their overall influence on transcription regulation make them an exciting topic for research and the main subject of investigation in the present thesis.

THESIS CONTEXT AND RATIONALE

The peculiar lifestyle of Mtb to remain dormant during decades inside macrophages of the human host is coupled with phenotypic drug resistance. Subsequent resuscitation to active growth causing tuberculosis implies how sophisticated and adapted gene expression

mechanisms in this pathogen have to be for successful virulence. Studying gene regulation, especially in the context of activation of virulence genes, helps to understand under which conditions they are activated or repressed. Traditional transcription factors control only a single gene or at most a dozen, but do not influence chromosome topology. Nucleoid associated proteins on the other hand have a major impact on total gene expression. Many regulatory pathways are not fully understood or completely unclear, and improving our knowledge of gene regulation is a key objective if we want to understand how the virulence genes are activated and when the metabolic pathways inside the host are active.

Recent progress was made in understanding transcription in Mtb by re-assigning the transcriptional start sites (Shell, 2015) or investigating mycobacterial promoters (Newton-Foot, 2013). Three of Mtb's NAPs were already defined and characterized as global transcription factors; EspR (Rosenberg, 2011; Blasco, 2012), Lsr2 (Bartek, 2014) and HupB (Pandey, 2014). Missing were H-NS and mIHF of Mtb, for which neither a defined target regulon nor binding sites were available.

Research was driven by the increasingly affordable methods of ChIP- and RNA-sequencing, which allowed assigning chromosome wide transcription factor binding profiles and high-resolution gene expression data. The multiplexing of several libraries reduced costs and opened the way to faster analysis of transcription factors. During my PhD work, I used genetic tools like knockout mutant construction or conditional knockdown of the gene of interest directly in the virulent Mtb strain H37Rv, which allowed for in-depth characterization of the impact on the bacterial cell. Integrated with ChIP-seq and RNA-seq, which defines its binding profile and target regulon, it was possible to characterize the proteins from multiple angles.

Chapter 2 describes how deletion of the *rv3852* gene, originally thought to code for the NAP H-NS in Mtb, was scrutinized *in vitro*, *ex vivo* and *in vivo*, to finally disprove that Rv3852 acts as a NAP. Construction of an unmarked in-frame deletion of the *rv3852* gene from the H37Rv chromosome gave us a tool to thoroughly characterize the role of Rv3852 in

the resulting $\Delta rv3852$ strain. Surprisingly, we were not able to associate a phenotype in any *in vitro* experiment to the deletion of the *rv3852* gene. Growth dynamics, nucleoid position and compaction, drug susceptibility and sensitivity to pH were not affected in $\Delta rv3852$. We further found that $\Delta rv3852$ does not have a defined regulon, in particular no induction of a broad transcriptional change was observed upon deletion of the gene. Further, the $\Delta rv3852$ strain was not reduced in virulence in an *in vivo* mouse infection model. We concluded that $\Delta rv3852$ does not belong into the group of NAPs. The results were published in the Journal of Bacteriology (Odermatt, 2017).

In **chapter 3**, the function of mIHF is in focus. Unlike *rv3852*, mIHF proved essential, as deletion of the *mihF* gene from the bacterial chromosome was not possible. We therefore relied on the TetR / PIP-OFF system to downregulate expression of *mihF* by addition of the non-toxic tetracycline analogue anhydrotetracycline (ATc). The initial annotation of *mihF* not only included the coding sequence, but also an upstream region almost as long as the gene itself, exacerbating the construction of a functional conditional knockdown (cKD) mutant. **Chapter 3.1** describes how we determined the new translational start site of *mihF*, and subsequently reports the results obtained from investigation of the *mihF*-cKD mutant strain. We found that depletion of the mIHF protein led to elongated cells devoid of septa. mIHF depletion eventually led to cell death, corroborating the essentiality of mIHF. We therefore analysed the total nucleic acid and protein synthesis in mIHF depleted cells, and found they synthesize less macromolecules. Contrary to EspR, we did not find a consensus motif for mIHF binding sites, which were distributed around the chromosome at 150 broad binding sites. Several differentially regulated genes upon mIHF depletion were involved in virulence and the response to host-interaction. We concluded that mIHF has a dual role as transcriptional activator specifically for virulence related pathways, as well as a housekeeping function involved in protein and nucleic acid synthesis. Results were summarized in the manuscript “Essentiality of *mihF* and impact of gene silencing on Mtb physiology and transcriptional

landscape” that will be submitted for publication soon.

Apart from the function of mIHF in the bacterial cell, we also investigated its structure and DNA-binding mechanism. Most NAPs bind DNA at one protein domain, then dimerize and therefore introduce bridges or loops into the DNA. It was suggested that sIHF, a close relative of mIHF from *Streptomyces coelicolor*, is interacting with the chromosome in monomeric form (Swiercz, 2013). In **Chapter 3.2** we used an integrated approach to describe the binding mode of mIHF by atomic force microscopy (AFM). As mIHF has a small size of 12 kDa, was highly soluble in water and stable at room temperature, we assessed its structure by nuclear magnetic resonance (NMR) spectroscopy. Loop formation in AFM studies showed that mIHF preferentially introduces left-handed loops, decondensing and opening the supercoiled chromosome. mIHF is a globular protein consisting of one long, protruding alpha helix and a core of five short helices. Suggested by experts in the field, whom I met on a conference, we tested if mIHF was phosphorylated by the serine/threonine-protein kinases PknB and PknG. Both were able to phosphorylate the recombinant mIHF extracted from *E. coli* on different residues, suggesting a role of post-translational modification for mIHF activity. Last experiments, especially to investigate the role of phosphorylation on DNA binding, and further analysis have to be carried out before this work can be submitted for publication.

The difficulty of treating tuberculosis due to increasing drug resistance of Mtb leads to an ongoing effort to find new targets and active compounds. Chrysomycin was found to specifically bind to the NAP Lsr2 and inhibit DNA binding thereof. **Chapter 4** contains additional work done on NAPs and in particular summarizes how we found that chrysomycin non-specifically intercalates into DNA and therefore was not a suitable candidate for further drug-development. These findings will not be published in a paper but nonetheless were important results I created in my thesis.

Finally, in **chapter 5**, I critically review the work done during my PhD thesis and suggest further experiments to address unsolved questions

about the NAPs in Mtb, followed by a short summary of the obtained results and integration of these into known mechanisms of transcriptional regulation in Mtb.

REFERENCES

- Abdallah, A.M., Verboom, T., Weerdenburg, E.M., Gey van Pittius, N.C., Mahasha, P.W., Jiménez, C., et al. (2009) PPE and PE_PGRS proteins of *Mycobacterium marinum* are transported via the type VII secretion system ESX-5. *Mol. Microbiol.* **73**: 329–40.
- Ali, S.S., Whitney, J.C., Stevenson, J., Robinson, H., Howell, P.L., and Navarre, W.W. (2013) Structural insights into the regulation of foreign genes in salmonella by the Hha/H-NS complex. *J. Biol. Chem.* **288**: 13356–13369.
- Badrinarayanan, A., Le, T.B.K., and Laub, M.T. (2015) Bacterial Chromosome Organization and Segregation. *Annu. Rev. Cell Dev. Biol.* **31**: 171–199.
- Balhana, R.J.C., Singla, A., Sikder, M.H., Withers, M., and Kendall, S.L. (2015) Global analyses of TetR family transcriptional regulators in mycobacteria indicates conservation across species and diversity in regulated functions. *BMC Genomics* **16**: 479.
- Bartek, I.L., Woolhiser, L.K., Baughn, A.D., Basaraba, R.J., Jacobs, W.R., Lenaerts, A.J., and Voskuil, M.I. (2014) Mycobacterium tuberculosis Lsr2 Is a Global Transcriptional Regulator. *MBio* **5**: e01106-14.
- Basu, D., Khare, G., Singh, S., Tyagi, A., Khosla, S., and Mande, S.C. (2009) A novel nucleoid-associated protein of *Mycobacterium tuberculosis* is a sequence homolog of GroEL. *Nucleic Acids Res.* **37**: 4944–54.
- Baughn, A.D. and Rhee, K.Y. (2014) Metabolomics of Central Carbon Metabolism in *Mycobacterium tuberculosis*. In, *Molecular Genetics of Mycobacteria.*, pp. 323–339.
- Beckham, K.S.H., Ciccarelli, L., Bunduc, C.M., Mertens, H.D.T., Ummels, R., Lugmayr, W., et al. (2017) Structure of the mycobacterial ESX-5 type VII secretion system membrane complex by single-particle analysis. *Nat. Microbiol.* **2**: 17047.
- Benjak, A., Uplekar, S., Zhang, M., Piton, J., Cole, S.T., and Sala, C. (2016) Genomic and transcriptomic analysis of the streptomycin-dependent *Mycobacterium tuberculosis* strain 18b. *BMC Genomics* **17**: 1–14.
- Betts, J.C., Lukey, P.T., Robb, L.C., McAdam, R.A., and Duncan, K. (2002) Evaluation of a nutrient starvation model of *Mycobacterium tuberculosis* persistence by gene and protein expression profiling. *Mol. Microbiol.* **43**: 717–731.
- Beveridge, T.J. (2008) Wall Ultrastructure and Periplasm. In, *The Mycobacterial Cell Envelope2.*, pp. 13–23.
- Bhowmick, Bhowmick, T., Ghosh, S., Dixit, K., Ganesan, V., Ramagopal, U. a, et al. (2014) Targeting *Mycobacterium tuberculosis* nucleoid-associated protein HU with structure-based inhibitors. *Nat. Commun.* **5**: 4124.
- Blasco, B., Chen, J.M., Hartkoorn, R., Sala, C., Uplekar, S., Rougemont, J., et al. (2012) Virulence regulator EspR of *Mycobacterium tuberculosis* is a nucleoid-associated protein. *PLoS Pathog.* **8**: e1002621.
- Blasco, B., Japaridze, A., Stenta, M., Wicky, B.I.M., Dietler, G., Dal Peraro, M., et al. (2014) Functional dissection of intersubunit interactions in the EspR virulence regulator of *Mycobacterium tuberculosis*. *J. Bacteriol.* **196**: 1889–900.
- Boitel, B., Ortiz-Lombardía, M., Durán, R., Pompeo, F., Cole, S.T., Cerveñansky, C., and Alzari, P.M. (2003) PknB kinase activity is regulated by phosphorylation in two Thr residues and dephosphorylation by PstP, the cognate phospho-Ser/Thr phosphatase, in *Mycobacterium tuberculosis*. *Mol. Microbiol.* **49**: 1493–1508.
- Boritsch, E.C., Khanna, V., Pawlik, A., Honoré, N., Navas, V.H., Ma, L., et al. (2016) Key experimental evidence of chromosomal DNA transfer among selected tuberculosis-causing mycobacteria. *Proc. Natl. Acad. Sci.* 201604921.
- Braunstein, M., Espinosa, B.J., Chan, J., Belisle, J.T., and Jacobs, W.R. (2003) SecA2 functions in the secretion of superoxide dismutase A and in the virulence of *Mycobacterium tuberculosis*. *Mol. Microbiol.* **48**: 453–464.
- Brennan, M.J. (2017) The Enigmatic PE / PPE Multigene Family of Mycobacteria and Tuberculosis Vaccination. *Infect. Immun.* **85**: 1–8.
- Britton, W.J. and Triccas, J.A. (2008) The Constituents of the Cell Envelope and Their Impact on the Host Immune System. In, *The Mycobacterial Cell Envelope2.*, pp. 249–270.
- Browning, D.F. and Busby, S.J.W. (2016) Local and global regulation of transcription initiation in bacteria. *Nat. Rev. Microbiol.* **14**: 638–650.
- Browning, D.F., Grainger, D.C., and Busby, S.J. (2010) Effects of nucleoid-associated proteins on bacterial chromosome structure and gene expression. *Curr. Opin. Microbiol.* **13**: 773–80.
- de Bruyn, G. and Garner, P. (2003) *Mycobacterium vaccae* immunotherapy for treating tuberculosis. *Cochrane Database Syst Rev*

- CD001166.
- Cambau, E. and Drancourt, M. (2014) Steps towards the discovery of *Mycobacterium tuberculosis* by Robert Koch, 1882. *Clin. Microbiol. Infect.* **20**: 196–201.
- Camus, J.-C., Pryor, M.J., Médigue, C., and Cole, S.T. (2002) Re-annotation of the genome sequence of *Mycobacterium tuberculosis* H37Rv. *Microbiology* **148**: 2967–2973.
- Cao, G., Howard, S.T., Zhang, P., Wang, X., Chen, X.L., Samten, B., and Pang, X. (2015) EspR, a regulator of the ESX-1 secretion system in *Mycobacterium tuberculosis*, is directly regulated by the two-component systems MprAB and PhoPR. *Microbiology* **161**: 477–489.
- Cardona, P.J. (2006) RUT1: A new chance to shorten the treatment of latent tuberculosis infection. *Tuberculosis* **86**: 273–289.
- De Carvalho, L.P.S., Fischer, S.M., Marrero, J., Nathan, C., Ehrh, S., and Rhee, K.Y. (2010) Metabolomics of *Mycobacterium tuberculosis* reveals compartmentalized co-catabolism of carbon substrates. *Chem. Biol.* **17**: 1122–1131.
- Chen, J.M., German, G.J., Alexander, D.C., Ren, H., Tan, T., and Liu, J. (2006) Roles of Lsr2 in Colony Morphology and Biofilm Formation of *Mycobacterium smegmatis*. *J. Bacteriol.* **188**: 633–641.
- Chen, J.M., Ren, H., Shaw, J.E., Wang, Y.J., Li, M., Leung, A.S., et al. (2008) Lsr2 of *Mycobacterium tuberculosis* is a DNA-bridging protein. *Nucleic Acids Res.* **36**: 2123–35.
- Chen, T., He, L., Deng, W., and Xie, J. (2012) The *Mycobacterium* DosR regulon structure and diversity revealed by comparative genomic analysis. *J. Cell. Biochem.* **114**: 1–6.
- Colangeli, R., Haq, a, Arcus, V.L., Summers, E., Magliozzo, R.S., McBride, a, et al. (2009) The multifunctional histone-like protein Lsr2 protects mycobacteria against reactive oxygen intermediates. *Proc. Natl. Acad. Sci. U. S. A.* **106**: 4414–8.
- Cole, S.T. (2012) Infectious diseases: Transporter targeted in tuberculosis. *Nat. Chem. Biol.* **8**: 326–327.
- Cole, S.T. (2016) Inhibiting *Mycobacterium tuberculosis* within and without. *Phil. Trans. R. Soc. B* **371**.
- Cole, S.T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., et al. (1998) Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* **396**: 651–653.
- Cole, S.T., Davis Eisenach, K., McMurray, D.N., and Jacobs, W.R. (2005) Mycobacteriophages and Tuberculosis. In, *Tuberculosis and the Tubercle Bacillus*, pp. 203–218.
- Content, J. and Peirs, P. (2008) The ABC Transporter Systems. In, *The Mycobacterial Cell Envelope*, pp. 195–199.
- Daffé (2008) The Global Architecture of the Mycobacterial Cell Envelope. In, *The Mycobacterial Cell Envelope*, pp. 3–11.
- Daleke, M.H., Cascioferro, A., de Punder, K., Ummels, R., Abdallah, A.M., Van Der Wel, N., et al. (2011) Conserved Pro-Glu (PE) and Pro-Pro-Glu (PPE) protein domains target LipY lipases of pathogenic mycobacteria to the cell surface via the ESX-5 pathway. *J. Biol. Chem.* **286**: 19024–34.
- Daniel, V. and Daniel, T. (1999) Old Testament biblical references to tuberculosis. *Clin Infect Dis* **29**: 1557–1558.
- Dillon, S.C. and Dorman, C.J. (2010) Bacterial nucleoid-associated proteins, nucleoid structure and gene expression. *Nat. Rev. Microbiol.* **8**: 185–95.
- Dorman, C.J. (2013) Co-operative roles for DNA supercoiling and nucleoid-associated proteins in the regulation of bacterial transcription. *Biochem. Soc. Trans.* **41**: 542–7.
- Dorman, C.J. (2004) H-NS: a universal regulator for a dynamic genome. *Nat. Rev. Microbiol.* **2**: 391–400.
- Dorman, C.J. and Kane, K. a (2009) DNA bridging and antibridging: a role for bacterial nucleoid-associated proteins in regulating the expression of laterally acquired genes. *FEMS Microbiol. Rev.* **33**: 587–92.
- Draper, P. and Daffé, M. (2005) The Cell Envelope of *Mycobacterium tuberculosis* with Special Reference to the Capsule and Outer Permeability Barrier. In, *Tuberculosis and the Tubercle Bacillus*, pp. 261–273.
- Epstein, S.S. (2009) Microbial awakenings. *Nature* **457**: 1083–1083.
- Eum, S.Y., Kong, J.H., Hong, M.S., Lee, Y.J., Kim, J.H., Hwang, S.H., et al. (2010) Neutrophils are the predominant infected phagocytic cells in the airways of patients with active pulmonary TB. *Chest* **137**: 122–128.
- Farer, L., Lowell, A., and Meador, M. (1979) Extrapulmonary tuberculosis in the United States. *Am J Epidemiol* **109**: 205–217.
- Feltcher, M.E., Sullivan, J.T., and Braunstein, M. (2010) Protein export systems of *Mycobacterium tuberculosis*: novel targets for drug development? *NIH Public Access* 1581–1597.
- Fitzgerald, L.E., Abendaño, N., Juste, R.A., and Alonso-Hearn, M. (2014) Three-dimensional in vitro models of granuloma to study

- bacteria-host interactions, drug-susceptibility, and resuscitation of dormant mycobacteria. *Biomed Res. Int.* **2014**:
- Fortune, S.M., Jaeger, A., Sarracino, D.A., Chase, M.R., Sassetti, C.M., Sherman, D.R., et al. (2005) Mutually dependent secretion of proteins required for mycobacterial virulence. *Proc. Natl. Acad. Sci. U. S. A.* **102**: 10676–81.
- Gagneux, S. (2012) Host-pathogen coevolution in human tuberculosis. *Philos. Trans. R. Soc. B Biol. Sci.* **367**: 850–859.
- Gengenbacher, M. and Kaufmann, S.H.E. (2012) *Mycobacterium tuberculosis*: success through dormancy. *FEMS Microbiol. Rev.* **36**: 514–32.
- Gengenbacher, M., Nieuwenhuizen, N.E., and Kaufmann, S.H.E. (2017) BCG — old workhorse, new skills. *Curr. Opin. Immunol.* **47**: 8–16.
- Ghosh, S., Indi, S.S., and Nagaraja, V. (2013) Regulation of lipid biosynthesis, sliding motility and biofilm formation by a membrane-anchored nucleoid associated protein of *Mycobacterium tuberculosis*. *J. Bacteriol.* **195**: 1769–78.
- Gillissen, A. (2016) Noch immer brandgefährlich – die Tuberkulose. *Fortschr. Med.* **6**: 50–57.
- Gomez, J.E. and McKinney, J.D. (2004) *M. tuberculosis* persistence, latency, and drug tolerance. *Tuberculosis* **84**: 29–44.
- Gonzalo-Asensio, J., Soto, C.Y., Arbués, A., Sancho, J., Menéndez, M.D.C., García, M.J., et al. (2008) The *Mycobacterium tuberculosis* *phoPR* operon is positively autoregulated in the virulent strain H37Rv. *J. Bacteriol.* **190**: 7068–7078.
- Gordon, B.R.G., Imperial, R., Wang, L., Navarre, W.W., and Liu, J. (2008) Lsr2 of *Mycobacterium* represents a novel class of H-NS-like proteins. *J. Bacteriol.* **190**: 7052–9.
- Gordon, B.R.G., Li, Y., Wang, L., Sintsova, A., van Bakel, H., Tian, S., et al. (2010) Lsr2 is a nucleoid-associated protein that targets AT-rich sequences and virulence genes in *Mycobacterium tuberculosis*. *Proc. Natl. Acad. Sci. U. S. A.* **107**: 5154–9.
- Grainger, D.C., Hurd, D., Harrison, M., Holdstock, J., and Busby, S.J.W. (2005) Studies of the distribution of *Escherichia coli* cAMP-receptor protein and RNA polymerase along the *E. coli* chromosome. *Proc. Natl. Acad. Sci. U. S. A.* **102**: 17693–8.
- Griffin, J.E., Gawronski, J.D., Dejesus, M. a, Ioerger, T.R., Akerley, B.J., and Sassetti, C.M. (2011) High-resolution phenotypic profiling defines genes essential for mycobacterial growth and cholesterol catabolism. *PLoS Pathog.* **7**: e1002251.
- Gröschel, M.I., Sayes, F., Simeone, R., Majlessi, L., and Brosch, R. (2016) ESX secretion systems: mycobacterial evolution to counter host immunity. *Nat. Rev. Microbiol.* **14**: 677–691.
- Hammel, M., Amlanjyoti, D., Reyes, F.E., Chen, J.-H., Parpana, R., Tang, H.Y.H., et al. (2016) HU multimerization shift controls nucleoid compaction. *Sci. Adv.* **2**: e1600650.
- Hartkoorn, R.C., Sala, C., Uplekar, S., Busso, P., Rougemont, J., and Cole, S.T. (2012) Genome-wide definition of the SigF regulon in *Mycobacterium tuberculosis*. *J. Bacteriol.* **194**: 2001–9.
- Hawn, T.R., Day, T.A., Scriba, T.J., Hatherill, M., Hanekom, W.A., Evans, T.G., et al. (2014) Tuberculosis vaccines and prevention of infection. *Microbiol. Mol. Biol. Rev.* **78**: 650–71.
- Hershberg, R., Lipatov, M., Small, P.M., Sheffer, H., Niemann, S., Homolka, S., et al. (2008) High functional diversity in *Mycobacterium tuberculosis* is driven by genetic drift and human demography. *PLoS Biol.* **6**: 2658–2671.
- Hett, E.C. and Rubin, E.J. (2008) Bacterial Growth and Cell Division: a *Mycobacterium* Perspective. *Microbiol. Mol. Biol. Rev.* **72**: 126–156.
- Hopewell, P.C. and Jasmer, R.M. (2005) Overview of clinical tuberculosis. In, *Tuberculosis and the Tubercle Bacillus*, pp. 15–31.
- Iland, C.N. and Baines, S. (1949) The Effect of Penicillin on the Tubercle Bacillus: Tubercle Penicillinase. *J. Pathol.* **61**: 329–335.
- iM4TB *Innov. Med. Tuberc.* www.im4tb.org.
- Johansson, J., Karlsson, A., Bylund, J., and Welin, A. (2015) Phagocyte interactions with *Mycobacterium tuberculosis* - Simultaneous analysis of phagocytosis, phagosome maturation and intracellular replication by imaging flow cytometry. *J. Immunol. Methods* **427**: 73–84.
- Johnson, R.C., Johnson, L.M., Schmidt, J.W., and Gardner, J.F. (2005) Major nucleoid proteins in the structure and function of the *Escherichia coli* chromosome. *Bact. Chromosom.* 65–132.
- Kahramanoglou, C., Cortes, T., Matange, N., Hunt, D.M., Visweswariah, S.S., Young, D.B., and Buxton, R.S. (2014) Genomic mapping of cAMP receptor protein (CRPMT) in *Mycobacterium tuberculosis*: Relation to transcriptional start sites and the role of CRPMT as a transcription factor. *Nucleic Acids Res.* **42**: 8320–8329.
- Kalia, N.P., Hasenoehrl, E.J., Ab Rahman, N.B., Koh, V.H., Ang, M.L.T., Sajorda, D.R., et al. (2017) Exploiting the synthetic lethality between terminal respiratory oxidases to kill

- Mycobacterium tuberculosis* and clear host infection. *Proc. Natl. Acad. Sci.* **114**: 201706139.
- Kamashev, D., Balandina, A., Mazur, A.K., Arimondo, P.B., and Rouviere-Yaniv, J. (2008) HU binds and folds single-stranded DNA. *Nucleic Acids Res.* **36**: 1026–1036.
- Kappelman, J., Cihat Alcicek, M., Kazanci, N., Schultz, M., Özkul, M., and Sen, S. (2008) Brief Communication: First *Homo erectus* from Turkey and Implications for Migrations into Temperate Eurasia. *Am. J. Phys. Anthr.* **135**: 110–116.
- Kashino, S.S., Napolitano, D.R., Skobe, Z., and Campos-Neto, A. (2008) Guinea pig model of *Mycobacterium tuberculosis* latent/dormant infection. *Microbes Infect.* **10**: 1469–1476.
- Kaufmann, S.H. (2001) How can immunology contribute to the control of tuberculosis? *Nat. Rev. Immunol.* **1**: 20–30.
- Kendall, S.L., Burgess, P., Balhana, R., Withers, M., Ten Bokum, A., Lott, J.S., et al. (2010) Cholesterol utilization in mycobacteria is controlled by two TetR-type transcriptional regulators: kstR and kstR2. *Microbiology* **156**: 1362–1371.
- Keren, I., Minami, S., Rubin, E., and Lewis, K.C.-3119538 (2011) Characterization and transcriptome analysis of *Mycobacterium tuberculosis* persists. *MBio* **2**: e00100-11.
- Koch, R. (1882) Die Aetiologie der Tuberculose (Nach einem in der physiologischen Gesellschaft zu Berlin am 24. März gehaltenem Vortrage). *Berliner klin Wochenschr* **19**: 221–30.
- Kunisch, R., Kamal, E., and Lewin, A. (2012) The role of the mycobacterial DNA-binding protein 1 (MDP1) from *Mycobacterium bovis* BCG in host cell interaction. *BMC Microbiol.* **12**: 165.
- Laal, S. and Skeiky, Y.A.W. (2005) Diagnostics: Immune-Based Methods. In, *Tuberculosis and the tubercle bacillus.*, pp. 71–83.
- Lawn, S.D. and Zumla, A.I. (2011) Tuberculosis. *Lancet* **378**: 57–72.
- Lee, J.-H., Karakousis, P.C., and Bishai, W.R. (2008) Roles of SigB and SigF in the *Mycobacterium tuberculosis* sigma factor network. *J. Bacteriol.* **190**: 699–707.
- Lee, M.H. and Hatfull, G.F. (1993) Mycobacteriophage L5 Integrase-Mediated Site-Specific Integration In Vitro. *J. Bacteriol.* **175**: 6836–6841.
- Lemaitre, B., Nicolas, E., Michaut, L., Reichhart, J.M., and Hoffmann, J.A. (1996) The dorsoventral regulatory gene cassette spatzle/Toll/Cactus controls the potent antifungal response in *Drosophila* adults. *Cell* **86**: 973–983.
- Lew, J.M., Kapopoulou, A., Jones, L.M., and Cole, S.T. (2011) TubercuList - 10 years after. *Tuberculosis* **91**: 1–7.
- Lewin, A., Baus, D., Kamal, E., Bon, F., Kunisch, R., Maurischat, S., et al. (2008) The mycobacterial DNA-binding protein 1 (MDP1) from *Mycobacterium bovis* BCG influences various growth characteristics. *BMC Microbiol.* **8**: 91.
- Liu, Y., Wang, H., Cui, T., Zhou, X., Jia, Y., Zhang, H., et al. (2016) NapM, a new nucleoid-associated protein, broadly regulates gene expression and affects mycobacterial resistance to anti-tuberculosis drugs. *Mol. Microbiol.* **101**: 167–181.
- Lou, Y., Rybníček, J., Sala, C., and Cole, S.T. (2016) EspC forms a filamentous structure in the cell envelope of *Mycobacterium tuberculosis* and impacts ESX-1 secretion. *Mol. Microbiol.* **103**: 1–39.
- Makarov, V., Lechartier, B., Zhang, M., Neres, J., van der Sar, A.M., Raadsen, S.A., et al. (2014) Towards a new combination therapy for tuberculosis with next generation benzothiazinones. *EMBO Mol. Med.* **6**: 372–383.
- Målen, H., Berven, F.S., Fladmark, K.E., and Wiker, H.G. (2007) Comprehensive analysis of exported proteins from *Mycobacterium tuberculosis* H37Rv. *Proteomics* **7**: 1702–1718.
- Manganelli, R. (2014) Sigma Factors : Key Molecules in *Mycobacterium tuberculosis* Physiology and Virulence. *Microbiol. Spectr.* 1–23.
- Manganelli, R., Dubnau, E., Tyagi, S., Kramer, F.R., and Smith, I. (1999) Differential expression of 10 sigma factor genes in *Mycobacterium tuberculosis*. *Mol. Microbiol.* **31**: 715–724.
- Marais, B.J. (2016) The global tuberculosis situation and the inexorable rise of drug-resistant disease. *Adv. Drug Deliv. Rev.* 1–7.
- Marchetti, G., Gori, A., Catozzi, L., Vago, L., Nebuloni, M., Rossi, M.C., et al. (1998) Evaluation of PCR in detection of *Mycobacterium tuberculosis* from formalin-fixed, paraffin-embedded tissues: Comparison of four amplification assays. *J. Clin. Microbiol.* **36**: 1512–1517.
- Martin, C., Williams, A., Hernandez-Pando, R., Cardona, P.J., Gormley, E., Bordat, Y., et al. (2006) The live *Mycobacterium tuberculosis* phoP mutant strain is more attenuated than BCG and confers protective immunity against tuberculosis in mice and guinea pigs. *Vaccine* **24**: 3408–3419.
- McShane, H., Jacobs, W.R., Fine, P.E., Reed, S.G.,

- McMurray, D.N., Behr, M., et al. (2012) BCG: Myths, realities, and the need for alternative vaccine strategies. *Tuberculosis* **92**: 283–288.
- Mortimer, T.D. and Pepperell, C.S. (2014) Genomic signatures of Distributive conjugal transfer among mycobacteria. *Genome Biol. Evol.* **6**: 2489–2500.
- Mortimer, T.D., Weber, A.M., and Pepperell, C.S. (2017) Evolutionary thrift: mycobacteria repurpose plasmid diversity during adaptation of type VII secretion systems. *Genome Biol. Evol.* **9**: 398–413.
- Murray, J.F. (2015) The Industrial Revolution and the decline in death rates from tuberculosis. *Int. J. Tuberc. Lung Dis.* **19**: 502–503.
- Newton-Foot, M. and Gey van Pittius, N.C. (2013) The complex architecture of mycobacterial promoters. *Tuberculosis (Edinb.)* **93**: 60–74.
- Newton-Foot, M., Warren, R.M., Sampson, S.L., van Helden, P.D., and Gey van Pittius, N.C. (2016) The plasmid-mediated evolution of the mycobacterial ESX (Type VII) secretion systems. *BMC Evol. Biol.* **16**: 62.
- Niki, M., Niki, M., Tateishi, Y., Ozeki, Y., Kirikae, T., Lewin, A., et al. (2012) A novel mechanism of growth phase-dependent tolerance to isoniazid in mycobacteria. *J. Biol. Chem.* **287**: 27743–27752.
- Oberto, J., Drica, K., and Rouvière-Yaniv, J. (1994) Histones, HMG, HU, IHF: Mème combat. *Biochimie* **76**: 901–908.
- Oberto, J., Nabti, S., Jooste, V., Mignot, H., and Rouvière-Yaniv, J. (2009) The HU regulon is composed of genes responding to anaerobiosis, acid stress, high osmolarity and SOS induction. *PLoS One* **4**: e4367.
- Odermatt, N.T., Sala, C., Benjak, A., Kolly, G.S., Vocat, A., Lupien, A., and Cole, S.T. (2017) Rv3852 (H-NS) of *Mycobacterium tuberculosis* Is Not Involved in Nucleoid Compaction and Virulence Regulation. *J. Bact.* **199**: 1–12.
- Pandey, S.D., Choudhury, M., and Sritharan, M. (2014) Transcriptional regulation of *Mycobacterium tuberculosis* hupB gene expression. *Microbiology* **160**: 1637–1647.
- Pandey, S.D., Choudhury, M., Yousuf, S., Wheeler, P.R., Gordon, S. V., Ranjan, A., and Sritharan, M. (2014) Iron-Regulated Protein HupB of *Mycobacterium tuberculosis* Positively Regulates Siderophore Biosynthesis and Is Essential for Growth in Macrophages. *J. Bacteriol.* **196**: 1853–65.
- Pang, X., Samten, B., Cao, G., Wang, X., Tvinnereim, A.R., Chen, X.L., and Howard, S.T. (2013) MprAB regulates the espA operon in *Mycobacterium tuberculosis* and modulates ESX-1 function and host cytokine response. *J. Bacteriol.* **195**: 66–75.
- Parish, T. (2014) Two-Component Regulatory Systems of Mycobacteria. In, *Molecular Genetics of Mycobacteria*, pp. 209–223.
- Parrish, N.M., Dick, J.D., and Bishai, W.R. (1998) Mechanisms of latency in *Mycobacterium tuberculosis*. *Trends Microbiol.* **6**: 107–112.
- Pedulla, M.L. and Hatfull, G.F. (1998) Characterization of the mHf gene of *Mycobacterium smegmatis*. *J. Bacteriol.* **180**: 5473–7.
- Pedulla, M.L., Lee, M.H., Lever, D.C., and Hatfull, G.F. (1996) A novel host factor for integration of mycobacteriophage L5. *Proc. Natl. Acad. Sci. U. S. A.* **93**: 15411–6.
- Pieters, J. and McKinney, J.D. (2013) Pathogenesis of *Mycobacterium tuberculosis* and its Interaction with the Host Organism.
- Prabhakar, S., Annapurna, P.S., Jain, N.K., Dey, a B., Tyagi, J.S., and Prasad, H.K. (1998) Identification of an immunogenic histone-like protein (HLPMT) of *Mycobacterium tuberculosis*. *Tuber. Lung Dis.* **79**: 43–53.
- Qu, Y., Lim, C.J., Whang, Y.R., Liu, J., and Yan, J. (2013) Mechanism of DNA organization by *Mycobacterium tuberculosis* protein Lsr2. *Nucleic Acids Res.* **41**: 5263–5272.
- Raghavan, S., Manzanillo, P., Chan, K., Dovey, C., and Cox, J.S. (2008) Secreted transcription factor controls *Mycobacterium tuberculosis* virulence. *Nature* **454**: 717–721.
- Rickman, L., Scott, C., Hunt, D.M., Hutchinson, T., Menéndez, M.C., Whalan, R., et al. (2005) A member of the cAMP receptor protein family of transcription regulators in *Mycobacterium tuberculosis* is required for virulence in mice and controls transcription of the rpfA gene coding for a resuscitation promoting factor. *Mol. Microbiol.* **56**: 1274–1286.
- Roberts, C.A., Pfister, L.A., and Mays, S. (2009) Letter to the editor: Was tuberculosis present in *Homo erectus* in Turkey? *Am. J. Phys. Anthropol.* **139**: 442–444.
- Rodriguez-Campos, S., Smith, N.H., Boniotti, M.B., and Aranaz, A. (2014) Overview and phylogeny of *Mycobacterium tuberculosis* complex organisms: Implications for diagnostics and legislation of bovine tuberculosis. *Res. Vet. Sci.* **97**: S5–S19.
- Rodriguez, G.M. and Neyrolles, O. (2014) Metallobiology of Tuberculosis. In, *Molecular Genetics of Mycobacteria2*, pp. 377–378.
- Rosenberg, O.S., Dovey, C., Tempesta, M., Robbins, R.A., Finer-Moore, J.S., Stroud, R.M., and Cox, J.S. (2011) EspR, a key regulator of *Mycobacterium tuberculosis* virulence, adopts a unique dimeric structure among helix-turn-

- helix proteins. *Proc. Natl. Acad. Sci. U. S. A.* **108**: 13450–5.
- Rouvière-Yaniv, J., Yaniv, M., and Germond, J.E. (1979) E. coli DNA binding protein HU forms nucleosome-like structure with circular double-stranded DNA. *Cell* **17**: 265–274.
- Russell, D.G., Cardona, P., Kim, M., and Allain, S. (2010) Foamy macrophages and the progression of the human TB granuloma. *Nat. Immunol.* **10**: 943–948.
- Sala, C., Dhar, N., Hartkoorn, R.C., Zhang, M., Ha, Y.H., Schneider, P., and Cole, S.T. (2010) Simple model for testing drugs against nonreplicating *Mycobacterium tuberculosis*. *Antimicrob. Agents Chemother.* **54**: 4150–8.
- Sampson, S.L. (2011) Mycobacterial PE/PPE proteins at the host-pathogen interface. *Clin. Dev. Immunol.* **2011**: ID 497203.
- Sassetti, C.M., Boyd, D.H., and Rubin, E.J. (2003) Genes required for mycobacterial growth defined by high density mutagenesis. *Mol. Microbiol.* **48**: 77–84.
- Schlesinger, L. (1996) Entry of *Mycobacterium tuberculosis* into mononuclear phagocytes. *Curr Top Microbiol Immunol* **215**: 71–96.
- Schubert, O.T., Ludwig, C., Kogadeeva, M., Zimmermann, M., Rosenberger, G., Gengenbacher, M., et al. (2015) Absolute Proteome Composition and Dynamics during Dormancy and Resuscitation of *Mycobacterium tuberculosis*. *Cell Host Microbe* **18**: 96–108.
- Shell, S.S., Wang, J., Lapierre, P., Mir, M., Chase, M.R., Pyle, M.M., et al. (2015) Leaderless Transcripts and Small Proteins Are Common Features of the Mycobacterial Translational Landscape. *PLoS Genet.* **11**: 1–31.
- Siddiqi, K., Lambert, M.-L., Walley, J., Lauzardo, M., Ashkin, D., Maher, D., et al. (2003) Clinical diagnosis of smear-negative pulmonary tuberculosis in low-income countries: the current evidence. *Lancet. Infect. Dis.* **3**: 288–96.
- Siegrist, M.S., Unnikrishnan, M., McConnell, M.J., Borowsky, M., Cheng, T.-Y., Siddiqi, N., et al. (2009) Mycobacterial Esx-3 is required for mycobactin-mediated iron acquisition. *Proc. Natl. Acad. Sci.* **106**: 18792–18797.
- Sielaff, B., Lee, K.S., and Tsai, F.T.F. (2011) Structural and functional conservation of mycobacterium tuberculosis GroEL paralogs suggests that GroEL1 is a chaperonin. *J. Mol. Biol.* **405**: 831–839.
- Simeone, R., Bottai, D., and Brosch, R. (2009) ESX/type VII secretion systems and their role in host-pathogen interaction. *Curr. Opin. Microbiol.* **12**: 4–10.
- Sivakolundu, S., Mannela, U.D., Jain, S., Srikantam, A., Peri, S., Pandey, S.D., and Sritharan, M. (2013) Serum iron profile and ELISA-based detection of antibodies against the iron-regulated protein HupB of *Mycobacterium tuberculosis* in TB patients and household contacts in Hyderabad (Andhra Pradesh), India. *Trans. R. Soc. Trop. Med. Hyg.* **107**: 43–50.
- Smith, I., Bishai, W.R., and Nagaraja, V. (2005) Control of Mycobacterial Transcription. In, *Tuberculosis and the Tubercle Bacillus.*, pp. 2019–231.
- Song, D. and Loparo, J.J. (2015) Building bridges within the bacterial chromosome. *Trends Genet.* **31**: 164–173.
- Spertini, F., Audran, R., Chakour, R., Karoui, O., Steiner-Monrad, V., Thierry, A.-C., et al. (2015) Safety of human immunisation with a live-attenuated *Mycobacterium tuberculosis* vaccine: a randomised, double-blind, controlled phase I trial. *Lancet. Respir. Med.* **3**: 953–962.
- Stewart, G.R., Wernisch, L., Stabler, R., Mangan, J. a., Hinds, J., Laing, K.G., et al. (2002) Dissection of the heat-shock response in *Mycobacterium tuberculosis* using mutants and microarrays. *Microbiology* **148**: 3129–3138.
- Stiftung Historisches Lexikon der Schweiz and Ritzmann, I. (2001) Historisches Lexikon der Schweiz.
- Subbian, S., Tsenova, L., Kim, M.J., Wainwright, H.C., Visser, A., Bandyopadhyay, N., et al. (2015) Lesion-specific immune response in granulomas of patients with pulmonary tuberculosis: A pilot study. *PLoS One* **10**: 1–21.
- Sureka, K., Dey, S., Datta, P., Singh, A.K., Dasgupta, A., Rodrigue, S., et al. (2007) Polyphosphate kinase is involved in stress-induced mprAB-sigE-rel signalling in mycobacteria. *Mol. Microbiol.* **65**: 261–276.
- Swiercz, J.P., Nanji, T., Gloyd, M., Guarné, A., and Elliot, M. a (2013) A novel nucleoid-associated protein specific to the actinobacteria. *Nucleic Acids Res.* **41**: 4171–84.
- Tiberi, S., Scardigli, A., Centis, R., D’Ambrosio, L., Muñoz-Torrico, M., Zumla, A., et al. (2017) Classifying new anti-tuberculosis drugs: rationale and future perspectives. *Int. J. Infect. Dis.* **56**: 181–184.
- Tiemersma, E.W., van der Werf, M.J., Borgdorff, M.W., Williams, B.G., and Nagelkerke, N.J.D. (2011) Natural history of tuberculosis: Duration and fatality of untreated pulmonary

- tuberculosis in HIV negative patients: A systematic review. *PLoS One* **6**: e17601.
- Todar, K. (2017) Todar's Online Textbook of Bacteriology. www.textbookofbacteriology.net.
- Torrelles, J.B. and Schlesinger, L.S. (2017) Integrating Lung Physiology, Immunology, and Tuberculosis. *Trends Microbiol.* **25**: 688–697.
- Tran, V., Behr, M.A., and Liu, J. (2014) BCG Vaccines. *Mol. Genet. Mycobacteria, Second Ed.* 49–59.
- Ummels, R., Abdallah, A.M., Kuiper, V., Aâjoud, A., Sparrius, M., Naeem, R., et al. (2014) Identification of a Novel Conjugative Plasmid in Mycobacteria That Requires Both Type IV and Type VII Secretion. *MBio* **5**: e01744-14.
- Warner, D.F. (2014) Mycobacterium tuberculosis metabolism. *Cold Spring Harb Persp Biol* 1–24.
- Wassermann, R., Gulen, M.F., Cole, S.T., Ablasser, A., Wassermann, R., Gulen, M.F., et al. (2015) *Mycobacterium tuberculosis* Differentially Activates cGAS- and Inflammasome-Dependent Intracellular Immune Responses through ESX-1. *Cell Host Microbe* **17**: 1–12.
- Watson, R.O., Bell, S.L., MacDuff, D.A., Kimmey, J.M., Diner, E.J., Olivas, J., et al. (2015) The Cytosolic Sensor cGAS Detects *Mycobacterium tuberculosis* DNA to Induce Type I Interferons and Activate Autophagy. *Cell Host Microbe* **17**: 811–819.
- van der Wel, N., Hava, D., Houben, D., Fluitsma, D., van Zon, M., Pierson, J., et al. (2007) *M. tuberculosis* and *M. leprae* Translocate from the Phagolysosome to the Cytosol in Myeloid Cells. *Cell* **129**: 1287–1298.
- Werlang, I.C.R., Schneider, C.Z., Mendonça, J.D., Palma, M.S., Basso, L.A., and Santos, D.S. (2009) Identification of Rv3852 as a nucleoid-associated protein in *Mycobacterium tuberculosis*. *Microbiology* **155**: 2652–63.
- Whitworth, H.S., Scott, M., Connell, D.W., Dongés, B., and Lalvani, A. (2013) IGRAs - The gateway to T cell based TB diagnosis. *Methods* **61**: 52–62.
- WHO (2016) Global Tuberculosis Report 2016. *Cdc* 2016 214.
- WHO (2014) Tuberculosis Fact Sheet. *N°104*.
- Working Group for new TB Drugs <https://www.newtbdrugs.org/pipeline/clinical>
- World Health Organization (2016) WHO treatment guidelines for drug-resistant tuberculosis. *Geneva, Switz.* 61.
- Xu, J., Laine, O., Masciocchi, M., Manoranjan, J., Smith, J., Du, S.J., et al. (2007) A unique Mycobacterium ESX-1 protein co-secretes with CFP-10/ESAT-6 and is necessary for inhibiting phagosome maturation. *Mol. Microbiol.* **66**: 787–800.
- Yeruva, V.C., Duggirala, S., Lakshmi, V., Kolarich, D., Altmann, F., and Sriharan, M. (2006) Identification and characterization of a major cell wall-associated iron-regulated envelope protein (Irep-28) in Mycobacterium tuberculosis. *Clin. Vaccine Immunol.* **13**: 1137–1142.
- Zhang, M., Sala, C., Hartkoorn, R.C., Dhar, N., Mendoza-Losana, A., and Cole, S.T. (2012) Streptomycin-starved Mycobacterium tuberculosis 18b, a drug discovery tool for latent tuberculosis. *Antimicrob. Agents Chemother.* **56**: 5782–5789.
- Zumla, A., Raviglione, M., Hafner, R., and Fordham von Reyn, C. (2013) Tuberculosis. *N. Engl. J. Med.* **368**: 745–755.
- Zumla, A.I., Gillespie, S.H., Hoelscher, M., Philips, P.P.J., Cole, S.T., Abubakar, I., et al. (2014) New antituberculosis drugs, regimens, and adjunct therapies: Needs, advances, and future prospects. *Lancet Infect. Dis.* **14**: 327–340.

CHAPTER 2

Rv3852 (H-NS) of *Mycobacterium tuberculosis* Is Not Involved in Nucleoid Compaction and Virulence Regulation

Nina T. Odermatt, Claudia Sala, Andrej Benjak, Gaëlle S. Kolly, Anthony Vocat,
Andréanne Lupien, Stewart T. Cole

École Polytechnique Fédérale de Lausanne, Global Health Institute, Station 19,
Lausanne, Switzerland

2017. Journal of Bacteriology 199:e00129-17

Contributions: design of experiments, phenotype assessment, light microscopy, transcriptome
analysis, *ex vivo* infection assay, data analysis, manuscript preparation



Rv3852 (H-NS) of *Mycobacterium tuberculosis* Is Not Involved in Nucleoid Compaction and Virulence Regulation

Nina T. Odermatt, Claudia Sala, Andrej Benjak, Gaëlle S. Kolly, Anthony Vocat, Andréanne Lupien, Stewart T. Cole

Global Health Institute, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

ABSTRACT A handful of nucleoid-associated proteins (NAPs) regulate the vast majority of genes in a bacterial cell. H-NS, the histone-like nucleoid-structuring protein, is one of these NAPs and protects *Escherichia coli* from foreign gene expression. Though lacking any sequence similarity with *E. coli* H-NS, Rv3852 was annotated as the H-NS ortholog in *Mycobacterium tuberculosis*, as it resembles human histone H1. The role of Rv3852 was thoroughly investigated by immunoblotting, subcellular localization, construction of an unmarked *rv3852* deletion in the *M. tuberculosis* genome, and subsequent analysis of the resulting Δ *rv3852* strain. We found that Rv3852 was predominantly present in the logarithmic growth phase with a decrease in protein abundance in stationary phase. Furthermore, it was strongly associated with the cell membrane and not detected in the cytosolic fraction, nor was it secreted. The Δ *rv3852* strain displayed no growth defect or morphological abnormalities. Quantitative measurement of nucleoid localization in the Δ *rv3852* mutant strain compared to that in the parental H37Rv strain showed no difference in nucleoid position or spread. Infection of macrophages as well as severe combined immunodeficient (SCID) mice demonstrated that loss of Rv3852 had no detected influence on the virulence of *M. tuberculosis*. We thus conclude that *M. tuberculosis* Rv3852 is not involved in pathogenesis and is not a typical NAP. The existence of an as yet undiscovered Rv3852 ortholog cannot be excluded, although this role is likely played by the well-characterized Lsr2 protein.

IMPORTANCE *Mycobacterium tuberculosis* is the causative agent of the lung infection tuberculosis, claiming more than 1.5 million lives each year. To understand the mechanisms of latent infection, where *M. tuberculosis* can stay dormant inside the human host, we require deeper knowledge of the basic biology and of the regulatory networks. In our work, we show that Rv3852, previously annotated as H-NS, is not a typical nucleoid-associated protein (NAP) as expected from its initial annotation. Rv3852 from *M. tuberculosis* has neither influence on nucleoid shape or compaction nor a role in virulence. Our findings reduce the repertoire of identified nucleoid-associated proteins in *M. tuberculosis* to four transcription regulators and underline the importance of genetic studies to assign a function to bacterial genes.

KEYWORDS tuberculosis, *Mycobacterium tuberculosis*, H-NS, NAP, global regulation, *rv3852*

The chromosome must be heavily compacted to fit into a bacterial cell about 1,000 times smaller than the length of its DNA. Similarly to eukaryotes, where DNA is wrapped around histones for condensation, prokaryotes possess nucleoid-associated proteins (NAPs), which influence DNA topology. Typically, NAPs interact with the chromosome at several hundred binding sites in a sequence-specific or non-sequence-specific manner (1). As a result, NAPs can bridge, loop, and bend DNA (2), thus

Received 28 February 2017 Accepted 25 May 2017

Accepted manuscript posted online 30 May 2017

Citation Odermatt NT, Sala C, Benjak A, Kolly GS, Vocat A, Lupien A, Cole ST, 2017, Rv3852 (H-NS) of *Mycobacterium tuberculosis* is not involved in nucleoid compaction and virulence regulation. *J Bacteriol* 199:e00129-17. <https://doi.org/10.1128/JB.00129-17>.

Editor Olaf Schneewind, University of Chicago

Copyright © 2017 American Society for Microbiology. All Rights Reserved.

Address correspondence to Stewart T. Cole, stewart.cole@epfl.ch.

impacting global transcription regulation through activation or silencing of target genes (3). It is estimated that, together, a few NAPs control the expression of more than 50% of the genes in *Escherichia coli* (4).

In the well-characterized *E. coli* cell, at least 12 NAPs have been identified, among them the factor for inversion stimulation (Fis), the integration host factor (IHF), and the histone-like nucleoid-structuring protein (H-NS) (5). H-NS is also known as the “genome guardian” (2), as it silences xenogeneic genes acquired through horizontal gene transfer (6). *E. coli* H-NS predominantly forms dimers (7), has a preference for AT-rich bent DNA (8), and is expressed throughout the bacterial growth cycle (2).

Mycobacterium tuberculosis is considered to be the world’s most successful pathogen, causing 1.5 million deaths each year (9). Once *M. tuberculosis* has entered the host, it can establish a latent infection by persisting inside macrophages for several decades before regrowth leads to active tuberculosis (10). Several transcription factors contribute to gene regulation in *M. tuberculosis*, and one of these transcription factors, EspR, proved to behave as a NAP by binding to more than 160 loci, including virulence genes (11). In addition to EspR, four NAPs have been reported so far in *M. tuberculosis*: HupB, Lsr2, mycobacterial IHF (mIHF), and Rv3852 (12). The latter was named H-NS due to its N-terminal similarity to human histone H1 (12).

Rv3852 is a small protein of 134 amino acids, with a proposed molecular mass of 13.8 kDa, and is predicted to be nonessential by transposon mutagenesis (13). Rv3852 from *M. tuberculosis* has a very low sequence similarity with *E. coli* H-NS, but it is highly conserved among pathogenic mycobacteria. *Mycobacterium bovis* (14), *Mycobacterium marinum* (15), and even *Mycobacterium leprae* with its downsized genome (16) possess a copy of the *rv3852* gene, while it is absent from the nonpathogenic model organism *Mycobacterium smegmatis* (17). The mycobacterial proteins annotated as H-NS contain tetrapeptide repeats (PAKK and KAAK) that are critical for DNA binding by histones of the H1/H5 family, and this may explain the attribution of the *hns* gene name. The number of PAKK repeats varies between mycobacterial species (18), but all of the proteins contain a single conserved transmembrane domain at the C terminus.

Curiously, Werlang and colleagues reported that *M. tuberculosis* Rv3852 does not complement *hns* mutations in *E. coli* (19), whereas Lsr2, another NAP, does (20) and is therefore a functional homologue of *E. coli* H-NS. Ectopic expression of *M. tuberculosis* *rv3852* in *M. smegmatis* led to a less-compacted nucleoid, altered biofilm formation, and decreased sliding motility (21). Unusually for a NAP, the C-terminal domain of Rv3852 was identified as a transmembrane helix, suggesting a role in anchoring DNA to the membrane. RNA profiling showed that expression of *groEL1*, a heat shock chaperone, and *kasA*, involved in fatty acid biosynthesis, are downregulated upon expression of *rv3852* in *M. smegmatis* (21).

Here, we report the construction and in-depth characterization of an *M. tuberculosis* *rv3852* deletion mutant. By means of biological, molecular, and microscopy techniques, we show that Rv3852 is not involved in nucleoid compaction or virulence and plays only a minor role in transcriptional control.

RESULTS

Generation of an *rv3852* deletion mutant in *M. tuberculosis*. To probe the essentiality of the *rv3852* gene in *M. tuberculosis* strain H37Rv, we planned the construction of a deletion mutant by a two-step allelic exchange method based on the suicide vector pJG1100 (22, 23). Since *rv3852* was reported to be part of an operon with the downstream gene *rraA* (24), we first mapped the 5′ end of the bicistronic mRNA to ensure that the promoter region was not deleted and that *rraA* expression was not affected in the mutant. The 5′ rapid amplification of cDNA ends (5′-RACE) experiment identified a single 5′ end 47 bp upstream of the *rv3852* start codon (see Fig. S1 in the supplemental material). An in-frame deletion of the *rv3852* coding sequence was then designed. Homologous recombination with the pJG1100-derived plasmid pCS21 was induced in strain H37Rv and confirmed by colony PCR (data not shown). Deletion of *rv3852* was demonstrated in two sucrose-resistant clones (clones 1 and 17), which were

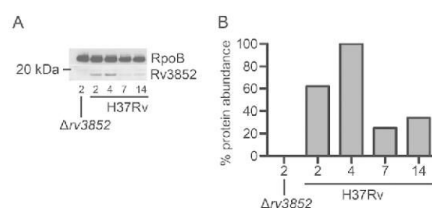


FIG 1 Analysis of Rv3852 abundance *in vitro*. (A) Immunoblot analysis of total protein extracts from *M. tuberculosis* H37Rv and $\Delta rv3852$ mutant strain at different time points. (B) The corresponding graph representing Rv3852 band intensity, normalized to RpoB, and relative to the highest value at day 4. Samples were taken on day 2 (OD_{600} of 0.21) for $\Delta rv3852$ mutant and on days 2, 4, 7, and 14 (corresponding to OD_{600} values of 0.28, 0.36, 0.9, and 2.4, respectively) for strain H37Rv. The experiment was repeated twice. A representative image is shown.

further analyzed by reverse transcription-PCR (RT-PCR) (data not shown) and by Southern blotting (Fig. S2), thus validating the deletion procedure. The resulting mutant (clone 1) was named $\Delta rv3852$ mutant and used for subsequent analyses.

The $\Delta rv3852$ mutant was transformed with plasmid pC524 (Table S1), harboring the native *rv3852* gene under the control of the constitutive *ptr* promoter, thereby obtaining the complemented mutant ($\Delta rv3852/rv3852$ [Fig. S3]). A control strain ($\Delta rv3852/pGA44$) was constructed by transforming $\Delta rv3852$ with the empty vector pGA44 (25).

***In vitro* phenotype of $\Delta rv3852$ deletion mutant.** To assess the *in vitro* growth dynamics of the $\Delta rv3852$ deletion mutant compared to its H37Rv parent, growth curves were generated in Middlebrook 7H9 complete medium. None of the strains showed any growth defects, either during exponential phase or in early stationary phase (Fig. S4A). In addition, the mutant strain carrying either the empty vector or the complementing plasmid did not display any abnormal phenotype, indicating that the lack of Rv3852 or ectopic expression has no influence on growth of *M. tuberculosis in vitro*.

The effect of external pH was evaluated by comparing growth in low-pH media, buffered at pH 5 and pH 6. As PhoPR is required to slow growth at acidic pH (26), the $\Delta phoP$ strain (27) was used as a control. At neutral pH, the $\Delta phoP$ mutant showed a slight growth defect (Fig. S4B). At pH 6, all strains had lower growth rates, while the $\Delta phoP$ mutant grew slightly faster than the others (Fig. S4C). No difference was noticed for the $\Delta rv3852$ mutant relative to the wild-type H37Rv strain under any of the conditions tested. In more-acidic medium (pH 5), the optical density did not reach values higher than 0.1 for any of the strains after 10 days (data not shown).

Since Rv3852 had been hypothesized to play a role as a NAP (19), thus contributing to shaping the chromosome, the MICs of drugs that have an effect on the topology of DNA were measured (28). We used the fluoroquinolones moxifloxacin and novobiocin targeting the GyrA and GyrB subunits of DNA gyrase, respectively, for this purpose together with rifampin, inhibiting RNA synthesis, as a control. The resazurin microplate assay (REMA) was carried out in Middlebrook 7H9 medium at pH 7. Results showed that the MIC was the same for all three strains tested, namely, wild-type H37Rv, $\Delta rv3852$ mutant, and the complemented mutant (Fig. S5). We concluded that lack of H-NS does not noticeably influence the susceptibility of *M. tuberculosis* to DNA gyrase-targeting drugs.

Rv3852 is most abundant in exponential phase and localizes to the cell membrane. Some NAPs are present throughout the whole bacterial cell cycle, while others are present only at a specific growth stage. In order to analyze to which category *M. tuberculosis* Rv3852 belongs, total protein extracts were prepared at different time points from wild-type H37Rv and probed by immunoblotting. The extract from the $\Delta rv3852$ strain was used as a negative control, and the constitutively expressed RNA polymerase subunit β , RpoB, was used as an internal standard. Rv3852 was identified at approximately 19 kDa (Fig. 1A) on a denaturing gel. Figure 1A and B demonstrate that Rv3852 is present in early exponential phase and reaches its maximum level in

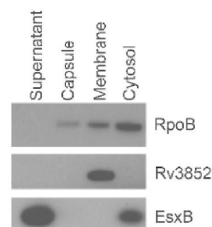


FIG 2 Localization of Rv3852 in subcellular fractions. Immunoblot of secreted (supernatant), capsular, membrane, and cytosolic protein fractions from strain H37Rv. RpoB was used as a lysis control for the supernatant, and EsxB was used as a control for the secreted fraction. Note that Rv3852 is found exclusively in the membrane fraction.

mid-exponential phase on day 4. Its abundance declines at the onset of stationary phase and remains at about 30% of the initial level in late stationary phase. This pattern suggests that Rv3852 plays its main role during the logarithmic phase of growth.

To localize Rv3852 subcellularly, total proteins from the wild-type strain were fractionated into secreted (culture supernatant), capsular, membrane, and cytosolic fractions. RpoB was used as a lysis control and proved that cells did not lyse, as it was not detected in the culture supernatant. EsxB, which was used as a secretion control, was present in the cytosol and in the supernatant. Importantly, EsxB was not seen in the membrane and capsule, indicating a clear separation of these fractions from the cytosol. Interestingly, Rv3852 was found only in the cell membrane and in none of the other subcellular compartments (Fig. 2).

Effect of Rv3852 on the nucleoid. To evaluate the effects of Rv3852 on nucleoid compaction and shape, the Δ r3852 and H37Rv strains were stained with SYTO9, which specifically binds to DNA, and analyzed by fluorescence microscopy. Representative images are shown in Fig. 3A. Nucleoid position relative to the cell, the number of nucleoid peaks, and nucleoid spread were evaluated in 54 mutant cells and 74 wild-type cells. No significant difference was observed for any of the tested parameters (Fig. 3B) (Student's *t* test). As chloramphenicol is known to contract the nucleoid (29), H37Rv cells treated with chloramphenicol were used as a control to detect any change in nucleoid morphology ($n = 50$). Results confirmed the previous findings (29).

Δ r3852 mutant displays normal cell morphology by scanning electron microscopy. To evaluate a potential morphological difference between wild-type H37Rv and its Δ r3852 mutant, high-resolution scanning electron microscopy (SEM) analysis was

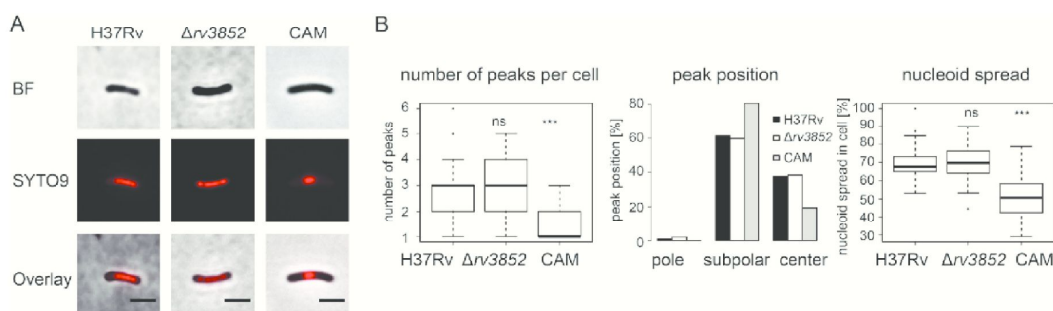


FIG 3 Impact of *r3852* deletion on nucleoid structure and shape. (A) Micrographs of the H37Rv strain, Δ r3852 mutant, and H37Rv treated with chloramphenicol (CAM) in bright-field (BF) channel, fluorescent channel (SYTO9), and overlay. Bars, 1 μ m. (B) No difference was observed between the Δ r3852 mutant and H37Rv strain in the number of peaks per cell, peak position inside the cell, and nucleoid spread (Student's *t* test), while H37Rv treated with chloramphenicol showed a lower number of peaks and a reduced nucleoid spread relative to nontreated H37Rv cells. Dots indicate outliers. ns, not significant; ***, $P < 0.001$.

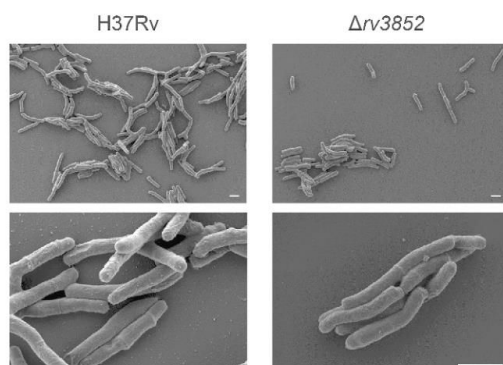


FIG 4 Scanning electron micrographs of H37Rv and Δ rv3852 bacterial cells. Bars, 1 μ m.

performed. Representative images are shown in Fig. 4. No obvious morphological difference was observed. Wild-type and mutant bacteria were similar in size, shape, and thickness. Cell length was measured on 104 and 106 cells for Δ rv3852 and H37Rv strains, respectively. The mean values were $1.93 \pm 0.52 \mu\text{m}$ for the Δ rv3852 mutant and $2.04 \pm 0.55 \mu\text{m}$ for strain H37Rv (no significant difference by Student's *t* test).

Global transcriptomic analysis. Transcriptome sequencing (RNA-seq) studies were undertaken to examine the impact of *rv3852* deletion on the global transcription profile. The analysis was performed on RNA extracted from exponential-phase cultures, where Rv3852 was shown to be most abundant. Between 16 and 18 million reads from each replicate mapped uniquely to the H37Rv genome and allowed quantification of gene expression as reported in Data Set S1. A false-discovery rate (FDR) of <1% and a cutoff fold change at \log_2 of 2 were applied to identify the deregulated genes. Surprisingly, only three features met these criteria. The most downregulated gene was *rv3852* with 0.04-fold expression in the mutant compared to the wild type, confirming loss of the gene in the Δ rv3852 strain. The *rv3852* upstream intergenic region (5' untranslated region [5'-UTR]) was also found to be highly downregulated with 0.05-fold expression, which suggests that Rv3852 controls its own expression. On the other hand, the *leuC-leuD* transcriptional unit was upregulated fourfold relative to the H37Rv parental strain.

We then relaxed the stringency conditions and set the cutoff at twofold change, resulting in identification of an additional 19 features (Table 1). Of the 22 deregulated features, most cluster at three different loci in the chromosome. Upstream of *leuC*, *rv2989* (probable transcriptional regulatory protein) and *rv2990* were expressed at a higher level in the mutant than in the wild type. Furthermore, the intergenic regions between *mce1R* and *fadD5* (both strands) genes and *mce1R* (transcriptional regulator) and *TB18.5* genes were found to be transcribed at a higher level in the mutant than in the wild type. A 2.2-fold change was observed for the *fadD11* mRNA, similar to the downstream gene *frdA*, which was 2.7-fold upregulated. Interestingly, while the *rv3852* 5'-UTR was almost undetected, the downstream *rraA* gene was found to be slightly more expressed in the mutant than in the wild-type strain. Another downregulated gene was the initiator tRNA *metU*. Overall, three genes encode probable transcriptional regulators.

It was reported that expression of *kasA* and *groEL1* was reduced fourfold upon ectopic expression of *M. tuberculosis* Rv3852 in *M. smegmatis* (21). We did not observe the same deregulation in our experiments, where *kasA*, *kasB*, the intergenic region between the two genes and *groEL1* were expressed at similar levels in the mutant and wild-type strains (Data Set S1). The same genes were tested by quantitative RT-PCR (qRT-PCR) on biologically different samples. Results confirmed the RNA-seq data, with an overall differential expression below 1.5-fold (data not shown).

TABLE 1 RNA-seq results with a false-discovery rate of <1% and a twofold change cutoff

Accession no. ^a	Fold change ^b	Feature name	Feature ^c
rv3852	0.04	rv3852	cds
rv3851-rv3852	0.05	rv3851- rv3852	igr
MTB000043	0.4	metU	tRNA
rv2450c-rv2451	0.4	rpfe-rv2451	igr
rv2028c	0.5	rv2028c	cds
rv2625c	0.5	rv2625c	cds
rev_rv0165c-rv0166 ^d	2.0	rev_mce1R-fadD5	igr
rv0232	2.1	rv0232	cds
rv1550	2.2	fadD11	cds
rv2485c	2.2	lipQ	cds
rv3853	2.3	rraA	cds
rev_rv0163-rv0164	2.3	rev_rv0163-TB18.5	igr
rev_rv0234c	2.3	rev_gabD1	cds
rev_rv0164	2.4	rev_TB18.5	cds
rv0165c	2.5	mce1R	cds
rv1552	2.7	frdA	cds
rev_rv2990c	3.1	rev_rv2990c	cds
rv2989	3.2	rv2989	cds
rv0165c-rv0166	3.3	mce1R-fadD5	igr
rv2987c-rv2988c	3.6	leuD-leuC	igr
rv2988c	3.9	leuC	cds
rv2987c	4.3	leuD	cds

^aTubercuList accession number (45).^bFold change in the mutant strain compared to the wild type.^ccds, coding sequence; igr, intergenic region.^drev, antisense feature.

Rv3852 does not influence virulence *ex vivo*. To test whether the lack of Rv3852 affected the *ex vivo* virulence of *M. tuberculosis*, macrophage infection assays were carried out. Differentiated THP-1 macrophages were infected at a multiplicity of infection (MOI) of 2.5 for 3 days. While the attenuated Δ RD1 strain (30), used as a control, was considerably less virulent, no significant difference was observed between the wild-type H37Rv strain, Δ rv3852 mutant strain, and the complemented strain (Fig. 5). The Δ rv3852 strain was as virulent as the wild type and effectively killed macrophages at the same level as H37Rv did.

Deletion of rv3852 has no impact on *in vivo* virulence. A more sensitive mouse infection model was chosen to investigate a potentially weak phenotype of the Δ rv3852 mutant. Severe combined immunodeficient (SCID) mice were infected with strain H37Rv, Δ rv3852 mutant strain, and the complemented strain by the aerosol route. The bacterial burden was evaluated by counting the CFU per lung and spleen after 1, 14 (acute infection), and 25 (onset of the chronic infection) days postinfection. Thirteen

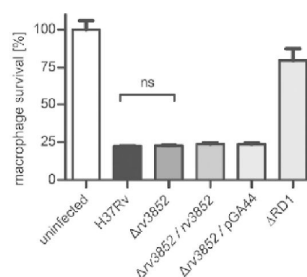


FIG 5 Virulence of the Δ rv3852 mutant compared to strain H37Rv in an *ex vivo* model. Macrophage survival was measured upon infection with H37Rv, Δ rv3852, and control strains. The Δ RD1 strain carries a deletion of the RD1 region and is attenuated. Data were normalized to the noninfected condition and are presented as means plus standard deviations (SD) (error bars) from five independent replicates. ns, not significant (Student's *t* test).

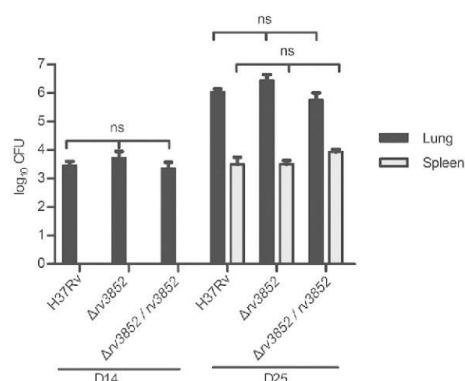


FIG 6 Virulence of the $\Delta rv3852$ mutant compared to strain H37Rv *in vivo*. The bacterial burden in the lung and in the spleen 14 days (D14) and 25 days (D25) postinfection with strain H37Rv, $\Delta rv3852$ mutant, and the complemented $\Delta rv3852/rv3852$ strain. Values are means plus SD of \log_{10} CFU for five SCID mice in each group at each time point. The values for the CFU level 14 days and 25 days postinfection for the $\Delta rv3852$ mutant or complemented $\Delta rv3852/rv3852$ strain relative to the H37Rv strain in lungs and spleens were not significantly different (ns) by Student's *t* test.

CFU for the mutant and the complemented strain and 33 CFU for the wild type were detected 1 day postinfection. Two weeks later, the number of bacteria reached 10^3 to 10^4 in the lungs but was below the detection limit of 20 CFU/organ in the spleens. After 25 days, the bacterial load in the lungs increased to approximately 10^6 CFU for all three strains, whereas 10^4 CFU were counted in the spleens (Fig. 6). Overall, no significant difference was observed between the strains. At days 14 and 25, spleen size and macroscopic lesions of the lungs were evaluated. As shown in Fig. S6, the spleen size increased marginally from 1 day postinfection to the 25-day time point. However, again, no difference was noted between the three strains (Fig. S6A). Similarly, macroscopic lesions in the lungs were visible at day 25 in animals infected with the wild-type strain and animals infected with the $\Delta rv3852$ strain (Fig. S6B).

DISCUSSION

As part of our systematic investigation of the NAPs of *M. tuberculosis*, which began with the molecular characterization of EspR, a nonessential NAP that impacts virulence (11, 18, 31), we have extended the study to the Rv3852 protein. The *rv3852* open reading frame was removed from the *M. tuberculosis* chromosome by means of allelic exchange, indicating that Rv3852 is not essential for growth. Extensive characterization of the $\Delta rv3852$ mutant was performed. Expression of the downstream gene *rraA*, which forms an operon with *rv3852* (24), was analyzed by RNA-seq and was at the level of the parental strain. A wide range of phenotypic tests revealed no differences in the behavior of the $\Delta rv3852$ mutant compared to the parent strain be it *in vitro*, *ex vivo*, or during infection of mice.

With respect to NAP activity, loss of Rv3852 had no effect on nucleoid shape or size (Fig. 3), contrary to the findings of Ghosh and colleagues upon overexpression of *M. tuberculosis* H-NS in *M. smegmatis* (21). Moreover, unlike Ghosh et al. (21), who reported deregulation of the fatty acid biosynthetic genes *kasA* and *kasB* and of the major chaperone gene *groEL1*, we observed no change in expression of these genes when comparing the global transcriptome of the wild-type strain H37Rv to that of the $\Delta rv3852$ mutant by RNA-seq. Besides, we found only two genes involved in fatty acid metabolism (*fadD11* and *lipQ*) and three coding for probable transcriptional regulators (*rv0232*, *mce1R*, and *rv2989*) that were slightly deregulated in expression. Overall, our experimental results rule out a role for Rv3852 as a global transcriptional regulator. The discrepancy between our findings and those of Ghosh et al. (21) could be ascribed to

the ectopic expression of *M. tuberculosis* H-NS in *M. smegmatis*, a species that has no ortholog of *rv3852* in its genome (17). The expression system employed by Ghosh and coworkers may have altered the regulatory network in the bacterium, thus generating an aberrant phenotype.

By immunoblotting (Fig. 1), we demonstrated that Rv3852 is mainly expressed during the logarithmic phase of growth, which is consistent with the findings of two recent investigations. In one investigation, the transcription of *rv3852* was found to be reduced when *M. tuberculosis* entered a nonreplicating state (32), while in the other, *rv3852* expression levels were sevenfold higher in exponential phase than in stationary phase (33). These findings are consistent with those of Sharadamma et al. (34), who showed that *in vitro* Rv3852 binds to DNA replication/repair intermediates and to Holliday junctions, which are mainly generated during replication and exponential growth. Another important feature of the *M. tuberculosis* Rv3852 protein is its subcellular localization, since it was detected only in the cell membrane and not in the other subcellular compartments or in the culture supernatant (Fig. 2), as reported earlier in a proteomic survey (35). This localization is compatible with a role for Rv3852 in chromosome segregation during cell division, as proposed by Ghosh et al. (21). However, although there is a clear transmembrane domain, encompassing amino acid residues 110 to 128, immediately preceding the C terminus of the Rv3852 protein, the bulk of the protein is predicted to be in the periplasm. Further topological investigation of Rv3852 is required.

Recently, during a screen for inhibitors that disrupt the intrabacterial pH in an acidic environment, the natural product agrimophol was found to show activity against *M. tuberculosis in vitro* (36). Agrimophol, which is used in traditional Chinese medicine to treat pulmonary infections, was thought to target Rv3852. However, deletion of the *rv3852* gene in *M. tuberculosis* did not mimic the effect of agrimophol or impact virulence in the macrophage infection model (36). These results confirm our findings *in vitro* where we proved that Rv3852 does not respond to changes in external pH and demonstrated that Rv3852 does not play a role upon intracellular infection.

Taken together, our combined data indicate that Rv3852 does not act as a NAP in any of the experimental conditions tested, thus raising doubts about its functional attribution. Other investigators have proposed that Lsr2, a bona fide NAP (20), is the true ortholog of the *E. coli* H-NS, and this appears to be the case. Finally, the nature of the true function of *rv3852* remains to be elucidated.

MATERIALS AND METHODS

Strains, media, and chemicals. *M. tuberculosis* H37Rv, *rv3852* deletion mutant ($\Delta rv3852$), and *phoP* deletion mutant ($\Delta phoP$) (27) were grown at 37°C either in Middlebrook 7H9 broth (Difco) supplemented with 10% albumin-dextrose-catalase, 0.2% glycerol, and 0.05% Tween 80 or in Sauton's liquid medium supplemented with 0.005% Tween 80. Where necessary, the medium was buffered with 10% morpholineethanesulfonic acid (MES) (pH 5 or 6) or morpholinepropanesulfonic acid (MOPS) (pH 7). Cultures were plated on Middlebrook 7H10 (Difco) agar supplemented with 10% oleic acid-albumin-dextrose-catalase and 0.2% glycerol. Hygromycin (50 $\mu\text{g ml}^{-1}$), kanamycin (25 $\mu\text{g ml}^{-1}$), streptomycin (25 $\mu\text{g ml}^{-1}$), chloramphenicol (30 $\mu\text{g ml}^{-1}$), or 2.5% sucrose was added when needed. For cloning procedures, One Shot TOP10 chemically competent *Escherichia coli* (Invitrogen) was grown in Luria-Bertani (LB) broth or on LB agar with hygromycin (200 $\mu\text{g ml}^{-1}$), kanamycin (50 $\mu\text{g ml}^{-1}$), or spectinomycin (25 $\mu\text{g ml}^{-1}$). All chemicals were purchased from Sigma-Aldrich, unless otherwise stated.

Plasmid construction. Two 900-bp-long fragments corresponding to the upstream and downstream regions of *rv3852* were generated by PCR amplification using the U-fwd (fwd stands for forward) and U-rev (rev stands for reverse) primers and the D-fwd and D-rev primers, listed in Table S2 in the supplemental material. Fragments were ligated with the AvrII restriction site and cloned into the PacI and AscI sites of pJG1100 (22, 23).

To complement the *rv3852* deletion strain, pCS24 was constructed, where the *rv3852* gene was amplified with primers *rv3852*-F (F stands for forward) and *rv3852*-R (R stands for reverse) and cloned in front of the *ptr* promoter into pGA44 (25), which stably integrates at the *LS attB* site.

For plasmid generation purposes, all PCR products were ligated into the pCR-Blunt II-TOPO vector (Invitrogen) according to the manufacturer's recommendations. The plasmids and oligonucleotide sequences used in this study are given in Tables S1 and S2, respectively. All enzymes were purchased from New England Biolabs.

Construction of the unmarked in-frame $\Delta rv3852$ deletion strain. Deletion of *rv3852* was accomplished by homologous recombination using the pJG1100-derived vector pCS21. After transformation of

M. tuberculosis H37Rv, the first recombination event was selected on Middlebrook 7H10 medium containing hygromycin and kanamycin. Colonies were screened by colony PCR using primers 329-330 and 331-332 (Table S2). Positive clones were plated on Middlebrook 7H10 medium supplemented with sucrose to select for the second crossing over and loss of pJG1100. The resulting clones were tested by colony PCR with primers CS-351-rv3852F and CS-352-rv3852R by reverse transcription-PCR for loss of rv3852 expression and further confirmed by Southern blotting. The probe used for Southern blotting was amplified with primers U-fwd and D-rev.

Generation of the complemented strain was achieved by transformation of the rv3852 deletion strain with pCS24, providing the integrase in *trans* with pGA80 (25), selecting on streptomycin. The empty vector pGA44 was transformed in the same manner to generate the Δ rv3852 empty vector control strain.

Growth curve measurements and MIC determination. To characterize the growth of the Δ rv3852 mutant, the strains were grown to mid-logarithmic phase and then diluted to an optical density at 600 nm (OD_{600}) of 0.05 in Middlebrook 7H9 medium. The OD_{600} was recorded at different time points to obtain the growth curves.

MIC determination using the resazurin reduction microplate assay (REMA) was performed as previously described (37). Briefly, *M. tuberculosis* was grown to mid-log phase, diluted to an OD_{600} of 1×10^{-4} , and added to a 96-well plate at a volume of 100 μ l. The drug being tested was added at twofold serial dilutions to the bacteria, and the plates were then incubated at 37°C for 7 days before the addition of 0.025% resazurin. After overnight incubation, the fluorescence of the resazurin metabolite resorufin was determined (excitation at 560 nm and emission at 590 nm) by using a Tecan Infinite M200 microplate reader. Results were plotted in GraphPad Prism 5, and MIC was determined by the Gompertz equation.

Genomic DNA extraction and Southern blotting. Mycobacterial genomic DNA was extracted using standard protocols (38). To confirm successful allelic exchange of rv3852, genomic DNA was digested with NcoI or NcoI plus AvrII restriction enzymes. DNA fragments were separated by 0.8% agarose gel electrophoresis before capillary blotting onto a Hybond-N+ nylon membrane (GE Healthcare) and hybridization with a probe corresponding to the same upstream and downstream region of rv3852 cloned into pJG1100. Hybridization was carried out using the ECL direct nucleic acid labeling and detection system (GE Healthcare) as recommended by the manufacturer.

Total RNA extraction. *M. tuberculosis* cultures were harvested by centrifugation, and pellets were resuspended in TRIzol reagent (Thermo Fisher Scientific) and stored at -80° until further processing. Total RNA was extracted by bead beating as previously described (39). The integrity of RNA was checked by agarose gel electrophoresis, and the purity and amount of RNA were assessed using a Nanodrop instrument and Qubit fluorometric quantitation assay kit (Thermo Fisher Scientific), respectively. SuperScript III first-strand synthesis system (Invitrogen) was used to generate randomly primed cDNA from 500 ng of RNA, according to the manufacturer's recommendations.

Library preparation for RNA-seq analysis and Illumina high-throughput sequencing. A total of 300 ng of total RNA was mixed with 5 \times fragmentation buffer (Applied Biosystems), incubated at 70°C for 4 min, and then transferred immediately to ice. RNA was purified using RNeasy XP beads (Beckman Coulter), according to the manufacturer's recommendations, and subsequently treated with Antarctic phosphatase (New England Biolabs). RNA was then rephosphorylated at the 5' end with polynucleotide kinase (New England Biolabs) and purified with RNeasy MinElute columns (Qiagen). To ensure strand specificity, v1.5 small RNA (sRNA) adapters (Illumina) were ligated at the 5' and 3' ends using RNA ligase. Reverse transcription was carried out using SuperScript III reverse transcriptase (Invitrogen) and SRA RT primer (Illumina). Twelve cycles of PCR amplification using Phusion DNA polymerase were then performed, and the resulting library was purified with AMPure beads (Beckman Coulter) per the manufacturer's instructions and sequenced on the Illumina HiSeq 2000 instrument using the TruSeq SR cluster generation kit v3 and TruSeq SBS kit v3. Data were processed with the Illumina Pipeline software v1.82. RNA-seq experiments were performed on two biological replicates per strain.

RNA-seq analysis. Reads were mapped against the H37Rv reference strain genome sequence (NC_000962.2) using Bowtie 2 (40). Read counting over features was done using *htseq-count* (41) and the annotation from Tuberculist. To take advantage of strand specificity of the libraries and to maximize detection of putative unknown transcriptionally active regions, we also included the reverse orientation of each feature, as well as intergenic regions. Differential gene expression analysis was done using DESeq (42).

Quantitative PCR. Primers listed in Table S2 were used for quantification of gene expression. Primers CS-057 and CS-058 for *sigA* were used to normalize the amount of cDNA template added to each sample. qRT-PCR was carried out in duplicate using 7900HT sequence detection system and Power SYBR green PR master mix (Applied Biosystems) according to the manufacturer's recommendations. The $\Delta\Delta C_T$ method was used for quantification.

5' RACE. Two micrograms of *M. tuberculosis* H37Rv RNA and 1 μ g of primer CS-397 were incubated at 70°C for 5 min and then at 55°C for 1 h in the presence of 1 \times cDNA synthesis buffer, 1 mM each deoxynucleoside triphosphate (dNTP), 40 U RNase inhibitor, 25 U Transcriptor reverse transcriptase (5'/3' RACE kit, second generation; Roche). cDNA was then purified with the High Pure PCR product purification kit (Roche) and used in the subsequent poly(A) tailing reaction (30 min at 37°C in the presence of 0.2 mM dATP and 80 U terminal transferase; Roche). Seminested PCR amplification on poly(A)-tailed cDNA was performed using an oligo(dT) anchor primer (CS-81) and primer CS-395. Only one amplification product was obtained and directly sequenced.

Protein extraction, immunoblot analysis, and subcellular fractionation. *M. tuberculosis* cultures grown in Middlebrook 7H9 medium were pelleted at different time points by centrifugation, washed once in Tris-buffered saline (TBS) (20 mM Tris-HCl [pH 7.5], 150 mM NaCl), and stored at -80° C until

further processing. The cells were sonicated in TBS supplemented with a protease inhibitor tablet (Roche) for 15 min, and the protein solution was then sterilized by filtration through a 0.2- μ m filter to remove any residual intact cells. Protein samples were quantified using the Qubit fluorometric quantitation instrument (Thermo Fisher Scientific). Equal amounts of protein preparations were loaded on SDS-polyacrylamide 12 to 15% NuPAGE gels (Invitrogen) and transferred onto polyvinylidene difluoride (PVDF) membranes using a semidry electrophoresis transfer apparatus (Bio-Rad). The membranes were incubated in TBS-Tween blocking buffer (25 mM Tris [pH 7.5], 150 mM NaCl, 0.05% Tween 20) with 5% (wt/vol) skimmed milk powder for 3 h prior to overnight incubation with primary antibody. The membranes were washed in TBS-Tween blocking buffer three times and then incubated with secondary antibody for 2 h before washing. Signals were detected using chemiluminescent peroxidase substrate 1 (Sigma-Aldrich).

Primary anti-Rv3852 antibody was produced by Eurogentec against Rv3852 peptides and used at a concentration of 1:4,000 in immunoblots. Horseradish peroxidase (HRP)-conjugated anti-rabbit (Sigma-Aldrich) secondary antibody was used at a 1:150,000 dilution. Anti-RpoB antibodies (NeoClone) were used as an internal loading control. The band intensity of immunoblots was analyzed with Fiji/ImageJ and normalized to the intensity of the RpoB signal.

For cell fractionation, 50-ml cultures of H37Rv strain and Δ r3852 mutant were grown in Sauton's medium with 0.005% Tween 80. Fractions were obtained as described previously (43). Briefly, cells were collected by centrifugation, supernatant was filtered and concentrated 100 times to obtain the secreted fraction. The pellet was treated with 0.25% Genapol-X080 for 30 min followed by centrifugation at $14,000 \times g$ for 10 min, and the proteins of the resulting supernatant were precipitated with trichloroacetic acid (TCA), yielding the capsular fraction. The remaining pellet was subjected to sonication to break the cells, sterilized by filtration through a 0.2- μ m filter, and centrifuged at 45,000 rpm for 1 h in an ultracentrifuge. The supernatant contained the cytosolic fraction, while the pellet was enriched with membrane proteins.

Ex vivo virulence assay. THP-1 cells were grown in RPMI 1640 medium containing 10% fetal bovine serum (HyClone) and incubated at 37°C with 5% CO₂. To investigate the virulence of the Δ r3852 mutant compared to strain H37Rv, the strains were grown to mid-exponential phase and used to infect macrophages according to a previously published protocol (25, 44). Briefly, 100,000 THP-1 cells per well were seeded into 96-well plates, differentiated into macrophages by adding 100 nM phorbol myristate acetate (PMA) overnight and then infected with the various *M. tuberculosis* strains at a multiplicity of infection (MOI) of 2.5. Macrophage survival was measured 3 days postinfection by exposing the infected cells to PrestoBlue cell viability reagent (Life Technology) for 1 h. Fluorescence was read using a Tecan Infinite M200 microplate reader, and statistical significance was calculated by an unpaired Student's *t* test method.

In vivo virulence studies. Log-phase bacterial cultures of wild-type strain H37Rv, Δ r3852 deletion mutant, and the complemented strain were grown to mid-exponential phase and centrifuged briefly to exclude cell clumps, and the supernatant was diluted in phosphate-buffered saline (PBS) (supplemented with 0.05% Tween 80) to an OD₆₀₀ of 0.3. Seven-week-old female Fox Chase SCID mice (Charles River Laboratories) were infected via the aerosol route with the bacterial suspension by using a custom-built aerosol exposure chamber (Mechanical Engineering Shops, University of Wisconsin, Madison) to deliver approximately 300 CFU to the lungs of each mouse.

Three mice per group were sacrificed on day 1 postinfection, and lung homogenates were plated to assess the level of infection. Five mice per group and time point were sacrificed at week 2 and week 3.5 postinfection, and lung and spleen homogenates were plated on Middlebrook 7H10 medium plates supplemented with cycloheximide (10 μ g ml⁻¹) and ampicillin (50 μ g ml⁻¹). Spleen size was measured at the same time points. Experimental procedures involving animals were approved by the Swiss Cantonal and Federal Authorities (authorization 3082).

Fluorescence microscopy and image analysis. *M. tuberculosis* H37Rv and Δ r3852 mutant strain were grown to early logarithmic phase in Middlebrook 7H9 medium, washed once in TBS-Tween and stained with 10 μ M SYTO9 (Thermo Fisher Scientific) for 25 min. Cells were mounted on an agarose pad and visualized with an Olympus IX81 fluorescence microscope under a 100 \times objective.

The resulting images were analyzed in Fiji/ImageJ. Each cell was visually inspected to be a single, straight cell before plotting a straight-line intensity profile in bright-field and fluorescence channels. Cell boundaries were determined as 10% decrease in intensity relative to the minimum value of the bright-field picture. Nucleoid boundaries were determined as 25% increase in fluorescence relative to the maximum value. All profiles were plotted in R, again visually analyzed to exclude any wrong boundary determination and used to define cell margins, nucleoid localization, and spread (as a percentage of nucleoid length relative to cell length). To assign a position to the nucleoid peaks relative to the individual cell length, each cell was divided into five even sections. The middle of the cell being the center, the two most polar sections were merged, as well as the two intermediary sections between the pole and the center (subpolar).

Scanning electron microscopy. Whole-cell scanning electron microscopy was done as previously published (25). Briefly, the Δ r3852 mutant and wild-type H37Rv strains were grown in Middlebrook 7H9 medium until mid-exponential phase, pelleted, washed in PBS, and resuspended to an OD₆₀₀ of 0.5. The samples were then fixed on a coverslip in a solution of 1.25% glutaraldehyde and 1% tannic acid in phosphate buffer (0.1 M; pH 7.4) for 1 to 2 h. The samples were washed in cacodylate buffer prior to fixation for 30 min in 1% osmium tetroxide in cacodylate buffer and washed again twice for 3 min each time in water. Gradual dehydration was carried out in ethanol. Finally, a metal coat (Au-Pd alloy) was

applied to the critically dried samples. Images were obtained with a Zeiss Merlin scanning electron microscope.

Statistics. Statistical analysis was performed with unpaired *t* test in Prism version 5.0 (GraphPad, San Diego, CA).

Accession number(s). Raw sequencing data have been deposited in the GEO database under accession number [GSE95181](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE95181).

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at <https://doi.org/10.1128/JB.00129-17>.

SUPPLEMENTAL FILE 1, XLSX file, 1.8 MB.

SUPPLEMENTAL FILE 2, PDF file, 0.8 MB.

ACKNOWLEDGMENTS

We thank the staff at the École Polytechnique Fédérale de Lausanne (EPFL) Electron Microscopy Core Facility, EPFL Bioimaging and Optics Platform, and Genomic Technologies Facility at the University of Lausanne for technical assistance and Jacques Rougemont for setting up a basic script for cell boundary determination.

This work was supported by the Swiss National Science Foundation (grant 31003A-162641).

REFERENCES

- Luijsterburg MS, Noom MC, Wuite GJL, Dame RT. 2006. The architectural role of nucleoid-associated proteins in the organization of bacterial chromatin: a molecular perspective. *J Struct Biol* 156:262–272. <https://doi.org/10.1016/j.jsb.2006.05.006>.
- Dillon SC, Dorman CJ. 2010. Bacterial nucleoid-associated proteins, nucleoid structure and gene expression. *Nat Rev Microbiol* 8:185–195. <https://doi.org/10.1038/nrmicro2261>.
- Gengenbacher M, Kaufmann SHE. 2012. *Mycobacterium tuberculosis*: success through dormancy. *FEMS Microbiol Rev* 36:514–532. <https://doi.org/10.1111/j.1574-6976.2012.00331.x>.
- Martinez-Antonio A, Collado-Vides J. 2003. Identifying global regulators in transcriptional regulatory networks in bacteria. *Curr Opin Microbiol* 6:482–489. <https://doi.org/10.1016/j.mib.2003.09.002>.
- Browning DF, Grainger DC, Busby SJ. 2010. Effects of nucleoid-associated proteins on bacterial chromosome structure and gene expression. *Curr Opin Microbiol* 13:773–780. <https://doi.org/10.1016/j.mib.2010.09.013>.
- Lang B, Blot N, Bouffartigues E, Buckle M, Geertz M, Gualerzi CO, Mavathur R, Muskhelishvili G, Pon CL, Rimsky S, Stella S, Babu MM, Travers A. 2007. High-affinity DNA binding sites for H-NS provide a molecular basis for selective silencing within proteobacterial genomes. *Nucleic Acids Res* 35:6330–6337. <https://doi.org/10.1093/nar/gkm712>.
- Falconi M, Gualtieri MT, Losso MA. 1988. Proteins from the prokaryotic nucleoid: primary and quaternary structure of the 15-kD *Escherichia coli* DNA binding protein H-NS. *Mol Microbiol* 2:323–329. <https://doi.org/10.1111/j.1365-2958.1988.tb00035.x>.
- Dorman CJ. 2004. H-NS: a universal regulator for a dynamic genome. *Nat Rev Microbiol* 2:391–400. <https://doi.org/10.1038/nrmicro883>.
- World Health Organization. 2014. WHO tuberculosis fact sheet. WHO fact sheet no. 104. World Health Organization, Geneva, Switzerland.
- Wayne LG. 1994. Dormancy of *Mycobacterium tuberculosis* and latency of disease. *Eur J Clin Microbiol Infect Dis* 13:908–914. <https://doi.org/10.1007/BF02111491>.
- Blasco B, Chen JM, Hartkoorn R, Sala C, Uplekar S, Rougemont J, Pojer F, Cole ST. 2012. Virulence regulator EspR of *Mycobacterium tuberculosis* is a nucleoid-associated protein. *PLoS Pathog* 8:e1002621. <https://doi.org/10.1371/journal.ppat.1002621>.
- Cole ST, Brosch R, Parkhill J, Garnier T. 1998. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* 396:651–653. <https://doi.org/10.1038/24206>.
- Griffin JE, Gawronska JD, Dejesus MA, Joerges TR, Akerley BJ, Sassetti CM. 2011. High-resolution phenotypic profiling defines genes essential for mycobacterial growth and cholesterol catabolism. *PLoS Pathog* 7:e1002251. <https://doi.org/10.1371/journal.ppat.1002251>.
- Garnier T, Eiglmeyer K, Camus JC, Medina N, Mansoor H, Pryor M, Duthoy S, Grondin S, Lacroix C, Monsemp C. 2003. The complete genome sequence of *Mycobacterium bovis*. *Proc Natl Acad Sci U S A* 100:7877–7882. <https://doi.org/10.1073/pnas.1130426100>.
- Stinear TP, Seemann T, Harrison PF, Jenkin GA, Davies JK, Johnson PDR, Abdellah Z, Arrowsmith C, Chillingworth T, Churcher C, Clarke K, Cronin A, Davis P, Goodhead I, Holroyd N, Jagels K, Lord A, Moule S, Mungall K, Norbertczak H, Quail MA, Rabinowitsch E, Walker D, White B, Whitehead S, Small PLC, Brosch R, Ramakrishnan L, Fischbach MA, Parkhill J, Cole ST. 2008. Insights from the complete genome sequence of *Mycobacterium marinum* on the evolution of *Mycobacterium tuberculosis*. *Genome Res* 18:729–741. <https://doi.org/10.1101/gr.075069.107>.
- Cole ST, Eiglmeyer K, Parkhill J, James KD, Thomson NR, Wheeler PR, Honoré N, Garnier T, Churcher C, Harris D, Mungall K, Basham D, Brown D, Chillingworth T, Connor R, Davies RM, Devlin K, Duthoy S, Feltwell T, Fraser A, Hamlin N, Holroyd S, Homsby T, Jagels K, Lacroix C, Maclean J, Moule S, Murphy L, Oliver K, Quail MA, Rajandream MA, Rutherford KM, Rutter S, Seeger K, Simon S, Simmonds M, Skelton J, Squares R, Squares S, Stevens K, Taylor K, Whitehead S, Woodward JR, Barrell BG. 2001. Massive gene decay in the leprosy bacillus. *Nature* 409:1007–1011. <https://doi.org/10.1038/35059006>.
- Mohan A, Padladpu J, Baloni P, Chandra N. 2015. Complete genome sequences of a *Mycobacterium smegmatis* laboratory strain (MC² 155) and isoniazid-resistant (4XR1/R2) mutant strains. *Genome Announc* 3:e01520-14. <https://doi.org/10.1128/genomeA.01520-14>.
- Raghavan S, Manzanillo P, Chan K, Dovey C, Cox JS. 2008. Secreted transcription factor controls *Mycobacterium tuberculosis* virulence. *Nature* 454:717–721. <https://doi.org/10.1038/nature07219>.
- Werlang ICR, Schneider CZ, Mendonça JD, Palma MS, Basso LA, Santos DS. 2009. Identification of Rv3852 as a nucleoid-associated protein in *Mycobacterium tuberculosis*. *Microbiology* 155:2652–2663. <https://doi.org/10.1099/mic.0.030148-0>.
- Gordon BRG, Imperial R, Wang L, Navarre WW, Liu J. 2008. Lsr2 of *Mycobacterium* represents a novel class of H-NS-like proteins. *J Bacteriol* 190:7052–7059. <https://doi.org/10.1128/JB.00733-08>.
- Ghosh S, Indi SS, Nagaraja V. 2013. Regulation of lipid biosynthesis, sliding motility and biofilm formation by a membrane-anchored nucleoid associated protein of *Mycobacterium tuberculosis*. *J Bacteriol* 195:1769–1778. <https://doi.org/10.1128/JB.02081-12>.
- Gomez JE, Bishai WR. 2000. whmD is an essential mycobacterial gene required for proper septation and cell division. *Proc Natl Acad Sci U S A* 97:8554–8559. <https://doi.org/10.1073/pnas.140225297>.
- Gagneux S, Burgos MV, DeRiemer K, Enciso A, Muñoz S, Hopewell PC, Small PM, Pym AS. 2006. Impact of bacterial genetics on the transmission of isoniazid-resistant *Mycobacterium tuberculosis*. *PLoS Pathog* 2:e61. <https://doi.org/10.1371/journal.ppat.0020061>.
- Cortes T, Schubert OT, Rose G, Arnvig KB, Comas I, Aebbersold R, Young

- DB. 2013. Genome-wide mapping of transcriptional start sites defines an extensive leaderless transcriptome in *Mycobacterium tuberculosis*. *Cell Rep* 5:1121–1131. <https://doi.org/10.1016/j.celrep.2013.10.031>.
25. Kolly GS, Boldrin F, Sala C, Dhar N, Hartkoorn RC, Ventura M, Serafini A, McKinney JD, Manganelli R, Cole ST. 2014. Assessing the essentiality of the decaprenyl-phospho-*o*-arabinofuranose pathway in *Mycobacterium tuberculosis* using conditional mutants. *Mol Microbiol* 92:194–211. <https://doi.org/10.1111/mmi.12546>.
 26. Baker JJ, Johnson BK, Abramovitch RB. 2014. Slow growth of *Mycobacterium tuberculosis* at acidic pH is regulated by phoPR and host-associated carbon sources. *Mol Microbiol* 94:56–69. <https://doi.org/10.1111/mmi.12688>.
 27. Gonzalo-Asensio J, Soto CY, Arbués A, Sancho J, Menéndez MDC, García MJ, Gicquel B, Martin C. 2008. The *Mycobacterium tuberculosis* *phoPR* operon is positively autoregulated in the virulent strain H37Rv. *J Bacteriol* 190:7068–7078. <https://doi.org/10.1128/JB.00712-08>.
 28. Stuger R, Woldringh CL, Van der Weijden CC, Vischer NOE, Bakker BM, Van Spanning RJM, Snoep JL, Weterhoff HV. 2002. DNA supercoiling by gyrase is linked to nucleoid compaction. *Mol Biol Rep* 29:79–82. <https://doi.org/10.1023/A:1020318705894>.
 29. Manasherob R, Zaritsky A, Metzler Y, Ben-Dov E, Itsko M, Fishov I. 2003. Compaction of the *Escherichia coli* nucleoid caused by Cyt1Aa. *Microbiology* 149:3553–3564. <https://doi.org/10.1099/mic.0.26271-0>.
 30. Hsu T, Hingley-Wilson SM, Chen B, Chen M, Dai AZ, Morin PM, Marks CB, Padiyar J, Goulding C, Gingery M, Eisenberg D, Russell RG, Derrick SC, Collins FM, Morris SL, King CH, Jacobs WR. 2003. The primary mechanism of attenuation of bacillus Calmette-Guerin is a loss of secreted lytic function required for invasion of lung interstitial tissue. *Proc Natl Acad Sci U S A* 100:12420–12425. <https://doi.org/10.1073/pnas.1635213100>.
 31. Blasco B, Stenta M, Alonso-Sarduy L, Dietler G, Peraro MD, Cole ST, Pojer F. 2011. Atypical DNA recognition mechanism used by the *EspR* virulence regulator of *Mycobacterium tuberculosis*. *Mol Microbiol* 82:251–264. <https://doi.org/10.1111/j.1365-2958.2011.07813.x>.
 32. Benjak A, Uplekar S, Zhang M, Piton J, Cole ST, Sala C. 2016. Genomic and transcriptomic analysis of the streptomycin-dependent *Mycobacterium tuberculosis* strain 18b. *BMC Genomics* 17:190. <https://doi.org/10.1186/s12864-016-2528-2>.
 33. Uplekar S, Rougemont J, Cole ST, Sala C. 2013. High-resolution transcriptome and genome-wide dynamics of RNA polymerase and NusA in *Mycobacterium tuberculosis*. *Nucleic Acids Res* 41:961–977. <https://doi.org/10.1093/nar/gks1260>.
 34. Sharadamma N, Harshavardhana Y, Singh P, Muniyappa K. 2010. *Mycobacterium tuberculosis* nucleoid-associated DNA-binding protein H-NS binds with high-affinity to the Holliday junction and inhibits strand exchange promoted by RecA protein. *Nucleic Acids Res* 38:3555–3569. <https://doi.org/10.1093/nar/gkq064>.
 35. de Souza G, Leversen N, Målen H, Wiker HG. 2011. Bacterial proteins with cleaved or uncleaved signal peptides of the general secretory pathway. *J Proteomics* 75:502–510. <https://doi.org/10.1016/j.jprot.2011.08.016>.
 36. Zhao N, Sun M, Burns-Huang K, Jiang X, Ling Y, Darby C, Ehrst S, Liu G, Nathan C. 2015. Identification of Rv3852 as an agrimophol-binding protein in *Mycobacterium tuberculosis*. *PLoS One* 10:e0126211. <https://doi.org/10.1371/journal.pone.0126211>.
 37. Palomino J, Martin A, Camacho M, Guerra H, Swings J, Portaels F. 2002. Resazurin microtiter assay plate: simple and inexpensive method for detection of drug resistance in *Mycobacterium tuberculosis*. *Antimicrob Agents Chemother* 46:2720–2722. <https://doi.org/10.1128/AAC.46.8.2720-2722.2002>.
 38. Pelicic V, Jackson M, Reyrat J-M, Jacobs WR, Gicquel B, Guilhot C. 1997. Efficient allelic exchange and transposon mutagenesis in *Mycobacterium tuberculosis*. *Proc Natl Acad Sci U S A* 94:10955–10960. <https://doi.org/10.1073/pnas.94.20.10955>.
 39. Jungwirth B, Sala C, Kohl TA, Uplekar S, Baumbach J, Cole ST, Pühler A, Tauch A. 2013. High-resolution detection of DNA binding sites of the global transcriptional regulator GlxR in *Corynebacterium glutamicum*. *Microbiology* 159:12–22. <https://doi.org/10.1099/mic.0.062059-0>.
 40. Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9:357–359. <https://doi.org/10.1038/nmeth.1923>.
 41. Anders S, Pyl PT, Huber W. 2015. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* 31:166–169. <https://doi.org/10.1093/bioinformatics/btu638>.
 42. Anders S, Huber W. 2010. Differential expression analysis for sequence count data. *Genome Biol* 11:R106. <https://doi.org/10.1186/gb-2010-11-10-r106>.
 43. Lou Y, Rybníček J, Sala C, Cole ST. 2017. *EspC* forms a filamentous structure in the cell envelope of *Mycobacterium tuberculosis* and impacts ESX-1 secretion. *Mol Microbiol* 103:26–38. <https://doi.org/10.1111/mmi.13575>.
 44. Chen JM, Boy-Röttger S, Dhar N, Sweeney N, Buxton RS, Pojer F, Rosenkrands I, Cole ST. 2012. *EspD* is critical for the virulence-mediating ESX-1 secretion system in *Mycobacterium tuberculosis*. *J Bacteriol* 194:884–893. <https://doi.org/10.1128/JB.06417-11>.
 45. Kapopoulou A, Lew JM, Cole ST. 2011. The MycoBrowser portal: a comprehensive and manually annotated resource for mycobacterial genomes. *Tuberculosis* 91:8–13. <https://doi.org/10.1016/j.tube.2010.09.006>.

SUPPLEMENTARY MATERIAL

Figure S1

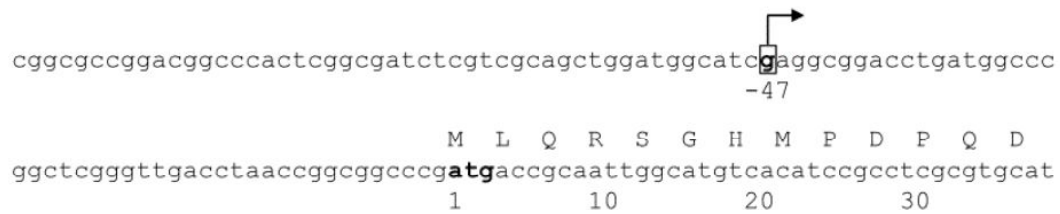


Figure S1 Sequence of the *rv3852* promoter region and beginning of the *rv3852* gene. The position of the transcriptional start site lies 47 bp upstream of the translational start site (atg in bold) and is marked with an arrow. Translation from the start of *rv3852* is indicated with one letter amino acid code.

Figure S2

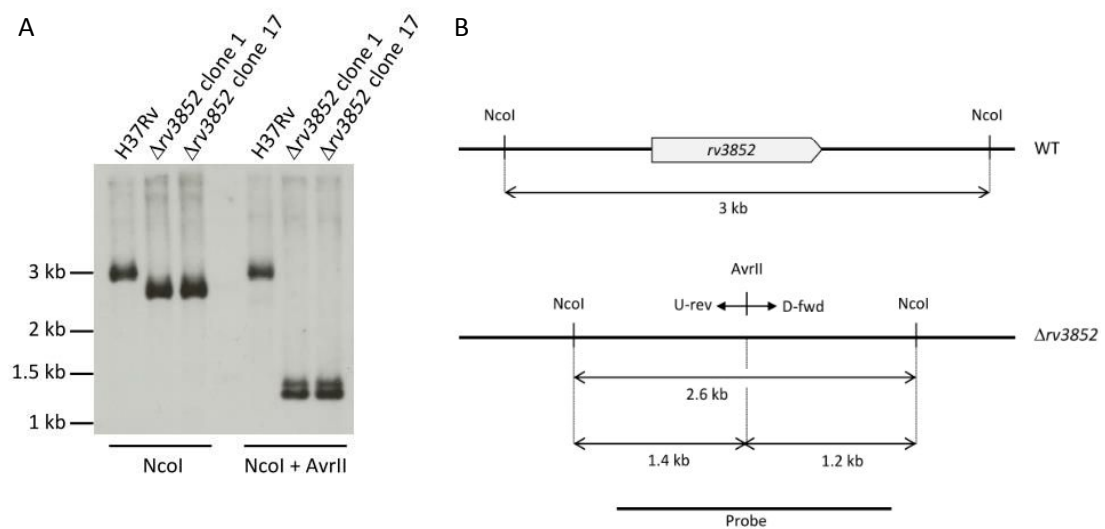


Figure S2 Validation of the Δ *rv3852* mutants. (A) Southern blot analysis of the Δ *rv3852* deletion strains. (B) Restriction profile of the wild type and Δ *rv3852* genomic DNA region surrounding *rv3852*. The *AvrII* restriction site was introduced by primers U-rev and D-fwd. The position of the probe used in Southern blot analysis is shown.

Figure S3

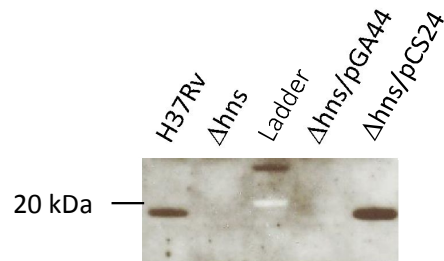


Figure S3 Validation of H-NS expression in the complemented mutant (pCS24). pGA44 is the empty vector.

Figure S4

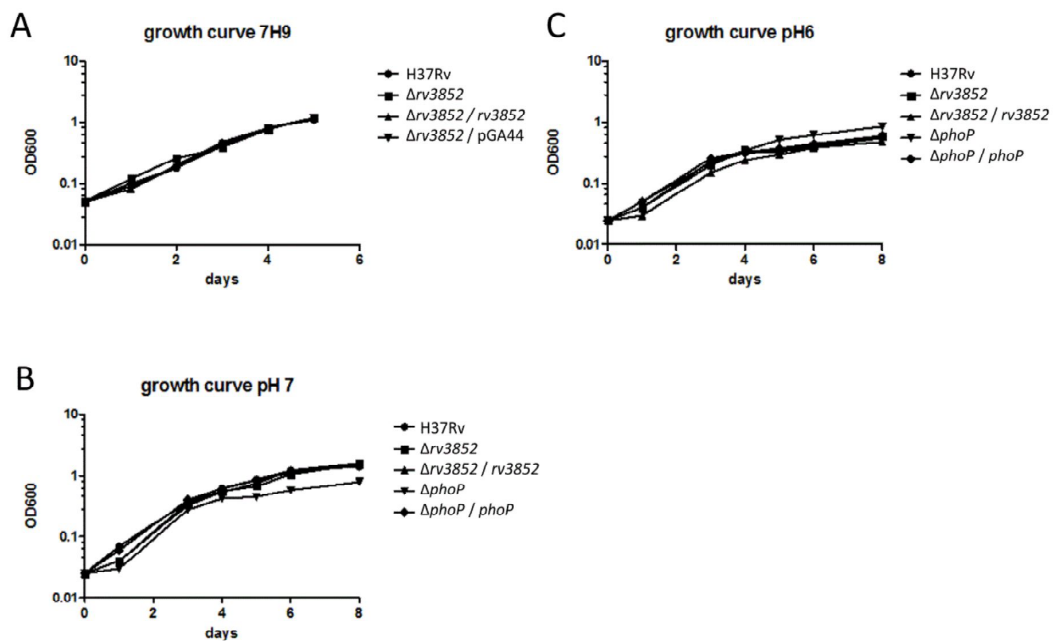


Figure S4 *In vitro* growth curves. (A) Growth curves of *M. tuberculosis* strains H37Rv, $\Delta rv3852$, $\Delta rv3852 / rv3852$ and $\Delta rv3852 / pGA44$ in 7H9 medium. (B) Growth curves of H37Rv, $\Delta rv3852$, $\Delta rv3852 / rv3852$ with a $\Delta phoP$ mutant and the complemented strain $\Delta phoP / phoP$ as controls in media buffered at pH = 7 and (C) pH = 6.

Figure S5

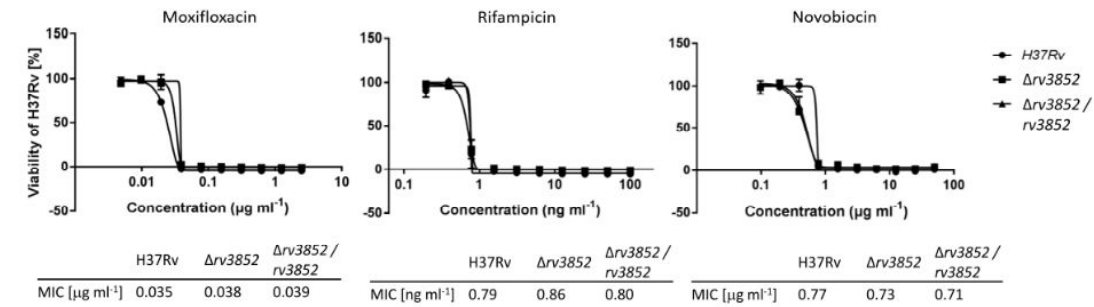
Figure S5 MIC determination for moxifloxacin, rifampicin and novobiocin in H37Rv, Δrv3852 and $\Delta\text{rv3852} / \text{rv3852}$.

Figure S6

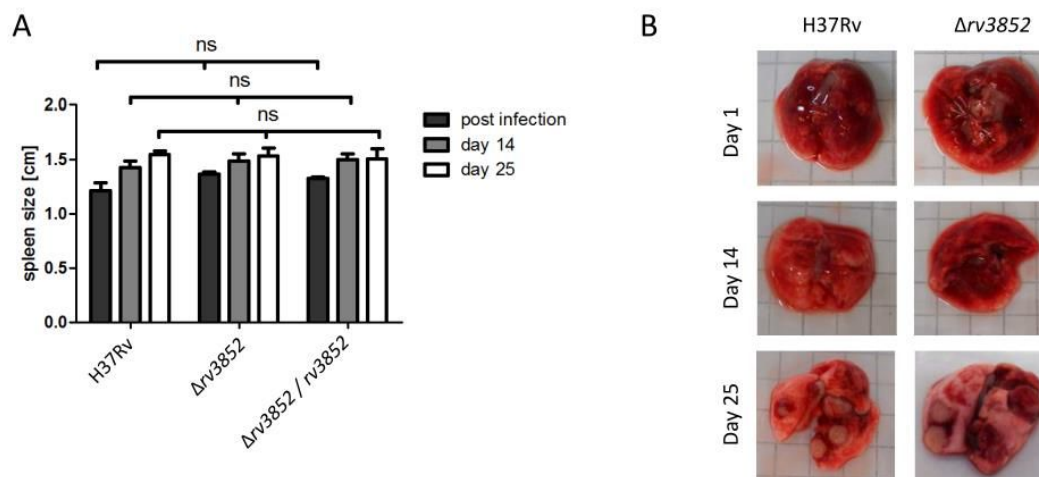


Figure S6 *In vivo* analysis of Δrv3852 deletion strain. (A) Spleen size. Columns correspond to the length of the spleen 1 (post infection), 14 (D14) and 25 (D25) days after infection with H37Rv, Δrv3852 and $\Delta\text{rv3852} / \Delta\text{rv3852}$. Bars represent the mean \pm SD from 3 (post infection) and 5 (day 14 and 25) SCID mice per group and time point. No difference was observed for the spleen size for Δrv3852 or $\Delta\text{rv3852} / \text{rv3852}$ relative to H37Rv for any of the time points (ns: not significant in Student's t-test). (B) Representative photographs of lungs of SCID mice infected with Δrv3852 or with H37Rv 1, 14 and 25 days after infection.

Table S1 Plasmids used in this study

Name	Description and main features	Reference
pJG1100	Suicide vector for mutant construction, Hyg ^R , Kan ^R , <i>sacB</i>	(1, 2)
pCS21	Suicide vector for mutant construction derived from pJG1100, Hyg ^R , Kan ^R , <i>sacB</i>	This study
pCS24	<i>rv3852</i> cloned in pGA44, integrative vector at L5 attB site, Str ^R	This study
pGA44	Integrative vector at L5 attB site, Str ^R	(3)
pGA80	pMV261 derivative, oriM has been removed, L5 <i>int</i> is expressed by the <i>hsp60</i> promoter, Kan ^R	(3)

1. Gomez JE, Bishai WR. 2000. whmD is an essential mycobacterial gene required for proper septation and cell division. *Proc Natl Acad Sci USA* 97:8554–8559.
2. Gagneux S, Burgos M V., DeRiemer K, Enciso A, Muñoz S, Hopewell PC, Small PM, Pym AS. 2006. Impact of bacterial genetics on the transmission of isoniazid-resistant *Mycobacterium tuberculosis*. *PLoS Pathog* 2:0603–0610.
3. Kolly GS, Boldrin F, Sala C, Dhar N, Hartkoorn RC, Ventura M, Serafini A, McKinney JD, Manganelli R, Cole ST. 2014. Assessing the essentiality of the decaprenyl-phospho-d-arabinofuranose pathway in *Mycobacterium tuberculosis* using conditional mutants. *Mol Microbiol* 92:194–211.

Table S2 Primers used in this study

Primer name	Sequence 5' – 3'
U-fwd	ACGTTCTTAATTAAGCGCTCAAGGCGCAGCTGAAGG
U-rev	ACGTCCTAGGCATGTGGCCGCTCCTTTGCAGC
D-fwd	ACGTCCTAGGTGAACGCGCTGGCACCATAGTG
D-rev	ACGTGGCGCGCCGACGGCGACTATCGACGGACAAC
<i>rv3852</i> -F	ACGTGTCCTAGGCCAGACCCGCAGGATCGACCC
<i>rv3852</i> -R	ACGTGTGGCGCGCCTCAGCGGCGGCGCAGTTG
CS-329-U1	CGCATCAAGGCCGCCAG
CS-330-D1	GGCCCGATGTCGTCGACGAGG
CS-331-U2	GATGGGACCGAGCGTTTGCGATCTAG
CS-332-D2	GGCTGACATCATCATTACCGCAACGGG
CS-351- <i>rv3852</i> F	CGATCTAGTGAGCGACGACA
CS-352- <i>rv3852</i> R	ATTGGCGGAAGTGTAGGTCA
CS-395- <i>rv3852</i> R1	CTTGCTGGTGCTTTCTTGCGCG
CS-396- <i>rv3852</i> R2	GCGGACTTAGCACCCCTGGCG
CS-397- <i>rv3852</i> R3	CGATCCGCTGCTGCAAACTGAC
CS-080-dTanchor	GACCACGCGTATCGATGTCGACTTTTTTTTTTTTTTTT
CS-081-anchor	GACCACGCGTATCGATGTCGAC
CS-395- <i>rv3852</i> R1	CTTGCTGGTGCTTTCTTGCGCG
CS-397- <i>rv3852</i> R3	CGATCCGCTGCTGCAAACTGAC
CS-057- <i>sigA</i> -F	AAACAGATCGGCAAGGTAGC
CS-058- <i>sigA</i> -R	CTGGATCAGGTCGAGAAACG
NO-128- <i>kasA</i> -F	GTGGGATCTAGCGGTCAAGA
NO-129- <i>kasA</i> -R	AGCTCTCGACAATCCTCTCG
NO-130- <i>kasB</i> -F	CGGACATCTGCTTGAGGAAT
NO-131- <i>kasB</i> -R	GCGCATATCGTCGTAAGTGA

Rv3852 – 2

CHAPTER 3.1

Essentiality of *mihF* and impact of gene silencing on *Mycobacterium tuberculosis* physiology and transcriptional landscape

Nina T. Odermatt, Claudia Sala, Andrej Benjak, Stewart T. Cole

École Polytechnique Fédérale de Lausanne, Global Health Institute, Station 19,
1015 Lausanne, Switzerland

2017. Manuscript in preparation

Contributions: design of experiments, mutant construction, phenotype assessment, light microscopy, transcriptome analysis, bioinformatics, data analysis, manuscript preparation

ABSTRACT

The airborne pathogen *Mycobacterium tuberculosis* experiences a range of stresses when invading or exiting human cells and during latent infection. Tight control of gene expression is thus crucial for tubercle bacilli to adapt to these rapidly changing and hostile environments. Nucleoid associated proteins (NAPs) alter chromosome topology and act as global transcriptional regulators thereby influencing gene expression. Here, we report the in-depth functional analysis of the mycobacterial integration host factor miHF, one of the four known NAPs in *M. tuberculosis*.

Construction of a conditional knockdown mutant confirmed the essentiality of *mihF* in *M. tuberculosis*. Depletion of the miHF protein resulted in elongated cells devoid of septa, containing only one nucleoid, and was bactericidal. Nucleic acid synthesis as well as protein production were abrogated after two and four days of miHF depletion, respectively. ChIP-sequencing identified 153 broad peaks distributed around the chromosome, but predominantly situated upstream of transcriptional start sites. No consensus motif could be found but miHF bound to AT-rich loci (42%) in *M. tuberculosis* that often harbour binding sites for EspR, another NAP. Total transcriptome analysis by RNA-sequencing upon miHF depletion revealed a pleiotropic effect with 150 up- and 59 downregulated genes, and a clear relationship between miHF binding and differential gene expression. Downregulated transcripts included the NAPs EspR and Lsr2, many tRNAs as well as components of virulence pathways. Most strikingly, the *espACD* virulence operon required for ESX-1 secretion and several other ESX secretion system genes were downregulated, together with genes for biofilm formation, while DNA-related pathways were upregulated. Overall, miHF was confirmed as a global transcriptional regulator, which, among others, impacts genes required for pathogenesis in humans.

INTRODUCTION

Bacterial gene expression is tightly controlled and influenced by environmental changes. In the case of *Mycobacterium tuberculosis*, which has to adapt to a hostile milieu upon phagocytosis by a macrophage, these stimuli include acidic pH, reactive oxygen species, nutrient limitation, fatty acid availability or presence of lytic enzymes (Manganelli, 1999). It is therefore crucial to carefully regulate gene expression for bacterial growth and survival. Traditional transcription factors have one or more target genes that they directly activate or repress, whereas nucleoid associated proteins (NAPs) act at the global level by shaping the architecture of the DNA. Indeed, they can bend, bridge and loop the DNA, therefore regulating the vast majority of the genes in a bacterium (Dillon, 2010). In *Escherichia coli*, twelve NAPs have been characterized so far. Their cellular abundance fluctuates during the different growth phases (Azam, 1999) and each NAP targets a specific set of genes (Browning, 2010). On the other hand, only four NAPs have been reported to date in the human pathogen *M. tuberculosis*. HupB, is essential for growth in macrophages and for iron acquisition (Pandey, 2014); Lsr2, is an *E. coli* H-NS-like protein that preferentially binds to AT-rich sequences (Gordon, 2008); EspR regulates secretion of the main virulence factors of *M. tuberculosis* (Blasco, 2012) by controlling expression of the *espACD* operon among others (Garces, 2010); and finally, the mycobacterial integration host factor miHF.

miHF was discovered as essential for mycobacterial phage L5 integration into the *M. smegmatis* genome (Lee, 1993) and was therefore named after the *E. coli* ortholog, although the two genes and the respective products do not share any sequence similarity. miHF is highly conserved among the *Mycobacterium* genus and even *M. leprae*, with its reduced genome, possesses a copy of *mihF* (Cole, 2001). The *mihF* (*rv1388*) gene in *M. tuberculosis* was initially predicted to be 573 bp-long and to encode a ~20 kDa protein (Cole, 1998). More recently, Mishra and colleagues suggested that, based on comparative genomics, miHF of *M. tuberculosis* is 80 amino

acids shorter than the proposed 190 residue-long protein (Mishra, 2013). They also showed that mIHF binds to linear and supercoiled DNA and enhances topoisomerase activity (Mishra, 2013). Surprisingly, mIHF was identified among the top ten most abundant proteins of *M. tuberculosis* (Schubert, 2015).

The *mihF* gene was predicted to be essential for *in vitro* growth with glycerol or cholesterol as carbon sources by Himar-1 based transposon mutagenesis (Griffin, 2011), thus suggesting an important regulatory role in *M. tuberculosis* metabolism. In this study, we investigated the biology of mIHF thoroughly by constructing a conditional knockdown (cKD) mutant and by showing that mIHF is indeed essential for growth. mIHF-depleted cells displayed an aberrant phenotype, and a marked growth defect before dying. The pleiotropic effect observed upon mIHF depletion demonstrated that this NAP has a broad impact on *M. tuberculosis* gene regulation and controls expression of housekeeping as well as virulence genes.

RESULTS

mIHF is an abundant cytosolic protein

While some NAPs are present throughout the entire growth cycle of a bacterium, others

peak at certain stages only. The mIHF protein was identified at ca. 12 kDa on immunoblots, and time-course analysis of protein levels showed that it is constantly present from exponential to stationary phase with little fluctuations (Fig. 1a). To investigate the localization of mIHF inside the bacterial cell, *M. tuberculosis* H37Rv cell extracts were fractionated for subsequent immunoblotting. RpoB, Rv3852 (Odermatt, 2017) and EsxB were used as positive controls for the cytosol, membrane and secreted fractions, respectively. The mIHF protein was detected in the cytosol only (Fig 1b).

Transcription start site identification and conditional knockdown mutant generation

To explore the regulatory function of mIHF, we planned to construct an *mihF* knockout mutant. As it was suggested that mIHF of *M. tuberculosis* is shorter than originally annotated (Mishra, 2013), the transcripts were analysed by rapid amplification of cDNA ends (5'-RACE). Three potential transcription start sites (TSS) were detected: 91 bp upstream (TSS1), 127 bp downstream (TSS2) and 167 bp (TSS3) downstream of the currently annotated translation start site (Fig. S1). In addition, a TANNT-10 motif, shared by most promoters of *M. tuberculosis* (Cortes, 2013), was identified upstream of TSS3. Two potential translation

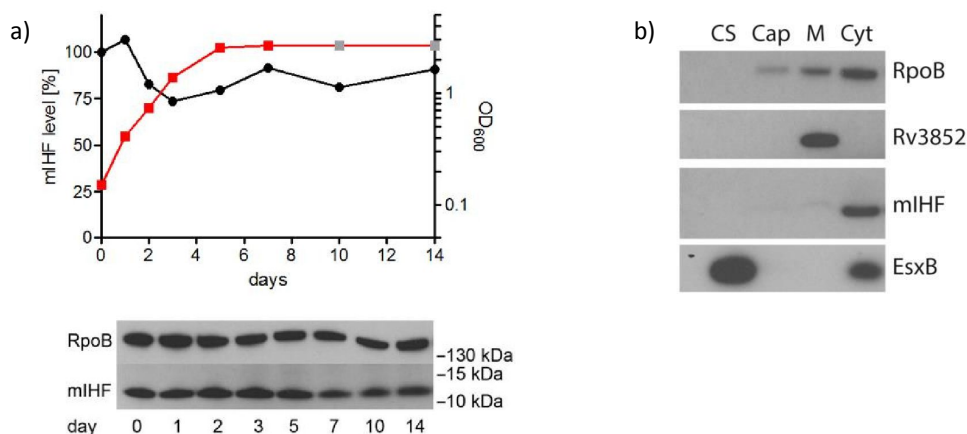


Fig. 1. Expression and localization of mIHF. a) Time-course analysis by immunoblot of mIHF levels (black bullets, left y-axis), and optical density at 600 nm (OD₆₀₀, red squares, right y-axis) of H37Rv from exponential to stationary phase. Protein levels were calculated by density analysis of the image below, relative to RpoB and to the first time point (day 0). OD₆₀₀ at days 10 and 14 (grey square) were set to the same value as day 7, as the culture formed aggregates typical of *M. tuberculosis* in stationary phase, which prohibited proper measurement of OD₆₀₀. b) Immunoblot of culture supernatant (CS), capsular (Cap), membrane (M) and cytosolic (Cyt) fractions of H37Rv. Antibodies used are indicated to the right.

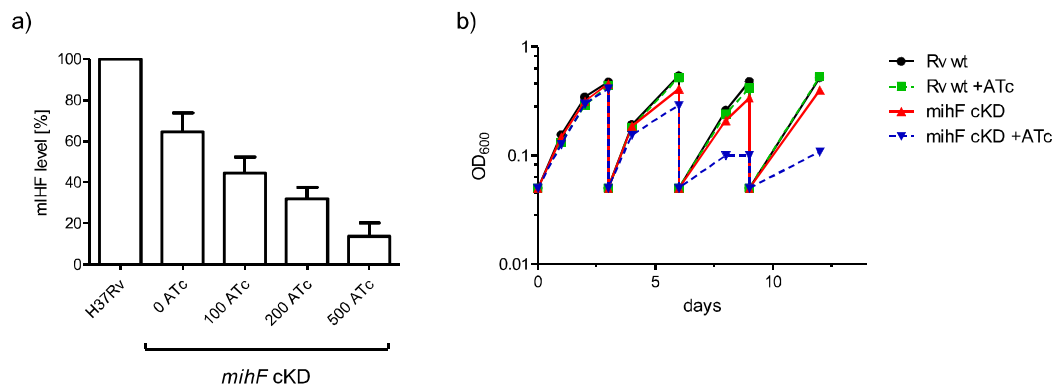


Fig. 2. *miHF* silencing by ATc. a) miHF levels detected by immunoblot normalized to RpoB and relative to H37Rv wildtype. ATc concentration is indicated in ng ml⁻¹. b) Growth curves of *miHF*-cKD and H37Rv wild type strains with 600 ng/ml ATc and without ATc. Cultures were diluted to OD₆₀₀ = 0.05 every 3 days.

start sites at M80 and V86 were detected downstream of TSS3 and the corresponding translational products were referred to as miHF-80 and miHF-86, respectively. A purine-rich heptamer (GGAGGAA) was recognized as the putative Shine-Dalgarno sequence centred at -6 bp upstream of *miHF*-86 (Fig. S1).

We employed gene replacement to remove the annotated, full-length *miHF* gene from the chromosome. After confirming the nature of the merodiploid strain, the second crossing over event, which led to the in-frame deletion of the gene, was only successful when a second copy of *miHF* was provided *in trans* under control of the TET-PIP OFF system (Boldrin, 2010). Furthermore, it was necessary to include the two downstream genes, *gmk* and *rpoZ*, on the complementing vector to obtain deletion of chromosomal *miHF* and therefore the conditional mutant *miHF*-cKD. The mutation

introduced was confirmed by Southern blot (Fig. S2).

Addition of anhydrotetracycline (ATc) to the *miHF*-cKD mutant complemented with the full-length *miHF* caused no variation in miHF protein abundance, nor in growth dynamics (data not shown). Only after replacing the complementing plasmid with pNO62 or with pNO63, carrying *miHF*-80 and *miHF*-86 respectively, was it possible to control the expression of *miHF* by ATc. As no difference was observed between these two versions of the *miHF*-cKD mutant in terms of *in vitro* growth (data not shown), we used the shorter form, *miHF*-86, for further experiments.

miHF is essential for growth and survival of *M. tuberculosis*

Addition of increasing concentrations of ATc to *miHF*-cKD allowed titration of miHF levels (Fig. 2a). Interestingly, the conditional mutant without ATc expressed lower levels of miHF than the wild type strain (approximately 30% less), indicating that the controllable *ptr* promoter is weaker than the natural one. This was reduced to less than 10% upon treatment with 500 ng ml⁻¹ ATc. Of note, it was necessary to dilute the *miHF*-cKD cultures at least twice to see a reduction of the growth rate of the mutant compared to the parental strain and to the uninduced *miHF*-cKD strain (Fig 2b). As expected, ATc had no impact on the growth of the H37Rv wild type strain.

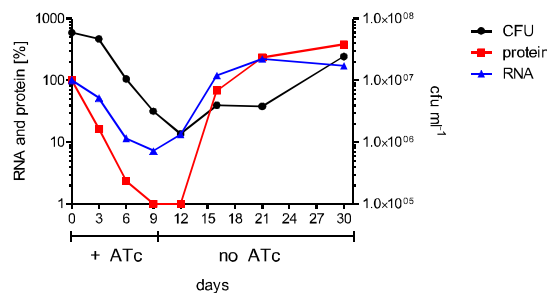


Fig. 3. *miHF* mRNA, miHF protein levels and colony forming units in the presence of ATc (day 0 – 9, diluted every three days) and under permissive conditions (from day 9).

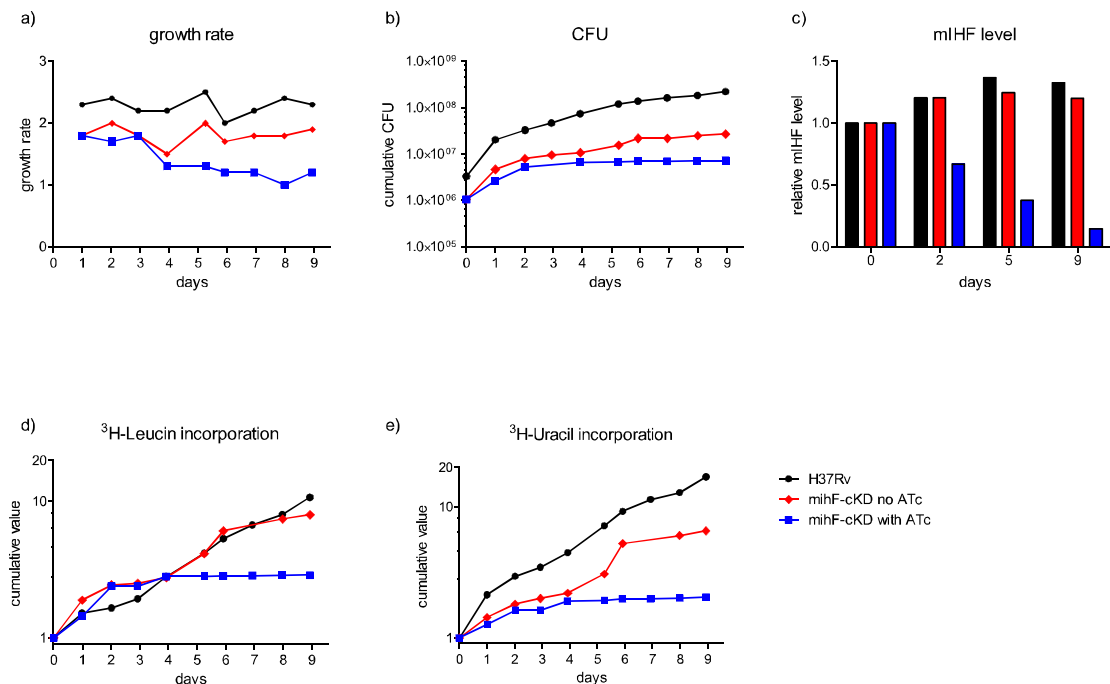


Fig. 4. Tritium labelled-leucine and -uracil incorporation. a) Growth rate of H37Rv wild type and *mihF*-cKD with and without ATc. b) Cumulative CFU ml⁻¹ of the three strains. c) mIHF levels at days 0, 2, 5 and 9 measured by immunoblot, relative to day 0.

Fig. 3 shows the phenotype of the *mihF*-cKD strain upon repression of *mihF* expression for 9 days, followed by three weeks under permissive conditions. While *mihF* RNA and mIHF protein levels rapidly decreased during the first three days, the number of colony forming units (CFU) dropped at day 6, proving that depletion of mIHF had a bactericidal effect. When ATc was removed from the bacterial culture, RNA production resumed, whereas a three-day lag was noticed for mIHF protein synthesis. Similarly, viable counts increased after a three-day lag under permissive conditions.

Impact of mIHF depletion on macromolecular synthesis

To examine *de novo* protein and nucleic acid (RNA plus DNA) production in mIHF-depleted cells, incorporation of ³H-labelled leucine and uracil by *mihF*-cKD plus ATc was monitored over nine days. Fig. 4a shows that the growth rate of the mutant strain without ATc was lower than that of H37Rv, and then further decreased in the presence of ATc. Likewise, the number of CFU increased from 5×10^6 to 2×10^8 CFU ml⁻¹ for

H37Rv after nine days, while *mihF*-cKD without ATc started at 1.05×10^6 and attained 2.68×10^7 CFU ml⁻¹ (Fig. 4b). In contrast, *mihF*-cKD in the presence of ATc reached a plateau after four days and only marginally increased to 7×10^6 CFU ml⁻¹ at day 9. While mIHF protein levels slightly augmented in H37Rv and in *mihF*-cKD without ATc as measured by immunoblot, *mihF*-cKD plus ATc showed lower mIHF levels at day 2 already and mIHF abundance dropped to ~40% at day five (Fig. 4c), when growth was arrested and nucleic acid as well as protein synthesis came to a halt (Fig. 4d and 4e). Depletion of mIHF had therefore a pleiotropic effect on *M. tuberculosis* physiology, as it severely compromised DNA, RNA and protein synthesis.

mIHF-depleted bacteria are longer than wild type and do not form septa

Scanning electron microscopy allowed investigation of the ultrastructure of *mihF*-cKD mutant cells. mIHF depleted cells were more than twice as long as the H37Rv parental strain and *mihF*-cKD mutant grown without ATc (Fig. 5a). The average length for the wild type strain

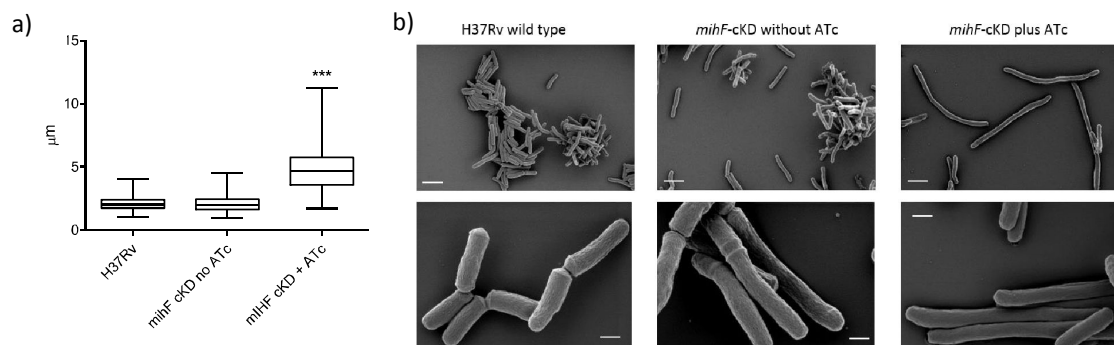


Fig. 5. Electron microscopy analysis. a) Cell length measured by surface electron microscopy. *miHF*-cKD cells grown with ATc for 9 days are twice as long as H37Rv wild type and *miHF*-cKD grown in the absence of ATc (ANOVA, *** = $p < 0.0001$). b) Scanning electron micrographs of H37Rv parental strain and *miHF*-cKD strains with and without ATc. Bar represents 2 μm on the upper panel and 200 nm on the lower panel.

was 2.13 μm, whilst *miHF*-cKD cells with ATc were 4.98 μm long. In addition, most *miHF*-cKD plus ATc cells longer than 3 μm did not show any ridge formation, which is indicative of the lack of an underlying septum (Dahl, 2004). Conversely, ridges were clearly visible in the wild type and in *miHF*-cKD cells without ATc (compare lower panel of H37Rv and *miHF*-cKD plus ATc in Fig. 5b). Apart from the elongated, septum-less phenotype, no other abnormal shape of was observed (i.e. no branching, swelling or bending) after depletion of miHF in the *miHF*-cKD mutant.

Fluorescence microscopy confirmed that most of the long miHF-depleted cells did not generate septa and additionally showed that

they contained only one nucleoid (Fig. 6). On the other hand, in those cells that contained a septum (wild type and *miHF*-cKD without ATc), DNA was present in each compartment (Fig. 6).

miHF has a broad impact on the *M. tuberculosis* transcriptional profile

To characterize the miHF-dependent regulon, global gene expression analysis was performed by RNA-seq. To rule out any effect of the inducer ATc on the transcriptome, we compared the transcription profile of H37Rv wild type strain in the presence and absence of ATc. Using a false discovery rate (FDR) below 1% and a fold-change cut-off above two, no differentially expressed genes were found, which confirmed that ATc itself did not affect gene expression (Supplementary Dataset 1). We further compared the untreated *miHF*-cKD strain versus H37Rv parental strain, to see if the integration of the complementing plasmid had any unusual impact on transcription. One gene (*serT*) was found to be 4-fold induced and seven were repressed up to 7-fold (*rv2463*, *rv0157*, *rv2383*, *rv1285*, *rv1288*, *rv1287*, *rv1196*), thus indicating that no major change in gene expression took place in the conditional mutant strain. Transcription of the *miHF*-cKD mutant strain in the presence and absence of ATc was then analysed.

Using the same cut-offs (FDR < 1% and a fold-change > 2), 679 downregulated and 464 upregulated genes were detected, thereby confirming that depletion of miHF had a vast impact on the global transcriptome

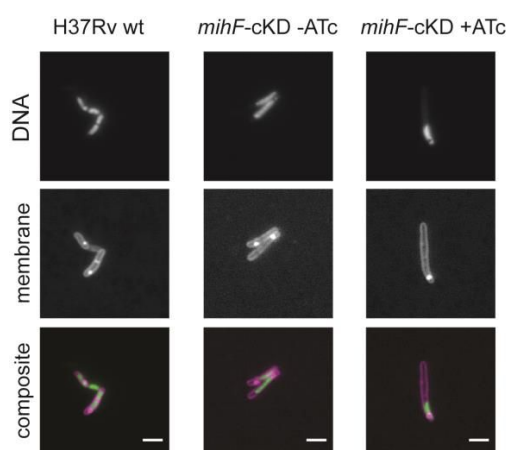


Fig. 6. Fluorescence microscopy images of H37Rv wild type, *miHF*-cKD with and without ATc. DNA was stained with SYTO9, membranes with FM4-64. Scale bars represent 2 μm.

(Supplementary Dataset 1). By increasing the minimal fold-change to 4, the number of deregulated genes was limited to 150 downregulated and 59 upregulated (Supplementary Dataset 1). Fig. 7 illustrates the distribution of the deregulated genes throughout the *M. tuberculosis* genome: no clustering was observed. The most interesting of these are listed in Table 1 and discussed in greater detail below.

Most striking was the most repressed operon, *espACD*, required for ESX-1 function and secretion of the main virulence factor EsxA in *M. tuberculosis*. Additionally, various genes that encode the type 7 secretion systems ESX-1, ESX-2 and ESX-5, including *esxA* and *esxB*, were expressed at a lower level. mIHF has therefore a major impact on expression of

virulence-related genes. The second most downregulated operon was *rv1168-rv1169* (13.6 and 18.3-fold lower expression, respectively), coding for PPE17 and LipX.

Most of the stable RNAs, notably 23 out of 45 tRNAs and several small regulatory RNAs, were more than 2-fold repressed, as well as many ribosomal protein genes. The gene for the replication initiator protein DnaA and for the beta chain of the DNA polymerase III, *DnaN*, were repressed > 4 times. These results indicate that mIHF is involved in controlling transcription of housekeeping genes.

Conversely, the most upregulated operon was *rv0196-rv0197*, coding for a possible transcriptional regulatory protein and a possible oxidoreductase, followed by two conserved hypothetical proteins and the *mce3*

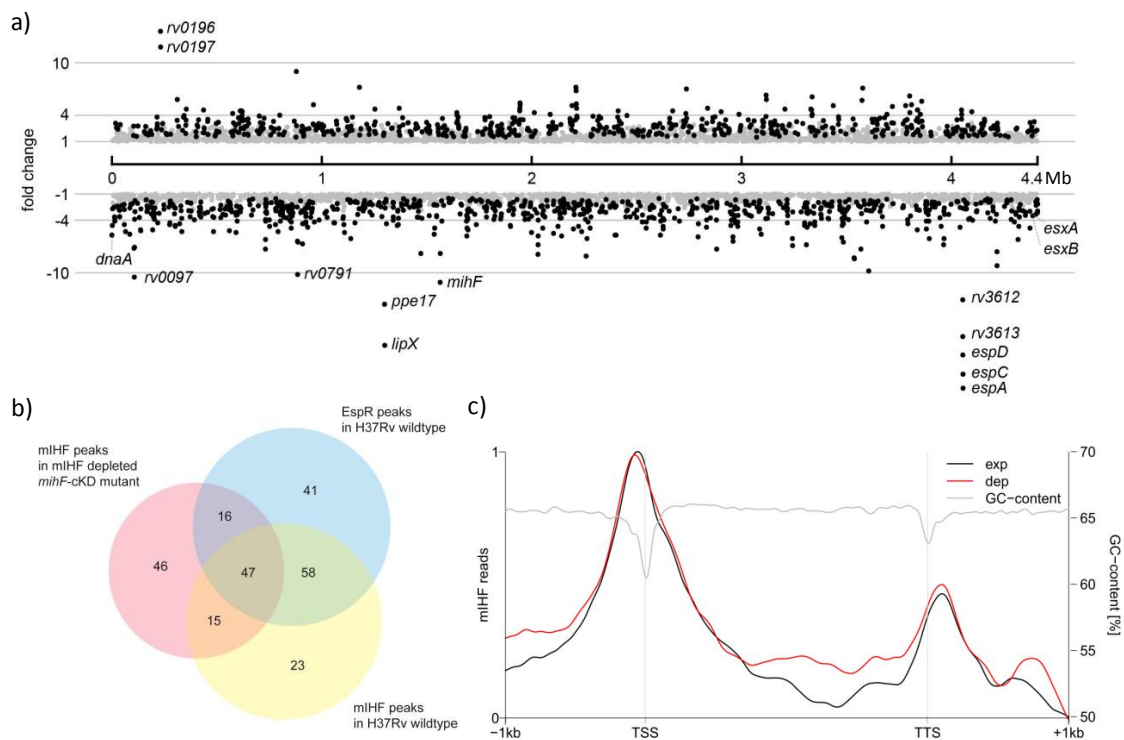


Fig. 7. RNA-seq and ChIP-seq analysis of mIHF binding sites. a) Differentially expressed genes between mIHF-depleted and non-depleted *M. tuberculosis* cultures plotted at their genomic position. Grey dots represent genes with a false discovery rate (FDR) higher than 1 %, black dots are significantly deregulated genes with FDR < 1. Interesting genes discussed further in the text are labelled. b) Overlap of peaks between EspR binding sites and mIHF in exponentially growing and mIHF depleted cells. c) Global profile of mIHF binding peaks relative to the transcriptional start site (TSS) and transcriptional termination site (TTS), range 0 to 1 on left y-axis indicates lowest to highest accumulation of mIHF reads, black profile for mIHF peaks in exponentially growing cells and red profile for mIHF peaks in mIHF-depleted cells. GC content of the H37Rv genome is plotted in grey (right y-axis).

Table 1. Selected top-scoring and differentially expressed genes in *miHF*-cKD upon ATc treatment.

	gene	description	FC	
stable RNAs				
MTB000049	<i>serT</i>	tRNA-Ser (GCT)	-9.2	
MTB000048	<i>argU</i>	tRNA-Arg (ACG)	-7.6	
MTB000007	<i>metT</i>	tRNA-Met (CAT)	-7.3	
replication and transcription				
Rv0001	<i>dnaA</i>	Chromosomal replication initiator protein DnaA	-5.68	
Rv0002	<i>dnaN</i>	DNA polymerase III (beta chain) DnaN (DNA nucleotidyltransferase)	-4.6	
Rv3711c	<i>dnaQ</i>	Probable DNA polymerase III (epsilon subunit) DnaQ	3.2	
Rv3370c	<i>dnaE2</i>	Probable DNA polymerase III (alpha chain) DnaE2	5.1	
Rv3202c	<i>Rv3202c</i>	Possible ATP-dependent DNA helicase	7.1	
virulence				
Rv3616c	<i>espA</i>	ESX-1 secretion-associated protein A, EspA	-23.2	
Rv3615c	<i>espC</i>	ESX-1 secretion-associated protein EspC	-21.6	
Rv3614c	<i>espD</i>	ESX-1 secretion-associated protein EspD	-19.4	
Rv3874	<i>esxB</i>	10 kDa culture filtrate antigen EsxB (LHP) (CFP10)	-3.4	
Rv0167	<i>yrbE1A</i>	Conserved integral membrane protein YrbE1A	-3.4	
Rv3875	<i>esxA</i>	6 kDa early secretory antigenic target EsxA (ESAT-6)	-3.1	
Rv0168	<i>yrbE1B</i>	Conserved integral membrane protein YrbE1B	-2.7	
Rv1963c	<i>mce3R</i>	Probable transcriptional repressor (probably TetR-family) Mce3R	1.4	
Rv1964	<i>yrbE3A</i>	Conserved hypothetical integral membrane protein YrbE3A	2.1	
Rv3660c	<i>Rv3660c</i>	Conserved hypothetical protein	2.6	
Rv1965	<i>yrbE3B</i>	Conserved hypothetical integral membrane protein YrbE3B	3.5	
Rv1971	<i>mce3F</i>	Mce-family protein Mce3F	3.9	
Rv1966	<i>mce3A</i>	Mce-family protein Mce3A	5.2	
Rv1969	<i>mce3D</i>	Mce-family protein Mce3D	5.3	
Rv1968	<i>mce3C</i>	Mce-family protein Mce3C	6.8	
Rv1967	<i>mce3B</i>	Mce-family protein Mce3B	7.2	
regulation				
Rv1388	<i>miHF</i>	Putative integration host factor MihF	-11.1	
Rv3597c	<i>lsr2</i>	Iron-regulated H-NS-like protein Lsr2	-4.8	
Rv1221	<i>sigE</i>	Alternative RNA polymerase sigma factor SigE	-4.5	
Rv3849	<i>espR</i>	ESX-1 transcriptional regulatory protein EspR	-3.6	
Rv0757	<i>phoP</i>	Possible TCS response transcriptional positive regulator PhoP	-2.7	
Rv2986c	<i>hupB</i>	DNA-binding protein HU homolog HupB	-1.1	
Rv0758	<i>phoR</i>	Possible TCS response sensor kinase membrane associated PhoR	1.0	
Rv0196	<i>Rv0196</i>	Possible transcriptional regulatory protein	13.6	
lipid metabolism				
Rv1169c	<i>lipX</i>	PE family protein. Possible lipase LipX.	-18.3	
Rv1168c	<i>PPE17</i>	PPE family protein PPE17	-13.6	
Rv0469	<i>umaA</i>	Possible mycolic acid synthase UmaA	-5.2	
Rv0166	<i>fadD5</i>	Probable fatty-acid-CoA ligase FadD5	-3.4	
others				
Rv3613c	<i>Rv3613c</i>	Hypothetical protein	-17.3	
Rv3612c	<i>Rv3612c</i>	Conserved hypothetical protein	-13.1	
Rv0197	<i>Rv0197</i>	Possible oxidoreductase	11.8	

FDR < 1%, fold change = FC, in the heat map red squares represent downregulated and blue squares upregulated genes.

operon. Some genes, which encode transcription factors, were found to be deregulated. For instance *lsr2*, *sigE* and *espR*, which may act downstream of mIHF, thus amplifying the regulatory signal. The *mihF* gene itself was more than 11-fold repressed, confirming the silencing effect of ATc.

Gene ontology (GO) analysis associates a function to each gene in the genome. Enrichment of a GO term indicates that a certain biological process, molecular function or cellular component is highly represented among the differentially expressed genes. In the case of ATc-treated *mihF*-cKD, the most downregulated genes showed a significant enrichment with $p < 0.05$ in several GO terms associated with host-pathogen interaction (Supplementary Dataset 4). Also, the type-7 secretion systems and DNA replication initiation are listed, as well as the cellular components “ribosome” and “extracellular region”. Molecular functions are present with transcription factors, DNA binding proteins, ribosomal constituents or rRNA binding. The GO categories related to upregulated genes were associated with DNA damage such as “DNA duplex unwinding”, “DNA repair” and “exonuclease activity”. Overall, RNA-seq studies revealed that mIHF affects transcription of a large subset of *M. tuberculosis* genes, thus contributing to explaining the essentiality demonstrated above.

mIHF is a nucleoid associated protein in *M. tuberculosis*

mIHF binding to the *M. tuberculosis* chromosome was investigated in exponentially growing H37Rv wild type cells by chromatin immunoprecipitation followed by high-throughput sequencing (ChIP-seq) and in the mIHF-depleted *mihF*-cKD mutant. Upon analysing the ChIP-seq results in exponentially growing wild type cells (Supplementary Dataset 1), we noticed that many mIHF peaks overlapped the previously published EspR binding sites (Blasco, 2012). For the sake of consistency, EspR ChIP-seq was then repeated on the same samples used for mIHF ChIP-seq. 162 EspR binding sites were detected, whereas 128 had been previously published (Blasco, 2012) and the others were found close by (Fig S3a for Pearson’s correlation coefficient, PCC).

mIHF bound to 153 loci in exponential phase (Supplementary Dataset 1), whereas 124 sites were contacted upon depletion of mIHF (Supplementary Dataset 1). The overlap between mIHF binding sites in exponentially growing cells and upon mIHF depletion was seen in 62 binding sites, 47 of which were shared with EspR as well (Fig. 7b). The overlap between the mIHF binding sites in exponential phase and after depletion had a PCC of 0.67, while the mIHF and EspR binding sites overlapped to a higher degree with a PCC of 0.6 in exponential phase and a PCC of 0.44 after mIHF depletion (Fig. S3).

While the GC content of the H37Rv genome is ~66%, the GC content of regions where EspR and/or mIHF bound ranged between 58% and 60%, were all significantly lower than the genome average ($p < 0.0001$, one-way ANOVA with Bonferroni’s multiple comparison test). This is consistent with the preferential location of mIHF binding near the transcriptional start site (TSS, Fig 8b). Indeed, mIHF-bound loci were mostly situated immediately upstream of the TSS and only a minor fraction was found after the transcriptional termination site (Fig. 7c). No difference in binding relative to gene position was observed before or after mIHF depletion. Moreover, mIHF binding sites were broad rather than sharp peaks in exponentially growing as well as in mIHF-depleted cells.

Enriched loci were distributed all over the genome, with some regions harbouring several peaks next to each other. Fig. 8 shows the binding profile for mIHF (black for exponential, red after depletion) with zooms of regions of interest. The genomic region around the top upregulated genes *rv0196-rv0197* was not bound by mIHF (Fig. 8e), while the long intergenic region upstream of the most downregulated gene *espA* showed peaks in both conditions (Fig. 8e).

Interestingly, despite its name as integration host factor, mIHF did not bind close to any of the two prophages PhiRv1 (spanning *rv1573-rv1587*) or PhiRv2 (*rv2650-rv2659*). On the other hand, genomic islands (GI) defined by Becq et al. (Becq, 2007), which represent 4.5% of the whole *M. tuberculosis* genome, were contacted by mIHF in both growth conditions. Almost one quarter (23%) of mIHF binding in

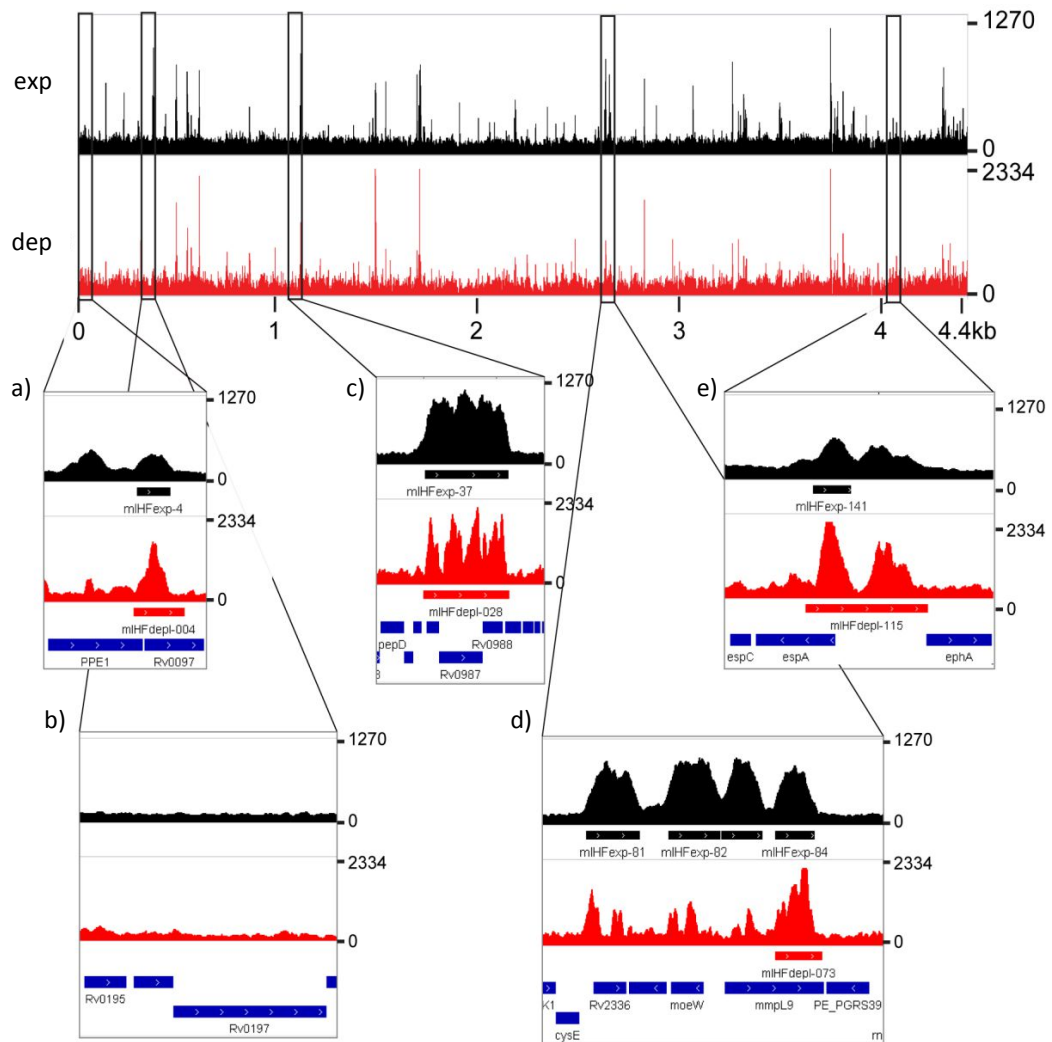


Fig. 8. miHF contacted regions on the H37Rv chromosome. The upper panel shows the global distribution of miHF binding in exponentially growing cells (black) and in miHF-depleted conditions (red), respectively, number of reads are indicated on the right. Zoomed images show examples of miHF binding. a) miHF peak at *rv0097*, 10.5-fold downregulated upon miHF depletion. b) No peak is located close to the top upregulated genes *rv0196* and *rv0197*. c) Third top-scoring peak cluster close to *mprA* / *mprB*. d) Accumulation of peaks around *mmpL9*. e) *espA* extended promoter region.

exponential phase, and 18% after depletion, was to a GI.

The impact of miHF as a potential direct activator or repressor was defined for features where miHF peaks either overlapped the genes or were within 500 bp of the coding sequence boundaries. We observed that genes associated with an miHF binding site were more frequently deregulated as compared to genes unlinked to an miHF-enriched locus (ANOVA, Bonferroni's multiple comparison test, $p < 0.0001$). Genes associated with miHF following depletion had a

mean expression level of -1.91, representing an almost 2-fold downregulation, compared to genes unaffected by miHF depletion with a mean expression of 0.05, indicating no transcriptional change. Similarly, genes in close proximity to miHF binding sites in exponential phase were expressed at -2.0, and more distant genes at 0.02. This indicates that miHF binding directly affects gene regulation.

DISCUSSION

NAPs have a major effect on bacterial gene expression by acting as global transcription regulators and chromosome architects. In *M. tuberculosis*, three NAPs (EspR, Lsr2 and HupB) were characterized in depth previously, while the proposed function of a fourth candidate NAP, Rv3852, was disproved (Odermatt, 2017). Here, we report the systematic investigation of mIHF and its potential importance for *M. tuberculosis* pathogenesis by direct control of genes involved in virulence, fatty acid metabolism and biofilm formation. Like EspR (Blasco, 2012), mIHF is solely located in the cytosol, where it associates with DNA.

A conditional mutant was constructed by deleting the entire annotated *mihF* open reading frame and replacing it with an *mihF* gene under the control of the tunable *ptr* promoter. The resultant ATc-dependent silencing of *mihF* expression confirmed the predicted essentiality and demonstrated that mIHF is required not only for multiplication but also for bacterial survival. Indeed, a dramatic effect on cell morphology and global physiology was noticed upon depletion of the protein. At least two passages were necessary to decrease mIHF levels, thus corroborating the reported abundance of mIHF in exponential phase and during hypoxia-induced dormancy (Schubert, 2015). Defects in cell elongation, cell shape and septum formation were observed, as well as decreased DNA replication.

The numerous differentially regulated genes identified upon mIHF depletion indicated a pleiotropic effect with a strong bias towards virulence-related genes. Most prominent was the 20-fold repression of one of the main virulence operons, *espACD*, critical for secretion of EsxA / EsxB and full virulence of *M. tuberculosis* (Chen, 2012), as well as the lower expression of multiple Esx protein genes (*esxA*, *esxB*, *esxC*, *esxD*, *esxM*, *esxN*) indicating that mIHF plays a central role in successful pathogenesis by affecting secretion of various virulence factors.

In the same vein, expression of LipX (*rv1169c*), which is necessary for *ex vivo* survival (Narayana, 2007), was repressed 18-fold. Downregulation of *lipX* in *M. tuberculosis*

reduces lipid levels, including PDIM (phthiocerol dimycocerosate), and leads to enhanced cellular aggregation and biofilm formation (Rastogi, 2017). PDIMs are required for phagosomal escape (Quigley, 2017) and thus another important virulence factor is controlled by mIHF. Biofilms are associated with persistent infections of *M. tuberculosis* (Richards, 2014) and contain mycolic acids in their extracellular matrix (Ojha, 2008). In addition to *lipX*, several genes encoding mycolic acid synthases (e.g. *umaA*, *mmaA3*) were expressed at lower levels in mIHF depleted cells, implying that mIHF also affects biofilm formation and mycolic acid synthesis, consistent with the results of our GO survey. Moreover, *ppe17*, lying downstream of *lipX*, was repressed 13.5-fold. PPE17 was shown to be upregulated in macrophages (Schnappinger, 2003; Donà, 2013), again indicating the importance of mIHF for successful infection.

Apart from control of virulence genes, several RNA, DNA and protein synthesis pathways were also found to be deregulated. For example, the gene for the epsilon subunit (*dnaQ*) of DNA polymerase III was 3.2-fold upregulated, while the *dnaN* was 4.6-fold, and *dnaA* was 5.7-fold repressed. DnaN acts as a bridge between the alpha and epsilon subunit and plays a regulatory role on the replicase (Gu, 2016). Absence of the DnaA (chromosomal replication initiator) and DnaN proteins should reduce initiation of DNA replication, and this is consistent with the lower nucleic acid synthesis observed by uracil-incorporation. We therefore assume that mIHF is critical for DNA replication initiation and maintenance, which is again supported by the GO analysis.

Most of the sigma factors were not differentially expressed upon mIHF depletion, except for *sigE*, which was more than fourfold downregulated. This alternative sigma factor is activated in various stress conditions, such as growth inside macrophages (Manganelli, 2014), and is required for interrupting phagosome maturation (Casonato, 2014). SigE positively regulates *mprAB* and induces the stringent response in *M. smegmatis* (Sureka, 2007). As a consequence of the downregulation of *sigE*, neither the stringent response nor the SOS response to DNA damage were induced upon mIHF depletion, which implies that mIHF does

not play a direct role in facing starvation. Concerning protein synthesis, the low amount of tRNAs and downregulation of ribosomal protein genes were reflected in the poor incorporation of radiolabelled leucine by mIHF-depleted cells. GO analysis confirmed the influence of mIHF in protein synthesis with cellular components associated with the ribosome as well as molecular functions linked to structural constituents of the ribosome.

Surprisingly, despite the observed phenotype in *mihF*-cKD, no known gene involved in cell division or septum formation was found to be heavily deregulated. While *ftsZ*, coding for the cell division initiator protein, and *wag31*, whose product localizes to the septum and is supposedly involved in lipid II biosynthesis (Kieser, 2014), were marginally downregulated, *ftsK* and *ftsQ*, necessary for septum formation (Slayden, 2006), were slightly upregulated.

Very good reproducibility was noticed between the published EspR ChIP-seq data (Blasco, 2012) and those generated here. The minor differences can be attributed to experimental variation and to the different software used for alignment and quality control. The mIHF binding profiles in exponentially growing cells as well as in non-permissive conditions were essentially the same but differed in peak width since after mIHF depletion peaks were generally narrower, probably due to the lower amount of mIHF. The strong, but not complete, correlation between mIHF peaks in exponential phase (153) and after depletion (124) suggested that mIHF probably binds to a core set of genes with higher affinity. The overall number of mIHF binding sites (205) was similar to that of EspR (165), but lower than the number of Lsr2 contact regions, which was approximately 800 (Gordon, 2010). mIHF may interact with EspR, as suggested by their binding to the same loci in growing cells. It is especially striking that Crp, EspR, Lsr2, mIHF, and MprA bind to the same region preceding *espACD*, thereby indicating its critical role in integrating different regulatory signals.

Overall, genes harbouring an mIHF binding site within their coding sequence or located less than 500 bp away showed downregulation

upon mIHF depletion. While mIHF peaks were still detected following depletion, these genes were not transcribed as efficiently as in permissive conditions. Hence, mIHF mainly acts as a direct transcriptional activator. The most repressed operon, *espACD*, and *rv0097* (9th most downregulated gene) both had a mIHF binding site in their upstream region. Conversely, *rv1168* – *rv1169*, the second strongest repressed operon, or *rv0791* (downregulated 10-fold), were not associated with an mIHF binding site. Similarly, only a few of the lowly expressed tRNA genes had an mIHF peak in close proximity, suggesting that apart from direct transcriptional control by binding close to TSS, mIHF can impact expression indirectly through a regulatory cascade or by mediating long-range control. The role of mIHF as the integration host factor remains elusive, as no binding to the two prophages PhiRv1 or PhiRv2 was found. However, it is possible that mIHF had such a role in the early evolution of *M. tuberculosis*, but, with the gradual loss of mobile elements and absence of horizontal gene transfer (Boritsch, 2016), it adapted its binding preferences. GI represent only ca. 4.5% of the genome, but contain a much higher fraction of the mIHF binding sites (23%), indicating a possible association of mIHF with horizontally acquired genes.

No common motif was found for the mIHF binding sites, confirming the non-specific binding of mIHF to DNA demonstrated *in vitro* by Mishra and co-workers (Mishra, 2013). The broad peaks of mIHF further indicate binding to a region rather than to a specific motif. As EspR does display sequence specificity, it might bind first to the DNA, then recruit mIHF to bind adjacent to EspR but this requires further investigation.

METHODS

Strains, media and chemicals

M. tuberculosis H37Rv and *mihF*-cKD strains were grown at 37°C either in Middlebrook 7H9 broth (Difco) supplemented with 10% albumin-dextrose-catalase, 0.2 % glycerol and 0.05% Tween 80 or in Sauton's liquid medium supplemented with 0.005% Tween 80. Cultures

were plated on Middlebrook 7H10 (Difco) agar supplemented with 10% oleic acid-albumin-dextrose-catalase and 0.2% glycerol. Hygromycin (50 $\mu\text{g ml}^{-1}$), kanamycin (25 $\mu\text{g ml}^{-1}$), streptomycin (25 $\mu\text{g ml}^{-1}$), 2.5 % sucrose or Anhydrotetracycline (ATc, Clontech, 600 ng ml^{-1}) were added when needed. For cloning procedures, One shot® TOP10 chemically competent *Escherichia coli* (Invitrogen) were grown in Luria–Bertani (LB) broth or on LB agar with hygromycin (200 $\mu\text{g ml}^{-1}$), kanamycin (50 $\mu\text{g ml}^{-1}$) or spectinomycin (25 $\mu\text{g ml}^{-1}$). All chemicals were purchased from Sigma-Aldrich, unless otherwise stated.

Plasmid and conditional knockdown mutant construction

1 kb up- and downstream regions of full length *miHF* were generated by PCR amplification using primers *miHF*-UF / *miHF*-UR and *miHF*-DF / *miHF*-DR (listed in Supplementary Table 1) respectively. Fragments were ligated in-frame with the *AvrII* site and cloned into the *PacI* and *Ascl* sites of pJG1100, resulting in the suicide vector pCS35. The complementing plasmid pCS31 was constructed by cloning the full-length *miHF* gene, amplified with primers *miHF*-F / *miHF*-R, into pGA44 under control of the repressible *ptr* promoter (Kolly, 2014). The two genes located downstream of *miHF* (*gmk* and *rpoZ*), were amplified with primers *gmk*-F / *rpoZ*-R. The constitutively active promoter *PfurA102* (Sala, 2003) was amplified with primers *PfurA102*-F / *PfurA102*-R. An overlap PCR was performed to fuse *PfurA102* with *gmk-rpoZ*, and then cloned into pCS31, resulting in pNO12. The complementing vectors harbouring *miHF*-80 and *miHF*-86, as well as *PfurA102-gmk-rpoZ*, were constructed in pGA118, a derivative of pGA44, which carries a hygromycin instead of a streptomycin resistance cassette, resulting in pNO62 and pNO63, respectively.

Deletion of the full-length *miHF* gene was obtained by homologous recombination using plasmid pCS35. After transformation of *M. tuberculosis* H37Rv, the first recombination event was selected on 7H10 plates, supplemented with hygromycin and kanamycin. Colonies were screened by colony PCR using CS-402 / CS-403 and CS-404 / CS-405 primer pairs. The merodiploid strain was generated by integration of plasmid pNO12 at the L5 attB site

and transformants were plated on 7H10 with hygromycin, kanamycin and streptomycin. Finally, deletion of the wild type gene by allelic exchange and generation of the *miHF* cKD strain was accomplished by plating the bacteria on 7H10 supplemented with streptomycin and 2.5% sucrose. The resulting colonies were tested by PCR with primers CS-403 / CS-415 for deletion of *miHF* from its native locus and confirmed by Southern blot.

Genomic DNA extraction and Southern Blot

Mycobacterial genomic DNA was extracted using standard protocols. To confirm successful allelic exchange of *miHF*, genomic DNA was digested with *AvrII* and *PvuII* restriction enzymes. DNA fragments were separated by 0.8 % agarose gel electrophoresis before capillary blotting onto a Hybond-N+ nylon membrane (GE Healthcare) and hybridization with a probe corresponding to the same upstream and downstream regions of *miHF* cloned into pJG1100. Hybridization was carried out using the ECL Direct Nucleic Acid Labelling and Detection System (GE Healthcare) as recommended by the manufacturer.

Growth curve measurements, colony forming unit counts, ^3H -leucine and ^3H -uracil incorporation

To characterize the growth of the *miHF* cKD mutant, the strains were grown to mid-logarithmic phase and then diluted to an optical density at 600 nm (OD_{600}) of 0.05 in 7H9 medium. ATc was added to 600 ng ml^{-1} and the OD_{600} was recorded at different time points to obtain the growth curves. As ATc is light sensitive and depleted over time, cultures were diluted to $\text{OD}_{600} = 0.05$ every three days.

Nucleic acid and protein synthesis was measured by incorporating tritium-labelled leucine and uracil as previously described (Wayne, 1977). As mycobacteria do not incorporate exogenous thymidine, but can use uracil for RNA as well as for DNA synthesis after methylation (Wayne, 1977), only uracil was used to assess total nucleic acid production. H37Rv wildtype with ATc, *miHF*-cKD without and in presence of ATc (600 ng ml^{-1}) were grown to mid-exponential phase, diluted to $\text{OD}_{600} = 0.1$ and then subsequently diluted again to the same $\text{OD} = 0.1$ every day. Of these cultures, 1

ml was incubated daily with 1 μCi ^3H -uracil and ^3H -leucin, respectively. After 24 hours, the sample was washed once in PBS supplemented with 0.05 % Tween 80, the cells were harvested by centrifugation and stored at -80°C until further processing. Counts per minute were measured by suspending the sample in 5 ml Ecoscint XR (National diagnostics) on a Beckman Coulter LS6500 Multi-Purpose Scintillation Counter. At every sampling point, colony-forming units (CFU) per millilitre of culture was evaluated and protein samples were taken at days 0, 2, 5 and 9. Cumulative protein and nucleic acid incorporation was derived as the sum of the daily incorporation multiplied by the CFU for the total incorporation relative to day 0, multiplied by the dilution factor, to account for the growth rate during the 24-hour incubation. Colony forming units were evaluated from serial dilutions of *M. tuberculosis* cultures plated on 7H10 plates and cumulative CFU was calculated similarly by summing the previous CFU with the daily CFU multiplied by the dilution factor necessary to reach OD = 0.1 to normalize all of the three samples. Growth rate was calculated as the daily OD₆₀₀ divided by the target OD₆₀₀ of 0.1.

Scanning electron microscopy

For surface scanning electron microscopy, *mihF*-cKD mutant without ATc and H37Rv wild type strains were grown in 7H9 until mid-exponential phase, pelleted, washed in PBS and resuspended to OD₆₀₀ = 0.5. *mihF*-cKD mutant with ATc was diluted three times in fresh ATc-containing medium and then subjected to the same protocol. Samples were then fixed on a coverslip in a solution of 1.25 % glutaraldehyde, 1 % tannic acid in phosphate buffer (0.1 M, pH = 7.4) for 1 h, washed in PBS prior to fixing for 30 min in 1 % osmium tetroxide. The samples were then dehydrated in a graded alcohol series and dried by passing them through the supercritical point of carbon dioxide (Leica Microsystems CPD300), and coated with a 2nm layer of osmium metal using an osmium plasma coater (Filgen OPC60). Scanning electron microscopy images were taken using field emission scanning electron microscope (Merlin, Zeiss NTS) using an acceleration voltage of 2kV and the in-lens secondary electron detector.

Fluorescence microscopy

H37Rv and *mihF*-cKD without ATc were grown to exponential phase in 7H9 Middlebrook media, while *mihF*-cKD plus ATc was diluted three times in fresh medium with 600 ng ml⁻¹ ATc. Samples were incubated with SYTO9 (4 μM), which stains the DNA, and FM4-64 (5 μg ml⁻¹) to stain the membrane for 20 minutes at 37°. Bacteria were mounted on an agarose pad and imaged with an Olympus IX81 microscope under a 100x objective. Representative images were selected and single channel and composite images were adjusted for brightness and contrast in ImageJ.

Total RNA extraction and 5' rapid amplification of cDNA ends

M. tuberculosis H37Rv and *mihF*-cKD cultures were harvested by centrifugation, pellets were resuspended in TRIzol Reagent (ThermoFisher) and stored at -80°C until further processing. Total RNA was extracted by bead-beating as previously described (Jungwirth, 2012). Integrity of RNA was checked by agarose gel electrophoresis, purity and amount of RNA were assessed using a Nanodrop instrument and Qubit Fluorometric Quantitation (ThermoFisher) respectively. SuperScript III First-Strand Synthesis System (Invitrogen) was used to generate randomly primed cDNA from 500 ng of RNA, according to the manufacturer's recommendations.

Primers CS-057 / CS-058 for *sigA* were used to normalize the amount of cDNA template added to each sample.

For the 5'-RACE two micrograms of *M. tuberculosis* H37Rv RNA and 1 μg of primer NO-095 were incubated at 70°C for 5 min and then at 55°C for one hour in the presence of 1x cDNA synthesis buffer, 1 mM dNTPs, 40 U RNase inhibitor, 25 U Transcriptor Reverse Transcriptase (5'/3' RACE Kit, 2nd Generation, Roche). cDNA was then purified with the High Pure PCR Product Purification kit (Roche), and used in the subsequent poly(A) tailing reaction (30 min at 37°C in the presence of 0.2 mM dATP and 80 U Terminal Transferase, Roche). Semi-nested PCR amplification on poly(A)-tailed cDNA was performed using an oligo dT-anchor primer (CS-080) and primer NO-094. Three amplification products were obtained, cloned into pTOPO (Invitrogen) and sequenced.

RNA sequencing and analysis

RNA was extracted from biological duplicate samples from exponential phase H37Rv and three-times diluted *mihF*-cKD as described above. The ribosomal RNA was depleted with the Ribo-Zero rRNA Removal Kit for Gram-positive Bacteria (Illumina), following the manufacturer's instructions. Libraries were prepared by the Lausanne Genomic Technologies Facility, using the Truseq Stranded mRNA Library Prep kit reagents (Illumina) according to the manufacturer's recommendations. The multiplexed libraries were sequenced on a HiSeq 2500 instrument using TruSeq SBS Kit V4 reagents as single-end 100nt-long reads. Sequencing data were processed using the Illumina Pipeline Software version 1.84. Reads were adapter- and quality-trimmed with Trimmomatic v0.33 (Bolger, 2014). The quality settings were "SLIDINGWINDOW:5:15 MINLEN:40". Reads were aligned with Bowtie2 (Langmead, 2012). Counting reads over annotated features was done with featureCounts (Liao, 2014). Annotation was taken from TubercuList release R27 (<http://tuberculist.epfl.ch/>). Differential gene expression analysis was done using DESeq2 (Love, 2014). "Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2." Genome Biology, 15, pp. 550. doi: 10.1186/s13059-014-0550-8.). Raw and processed data will be deposited in the GEO database (<https://www.ncbi.nlm.nih.gov/geo/>, submission in process).

For gene ontology (GO) enrichment analysis, a cut-off of 4-fold differentially expressed genes was chosen. The GO annotation was retrieved from BioCyc (Caspi, 2016) and analysis was performed with TopGO (Alexa, 2016). The conservative weighted algorithm and Fisher's exact test were used to calculate p-values for enrichment of GO terms in biological processes, molecular functions and cellular components in up- and downregulated genes upon mIHF depletion. The tree was pruned to a node size of 5 to exclude statistical artefacts of small sized GO terms.

Protein extraction, immunoblot analysis and subcellular fractionation

M. tuberculosis cultures grown in 7H9 were pelleted at different time points by

centrifugation, washed once in Tris-Buffered Saline (TBS, 20 mM Tris-HCl pH 7.5, 150 mM NaCl) and stored at -80°C until further processing. Cells were sonicated in TBS supplemented with a protease inhibitor tablet (cOmplete, mini, EDTA free, Roche) for 15 minutes and the protein solution was then sterilized by filtration through a 0.2 µm filter to remove any residual intact cells. Protein samples were quantified using the Qubit Fluorometric Quantitation device (ThermoFisher). Equal amounts of protein preparations were loaded on SDS-PAGE 12–15% NuPAGE gels (Invitrogen) and transferred onto PVDF membranes using a semidry electrophoresis transfer apparatus (Invitrogen). Membranes were incubated in TBS-Tween blocking buffer (25 mM Tris pH 7.5, 150 mM NaCl, 0.05 % Tween 20) with 5 % w/v skimmed milk powder for 3 hours at 4°C prior to overnight incubation with primary antibody. Membranes were washed in TBS-Tween three times, and then incubated with secondary antibody for 2 hours before washing. Signals were detected using Chemiluminescent Peroxidase Substrate 1 (Sigma-Aldrich).

Primary anti-mIHF antibody was produced by Alere against recombinant mIHF-86 and used at a concentration of 1:2,000 in immunoblots. Horseradish peroxidase (HRP) conjugated Goat anti-mouse Kappa (SouthernBiotech) secondary antibody was used at a 1:12,000 dilution. Anti-RpoB antibodies (NeoClone) were used to detect RpoB, the internal loading control. Band intensity of immunoblots was analysed with Fiji / ImageJ and normalized to the intensity of the RpoB signal.

Cell fractions were obtained as described previously (Lou, 2016). Briefly, H37Rv was grown in Sauton's medium with 0.005 % Tween 80 to mid-exponential phase, cells were collected by centrifugation, and supernatant was filtered and concentrated 100 x to obtain the secreted fraction. The pellet was treated with 0.25 % Genapol-X080 for 30 min followed by centrifugation at 14,000 g for 10 minutes and the proteins of the resulting supernatant precipitated with TCA, yielding the capsular fraction. The remaining pellet was subjected to sonication to break the cells, sterilized by filtration through a 0.2 µm filter followed by ultra-centrifugation at 45,000 rpm for one hour.

The supernatant contained the cytosolic fraction, while the pellet was enriched with membrane proteins.

Chromatin Immuno Precipitation (ChIP) and library construction

ChIP was performed as previously described (Hartkoorn, 2012). Briefly, exponentially growing *M. tuberculosis* H37Rv liquid cultures (for input control, EspR ChIP and MexpBS ChIP) or mIHF depleted *mihF*-cKD mutant (MdepBS) were cross-linked with formalin 1 % for 10' and quenched with glycine 125 mM for 10', washed twice in Tris-buffered saline (TBS, pH 7.5) and sonicated on a Diagenode Bioruptor with 30'' on/off cycles for 10' on high settings to shear DNA to 200 – 500 bp fragments. Immunoprecipitation was performed with monoclonal anti-mIHF (Alere) or polyclonal rat anti-EspR antibodies (Statens Serum Institut, Copenhagen, Denmark) in 1 ml immunoprecipitation (IP) buffer (containing 50 mM HEPES-KOH pH 7.5, 150 mM NaCl, 1 mM ethylenediaminetetraacetic acid (EDTA), 1 % Triton X-100, 0.1 % (w/v) sodium deoxycholate, 0.1 % sodium dodecyl sulphate (SDS) and one protease inhibitor cocktail tablet (Roche)) overnight at 4°C. 100 µl Dynabeads sheep anti-rag IgG (DynaL Biotech) for EspR and 100 µl per sample anti-ProteinL magnetic beads (Pierce) for mIHF were pre-saturated with 1mg ml⁻¹ bovine serum albumin and 0.1 mg ml⁻¹ salmon sperm DNA, then incubated with each corresponding sample for four hours. The IP was washed 5 times with increasing stringency buffers as described previously (Blasco, 2012). Input control was not incubated with any antibody.

Libraries were prepared with the NEBNext Ultra II DNA kit (NEB) by the Lausanne Genomic Technologies Facility following the manufacturer's instructions, multiplexed and sequenced on a HiSeq 2500 instrument.

ChIP-seq data analysis

Alignment was performed with Bowtie2 (Langmead, 2012) against the H37Rv genome (NCBI NC_000962.2) and resulted in 8.1 M uniquely aligned reads for EBS, 1.1 M for MdepBS, 7.3 M for MexpBS and 3.8 M for the input, respectively. For MexpBS, two different concentrations of antibodies were tested, and these two datasets were pooled, as the Spearman correlation of 0.96 was excellent (Fig Sx). HOMER (Heinz, 2010) was used for peak calling using the dynamic peak size algorithm with the input as a control. The enrichment is calculated relative to the input sequence, which might not reflect the real peak size. Resulting peaks were manually curated, subsequently annotated and further analysed with BEDtools (Quinlan, 2014) and deepTools (Ramirez, 2016). Profile plots of MdepBS and MexpBS, as well as GC content of H37Rv were generated with deepTools at binsize = 25 and visualized with IGV (Robinson, 2011). Coordinates of transcriptional start sites were taken from Cortes et al. (Cortes, 2013).

Files

Supplementary Dataset 1 contains RNA-seq H37Rv vs H37Rv+ATc data, RNA-seq *mihF*-cKD vs *mihF*-cKD+ATc data, EspR binding sites, mIHF exponentially growing cells binding sites, mIHF depleted binding sites, GO analysis

ACKNOWLEDGEMENTS

We thank the Lausanne Genomic Technologies Facility at the University of Lausanne, the BIOEM platform at EPFL for technical assistance and Joe Buechler from Alere for production of antibodies. This work was supported by the Swiss National Science Foundation (grant 31003A-162641 to STC).

REFERENCES

- Alexa, A. and Rahnenfuhrer, J. (2016) topGO: Enrichment Analysis for Gene Ontology.
- Azam, T.A., Iwata, A., Nishimura, A., Azam, T.A.L.I., and Ueda, S. (1999) Growth Phase-Dependent Variation in Protein Composition of the *Escherichia coli* Nucleoid Growth Phase-Dependent Variation in Protein Composition of the *Escherichia coli* Nucleoid. *J. Bacteriol.* **181**: 6361–6370.
- Becq, J., Gutierrez, M.C., Rosas-Magallanes, V., Rauzier, J., Gicquel, B., Neyrolles, O., and Deschavanne, P. (2007) Contribution of horizontally acquired genomic islands to the evolution of the tubercle bacilli. *Mol. Biol. Evol.* **24**: 1861–71.
- Blasco, B., Chen, J.M., Hartkoorn, R., Sala, C., Uplekar, S., Rougemont, J., et al. (2012) Virulence regulator EspR of *Mycobacterium tuberculosis* is a nucleoid-associated protein. *PLoS Pathog.* **8**: e1002621.
- Boldrin, F., Casonato, S., Dainese, E., Sala, C., Dhar, N., Palù, G., et al. (2010) Development of a repressible mycobacterial promoter system based on two transcriptional repressors. *Nucleic Acids Res.* **38**: e134.
- Bolger, A.M., Lohse, M., and Usadel, B. (2014) Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **30**: 2114–2120.
- Boritsch, E.C., Khanna, V., Pawlik, A., Honoré, N., Navas, V.H., Ma, L., et al. (2016) Key experimental evidence of chromosomal DNA transfer among selected tuberculosis-causing mycobacteria. *Proc. Natl. Acad. Sci.* **201604921**.
- Browning, D.F., Grainger, D.C., and Busby, S.J. (2010) Effects of nucleoid-associated proteins on bacterial chromosome structure and gene expression. *Curr. Opin. Microbiol.* **13**: 773–80.
- Casonato, S., Provvedi, R., Dainese, E., Palù, G., and Manganelli, R. (2014) *Mycobacterium tuberculosis* Requires the ECF Sigma Factor SigE to Arrest Phagosome Maturation. *PLoS One* **9**: e108893.
- Caspi, R., Billington, R., Ferrer, L., Foerster, H., Fulcher, C.A., Keseler, I.M., et al. (2016) The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. *Nucleic Acids Res.* **44**: D471–D480.
- Chen, J.M., Boy-Röttger, S., Dhar, N., Sweeney, N., Buxton, R.S., Pojer, F., et al. (2012) EspD is critical for the virulence-mediating ESX-1 secretion system in *Mycobacterium tuberculosis*. *J. Bacteriol.* **194**: 884–93.
- Cole, S.T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., et al. (1998) Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* **396**: 651–653.
- Cole, S.T., Eiglmeier, K., Parkhill, J., James, K.D., Thomson, N.R., Wheeler, P.R., et al. (2001) Massive gene decay in the leprosy bacillus. *Nature* **409**: 1007–1011.
- Cortes, T., Schubert, O.T., Rose, G., Arnvig, K.B., Comas, I., Aebbersold, R., and Young, D.B. (2013) Genome-wide Mapping of Transcriptional Start Sites Defines an Extensive Leaderless Transcriptome in *Mycobacterium tuberculosis*. *Cell Rep.* **1121–1131**.
- Dahl, J.L. (2004) Electron microscopy analysis of *Mycobacterium tuberculosis* cell division. *FEMS Microbiol. Lett.* **240**: 15–20.
- Dillon, S.C. and Dorman, C.J. (2010) Bacterial nucleoid-associated proteins, nucleoid structure and gene expression. *Nat. Rev. Microbiol.* **8**: 185–95.
- Donà, V., Ventura, M., Sali, M., Cascioferro, A., Provvedi, R., Palù, G., et al. (2013) The PPE Domain of PPE17 Is Responsible for Its Surface Localization and Can Be Used to Express Heterologous Proteins on the Mycobacterial Surface. *PLoS One* **8**: 1–8.
- Garces, A., Atmakuri, K., Chase, M.R., Woodworth, J.S., Krastins, B., Rothchild, A.C., et al. (2010) EspA acts as a critical mediator of ESX1-dependent virulence in *Mycobacterium tuberculosis* by affecting bacterial cell wall integrity. *PLoS Pathog.* **6**: 15–16.
- Gomez, J.E. and Bishai, W.R. (2000) whmD is an essential mycobacterial gene required for proper septation and cell division. *Proc Natl Acad Sci USA* **97**: 8554–8559.
- Gordon, B.R.G., Imperial, R., Wang, L., Navarre, W.W., and Liu, J. (2008) Lsr2 of *Mycobacterium* represents a novel class of H-NS-like proteins. *J. Bacteriol.* **190**: 7052–9.
- Gordon, B.R.G., Li, Y., Wang, L., Sintsova, A., van Bakel, H., Tian, S., et al. (2010) Lsr2 is a nucleoid-associated protein that targets AT-rich sequences and virulence genes in *Mycobacterium tuberculosis*. *Proc. Natl. Acad. Sci. U. S. A.* **107**: 5154–9.
- Griffin, J.E., Gawronski, J.D., Dejesus, M. a, Ioerger, T.R., Akerley, B.J., and Sassetti, C.M. (2011) High-resolution phenotypic profiling defines genes essential for mycobacterial growth and

- cholesterol catabolism. *PLoS Pathog.* **7**: e1002251.
- Gu, S., Li, W., Zhang, H., Fleming, J., Yang, W., Wang, S., et al. (2016) The $\beta 2$ clamp in the *Mycobacterium tuberculosis* DNA polymerase III $\alpha\beta\epsilon$ replicase promotes polymerization and reduces exonuclease activity. *Sci. Rep.* **6**: 18418.
- Hartkoorn, R.C., Sala, C., Uplekar, S., Busso, P., Rougemont, J., and Cole, S.T. (2012) Genome-wide definition of the SigF regulon in *Mycobacterium tuberculosis*. *J. Bacteriol.* **194**: 2001–9.
- Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., et al. (2010) Simple Combinations of Lineage-Determining Transcription Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell Identities. *Mol. Cell* **38**: 576–589.
- Jungwirth, B., Sala, C., Kohl, T.A., Uplekar, S., Baumbach, J., Cole, S.T., et al. (2012) High-resolution detection of DNA binding sites of the global transcriptional regulator GlxR in *Corynebacterium glutamicum*. *Microbiology* **159**: 12–22.
- Kieser, K.J. and Rubin, E.J. (2014) How sisters grow apart: mycobacterial growth and division. *Nat. Rev. Microbiol.* **12**: 550–562.
- Kolly, G.S., Boldrin, F., Sala, C., Dhar, N., Hartkoorn, R.C., Ventura, M., et al. (2014) Assessing the essentiality of the decaprenyl-phospho-d-arabinofuranose pathway in *Mycobacterium tuberculosis* using conditional mutants. *Mol. Microbiol.* **92**: 194–211.
- Langmead, B. and Salzberg, S.L. (2012) Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**: 357–359.
- Lee, M.H. and Hatfull, G.F. (1993) Mycobacteriophage L5 Integrase-Mediated Site-Specific Integration In Vitro. *J. Bacteriol.* **175**: 6836–6841.
- Liao, Y., Smyth, G.K., and Shi, W. (2014) FeatureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**: 923–930.
- Lou, Y., Rybníček, J., Sala, C., and Cole, S.T. (2016) EspC forms a filamentous structure in the cell envelope of *Mycobacterium tuberculosis* and impacts ESX-1 secretion. *Mol. Microbiol.* **103**: 1–39.
- Love, M.I., Huber, W., and Anders, S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**: 550.
- Manganelli, R. (2014) Sigma Factors : Key Molecules in *Mycobacterium tuberculosis* Physiology and Virulence. *Microbiol. Spectr.* 1–23.
- Manganelli, R., Dubnau, E., Tyagi, S., Kramer, F.R., and Smith, I. (1999) Differential expression of 10 sigma factor genes in *Mycobacterium tuberculosis*. *Mol. Microbiol.* **31**: 715–724.
- Mishra, A., Vij, M., Kumar, D., Taneja, V., Mondal, A.K., Bothra, A., et al. (2013) Integration Host Factor of *Mycobacterium tuberculosis*, mIHF, Compacts DNA by a Bending Mechanism. *PLoS One* **8**: e69985.
- Narayana, Y., Joshi, B., Katoch, V.M., Mishra, K.C., and Balaji, K.N. (2007) Differential B-cell responses are induced by *Mycobacterium tuberculosis* PE antigens Rv1169c, Rv0978c, and Rv1818c. *Clin. Vaccine Immunol.* **14**: 1334–1341.
- Odermatt, N.T., Sala, C., Benjak, A., Kolly, G.S., Vocat, A., Lupien, A., and Cole, S.T. (2017) Rv3852 (H-NS) of *Mycobacterium tuberculosis* Is Not Involved in Nucleoid Compaction and Virulence Regulation. *J. Bact.* **199**: 1–12.
- Ojha, A.K., Baughn, A.D., Sambandan, D., Hsu, T., Trivelli, X., Guerardel, Y., et al. (2008) Growth of *Mycobacterium tuberculosis* biofilms containing free mycolic acids and harbouring drug-tolerant bacteria. *Mol. Microbiol.* **69**: 164–174.
- Pandey, S.D., Choudhury, M., Yousuf, S., Wheeler, P.R., Gordon, S. V., Ranjan, A., and Sritharan, M. (2014) Iron-Regulated Protein HupB of *Mycobacterium tuberculosis* Positively Regulates Siderophore Biosynthesis and Is Essential for Growth in Macrophages. *J. Bacteriol.* **196**: 1853–65.
- Quigley, J., Hughitt, V.K., Velikovskiy, C.A., Mariuzza, R.A., El-Sayed, N.M., and Briken, V. (2017) The cell wall lipid PDIM contributes to phagosomal escape and host cell exit of *Mycobacterium tuberculosis*. *MBio* **8**: 1–12.
- Quinlan, A.R. (2014) BEDTools: The Swiss-Army tool for genome feature analysis.
- Ramirez, F., Ryan, D.P., Gruning, B., Bhardwaj, V., Kilpert, F., Richter, A.S., et al. (2016) deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* **44**: 160–165.
- Rastogi, S., Singh, A.K., Pant, G., Mitra, K., Sashidhara, K. V., and Krishnan, M.Y. (2017) Down-regulation of PE11, a cell wall associated esterase, enhances the biofilm growth of *Mycobacterium tuberculosis* and reduces cell wall virulence lipid levels. 52–61.
- Richards, J.P. and Ojha, A.K. (2014) Mycobacterial Biofilms. 1–11.
- Robinson, J.T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesiro, J.P. (2011) Integrative Genomic

- Viewer. *Nature* **29**: 24–26.
- Sala, C., Forti, F., Florio, E. Di, Canneva, F., Milano, A., Riccardi, G., and Ghisotti, D. (2003) *Mycobacterium tuberculosis* FurA Autoregulates Its Own Expression. *J. Bacteriol.* **185**: 5357–5362.
- Schnappinger, D., Ehrt, S., Voskuil, M.I., Liu, Y., Mangan, J. a, Monahan, I.M., et al. (2003) Transcriptional Adaptation of *Mycobacterium tuberculosis* within Macrophages: Insights into the Phagosomal Environment. *J. Exp. Med.* **198**: 693–704.
- Schubert, O.T., Ludwig, C., Kogadeeva, M., Zimmermann, M., Rosenberger, G., Gengenbacher, M., et al. (2015) Absolute Proteome Composition and Dynamics during Dormancy and Resuscitation of *Mycobacterium tuberculosis*. *Cell Host Microbe* **18**: 96–108.
- Slayden, R.A., Knudson, D.L., and Belisle, J.T. (2006) Identification of cell cycle regulators in *Mycobacterium tuberculosis* by inhibition of septum formation and global transcriptional analysis. *Microbiology* **152**: 1789–1797.
- Sureka, K., Dey, S., Datta, P., Singh, A.K., Dasgupta, A., Rodrigue, S., et al. (2007) Polyphosphate kinase is involved in stress-induced mprAB-sigE-rel signalling in mycobacteria. *Mol. Microbiol.* **65**: 261–276.
- Wayne, L.G. (1977) Synchronized replication of *Mycobacterium tuberculosis*. *Infect. Immun.* **17**: 528–530.

SUPPLEMENTARY MATERIAL

Supplementary Table 1. Primers used in this study.

Primer name	Sequence (5' → 3')
CS-080-dTanchor	GAC CAC GCG TAT CGA TGT CGA CTT TTT TTT TTT TTT TTV
CS-364- <i>mihF</i> UF	ACG TTC TTA ATT AAC CCG CAG TTG CTG TCG GCT GCC C
CS-365- <i>mihF</i> UR	ACG TTC CCT AGG CAA TGG ATT TCG GCT TTC CGG ACC TGC
CS-366- <i>mihF</i> DF	ACG TTC CCT AGG TAA CCC CGC CGG CCG ACG ATG
CS-367- <i>mihF</i> DR	ACG TTC GGC GCG CCC CCT CGC CAA GCT GGT TGT AG
CS-370- <i>mihF</i> FF	ACG TGT CCT AGG TTA GGC AAC ACT ATT CAT GTC CCA TGC
CS-371- <i>mihF</i> FR	ACG TGT GGC GCG CCT TAG GCG GAG CCG AAC TTT TCC AGC
CS-402-U1	CAG ATC GCT ACC TGG CTG GGC
CS-403-D1	GGT ACG CCC ACC ACA GGC
CS-404-U2	CAC CAG TCG TCG GCC GTC GTC
CS-405-D2	GCG GAC CGA CAT ATT CGA GGA TG
CS-415- <i>mihF</i> FF	GAC GCA ACC ATG GCG CGA C
CS-423-PfurA102F	ACG TGA TTC GAA CCA CCA TGC AGG CCC GGC
CS-424-PfurA102R	GGT CCC TCG CCG ACG CTC ACA CTA GAC AAT ATG ACT CCC
CS-425- <i>gmK</i> F	GGG AGT CAT ATT GTC TAG TGT GAG CGT CGG CGA GGG ACC
CS-426- <i>rpoZ</i> R	ACG TCC ATG GCT ACT CGC CCT CGG TGT GCT CG
<i>mihF</i> -80-F	CGA TCC TAG GAT GAG AGA CGG AGG AAT CGT GG
<i>mihF</i> -86-F	GCA TCC TAG GGT GGC CCT TCC CCA GTT GAC CG
NO-094-rev-120	GCG TCC TTG AGG ACC TGG G
NO-095-rev-180	GGC AAG GCC TCA AGC AGC G

Supplementary Table 2. Plasmids used in this study.

Plasmid name	Description	Reference
pJG1100	Suicide vector for mutant construction, Hyg ^R , Kan ^R , sacB	(Gomez, 2000)
pGA44	Integrative vector at L5 attB site, carrying the TET-PIP OFF expression system, Str ^R	(Kolly, 2014)
pGA80	pMV261-derived vector, carrying the L5 <i>int</i> gene for expression <i>in trans</i> , lacking oriM, Kan ^R .	(Kolly, 2014)
pGA118	Integrative vector at L5 attB site, carrying the TET-PIP OFF expression system, Hyg ^R	This study
pCS35	Suicide vector for mutant construction derived from pJG1100, Hyg ^R , Kan ^R , sacB	This study
pCS31	Full-length <i>mihF</i> cloned in pGA44, integrative vector at L5 attB site, Str ^R	This study
pNO71	Empty vector pGA44 carrying <i>PfurA102-gmk-rpoZ</i>	This study
pNO12	Full-length <i>mihF</i> , <i>PfurA102-gmk-rpoZ</i> cloned in pGA44, integrative vector at L5 attB site, Str ^R	This study
pNO62	<i>mihF</i> -80, <i>PfurA102-gmk-rpoZ</i> cloned in pGA118, integrative vector at L5 attB site, Hyg ^R	This study
pNO63	<i>mihF</i> -86, <i>PfurA102-gmk-rpoZ</i> cloned in pGA118, integrative vector at L5 attB site, Hyg ^R	This study

Fig. S1

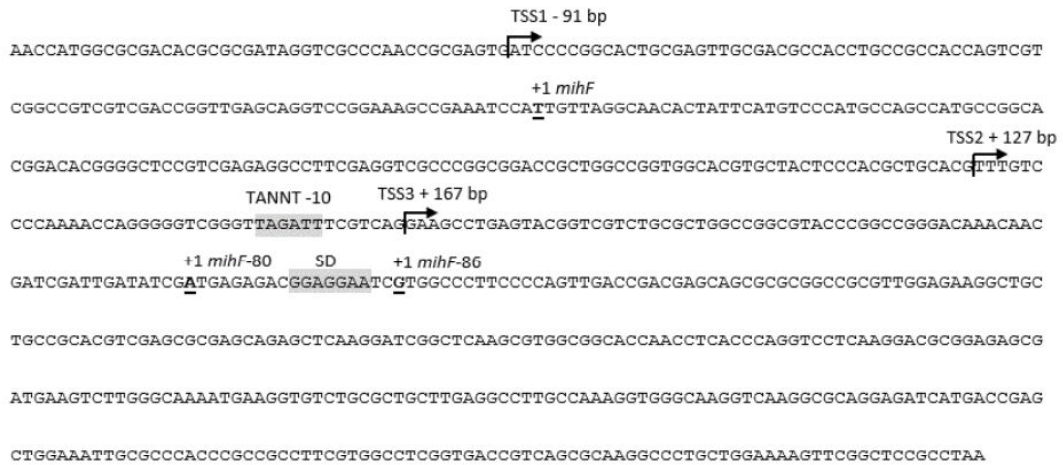


Fig. S1. *mihF* gene locus with the annotated translation start site (*mihF*), and the two proposed short forms of *mihF* (*mihF*-80 and *mihF*-86). The three identified transcription start sites (TSS) by 5'-RACE are marked by arrows and the potential AG-rich Shine-Dalgarno (SD) ribosomal binding site and TANNT -10 motifs are highlighted in grey. Coordinate +1 *mihF* corresponds to the first base of the full-length *mihF* translation start site.

Fig. S2

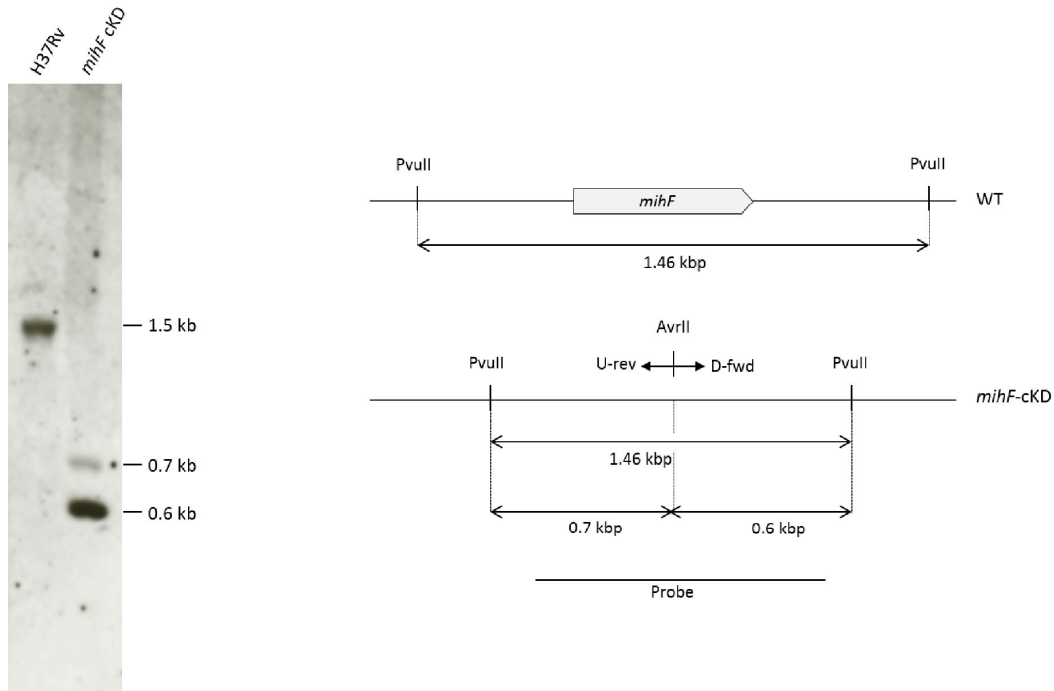


Fig. S2. Southern blot analysis of *mihF*-cKD and H37Rv parental strains and genomic locus with restriction sites and probe used for Southern blot.

Fig. S3

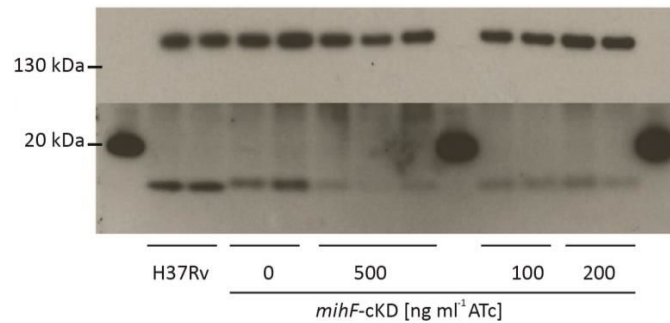


Fig. S3. Immunoblot of miHF levels upon *miHF* silencing by ATc. Duplicate or triplicate (in case of *miHF*-cKD with 500 ng ml⁻¹ ATc) samples were taken after 2 dilutions (9 days).

Fig. S4

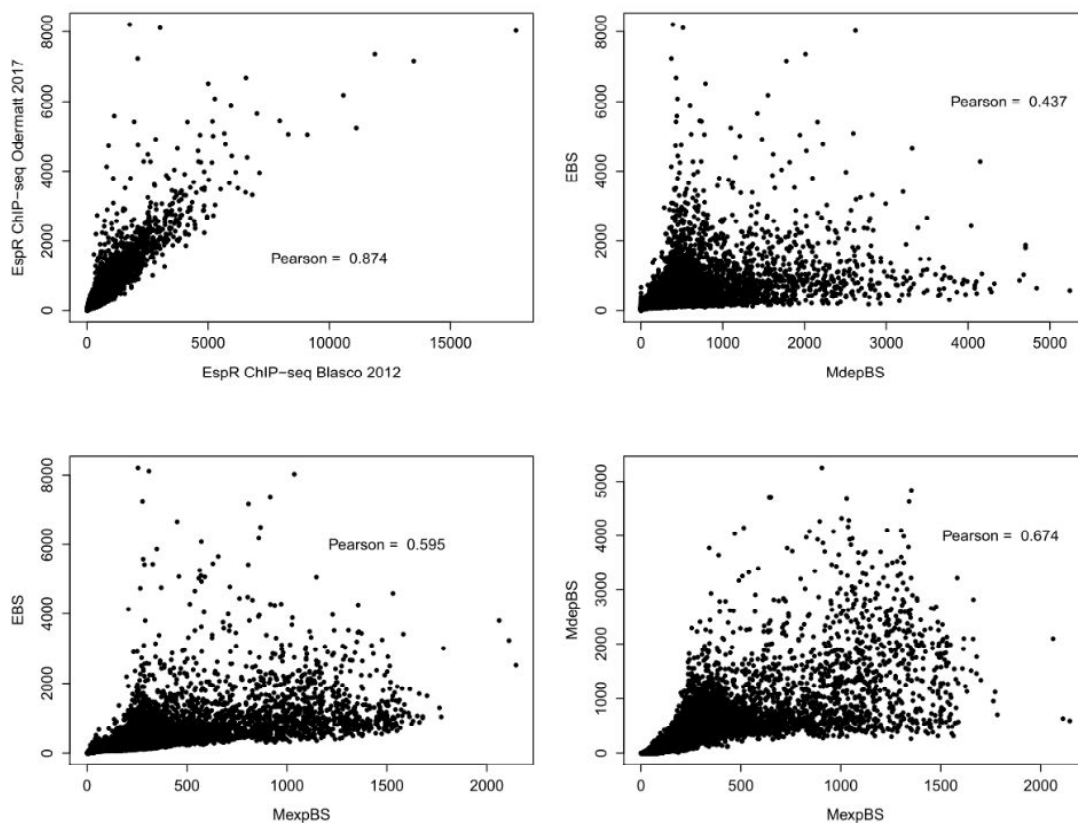


Fig. S4. Correlation between different ChIP-seq experiments. a) Correlation between the 2012 EspR ChIP-Seq by Blasco et al. and the EspR ChIP-seq conducted now has a Pearson index of 0.87. b) and c) Correlation between miHF binding sites in miHF depleted condition (MdepBS) versus EspR binding sites (Pearson index = 0.44) and between exponentially growing cells (MexpBS) and EspR (Pearson index = 0.6). d) Correlation between MexpBS and MdepBS with a Pearson index = 0.67.

Fig. S5

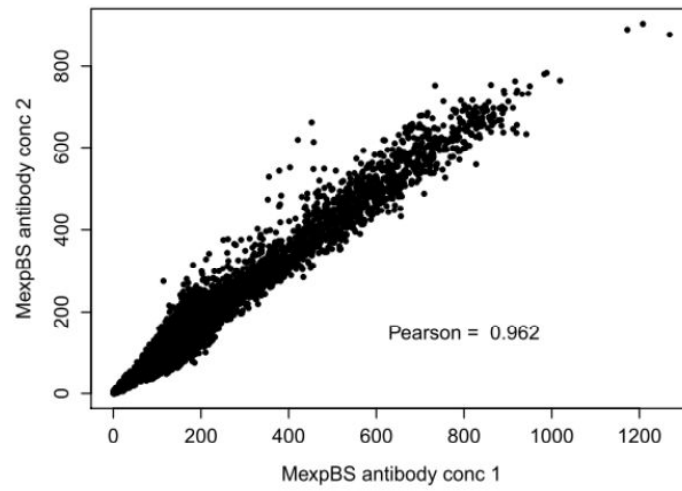


Fig. S4. Correlation between mIHF binding sites in exponentially growing cells in two independent experiments conducted with different concentrations of primary anti-mIHF antibody. Data was pooled because of the excellent correlation (Pearson index = 0.96).

EXCERPT FROM DATASET S1

RNA-seq miHF-ckD + ATc, Fold change (FC) > 2

Locus	FC	pvalue	name	Rv0170	-3.59	0.000	mce1B
Rv0001	-5.68	0.000	dnaA				
Rv0002	-4.57	0.000	dnaN	Rv0171	-2.66	0.000	mce1C
Rv0003	-2.69	0.000	recF	Rv0172	-3.89	0.000	mce1D
Rv0005	-2.01	0.009	gyrB	Rv0173	-3.24	0.000	lprK
MTB000001	-2.50	0.003	ileT	Rv0174	-2.60	0.000	mce1F
MTB000002	-2.57	0.003	alaT	Rv0177	-2.13	0.025	Rv0177
Rv0009	-3.62	0.000	ppiA	Rv0182c	2.41	0.000	sigG
Rv0010c	-3.76	0.000	Rv0010c	Rv0184	3.79	0.000	Rv0184
Rv0011c	-4.09	0.000	Rv0011c	Rv0185	2.67	0.000	Rv0185
Rv0015c	3.09	0.000	pknA	Rv0186	3.02	0.000	bglS
Rv0018c	2.56	0.000	pstP	Rv0186A	-4.30	0.000	mymT
MTB000003	-2.39	0.001	leuT	Rv0193c	-3.21	0.000	Rv0193c
Rv0023	2.44	0.001	Rv0023	Rv0196	13.64	0.000	Rv0196
Rv0024	2.40	0.000	Rv0024	Rv0197	11.77	0.000	Rv0197
Rv0029	-2.13	0.004	Rv0029	Rv0199	2.34	0.000	Rv0199
Rv0031	2.24	0.004	Rv0031	Rv0200	2.23	0.008	Rv0200
Rv0032	-2.79	0.002	bioF2	Rv0201c	2.70	0.014	Rv0201c
Rv0033	-3.58	0.000	acpA	Rv0204c	3.05	0.000	Rv0204c
Rv0034	-2.87	0.000	Rv0034	Rv0207c	-2.24	0.000	Rv0207c
Rv0035	-3.61	0.000	fadD34	Rv0208c	-2.73	0.000	Rv0208c
Rv0036c	2.74	0.007	Rv0036c	Rv0209	2.55	0.013	Rv0209
Rv0037c	2.94	0.001	Rv0037c	Rv0210	2.11	0.030	Rv0210
Rv0042c	-2.32	0.000	Rv0042c	Rv0227c	-2.49	0.000	Rv0227c
Rv0053	-2.19	0.000	rpsF	Rv0230c	-4.51	0.000	php
Rv0055	-4.46	0.000	rpsR1	Rv0231	-4.34	0.000	fadE4
Rv0056	-2.78	0.000	rplI	Rv0232	-2.02	0.000	Rv0232
Rv0057	-2.14	0.029	Rv0057	Rv0237	2.28	0.025	lpqI
Rv0061c	-5.33	0.000	Rv0061c	Rv0239	-2.19	0.000	vapB24
Rv0076c	-5.55	0.000	Rv0076c	MTB0000051	-2.37	0.000	F6
Rv0077c	-2.60	0.000	Rv0077c	Rv0247c	-3.27	0.000	Rv0247c
Rv0078B	-2.37	0.004	Rv0078B	Rv0248c	-2.56	0.000	Rv0248c
Rv0079	2.00	0.014	Rv0079	Rv0249c	-4.60	0.000	Rv0249c
Rv0080	2.48	0.023	Rv0080	Rv0250c	-4.71	0.000	Rv0250c
Rv0081	-2.17	0.000	Rv0081	Rv0251c	-2.61	0.000	hsp
Rv0084	2.35	0.029	hycD	Rv0259c	3.42	0.001	Rv0259c
Rv0085	2.22	0.035	hycP	Rv0260c	5.78	0.000	Rv0260c
Rv0088	-2.12	0.000	Rv0088	Rv0264c	2.12	0.010	Rv0264c
Rv0094c	2.75	0.000	Rv0094c	Rv0266c	2.72	0.000	oplA
Rv0096	-7.17	0.000	PPE1	Rv0268c	-4.38	0.000	Rv0268c
Rv0097	-10.51	0.000	Rv0097	Rv0272c	-2.36	0.000	Rv0272c
Rv0098	-7.11	0.000	fcoT	Rv0273c	-2.91	0.000	Rv0273c
Rv0099	-3.88	0.000	fadD10	Rv0277c	-3.79	0.000	vapC25
Rv0100	-3.02	0.000	Rv0100	Rv0277A	-5.33	0.000	vapB25
Rv0104	-2.07	0.001	Rv0104	Rv0279c	2.27	0.002	PE_PGRS4
Rv0106	2.77	0.000	Rv0106	Rv0281	3.78	0.000	Rv0281
Rv0108c	-3.68	0.000	Rv0108c	Rv0283	3.96	0.000	eccB3
Rv0112	-5.10	0.000	gca	Rv0289	2.85	0.000	espG3
Rv0113	-5.21	0.000	gmhA	Rv0290	3.45	0.000	eccD3
Rv0114	-2.71	0.001	gmhB	Rv0291	4.67	0.000	mycP3
Rv0127	2.12	0.001	mak	Rv0292	3.10	0.004	eccE3
Rv0130	2.27	0.014	htdZ	Rv0304c	-2.91	0.000	PPE5
Rv0133	2.72	0.001	Rv0133	Rv0305c	-3.60	0.000	PPE6
Rv0139	2.57	0.009	Rv0139	Rv0306	-2.12	0.008	Rv0306
Rv0140	-3.02	0.000	Rv0140	Rv0307c	-2.04	0.000	Rv0307c
Rv0141c	-2.36	0.002	Rv0141c	Rv0309	-2.95	0.000	Rv0309
Rv0144	-2.72	0.000	Rv0144	MTB0000004	-2.69	0.005	glyU
Rv0151c	-2.39	0.006	PE1	Rv0322	2.12	0.008	udgA
Rv0157	2.24	0.000	pntB	Rv0327c	-3.95	0.000	cyp135A1
Rv0157A	-2.47	0.001	Rv0157A	Rv0336	2.11	0.001	Rv0336
Rv0159c	-3.26	0.000	PE3	Rv0339c	2.89	0.000	Rv0339c
Rv0165c	-5.07	0.000	mce1R	Rv0341	2.68	0.000	iniB
Rv0166	-3.41	0.000	fadD5	Rv0343	3.47	0.000	iniC
Rv0167	-3.38	0.000	yrbE1A	Rv0346c	-3.06	0.000	ansP2
Rv0168	-2.75	0.000	yrbE1B	Rv0347	-2.51	0.001	Rv0347
Rv0169	-3.67	0.000	mce1A	Rv0348	-3.44	0.000	Rv0348

3.1 – mIHF Function

Rv0349	-2.64	0.000	Rv0349	Rv0520	-3.05	0.000	Rv0520
Rv0350	-3.06	0.000	dnaK	Rv0523c	-2.65	0.000	Rv0523c
Rv0351	-2.51	0.000	grpE	Rv0524	2.51	0.001	hemL
Rv0352	-2.76	0.000	dnaI1	Rv0525	3.50	0.001	Rv0525
Rv0354c	-3.63	0.000	PPE7	Rv0526	3.78	0.000	Rv0526
Rv0357c	2.00	0.000	purA	Rv0527	3.43	0.001	ccdA
Rv0358	4.64	0.000	Rv0358	Rv0528	4.50	0.000	Rv0528
Rv0359	2.71	0.004	Rv0359	Rv0529	3.96	0.000	ccsA
Rv0361	2.04	0.000	Rv0361	Rv0530	4.39	0.000	Rv0530
Rv0362	2.52	0.000	mgtE	Rv0532	3.38	0.000	PE_PGR56
Rv0364	-3.22	0.000	Rv0364	Rv0536	2.76	0.006	galE3
Rv0368c	2.05	0.057	Rv0368c	Rv0537c	2.47	0.002	Rv0537c
Rv0379	-3.10	0.000	secE2	Rv0539	2.66	0.002	Rv0539
Rv0380c	3.30	0.000	Rv0380c	Rv0540	3.46	0.000	Rv0540
Rv0381c	3.53	0.002	Rv0381c	Rv0541c	2.29	0.008	Rv0541c
Rv0387c	2.52	0.008	Rv0387c	Rv0552	2.77	0.002	Rv0552
Rv0390	2.51	0.016	Rv0390	Rv0553	2.42	0.011	menC
Rv0397A	-2.74	0.000	Rv0397A	Rv0557	3.58	0.000	mgtA
Rv0401	2.66	0.020	Rv0401	Rv0561c	3.07	0.000	Rv0561c
Rv0402c	-3.59	0.000	mmpL1	Rv0566c	-2.22	0.000	Rv0566c
Rv0403c	-2.67	0.000	mmpS1	MTB000005	-2.01	0.100	tyrT
Rv0404	-5.30	0.000	fadB30	Rv0570	2.05	0.011	nrdZ
Rv0405	-3.68	0.000	pkc6	Rv0572c	-3.42	0.000	Rv0572c
Rv0411c	3.06	0.000	glnH	Rv0575c	2.42	0.006	Rv0575c
Rv0412c	2.19	0.000	Rv0412c	Rv0578c	-2.82	0.000	PE_PGR57
Rv0413	2.46	0.012	mutT3	Rv0593	2.21	0.000	lprL
Rv0421c	3.20	0.003	Rv0421c	Rv0599c	-2.36	0.001	vapB27
Rv0425c	2.10	0.000	ctpH	Rv0600c	2.23	0.059	Rv0600c
Rv0426c	-2.77	0.000	Rv0426c	Rv0603	2.68	0.018	Rv0603
Rv0430	-3.71	0.000	Rv0430	Rv0604	2.36	0.014	lpqO
Rv0431	-2.28	0.001	Rv0431	Rv0606	3.91	0.000	Rv0606
Rv0435c	2.54	0.005	Rv0435c	Rv0607	2.10	0.000	Rv0607
Rv0440	-3.33	0.000	groEL2	Rv0612	-2.10	0.004	Rv0612
Rv0448c	2.08	0.005	Rv0448c	Rv0613c	3.15	0.000	Rv0613c
Rv0450c	-2.72	0.002	mmpL4	Rv0616A	2.57	0.000	vapB29
Rv0451c	-3.41	0.000	mmpS4	Rv0617	2.21	0.001	vapC29
Rv0452	2.17	0.015	Rv0452	Rv0618	2.60	0.024	galTa
Rv0455c	-2.16	0.026	Rv0455c	Rv0620	3.05	0.009	galK
Rv0456B	-2.45	0.000	mazE1	Rv0626	2.54	0.030	vapB5
Rv0466	-2.34	0.000	Rv0466	Rv0627	2.00	0.065	vapC5
Rv0468	-2.78	0.000	fadB2	Rv0629c	3.04	0.000	recD
Rv0469	-5.16	0.000	umaA	Rv0630c	2.47	0.000	recB
Rv0470c	-2.45	0.001	pcaA	Rv0631c	3.66	0.000	recC
Rv0474	-2.27	0.000	Rv0474	Rv0634A	-2.11	0.000	Rv0634A
Rv0478	2.08	0.001	deoC	MTB000006	-2.56	0.005	thrT
Rv0479c	-2.74	0.000	Rv0479c	MTB000007	-7.31	0.000	metT
Rv0483	-3.04	0.000	lprQ	Rv0634B	-6.22	0.000	rpmG2
Rv0485	2.07	0.000	Rv0485	Rv0635	-3.93	0.000	hadA
MTB000060	2.31	0.030	mcr19	Rv0636	-3.10	0.000	hadB
Rv0486	4.18	0.000	mshA	Rv0637	-4.11	0.000	hadC
Rv0490	3.13	0.000	senX3	Rv0638	-2.46	0.002	secE1
Rv0491	2.20	0.000	regX3	Rv0639	-2.14	0.005	nusG
Rv0492c	2.23	0.025	Rv0492c	Rv0642c	-3.02	0.000	mmaA4
Rv0496	2.25	0.000	Rv0496	Rv0643c	-3.22	0.000	mmaA3
Rv0498	2.85	0.000	Rv0498	Rv0644c	-2.08	0.000	mmaA2
Rv0499	2.57	0.007	Rv0499	Rv0648	2.06	0.064	Rv0648
Rv0500A	-3.16	0.000	Rv0500A	Rv0654	-3.96	0.000	Rv0654
Rv0500B	-5.36	0.000	Rv0500B	Rv0655	-4.72	0.000	mkI
Rv0501	-3.42	0.000	galE2	Rv0661c	2.33	0.048	vapC7
Rv0502	-3.10	0.000	Rv0502	Rv0665	2.61	0.001	vapC8
Rv0503c	-2.39	0.001	cmaA2	Rv0670	3.30	0.000	end
Rv0506	-4.10	0.000	mmpS2	Rv0671	2.15	0.001	lpqP
Rv0507	-3.36	0.000	mmpL2	Rv0674	2.44	0.000	Rv0674
Rv0510	2.05	0.000	hemC	Rv0675	2.03	0.014	echA5
Rv0511	3.71	0.000	hemD	Rv0679c	-4.51	0.000	Rv0679c
Rv0512	3.06	0.001	hemB	Rv0680c	-2.69	0.000	Rv0680c
Rv0513	2.06	0.001	Rv0513	Rv0682	-2.92	0.000	rpsL
Rv0516c	-2.93	0.000	Rv0516c	Rv0683	-2.91	0.000	rpsG
Rv0517	-2.53	0.000	Rv0517	Rv0684	-2.62	0.000	fusA1
Rv0518	-2.46	0.000	Rv0518	Rv0686	-2.12	0.009	Rv0686
Rv0519c	4.51	0.000	Rv0519c	Rv0690c	2.15	0.009	Rv0690c

mIHF Function – 3.1

Rv0691A	-5.34	0.000	Rv0691A	MTB000011	-4.18	0.000	gluT
Rv0692	-2.16	0.000	Rv0692	MTB000012	-3.25	0.000	aspT
Rv0696	2.49	0.000	Rv0696	MTB000013	-3.57	0.000	pheU
Rv0699	4.02	0.001	Rv0699	Rv0833	2.12	0.080	PE_PGRS13
Rv0700	-3.05	0.000	rpsJ	Rv0835	-2.22	0.002	lpqQ
Rv0701	-3.09	0.000	rplC	Rv0836c	-2.58	0.001	Rv0836c
Rv0703	-2.45	0.000	rplW	Rv0837c	-4.78	0.000	Rv0837c
Rv0704	-4.06	0.000	rplB	Rv0839	-2.28	0.000	Rv0839
Rv0705	-3.37	0.000	rpsS	Rv0843	2.48	0.000	Rv0843
Rv0706	-3.04	0.000	rplV	Rv0850	-3.65	0.000	Rv0850
Rv0707	-2.33	0.000	rpsC	Rv0851c	2.13	0.010	Rv0851c
Rv0709	-2.06	0.002	rpmC	Rv0854	-3.11	0.000	Rv0854
Rv0710	-2.38	0.000	rpsQ	Rv0856	-2.82	0.000	Rv0856
Rv0714	-2.79	0.000	rplN	Rv0859	-2.64	0.000	fadA
Rv0715	-3.34	0.000	rplX	Rv0861c	3.01	0.000	ercc3
Rv0716	-2.28	0.000	rplE	Rv0862c	5.18	0.000	Rv0862c
Rv0718	-3.29	0.000	rpsH	Rv0865	2.21	0.017	mog
Rv0719	-2.89	0.000	rplF	Rv0866	2.88	0.003	moaE2
Rv0726c	-3.64	0.000	Rv0726c	Rv0867c	-4.21	0.000	rpfA
Rv0729	2.08	0.000	xyfB	Rv0868c	-2.06	0.005	moaD2
Rv0732	-2.09	0.012	secY	Rv0872c	2.63	0.005	PE_PGRS15
Rv0735	2.08	0.017	sigL	Rv0874c	2.31	0.001	Rv0874c
Rv0737	2.41	0.021	Rv0737	Rv0884c	2.80	0.000	serC
Rv0739	-3.72	0.000	Rv0739	Rv0886	2.94	0.000	fprB
Rv0740	-2.57	0.000	Rv0740	Rv0887c	3.04	0.001	Rv0887c
Rv0741	-2.43	0.001	Rv0741	Rv0888	-3.98	0.000	Rv0888
Rv0746	2.65	0.022	PE_PGRS9	Rv0889c	2.59	0.000	citA
Rv0748	-2.60	0.004	vapB31	Rv0895	2.69	0.000	Rv0895
Rv0749	-2.43	0.001	vapC31	Rv0897c	2.94	0.000	Rv0897c
Rv0749A	-2.95	0.003	Rv0749A	Rv0898c	2.92	0.005	Rv0898c
Rv0750	-2.38	0.000	Rv0750	Rv0902c	2.26	0.001	prfB
Rv0752c	-2.73	0.000	fadE9	Rv0905	2.79	0.000	echA6
Rv0753c	-2.88	0.000	mmsA	Rv0906	2.73	0.000	Rv0906
Rv0755c	-4.04	0.000	PPE12	Rv0909	-3.21	0.000	Rv0909
Rv0755A	-3.39	0.000	Rv0755A	Rv0917	-2.99	0.000	betP
MTB000009	-2.46	0.006	thrV	MTB000014	-3.68	0.000	argT
Rv0757	-2.72	0.000	phoP	Rv0922	2.14	0.000	Rv0922
Rv0760c	-4.94	0.000	Rv0760c	Rv0938	2.20	0.000	ligD
Rv0761c	-4.23	0.000	adhB	Rv0939	2.06	0.000	Rv0939
Rv0762c	-2.80	0.000	Rv0762c	Rv0940c	-2.45	0.002	Rv0940c
Rv0763c	-3.55	0.000	Rv0763c	Rv0942	-2.56	0.005	Rv0942
Rv0764c	-2.22	0.001	cyp51	Rv0947c	-2.74	0.000	Rv0947c
Rv0765c	-2.00	0.010	Rv0765c	Rv0948c	-4.34	0.000	Rv0948c
Rv0766c	-3.51	0.000	cyp123	Rv0950c	-4.28	0.000	Rv0950c
Rv0767c	-2.04	0.033	Rv0767c	Rv0958	3.33	0.000	Rv0958
Rv0768	-2.53	0.000	aldA	Rv0959	2.97	0.000	Rv0959
Rv0769	-3.16	0.000	Rv0769	Rv0960	-2.10	0.000	vapC9
Rv0771	-2.48	0.000	Rv0771	Rv0961	-2.73	0.000	Rv0961
Rv0776c	3.32	0.001	Rv0776c	Rv0969	2.35	0.003	ctpV
Rv0779c	2.26	0.057	Rv0779c	Rv0985c	-3.81	0.000	mscL
Rv0781	-2.09	0.012	ptrBa	Rv0986	-2.88	0.004	Rv0986
Rv0784	9.02	0.000	Rv0784	Rv0987	-3.59	0.000	Rv0987
Rv0785	2.81	0.000	Rv0785	Rv0988	-4.53	0.000	Rv0988
Rv0789c	-2.49	0.002	Rv0789c	Rv0990c	-2.40	0.009	Rv0990c
Rv0790c	-6.45	0.000	Rv0790c	Rv0991c	-5.40	0.000	Rv0991c
Rv0791c	-10.22	0.000	Rv0791c	Rv0994	2.17	0.006	moeA1
Rv0792c	-6.52	0.000	Rv0792c	Rv0995	2.23	0.025	rimJ
Rv0800	2.55	0.009	pepC	Rv1000c	3.63	0.002	Rv1000c
Rv0804	2.08	0.083	Rv0804	Rv1001	3.55	0.000	arcA
Rv0806c	-3.13	0.000	cpsY	Rv1003	2.07	0.051	Rv1003
Rv0808	2.19	0.003	purF	Rv1004c	3.38	0.000	Rv1004c
Rv0810c	-2.91	0.000	Rv0810c	Rv1005c	2.67	0.000	pabB
Rv0812	2.75	0.000	Rv0812	Rv1013	2.16	0.000	pkS16
Rv0814c	-6.66	0.000	sseC2	MTB000016	-6.10	0.000	glnT
Rv0815c	-4.44	0.000	cysA2	Rv1020	2.75	0.000	mfd
Rv0824c	-2.11	0.013	desA1	Rv1023	2.28	0.000	eno
MTB000053	-2.30	0.004	ASdes	Rv1024	3.79	0.001	Rv1024
Rv0825c	-3.09	0.000	Rv0825c	Rv1025	3.02	0.006	Rv1025
Rv0828c	-2.66	0.001	Rv0828c	Rv1026	2.27	0.013	Rv1026
Rv0829	2.40	0.001	Rv0829	Rv1033c	-4.28	0.000	trcR
Rv0831c	-4.45	0.000	Rv0831c	Rv1034c	2.32	0.011	Rv1034c

3.1 – mIHF Function

Rv1048c	-3.35	0.000	Rv1048c	Rv1264	3.15	0.001	Rv1264
Rv1054	2.01	0.011	Rv1054	MTB0000063	-2.27	0.006	mcr11
Rv1055	-2.05	0.011	Rv1055	Rv1268c	-2.53	0.006	Rv1268c
MTB0000017	-2.70	0.001	leuX	Rv1270c	-2.52	0.000	lprA
Rv1057	7.20	0.000	Rv1057	Rv1272c	2.15	0.000	Rv1272c
Rv1063c	2.44	0.002	Rv1063c	Rv1273c	2.09	0.000	Rv1273c
Rv1066	2.58	0.017	Rv1066	Rv1277	2.03	0.000	Rv1277
Rv1067c	2.19	0.018	PE_PGRS19	Rv1285	-2.15	0.001	cysD
Rv1072	-2.43	0.000	Rv1072	Rv1286	-2.49	0.000	cysN
Rv1073	-2.94	0.000	Rv1073	MTB0000018	-2.25	0.053	argV
Rv1076	-2.01	0.002	lipU	Rv1292	2.44	0.000	argS
Rv1078	-2.58	0.001	pra	Rv1293	3.01	0.000	lysA
Rv1080c	-2.58	0.000	greA	Rv1294	2.98	0.000	thrA
Rv1083	-2.21	0.012	Rv1083	Rv1295	2.92	0.000	thrC
Rv1085c	2.35	0.017	Rv1085c	Rv1297	-2.40	0.000	rho
Rv1087A	2.97	0.003	Rv1087A	Rv1298	-3.66	0.000	rpmE
Rv1094	-2.17	0.001	desA2	Rv1301	2.59	0.000	Rv1301
Rv1095	2.53	0.002	phoH2	Rv1303	-2.05	0.000	Rv1303
Rv1098c	2.37	0.001	fum	Rv1304	-2.98	0.002	atpB
Rv1101c	-3.06	0.000	Rv1101c	Rv1305	-3.66	0.000	atpE
Rv1113	-3.40	0.001	vapB32	Rv1306	-2.36	0.000	atpF
Rv1115	-2.22	0.001	Rv1115	Rv1308	-2.09	0.000	atpA
Rv1116	-2.33	0.005	Rv1116	Rv1313c	-2.33	0.000	Rv1313c
Rv1117	2.04	0.032	Rv1117	MTB0000064	-7.77	0.000	mcr3
Rv1120c	2.16	0.066	Rv1120c	MTB0000019	-2.03	0.097	rrs
Rv1127c	2.88	0.000	ppdK	MTB0000020	-2.87	0.005	rri
Rv1128c	2.15	0.001	Rv1128c	MTB0000021	-3.78	0.001	rri
Rv1129c	4.68	0.000	Rv1129c	Rv1317c	2.11	0.002	alkA
Rv1130	3.25	0.000	prpD	Rv1329c	2.42	0.001	dinG
Rv1133c	2.18	0.019	metE	Rv1330c	3.19	0.000	pncB1
Rv1135c	-2.17	0.006	PPE16	Rv1336	2.15	0.002	cysM
Rv1141c	-2.27	0.001	echA11	Rv1340	2.23	0.002	rphA
Rv1155	-2.46	0.000	Rv1155	Rv1341	2.58	0.006	Rv1341
MTB0000072	2.83	0.011	mcr10	Rv1345	2.82	0.000	mbtM
Rv1159A	-3.58	0.000	Rv1159A	Rv1346	2.75	0.000	mbtN
Rv1163	2.37	0.004	narJ	Rv1347c	-3.48	0.000	mbtK
Rv1168c	-13.56	0.000	PPE17	MTB0000022	-2.19	0.018	leuW
Rv1169c	-18.29	0.000	lipX	Rv1349	2.22	0.000	irtB
Rv1171	-2.61	0.000	Rv1171	Rv1362c	-2.26	0.002	Rv1362c
Rv1172c	-2.59	0.000	PE12	Rv1363c	-2.04	0.005	Rv1363c
Rv1176c	-4.37	0.000	Rv1176c	MTB0000065	-5.60	0.000	mcr15
Rv1181	3.04	0.000	pkS4	Rv1364c	-2.21	0.001	Rv1364c
Rv1185c	-2.37	0.005	fadD21	Rv1365c	-4.91	0.000	rsfA
Rv1189	2.48	0.001	sigI	Rv1374c	-3.48	0.000	Rv1374c
Rv1190	2.52	0.000	Rv1190	MTB0000076	-3.58	0.000	MTS1082
Rv1192	2.53	0.001	Rv1192	Rv1375	-2.29	0.000	Rv1375
Rv1193	2.54	0.000	fadD36	Rv1377c	2.50	0.006	Rv1377c
MTB0000075	-4.02	0.000	ncrMT1234	Rv1378c	2.92	0.000	Rv1378c
Rv1197	-3.62	0.000	esxK	Rv1379	2.24	0.044	pyrR
Rv1198	-2.52	0.000	esxL	Rv1381	2.08	0.000	pyrC
Rv1199c	-2.02	0.001	Rv1199c	Rv1383	2.51	0.000	carA
Rv1208	2.07	0.006	gpgS	Rv1387	-2.15	0.009	PPE20
Rv1209	2.94	0.011	Rv1209	Rv1388	-11.14	0.000	mihF
Rv1216c	-4.58	0.000	Rv1216c	Rv1389	-7.84	0.000	gmk
Rv1218c	-2.97	0.000	Rv1218c	Rv1396c	-4.11	0.000	PE_PGRS25
Rv1219c	-2.85	0.000	Rv1219c	Rv1399c	3.63	0.000	nlhH
Rv1220c	2.53	0.020	Rv1220c	Rv1400c	2.59	0.002	lipI
Rv1221	-4.46	0.000	sigE	Rv1402	2.10	0.037	priA
MTB0000062	-3.77	0.000	mpr6	Rv1404	-2.30	0.002	Rv1404
Rv1225c	2.15	0.000	Rv1225c	Rv1406	2.55	0.004	fmt
Rv1226c	4.82	0.000	Rv1226c	Rv1407	2.90	0.008	fmu
Rv1227c	2.66	0.000	Rv1227c	Rv1408	2.78	0.000	rpe
Rv1237	-2.11	0.000	sugB	Rv1409	2.01	0.000	ribG
Rv1238	-2.38	0.003	sugC	Rv1413	2.42	0.015	Rv1413
Rv1243c	3.71	0.000	PE_PGRS23	Rv1421	2.32	0.000	Rv1421
Rv1244	2.91	0.001	lpqZ	Rv1422	2.49	0.005	Rv1422
Rv1247c	-2.86	0.000	relB	Rv1432	2.22	0.036	Rv1432
Rv1255c	-2.18	0.000	Rv1255c	Rv1443c	2.19	0.050	Rv1443c
Rv1257c	2.16	0.044	Rv1257c	Rv1457c	3.13	0.000	Rv1457c
Rv1262c	2.07	0.027	Rv1262c	Rv1458c	4.30	0.000	Rv1458c
Rv1263	3.13	0.000	amiB2	Rv1459c	4.03	0.000	Rv1459c

mIHF Function – 3.1

Rv1460	3.34	0.000	Rv1460	Rv1646	-3.27	0.000	PE17
Rv1462	3.78	0.000	Rv1462	Rv1647	2.97	0.000	Rv1647
Rv1463	2.93	0.000	Rv1463	Rv1648	2.44	0.000	Rv1648
Rv1464	2.69	0.000	csd	Rv1650	2.10	0.010	pheT
Rv1465	2.50	0.000	Rv1465	Rv1652	2.43	0.015	argC
Rv1466	2.64	0.002	Rv1466	Rv1655	2.21	0.063	argD
Rv1470	-3.28	0.000	trxA	Rv1665	-2.89	0.000	pks11
Rv1475c	-2.15	0.003	acn	Rv1667c	2.53	0.000	Rv1667c
Rv1476	-2.28	0.001	Rv1476	Rv1668c	3.16	0.004	Rv1668c
Rv1480	2.06	0.033	Rv1480	Rv1670	-2.30	0.001	Rv1670
Rv1491c	2.16	0.000	Rv1491c	Rv1671	-2.11	0.000	Rv1671
Rv1495	2.27	0.003	mazF4	Rv1673c	-4.59	0.000	Rv1673c
Rv1496	2.51	0.000	Rv1496	Rv1674c	-4.75	0.000	Rv1674c
Rv1498c	-4.05	0.000	Rv1498c	Rv1679	2.17	0.001	fadE16
Rv1498A	-3.45	0.000	Rv1498A	Rv1683	3.02	0.000	Rv1683
Rv1500	-5.76	0.000	Rv1500	Rv1684	3.71	0.001	Rv1684
Rv1501	-4.86	0.000	Rv1501	Rv1685c	-2.66	0.000	Rv1685c
Rv1502	-3.62	0.000	Rv1502	Rv1686c	-3.34	0.000	Rv1686c
Rv1503c	-2.84	0.003	Rv1503c	MTB000054	-2.79	0.001	G2
Rv1504c	-2.30	0.003	Rv1504c	Rv1690	-2.42	0.007	lprJ
Rv1505c	-3.71	0.000	Rv1505c	Rv1692	2.00	0.000	Rv1692
Rv1506c	-3.90	0.000	Rv1506c	Rv1693	2.32	0.001	Rv1693
Rv1507c	-5.11	0.000	Rv1507c	Rv1694	3.25	0.000	tlyA
Rv1507A	-3.07	0.000	Rv1507A	Rv1695	3.38	0.000	ppnK
Rv1508c	-2.16	0.007	Rv1508c	Rv1696	3.68	0.000	recN
Rv1508A	-2.72	0.001	Rv1508A	Rv1712	2.99	0.001	cmk
Rv1509	-2.26	0.019	Rv1509	Rv1713	3.35	0.000	engA
Rv1515c	-2.04	0.017	Rv1515c	Rv1715	4.48	0.000	fadB3
Rv1521	-2.97	0.000	fadD25	Rv1716	4.79	0.000	Rv1716
Rv1528c	-3.31	0.000	papA4	Rv1717	5.44	0.000	Rv1717
Rv1535	-4.55	0.000	Rv1535	Rv1718	5.24	0.000	Rv1718
Rv1540	2.14	0.001	Rv1540	Rv1719	3.50	0.000	Rv1719
Rv1548c	-2.12	0.020	PPE21	Rv1724c	-2.38	0.003	Rv1724c
Rv1554	-2.20	0.025	frdC	Rv1725c	2.44	0.004	Rv1725c
Rv1562c	2.08	0.000	treZ	MTB000055	2.05	0.085	AS1726
Rv1567c	-2.52	0.001	Rv1567c	Rv1727	-2.09	0.066	Rv1727
Rv1568	2.56	0.013	bioA	Rv1732c	3.48	0.002	Rv1732c
Rv1569	3.74	0.001	bioF1	MTB000077	-2.95	0.004	MTS1338
Rv1570	3.48	0.000	bioD	Rv1734c	-3.27	0.001	Rv1734c
Rv1571	3.67	0.000	Rv1571	Rv1735c	-2.38	0.002	Rv1735c
Rv1575	2.44	0.000	Rv1575	Rv1738	-3.38	0.000	Rv1738
Rv1579c	-2.40	0.000	Rv1579c	Rv1739c	-2.04	0.001	Rv1739c
Rv1580c	-2.43	0.000	Rv1580c	Rv1740	-3.92	0.000	vapB34
Rv1581c	-2.95	0.000	Rv1581c	Rv1741	-3.19	0.000	vapC34
Rv1582c	-3.64	0.000	Rv1582c	Rv1751	2.31	0.000	Rv1751
Rv1584c	-2.56	0.000	Rv1584c	Rv1752	2.51	0.000	Rv1752
Rv1595	2.44	0.024	nadB	Rv1753c	-3.23	0.001	PPE24
Rv1596	2.17	0.043	nadC	Rv1754c	-3.87	0.000	Rv1754c
Rv1599	3.84	0.001	hisD	Rv1755c	-2.37	0.007	plcD
Rv1601	2.47	0.000	hisB	Rv1759c	2.19	0.000	wag22
Rv1604	2.05	0.000	impA	Rv1760	-2.46	0.000	Rv1760
Rv1605	2.85	0.000	hisF	Rv1761c	-2.02	0.001	Rv1761c
Rv1606	2.15	0.000	hisI	Rv1762c	-2.31	0.000	Rv1762c
Rv1608c	-2.06	0.001	bcpB	Rv1765A	-5.35	0.000	Rv1765A
Rv1610	3.11	0.000	Rv1610	Rv1766	-3.54	0.000	Rv1766
Rv1615	-2.60	0.000	Rv1615	Rv1773c	2.19	0.041	Rv1773c
Rv1620c	-3.91	0.000	cydC	Rv1779c	4.28	0.000	Rv1779c
Rv1622c	-2.41	0.000	cydB	Rv1780	-2.77	0.000	Rv1780
Rv1623c	-3.51	0.000	cydA	Rv1782	2.20	0.000	eccB5
Rv1624c	-2.91	0.000	Rv1624c	Rv1786	-3.17	0.000	Rv1786
MTB000023	-2.93	0.006	leuV	Rv1791	-7.93	0.000	PE19
Rv1626	-2.09	0.002	Rv1626	Rv1792	-6.82	0.000	esxM
Rv1631	2.21	0.015	coaE	Rv1793	-5.75	0.000	esxN
Rv1633	2.75	0.000	uvrB	Rv1794	-3.98	0.000	Rv1794
Rv1634	2.06	0.000	Rv1634	Rv1802	-2.43	0.002	PPE30
Rv1638	2.34	0.000	uvrA	Rv1804c	-2.21	0.013	Rv1804c
Rv1638A	-2.17	0.000	Rv1638A	Rv1807	-2.38	0.000	PPE31
Rv1639c	-2.33	0.000	Rv1639c	Rv1808	-2.36	0.000	PPE32
Rv1641	-3.87	0.000	infC	Rv1809	-3.30	0.000	PPE33
Rv1644	2.06	0.090	tsnR	Rv1810	-4.90	0.000	Rv1810
Rv1645c	-3.61	0.000	Rv1645c	Rv1813c	-3.48	0.000	Rv1813c

3.1 – mIHF Function

Rv1814	-2.88	0.001	erg3	Rv1981c	-2.18	0.001	nrdF1
Rv1819c	-3.62	0.000	bacA	Rv1987	-3.35	0.000	Rv1987
Rv1829	-2.09	0.007	Rv1829	Rv1989c	-2.18	0.000	Rv1989c
Rv1830	-2.41	0.000	Rv1830	Rv1995	-5.32	0.000	Rv1995
Rv1835c	2.72	0.000	Rv1835c	Rv2000	-2.71	0.000	Rv2000
Rv1840c	4.09	0.000	PE_PGRS34	Rv2001	-3.29	0.000	Rv2001
Rv1841c	2.30	0.000	Rv1841c	Rv2007c	-3.90	0.000	fdxA
Rv1842c	2.64	0.000	Rv1842c	Rv2011c	-3.02	0.000	Rv2011c
Rv1845c	2.32	0.046	blaR	Rv2012	-8.15	0.000	Rv2012
Rv1851	3.18	0.000	ureF	Rv2016	-6.05	0.000	Rv2016
Rv1852	2.74	0.000	ureG	Rv2017	-3.03	0.000	Rv2017
Rv1854c	2.41	0.000	ndh	Rv2018	-3.01	0.000	Rv2018
Rv1855c	-2.21	0.001	Rv1855c	Rv2019	-2.73	0.000	Rv2019
Rv1856c	-5.51	0.000	Rv1856c	Rv2021c	-2.79	0.001	Rv2021c
Rv1861	-2.96	0.000	Rv1861	Rv2022c	-2.32	0.003	Rv2022c
Rv1862	2.15	0.040	adhA	Rv2023A	-4.23	0.000	Rv2023A
Rv1867	2.50	0.000	Rv1867	Rv2027c	-2.75	0.000	dosT
Rv1868	2.69	0.000	Rv1868	Rv2028c	-2.83	0.000	Rv2028c
Rv1874	-2.88	0.000	Rv1874	Rv2029c	-2.30	0.000	pfbB
Rv1875	-2.64	0.000	Rv1875	Rv2030c	-2.05	0.001	Rv2030c
Rv1876	-4.69	0.000	bfrA	Rv2031c	-4.91	0.000	hspX
Rv1883c	-2.59	0.000	Rv1883c	Rv2033c	2.00	0.035	Rv2033c
Rv1885c	-3.29	0.000	Rv1885c	Rv2035	-2.24	0.002	Rv2035
Rv1886c	-3.77	0.000	fbpB	Rv2036	-3.63	0.000	Rv2036
Rv1888c	-3.17	0.000	Rv1888c	Rv2047c	3.88	0.000	Rv2047c
MTB000056	2.94	0.001	AS1890	Rv2048c	2.42	0.000	pkS12
Rv1891	-2.15	0.000	Rv1891	Rv2049c	2.15	0.076	Rv2049c
Rv1894c	-3.74	0.000	Rv1894c	Rv2050	-2.23	0.001	Rv2050
Rv1895	-5.78	0.000	Rv1895	Rv2056c	3.41	0.000	rpsN2
Rv1900c	2.03	0.004	lipJ	Rv2059	2.24	0.000	Rv2059
Rv1903	-2.44	0.008	Rv1903	Rv2062c	2.30	0.002	cobN
Rv1904	-2.06	0.000	Rv1904	Rv2063A	3.86	0.001	mazF7
Rv1908c	-2.44	0.000	katG	Rv2064	2.88	0.000	cobG
Rv1909c	-2.76	0.003	furA	Rv2065	2.45	0.037	cobH
Rv1910c	-3.12	0.000	Rv1910c	Rv2067c	-3.04	0.000	Rv2067c
Rv1917c	-2.14	0.019	PPE34	Rv2071c	2.64	0.000	cobM
Rv1918c	-2.17	0.023	PPE35	Rv2072c	3.68	0.000	cobL
Rv1919c	-2.15	0.001	Rv1919c	Rv2073c	2.32	0.044	Rv2073c
Rv1926c	-2.58	0.000	mpt63	Rv2076c	-2.64	0.000	Rv2076c
Rv1928c	-2.03	0.033	Rv1928c	Rv2079	-2.01	0.027	Rv2079
Rv1930c	2.61	0.008	Rv1930c	Rv2080	-2.43	0.003	lppJ
Rv1931c	3.60	0.000	Rv1931c	Rv2081c	-2.99	0.000	Rv2081c
Rv1933c	3.20	0.003	fadE18	Rv2090	2.50	0.012	Rv2090
Rv1934c	3.48	0.002	fadE17	Rv2092c	2.44	0.001	helY
Rv1935c	3.02	0.000	echA13	Rv2094c	-3.10	0.000	tatA
Rv1936	2.99	0.000	Rv1936	Rv2098c	2.62	0.003	PE_PGRS36
Rv1937	2.98	0.000	Rv1937	Rv2108	-2.30	0.022	PPE36
Rv1938	2.33	0.000	ephB	Rv2112c	2.08	0.000	dop
Rv1939	5.09	0.000	Rv1939	Rv2114	-2.53	0.000	Rv2114
Rv1940	2.87	0.001	ribA1	Rv2115c	-2.38	0.001	mpa
Rv1941	3.23	0.000	Rv1941	Rv2125	2.13	0.008	Rv2125
Rv1947	2.04	0.062	Rv1947	Rv2127	-2.04	0.000	ansP1
Rv1948c	-4.97	0.000	Rv1948c	Rv2132	-2.44	0.000	Rv2132
Rv1949c	-2.31	0.005	Rv1949c	Rv2133c	2.64	0.000	Rv2133c
Rv1950c	-2.81	0.001	Rv1950c	Rv2134c	2.14	0.017	Rv2134c
Rv1954c	-2.57	0.003	Rv1954c	Rv2136c	-2.34	0.000	Rv2136c
Rv1956	-2.04	0.002	higA	Rv2137c	-5.73	0.000	Rv2137c
Rv1957	-3.36	0.000	Rv1957	Rv2138	-2.53	0.000	lppL
Rv1958c	-3.34	0.000	Rv1958c	Rv2139	2.05	0.041	pyrD
Rv1962A	-2.55	0.000	vapB35	MTB000025	-5.72	0.000	leuU
Rv1964	2.10	0.000	yrbE3A	Rv2142c	-2.39	0.002	parE2
Rv1965	3.48	0.000	yrbE3B	Rv2142A	-2.16	0.004	parD2
Rv1966	5.20	0.000	mce3A	Rv2145c	-2.10	0.012	wag31
Rv1967	7.17	0.000	mce3B	Rv2149c	2.19	0.009	yfiH
Rv1968	6.76	0.000	mce3C	Rv2151c	3.79	0.000	ftsQ
Rv1969	5.30	0.000	mce3D	Rv2152c	2.74	0.000	murC
Rv1970	4.80	0.000	lprM	Rv2153c	4.52	0.000	murG
Rv1971	3.91	0.000	mce3F	Rv2154c	2.30	0.000	ftsW
Rv1972	3.71	0.000	Rv1972	Rv2156c	2.15	0.000	murX
Rv1979c	-2.31	0.003	Rv1979c	Rv2157c	3.34	0.000	murF
Rv1980c	-4.72	0.000	mpt64	Rv2158c	2.42	0.015	murE

mIHF Function – 3.1

Rv2159c	2.14	0.053	Rv2159c	Rv2375	-2.964	0.000	Rv2375
Rv2160A	2.44	0.027	Rv2160A	Rv2378c	3.140	0.000	mbtG
Rv2160c	2.34	0.015	Rv2160c	Rv2379c	3.654	0.000	mbtF
Rv2163c	2.18	0.000	pbpB	Rv2380c	2.459	0.000	mbtE
Rv2164c	3.05	0.000	Rv2164c	Rv2381c	2.955	0.000	mbtD
Rv2165c	2.30	0.000	Rv2165c	Rv2382c	2.704	0.000	mbtC
Rv2182c	-2.06	0.018	Rv2182c	Rv2383c	4.514	0.000	mbtB
Rv2185c	-2.37	0.009	TB16.3	Rv2387	-3.759	0.000	Rv2387
Rv2187	-3.61	0.000	fadD15	Rv2388c	2.247	0.011	hemN
Rv2190c	-2.94	0.000	Rv2190c	Rv2393	2.896	0.001	che1
Rv2191	3.97	0.000	Rv2191	Rv2401	-2.690	0.000	Rv2401
Rv2192c	2.53	0.014	trpD	Rv2406c	-4.136	0.000	Rv2406c
Rv2193	-2.82	0.000	ctaE	Rv2413c	3.269	0.004	Rv2413c
Rv2196	-2.16	0.000	qcrB	Rv2414c	2.465	0.033	Rv2414c
Rv2200c	-2.00	0.032	ctaC	Rv2425c	2.683	0.000	Rv2425c
Rv2204c	-3.62	0.000	Rv2204c	Rv2426c	3.353	0.000	Rv2426c
Rv2208	2.46	0.007	cobS	Rv2427c	2.835	0.000	proA
Rv2209	-2.28	0.001	Rv2209	Rv2427A	2.871	0.000	oxyR'
Rv2214c	2.18	0.000	ephD	Rv2429	-2.098	0.014	ahpD
Rv2216	2.60	0.004	Rv2216	Rv2432c	-2.558	0.004	Rv2432c
Rv2220	-3.24	0.000	glnA1	Rv2433c	2.807	0.006	Rv2433c
Rv2221c	2.37	0.007	glnE	Rv2435c	-2.258	0.014	Rv2435c
MTB000026	-3.68	0.001	rnpB	Rv2436	3.266	0.000	rbsK
Rv2227	-2.84	0.000	Rv2227	Rv2437	4.196	0.000	Rv2437
Rv2231c	2.31	0.003	cobC	Rv2439c	7.012	0.000	proB
Rv2240c	2.67	0.006	Rv2240c	Rv2440c	3.646	0.000	obg
Rv2244	-2.37	0.001	acpM	Rv2447c	2.095	0.003	folC
Rv2247	2.16	0.001	accD6	Rv2449c	2.735	0.000	Rv2449c
Rv2252	-4.24	0.000	Rv2252	Rv2451	-2.095	0.064	Rv2451
Rv2253	-3.40	0.000	Rv2253	Rv2452c	-3.320	0.002	Rv2452c
Rv2259	2.25	0.001	mscR	Rv2458	2.619	0.000	mmuM
Rv2262c	2.37	0.020	Rv2262c	Rv2459	-3.088	0.000	Rv2459
Rv2267c	-2.58	0.006	Rv2267c	Rv2460c	-3.238	0.000	clpP2
Rv2271	-4.88	0.000	Rv2271	Rv2461c	-3.319	0.000	clpP1
Rv2272	-4.17	0.000	Rv2272	MTB000031	-2.77	0.003	glyV
Rv2277c	-2.19	0.009	Rv2277c	Rv2466c	-2.396	0.000	Rv2466c
Rv2286c	-3.70	0.000	Rv2286c	Rv2468c	2.919	0.010	Rv2468c
Rv2289	-2.65	0.004	cdh	Rv2481c	-2.610	0.000	Rv2481c
Rv2291	2.12	0.001	sseB	Rv2485c	-5.885	0.000	lipQ
Rv2292c	-2.78	0.000	Rv2292c	MTB000032	-4.39	0.000	argW
Rv2293c	-2.54	0.000	Rv2293c	Rv2491	-2.229	0.020	Rv2491
Rv2301	-2.29	0.000	cut2	Rv2492	-2.444	0.014	Rv2492
Rv2302	-3.66	0.000	Rv2302	Rv2495c	2.765	0.000	bkdC
Rv2307D	2.63	0.000	Rv2307D	Rv2496c	2.410	0.000	bkdB
Rv2308	2.86	0.000	Rv2308	Rv2497c	2.766	0.000	bkdA
MTB000028	-2.53	0.001	metV	Rv2498c	3.551	0.000	citE
Rv2309A	-3.77	0.000	Rv2309A	Rv2504c	-2.364	0.000	scoA
Rv2311	-2.33	0.004	Rv2311	Rv2507	2.297	0.000	Rv2507
Rv2314c	2.90	0.000	Rv2314c	Rv2508c	2.102	0.030	Rv2508c
Rv2321c	2.28	0.025	rocD2	Rv2509	2.229	0.005	Rv2509
Rv2322c	2.81	0.002	rocD1	MTB000033	-2.10	0.011	hisT
Rv2326c	2.71	0.000	Rv2326c	Rv2513	-2.631	0.003	Rv2513
Rv2327	2.23	0.013	Rv2327	Rv2516c	-4.779	0.000	Rv2516c
Rv2335	3.05	0.000	cysE	Rv2517c	-3.968	0.000	Rv2517c
Rv2337c	-3.36	0.000	Rv2337c	Rv2518c	3.082	0.000	ldtB
Rv2338c	-5.70	0.000	moeW	MTB000034	-2.86	0.002	lysU
Rv2339	-4.08	0.000	mmpL9	Rv2520c	-2.505	0.000	Rv2520c
MTB000029	-2.88	0.002	asnT	Rv2522c	4.840	0.000	Rv2522c
Rv2345	2.56	0.000	Rv2345	Rv2523c	2.300	0.007	acpS
Rv2346c	-3.80	0.000	esxO	Rv2527	-3.733	0.000	vapC17
Rv2347c	-4.73	0.000	esxP	Rv2530A	-2.590	0.000	vapB39
Rv2348c	-5.04	0.000	Rv2348c	Rv2537c	2.453	0.005	aroD
Rv2349c	-2.20	0.000	plcC	Rv2538c	3.286	0.001	aroB
Rv2350c	-2.34	0.002	plcB	Rv2539c	3.223	0.004	aroK
Rv2351c	-3.96	0.000	plcA	Rv2540c	3.738	0.001	aroF
Rv2352c	-2.06	0.007	PPE38	Rv2544	-2.270	0.027	lppB
Rv2353c	-2.05	0.040	PPE39	Rv2548	2.325	0.004	vapC19
Rv2356c	-2.54	0.007	PPE40	Rv2552c	2.292	0.002	aroE
Rv2357c	-2.48	0.003	glyS	Rv2554c	3.822	0.002	Rv2554c
Rv2362c	2.24	0.008	recO	Rv2560	-3.976	0.000	Rv2560
Rv2372c	2.699	0.010	Rv2372c	MTB000068	-7.27	0.000	mpr11

3.1 – mIHF Function

Rv2561	-2.571	0.000	Rv2561	Rv2780	-2.846	0.000	ald
Rv2562	-2.184	0.001	Rv2562	Rv2782c	2.377	0.016	pepR
Rv2564	-2.266	0.000	glnQ	Rv2784c	-2.627	0.000	lppU
Rv2567	2.539	0.001	Rv2567	Rv2785c	-5.631	0.000	rpsO
Rv2570	2.997	0.000	Rv2570	Rv2789c	2.069	0.000	fadE21
Rv2577	2.359	0.000	Rv2577	Rv2791c	3.412	0.000	Rv2791c
Rv2578c	4.416	0.000	Rv2578c	Rv2792c	2.714	0.001	Rv2792c
Rv2580c	2.017	0.000	hisS	Rv2793c	3.193	0.001	truB
Rv2589	2.889	0.000	gabT	Rv2794c	2.730	0.000	pptT
Rv2592c	2.769	0.001	ruvB	Rv2797c	3.197	0.002	Rv2797c
Rv2593c	3.474	0.002	ruvA	Rv2802c	2.079	0.034	Rv2802c
Rv2598	2.066	0.050	Rv2598	Rv2805	2.012	0.018	Rv2805
Rv2601	2.681	0.007	speE	Rv2808	-3.821	0.000	Rv2808
Rv2601A	-2.074	0.000	vapB41	Rv2809	-2.096	0.012	Rv2809
Rv2609c	2.296	0.011	Rv2609c	Rv2811	3.850	0.000	Rv2811
Rv2611c	2.469	0.001	Rv2611c	Rv2812	6.297	0.000	Rv2812
Rv2614c	2.248	0.000	thrS	Rv2813	5.806	0.000	Rv2813
Rv2615c	2.320	0.024	PE_PGRS45	Rv2816c	-3.783	0.000	Rv2816c
Rv2616	2.056	0.077	Rv2616	Rv2818c	-2.293	0.007	Rv2818c
Rv2617c	-3.053	0.000	Rv2617c	Rv2819c	-2.608	0.002	Rv2819c
Rv2620c	-2.275	0.000	Rv2620c	Rv2821c	-2.704	0.000	Rv2821c
Rv2622	-4.583	0.000	Rv2622	Rv2822c	-3.203	0.001	Rv2822c
Rv2624c	-3.235	0.000	Rv2624c	Rv2826c	2.374	0.009	Rv2826c
Rv2625c	-3.999	0.000	Rv2625c	Rv2827c	2.008	0.074	Rv2827c
Rv2626c	-3.304	0.000	hrp1	Rv2828c	2.059	0.045	Rv2828c
Rv2632c	-5.919	0.000	Rv2632c	Rv2832c	3.177	0.000	ugpC
Rv2633c	-6.422	0.000	Rv2633c	Rv2833c	2.920	0.000	ugpB
Rv2636	-3.412	0.000	Rv2636	Rv2834c	2.475	0.000	ugpE
Rv2638	-3.846	0.000	Rv2638	Rv2838c	2.193	0.000	rbfA
Rv2641	-2.747	0.000	cadI	Rv2840c	-3.256	0.000	Rv2840c
MTB000035	-2.00	0.011	valT	Rv2843	2.904	0.001	Rv2843
MTB000036	-2.47	0.000	glyT	Rv2844	2.839	0.007	Rv2844
MTB000037	-4.11	0.000	cysU	Rv2848c	2.970	0.000	cobB
MTB000038	-4.26	0.000	valU	Rv2850c	3.351	0.000	Rv2850c
Rv2662	-2.678	0.000	Rv2662	Rv2858c	2.209	0.000	aldC
Rv2663	-4.199	0.000	Rv2663	Rv2859c	2.851	0.000	Rv2859c
Rv2664	-4.742	0.000	Rv2664	Rv2867c	2.342	0.016	Rv2867c
Rv2665	-3.448	0.000	Rv2665	Rv2873	-2.126	0.001	mpt83
Rv2672	4.096	0.000	Rv2672	Rv2877c	2.136	0.040	Rv2877c
Rv2677c	2.225	0.000	hemY	Rv2882c	-2.098	0.025	frr
Rv2678c	2.570	0.001	hemE	Rv2885c	2.816	0.000	Rv2885c
Rv2680	2.144	0.006	Rv2680	Rv2886c	2.513	0.000	Rv2886c
Rv2681	2.834	0.000	Rv2681	Rv2888c	2.012	0.023	amiC
Rv2687c	2.459	0.007	Rv2687c	Rv2890c	-2.579	0.000	rpsB
Rv2692	2.068	0.041	ceoC	Rv2896c	2.629	0.020	Rv2896c
Rv2696c	2.158	0.015	Rv2696c	Rv2897c	4.662	0.000	Rv2897c
Rv2697c	2.434	0.011	dut	Rv2899c	2.199	0.002	fdhD
Rv2699c	-2.922	0.000	Rv2699c	Rv2900c	2.021	0.000	fdhF
Rv2706c	-3.210	0.000	Rv2706c	Rv2901c	-2.374	0.005	Rv2901c
Rv2707	-2.210	0.002	Rv2707	Rv2904c	-2.256	0.000	rplS
Rv2720	3.477	0.000	lexA	Rv2908c	-2.103	0.000	Rv2908c
Rv2723	-3.411	0.000	Rv2723	Rv2909c	-2.941	0.000	rpsP
Rv2724c	-2.300	0.001	fadE20	Rv2910c	-4.251	0.000	Rv2910c
Rv2726c	2.640	0.000	dapF	Rv2911	2.170	0.042	dacB2
Rv2727c	2.074	0.004	miaA	Rv2912c	-2.675	0.003	Rv2912c
Rv2728c	2.710	0.009	Rv2728c	Rv2913c	-3.468	0.000	Rv2913c
Rv2732c	2.792	0.004	Rv2732c	Rv2917	2.830	0.000	Rv2917
Rv2733c	2.911	0.002	Rv2733c	Rv2918c	2.973	0.000	glnD
Rv2734	-3.531	0.000	Rv2734	Rv2923c	2.701	0.004	Rv2923c
Rv2735c	-2.032	0.033	Rv2735c	Rv2928	-6.131	0.000	tesA
Rv2738c	2.184	0.018	Rv2738c	Rv2929	-6.935	0.000	Rv2929
Rv2739c	2.705	0.017	Rv2739c	Rv2930	-3.053	0.000	fadD26
Rv2744c	-2.072	0.000	35kd_ag	Rv2934	3.261	0.000	ppsD
Rv2748c	3.361	0.000	ftsK	Rv2941	-3.050	0.000	fadD28
Rv2749	2.311	0.049	Rv2749	Rv2946c	3.053	0.000	pks1
Rv2753c	2.050	0.057	dapA	Rv2948c	-2.533	0.003	fadD22
Rv2760c	2.205	0.001	vapB42	Rv2949c	-4.572	0.000	Rv2949c
Rv2764c	-2.516	0.001	thyA	Rv2950c	-4.283	0.000	fadD29
Rv2775	-3.002	0.000	Rv2775	Rv2951c	-2.129	0.000	Rv2951c
Rv2777c	-2.059	0.001	Rv2777c	Rv2952	-2.173	0.003	Rv2952
Rv2779c	-5.013	0.000	Rv2779c	Rv2954c	-2.027	0.003	Rv2954c

mIHF Function – 3.1

Rv2956	-6.815	0.000	Rv2956	Rv3119	-4.403	0.000	moaE1
Rv2957	-3.103	0.000	Rv2957	Rv3120	-2.758	0.000	Rv3120
Rv2958c	-2.561	0.002	Rv2958c	Rv3121	-3.750	0.000	cyp141
Rv2959c	-4.840	0.000	Rv2959c	Rv3123	2.169	0.063	Rv3123
Rv2960c	-2.030	0.001	Rv2960c	Rv3124	-3.510	0.001	moaR1
Rv2963	-3.218	0.000	Rv2963	Rv3128c	-3.738	0.000	Rv3128c
Rv2966c	2.479	0.020	Rv2966c	Rv3129	-3.026	0.000	Rv3129
Rv2970A	-4.387	0.000	Rv2970A	Rv3136A	-2.754	0.000	Rv3136A
Rv2971	-2.993	0.000	Rv2971	Rv3141	2.617	0.005	fadB4
Rv2972c	2.612	0.000	Rv2972c	Rv3142c	-3.931	0.000	Rv3142c
Rv2974c	4.694	0.000	Rv2974c	Rv3143	-2.487	0.003	Rv3143
Rv2977c	2.157	0.000	thiL	Rv3149	2.065	0.023	nuoE
Rv2978c	4.579	0.000	Rv2978c	Rv3159c	-2.055	0.036	PPE53
Rv2979c	3.952	0.000	Rv2979c	Rv3160c	2.383	0.000	Rv3160c
Rv2980	6.145	0.000	Rv2980	Rv3162c	3.579	0.000	Rv3162c
Rv2981c	4.134	0.000	ddlA	Rv3163c	2.317	0.000	Rv3163c
Rv2982c	2.585	0.000	gpdA2	Rv3168	-8.279	0.000	Rv3168
Rv2987c	-3.207	0.000	leuD	Rv3169	-8.394	0.000	Rv3169
Rv2988c	-2.420	0.000	leuC	Rv3171c	2.151	0.046	hpx
MTB000039	-4.22	0.000	gluU	Rv3173c	2.920	0.002	Rv3173c
MTB000040	-4.63	0.000	glnU	Rv3180c	2.144	0.072	Rv3180c
Rv2992c	2.287	0.000	gltS	Rv3191c	-2.321	0.003	Rv3191c
Rv3006	2.143	0.000	lppZ	Rv3196A	-2.483	0.002	Rv3196A
Rv3011c	2.447	0.013	gatA	Rv3197A	-7.314	0.000	whiB7
Rv3012c	3.044	0.005	gatC	Rv3198A	-3.237	0.000	Rv3198A
Rv3013	2.052	0.092	Rv3013	Rv3201c	5.657	0.000	Rv3201c
Rv3014c	2.318	0.007	ligA	Rv3202c	7.061	0.000	Rv3202c
Rv3017c	-2.719	0.001	esxQ	Rv3203	2.304	0.024	lipV
Rv3018c	-3.330	0.000	PPE46	Rv3208A	-2.200	0.000	TB9.4
Rv3018A	-2.446	0.008	PE27A	Rv3216	-2.836	0.001	Rv3216
Rv3019c	-2.199	0.010	esxR	Rv3219	-4.154	0.000	whiB1
Rv3020c	-3.363	0.000	esxS	Rv3224A	2.136	0.067	Rv3224A
Rv3021c	-2.560	0.000	PPE47	Rv3224B	2.614	0.019	Rv3224B
Rv3022A	-3.796	0.000	PE29	Rv3227	2.215	0.016	aroA
Rv3026c	2.579	0.000	Rv3026c	Rv3229c	-9.816	0.000	desA3
Rv3031	3.388	0.000	Rv3031	Rv3230c	-2.155	0.002	Rv3230c
Rv3032	2.140	0.005	Rv3032	Rv3232c	-2.897	0.001	ppk2
Rv3035	2.288	0.023	Rv3035	Rv3235	-2.062	0.000	Rv3235
Rv3039c	2.312	0.029	echA17	Rv3241c	-2.942	0.000	Rv3241c
Rv3040c	2.775	0.001	Rv3040c	Rv3242c	3.541	0.001	Rv3242c
Rv3041c	2.343	0.011	Rv3041c	Rv3243c	2.040	0.008	Rv3243c
Rv3043c	-2.281	0.000	ctaD	Rv3244c	2.099	0.000	lpqB
Rv3047c	-2.237	0.001	Rv3047c	Rv3245c	3.178	0.000	mtrB
Rv3051c	-3.775	0.000	nrdE	Rv3250c	-2.662	0.000	rubB
Rv3052c	-2.179	0.000	nrdI	Rv3251c	-2.171	0.002	rubA
Rv3053c	-3.844	0.000	nrdH	Rv3255c	2.307	0.000	manA
Rv3056	-2.012	0.006	dinP	Rv3256c	3.544	0.002	Rv3256c
Rv3058c	-3.001	0.000	Rv3058c	Rv3260c	-2.246	0.000	whiB2
Rv3059	-2.369	0.000	cyp136	Rv3264c	2.345	0.000	manB
Rv3064c	-2.417	0.000	Rv3064c	Rv3265c	2.974	0.000	wbbL1
Rv3067	2.547	0.000	Rv3067	Rv3266c	3.459	0.001	rmlD
MTB000041	-2.18	0.022	alaU	Rv3268	2.656	0.012	Rv3268
Rv3072c	2.516	0.002	Rv3072c	Rv3271c	2.433	0.030	Rv3271c
Rv3074	5.069	0.000	Rv3074	Rv3276c	3.669	0.000	purK
Rv3080c	2.184	0.004	pknK	Rv3277	-3.003	0.000	Rv3277
Rv3092c	2.947	0.000	Rv3092c	Rv3281	-2.529	0.001	accE5
Rv3093c	3.829	0.000	Rv3093c	Rv3284	2.416	0.016	Rv3284
Rv3094c	3.519	0.000	Rv3094c	Rv3288c	-2.306	0.000	usfY
Rv3098A	-2.363	0.000	Rv3098A	Rv3290c	-3.594	0.000	lat
MTB000042	-2.79	0.000	ssr	Rv3291c	-4.727	0.000	lrpA
Rv3101c	-2.557	0.002	ftsX	Rv3296	2.222	0.000	lhr
Rv3103c	2.173	0.002	Rv3103c	Rv3297	3.721	0.000	nei
Rv3109	-5.979	0.000	moaA1	Rv3303c	2.942	0.000	lpdA
Rv3110	-3.794	0.000	moaB1	Rv3306c	4.080	0.000	amiB1
Rv3111	-3.062	0.000	moaC1	Rv3307	2.786	0.014	deoD
Rv3112	-4.580	0.000	moaD1	Rv3309c	2.512	0.002	upp
Rv3113	-2.770	0.001	Rv3113	Rv3323c	-2.399	0.004	moaX
Rv3114	-2.390	0.007	Rv3114	Rv3330	3.128	0.000	dacB1
Rv3116	-4.186	0.000	moeB2	Rv3331	2.133	0.009	sugI
Rv3117	-4.402	0.000	cysA3	Rv3332	5.015	0.000	nagA
Rv3118	-6.455	0.000	sseC1	Rv3334	-2.112	0.000	Rv3334

3.1 – mIHF Function

Rv3338	3.098	0.003	Rv3338	Rv3512	2.481	0.033	PE_PGRS56
Rv3339c	-2.030	0.017	icd1	Rv3513c	-3.717	0.000	fadD18
Rv3345c	2.270	0.057	PE_PGRS50	Rv3515c	-4.702	0.000	fadD19
Rv3351c	-2.571	0.002	Rv3351c	Rv3526	-6.782	0.000	kshA
Rv3352c	2.283	0.055	Rv3352c	Rv3527	-2.978	0.000	Rv3527
Rv3356c	3.025	0.002	folD	Rv3528c	-6.095	0.000	Rv3528c
Rv3357	-2.514	0.000	relJ	Rv3529c	-2.595	0.000	Rv3529c
Rv3358	-2.331	0.001	relK	Rv3530c	-3.329	0.000	Rv3530c
Rv3364c	2.326	0.028	Rv3364c	Rv3534c	2.075	0.002	hsaF
Rv3367	3.374	0.001	PE_PGRS51	Rv3535c	2.107	0.003	hsaG
Rv3368c	2.654	0.015	Rv3368c	Rv3536c	2.278	0.011	hsaE
Rv3369	2.219	0.027	Rv3369	Rv3540c	2.733	0.000	ltp2
Rv3370c	5.062	0.000	dnaE2	Rv3542c	2.139	0.000	Rv3542c
Rv3373	2.094	0.001	echA18	Rv3547	-3.336	0.000	ddn
Rv3374	2.603	0.001	echA18.1	Rv3558	-2.184	0.001	PPE64
Rv3375	3.508	0.000	amiD	Rv3563	2.185	0.007	fadE32
Rv3377c	-3.947	0.000	Rv3377c	Rv3564	2.720	0.005	fadE33
Rv3378c	-4.351	0.000	Rv3378c	Rv3566A	2.417	0.001	Rv3566A
Rv3384c	2.306	0.051	vapC46	Rv3577	2.456	0.002	Rv3577
Rv3387	6.238	0.000	Rv3387	Rv3579c	2.192	0.000	Rv3579c
Rv3388	4.255	0.000	PE_PGRS52	Rv3580c	2.323	0.000	cysS1
Rv3390	-3.511	0.000	lpqD	Rv3583c	-2.096	0.006	Rv3583c
Rv3393	2.854	0.001	iunH	Rv3585	3.479	0.000	radA
Rv3394c	5.168	0.000	Rv3394c	Rv3586	2.634	0.000	Rv3586
Rv3395c	2.566	0.000	Rv3395c	Rv3590c	2.623	0.020	PE_PGRS58
Rv3395A	-2.320	0.004	Rv3395A	Rv3597c	-4.808	0.000	lsr2
Rv3404c	-2.177	0.008	Rv3404c	Rv3598c	-2.539	0.003	lysS
Rv3411c	2.054	0.001	guaB2	Rv3605c	4.593	0.000	Rv3605c
Rv3412	-2.366	0.000	Rv3412	Rv3606c	3.563	0.000	folK
Rv3413c	3.393	0.000	Rv3413c	Rv3607c	4.087	0.000	folB
Rv3415c	-2.455	0.007	Rv3415c	Rv3608c	2.321	0.000	folP1
Rv3417c	-2.779	0.000	groEL1	Rv3612c	-13.078	0.000	Rv3612c
Rv3418c	-5.147	0.000	groES	Rv3613c	-17.256	0.000	Rv3613c
Rv3419c	4.641	0.000	gcp	Rv3614c	-19.410	0.000	espD
Rv3420c	3.663	0.000	rimI	Rv3615c	-21.588	0.000	espC
Rv3421c	4.597	0.000	Rv3421c	Rv3616c	-23.151	0.000	espA
Rv3422c	2.330	0.000	Rv3422c	Rv3617	-2.661	0.000	ephA
Rv3423c	2.156	0.000	alr	Rv3626c	2.294	0.000	Rv3626c
Rv3424c	-2.692	0.003	Rv3424c	Rv3627c	2.094	0.000	Rv3627c
Rv3428c	3.211	0.000	Rv3428c	Rv3628	-2.709	0.000	ppa
Rv3429	-2.859	0.002	PPE59	Rv3629c	-2.024	0.000	Rv3629c
Rv3433c	4.073	0.000	Rv3433c	Rv3631	-2.159	0.003	Rv3631
Rv3436c	2.220	0.007	glmS	Rv3633	-2.697	0.000	Rv3633
Rv3439c	5.571	0.000	Rv3439c	Rv3636	-2.090	0.003	Rv3636
Rv3440c	3.331	0.001	Rv3440c	Rv3640c	2.850	0.000	Rv3640c
Rv3442c	2.693	0.000	rpsI	MTB0000044	-3.89	0.000	thrU
Rv3443c	2.708	0.000	rplM	Rv3648c	-3.782	0.000	cspA
Rv3448	2.174	0.049	eccD4	Rv3652	2.390	0.026	PE_PGRS60
Rv3450c	2.051	0.014	eccB4	Rv3653	2.193	0.065	PE_PGRS61
Rv3453	-2.314	0.013	Rv3453	Rv3655c	2.492	0.029	Rv3655c
Rv3456c	-2.198	0.000	rplQ	Rv3656c	2.392	0.000	Rv3656c
Rv3457c	-2.701	0.000	rpoA	Rv3658c	3.093	0.007	Rv3658c
Rv3458c	-2.769	0.000	rpsD	Rv3659c	4.383	0.000	Rv3659c
Rv3460c	-3.094	0.000	rpsM	Rv3660c	2.601	0.022	Rv3660c
Rv3461c	-4.087	0.000	rpmJ	Rv3661	-2.019	0.015	Rv3661
Rv3462c	-5.040	0.000	infA	MTB0000078	-4.02	0.000	MTS2823
Rv3463	-3.052	0.000	Rv3463	Rv3662c	2.262	0.038	Rv3662c
Rv3465	-2.205	0.001	rmlC	Rv3663c	2.406	0.000	dppD
Rv3466	-2.812	0.000	Rv3466	Rv3664c	2.413	0.004	dppC
Rv3472	-3.350	0.000	Rv3472	Rv3666c	2.583	0.000	dppA
Rv3476c	-4.140	0.000	kgpP	Rv3676	-2.148	0.000	crp
Rv3477	-6.132	0.000	PE31	Rv3678A	2.008	0.067	Rv3678A
Rv3478	-2.831	0.000	PPE60	Rv3679	-3.017	0.000	Rv3679
Rv3482c	2.066	0.000	Rv3482c	Rv3680	-3.033	0.000	Rv3680
Rv3487c	-2.125	0.002	lipF	Rv3681c	-2.907	0.000	whiB4
Rv3491	-2.441	0.000	Rv3491	Rv3683	2.127	0.012	Rv3683
Rv3492c	-2.455	0.000	Rv3492c	Rv3687c	-2.121	0.024	rsfB
Rv3502c	-2.170	0.001	Rv3502c	Rv3694c	2.091	0.000	Rv3694c
Rv3503c	-2.497	0.000	fdxI	Rv3697c	-2.007	0.001	vapC48
Rv3504	-2.365	0.000	fadE26	Rv3700c	3.009	0.003	Rv3700c
Rv3511	-2.091	0.002	PE_PGRS55	Rv3702c	2.457	0.003	Rv3702c

mlHF Function – 3.1

Rv3704c	2.624	0.006	gshA	Rv3834c	2.412	0.001	serS
Rv3708c	2.225	0.004	asd	Rv3835	2.488	0.003	Rv3835
Rv3709c	2.839	0.000	ask	Rv3838c	2.530	0.001	pheA
Rv3711c	3.203	0.000	dnaQ	MTB000080	-6.16	0.000	MTS2975
Rv3712	3.539	0.000	Rv3712	Rv3846	-4.727	0.000	sodA
Rv3713	3.296	0.001	cobQ2	Rv3847	-2.612	0.000	Rv3847
MTB000059	-2.02	0.019	C8	Rv3849	-3.628	0.000	espR
Rv3727	-3.154	0.000	Rv3727	Rv3854c	-2.850	0.000	ethA
Rv3734c	-3.801	0.000	tgs2	Rv3855	-2.206	0.001	ethR
Rv3737	2.491	0.000	Rv3737	Rv3864	-2.303	0.004	espE
Rv3743c	2.227	0.010	ctpJ	Rv3865	-2.927	0.001	espF
Rv3746c	-2.244	0.007	PE34	Rv3874	-3.441	0.001	esxB
Rv3749c	-2.484	0.002	Rv3749c	Rv3875	-3.083	0.000	esxA
Rv3750c	-2.924	0.002	Rv3750c	Rv3883c	2.105	0.000	mycP1
MTB000047	-2.04	0.046	serX	Rv3886c	2.192	0.000	mycP2
Rv3752c	2.334	0.049	Rv3752c	Rv3890c	-2.988	0.000	esxC
Rv3756c	2.794	0.001	proZ	Rv3891c	-4.929	0.000	esxD
Rv3758c	2.524	0.000	proV	Rv3892c	3.164	0.001	PPE69
Rv3763	-4.109	0.000	lpqH	Rv3895c	2.265	0.000	eccB2
Rv3768	-3.324	0.000	Rv3768	Rv3896c	2.646	0.001	Rv3896c
MTB000048	-7.59	0.000	argU	Rv3902c	-3.497	0.000	Rv3902c
MTB000049	-9.17	0.000	serT	Rv3904c	2.716	0.010	esxE
Rv3774	3.092	0.000	echA21	Rv3905c	2.948	0.006	esxF
Rv3775	2.619	0.001	lipE	Rv3906c	2.150	0.020	Rv3906c
Rv3776	2.303	0.000	Rv3776	Rv3909	3.270	0.000	Rv3909
Rv3796	-2.431	0.010	Rv3796	Rv3910	2.145	0.001	Rv3910
Rv3798	-2.416	0.000	Rv3798	Rv3914	-3.301	0.000	trxC
Rv3802c	-2.002	0.003	Rv3802c	Rv3915	-2.291	0.000	Rv3915
Rv3818	-2.101	0.006	Rv3818	Rv3919c	2.291	0.000	gid
Rv3820c	-3.909	0.000	papA2	Rv3921c	-3.159	0.000	Rv3921c
Rv3822	-2.941	0.001	Rv3822	Rv3922c	-3.035	0.000	Rv3922c
Rv3824c	-2.608	0.004	papA1	Rv3923c	-2.945	0.000	rnpA
Rv3826	-4.211	0.000	fadD23	Rv3924c	-2.134	0.001	rpmH
Rv3831	-3.115	0.000	Rv3831				
Rv3833	-3.328	0.000	Rv3833				

mIHF peaks in exponentially growing cells

PeakName	Enr	FeatureName	FeatureGene	FC					
mlHFexp-001	6.09	dnaA	Rv0001	-5.7	mlHFexp-042	6.68	atpB	Rv1304	-3.0
mlHFexp-002	9.36	Rv0031	Rv0031	2.2	mlHFexp-042	6.68	atpE	Rv1305	-3.7
mlHFexp-002	9.36	bioF2	Rv0032	-2.8	mlHFexp-042	6.68	atpF	Rv1306	-2.4
mlHFexp-003	6.48	bioF2	Rv0032	-2.8	mlHFexp-043	6.67	murA	Rv1315	-1.3
mlHFexp-004	6.41	PPE1	Rv0096	-7.2	mlHFexp-043	6.67	mcr3	MTB000064	-7.8
mlHFexp-004	6.41	Rv0097	Rv0097	-10.5	mlHFexp-043	6.67	mcr3	MTB000064	-7.8
mlHFexp-005	8.84	Rv0111	Rv0111	-1.4	mlHFexp-044	7.04	rrs	MTB000019	-2.0
mlHFexp-005	8.84	gca	Rv0112	-5.1	mlHFexp-043	6.67	rrs	MTB000019	-2.0
mlHFexp-005	8.84	gmhA	Rv0113	-5.2	mlHFexp-045	7.3	rriI	MTB000020	-2.9
mlHFexp-005	8.84	gmhB	Rv0114	-2.7	mlHFexp-044	7.04	rriI	MTB000020	-2.9
mlHFexp-006	6.38	PE1	Rv0151c	-2.4	mlHFexp-045	7.3	rriF	MTB000021	-3.8
mlHFexp-007	7.63	Rv0193c	Rv0193c	-3.2	mlHFexp-045	7.3	ogt	Rv1316c	-1.3
mlHFexp-007	7.63	Rv0194	Rv0194	1.2	mlHFexp-046	8.92	Rv1356c	Rv1356c	-1.8
mlHFexp-008	7.05	Rv0245	Rv0245	1.4	mlHFexp-046	8.92	Rv1357c	Rv1357c	-1.2
mlHFexp-008	7.05	Rv0246	Rv0246	-1.9	mlHFexp-047	6.7	Rv1461	Rv1461	-1.1
mlHFexp-009	6.83	Rv0246	Rv0246	-1.9	mlHFexp-048	7.7	Rv1490	Rv1490	-1.1
mlHFexp-009	6.83	Rv0247c	Rv0247c	-3.3	mlHFexp-049	8.87	Rv1500	Rv1500	-5.8
mlHFexp-010	6.21	Rv0247c	Rv0247c	-3.3	mlHFexp-050	6.54	Rv1500	Rv1500	-5.8
mlHFexp-011	6.7	Rv0249c	Rv0249c	-4.6	mlHFexp-049	8.87	Rv1501	Rv1501	-4.9
mlHFexp-011	6.7	Rv0250c	Rv0250c	-4.7	mlHFexp-051	8.8	Rv1501	Rv1501	-4.9
mlHFexp-014	9.6	PPE5	Rv0304c	-2.9	mlHFexp-050	6.54	Rv1501	Rv1501	-4.9
mlHFexp-013	8.92	PPE5	Rv0304c	-2.9	mlHFexp-051	8.8	Rv1502	Rv1502	-3.6
mlHFexp-015	8.9	PPE5	Rv0304c	-2.9	mlHFexp-051	8.8	Rv1503c	Rv1503c	-2.8
mlHFexp-012	7.42	PPE5	Rv0304c	-2.9	mlHFexp-052	6.71	Rv1503c	Rv1503c	-2.8
mlHFexp-016	9.72	PPE6	Rv0305c	-3.6	mlHFexp-051	8.8	Rv1504c	Rv1504c	-2.3
mlHFexp-015	8.9	PPE6	Rv0305c	-3.6	mlHFexp-053	7.75	Rv1504c	Rv1504c	-2.3
mlHFexp-017	6.33	PPE6	Rv0305c	-3.6	mlHFexp-052	6.71	Rv1504c	Rv1504c	-2.3
mlHFexp-020	8.8	PPE8	Rv0355c	-1.6	mlHFexp-053	7.75	Rv1505c	Rv1505c	-3.7
mlHFexp-018	6.32	PPE8	Rv0355c	-1.6	mlHFexp-053	7.75	Rv1506c	Rv1506c	-3.9
mlHFexp-019	6.26	PPE8	Rv0355c	-1.6	mlHFexp-053	7.75	Rv1507c	Rv1507c	-5.1
mlHFexp-021	9.61	mmpS1	Rv0403c	-2.7	mlHFexp-054	6.75	Rv1507c	Rv1507c	-5.1
mlHFexp-021	9.61	fadD30	Rv0404	-5.3	mlHFexp-054	6.75	Rv1507A	Rv1507A	-3.1
mlHFexp-022	9.58	fadD30	Rv0404	-5.3	mlHFexp-055	7.71	Rv1508A	Rv1508A	-2.7
mlHFexp-023	8.6	fadD30	Rv0404	-5.3	mlHFexp-055	7.71	Rv1509	Rv1509	-2.3
mlHFexp-023	8.6	pkS6	Rv0405	-3.7	mlHFexp-056	8.11	fadD25	Rv1521	-3.0
mlHFexp-024	7.23	pkS6	Rv0405	-3.7	mlHFexp-057	7.37	mmpL12	Rv1522c	-1.6
mlHFexp-024	7.23	Rv0406c	Rv0406c	1.3	mlHFexp-058	7.11	pkS5	Rv1527c	-1.1
mlHFexp-025	10.31	mmpL4	Rv0450c	-2.7	mlHFexp-058	7.11	papA4	Rv1528c	-3.3
mlHFexp-027	8.77	mmpL4	Rv0450c	-2.7	mlHFexp-059	9.21	PPE21	Rv1548c	-2.1
mlHFexp-026	8.14	mmpL4	Rv0450c	-2.7	mlHFexp-060	8.03	pkS9	Rv1664	-1.3
mlHFexp-027	8.77	mmpS4	Rv0451c	-3.4	mlHFexp-061	7.5	pkS9	Rv1664	-1.3
mlHFexp-027	8.77	Rv0452	Rv0452	2.2	mlHFexp-061	7.5	pkS11	Rv1665	-2.9
mlHFexp-028	6.91	Rv0455c	Rv0455c	-2.2	mlHFexp-063	7.28	PPE24	Rv1753c	-3.2
mlHFexp-029	6.06	mazF1	Rv0456A	1.6	mlHFexp-062	7.26	PPE24	Rv1753c	-3.2
mlHFexp-029	6.06	mazE1	Rv0456B	-2.4	mlHFexp-064	6.28	PPE24	Rv1753c	-3.2
mlHFexp-029	6.06	Rv0457c	Rv0457c	-1.3	mlHFexp-065	6.44	Rv1754c	Rv1754c	-3.9
mlHFexp-030	6.8	umaA	Rv0469	-5.2	mlHFexp-064	6.28	Rv1754c	Rv1754c	-3.9
mlHFexp-031	6.21	Rv0488	Rv0488	-1.2	mlHFexp-066	6.48	lppT	Rv1799	-1.5
mlHFexp-031	6.21	gpm1	Rv0489	-1.0	mlHFexp-067	6.31	PPE28	Rv1800	-1.0
mlHFexp-033	9.87	mmpL2	Rv0507	-3.4	mlHFexp-069	9.18	PPE34	Rv1917c	-2.1
mlHFexp-032	7.9	mmpL2	Rv0507	-3.4	mlHFexp-068	6.4	PPE34	Rv1917c	-2.1
mlHFexp-034	9.02	PPE12	Rv0755c	-4.0	mlHFexp-070	6.32	PPE34	Rv1917c	-2.1
mlHFexp-035	7.2	Rv0805	Rv0805	-1.8	mlHFexp-072	9.42	PPE35	Rv1918c	-2.2
mlHFexp-035	7.2	cpsY	Rv0806c	-3.1	mlHFexp-071	8.31	PPE35	Rv1918c	-2.2
mlHFexp-036	11.35	pknD	Rv0931c	1.2	mlHFexp-070	6.32	PPE35	Rv1918c	-2.2
mlHFexp-036	11.35	pstS2	Rv0932c	-1.8	mlHFexp-073	6.96	Rv1948c	Rv1948c	-5.0
mlHFexp-037	8.99	Rv0986	Rv0986	-2.9	mlHFexp-073	6.96	Rv1949c	Rv1949c	-2.3
mlHFexp-038	10.07	Rv0987	Rv0987	-3.6	mlHFexp-074	6.15	mce3R	Rv1963c	1.4
mlHFexp-039	9.42	Rv0987	Rv0987	-3.6	mlHFexp-075	7.22	Rv2023c	Rv2023c	-1.9
mlHFexp-037	8.99	Rv0987	Rv0987	-3.6	mlHFexp-075	7.22	Rv2023A	Rv2023A	-4.2
mlHFexp-039	9.42	Rv0988	Rv0988	-4.5	mlHFexp-075	7.22	Rv2024c	Rv2024c	1.2
mlHFexp-039	9.42	grcC2	Rv0989c	-1.9	mlHFexp-076	6.51	Rv2067c	Rv2067c	-3.0
mlHFexp-040	6.3	PE9	Rv1088	-1.8	mlHFexp-076	6.51	blaC	Rv2068c	1.4
mlHFexp-040	6.3	PE10	Rv1089	-1.3	mlHFexp-077	7.74	PE22	Rv2107	-1.1
mlHFexp-040	6.3	celA2a	Rv1089A	1.0	mlHFexp-077	7.74	PPE36	Rv2108	-2.3
mlHFexp-041	7.08	Rv1116	Rv1116	-2.3	mlHFexp-078	6.9	Rv2293c	Rv2293c	-2.5
mlHFexp-041	7.08	Rv1116A	Rv1116A	-1.3	mlHFexp-078	6.9	Rv2294	Rv2294	-1.5
mlHFexp-042	6.68	Rv1303	Rv1303	-2.0	mlHFexp-079	8.44	Rv2307A	Rv2307A	-1.4

miHF Function – 3.1

miHFexp-079	8.44	Rv2307B	Rv2307B	-1.5	miHFexp-113	7.09	sseC1	Rv3118	-6.5
miHFexp-079	8.44	Rv2307D	Rv2307D	2.6	miHFexp-113	7.09	moaE1	Rv3119	-4.4
miHFexp-080	6.25	Rv2308	Rv2308	2.9	miHFexp-113	7.09	Rv3120	Rv3120	-2.8
miHFexp-081	8.49	Rv2336	Rv2336	-2.0	miHFexp-114	7.18	cyp141	Rv3121	-3.7
miHFexp-081	8.49	Rv2337c	Rv2337c	-3.4	miHFexp-115	6.71	moaR1	Rv3124	-3.5
miHFexp-082	9.97	moaW	Rv2338c	-5.7	miHFexp-116	9.85	nuoN	Rv3158	-1.1
miHFexp-083	9.29	mmpL9	Rv2339	-4.1	miHFexp-116	9.85	PPE53	Rv3159c	-2.1
miHFexp-084	7.61	mmpL9	Rv2339	-4.1	miHFexp-117	6.58	moaB1	Rv3206c	-1.8
miHFexp-085	7.18	plcA	Rv2351c	-4.0	miHFexp-117	6.58	Rv3207c	Rv3207c	-1.2
miHFexp-086	9.9	PPE38	Rv2352c	-2.1	miHFexp-120	8.3	PPE54	Rv3343c	1.1
miHFexp-086	9.9	PPE39	Rv2353c	-2.0	miHFexp-118	8.22	PPE54	Rv3343c	1.1
miHFexp-086	9.9	Rv2354	Rv2354	1.2	miHFexp-122	8.05	PPE54	Rv3343c	1.1
miHFexp-087	8.94	PPE40	Rv2356c	-2.5	miHFexp-119	7.75	PPE54	Rv3343c	1.1
miHFexp-088	6.62	glyS	Rv2357c	-2.5	miHFexp-121	7.07	PPE54	Rv3343c	1.1
miHFexp-089	6.4	Rv2451	Rv2451	-2.1	miHFexp-125	7.47	PPE55	Rv3347c	-1.3
miHFexp-089	6.4	Rv2452c	Rv2452c	-3.3	miHFexp-124	6.82	PPE55	Rv3347c	-1.3
miHFexp-089	6.4	mobA	Rv2453c	1.6	miHFexp-123	6.61	PPE55	Rv3347c	-1.3
miHFexp-090	8.18	Rv2491	Rv2491	-2.2	miHFexp-128	7.2	PPE56	Rv3350c	-1.3
miHFexp-090	8.18	Rv2492	Rv2492	-2.4	miHFexp-127	6.92	PPE56	Rv3350c	-1.3
miHFexp-091	6.64	lppA	Rv2543	-1.7	miHFexp-126	6.37	PPE56	Rv3350c	-1.3
miHFexp-091	6.64	lppB	Rv2544	-2.3	miHFexp-129	6.07	Rv3376	Rv3376	-1.8
miHFexp-092	6.07	lppB	Rv2544	-2.3	miHFexp-130	7.56	Rv3377c	Rv3377c	-3.9
miHFexp-092	6.07	vapB18	Rv2545	-1.5	miHFexp-129	6.07	Rv3377c	Rv3377c	-3.9
miHFexp-093	6.63	PPE42	Rv2608	-1.5	miHFexp-130	7.56	Rv3378c	Rv3378c	-4.4
miHFexp-095	9.83	Rv2734	Rv2734	-3.5	miHFexp-131	6.24	Rv3378c	Rv3378c	-4.4
miHFexp-094	7.93	Rv2734	Rv2734	-3.5	miHFexp-131	6.24	dxs2	Rv3379c	1.4
miHFexp-095	9.83	Rv2735c	Rv2735c	-2.0	miHFexp-132	8.2	idsB	Rv3383c	-1.9
miHFexp-095	9.83	recX	Rv2736c	-1.0	miHFexp-133	5.55	guaB2	Rv3415c	-2.5
miHFexp-096	6.44	clgR	Rv2745c	1.2	miHFexp-133	5.55	Rv3412	Rv3416	-1.7
miHFexp-096	6.44	pgsA3	Rv2746c	-1.5	miHFexp-134	7.57	Rv3424c	Rv3424c	-2.7
miHFexp-096	6.44	argA	Rv2747	-1.9	miHFexp-134	7.57	PPE57	Rv3425	-1.6
miHFexp-097	6.85	Rv2816c	Rv2816c	-3.8	miHFexp-134	7.57	PPE58	Rv3426	-1.4
miHFexp-098	6.82	Rv2816c	Rv2816c	-3.8	miHFexp-136	6.27	PPE59	Rv3429	-2.9
miHFexp-099	7.56	tesA	Rv2928	-6.1	miHFexp-136	6.27	Rv3430c	Rv3430c	-1.1
miHFexp-099	7.56	Rv2929	Rv2929	-6.9	miHFexp-138	6.99	Rv3528c	Rv3528c	-6.1
miHFexp-100	8.47	fadD26	Rv2930	-3.1	miHFexp-137	6.58	Rv3528c	Rv3528c	-6.1
miHFexp-101	6.24	drvB	Rv2937	-1.0	miHFexp-139	7.28	PPE62	Rv3533c	-1.1
miHFexp-101	6.24	drvC	Rv2938	-1.1	miHFexp-140	7.4	lsr2	Rv3597c	-4.8
miHFexp-102	7.66	fadD28	Rv2941	-3.1	miHFexp-140	7.4	lysS	Rv3598c	-2.5
miHFexp-103	6.52	fadD28	Rv2941	-3.1	miHFexp-141	6.08	espA	Rv3616c	-23.2
miHFexp-103	6.52	mmpL7	Rv2942	-1.2	miHFexp-142	7.57	papA2	Rv3820c	-3.9
miHFexp-104	6.55	fadD22	Rv2948c	-2.5	miHFexp-142	7.57	Rv3821	Rv3821	-1.4
miHFexp-105	6.36	fadD22	Rv2948c	-2.5	miHFexp-143	6.43	Rv3822	Rv3822	-2.9
miHFexp-105	6.36	Rv2949c	Rv2949c	-4.6	miHFexp-144	7.22	mmpL8	Rv3823c	-2.0
miHFexp-106	7.5	fadD29	Rv2950c	-4.3	miHFexp-145	7.39	papA1	Rv3824c	-2.6
miHFexp-105	6.36	fadD29	Rv2950c	-4.3	miHFexp-145	7.39	pkS2	Rv3825c	1.8
miHFexp-107	6.06	Rv2951c	Rv2951c	-2.1	miHFexp-146	8.34	fadD23	Rv3826	-4.2
miHFexp-108	7.46	Rv2955c	Rv2955c	1.4	miHFexp-147	6.28	Rv3840	Rv3840	1.0
miHFexp-108	7.46	Rv2956	Rv2956	-6.8	miHFexp-147	6.28	bfrB	Rv3841	-1.9
miHFexp-109	6.15	Rv2957	Rv2957	-3.1	miHFexp-149	7.31	Rv3863	Rv3863	-1.1
miHFexp-110	7.16	Rv2958c	Rv2958c	-2.6	miHFexp-149	7.31	espE	Rv3864	-2.3
miHFexp-110	7.16	Rv2959c	Rv2959c	-4.8	miHFexp-150	6.81	espK	Rv3879c	1.5
miHFexp-110	7.16	Rv2960c	Rv2960c	-2.0	miHFexp-151	6.55	eccD2	Rv3887c	-1.6
miHFexp-111	9.3	Rv3108	Rv3108	-1.9	miHFexp-151	6.55	Rv3888c	Rv3888c	-2.0
miHFexp-111	9.3	moaA1	Rv3109	-6.0	miHFexp-152	6.46	Rv3900c	Rv3900c	-1.4
miHFexp-111	9.3	moaB1	Rv3110	-3.8	miHFexp-153	7.45	Rv3902c	Rv3902c	-3.5
miHFexp-111	9.3	moaC1	Rv3111	-3.1	miHFexp-153	7.45	Rv3903c	Rv3903c	-1.5
miHFexp-111	9.3	moaD1	Rv3112	-4.6					
miHFexp-111	9.3	Rv3113	Rv3113	-2.8					
miHFexp-112	7.37	moaB2	Rv3116	-4.2					

Reported features are 500 bp within the miHF binding site. Enr = enrichment of miHF peak. Fold change from RNA-seq miHF cKD + ATc is reported in column FC.

mIHF peaks in mIHF-depleted cells

PeakName	Enr	Feature Name	Feature Gene	FC					
mIHFdepl-001	6.04	dnaA	Rv0001	-5.7	mIHFdepl-035	9.84	mcr3	MTB000064	-7.8
mIHFdepl-002	7.13	Rv0010c	Rv0010c	-3.8	mIHFdepl-035	9.84	rrs	MTB000019	-2.0
mIHFdepl-002	7.13	Rv0011c	Rv0011c	-4.1	mIHFdepl-035	9.84	rriI	MTB000020	-2.9
mIHFdepl-002	7.13	Rv0012	Rv0012	-1.3	mIHFdepl-035	9.84	rriF	MTB000021	-3.8
mIHFdepl-003	6.44	bioF2	Rv0032	-2.8	mIHFdepl-035	9.84	ogt	Rv1316c	-1.3
mIHFdepl-004	7.67	PPE1	Rv0096	-7.2	mIHFdepl-036	6.84	Rv1353c	Rv1353c	1.5
mIHFdepl-004	7.67	Rv0097	Rv0097	-10.5	mIHFdepl-036	6.84	Rv1354c	Rv1354c	1.8
mIHFdepl-005	7.17	gca	Rv0112	-5.1	mIHFdepl-037	7.4	Rv1356c	Rv1356c	-1.8
mIHFdepl-005	7.17	gmhA	Rv0113	-5.2	mIHFdepl-038	6.04	PPE19	Rv1361c	-1.6
mIHFdepl-006	6.2	treS	Rv0126	1.2	mIHFdepl-038	6.04	Rv1362c	Rv1362c	-2.3
mIHFdepl-007	6.29	mce1A	Rv0169	-3.7	mIHFdepl-039	6.41	Rv1374c	Rv1374c	-3.5
mIHFdepl-007	6.29	mce1B	Rv0170	-3.6	mIHFdepl-039	6.41	MTS1082	MTB000076	-3.6
mIHFdepl-008	6.13	fadE5	Rv0244c	-1.4	mIHFdepl-039	6.41	MTS1082	MTB000076	-3.6
mIHFdepl-008	6.13	Rv0245	Rv0245	1.4	mIHFdepl-039	6.41	Rv1375	Rv1375	-2.3
mIHFdepl-009	6.17	PPE7	Rv0354c	-3.6	mIHFdepl-040	11.2	Rv1461	Rv1461	-1.1
mIHFdepl-009	6.17	PPE8	Rv0355c	-1.6	mIHFdepl-041	6.23	trxA	Rv1470	-3.3
mIHFdepl-010	6.13	PPE8	Rv0355c	-1.6	mIHFdepl-041	6.23	trxB1	Rv1471	-1.4
mIHFdepl-011	6.86	mmpL1	Rv0402c	-3.6	mIHFdepl-042	7.56	Rv1490	Rv1490	-1.1
mIHFdepl-012	8.3	fadD30	Rv0404	-5.3	mIHFdepl-043	6.91	Rv1498c	Rv1498c	-4.1
mIHFdepl-012	8.3	pkS6	Rv0405	-3.7	mIHFdepl-044	8.23	Rv1500	Rv1500	-5.8
mIHFdepl-013	11.17	pkS6	Rv0405	-3.7	mIHFdepl-044	8.23	Rv1501	Rv1501	-4.9
mIHFdepl-013	11.17	Rv0406c	Rv0406c	1.3	mIHFdepl-045	6.99	Rv1502	Rv1502	-3.6
mIHFdepl-014	7.17	mmpL4	Rv0450c	-2.7	mIHFdepl-045	6.99	Rv1503c	Rv1503c	-2.8
mIHFdepl-015	7.07	mmpL4	Rv0450c	-2.7	mIHFdepl-046	6.92	Rv1520	Rv1520	-1.3
mIHFdepl-015	7.07	mmpS4	Rv0451c	-3.4	mIHFdepl-046	6.92	fadD25	Rv1521	-3.0
mIHFdepl-016	9.37	fadB2	Rv0468	-2.8	mIHFdepl-047	6.02	pkS5	Rv1527c	-1.1
mIHFdepl-016	9.37	umaA	Rv0469	-5.2	mIHFdepl-048	7.4	PPE21	Rv1548c	-2.1
mIHFdepl-017	6.14	serB1	Rv0505c	1.1	mIHFdepl-049	6.2	pkS9	Rv1664	-1.3
mIHFdepl-017	6.14	mmpS2	Rv0506	-4.1	mIHFdepl-049	6.2	pkS11	Rv1665	-2.9
mIHFdepl-017	6.14	mmpL2	Rv0507	-3.4	mIHFdepl-050	6.09	PPE24	Rv1753c	-3.2
mIHFdepl-018	8.47	mmpL2	Rv0507	-3.4	mIHFdepl-051	6.54	PPE33	Rv1809	-3.3
mIHFdepl-019	7.5	mmaA4	Rv0642c	-3.0	mIHFdepl-051	6.54	Rv1810	Rv1810	-4.9
mIHFdepl-019	7.5	mmaA3	Rv0643c	-3.2	mIHFdepl-052	6.44	bfrA	Rv1876	-4.7
mIHFdepl-020	6.46	lpqP	Rv0671	2.1	mIHFdepl-052	6.44	Rv1877	Rv1877	1.3
mIHFdepl-020	6.46	fadE8	Rv0672	-1.6	mIHFdepl-053	6.3	nanT	Rv1902c	-1.9
mIHFdepl-021	8.08	mmpL5	Rv0676c	1.1	mIHFdepl-053	6.3	Rv1903	Rv1903	-2.4
mIHFdepl-022	6.32	Rv0739	Rv0739	-3.7	mIHFdepl-053	6.3	Rv1904	Rv1904	-2.1
mIHFdepl-022	6.32	Rv0740	Rv0740	-2.6	mIHFdepl-054	6.83	katG	Rv1908c	-2.4
mIHFdepl-023	7.96	PPE12	Rv0755c	-4.0	mIHFdepl-054	6.83	furA	Rv1909c	-2.8
mIHFdepl-024	7.11	PPE12	Rv0755c	-4.0	mIHFdepl-055	6.79	PPE35	Rv1918c	-2.2
mIHFdepl-024	7.11	Rv0755A	Rv0755A	-3.4	mIHFdepl-056	6.37	PPE35	Rv1918c	-2.2
mIHFdepl-024	7.11	thrV	MTB000009	-2.5	mIHFdepl-056	6.37	Rv1919c	Rv1919c	-2.2
mIHFdepl-024	7.11	Rv0756c	Rv0756c	-1.4	mIHFdepl-057	6.87	echA13	Rv1935c	3.0
mIHFdepl-025	6.09	rpfA	Rv0867c	-4.2	mIHFdepl-057	6.87	Rv1936	Rv1936	3.0
mIHFdepl-025	6.09	moaD2	Rv0868c	-2.1	mIHFdepl-058	8.79	Rv1948c	Rv1948c	-5.0
mIHFdepl-026	6.13	Rv0877	Rv0877	-1.4	mIHFdepl-058	8.79	Rv1949c	Rv1949c	-2.3
mIHFdepl-026	6.13	PPE13	Rv0878c	-1.6	mIHFdepl-059	6.14	Rv1954c	Rv1954c	-2.6
mIHFdepl-027	6.56	Rv0950c	Rv0950c	-4.3	mIHFdepl-059	6.14	higB	Rv1955	-1.8
mIHFdepl-027	6.56	sucC	Rv0951	-1.1	mIHFdepl-059	6.14	higA	Rv1956	-2.0
mIHFdepl-028	7.5	Rv0986	Rv0986	-2.9	mIHFdepl-059	6.14	Rv1957	Rv1957	-3.4
mIHFdepl-028	7.5	Rv0987	Rv0987	-3.6	mIHFdepl-059	6.14	Rv1958c	Rv1958c	-3.3
mIHFdepl-028	7.5	Rv0988	Rv0988	-4.5	mIHFdepl-060	6.1	mce3R	Rv1963c	1.4
mIHFdepl-028	7.5	grcC2	Rv0989c	-1.9	mIHFdepl-061	6.35	nrdF1	Rv1981c	-2.2
mIHFdepl-029	8.04	Rv1057	Rv1057	7.2	mIHFdepl-062	10.85	Rv2023c	Rv2023c	-1.9
mIHFdepl-030	6.06	vapC32	Rv1114	-1.4	mIHFdepl-062	10.85	Rv2023A	Rv2023A	-4.2
mIHFdepl-030	6.06	Rv1115	Rv1115	-2.2	mIHFdepl-062	10.85	Rv2024c	Rv2024c	1.2
mIHFdepl-031	6.09	PPE16	Rv1135c	-2.2	mIHFdepl-063	6.3	Rv2067c	Rv2067c	-3.0
mIHFdepl-032	6.01	Rv1179c	Rv1179c	1.5	mIHFdepl-064	8.97	PE22	Rv2107	-1.1
mIHFdepl-032	6.01	pkS3	Rv1180	-1.4	mIHFdepl-064	8.97	PPE36	Rv2108	-2.3
mIHFdepl-033	6.75	Rv1220c	Rv1220c	2.5	mIHFdepl-065	9.69	Rv2136c	Rv2136c	-2.3
mIHFdepl-033	6.75	sigE	Rv1221	-4.5	mIHFdepl-065	9.69	Rv2137c	Rv2137c	-5.7
mIHFdepl-034	6.45	atpB	Rv1304	-3.0	mIHFdepl-065	9.69	lppL	Rv2138	-2.5
mIHFdepl-034	6.45	atpE	Rv1305	-3.7	mIHFdepl-066	11.86	leuU	MTB000025	-5.7
mIHFdepl-034	6.45	atpF	Rv1306	-2.4	mIHFdepl-066	11.86	parE2	Rv2142c	-2.4
mIHFdepl-035	9.84	murA	Rv1315	-1.3	mIHFdepl-066	11.86	parD2	Rv2142A	-2.2
mIHFdepl-035	9.84	mcr3	MTB000064	-7.8	mIHFdepl-067	6.58	wag31	Rv2145c	-2.1

mlHF Function – 3.1

mlHFdepl-067	6.58	Rv2146c	Rv2146c	-1.3	mlHFdepl-093	8.9	Rv2960c	Rv2960c	-2.0
mlHFdepl-068	6.85	ctaC	Rv2200c	-2.0	mlHFdepl-094	6.95	PPE46	Rv3018c	-3.3
mlHFdepl-069	6.7	Rv2267c	Rv2267c	-2.6	mlHFdepl-095	8.13	PPE47	Rv3021c	-2.6
mlHFdepl-070	6	cyp128	Rv2268c	-1.2	mlHFdepl-096	6.44	Rv3108	Rv3108	-1.9
mlHFdepl-071	6.01	Rv2271	Rv2271	-4.9	mlHFdepl-096	6.44	moaA1	Rv3109	-6.0
mlHFdepl-071	6.01	Rv2272	Rv2272	-4.2	mlHFdepl-097	6.37	cyp141	Rv3121	-3.7
mlHFdepl-072	7.58	cdh	Rv2289	-2.6	mlHFdepl-098	6.07	moaR1	Rv3124	-3.5
mlHFdepl-072	7.58	lppO	Rv2290	-1.2	mlHFdepl-099	10.56	PPE53	Rv3159c	-2.1
mlHFdepl-073	8.58	mmpL9	Rv2339	-4.1	mlHFdepl-100	7.4	Rv3218	Rv3218	-1.4
mlHFdepl-074	7.37	PPE39	Rv2353c	-2.0	mlHFdepl-100	7.4	whiB1	Rv3219	-4.2
mlHFdepl-075	7.53	PPE40	Rv2356c	-2.5	mlHFdepl-100	7.4	Rv3220c	Rv3220c	1.6
mlHFdepl-076	7.17	glyS	Rv2357c	-2.5	mlHFdepl-101	6.36	Rv3231c	Rv3231c	1.5
mlHFdepl-077	7.76	Rv2387	Rv2387	-3.8	mlHFdepl-101	6.36	ppk2	Rv3232c	-2.9
mlHFdepl-078	9.58	Rv2452c	Rv2452c	-3.3	mlHFdepl-102	6.45	whiB2	Rv3260c	-2.2
mlHFdepl-078	9.58	mobA	Rv2453c	1.6	mlHFdepl-102	6.45	fbiA	Rv3261	1.4
mlHFdepl-079	6.14	lipQ	Rv2485c	-5.9	mlHFdepl-103	6.17	PPE54	Rv3343c	1.1
mlHFdepl-079	6.14	argW	MTB000032	-4.4	mlHFdepl-104	7.91	PPE54	Rv3343c	1.1
mlHFdepl-079	6.14	echA14	Rv2486	-1.2	mlHFdepl-105	6.2	PPE55	Rv3347c	-1.3
mlHFdepl-080	9.09	Rv2491	Rv2491	-2.2	mlHFdepl-105	6.2	Rv3348	Rv3348	1.1
mlHFdepl-080	9.09	Rv2492	Rv2492	-2.4	mlHFdepl-106	6.13	PPE56	Rv3350c	-1.3
mlHFdepl-080	9.09	vapB38	Rv2493	1.4	mlHFdepl-107	6.93	PPE56	Rv3350c	-1.3
mlHFdepl-081	6.3	Rv2512c	Rv2512c	-1.9	mlHFdepl-108	8.77	Rv3377c	Rv3377c	-3.9
mlHFdepl-081	6.3	Rv2513	Rv2513	-2.6	mlHFdepl-108	8.77	Rv3378c	Rv3378c	-4.4
mlHFdepl-082	6.27	valT	MTB000035	-2.0	mlHFdepl-108	8.77	dxs2	Rv3379c	1.4
mlHFdepl-082	6.27	glyT	MTB000036	-2.5	mlHFdepl-109	12.87	idsB	Rv3383c	-1.9
mlHFdepl-082	6.27	cysU	MTB000037	-4.1	mlHFdepl-110	6.42	Rv3424c	Rv3424c	-2.7
mlHFdepl-082	6.27	valU	MTB000038	-4.3	mlHFdepl-110	6.42	PPE57	Rv3425	-1.6
mlHFdepl-082	6.27	Rv2645	Rv2645	-1.9	mlHFdepl-111	8.35	PPE59	Rv3429	-2.9
mlHFdepl-083	6.38	Rv2734	Rv2734	-3.5	mlHFdepl-111	8.35	Rv3430c	Rv3430c	-1.1
mlHFdepl-083	6.38	Rv2735c	Rv2735c	-2.0	mlHFdepl-112	13.33	Rv3528c	Rv3528c	-6.1
mlHFdepl-084	7.07	Rv2778c	Rv2778c	-1.8	mlHFdepl-113	6.44	hsaA	Rv3570c	-1.1
mlHFdepl-084	7.07	Rv2779c	Rv2779c	-5.0	mlHFdepl-113	6.44	kshB	Rv3571	-1.8
mlHFdepl-084	7.07	ald	Rv2780	-2.8	mlHFdepl-114	6.34	lsr2	Rv3597c	-4.8
mlHFdepl-085	7.06	Rv2807	Rv2807	1.5	mlHFdepl-114	6.34	lysS	Rv3598c	-2.5
mlHFdepl-085	7.06	Rv2808	Rv2808	-3.8	mlHFdepl-115	7.14	espA	Rv3616c	-23.2
mlHFdepl-085	7.06	Rv2809	Rv2809	-2.1	mlHFdepl-115	7.14	ephA	Rv3617	-2.7
mlHFdepl-086	6.84	Rv2816c	Rv2816c	-3.8	mlHFdepl-116	6.96	PPE65	Rv3621c	1.3
mlHFdepl-086	6.84	Rv2817c	Rv2817c	-1.3	mlHFdepl-116	6.96	PE32	Rv3622c	-1.1
mlHFdepl-087	7.71	Rv2821c	Rv2821c	-2.7	mlHFdepl-116	6.96	lpqG	Rv3623	1.7
mlHFdepl-087	7.71	Rv2822c	Rv2822c	-3.2	mlHFdepl-117	6.1	Rv3748	Rv3748	-1.1
mlHFdepl-087	7.71	Rv2823c	Rv2823c	-1.5	mlHFdepl-117	6.1	Rv3749c	Rv3749c	-2.5
mlHFdepl-088	6.64	cdsA	Rv2881c	-1.6	mlHFdepl-117	6.1	Rv3750c	Rv3750c	-2.9
mlHFdepl-088	6.64	frr	Rv2882c	-2.1	mlHFdepl-118	6.45	mmpL8	Rv3823c	-2.0
mlHFdepl-089	6.26	tesA	Rv2928	-6.1	mlHFdepl-119	6.1	fadD23	Rv3826	-4.2
mlHFdepl-089	6.26	Rv2929	Rv2929	-6.9	mlHFdepl-120	7.58	Rv3840	Rv3840	1.0
mlHFdepl-089	6.26	fadD26	Rv2930	-3.1	mlHFdepl-120	7.58	bfrB	Rv3841	-1.9
mlHFdepl-090	6.43	drdB	Rv2937	-1.0	mlHFdepl-121	7.32	Rv3863	Rv3863	-1.1
mlHFdepl-090	6.43	drnC	Rv2938	-1.1	mlHFdepl-121	7.32	espE	Rv3864	-2.3
mlHFdepl-090	6.43	papA5	Rv2939	-1.2	mlHFdepl-122	6.12	eccCb1	Rv3871	1.1
mlHFdepl-091	7.42	fadD22	Rv2948c	-2.5	mlHFdepl-122	6.12	PE35	Rv3872	-1.4
mlHFdepl-091	7.42	Rv2949c	Rv2949c	-4.6	mlHFdepl-123	6.75	Rv3902c	Rv3902c	-3.5
mlHFdepl-091	7.42	fadD29	Rv2950c	-4.3	mlHFdepl-123	6.75	Rv3903c	Rv3903c	-1.5
mlHFdepl-092	6.52	Rv2956	Rv2956	-6.8	mlHFdepl-124	6.41	Rv3921c	Rv3921c	-3.2
mlHFdepl-092	6.52	Rv2957	Rv2957	-3.1	mlHFdepl-124	6.41	Rv3922c	Rv3922c	-3.0
mlHFdepl-093	8.9	Rv2958c	Rv2958c	-2.6					
mlHFdepl-093	8.9	Rv2959c	Rv2959c	-4.8					

Reported features are 500 bp within the mlHF binding site. Enr = enrichment of mlHF peak. Fold change from RNA-seq mlHF cKD + ATc is reported in column FC.

Gene Ontology analysis

	GO ID	GO terms enriched in downregulated genes	Annotated ¹	Significant ²	Expected ³	p-value
BP	GO:0001676	long-chain fatty acid metabolic process	11	4	0.33	0.0002
	GO:0052564	response to immune response of other organism invol	108	13	3.2	0.0003
	GO:0010039	response to iron ion	8	3	0.24	0.0013
	GO:0008610	lipid biosynthetic process	168	14	4.98	0.0025
	GO:0072330	monocarboxylic acid biosynthetic process	82	8	2.43	0.0025
	GO:0044315	protein secretion by the type VII SS	10	3	0.3	0.0026
	GO:0006777	Mo-molybdopterin cofactor biosynthetic process	13	3	0.39	0.0058
	GO:0009408	response to heat	14	3	0.42	0.0072
	GO:0009405	pathogenesis	124	9	3.68	0.0100
	GO:0071766	Actinobacterium-type cell wall biogenesis	47	5	1.39	0.0117
	GO:0006355	regulation of transcription, DNA-templated	226	12	6.7	0.0320
	GO:0046417	chorismate metabolic process	10	2	0.3	0.0335
CC	GO:0005840	ribosome	59	5	1.82	0.0250
	GO:0044421	extracellular region part	10	2	0.31	0.0360
MF	GO:0004792	thiosulfate sulfurtransferase activity	8	4	0.24	0.0001
	GO:0016877	ligase activity, forming carbon-sulfur bonds	19	5	0.58	0.0058
	GO:0070566	adenylyltransferase activity	18	3	0.55	0.0158
	GO:0035731	dinitrosyl-iron complex binding	7	2	0.21	0.0173
	GO:0004467	long chain fatty acid-CoA ligase activity	7	2	0.21	0.0173
	GO:0003735	structural constituent of ribosome	57	5	1.73	0.0277
	GO:0019843	rRNA binding	42	4	1.28	0.0368
	GO:0003700	transcription factor activity, sequence-specific DNA bir	84	6	2.55	0.0401
	GO:0043565	sequence-specific DNA binding	27	3	0.82	0.0467

	GO ID	GO terms enriched in upregulated genes	Annotated	Significant	Expected	p-value
BP	GO:0032508	DNA duplex unwinding	12	2	0.16	0.0110
	GO:1901606	alpha-amino acid catabolic process	19	2	0.26	0.0260
	GO:0009252	peptidoglycan biosynthetic process	22	2	0.3	0.0350
	GO:0008360	regulation of cell shape	26	2	0.35	0.0470
	GO:0006281	DNA repair	62	3	0.84	0.0500
CC	no significant GO term					
MF	GO:0004003	ATP-dependent DNA helicase activity	9	2	0.12	0.0059
	GO:0008194	UDP-glycosyltransferase activity	11	2	0.15	0.0089
	GO:0004527	exonuclease activity	21	2	0.28	0.0313
	GO:0003677	DNA binding	364	9	4.88	0.0452
	GO:0016811	hydrolase activity, acting on carbon-nit...	26	2	0.35	0.0464

1 = Total number of genes in *M. tuberculosis* with this GO term, 2 = Significant number of genes by Fisher's exact test found in the present dataset, 3 = Expected number of genes without any enrichment. BP = biological process, CC = cellular component, MF = molecular function

mIHF Function – 3.1

CHAPTER 3.2

Structural and DNA-binding properties of *Mycobacterium tuberculosis* mIHF

Nina T. Odermatt¹, Moreno Lelli², Torsten Herrmann², Luciano Abriata^{1/4},
Aleksandre Japaridze³, Hubert Voilquin³, Lyndon Emsley⁴, Giovanni Dietler³,
Matteo Dal Peraro¹, Stewart T. Cole¹

1 École Polytechnique Fédérale de Lausanne, School of Life Sciences,
Station 19, 1015 Lausanne, Switzerland

2 University of Lyon, Institute for Analytical Sciences – Center for High Field NMR,
CNRS UMR 5280, ENS Lyon, UCB Lyon 1, France

3 École Polytechnique Fédérale de Lausanne, School of Basic Life Sciences,
Route de la Sorge, 1015 Lausanne, Switzerland

4 École Polytechnique Fédérale de Lausanne, School of Basic Sciences,
Av. F.-A. Forel 2, 1015 Lausanne, Switzerland

2017. Manuscript in preparation

Contributions: design of experiments, gene cloning and protein purification, CD measurements,
structure and data analysis, manuscript preparation

ABSTRACT

Nucleoid associated proteins (NAP) in bacteria take part in active chromosome organization by supercoil management, three-dimensional DNA looping and direct transcriptional control. Mycobacterial integration host factor (mIHF, *rv1388*) is a NAP restricted to actinobacteria and essential for survival of the deadly human pathogen *Mycobacterium tuberculosis*. Its structure, mode of DNA-binding and if it shapes the chromosome are not known. Here, we describe the structure obtained by Nuclear Magnetic Resonance spectroscopy of mIHF and characterize it as a globular protein with a protruding alpha helix. No residues of high flexibility were identified, suggesting that mIHF is a rigid protein overall that does not undergo structural rearrangements. In solution, mIHF was identified as a monomer, but data suggest that the protein might form oligomers upon DNA binding. Two DNA binding domains (DBDs) were identified, lying on opposite sides of the core protein. DBD I has an extensive DNA-protein interface involving eight residues, while in DBD II only two residues interacted with DNA. Nevertheless, DBD II is required for DNA binding, as a mutant within this domain diminished the interaction with DNA. Further, DNA-binding of mIHF *in vitro* strongly stabilizes

the protein and increases its melting temperature. mIHF is able to introduce left-handed loops of ca. 300 bp size in supercoiled cosmids, thereby unwinding and relaxing the DNA. The protein kinases PknB and PknG phosphorylated mIHF *in vitro*, while only PknB targeted T83, the sole residue phosphorylated in mIHF protein extracted from *M. tuberculosis*. This post-translational modification might control the DNA binding activity of mIHF *in vivo*.

INTRODUCTION

Tuberculosis, caused by *Mycobacterium tuberculosis*, is the top-ranking cause of death by a single bacterium, claiming over 1.5 million lives each year (WHO, 2016). Strict control of gene expression is crucial for the bacterium to survive the hostile milieu represented by the host immune system and by low nutrient availability. Nucleoid associated proteins (NAPs) are small, highly abundant proteins in a bacterial cell with hundreds of binding sites on the chromosome. They influence DNA conformation and regulate a vast number of genes (Dillon, 2010). The mycobacterial integration host factor, mIHF, was identified as necessary for phage L5 integration into the *Mycobacterium smegmatis* genome (Pedulla, 1998) and was shown to bind and bend DNA non-specifically (Mishra, 2013). On the other

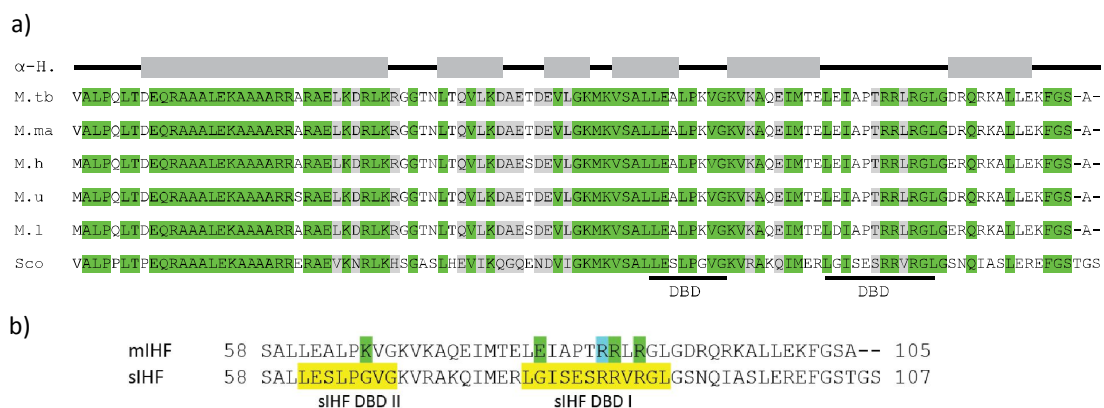


Fig. 1. Amino acid sequence alignment of IHF proteins. a) Green highlighted residues are conserved in all species, grey amino acids belong to the same chemical group. mIHF from *M. tuberculosis* (M.tb) is aligned with protein sequences from *M. marinum* (M.ma), *M. haemophilum* (M.h), *M. ulcerans* (M.u), *M. leprae* (M.l) and sIHF from *S. coelicolor* (Sco, SCO1480). Secondary structure prediction is shown on top, α -H. stands for α -Helix. DNA-binding domains (DBD) of sIHF are marked with bars on the bottom. b) DNA binding domains (DBD) of sIHF are highlighted in yellow. Arg86 used for mutant construction in mIHF is shaded in blue and specific amino acids discussed in the text in green.

hand, IHF of *E. coli* is known to not only mediate integration of phage DNA, but also for its pivotal role in gene regulation (Prieto, 2012) and DNA replication (Leonard, 2014). Binding of NAPs to the chromosome induces torsion, bending or looping and influences DNA topology, which is a central means of controlling gene expression and replication. The mIHF protein is encoded by *rv1388* and starts 255 bp downstream of the annotated translation initiation site in *Tuberculis* (Lew, 2011). The 105-amino acid

protein has a calculated molecular weight of 11.5 kDa and does not contain any Tryptophan, Tyrosine or Cysteine residues, rendering it undetectable by UV spectrophotometry. Amino acid sequence alignments (Fig. 1a) show that mIHF is highly conserved among Actinobacteria. Conversely, mIHF does not share any sequence similarity with its orthologues from Gram-negative bacteria.

Most NAPs act as dimers, with each monomer binding to a distinct DNA locus and

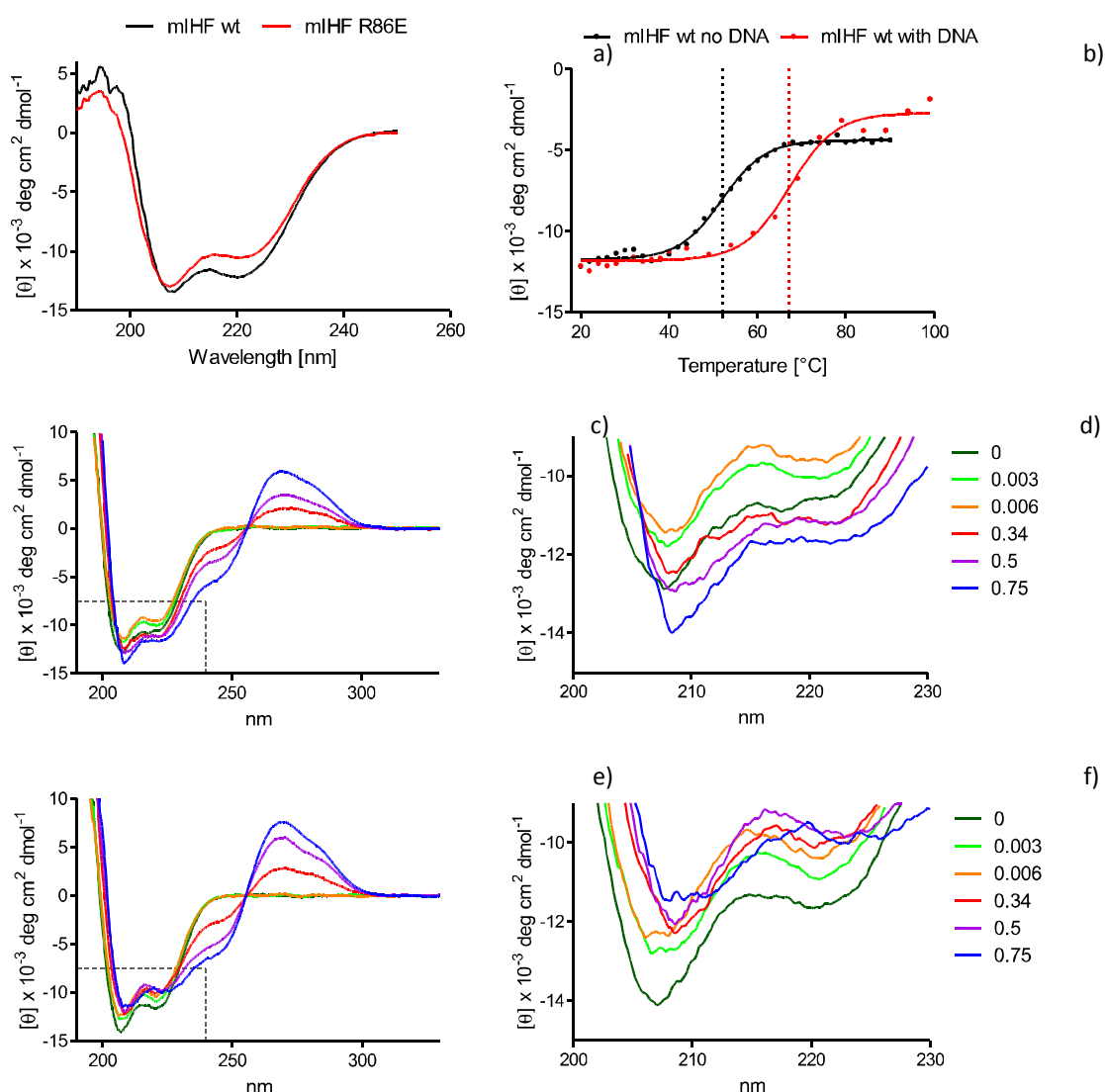


Fig. 2. Circular dichroism (CD) spectra of mIHF proteins. a) CD spectra of mIHF wt and mIHF R86E. b) Thermal unfolding of mIHF wt in absence and presence of dsDNA at 222 nm. Dotted line indicate melting temperature without (black) and in presence of DNA (red). c) CD spectra of mIHF wt with DNA titration and zoom of troughs representing alpha helices (d). DNA equivalents of 0, 0.003, 0.006, 0.34, 0.5 and 0.75 were used. e) CD spectra of mIHF R86E with DNA titration and zoom (f), as for mIHF wt.

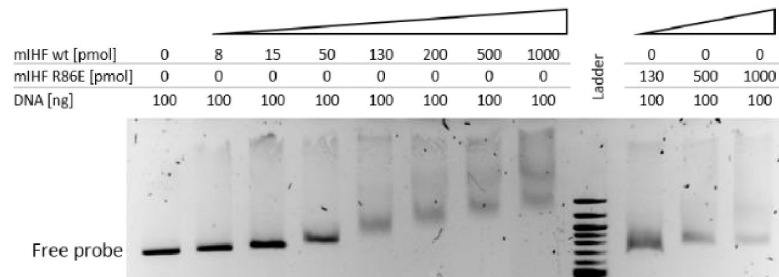


Fig. 3. Gel retardation assay obtained upon incubation of a 700 bp DNA fragment with miHF or with miHF R86E. DNA was incubated with increasing concentrations of miHF as detailed in the image and run on a 1% agarose gel. Post-run staining was performed with GelRed. Ladder: 100 bp molecular weight marker.

bringing these loci together upon dimerization as exemplified by EspR (Blasco, 2011), Lsr2 (Summers, 2012) or HupB (Bhowmick, 2014) in *M. tuberculosis*. On the contrary, the miHF homologue from *Streptomyces coelicolor*, siHF, has two separate DNA-binding domains (DBD) and binds DNA as a monomer (Swiercz, 2013). siHF contains a helix-two turns-helix (H2TH) motif, where the two turns between the helices contact DNA at one DBD of the protein. The second DBD lies on the opposite side of the protein, in a turn connecting two helices (Swiercz, 2013). This DNA binding mode identifies a new category of NAPs, characterized by two distinct DNA binding domains in the same protein rather than by dimerization of two proteins with one DBD each. siHF is thus unique and not related to other NAPs like IHF or H-NS from *E. coli* or Lsr2 from *M. tuberculosis* (Swiercz, 2013).

miHF is a NAP in *M. tuberculosis* and binds non-specifically to 150 loci around the genome. The *mihF* gene is essential and, upon depletion of the protein, bacteria show an abnormally long phenotype, stop growth and finally die. miHF regulates genes involved in lipid metabolism, metabolic pathways, translation and virulence (Odermatt et al., manuscript in preparation). In order to understand how miHF fulfils its role as an architectural DNA-shaping protein and as a transcription factor, we studied the structure of miHF and its interaction with DNA using a range of biophysical techniques including nuclear magnetic resonance (NMR) spectroscopy, and atomic force microscopy.

RESULTS

miHF is a highly soluble, alpha helical DNA-binding protein.

Alignment of miHF with siHF showed high homology levels between the two sequences, with 61% identity and 79% similarity. These data suggested that the two proteins might have a very similar structure. We therefore exploited the siHF structure with the determined DBDs (Swiercz, 2013) to identify the putative DBDs of miHF (Fig. 1b). Overall, both domains were conserved but several interesting differences between siHF and miHF were noticed, such as residues G67K and G81E, which changed the small, uncharged amino acid glycine present in siHF to the positively charged lysine and the negatively charged glutamic acid, respectively, in miHF. These changes might reflect the adaptation to specific DNA binding patterns in the respective genomes or modify sequence specificity. Furthermore, several positively charged arginine residues are present in the miHF and siHF DBD and are presumably responsible for the direct contact with the negatively charged DNA backbone. We chose Arg86, Arg88 and Gly89, conserved in both proteins, as targets of site-directed mutagenesis to investigate the DNA binding properties of the mutant proteins. Arginine was mutated to glutamic acid (R86E, R88E), and glycine to tryptophan (G89W). The mutant proteins containing R88E or G89W were not stable after purification, while miHF R86E was as stable as miHF wild type (wt) and therefore used for future experiments.

miHF was purified to homogeneity with a concentration of 60 mg ml⁻¹ and displayed

elevated stability at room temperature. The structural features of mIHF were characterized via circular dichroism (CD) spectroscopy. Fig. 2a illustrates the mean residue ellipticity of mIHF and of the R86E mutant. Both proteins showed a folded structure of predominately alpha helices, as reflected in the two pronounced negative peaks at 208 and 222 nm. mIHF R86E displayed a minor deviation from mIHF wt, suggesting that their overall secondary structures are very similar but not completely identical. The spectrum at 222 nm, as a function of temperature, was fitted to a sigmoidal non-linear model and the melting temperature was calculated as 52°C (Fig. 2b). The protein did not refold into its alpha helical structure at 20°C after heating to 95°C. Interestingly, the melting temperature increased to 67°C in the presence of DNA, suggesting higher stability of mIHF when bound to DNA (Fig. 2b).

To screen for conformational changes in the mIHF protein upon DNA binding, the wt and the mutant proteins were incubated with very low (0.003 and 0.006 equivalents) and increasingly higher amounts (0.34, 0.5, 0.75 equivalents) of a 40 bp dsDNA oligonucleotide. The troughs at 208 nm and at 222 nm became less pronounced in mIHF wt at low amounts of DNA, indicating a small conformational change with no emergence of a different secondary structure. At higher DNA equivalents, a deeper trough at 208 nm appeared (Fig. 2c and zoom

in Fig. 2d). mIHF R86E showed a similar pattern with a decrease at both 208 nm and 220 nm troughs, but no subsequent increase at high DNA concentrations. To assess how the mIHF protein shifted linear DNA, a 700 bp DNA fragment was incubated with mIHF wt and mIHF R86E for a gel retardation assay. While mIHF wt shifted the DNA effectively, R86E had a markedly lower retardation effect (Fig. 3). This confirmed that Arg86 is important for DNA binding by mIHF. The minor DNA shift caused by mIHF R86E suggested that the mutant protein is still capable of binding DNA, confirming the CD results, but with much lower affinity than mIHF wt. Thus, mIHF R86E represented a suitable control for other DNA-mIHF interaction experiments.

Solution structure of mIHF and protein dynamics

In order to define the exact binding mechanism and interaction with DNA, structure determination of mIHF was carried out by nuclear magnetic resonance (NMR) spectroscopy. High-dimensional automated projection spectroscopy (APSY) spectra for backbone assignment and 3D $^{15}\text{N}/^{13}\text{C}$ -HSQC-NOE spectroscopy (NOESYs) to obtain the structural restraints were acquired (Table 1). The backbone and total atom assignments were completed to 90.24% and 76.71%, respectively. mIHF was identified as a globular protein for residues 8 – 103 with an RMSD (root mean square deviation of atomic positions) of 1.45 Å or 0.99 Å for residues 14 – 103, excluding six more non-structured residues. The N-terminal domain (residues Val1-Thr7) did not form a defined secondary structure and was disordered (Fig. 4a). Overall, the NMR structure of mIHF resembled closely the sIHF X-ray structure with an overlay RMSD of 1.62 Å and a similar topology with six α -helices. The flexible N-terminal end of mIHF is constituted by residues Val1 to Arg11 followed by the first α -helix (α 1) comprising Ala12 to Leu32. This helix was predicted to form a coiled coil and might be the site of dimerization for mIHF. A minor kink at residues Ala24 and Arg25 caused a slight divergence from an optimal helix. α 1 is connected by a five-residue loop to α 2, a short helix from Leu43 to Glu50. α 3 only spans one turn from Lys54 to Val57 and is followed by α 4

Table 1. NMR structural statistics.

NMR constraints	
Distance	1818
Intra-residue ($ i-j = 0$)	429
Inter-residue	
Sequential ($ i-j =1$)	613
Medium range ($1 \leq i-j \leq 4$)	542
Long-range ($ i-j > 5$)	234
TALOS+ dihedral angle restraints	181
RMSD from mean structure (residue 14 – 103) [Å]	0.99
Ramachandran analysis	
most favoured	85.2%
additional allowed	13.6%
generously allowed	1.0%
disallowed	0.2%

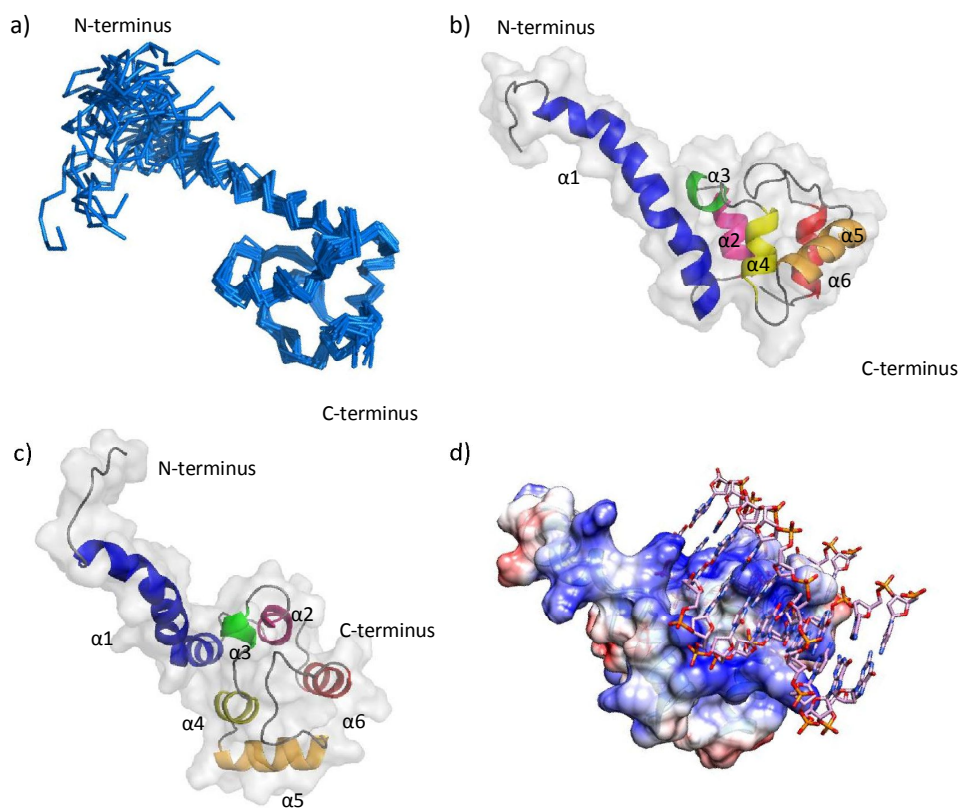


Fig. 4 Structure of mlHF. a) overlay of the 20 lowest energy structures of mlHF. b,c) Ribbon drawing of mlHF (90° rotated around z-axis). d) Surface charge of mlHF, blue representing positively and red negatively charged surface of mlHF. DNA was extrapolated from alignment with slHF.

from Leu60 – Lys66, $\alpha 5$ (Gln73 – Pro83) and finally $\alpha 6$ (Ala97 – Ala105) (Fig. 4b).

Structural assignment was difficult due to multiple ambiguous restraints potentially caused by dimerization of mlHF. To probe the internal dynamics of mlHF and its aggregation state, three relaxation parameters for each backbone amide were characterized: the longitudinal and transverse ^{15}N relaxation times R_1 and R_2 as well as the steady state nuclear Overhauser effect ($\{^1\text{H}\}$ - ^{15}N NOE). ^{15}N relaxation data are correlated with the molecular rotational correlation time, which in turn is related to the molecular hydrodynamic radius. The molecular rotational correlation time is derived from the T_1/T_2 ratio for the 44 most rigid residues (with $\{^1\text{H}\}$ - ^{15}N NOE > 0.65) and provided a rotational correlation time of 6.9 ± 1.7 ns, which is in line with a molecular weight of 10-12 kDa. Fig. 5 reports the acquired relaxation data (NOE, T_1 and T_2) mapped on the

mlHF amino acid sequence. The N-terminus and the two last C-terminal amino acids showed higher flexibility with low values in T_1 and NOE. Likewise, the parameters measured for the loop connecting $\alpha 1$ and $\alpha 2$ as well as for the residues between $\alpha 5$ and $\alpha 6$ indicated higher flexibility. The rest of the protein, i.e. $\alpha 1$ and $\alpha 2$ to $\alpha 5$, including the linker residues, was more rigid.

mlHF has two distinct DNA-binding domains

The interaction of mlHF with DNA was monitored through titration of 200 μM ^{13}C / ^{15}N -labelled mlHF with 16 bp dsDNA at 0.25, 0.5, 1 and 2.2 equivalents. After addition of DNA, the ^1H , ^{15}N HSQC spectrum was monitored and this allowed the identification of mlHF residues that participated in DNA binding. Analysis of the titration was based on the backbone ^1H and ^{15}N assignments delivered by the automated assignment and structure calculation procedure mentioned above. Unfortunately, a few ^1H , ^{15}N

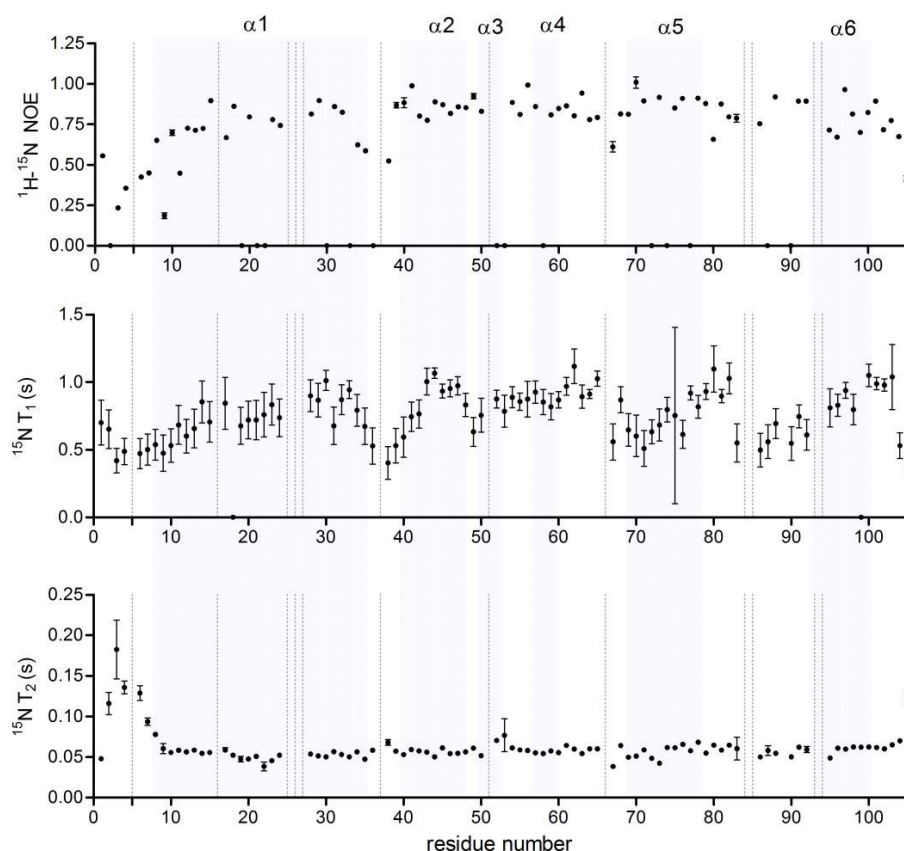


Fig. 5. NMR protein dynamics. Relaxation parameters heteronuclear $\{^1\text{H}\}$ - ^{15}N NOE (top panel), ^{15}N - T_1 (middle panel) and ^{15}N - T_2 (lower panel) are plotted as a function of the mIHF sequence. Dotted lines mark residues that could not be assigned. Data are represented \pm uncertainty of the fitted parameter.

crosspeaks remained unassigned by the automated procedure and therefore could not be integrated in the DBD analysis. Effects were evident already at 0.25 equivalents, with a series of ^1H , ^{15}N crosspeaks shifting and others disappearing almost entirely, namely Ala13, Lys33, Lys66, Gly68, Lys71, Ala72, Arg86 and Arg88 (crosspeak intensity ratio < 0.6). DNA binding and possibly other coupled effects occurred in this protein region, in a timescale of microseconds. Fig. 6a illustrates examples of strong peak shifting of an unassigned peak (i) and of Gly89 (ii), of a disappearing peak (iii, Lys66), a not affected peak (iv, Gly67) and a weakly shifting peak (v, Gly91). At 0.5, 1 and 2.2 equivalents, crosspeak shifts were increasingly larger and the intensity of all crosspeaks became weaker, which was compensated by increasing the number of scans. The global

decrease of intensity for all crosspeaks at higher DNA equivalents indicated an increase in protein size and pointed to possible oligomerization. Intensity drop at 0.25 equivalents and crosspeak shifts at 2.2 equivalents DNA are plotted versus the primary sequence of mIHF in Fig. 6b.

The mIHF DBD I was identified as crosspeaks with a strong intensity drop at 0.25 equivalents of DNA in close spatial proximity. DBD I includes residues from different secondary structure motifs, i.e. Lys28 and Lys33, part of $\alpha 1$, two residues in $\alpha 5$ (Lys71, Ala72) as well as Glu62, Leu64, Lys66 and Gly68, connecting $\alpha 4$ with $\alpha 5$ (highlighted in blue in Fig. 6b). Further, a second DBD was detected with only two residues showing an intensity drop (Arg86 and Arg88, orange in Fig. 6b). In contrast, strong interaction of sIHF with DNA was mapped to a loop

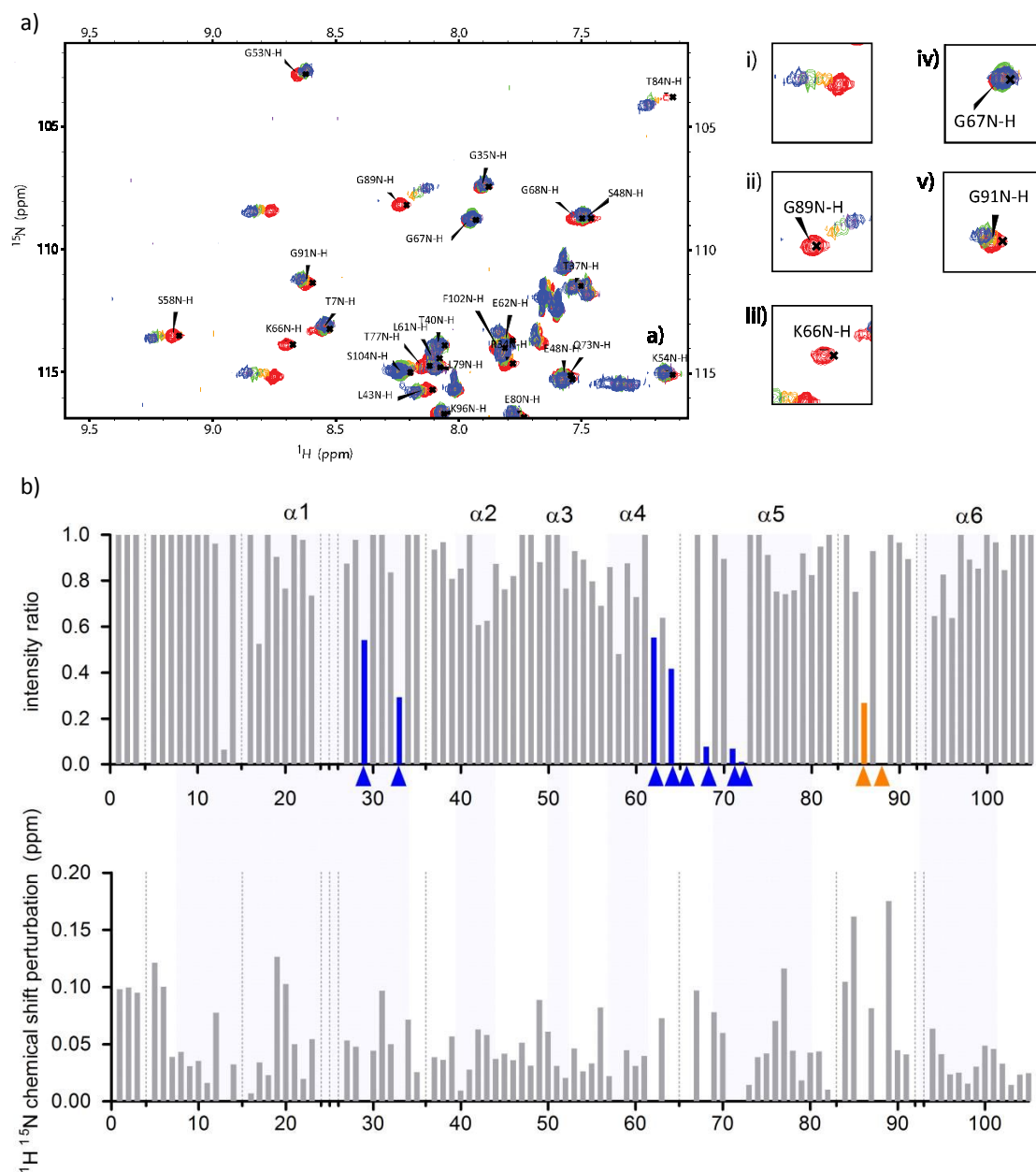


Fig. 6. Effect of DNA titration on mlHF. a) Crosspeak shifting during DNA titration. Overlay of ^1H - ^{15}N HSQC spectra without DNA (red), 0.25 (orange), 0.5 (green), 1 (purple) and 2.2 equivalents DNA (blue). Zoomed images i – v show examples of different effects on peaks. b) Upper panel: crosspeak intensity ratio of 0.25 equivalent DNA relative to mlHF without DNA. Residues of DBD I are highlighted with blue arrows and blue bars. Orange arrows point to residues that are part of DBD II. Residues with strong intensity drop were excluded from chemical shift perturbation calculations. Lower panel: chemical shift perturbation at 2.2 equivalents DNA. Dotted lines in both panels mark residues that could not be assigned.

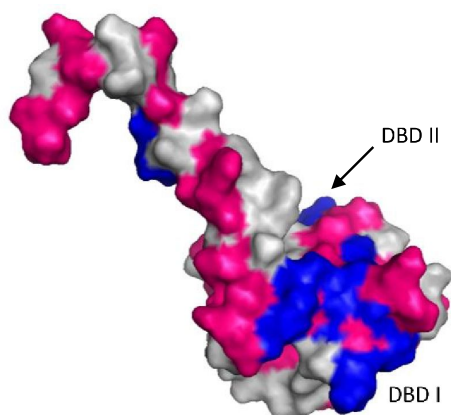


Fig. 7. Chemical shift perturbation mapped onto the mIHF NMR structure. Blue residues show sharp drop of intensity ratio at 0.25 equivalents DNA, pink residues have a high CSP and grey residues were not affected by DNA addition.

connecting $\alpha 5$ and $\alpha 6$, corresponding to DBD II in mIHF, and a less extensive interface the loop connecting $\alpha 4$ and $\alpha 5$ (Swiercz, 2013), corresponding to DBD I of mIHF. We therefore propose that mIHF, like sIHF also, has two distinct DBD, but with an inverted size of the interface.

The backbone chemical shift perturbation (CSP, Fig. 6b for 2.2 equivalents of DNA) exhibited high values not only for the inferred DBDs but also for other residues distributed over the whole mIHF protein. The widely distributed CSP suggested that other coupled effects took place upon DNA addition. Highly affected residues align on one side of $\alpha 1$, the site of the predicted coiled-coil interaction, as represented in Fig. 7.

mIHF introduces defined loops into linear DNA fragments

To study mIHF binding to DNA, purified mIHF protein was incubated with DNA substrates of various length, topology and GC content, and imaged using atomic force microscopy (AFM). The first substrate tested was a 1 kb linear fragment containing the *mihF* gene and the upstream promoter region with a GC content of 65%, representing the mean GC content in *M. tuberculosis*. The DNA fragment alone showed a looping probability of 10%, which increased upon addition of mIHF and reached 53% when

the maximum amount was added (Table S1, Fig. 8). The size of the loops was well defined with a mean length of 110 ± 10 nm, which corresponds to ca. 300 bp (Table S2). When the same experiment was repeated with mIHF R86E, loops in the linear DNA were observed as well but with a lower ratio compared to the wt protein. Only 28% looped fragments were detected upon binding of mIHF R86E at the same maximal concentration (Table S1, Fig. 8b). The average loop length was similar for mIHF

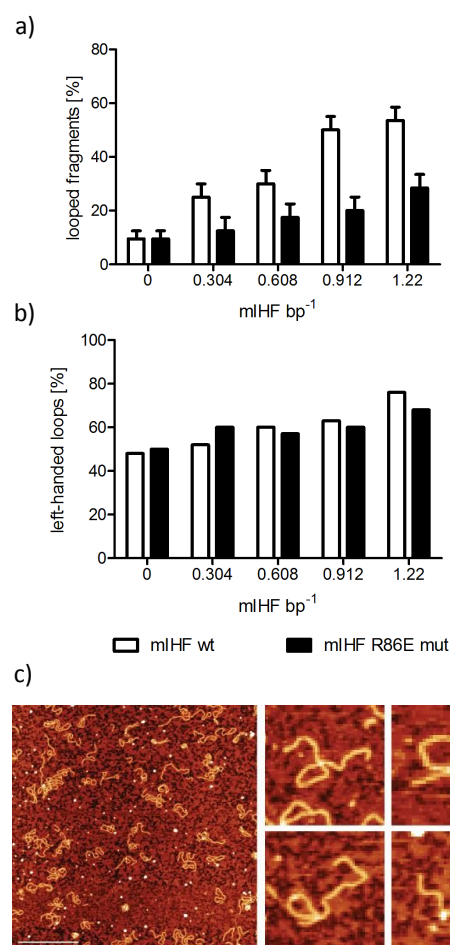


Fig. 8. Atomic force microscopy experiments. A 1 kb linear DNA fragment harbouring the *mihF* gene and promoter region was incubated with the mIHF protein. a) Graphs representing the mean percentage \pm standard deviation of looped fragments observed upon addition of increasing concentrations of mIHF wt and mIHF R86E and the percentage of left-handed loops (b). c) Image of DNA fragments and zoomed images of looped structures at 0.21 mIHF bp⁻¹. Scale bar represents 500 nm.

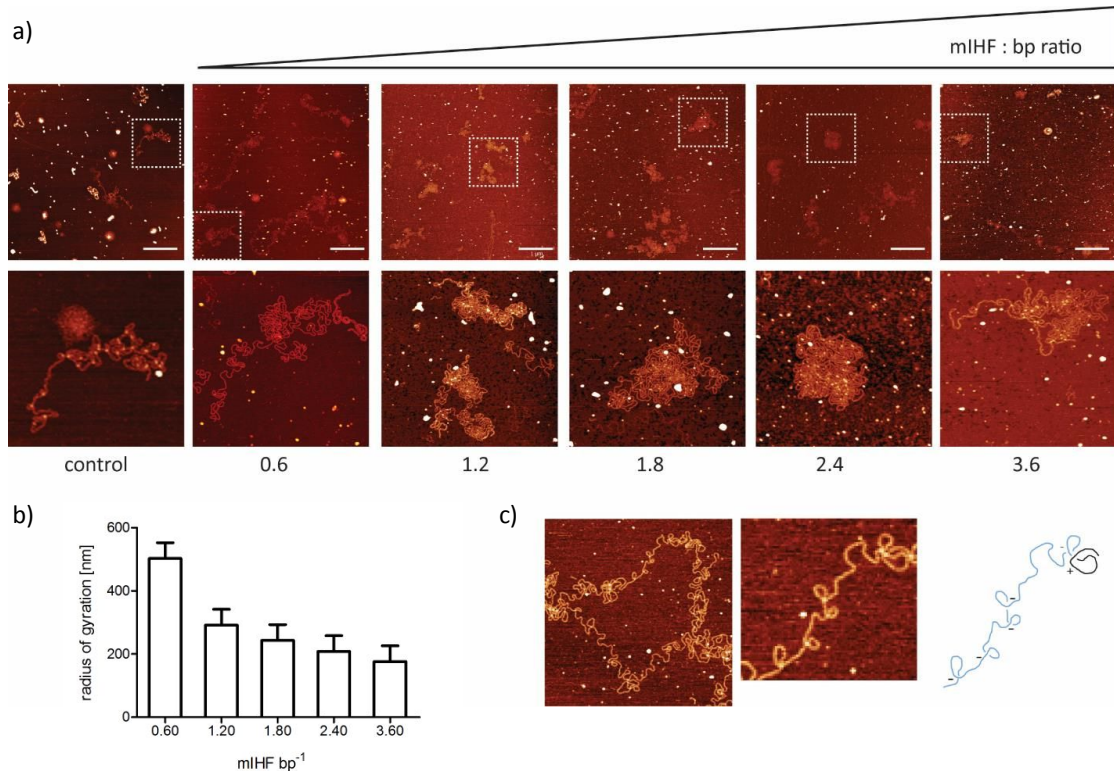


Fig. 9. Atomic force microscopy experiments with cosmid I95. a) Cosmid I95 alone and incubated with increasing concentrations of mIHF wt. The lower panels are zoomed images of the squares defined in the upper panels. Scale bar represents 1 μ m and mIHF bp⁻¹ ratios are indicated below images. b) Radius of gyration measured for cosmid I95 in the presence of mIHF. Bars represent mean \pm standard deviation. c) Cosmid I95 incubated with 1.2 mIHF bp⁻¹, zoom to an example structure and handedness of loops ("-" = left-handed, "+" = right-handed).

and mIHF R86E with 110.5 nm and 106.5 nm respectively (Table S2). These results indicated that mIHF R86E did exhibit DNA binding activity, but with lower affinity compared to the wt protein, confirming the agarose retardation assay results. Naked DNA had approximately 50% left- and right-handed loops, corresponding to positive and negative supercoiling, respectively. mIHF wt, as well as mIHF R86E, increased the number of left-handed loops in a concentration-dependent manner to 76% for the wt protein and to 68% for the mutant (Fig. 8c, Table S3).

Large cosmid unwind and compact in an mIHF dependent manner

The *M. tuberculosis* chromosome is a circular, supercoiled macromolecule of 4.4 Mb in length. Topology and size are not correctly represented by short linear DNA fragments or plasmids. Furthermore, bacterial chromatin is organized into microdomains of ca. 10 kb

(Dame, 2016), a compartmentalization that cannot be mimicked by plasmids. Therefore, we used the 42.6 kb I95 supercoiled cosmid (Bange, 1999) to reproduce more closely the features of the genomic DNA. The supercoiled I95 cosmid forms highly ordered structures, the so-called hyperplectonemes, only present in DNA molecules above 30 kb in size (Japaridze, 2017). Incubation of I95 with mIHF wt at low concentrations (0.6 mIHF bp⁻¹) led to unwinding of the cosmid from the hyperplectonemic form and reduced its complexity (Fig. 9). At higher mIHF concentrations (1.2 to 3.6 mIHF bp⁻¹), the cosmid collapsed and compacted again, as illustrated by the decrease in radius of gyration (smallest circle that contains the DNA polymer, Fig. 9b, Table S4). Re-compaction of the cosmid at high mIHF concentrations did not cause formation of any hyperplectonemes like the ones observed without any protein, but only reached first-order plectonemes. Similar to the loops observed on the linear DNA fragment,

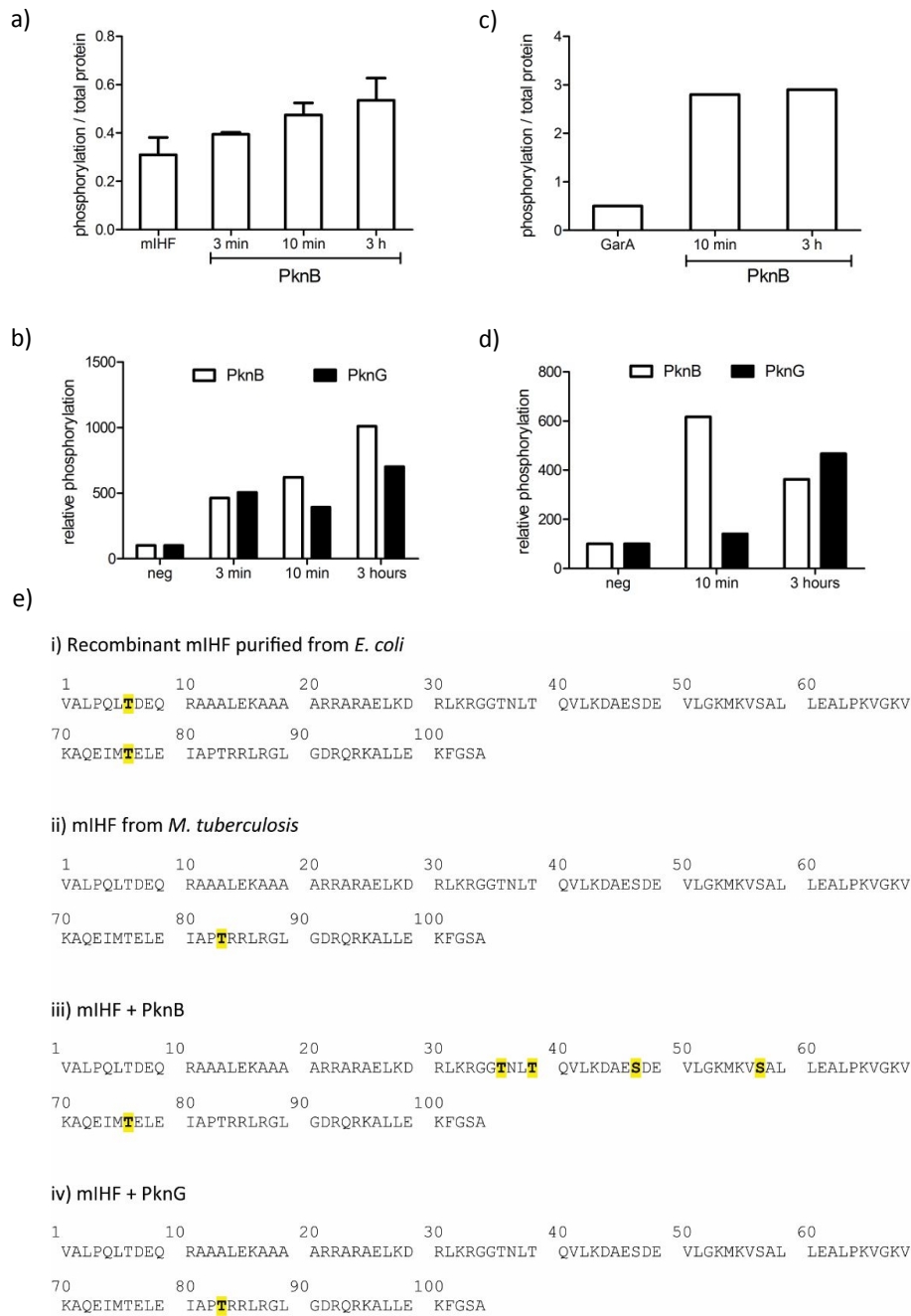


Fig. 10. Phosphorylation assay on mIHF and GarA. Incubation time of substrate with kinase is indicated below each bar. a) Band intensity of ProQ-stained mIHF relative to total mIHF proteins incubated with PknB. The graph represents the mean from two replicates \pm standard deviation normalized to the total protein content. b) Band intensity of ProQ-stained GarA relative to total GarA incubated with PknB. c) Ratio of phosphorylated mIHF peptides incubated with kinases PknB or PknG relative to negative control (mIHF not incubated with kinase), detected by mass spectrometry. d) Ratio of phosphorylated GarA peptides incubated with kinases PknB or PknG relative to negative control (GarA not incubated with kinase), detected by mass spectrometry. e) Phosphorylated residues (highlighted in bold yellow) of mIHF detected by mass spectrometry. Sequence i: protein as purified from *E. coli*. Sequence ii: mIHF incubated *in vitro* with PknB. Sequence iii: mIHF incubated *in vitro* with PknG. Sequence iv: published phosphorylation sites *in vivo*.

mIHF also introduced mostly left-handed loops in the cosmid. The R86E mutant simplified the topology of I95, but no collapse was visible at higher concentrations of the protein, in agreement with the reduced DNA-binding affinity observed with linear DNA.

mIHF is phosphorylated by serine/threonine protein kinases

Recombinant *M. tuberculosis* mIHF purified from *E. coli* was found to be phosphorylated at T7 and T77. In contrast, when extracted from *M. tuberculosis*, the protein was not phosphorylated at either tyrosine 7 nor 77, but at T83 (Prisic, 2010; Fortuin, 2015). To evaluate if the serine/threonine-protein kinases PknB or PknG were able to phosphorylate mIHF *in vitro*, a phosphorylation assay was carried out by in-gel staining of phosphorylated proteins. Incubation of mIHF with either PknB or PknG led to phosphorylation of different residues in a time-dependent manner (Fig. 10a and S1). GarA, a known substrate of PknB and PknG (Villarino, 2005; O'Hare, 2008) was used as a control for PknB phosphorylation and showed a 3-fold increase in phosphorylation intensity (Fig. 10b and S1). Already after ten minutes, both substrates, mIHF and GarA were phosphorylated by PknB. To confirm the phospho-stain experiments, mass spectrometry was conducted on mIHF and GarA incubated with PknB and PknG (Fig. 10c and 10d). PknB seemed to more efficiently phosphorylate mIHF than PknG, as was the case for the control substrate GarA. Phosphorylated residues detected by mass spectrometry are highlighted in Fig. 10e. Sequence i) shows recombinant mIHF with phosphorylated T7 and T77 and sequence ii) shows mIHF with the phosphorylated residue T83. Incubation of mIHF with PknB resulted in phosphorylation of several residues (T36, T39, S47, S57 and T77), while no modification was detected at T7. PknG phosphorylated mIHF only at T83, and neither T7 nor T77 were phosphorylated. This shows that both kinases target more residues *in vitro* than those detected *in vivo*, and that both PknB and PknG, can act as phosphatases.

DISCUSSION

Importance of mIHF in the *M. tuberculosis* physiology was suggested by Schubert and colleagues, whose mass spectrometry experiments detected the protein among the ten most abundant ones in the bacterium (Schubert, 2015). This initial suspicion was confirmed by our genetic studies, where we proved the essentiality of the gene and the role of the protein in controlling transcription of a variety of virulence factors and housekeeping genes (Odermatt et al., manuscript in preparation). Here we investigated the structure of mIHF and its DNA binding properties by means of physico-chemical approaches.

High similarity to the homologue sIHF from *S. coelicolor* was already evident from the primary amino acid sequence and further validated by the NMR structure, which differed for only 1.3 Å compared to the sIHF crystal structure. mIHF formed a globular protein with a protruding α -helix at the N-terminal end and the core of the protein consisted of five short α -helices. Upon DNA binding, several clearly defined crosspeaks disappeared from the NMR spectra, indicating a DNA binding affinity in the micromolar range. At higher DNA equivalents, additional crosspeaks showed either a weaker or a stronger shift, while others were not affected. It therefore seemed that titration of the protein by DNA had two effects acting on slightly different timescales. This segregation into two apparent effects was also observed spatially, as all except four residues whose crosspeaks lost intensity at 0.25 equivalents mapped close in space. Due to the automatic residue assignment, it might be possible that Lys21, whose peak disappeared completely after addition of 0.25 equivalents of DNA, but lies in the middle of the protruding α -helix, was wrongly assigned and exchanged with Lys69. Lys69 lies between Gly68 and Lys71, whose peaks disappeared with 0.25 equivalents of DNA, but did not seem to be affected at all.

By comparison with the X-ray structure of the related sIHF protein, crystallized with a dsDNA 16-mer in a 1:1 stoichiometry, the cluster of mIHF residues whose crosspeaks lost intensity at 0.25 equivalents corresponded to

the DNA binding site as observed only across unit cells in sIHF. In sIHF, this DNA binding site was proposed to be of secondary importance and characterized by weaker affinity for DNA. In contrast, the orthologous site in mIHF displayed higher affinity for DNA. Our data indicate that the “trans-unit cell pose” represents the true DNA binding region of mIHF.

The broad effect on chemical shift perturbation caused by addition of DNA, together with the broader and decreased intensity of peaks at higher protein:DNA ratios, suggested other coupled effects besides DNA binding. These may include oligomerization, aggregation or increased viscosity. However, the protein concentration at 200 μ M rules out viscosity as a possible side effect. Additionally, a 16 bp dsDNA is not long enough to harbour more than two bound mIHF proteins, which could lead to aggregation upon DNA binding. We therefore conclude that mIHF most probably dimerizes when bound to DNA. The apo-protein in solution had a T_c corresponding to 10–12 kDa, meaning that it mainly consisted of monomers, with possibly a minor fraction of dimers. Residues whose crosspeaks shifted during the titration with DNA are distributed across a large portion of the protein surface. We hypothesize that this might reflect secondary weaker DNA-protein interactions, such as protein dimerization or wrapping of DNA around the protein after initial binding at DBD I. The clearest pattern is a series of residues lying on one specific face of the N-terminal α 1 helix. α 1 was strongly predicted to form coiled coils, a typical structure for protein-protein interactions. We suspect that mIHF exists predominantly in its monomeric form in solution, and forms dimers by coiled-coil interactions upon DNA binding. The DNA might therefore constrain mIHF in its dimeric form. This hypothesis is supported by the increase in melting temperature of mIHF wt from 53°C to 67°C in the presence of DNA (Fig. 2b). Although no conformational change was observed, the protein was greatly stabilized when bound to DNA.

Structural studies established that mIHF is a very stable protein with two DNA binding sites that appears to dimerize upon DNA binding. To better characterize its effect on DNA, high-

resolution AFM images were acquired. Incubation of mIHF with a 1 kb linear DNA fragment introduced loops of a defined size. Such DNA loops are commonly associated with regulation of gene expression in bacteria, best described in the *lac* operon of *E. coli*, where LacI activates transcription by looping the promoter region and by constraining supercoils (Fulcrand, 2016). Although mIHF does appear to bind DNA in a sequence independent manner, it affects the expression of many genes in *M. tuberculosis* (Odermatt et al, manuscript in preparation). The mIHF protein might therefore have a direct effect on gene expression by looping the DNA and by bringing other regulatory elements closer to the target gene. Support of this is provided by the large number of loci occupied by mIHF and other NAPs, such as EspR (Odermatt et al, manuscript in preparation). 76% of the loops observed at the highest protein:DNA ratio (1.22 mIHF bp⁻¹) were characterized by left-handed orientation. Naturally, DNA in bacterial cells is tightly packed in right-handed supercoils, as exemplified by the I95 cosmid. Introduction of left-handed loops by mIHF relaxed and opened the cosmid, thus making it potentially accessible to replication and transcription machineries. Indeed, mIHF was found to act mainly as a transcriptional activator (Odermatt et al., manuscript in preparation). This pattern is similar to that caused by *E. coli* HU on cosmid I95, where ordered, regular, left-handed loops were observed (Japaridze, 2017). Similarly to mIHF, *E. coli* HU binds without sequence specificity to DNA, but prefers gapped and cruciform DNA (Pinson, 1999). On the contrary, *M. tuberculosis* EspR introduced loops and additional twists into the DNA (Blasco, 2012). Therefore, it appears that mIHF and EspR have divergent effects on the DNA topology of the tubercle bacillus.

Post-translational modification of mIHF such as phosphorylation of several residues was confirmed *in vitro*. The tested kinases, PknB and PknG, were suggested by experts in the field to be possibly involved in this process. Both of them successfully phosphorylated mIHF in a time-dependent manner. Most phosphorylated sites were not found in peptides extracted from *M. tuberculosis in vivo*, where only T83 was phosphorylated. This could be due to the

conditions tested in our experiments or to the higher degree of accessibility of certain residues *in vitro*, which otherwise interact with other proteins and are therefore hidden from kinases in the bacterial cell. Only PknG was able to phosphorylate T83, indicating that PknB could be the main kinase responsible for post-translational modification of mIHF. It was reported before that another NAP of *M. tuberculosis*, HupB, was phosphorylated by PknE, which led to a decrease in DNA binding efficiency (Gupta, 2014). Interestingly, PknG was described as a specific kinase activated upon infection and controls phagosome lysosome fusion (Richard-greenblatt, 2017), suggesting that mIHF is specifically post-translationally modified upon host infection.

mIHF was detected to activate genes belonging to different pathways such as DNA replication, biofilm formation and virulence genes (Odermatt et al., manuscript in preparation). Similar to HupB, phosphorylation of the mIHF protein might adapt its DNA-binding properties, changing its target regulon from housekeeping genes to virulence genes. The biological consequences of mIHF phosphorylation have to be further investigated.

Overall, we showed here that mIHF has a globular structure comprising six alpha helices. The protein has one DNA binding domain with strong affinity and a second, weaker, DNA interacting domain. It plays an important role in shaping the DNA structure by introducing left-handed loops that probably promote transcription by decondensing the DNA. Finally, mIHF undergoes post-translational modification at threonine residues which may have a regulatory, though presently unknown, effect.

METHODS

Sequence analysis and alignment

Amino acid sequences were obtained from NCBI and aligned with blastp (Coordinators, 2017). Secondary structure was predicted by PSIPRED (Jones, 2017) and coiled coils with COILS (Lupas, 1991).

Cloning, expression and purification of mIHF

Oligonucleotides were synthesized by Microsynth (Switzerland). Primers *mihF*-exp-F and *mihF*-exp-R were used to PCR-amplify the *mihF* gene starting from gtc 255 bp downstream of the annotated start site. Restriction sites NdeI and BamHI were included in the primers to clone the resulting fragment into pET28a vector to give the expression vector pNO72 containing the N-terminally 6 x His-tagged *mihF* gene. Recombinant His-tagged protein was expressed by growing *E. coli* BL21 DE3 (ThermoFisher) cells transformed with pNO72 in Luria-Bertani (LB) broth in the presence of 50 µg ml⁻¹ kanamycin. Protein expression was induced at an OD₆₀₀ = 0.6 with 500 µM IPTG overnight at 16°C with constant shaking at 180 rpm. Cells were harvested by centrifugation and stored at -80°C until further processing. The pellet was thawed in lysis buffer (50 mM Tris-HCl pH 7.5, 500 mM NaCl, 1 mM Imidazole, 10% (w/v) glycerol and 1% Tween 20) supplemented with 1 tablet cComplete mini protease inhibitor (Roche) and 10 µl DNase I (Roche) and lysed in a cell disrupter (EmulsiFlex, Avestin, Canada). Insoluble debris was removed after centrifugation and supernatant was loaded on a His-trap column pre-equilibrated with buffer A (500 mM NaCl, 50 mM Tris-HCl pH 7.5). Increasing concentrations of buffer B (500 mM NaCl, 50 mM Tris-HCl pH 7.5, 500 mM Imidazol) eluted the protein from the column and fractions were visualized by Coomassie blue staining of a sodium dodecyl sulphate (SDS) gel (NuPAGE 4 – 12% Bis-Tris Gel, invitrogen). Pooled fractions containing mIHF were dialyzed into Tris-buffer (150 mM NaCl, 50 mM Tris-HCl pH 7.5) overnight at 4°C. To remove the His-tag, 500 µl thrombin (Sigma) was added to the protein before dialysis. The cut mIHF protein was subsequently purified by passage over Ni-NTA Agarose (Machery-Nagel) and Benzamidine Sepharose (GE Healthcare) to remove uncut protein, His-tags and thrombin. A size-selection purification step was applied to remove unfolded protein and increase sample purity. mIHF was passed over a HiLoad 16/600 column and fractions were pooled and concentrated in centrifugal filters (Ultracel, Amicon) with 3 kDa cut-off. Purified mIHF was stored at -80°C in small aliquots until further

use. Concentration of mIHF was measured by Qubit (invitrogen).

Primers *mihF*-R86E-F and *mihF*-R86E-R were used in the QuickChange II Site-Directed Mutagenesis kit (Agilent) as recommended by the manufacturer. The single amino acid mutation R86E in pNO72 gave rise to pNO84.

For the production of isotope labelled mIHF (^{15}N - and ^{13}C - ammonium chloride and D-glucose from Cambridge Isotope Laboratories, USA), the exponentially growing culture was resuspended in minimal media containing 25 mM KH_2PO_4 , 50 mM $\text{Na}_2\text{HPO}_4 \cdot 2 \text{H}_2\text{O}$, 1 g l^{-1} ^{15}N -labelled ammonium chloride, 10 mM NaCl, 0.2 mM CaCl_2 , 1 mM MgSO_4 , 50 μM iron solution, 10 ml trace metal solution and 3 g l^{-1} ^{13}C labelled glucose. 1000 x iron solution contained 100 μM citric acid and 50 μM FeCl_3 . 100 x trace metal solution contained 68.2 μM $\text{MnCl}_2 \cdot 4 \text{H}_2\text{O}$, 3.7 μM $\text{ZnSO}_4 \cdot 7 \text{H}_2\text{O}$, 0.4 μM $\text{CoCl}_2 \cdot 6 \text{H}_2\text{O}$, 0.6 μM $\text{CuCl}_2 \cdot 2 \text{H}_2\text{O}$, 1.6 μM H_3BO_3 , 2.1 μM $\text{NiCl}_2 \cdot 6 \text{H}_2\text{O}$, 2.1 μM $\text{Na}_2\text{MoO}_4 \cdot 2 \text{H}_2\text{O}$, 1.9 μM $\text{Na}_2\text{SeO}_3 \cdot 5 \text{H}_2\text{O}$. *E. coli* was grown overnight in minimal medium at 16°C and protein was purified as described above.

Circular dichroism (CD) measurement

Circular dichroism of mIHF was measured at a concentration of 10 – 20 μM protein in a quartz cuvette of 1 mm path length and analyzed using a Jasco J-815 CD spectrometer. Data were converted to mean residual ellipticity and corrected for difference in concentration (Greenfield, 2007). Far UV (195 – 250 nm) was used for protein, and near UV (250 – 340 nm) for DNA absorbance. Thermal unfolding was monitored over the whole spectra. Molar ellipticity changes at 222 nm were used to fit a sigmoidal model and calculate the melting temperature (GraphPad Prism). DNA titration was carried out at 20°C with equivalents of 0, 0.003, 0.006, 0.34, 0.5 and 0.7 equivalents DNA. A 40 bp dsDNA oligonucleotide (5' – CTGGAGGAGCTGGCAGCAGCGTTTCCGGGTGAT GGCTGGT – 3') was used. As mIHF does not contain amino acids that absorb at 260 nm, the signal was undisturbed in the transition between near and far UV.

Gel retardation assay

100 ng linear DNA was incubated with increasing amounts of recombinant mIHF (1

pmol – 2 nmol) and mIHF R86E (0.1 nmol – 2 nmol) for 10 min at room temperature and run on a 0.75% agarose gel in 0.5 x TBS (25 mM Tris-HCl pH 7.5, 50 mM NaCl). Post-run staining was performed with GelRed (Biotium).

Nuclear Magnetic Resonance Spectroscopy and Structure Determination

NMR measurements were carried out on 1.34 mM uniformly ^{13}C - and ^{15}N -labelled mIHF samples in 100 μl using Shigemi NMR tubes. Protein solution was prepared in 90% H_2O (50 mM NaPi buffer pH 7.5, 100 mM NaCl) and 10% $^2\text{H}_2\text{O}$. 0.2% sodium azide was added to prevent sample degradation. The NMR spectra were acquired at 14.1 T, 18.8 T and 23.5 T (600.55, 800.13 and 1000.30 MHz of proton Larmor frequency, respectively) on Bruker Avance III spectrometers, all equipped with a cryogenic cooled probehead. The experiments for determining the ^1H , ^{13}C and ^{15}N protein resonance assignment were performed at 14.1 T, using the automated assignment APSY routine. Standard CBCACONH (Hiller, 2008), HACACONH (Fiorito, 2006), and HACANH (Hiller, 2005) APSY experiments were used; the $\pi/2$ hard pulses were 11.18 μs , 14.45 μs and 37.2 μs for ^1H , ^{13}C and ^{15}N , respectively. The number of projections acquired were 56, 56, and 41 with 16 scans, 8, and 8 scans for CBCACONH, HACACONH, and HACANH, respectively, and with a direct acquisition time of 91.7 ms and a recycle delay of 0.9 s. The ^1H - ^{15}N 3D and ^1H - ^{13}C 3D HSQC NOESY spectra to obtain the structural restraints for the solution structural determination were performed at 18.8 T and 288 K using standard NMR sequences. The $\pi/2$ hard pulses were 10.52 μs , 13.6 μs and 35.0 μs for ^1H , ^{13}C and ^{15}N , respectively. The mixing times were 100 ms and 120 ms for ^1H - ^{15}N 3D and ^1H - ^{13}C 3D HSQC NOESY, respectively. Up to 1536 x 72 x 100 complex points were acquired for direct ^1H , indirect ^{15}N and ^1H dimensions, respectively; for ^1H - ^{13}C 3D HSQC NOESY 1088 x 80 x 90 complex points were acquired for direct ^1H , indirect ^{13}C and ^1H dimensions, respectively. The number of scans were 8, with a direct acquisition time of 84.86 ms and a recycle delay of 1.20 s. Data were processed with 1024 x 128 x 256 complex points matrix and square cosine window function.

The analysis of the APSY experiments was performed with the software UNIO-MATCH for automated backbone resonance assignment (Volk, 2008) and UNIO-ATNOS/ASCAN (Fiorito, 2008) for automated side-chain chemical shift assignment. The input for automated NOESY peak picking and NOE assignments with UNIO-ATNOS/CANDID (T Herrmann, 2002; Torsten Herrmann, 2002) consisted of the UNIO-MATCH chemical shift assignments for the polypeptide backbone, the UNIO-ATNOS/ASCAN output of side-chain chemical shift assignments, and the NOESY datasets. Table 1 reports NMR structural statistics from NOE assignment and structure calculation using ANUIO-ATNOS/CANDID.

¹⁵N-Relaxation Data

The backbone mobility of mIHF was investigated through ¹⁵N-relaxation measurements, acquired on the same sample described above at 298 K, corrected for the different temperature. Standard experiments were used to measure the longitudinal and transverse ¹⁵N relaxation times (¹H-¹⁵N HSQC T₁ and HSQC T₂) and the {¹H}-¹⁵N NOE HSQC at 23.5T. For the experiments, the $\pi/2$ hard pulses were 11.93 μ s, 13.75 μ s and 26.5 μ s for ¹H, ¹³C and ¹⁵N, respectively. The build-up times for the ¹⁵N T₁ were 0.010 s, 0.010 s, 0.100 s, 0.300 s, 0.500 s, 0.700 s, 1.00 s, 1.30 s, 1.60 s, 1.60 s, with two experiments acquired twice. For ¹⁵N T₂ the build-up times were 17 ms, 34 ms, 51 ms, 68 ms, 68 ms, 102 ms, 136 ms, 170 ms, 204 ms with one experiment acquired twice. For each build-up point the number of scans were 32 with 1.50 s of recycle delay for the ¹H-¹⁵N HSQC T₁, and 2.50 s of recycle delay for the ¹H-¹⁵N HSQC T₂ to allow for the probe duty cycle. For the {¹H}-¹⁵N NOE HSQC 64 scans were acquired for each experiment, with a saturation time of 6.0 s. Each point in the HSQC plane relaxation data was processed with a 2048 x 512 complex points matrix and a square cosine windows function for both direct and indirect dimensions. Relaxation data were analysed with the Protein Dynamics Center 2.4.6 (2016 Nov/14) software provided by Bruker BioSpin.

DNA titration experiments monitored by NMR

The protein was prepared at 200 μ M in 50 mM phosphate buffer pH 7.5, 100 mM NaCl,

with 10% ²H₂O. Titration of this sample with dsDNA (sequence 5'-AGCTCGTCAACGCCTT-3', 5 mM stock solution, purchased from Microsynth) was monitored through ¹H,¹⁵N HSQC spectra collected at 298 K on an AVANCEII-800 MHz spectrometer equipped with a CPTC ¹H,¹³C,¹⁵N 5 mm cryoprobe. The HSQC spectra were collected with 128 increments in the ¹⁵N dimension. Because generalized crosspeak broadening was also observed on top of the crosspeak-specific intensity decreases, the number of scans was increased throughout the titration so as to keep a workable signal-to-noise ratio (from 8 scans without added DNA to 64 scans at 2.2 equivalents).

For data analysis, the backbone ¹H,¹⁵N assignments derived from the automatic structure calculation were used. All data for titrations was acquired and processed with Bruker's TopSpin 2.0, and analysed with NMRFAM-Sparky. Chemical shift perturbations (CSP) were calculated from weighed differences in ¹H and ¹⁵N chemical shifts ($\Delta\delta$ H and $\Delta\delta$ N).

Atomic Force Microscopy

The *mihF* gene and promoter region were obtained by amplification with primers *mihF*-prom-F and *mihF*-prom-R, cloning of the product into pTOPO (ThermoFisher), giving rise to pNO75, and subsequent plasmid purification. pNO75 was cut with SpeI and BamHI and the resulting linear fragment was purified with the PCR purification kit (Sigma-Aldrich). I95 cosmid DNA consists of a 42.6 kb region (coordinates 4,292,000 – 4,326,000) of the *M. tuberculosis* chromosome cloned into pYUB412 (8.6 kb) and amplified in *E. coli* (Bange, 1999). Cosmid DNA was extracted using a Midi Prep kit (Promega) and eluted in TE buffer (10 mM Tris, 1 mM EDTA).

DNA (either linear fragment, cosmid I95) at 0.5 ng μ l⁻¹ was mixed with varying concentrations of mIHF, brought to a final volume of 20 μ l in buffer II (5 mM MgCl₂, 5 mM Tris-HCl pH 7.5, 50 mM NaCl, used with pGA118) or APTES (3-Aminopropyltriethoxysilane) water (used with I95 and linear *mihF* gene + promoter), deposited on freshly cleaved mica and incubated for 5 min at 25°C.

AFM images were collected using a MultiMode SPM with a Nanoscope III controller (Veeco Instruments, Santa Barbara, CA, USA) operated in tapping-mode in air. The AFM cantilevers had a spring constant of 5 N/m (Bruker cantilevers, TAP150A) with resonance frequencies ranging between 120 and 160 kHz. All AFM images consist of 512×512 pixels with scan frequency ≤ 1 Hz. Each experiment was performed at least in duplicate and AFM images were obtained at several separate mica locations. Only DNA complexes that were completely visible in an AFM image were considered for statistical analysis. Images were simply flattened using the Gwyddion software (Version 2.22), and no further image processing was carried out (Nečas, 2012).

DNA molecules were traced using “DNA Trace” software previously described (Mikhaylov, 2013). Based on the statistics of tens of individual molecules the contour length, the radius of gyration as well as the bond correlation function were calculated by the software.

Phosphorylation assay

mIHF (0.6 nmol) or GarA (0.4 nmol) were mixed with PknB or PknG at 1:10 ratio (substrate:kinase), incubated at room temperature in phosphorylation buffer (50 mM HEPES pH 7.5, 1 mM DTT, 0.01% v/v Brij 35, 2 mM MnCl₂ and 100 μ M ATP) for 3 minutes, 10 minutes or 3 hours and run on an SDS-gel (NuPAGE 4 – 12% Bis-Tris Gel, Invitrogen). The gel was subsequently stained with the ProQ Diamond Phosphoprotein Gel Stain (ThermoFisher) and imaged with a Fusion FX gel imager (Vilber). After imaging, the gel was stained with Coomassie blue to detect all proteins present. Band intensity of the ProQ- and Coomassie-stained gel was analysed with ImageJ and plotted relative to the non-

phosphorylated band intensity. See Fig. S1 for pictures of the phospho-stained and total-protein stained gels.

LC-MS/MS

Bands of interest were excised from SDS-PAGE gels and in-gel digested with trypsin or chymotrypsin. After gel extraction, samples were dried by vacuum centrifugation and analysed by LC-MS/MS. Peptides were separated on a Dionex Ultimate 3000nano UPLC system in-line coupled to a high-resolution mass spectrometer (ThermoFisher Scientific). A pre-column (Acclaim PepMap C18 Trap) was used to capture the samples and separation was performed over a 90-min gradient using a capillary column (Acclaim PepMap C18 RSLC) at 250nl/min. Positive data-dependent acquisition mode was used and precursor peptides with charge > 2 were fragmented. Database search was performed with Proteome Discoverer 1.4 on Mascot, Sequest and MS-Amanda against Tuberculist R27 database. Met oxidation, Ser-Thr-Tyr phosphorylation and peptide N-acetylation were set as variable modifications while Cys carbamidomethylation was set as fixed modification. Scaffold 4.75 was used for data inspection.

ACKNOWLEDGMENTS

We would like to thank the mass spectrometry facility at EPFL for their technical support, Stefanie Boy-Röttger for the purification of GarA, Jérémie Piton for mutagenesis suggestions and Rajkumar Singh for general advice about protein work. This work was supported by the Swiss National Science Foundation (grant 31003A-162641 to STC).

REFERENCES

- Bange, F., Collins, F., and Jr., J.W. (1999) Survival of mice infected with *Mycobacterium smegmatis* containing large DNA fragments from *Mycobacterium tuberculosis*. *Tuber Lung Dis* **79**: 171–80.
- Bhowmick, Bhowmick, T., Ghosh, S., Dixit, K., Ganesan, V., Ramagopal, U. a, et al. (2014) Targeting *Mycobacterium tuberculosis* nucleoid-associated protein HU with structure-based inhibitors. *Nat. Commun.* **5**: 4124.
- Blasco, B., Chen, J.M., Hartkoorn, R., Sala, C., Uplekar, S., Rougemont, J., et al. (2012) Virulence regulator EspR of *Mycobacterium tuberculosis* is a nucleoid-associated protein. *PLoS Pathog.* **8**: e1002621.
- Blasco, B., Stenta, M., Alonso-Sarduy, L., Dietler, G., Peraro, M.D., Cole, S.T., and Pojer, F. (2011) Atypical DNA recognition mechanism used by the EspR virulence regulator of *Mycobacterium tuberculosis*. *Mol. Microbiol.* **82**: 251–64.
- Coordinators, N.R. (2017) Database Resources of the National Center for Biotechnology Information. *Nucleic Acids Res.* **45**: D12–D17.
- Dame, R.T. and Tark-Dame, M. (2016) Bacterial chromatin: converging views at different scales. *Curr. Opin. Cell Biol.* **40**: 60–65.
- Dillon, S.C. and Dorman, C.J. (2010) Bacterial nucleoid-associated proteins, nucleoid structure and gene expression. *Nat. Rev. Microbiol.* **8**: 185–95.
- Fiorito, F., Herrmann, T., Damberger, F.F., and Wüthrich, K. (2008) Automated amino acid side-chain NMR assignment of proteins using 13C- and 15N-resolved 3D [1H, 1H]-NOESY. *J. Biomol. NMR* **42**: 23–33.
- Fiorito, F., Hiller, S., Wider, G., and Wüthrich, K. (2006) Automated resonance assignment of proteins: 6D APSY-NMR. *J. Biomol. NMR* **35**: 27–37.
- Fortuin, S., Tomazella, G.G., Nagaraj, N., Sampson, S.L., Gey van Pittius, N.C., Soares, N.C., et al. (2015) Phosphoproteomics analysis of a clinical *Mycobacterium tuberculosis* Beijing isolate: Expanding the mycobacterial phosphoproteome catalog. *Front. Microbiol.* **6**: 1–12.
- Fulcrand, G., Dages, S., Zhi, X., Chapagain, P., Gerstman, B.S., Dunlap, D., and Leng, F. (2016) DNA supercoiling, a critical signal regulating the basal expression of the lac operon in *Escherichia coli*. *Sci. Rep.* **6**: 19243.
- Greenfield, N.J. (2007) Using circular dichroism collected as a function of temperature to determine the thermodynamics of protein unfolding and binding interactions. *Nat. Protoc.* **1**: 2527–2535.
- Gupta, M., Sajid, A., Sharma, K., Ghosh, S., Arora, G., Singh, R., et al. (2014) HupB, a nucleoid-associated protein of *Mycobacterium tuberculosis*, is modified by serine/threonine protein kinases in vivo. *J. Bacteriol.* **196**: 2646–2657.
- Herrmann, T., Guntert, P., and Wuthrich, K. (2002) Protein NMR structure determination with automated NOE-identification in the NOESY spectra using the new software ATNOS. *J. Biomol. NMR* **24**: 171–189.
- Herrmann, T., Güntert, P., and Wüthrich, K. (2002) Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA. *J. Mol. Biol.* **319**: 209–227.
- Hiller, S., Fiorito, F., Wuthrich, K., and Wider, G. (2005) Automated projection spectroscopy (APSY). *Proc. Natl. Acad. Sci.* **102**: 10876–10881.
- Hiller, S., Wider, G., and Wüthrich, K. (2008) APSY-NMR with proteins: Practical aspects and backbone assignment. *J. Biomol. NMR* **42**: 179–195.
- Japaridze, A., Muskhelishvili, G., Benedetti, F., Gavrilidou, A.F.M., Zenobi, R., De Los Rios, P., et al. (2017) Hyperplectonemes: A Higher Order Compact and Dynamic DNA Self-Organization. *Nano Lett.* **acs.nanolett.6b05294**.
- Jones, D.T., Buchan, D., Cozzetto, D., and Bryson, K. (2017) PSIPRED.
- Leonard, A.C. and Grimwade, J.E. (2014) Initiation of DNA Replication. *EcoSal Plus.* **4**: 1–44.
- Lew, J.M., Kapopoulou, A., Jones, L.M., and Cole, S.T. (2011) TubercuList - 10 years after.

- Tuberculosis* **91**: 1–7.
- Lupas, A., Van Dyke, M., and Stock, J. (1991) Predicting Coiled Coils from Protein Sequences. *Science* (80-.). **252**: 1162*1164.
- Mikhaylov, A., Sekatskii, S.K., and Dietler, G. (2013) DNA Trace: A Comprehensive Software for Polymer Image Processing. *J. Adv. Microsc. Res.* **8**: 241–245.
- Mishra, A., Vij, M., Kumar, D., Taneja, V., Mondal, A.K., Bothra, A., et al. (2013) Integration Host Factor of *Mycobacterium tuberculosis*, mIHF, Compacts DNA by a Bending Mechanism. *PLoS One* **8**: e69985.
- Nečas, D. and Klapetek, P. (2012) Gwyddion: an open-source software for SPM data analysis. *Open Phys.* **10**: 181–188.
- O'Hare, H.M., Durán, R., Cerveñansky, C., Bellinzoni, M., Wehenkel, A.M., Pritsch, O., et al. (2008) Regulation of glutamate metabolism by protein kinases in mycobacteria. *Mol. Microbiol.* **70**: 1408–1423.
- Pedulla, M.L. and Hatfull, G.F. (1998) Characterization of the mIHF gene of *Mycobacterium smegmatis*. *J. Bacteriol.* **180**: 5473–7.
- Pinson, V., Takahashi, M., and Rouviere-Yaniv, J. (1999) Differential binding of the *Escherichia coli* HU, homodimeric forms and heterodimeric form to linear, gapped and cruciform DNA. *J Mol Biol* **287**: 485–497.
- Prieto, A.I., Kahramanoglou, C., Ali, R.M., Fraser, G.M., Seshasayee, A.S.N., and Luscombe, N.M. (2012) Genomic analysis of DNA binding and gene regulation by homologous nucleoid-associated proteins IHF and HU in *Escherichia coli* K12. *Nucleic Acids Res.* **40**: 3524–37.
- Prisic, S., Dankwa, S., Schwartz, D., Chou, M.F., Locasale, J.W., Kang, C.-M., et al. (2010) Extensive phosphorylation with overlapping specificity by *Mycobacterium tuberculosis* serine/threonine protein kinases. *Proc. Natl. Acad. Sci.* **107**: 7521–7526.
- Richard-greenblatt, M. and Av-gay, Y. (2017) Epigenetic Phosphorylation Control of *Mycobacterium tuberculosis* Infection and Persistence. **5**: 1–23.
- Schubert, O.T., Ludwig, C., Kogadeeva, M., Zimmermann, M., Rosenberger, G., Gengenbacher, M., et al. (2015) Absolute Proteome Composition and Dynamics during Dormancy and Resuscitation of *Mycobacterium tuberculosis*. *Cell Host Microbe* **18**: 96–108.
- Summers, E.L., Meindl, K., Usón, I., Mitra, A.K., Radjainia, M., Colangeli, R., et al. (2012) The structure of the oligomerization domain of Lsr2 from *Mycobacterium tuberculosis* reveals a mechanism for chromosome organization and protection. *PLoS One* **7**: e38542.
- Swiercz, J.P., Nanji, T., Gloyd, M., Guarné, A., and Elliot, M. a (2013) A novel nucleoid-associated protein specific to the actinobacteria. *Nucleic Acids Res.* **41**: 4171–84.
- Villarino, A., Duran, R., Wehenkel, A., Fernandez, P., England, P., Brodin, P., et al. (2005) Proteomic identification of *M. tuberculosis* protein kinase substrates: PknB recruits GarA, a FHA domain-containing protein, through activation loop-mediated interactions. *J. Mol. Biol.* **350**: 953–963.
- Volk, J., Herrmann, T., and Wüthrich, K. (2008) Automated sequence-specific protein NMR assignment using the memetic algorithm MATCH. *J. Biomol. NMR* **41**: 127–138.
- WHO (2016) Global Tuberculosis Report 2016. *Cdc* 2016 214.

SUPPLEMENTARY MATERIAL

Table S1. Loop structures in linear 1 kb fragment containing *miHF* gene and *miHF* promoter region.

miHF [ng/μl]	miHF [pmol]	DNA [pmol]	miHF bp ⁻¹	miHF wild type		miHF R86E mutant	
				looped fragments [%]	n fragments	looped fragments [%]	n fragments
control	0	0.015	0	9.5 ± 3	550	9.5 ± 3	550
5	4.61	0.015	0.304	25.0 ± 5	155	12.5 ± 5	359
10	9.22	0.015	0.608	30.0 ± 5	515	17.5 ± 5	557
15	13.82	0.015	0.912	50.0 ± 5	535	20.0 ± 5	506
20	18.43	0.015	1.22	53.5 ± 5	505	28.5 ± 5	587

Table S2. Loop length of 1 kb fragment containing *miHF* gene and *miHF* promoter region upon miHF binding.

miHF bp ⁻¹	miHF wild type		miHF R86E mutant	
	n DNA fragments	loop length [nm]	n DNA fragments	loop length [nm]
0	48	131 ± 10	32	109.84 ± 3
0.30	31	133 ± 10	50	110.06 ± 2.5
0.61	47	112 ± 10	50	107.68 ± 2
0.91	50	102 ± 10	50	105.12 ± 2
1.22	50	95 ± 10	50	104.42 ± 3

Table S3. Loop orientation of 1 kb fragment containing *miHF* gene and *miHF* promoter region.

miHF bp ⁻¹	miHF wild type		miHF R86E mutant	
	n DNA fragments	left-handed loops [%]	n DNA fragments	left-handed loops [%]
0	40	48	50	50
0.30	40	52	45	60
0.61	40	60	45	57
0.91	40	63	43	60
1.22	50	76	50	68

Table S4. Radius of gyration of I95 cosmid upon miHF binding.

miHF [ng/μl]	miHF [pmol]	I95 [fmol]	miHF bp ⁻¹	miHF wild type	
				Radius of gyration miHF wt [nm]	looped fragments [%]
10	9.22	0.36	0.60	502.5 ± 50	5
20	18.43	0.36	1.20	291.5 ± 50	< 10
30	27.65	0.36	1.80	243.4 ± 50	< 10
40	36.87	0.36	2.40	208.3 ± 50	< 10
60	55.30	0.36	3.60	176.3 ± 50	< 10

Radius of gyration was not assessed for the hyperplectonemic I95 in the absence of miHF.

Table S5. Primers used in this study.

Name	Restriction site	Sequence 5' – 3'
<i>mihF</i> -exp-F	NdeI	cgtaCATATGgtggcccttccccagttgacc
<i>mihF</i> -exp-R	BamHI	cgctGGATCCttaggcggagccgaactttcc
<i>mihF</i> -R86E-F	--	cgaggccacgaagctcgcgggtgggcgca
<i>mihF</i> -R86E-R	--	tgcgcccaccgcgagcttcgtggcctcg
<i>mihF</i> -prom-F	SpeI	ggatAATATTgcaaagaggcggtcgtgaaagc
<i>mihF</i> -prom-R	BamHI	ctatGGATCCggcggagccgaactttcc

Table S6. Plasmids used in this study.

Name	Purpose and characteristics	Reference
pNO72	Expression of 6 x His-tagged mIHF. Derived from pET28a.	This study
pNO84	Expression of 6 x His-tagged mIHF R86E. Derived from pNO72.	This study
pNO75	AFM studies. pTOPO (ThermoFisher) containing <i>mihF</i> and its promoter region (coordinates 1,563,213 – 1,564,266 bp on H37Rv genome).	This study
I95	42.6 kb cosmid, 4292 – 4326 kbp region of <i>M. tuberculosis</i> H37Rv genome cloned in pYUB412.	(Bange, 1999)

Supplementary Figure 1

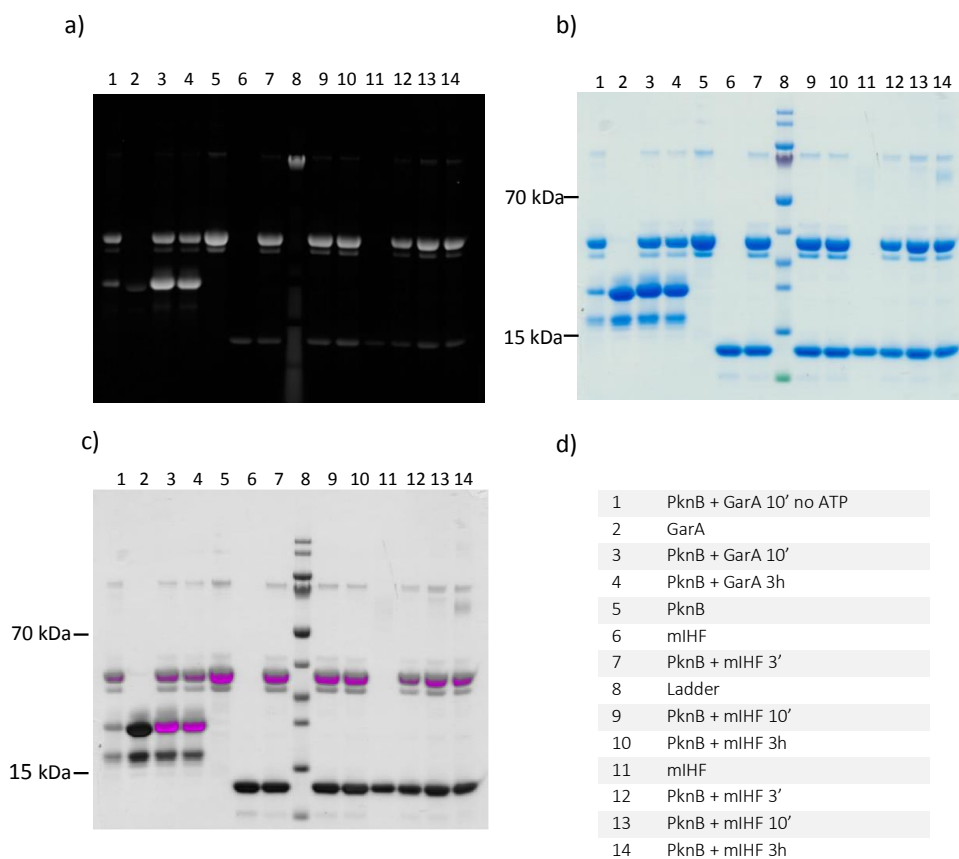


Fig. S1. Phosphorylation assay of mIHF and GarA. a) ProQ-stained phosphorylated proteins. See table in d) for correspondence between labels and reactions. b) The same gel was stained by Coomassie Blue to identify the loaded proteins. c) Overlay. Purple bands correspond to ProQ-stained proteins.

CHAPTER 4

Activity and Mode of Action of Chrysomycins in *Mycobacterium tuberculosis*

Nina T. Odermatt^a, Hui Guo^b, Lixin Zhang^b, Stewart T. Cole^a

^a École Polytechnique Fédérale de Lausanne, Global Health Institute, Station 19,
1015 Lausanne, Switzerland

^b Institute of Microbiology at the Chinese Academy of Sciences, Beijing, China

Contribution: Study design, MIC determination, intercalation assays, manuscript preparation

ABSTRACT

New drugs against *Mycobacterium tuberculosis* are in high demand, as the increasing rate of resistance renders many known medicines inefficient. The global transcriptional regulator Lsr2 was proposed as specific target of the natural compound chrysomycin. Here, we show that, despite of promising data proposing that chrysomycins bind specifically to Lsr2, this is indeed not the case. Chrysomycine was found to intercalate into the DNA and was toxic against eukaryotic cells. These results abrogated further development of chrysomycin as a lead compound against tuberculosis.

INTRODUCTION

Chrysomycin was first extracted from *Streptomyces* A-419, isolated from a soil sample in New York Botanical Garden, in 1955 and its activity successfully tested against several bacteriophages, fungi and bacteria, including *Mycobacterium smegmatis* (Strelitz et al. 1955). On the other hand, chrysomycin was not active on *Escherichia coli* (MIC > 50 µg ml⁻¹) nor on *Pseudomonas aeruginosa* (MIC > 25 µg ml⁻¹) (Strelitz et al. 1955). In 2013 Professor Lixin Zhang, from the Institute of Microbiology at the Chinese Academy of Sciences, Beijing, screened a natural compound library for activity against *M. bovis* BCG. One of the hits was chrysomycin, whose potency was subsequently tested and confirmed against various bacteria, including *Bacillus subtilis*, *M. tuberculosis*, *M. smegmatis* and *M. bovis* BGC. Prof. Zhang identified the marine species *Streptomyces qinglanensis* as capable of producing chrysomycin. Our laboratory started a collaboration with Prof. Zhang and his PhD student, Hui Guo, to investigate and prove the mode of action of chrysomycin against *M. tuberculosis*.

Hui Guo optimized chrysomycin production in *Streptomyces* sp. Strain MS751 and synthesised the derivatives chrysomycins B – H. He found the minimal inhibitory concentration (MIC) of chrysomycin A against *M. tuberculosis* H37Ra (a non-virulent relative to the laboratory strain H37Rv) to be below 1 µg ml⁻¹ and searched for the putative target of this compound by a pull-down screen of *M. bovis*

BCG whole-cell lysates. Three proteins were recovered after the pull-down. Mass spectrometry analysis identified two hypothetical proteins (Rv3169 and Rv2766) and Lsr2.

Surface-plasmon resonance experiments conducted in Prof. Zhang's laboratory confirmed that chrysomycins directly bound to the Lsr2 protein. Moreover, chrysomycin inhibited and reversed Lsr2-DNA binding in a dose-dependent manner and introduced an altered colony morphology phenotype in *M. smegmatis* similar to *lsr2* knockout strains.

Lsr2 is a highly conserved small nucleoid associated protein (NAP) that binds preferentially to AT-rich sequences and forms oligomers (Chen et al. 2008). It was proposed that Lsr2 had a similar function to H-NS in *E. coli*, as it can compensate for *hns* deletion in *E. coli*, and vice versa *E. coli hns* can complement an *lsr2* knockout in *M. smegmatis* (Gordon et al. 2008). Typical for a NAP, Lsr2 has over 800 binding sites on the *M. tuberculosis* chromosome (Gordon et al. 2010). While Lsr2 is dispensable in *M. smegmatis*, it was long thought to be essential in *M. tuberculosis* because of the lack of success to generate a knockout mutant (Gordon et al. 2010). However, Bartek and colleagues more recently deleted the *lsr2* gene from the *M. tuberculosis* genome and identified the differentially regulated genes in the mutant strain. Most of these were found to be involved in survival in high- as well as in low-oxygen environments (Bartek et al. 2014).

Zafirlukast, a drug used in prophylactic treatment of asthma, was reported to inhibit binding of Lsr2 to DNA, implying a similar mode of action as chrysomycin. Zafirlukast was never tested on virulent *M. tuberculosis* and showed only a modest MIC against *M. smegmatis*, therefore it was not used as an anti-tuberculosis drug (Pinault et al. 2013). Chrysomycins, on the other hand, have a low MIC in *M. tuberculosis* H37Ra and together with the proof that they specifically bind to Lsr2 and not only block interaction with DNA but also reverse Lsr2-DNA binding, suggested that chrysomycin was a promising compound to treat tuberculosis. As this mechanism had not been exploited before in tuberculosis treatment, we planned to completely characterize the effect of

chrysomycin on the virulent *M. tuberculosis* H37Rv strain and confirm the target of the compound by different methods for a future development as lead-compound for an anti-tuberculosis drug.

RESULTS

Mutagenicity and Cytotoxicity of Chrysomycin

Prof. Zhang's laboratory kindly provided us with chrysomycin A (ChryA) and chrysomycin B (ChryB) (See Figure 1 for chemical structure), the two most active compounds. Mutagenicity was assessed using the SOS-Chromotest. Neither of the compounds showed any mutagenic activity at 10 $\mu\text{g ml}^{-1}$ concentration. Toxicity was evaluated as median toxic dose 50 and 99 (TD_{50} and TD_{99}), the concentrations at which a compound exerts its toxic action against 50% and 99% of the tested cells, respectively. TD_{50} in HepG2 cells was 0.1 $\mu\text{g ml}^{-1}$ for both compounds, while TD_{99} was 10 $\mu\text{g ml}^{-1}$ for ChryA and 1.25 $\mu\text{g ml}^{-1}$ for ChryB. The concentrations causing toxic effects are therefore similar to MIC_{99} in H37Ra, thus questioning further development of the compounds.

Minimal Inhibitory Concentration upon Lsr2 Overexpression

MIC_{90} in *M. bovis* BCG, *M. tuberculosis* H37Ra and *M. smegmatis* were assessed in Prof. Zhang's laboratory and are reported in Table 1. ChryA was slightly more active than ChryB. MIC_{99} in *M. tuberculosis* H37Rv was

similar to MIC_{90} tested in the other species with 0.13 $\mu\text{g ml}^{-1}$ for ChryA and 1.56 $\mu\text{g ml}^{-1}$ for ChryB. The streptomycin-dependent strain 18b, used as a model of non-replicating persistence, was inhibited to 96% and 75%, respectively, by the two compounds (Table 1). Both are therefore active against growing as well as non-growing *M. tuberculosis*.

Table 1. MIC of Chrysomycin A (A) and B (B) against different strains.

	MIC_{90} [$\mu\text{g ml}^{-1}$]			MIC_{99} [$\mu\text{g ml}^{-1}$]
	BCG	Ra	smeg	Rv
A	0.1	0.39	0.78	0.13
B	6.25	1.56	1.56	1.56
Rif				0.007

SS18b	Inh_{max} [%]	$\text{Inh } 1 \mu\text{g ml}^{-1}$ [%]
A	96	45
B	75	5
Rif	46	33

MIC was measured on *M. bovis* BCG (BCG), *M. tuberculosis* H37Ra (Ra) and *M. smegmatis* (smeg) as well as on *M. tuberculosis* H37Rv (Rv, upper panel) and the non-replicating model strain SS18b (lower panel). Inh_{max} : maximal inhibition at 10 $\mu\text{g ml}^{-1}$. $\text{Inh } 1 \mu\text{g ml}^{-1}$: inhibition at 1 $\mu\text{g ml}^{-1}$. Rifampicin (Rif) was used as a control.

Chrysomycin was demonstrated to bind to Lsr2 and abrogate protein-DNA interaction by Prof. Zhang's group. To further validate these findings, we tested the MIC on different strains overexpressing *lsr2*, since higher amounts of

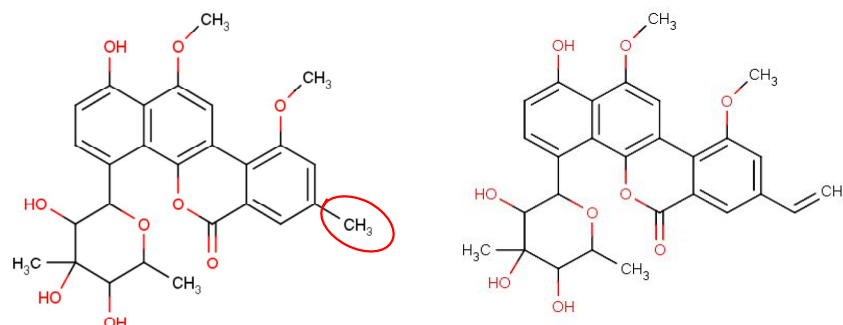


Figure 1. Chemical structures of chrysomycin A (left) and chrysomycin B (right). The chemical entities of the molecule that differ from the natural compound chrysomycine A to the derivatives is circled in red.

Table 2. MIC of Chrysomycin A (A) and B (B) against *Lsr2*-overexpressing strains.

	pMV261	pMV261:: <i>lsr2</i>	pMY769:: <i>lsr2</i>	pMY769:: <i>lsr2</i> + prist
A	0.35	0.3	0.156	0.156
B	3.1	3.1	1.56	1.56
Rif	0.001	0.002	0.001	0.001

MIC was tested on *M. tuberculosis* H37Rv transformed with the indicated plasmids. *Lsr2* expression by pMY769::*lsr2* was induced with 2 $\mu\text{g ml}^{-1}$ pristinamycin (prist). Rifampicin (Rif) was used as a control.

target protein should cause an increase in the MIC.

Two different overexpression systems were used to increase *Lsr2* levels: in the first one (plasmid pMV261::*lsr2*), *lsr2* was cloned downstream of a constitutively active and strong promoter (P_{hsp60}) in an episomal vector,

while in the second one transcription of *lsr2* was induced by pristinamycin (plasmid pMY769::*lsr2*). Quantitative reverse-transcription PCR (qRT-PCR) confirmed that *lsr2* mRNA levels were 4.9 times higher in pMV261::*lsr2*- and 4.3 times higher in pMY769::*lsr2*-containing strains compared to the empty vector or non-induced controls. Protein levels were estimated by immunoblot. Surprisingly, despite the qRT-PCR data, protein amounts were maximally 1.3-fold higher for both constructs (Fig. 2). MIC was evaluated for the strains carrying both overexpression systems, but did not show any difference between the wildtype and *lsr2* overexpressing strains (Table 2). The minor increase in *Lsr2* protein levels in the overexpressing strains might be the reason for the lack of effect on MIC values.

Transcriptional Response of *M. tuberculosis* Upon Incubation with ChryA

As we were not able to confirm *Lsr2* as the specific target of chrysomycins by *lsr2* overexpression, transcriptional response upon addition of chrysomycin was evaluated by RNA-sequencing (RNA-seq). A pleiotropic effect was observed in *M. tuberculosis* H37Rv treated with ChryA of several hundred differentially expressed genes.

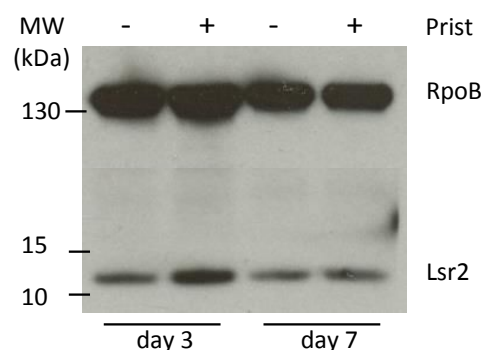


Figure 2. Pristinamycin-induced *Lsr2* expression detected by immunoblots. A representative image is shown. Pristinamycin absence and presence (2 $\mu\text{g } \mu\text{l}^{-1}$) is indicated by “-” and “+”, respectively.

Of the 201 genes listed upon applying a 4-fold cut-off and false discovery rate (FDR) < 0.01, 115 were downregulated and 86 were upregulated (Supplementary Table 1). Most of the upregulated genes are involved in DNA repair, such as *dnaE2* (DNA polymerase), *ruvC* (DNA repair at crossover junctions) and *recA* (the SOS-response induced recombinase A). Several of the downregulated genes were related to energy metabolism (subunits of the ATPase *atpG*, *atpC*, *atpD*), virulence factors (*espC*) and translation (ribosomal protein genes

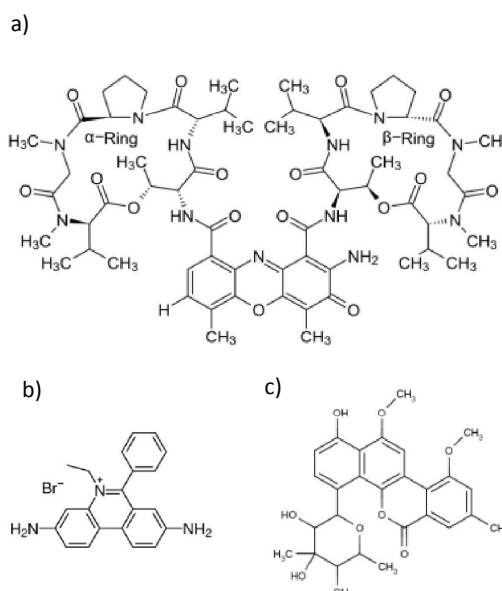


Figure 3. Chemical structures of DNA-intercalating compounds. a) Actinomycin D and b) Ethidium bromide share the tri-cyclic structure responsible for DNA intercalation, which is also present in a similar configuration in ChryA (c)

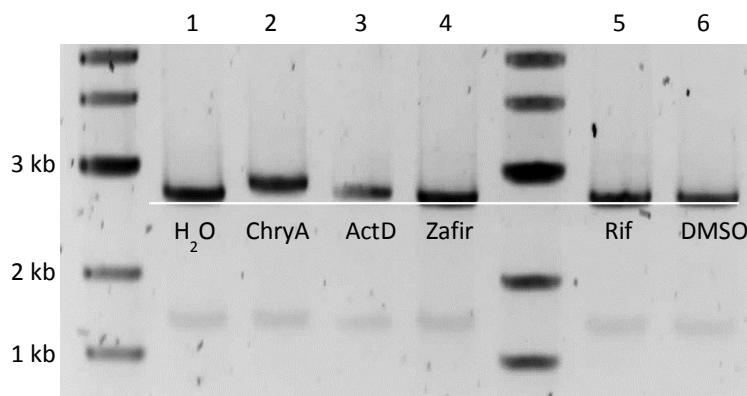


Figure 4. Agarose gel-shift assay to test for intercalation of ChryA into DNA. Linear pUC19 plasmid DNA was incubated with the indicated compounds and run on a 1% agarose gel. 1 kb ladder from New England Biolabs was used. The image was obtained by post-run staining with GelRed. Actinomycin D (ActD) was used as a positive control, whereas water, Zafirlukast (Zafir), Rifampicin (Rif) and the solvent of ChryA, DMSO, were used as negative controls. The white line represents progression of pUC19 relative to lane one, where the DNA was incubated with water.

rpmD, *rpsQ*, *rpsC*). No enrichment in a functional category was found. Overall, these data suggested that ChryA introduces unspecific DNA damage in *M. tuberculosis*.

ChryA Binds to DNA

To investigate the possibility that chrysomycins non-specifically intercalate into DNA, an agarose gel mobility shift assay was carried out. Linear plasmid DNA was incubated with different amounts of ChryA or with the solvent DMSO. Actinomycin D (ActD) was used as a positive control as it is a well-known intercalating agent (Figure 3 reports the chemical structures of ActD, ethidium bromide and ChryA) (Furlan et al. 2002). In addition, rifampicin was included as a negative control as well as zafirlukast, known to bind specifically to Lsr2 and inhibit interaction to DNA. The negative controls zafirlukast and rifampicin did not show any shift of the DNA. The retardation caused by ChryA was bigger than that observed with the positive control ActD (Fig. 4), thereby confirming the hypothesis that chrysomycin indeed intercalates into the DNA.

CONCLUSION

ChryA and ChryB were investigated for potential use as anti-tuberculosis drugs with a specific target in the global transcriptional regulator Lsr2. We were not able to confirm Lsr2 as their target by protein overexpression in *M. tuberculosis* nor by RNA-seq, as transcriptional analysis did not show any impact of ChryA on Lsr2-regulated genes. Further, ChryA caused plasmid retardation on agarose gel, suggesting that the compound intercalates into DNA, therefore explaining the toxicity that we measured in HepG2 cells at concentrations that equal the MIC. The latter data are consistent with those published by Strelitz and co-workers in 1955, when 5 mg of compound administered to mice caused loss of appetite and paralysis of the hind legs (Strelitz et al. 1955). Overall, these results advise against further development of ChryA and ChryB as leads for anti-tuberculosis drugs.

METHODS

Bacterial Strains and Growth Conditions

M. tuberculosis H37Rv was grown at 37°C in liquid Middlebrook 7H9 broth (Difco) supplemented with 10% Albumin-Dextrose-

Catalase (ADC), 0.2% glycerol and 0.05% Tween 80, or on solid Middlebrook 7H10 agar (Difco) supplemented with 10% Oleic acid-Albumin-Dextrose-Catalase (OADC) and 0.2% glycerol. Antibiotics were added where necessary (Hygromycin 50 $\mu\text{g ml}^{-1}$, kanamycin 20 $\mu\text{g ml}^{-1}$, streptomycin 20 $\mu\text{g ml}^{-1}$). All chemicals were purchased from Sigma-Aldrich, unless stated otherwise.

Resazurin Microtiter Assay (REMA)

REMA assays were performed in 7H9 broth in a 96 well-plate. *M. tuberculosis* was inoculated at $\text{OD}_{600} = 0.0001$ in 100 μl . Compounds were added at serial 2-fold dilutions, incubated for 7 days with the bacteria before addition of resazurin (0.025% w/v). After overnight incubation, fluorescence of resorufin was determined in a Tecan Infinite M200 microplate reader. Results were plotted in GraphPad Prism and MIC was determined by the Gompertz equation. Compounds were tested in duplicates.

The non-replicating model strain 18b is dependent on streptomycin for growth and streptomycin-starved bacteria (SS18b) were used in REMA (Zhang et al. 2012). Compounds were tested in triplicates. The maximal inhibition at 100 $\mu\text{g ml}^{-1}$ relative to the negative control not containing any compound is reported, as well as the inhibition at a specific concentration of the compounds.

SOS-Chromotest and Cytotoxicity Assay

The SOS-chromotest was carried out as previously described (Nair et al. 2000). The reporter strain *E. coli* PQ37 expresses beta-galactosidase when exposed to genotoxic agents. ChryA and ChryB were tested in duplicate at 10 $\mu\text{g ml}^{-1}$ for mutagenic activity. 4-nitroquinolone oxide (4-NQO, Sigma-Aldrich) was used as a positive control while isoniazid served as a negative control.

Cytotoxicity was measured as previously described (Magnet et al. 2010). HepG2 liver cells were incubated for 3 days at 37°C with serial dilutions of the compound to test in a 96-well microplate. Cell viability was determined by adding resazurin for 4 hours and by measuring fluorescence in a TECAN plate

reader. Viability was calculated as relative to untreated cells.

Plasmids, *lsr2* Expression and Validation

The inducible expression vector pMY769::*lsr2* was constructed by PCR amplification of the *lsr2* gene with *lsr2*-F 5'-gatccctagggcgaagaaagtaaccgtcaccttggtcg-3' and *lsr2*-R

5'-cgttgaactttcaggtcgccgctggtatgc-3', containing AvrII and HindIII restriction sites, respectively. The product was cloned into pMY769 and integration was validated by sequencing. Gene expression by the PIP-ON system (Forti et al. 2009) was induced by pristinamycin purified from pyostacin tablets by liquid chromatography (Hartkoorn et al. 2010). The overexpression plasmid pMV261::*lsr2* was constructed and kindly provided by Hui Guo. The *lsr2* gene is under control of the *hsp60* promoter and constitutively expressed at high level. *M. tuberculosis* H37Rv was transformed with both *lsr2* expression plasmids and the empty vector as a control, and plated on kanamycin and streptomycin for selection of pMV261 and pMY769 transformants, respectively, and checked by colony PCR. Overexpression of *lsr2* was confirmed by quantitative reverse-transcription PCR as described below, using primers 5'-cttgacgggggtgacctatga-3' and 5'-acattgtgccggttacgac-3' for *lsr2*. The gene encoding the housekeeping sigma factor *sigA* was used as an internal control with primers 5'-aaacagatcggaaggtagc-3' and 5'-ctggatcaggtcgagaaacg-3'.

Immunoblots were performed to assess the Lsr2 protein level in H37Rv transformed with pMV261::*lsr2* compared to the empty vector-carrying strain, and in H37Rv transformed with pMY769::*lsr2* in the presence and absence of the inducer pristinamycin, 3 and 7 days after induction. ImageJ was used to quantify the intensity of the bands. Pristinamycin was used at 1, 2 and 4 $\mu\text{g ml}^{-1}$.

RNA Preparation and Quantitative Reverse-Transcription PCR

M. tuberculosis H37Rv cultures were harvested by centrifugation, pellets were

resuspended in TRIzol Reagent (ThermoFisher) and stored at -80° C until further processing. Total RNA was extracted by bead-beating as previously described (Jungwirth et al. 2012). Integrity of RNA was checked by agarose gel electrophoresis, purity and amount of RNA were assessed by using a Nanodrop instrument and Qubit Fluorometric Quantitation (ThermoFisher) respectively. SuperScript III First-Strand Synthesis System (Invitrogen) was used to generate randomly primed cDNA from 500 ng of RNA, according to the manufacturer's recommendations.

Quantitative reverse-transcription PCR was carried out in duplicate on a 7900HT Sequence Detection System with Power SYBR Green PCR Master Mix (Applied Biosystems) according to the manufacturer's recommendations. The delta-delta Ct method was used for quantification of transcripts in different samples.

RNA-Sequencing

Exponentially growing *M. tuberculosis* H37Rv was incubated with ChryA at a concentration that equals 10 x MIC = 1 µg ml⁻¹ for 4 hours at 37°C. Biological duplicates were collected for the control (no compound added) and for the treated samples. Total RNA was subsequently extracted as described above and

subjected to library preparation as previously described (Odermatt et al. 2017). High-throughput sequencing according to the Illumina pipeline was carried out, followed by read mapping and analysis as previously described (Odermatt et al. 2017).

Gel-Shift Assay

Agarose gel mobility-shift assay was adapted after Furlan et al. (Furlan et al. 2002). Concentrations of 25, 50 and 100 ng ml⁻¹ of each compound were incubated with BamHI-digested pUC19 plasmid for 1 hour at 37°C. Samples were then electrophoresed on a 1% agarose gel and post-run stained with GelRed. Actinomycin D (ActD) was used as a positive control for intercalation into DNA and Rifampicin (Rif) as a non-intercalating negative control. Zafirlukast (Sigma-Aldrich), known to bind to Lsr2, was used as a second negative control.

ACKNOWLEDGMENTS

I would like to thank Professor Lixin Zhang and Hui Guo for sharing chrysomycins and the interactive collaboration on this project. Anthony Vocat for carrying out the SOS-chromotest and cytotoxicity assay, Andrej Benjak for help with the RNA-sequencing mapping, and Claudia Sala for general support.

SUPPLEMENTARY FILES

Supplementary Table 1. Differentially expressed genes (DGE) identified after incubation of H37Rv with ChryA.

gene-id	FC	FDR	gene name				
Rv0722	-24.1	2.6E-10	<i>rpmD</i>	Rv0589	-5.7	1.0E-21	<i>mce2A</i>
Rv0710	-22.9	1.8E-90	<i>rpsQ</i>	Rv0988	-5.7	6.0E-05	<i>Rv0988</i>
Rv3615c	-17.2	2.4E-83	<i>espC</i>	Rv1448c	-5.6	3.0E-32	<i>tal</i>
MTB000075	-15.9	9.6E-73	<i>ncrMT1234</i>	Rv1514c	-5.6	5.7E-08	<i>Rv1514c</i>
Rv0707	-15.5	5.6E-59	<i>rpsC</i>	Rv3119	-5.6	2.5E-21	<i>moaE1</i>
Rv1309	-13.6	8.7E-84	<i>atpG</i>	Rv2945c	-5.6	8.5E-26	<i>lppX</i>
Rv1311	-12.9	2.3E-68	<i>atpC</i>	Rv1195	-5.6	9.6E-39	<i>PE13</i>
Rv1310	-11.9	9.6E-86	<i>atpD</i>	Rv3281	-5.6	1.0E-35	<i>accE5</i>
Rv0709	-11.3	1.5E-58	<i>rpmC</i>	Rv3616c	-5.5	6.9E-41	<i>espA</i>
Rv3152	-11.3	4.3E-36	<i>nuoH</i>	Rv0684	-5.5	1.1E-35	<i>fusA1</i>
Rv0704	-11.1	4.3E-74	<i>rplB</i>	Rv2196	-5.4	2.8E-44	<i>qcrB</i>
Rv1438	-11.1	8.5E-49	<i>tpi</i>	Rv1107c	-5.4	5.8E-21	<i>xseB</i>
Rv3614c	-11.0	9.1E-73	<i>espD</i>	Rv2425c	-5.3	6.0E-20	<i>Rv2425c</i>
Rv1196	-10.9	5.9E-16	<i>PPE18</i>	Rv3051c	-5.3	6.2E-14	<i>nrdE</i>
Rv0718	-10.6	1.4E-65	<i>rpsH</i>	Rv0719	-5.3	2.1E-35	<i>rplF</i>
Rv0706	-9.9	1.2E-10	<i>rplV</i>	Rv2228c	-5.2	6.6E-19	<i>Rv2228c</i>
Rv0705	-9.7	4.8E-56	<i>rpsS</i>	Rv1612	-5.2	1.0E-34	<i>trpB</i>
Rv3155	-9.6	6.1E-14	<i>nuoK</i>	Rv1596	-5.2	1.8E-35	<i>nadC</i>
Rv0247c	-9.5	1.0E-37	<i>Rv0247c</i>	Rv0170	-5.1	1.2E-19	<i>mce1B</i>
Rv0173	-9.4	1.4E-55	<i>lprK</i>	Rv2475c	-5.1	1.4E-19	<i>Rv2475c</i>
Rv3920c	-9.3	2.5E-37	<i>Rv3920c</i>	Rv3492c	-5.1	8.5E-16	<i>Rv3492c</i>
Rv3148	-9.1	4.5E-53	<i>nuoD</i>	Rv1297	-5.0	1.6E-22	<i>rho</i>
Rv0723	-8.9	9.4E-29	<i>rplO</i>	Rv3799c	-5.0	2.5E-38	<i>accD4</i>
Rv2988c	-8.8	5.4E-06	<i>leuC</i>	Rv1465	-4.9	1.8E-11	<i>Rv1465</i>
Rv3613c	-8.8	6.1E-38	<i>Rv3613c</i>	Rv2084	-4.9	1.0E-17	<i>Rv2084</i>
Rv0703	-8.4	1.3E-51	<i>rplW</i>	Rv2244	-4.9	1.1E-36	<i>acpM</i>
Rv3612c	-8.3	3.5E-25	<i>Rv3612c</i>	Rv0908	-4.9	1.4E-23	<i>ctpE</i>
Rv0708	-8.1	1.5E-38	<i>rplP</i>	Rv0172	-4.9	2.7E-35	<i>mce1D</i>
Rv0720	-8.1	7.1E-06	<i>rplR</i>	Rv3921c	-4.9	1.2E-33	<i>Rv3921c</i>
Rv1198	-8.1	1.6E-48	<i>esxL</i>	Rv0097	-4.9	2.7E-31	<i>Rv0097</i>
Rv1447c	-7.6	2.4E-40	<i>zwf2</i>	Rv1772	-4.9	5.8E-10	<i>Rv1772</i>
Rv2947c	-7.5	5.7E-40	<i>pks15</i>	Rv2821c	-4.9	5.8E-13	<i>Rv2821c</i>
Rv2524c	-7.2	5.2E-24	<i>fas</i>	Rv3806c	-4.9	1.9E-21	<i>ubiA</i>
Rv1308	-7.2	1.4E-56	<i>atpA</i>	Rv0637	-4.8	3.1E-19	<i>hadC</i>
Rv0099	-7.1	3.4E-48	<i>fadD10</i>	Rv2818c	-4.8	7.7E-13	<i>Rv2818c</i>
Rv3824c	-6.8	9.5E-23	<i>papA1</i>	Rv1771	-4.7	9.3E-26	<i>Rv1771</i>
Rv2246	-6.8	1.2E-46	<i>kasB</i>	Rv1446c	-4.7	2.6E-17	<i>opcA</i>
Rv1437	-6.8	6.9E-39	<i>pgk</i>	Rv0526	-4.6	1.2E-14	<i>Rv0526</i>
Rv2245	-6.8	2.3E-51	<i>kasA</i>	Rv1445c	-4.6	1.7E-15	<i>devB</i>
Rv2247	-6.8	8.8E-33	<i>accD6</i>	Rv1307	-4.5	1.8E-16	<i>atpH</i>
Rv3157	-6.6	8.6E-42	<i>nuoM</i>	Rv3032	-4.5	3.8E-12	<i>Rv3032</i>
Rv1312	-6.5	1.6E-21	<i>Rv1312</i>	Rv0590	-4.3	1.6E-14	<i>mce2B</i>
Rv3150	-6.3	1.4E-39	<i>nuoF</i>	Rv3701c	-4.3	5.7E-14	<i>Rv3701c</i>
Rv3156	-6.3	3.0E-40	<i>nuoL</i>	Rv3115	-4.3	5.3E-19	<i>Rv3115</i>
Rv1613	-6.3	8.3E-40	<i>trpA</i>	Rv1466	-4.3	1.4E-17	<i>Rv1466</i>
Rv2276	-6.2	1.2E-27	<i>cyp121</i>	Rv1199c	-4.2	3.9E-20	<i>Rv1199c</i>
Rv0717	-6.1	2.6E-33	<i>rpsN1</i>	Rv1662	-4.2	2.3E-16	<i>pks8</i>
Rv2195	-6.1	1.4E-48	<i>qcrA</i>	Rv1098c	-4.2	5.8E-22	<i>fum</i>
Rv0351	-6.0	7.8E-08	<i>grpE</i>	Rv2392	-4.2	4.7E-21	<i>cysH</i>
Rv0640	-6.0	1.1E-37	<i>rplK</i>	Rv0952	-4.2	5.8E-26	<i>sucD</i>
Rv2949c	-6.0	4.5E-06	<i>Rv2949c</i>	Rv2476c	-4.1	7.3E-32	<i>gdh</i>
Rv0721	-5.9	4.5E-05	<i>rpsE</i>	Rv3147	-4.1	4.9E-21	<i>nuoC</i>
Rv0098	-5.8	4.7E-30	<i>fcoT</i>	Rv1362c	-4.1	3.0E-12	<i>Rv1362c</i>
Rv1197	-5.8	8.0E-45	<i>esxK</i>	Rv3301c	-4.1	2.6E-16	<i>phoY1</i>
Rv2948c	-5.7	5.8E-26	<i>fadD22</i>	Rv0696	-4.1	8.8E-15	<i>Rv0696</i>
				Rv2881c	-4.0	6.7E-23	<i>cdsA</i>

4 – Chrysomycins

Rv1513	-4.0	1.6E-12	<i>Rv1513</i>	Rv2225	5.7	2.4E-36	<i>panB</i>
Rv1047	-4.0	3.0E-11	<i>Rv1047</i>	Rv3642c	5.7	3.1E-06	<i>Rv3642c</i>
Rv3158	-4.0	3.4E-24	<i>nuoN</i>	Rv2720	5.7	9.6E-40	<i>lexA</i>
Rv3479	-4.0	3.1E-25	<i>Rv3479</i>	Rv1702c	5.9	8.3E-34	<i>Rv1702c</i>
Rv0029	4.0	9.5E-23	<i>Rv0029</i>	Rv1377c	5.9	5.0E-33	<i>Rv1377c</i>
MTB000019	4.0	2.5E-29	<i>rrs</i>	Rv0485	5.9	5.1E-42	<i>Rv0485</i>
Rv0834c	4.1	7.1E-11	<i>PE_PGRS14</i>	Rv3296	6.1	1.1E-39	<i>lhr</i>
Rv3036c	4.1	7.0E-07	<i>TB22.2</i>	Rv1765c	6.1	4.8E-37	<i>Rv1765c</i>
MTB000042	4.1	4.0E-06	<i>ssr</i>	Rv1986	6.2	1.4E-29	<i>Rv1986</i>
Rv2656c	4.1	8.0E-20	<i>Rv2656c</i>	Rv3384c	6.3	4.1E-29	<i>vapC46</i>
Rv0842	4.2	1.9E-21	<i>Rv0842</i>	Rv0275c	6.3	2.8E-20	<i>Rv0275c</i>
MTB000020	4.2	1.1E-08	<i>rrl</i>	Rv1277	6.4	1.3E-21	<i>Rv1277</i>
Rv1929c	4.2	7.4E-23	<i>Rv1929c</i>	Rv2734	6.4	3.0E-40	<i>Rv2734</i>
MTB000026	4.3	3.8E-03	<i>rnpB</i>	Rv2015c	6.4	1.5E-38	<i>Rv2015c</i>
Rv2369c	4.4	2.0E-22	<i>Rv2369c</i>	Rv0727c	6.5	5.2E-16	<i>fucA</i>
Rv3191c	4.4	2.4E-27	<i>Rv3191c</i>	Rv0116c	6.6	8.9E-19	<i>ldtA</i>
Rv1215c	4.4	4.8E-27	<i>Rv1215c</i>	Rv3585	6.7	1.9E-41	<i>radA</i>
Rv0441c	4.4	7.1E-24	<i>Rv0441c</i>	Rv2885c	6.7	1.4E-40	<i>Rv2885c</i>
Rv1317c	4.4	6.2E-25	<i>alkA</i>	Rv2737c	6.7	3.1E-36	<i>recA</i>
Rv1831	4.5	7.3E-29	<i>Rv1831</i>	Rv2594c	7.0	2.0E-46	<i>ruvC</i>
Rv1052	4.5	5.2E-20	<i>Rv1052</i>	Rv1833c	7.3	8.1E-15	<i>Rv1833c</i>
Rv3668c	4.5	1.4E-25	<i>Rv3668c</i>	Rv1945	7.6	2.4E-41	<i>Rv1945</i>
Rv1356c	4.5	8.1E-17	<i>Rv1356c</i>	Rv0750	7.8	2.5E-20	<i>Rv0750</i>
Rv2593c	4.6	1.4E-19	<i>ruvA</i>	Rv3466	7.8	2.0E-25	<i>Rv3466</i>
Rv3641c	4.6	3.0E-13	<i>fic</i>	Rv2979c	8.1	2.3E-46	<i>Rv2979c</i>
Rv3270	4.6	8.2E-31	<i>ctpC</i>	Rv2660c	8.5	2.2E-09	<i>Rv2660c</i>
Rv1390	4.6	2.9E-12	<i>rpoZ</i>	Rv3828c	8.7	1.2E-21	<i>Rv3828c</i>
Rv2657c	4.6	6.7E-21	<i>Rv2657c</i>	Rv0539	9.1	4.3E-13	<i>Rv0539</i>
Rv1148c	4.7	3.1E-28	<i>Rv1148c</i>	Rv3536c	9.9	1.4E-46	<i>hsaE</i>
Rv3190A	4.7	2.0E-26	<i>Rv3190A</i>	Rv2827c	9.9	1.5E-50	<i>Rv2827c</i>
Rv3095	4.7	3.0E-26	<i>Rv3095</i>	Rv3074	10.4	2.0E-64	<i>Rv3074</i>
Rv2876	4.9	1.4E-06	<i>Rv2876</i>	Rv0605	10.7	1.3E-59	<i>Rv0605</i>
Rv2896c	4.9	8.7E-23	<i>Rv2896c</i>	Rv0515	11.1	2.2E-73	<i>Rv0515</i>
Rv2100	4.9	1.5E-30	<i>Rv2100</i>	Rv3451	11.3	1.3E-61	<i>cut3</i>
Rv2251	5.0	2.4E-23	<i>Rv2251</i>	Rv1000c	12.3	1.2E-12	<i>Rv1000c</i>
Rv3581c	5.3	5.8E-35	<i>ispF</i>	Rv3289c	12.9	6.1E-59	<i>Rv3289c</i>
Rv0182c	5.3	3.4E-31	<i>sigG</i>	Rv2886c	14.3	2.1E-69	<i>Rv2886c</i>
Rv3555c	5.4	2.5E-24	<i>Rv3555c</i>	Rv0336	14.5	1.6E-87	<i>Rv0336</i>
Rv3424c	5.4	1.9E-26	<i>Rv3424c</i>	Rv3290c	16.9	4.0E-100	<i>lat</i>
Rv2579	5.4	5.4E-28	<i>dhaA</i>	Rv3776	20.0	1.8E-102	<i>Rv3776</i>
Rv3385c	5.4	9.4E-17	<i>vapB46</i>	Rv1378c	22.9	3.4E-112	<i>Rv1378c</i>
Rv0606	5.5	3.0E-29	<i>Rv0606</i>	Rv1588c	25.0	5.3E-95	<i>Rv1588c</i>
Rv3777	5.5	3.9E-36	<i>Rv3777</i>	Rv3370c	27.6	2.2E-59	<i>dnaE2</i>
Rv0921	5.6	2.1E-28	<i>Rv0921</i>	Rv2719c	29.5	4.1E-21	<i>Rv2719c</i>
MTB000058	5.6	2.8E-07	<i>B11</i>	Rv3395c	37.9	9.4E-99	<i>Rv3395c</i>

False discovery rate was set < 0.01 and cut-off of fold change (FC) > 4.

REFERENCES

- Bartek, I. L., Woolhiser, L. K., Baughn, A. D., Basaraba, R. J., Jacobs, W. R., Lenaerts, A. J., & Voskuil, M. I. (2014). Mycobacterium tuberculosis Lsr2 Is a Global Transcriptional Regulator. *mbio*, 5(3), e01106-14. doi:10.1128/mBio.01106-14.Editor
- Chen, J. M., Ren, H., Shaw, J. E., Wang, Y. J., Li, M., Leung, A. S., et al. (2008). Lsr2 of Mycobacterium tuberculosis is a DNA-bridging protein. *Nucleic acids research*, 36(7), 2123–35. doi:10.1093/nar/gkm1162
- Forti, F., Crosta, A., & Ghisotti, D. (2009). Pristinamycin-inducible gene regulation in mycobacteria. *Journal of biotechnology*, 140(3–4), 270–7. doi:10.1016/j.jbiotec.2009.02.001
- Furlan, R. L. A., Garrido, L. M., Brumatti, G., Amarante, G. P., Martins, R. A., Facciotti, M. C. R., & Padilla, G. (2002). A rapid and sensitive method for the screening of DNA intercalating antibiotics. *Biotechnology Letters*, 24, 1807–1813.
- Gordon, B. R. G., Imperial, R., Wang, L., Navarre, W. W., & Liu, J. (2008). Lsr2 of Mycobacterium represents a novel class of H-NS-like proteins. *Journal of bacteriology*, 190(21), 7052–9. doi:10.1128/JB.00733-08
- Gordon, B. R. G., Li, Y., Wang, L., Sintsova, A., van Bakel, H., Tian, S., et al. (2010). Lsr2 is a nucleoid-associated protein that targets AT-rich sequences and virulence genes in Mycobacterium tuberculosis. *Proceedings of the National Academy of Sciences of the United States of America*, 107(11), 5154–9. doi:10.1073/pnas.0913551107
- Hartkoorn, R. C., Sala, C., Magnet, S. J., Chen, J. M., Pojer, F., & Cole, S. T. (2010). Sigma factor F does not prevent rifampin inhibition of RNA polymerase or cause rifampin tolerance in Mycobacterium tuberculosis. *Journal of Bacteriology*, 192(20), 5472–5479. doi:10.1128/JB.00687-10
- Jungwirth, B., Sala, C., Kohl, T. A., Uplekar, S., Baumbach, J., Cole, S. T., et al. (2012). High-resolution detection of DNA binding sites of the global transcriptional regulator GlxR in *Corynebacterium glutamicum*. *Microbiology (Reading, England)*, 159(Pt 1), 12–22. doi:10.1099/mic.0.062059-0
- Magnet, S., Hartkoorn, R. C., Székely, R., Pató, J., Triccas, J. A., Schneider, P., et al. (2010). Leads for antitubercular compounds from kinase inhibitor library screens. *Tuberculosis*, 90(6), 354–360. doi:10.1016/j.tube.2010.09.001
- Nair, P. P., Davis, K. E., Shami, S., & Lagerholm, S. (2000). The induction of SOS function in Escherichia coli K-12/PQ37 by 4- nitroquinoline oxide (4-NQO) and fecapentaenes-12 and -14 is bile salt sensitive: Implications for colon carcinogenesis. *Mutation Research - Fundamental and Molecular Mechanisms of Mutagenesis*, 447(2), 179–185. doi:10.1016/S0027-5107(99)00205-5
- Odermatt, N. T., Sala, C., Benjak, A., Kolly, G. S., Vocat, A., Lupien, A., & Cole, S. T. (2017). Rv3852 (H-NS) of Mycobacterium tuberculosis is not involved in nucleoid compaction and virulence regulation. *Journal of bacteriology*, (May), JB.00129-17. doi:10.1128/JB.00129-17
- Pinault, L., Han, J.-S., Kang, C.-M., Franco, J., & Ronning, D. R. (2013). Zafirlukast inhibits complexation of Lsr2 with DNA and growth of Mycobacterium tuberculosis. *Antimicrobial agents and chemotherapy*, 57(5), 2134–40. doi:10.1128/AAC.02407-12
- Strelitz, F., Flon, H., & Asheshov, I. N. (1955). Chrysomycin: a new antibiotic substance for bacterial viruses. *Journal of bacteriology*, 69(3), 280–3. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=357526&tool=pmcentrez&rendertype=abstract>
- Zhang, M., Sala, C., Hartkoorn, R. C., Dhar, N., Mendoza-Losana, A., & Cole, S. T. (2012). Streptomycin-starved Mycobacterium tuberculosis 18b, a drug discovery tool for latent tuberculosis. *Antimicrobial Agents and Chemotherapy*, 56(11), 5782–5789. doi:10.1128/AAC.01125-12

Conclusions and Perspectives

I conducted my PhD studies within the framework of the Swiss National Science Foundation (SNF)-financed project entitled “Integrated investigation of the ESX-1 protein secretion system of *Mycobacterium tuberculosis*”. My role fell into Workpackage 1, whose aim was the detailed investigation of the global genetic regulation in Mtb, with special attention to the ESX-1 regulatory mechanisms. Nucleoid associated proteins (NAPs) represented the main subject of this work, since we aimed to complete the characterization of this family of regulators which, among others, include EspR (Raghavan, 2008). In the same context, I also collaborated with my colleagues in Prof. Cole’s laboratory and performed experiments that involved ESX-1 mutant strains, mainly related to investigation of EspL. I examined their virulence and ultrastructure by means of *ex vivo* model systems and electron microscopy, respectively (Data not presented in this thesis). In this final part of the PhD thesis, I would like to summarize my main findings and discuss future perspectives.

NAPs are important global transcriptional regulators. While most Gram-negative bacteria possess a dozen NAPs, they are seemingly underrepresented in Gram-positive bacteria (Dillon, 2010). When I started my PhD in 2013, five NAPs had been identified in Mtb: Lsr2, HupB, EspR, mIHF and Rv3852 (H-NS). My work showed that Rv3852 does not act as a NAP and defined the function and structure of mIHF.

Rv3852 is not H-NS

No abnormal phenotype, altered nucleoid position or change in virulence was observed in a Δ *rv3852* deletion mutant in Mtb H37Rv. Curiously, *M. leprae* with its reduced genome (Cole, 2001) also possesses a copy of the *rv3852* gene (named *ml0067*), which encodes a protein with 44% amino acid identity with Rv3852. Genes present on the 3.3 Mbp chromosome of *M. leprae* that did not undergo pseudogenization are usually considered of

important function. Often, their orthologues in Mtb are essential for bacterial survival or virulence. Regarding Rv3852 though, neither is the case. The lack of a phenotype in the mutant strain may be explained by the specific experimental conditions tested, although our experiments included *in vitro*, *ex vivo* and *in vivo* assays. Of interest, RNA-seq showed no impact of Rv3852 deletion on the bacterial transcriptome, thus ruling out a function as a universal repressor of horizontally acquired genes, as attributed to H-NS proteins (Dillon, 2010). Together, these results demonstrated that Rv3852 does not represent the real H-NS in Mtb, contrary to the initial annotation. Usually, gene and protein annotation relies on sequence similarity to known homologues. However, the case of our Rv3852 highlights the importance of genetic and functional studies to assign a precise role to a protein.

Our data are corroborated by the discovery that another protein, Lsr2, acts as H-NS in Mtb as it complements an *hns* knockout mutant in *E. coli* and, vice-versa, *E. coli* H-NS can compensate for the lack of Lsr2 in *M. smegmatis* (Gordon, 2008).

The mIHF regulon

ChIP-seq analysis showed that mIHF has binding sites located throughout the whole Mtb chromosome. Furthermore, RNA-seq revealed a large impact on gene regulation, and integration of both datasets proved that mIHF directly, as well as indirectly, regulates its target genes. This global transcription factor may affect expression of genes farther away than the tested 500 bp distance between the peak and the gene start. Atomic force microscopy showed that mIHF can change the conformation and topology of DNA by introduction of left-handed loops which open up supercoiled DNA. This structural change might bring other transcription factors closer to the target gene, or limit or promote access to RNA polymerase.

As most differentially expressed genes were downregulated, we concluded that mIHF mainly acts as a transcriptional activator. Strikingly, the *espACD* operon has an mIHF binding site in its upstream region and was additionally identified as the most downregulated transcript, directly connecting mIHF with the expression of major virulence factors. The coding region for the ESX-1 secretion system necessary for virulence is contacted more than once by mIHF, and several genes thereof were downregulated likewise. Gene ontology category analysis of the downregulated genes in the mIHF-depleted strain, i.e. activated by mIHF, showed that most were involved in host-pathogen interaction such as fatty-acid metabolism or nutrient acquisition-related pathways. Upregulated genes upon mIHF silencing, representing genes repressed by mIHF, mostly belonged to DNA repair pathways, showing that mIHF is required for normal DNA maintenance, most probably including replication. mIHF is therefore a NAP which acts as a specific transcription factor, despite the lack of a consensus motif.

In the case of mIHF too, we refined the annotation of the gene. The initial length of

mihF included a substantial 5'-end extension that is probably not part of the final protein. This may represent a 5'-UTR (untranslated region) with a regulatory role or may suggest that mIHF exists in two different isoforms. Mass spectrometry studies conducted in our and other laboratories support the presence of the shorter form of mIHF only.

ESX-1 regulation

The ESX-1 secretion system is controlled by a complex combination of key transcription factors. The NAPs EspR (Blasco, 2012) and Lsr2 (Gordon, 2010) were shown to bind to the extended promoter region of the *espACD* operon, a hot spot for regulatory proteins. The EspA, EspC and EspD proteins are essential for ESX-1 function and EsxA/B secretion. We proved that also mIHF associates with the *espACD* promoter region and positively regulates *espACD* expression. Furthermore,

mIHF binds upstream of *lsr2* and activates its transcription. Although Lsr2 represses the expression and antagonizes the activation of EspR, only a strong downregulation of *espACD*

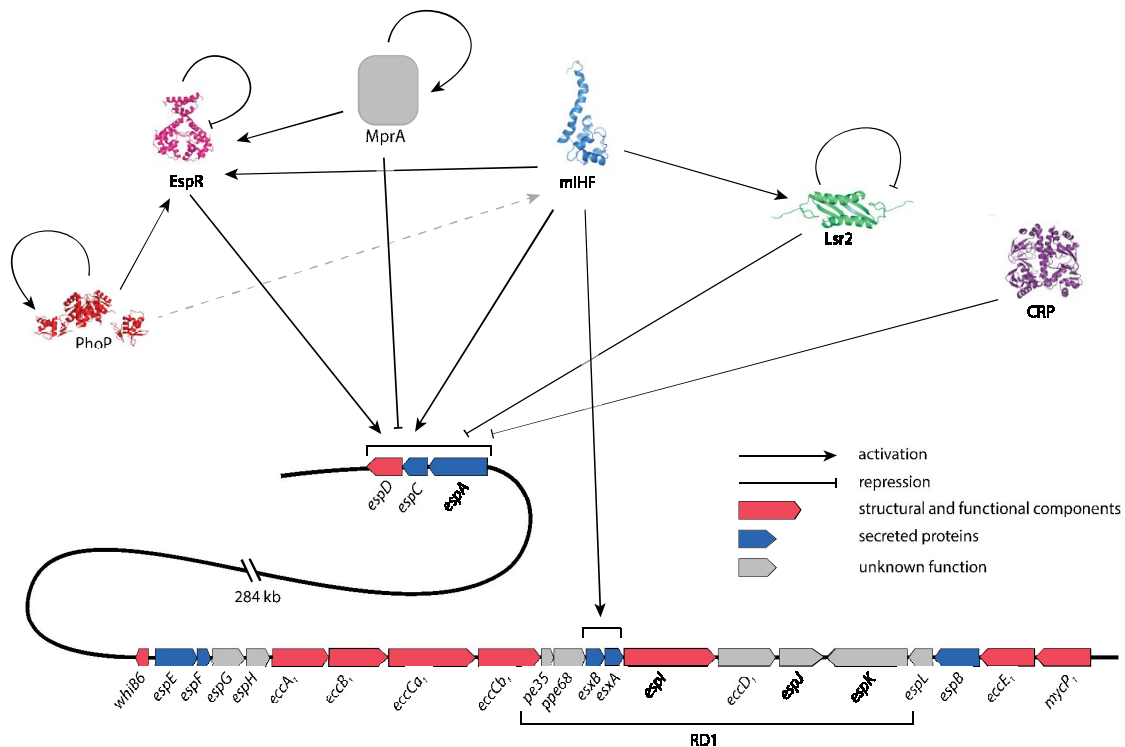


Figure 1: Regulatory network of the ESX-1 secretion system. The ESX-1 locus spanning *whiB6* to *mycP1* and the distantly located *espACD* operon is illustrated.

was observed. mIHF further binds at several loci inside the ESX-1 region, also close to *esxA/B* and activates their expression. PhoP controls EspR in a positive way, and binds upstream of *miHF*, although its regulatory impact on *miHF* expression is not clear (Solans, 2014). CRP and MprA both repress the *espACD* operon (Rickman, 2005; Pang, 2013). Figure 1 summarizes the current knowledge of ESX-1 regulation. Several projects in the laboratory of Prof. Cole aim to integrate the role of other transcription factors and investigate additional layers of regulation like post-translational modifications.

Potential interaction of mIHF with other transcription factors

Remarkably, a high degree of overlap between mIHF and EspR binding sites were observed in the ChIP-seq experiments. This suggests that the two proteins interact by binding to the same genetic loci, or act as antagonists. Several approaches can assess this hypothesis. A pulldown of protein complexes from Mtb by co-immunoprecipitation (co-IP) with subsequent mass spectrometry analysis can identify the interacting partners. To this purpose, I constructed mIHF-tagged strains, with an HA-tag at the N- or at the C-terminal of mIHF. Both strains are viable and express the tagged protein at high levels, as proved by immunoblotting (data not shown). Co-IP with anti-HA antibodies could be optimized following the protocol that was successfully used for EspC, which was shown to precipitate with EspA, in our laboratory (Lou, 2016). In addition to biochemical methods, an *in vivo* approach like the two-hybrid system in mycobacteria, known as mycobacterial protein fragment complementation (M-PFC) (Singh, 2006), allows to test protein-protein interaction. M-PFC was employed to screen for inhibitors of EspR homodimer formation (Blasco, 2014) and could easily be adapted to test putative mIHF-EspR interactions.

mIHF structure

Analysis of the mIHF protein structure revealed that it is a globular protein consisting of alpha helices only. The first helix, protruding from the core protein, is predicted to form

coiled coils, a site of dimerization by many proteins. The linker between the first helix and the rest of the protein is flexible to a certain degree as shown by NMR relaxation, which allows the protein to adopt small conformational changes. The NMR mIHF structure resembles closely the X-ray structure of sIHF (Swiercz, 2013), with a strong DNA-binding domain and potentially a second, weaker DNA-binding domain on the other side of the protein.

Based on the elution profile from size exclusion purification, we suspected that mIHF is present predominantly as a monomer. This was supported by the NMR relaxation data, which showed that the molecular rotational time of mIHF corresponded to a mostly monomeric form. Backbone assignment by NMR was particularly difficult though, as distinguishing between inter- and intramolecular interactions was problematic, suggesting that a small fraction of mIHF-dimers was present in solution. A shift in crosspeaks as well as broadening of signal was detected in NMR experiments upon DNA titration, indicating that mIHF dimerizes when bound to DNA. It is possible, that the mIHF protein weakly interacts by association and dissociation, thus representing an equilibrium state difficult to detect rather than a stable homodimer. Still, the strong indication for oligomerization of mIHF in presence of DNA is not absolute, as other coupled effects might contribute to crosspeak shifting. The aggregation state of mIHF has to be further explored.

Protein studies were additionally complicated by the fact that the small mIHF protein does not contain any tryptophan, tyrosine or cysteine residues that absorb at 280 nm, making it impossible to detect the protein by standard UV-absorption techniques. Every fraction of a protein purification experiment had to be stained with Coomassie blue to detect mIHF.

Overall, the achieved structure is of good quality, but the automated backbone assignment resulted in several ambiguously annotated residues. Especially for analysing the amino acids interacting with DNA, a complete annotation of peaks would have been ideal. We further planned to conduct ITC (isotherm titration calorimetry) studies to determine the

binding dynamics of mIHF, as it is not clear yet if mIHF has one or two distinct DNA-binding motifs. Different DNA substrates, e.g. with low and high GC content can be tested in different conditions, and the corresponding dissociation constant can be measured.

Phosphorylation can impact the DNA binding ability of transcription factors, as it was shown for HupB (Gupta, 2014). We expect a similar mechanism in mIHF, as *in vitro* phosphorylation of several residues by two kinases (PknB and PknG) was shown in our assays. DNA binding experiments like ITC can therefore be carried out to test what effect phosphorylation has on mIHF. Analysis of an Mtb total protein extract showed that the T83 residue was phosphorylated in mIHF (Prisic, 2010; Fortuin, 2015). Site-directed mutagenesis could be carried out on this amino acid to introduce T83E or T83V mutations, which would mimic either phosphorylated or non-phosphorylated threonine, respectively.

NAPs as drug targets

New drug targets are greatly desired in the TB drug research pipeline and the essentiality of most NAPs makes them an interesting target to explore.

Unfortunately, despite binding to Lsr2, zafirlukast did not have the expected effect against Mtb (Pinault, 2013), and chrysomycins were found to be toxic in eukaryotic cells. Better results were obtained with HupB, whose binding to DNA could be perturbed by small stilbene molecules (Bhowmick, 2014). These compounds showed good activity against Mtb H37Rv and only minor toxicity against eukaryotic cells (Suarez, 2017).

mIHF is essential for bacterial growth and survival, and depletion of the protein has a bactericidal effect. Therefore, mIHF would represent an appealing drug target, but the high abundance and stability of the protein may render this unrealistic. Additionally, we reported that multiple dilutions in fresh ATc-containing medium were needed before a phenotypic effect could be noticed, suggesting that large amounts or multiple administration of a possible drug would be required. Rather than targeting mIHF alone, the potential interaction between mIHF and EspR might be

considered for the design of specific inhibitors. With the availability of both structures, the proposed interaction could be modelled to rationally design intercalating molecules by a target-to-drug approach.

Three-dimensional chromosome conformation

ChIP-seq data reveals the binding profile of a transcription factor. Dimers of NAPs often bind to more than one DNA locus, but the 2D profile does not solve the spatial architecture of the genome. Bridging and looping of the DNA induce long-distance interactions between a gene and transcriptional regulators that are part of the regulatory circuits. Furthermore, functionally related genes often form an operon, are in close linear proximity, or are organized in clusters (e.g. the ESX loci coding for the T7SS in Mtb). In *Saccharomyces cerevisiae*, co-regulated genes were found to be in close spatial vicinity (Ben-Elazar, 2013), and also in *E. coli*, genes that are close in the 3D space were co-expressed and their products interacted (Xie, 2015). Similarly, a map of spatially interacting genetic loci helps to unravel the function of the many hypothetical proteins with unknown roles in Mtb.

Chromosome-conformation capture (3C) was the first technique to quantify the interaction of two specific DNA fragments, and more sophisticated methods were developed in recent years. ChIA-PET (chromatin interaction analysis with paired-end tag sequencing (Fullwood, 2009)) does not rely anymore on previously known interacting loci and can be used for *de novo* identification of genetic loci lying distantly on the 2D genome, but undergoing 3D interactions mediated by NAPs. EspR would be a good candidate for optimizing ChIA-PET in Mtb, as we showed high reproducibility for ChIP-seq experiments and EspR was shown to act as a homodimer (Blasco, 2012). Furthermore, Arg70 was identified as an important residue for dimerization (Blasco, 2014). The missense mutation of Arg70 could be introduced in Mtb and the resulting strain then used as a control in ChIA-PET experiments. Once the method has been optimized, ChIA-PET could be applied to mIHF and to other transcription factors.

Integration of the various datasets (RNA-seq, ChIP-seq and ChIA-PET) will provide a global view of the Mtb regulatory map in space

and time and will add tremendous new information to our relational database TubercuList (Lew, 2011).

REFERENCES

- Ben-Elazar, S., Yakhini, Z., and Yanai, I. (2013) Spatial localization of co-regulated genes exceeds genomic gene clustering in the *Saccharomyces cerevisiae* genome. *Nucleic Acids Res.* **41**: 2191–2201.
- Bhowmick, Bhowmick, T., Ghosh, S., Dixit, K., Ganesan, V., Ramagopal, U. a, et al. (2014) Targeting *Mycobacterium tuberculosis* nucleoid-associated protein HU with structure-based inhibitors. *Nat. Commun.* **5**: 4124.
- Blasco, B., Chen, J.M., Hartkoorn, R., Sala, C., Uplekar, S., Rougemont, J., et al. (2012) Virulence regulator EspR of *Mycobacterium tuberculosis* is a nucleoid-associated protein. *PLoS Pathog.* **8**: e1002621.
- Blasco, B., Japaridze, A., Stenta, M., Wicky, B.I.M., Dietler, G., Dal Peraro, M., et al. (2014) Functional dissection of intersubunit interactions in the EspR virulence regulator of *Mycobacterium tuberculosis*. *J. Bacteriol.* **196**: 1889–900.
- Cole, S.T., Eiglmeier, K., Parkhill, J., James, K.D., Thomson, N.R., Wheeler, P.R., et al. (2001) Massive gene decay in the leprosy bacillus. *Nature* **409**: 1007–1011.
- Dillon, S.C. and Dorman, C.J. (2010) Bacterial nucleoid-associated proteins, nucleoid structure and gene expression. *Nat. Rev. Microbiol.* **8**: 185–95.
- Fortuin, S., Tomazella, G.G., Nagaraj, N., Sampson, S.L., Gey van Pittius, N.C., Soares, N.C., et al. (2015) Phosphoproteomics analysis of a clinical *Mycobacterium tuberculosis* Beijing isolate: Expanding the mycobacterial phosphoproteome catalog. *Front. Microbiol.* **6**: 1–12.
- Fullwood, M.J., Wei, C.-L., Liu, E.T., and Ruan, Y. (2009) Next-generation DNA sequencing of paired-end tags (PET) for transcriptome and genome analyses. *Genome Res.* **19**: 521–32.
- Gordon, B.R.G., Imperial, R., Wang, L., Navarre, W.W., and Liu, J. (2008) Lsr2 of *Mycobacterium* represents a novel class of H-NS-like proteins. *J. Bacteriol.* **190**: 7052–9.
- Gordon, B.R.G., Li, Y., Wang, L., Sintsova, A., van Bakel, H., Tian, S., et al. (2010) Lsr2 is a nucleoid-associated protein that targets AT-rich sequences and virulence genes in *Mycobacterium tuberculosis*. *Proc. Natl. Acad. Sci. U. S. A.* **107**: 5154–9.
- Gupta, M., Sajid, A., Sharma, K., Ghosh, S., Arora, G., Singh, R., et al. (2014) HupB, a nucleoid-associated protein of *Mycobacterium tuberculosis*, is modified by serine/threonine protein kinases in vivo. *J. Bacteriol.* **196**: 2646–2657.
- Lew, J.M., Kapopoulou, A., Jones, L.M., and Cole, S.T. (2011) TubercuList - 10 years after. *Tuberculosis* **91**: 1–7.
- Lou, Y., Rybníček, J., Sala, C., and Cole, S.T. (2016) EspC forms a filamentous structure in the cell envelope of *Mycobacterium tuberculosis* and impacts ESX-1 secretion. *Mol. Microbiol.* **103**: 1–39.
- Pang, X., Samten, B., Cao, G., Wang, X., Tvinnereim, A.R., Chen, X.L., and Howard, S.T. (2013) MprAB regulates the espA operon in *Mycobacterium tuberculosis* and modulates ESX-1 function and host cytokine response. *J. Bacteriol.* **195**: 66–75.
- Pinault, L., Han, J.-S., Kang, C.-M., Franco, J., and Ronning, D.R. (2013) Zafirlukast inhibits complexation of Lsr2 with DNA and growth of *Mycobacterium tuberculosis*. *Antimicrob. Agents Chemother.* **57**: 2134–40.

Conclusions and Perspectives – 5

- Prisic, S., Dankwa, S., Schwartz, D., Chou, M.F., Locasale, J.W., Kang, C.-M., et al. (2010) Extensive phosphorylation with overlapping specificity by *Mycobacterium tuberculosis* serine/threonine protein kinases. *Proc. Natl. Acad. Sci.* **107**: 7521–7526.
- Raghavan, S., Manzanillo, P., Chan, K., Dovey, C., and Cox, J.S. (2008) Secreted transcription factor controls *Mycobacterium tuberculosis* virulence. *Nature* **454**: 717–721.
- Rickman, L., Scott, C., Hunt, D.M., Hutchinson, T., Menéndez, M.C., Whalan, R., et al. (2005) A member of the cAMP receptor protein family of transcription regulators in *Mycobacterium tuberculosis* is required for virulence in mice and controls transcription of the *rpfA* gene coding for a resuscitation promoting factor. *Mol. Microbiol.* **56**: 1274–1286.
- Singh, A., Mai, D., Kumar, A., and Steyn, A.J.C. (2006) Dissecting virulence pathways of *Mycobacterium tuberculosis* through protein-protein association. *Proc. Natl. Acad. Sci. U. S. A.* **103**: 11346–51.
- Solans, L., Gonzalo-Asensio, J., Sala, C., Benjak, A., Uplekar, S., Rougemont, J., et al. (2014) The PhoP-dependent ncRNA Mcr7 modulates the TAT secretion system in *Mycobacterium tuberculosis*. *PLoS Pathog.* **10**: e1004183.
- Suarez, M., Valencia, J., Cadena, C., Maiti, R., Datta, C., Puerto, G., et al. (2017) Diarylethenes Display In Vitro Anti-TB Activity and Are Efficient Hits Targeting the *Mycobacterium tuberculosis* HU Protein. *Molecules* **22**: 1245.
- Swiercz, J.P., Nanji, T., Gloyd, M., Guarné, A., and Elliot, M. a (2013) A novel nucleoid-associated protein specific to the actinobacteria. *Nucleic Acids Res.* **41**: 4171–84.
- Xie, T., Fu, L.-Y., Yang, Q.-Y., Xiong, H., Xu, H., Ma, B.-G., and Zhang, H.-Y. (2015) Spatial features for *Escherichia coli* genome organization. *BMC Genomics* **16**: 37.

CURRICULUM VITAE

Nina Theres Odermatt

Avenue de la Poste 27

1020 Renens

✉  nina.odermatt@outlook.com

📞 +41 79 769 36 89

 <https://www.linkedin.com/in/nina-odermatt>

Education & professional experience

PhD in Molecular Life Sciences

2013 - present

Supervised by Prof. Stewart T. Cole

École Polytechnique Fédérale de Lausanne (EPFL), Switzerland

PHD THESIS: “ Functional Characterization of Nucleoid Associated Proteins Acting as Global Transcription Factors in *Mycobacterium tuberculosis*”

- Design, conception and performance of experiments, evaluation and data analysis, including documentation and safety assessment
- Initiation and management of collaborations with professors, facilities and colleagues
- Teaching assistant in several undergraduate courses, participation at open doors events, organization of a PhD Winter School (the human gut microbiome in health and disease, a 1 week conference with 20 international speakers and 25 students)

Bachelor in Biology & Master in Microbiology, Minor in Chemistry

2007 – 2012

University of Zurich – Switzerland / Consiglio Nazionale delle Ricerche,

Verbania – Italy

MASTER’S THESIS: “*Crenarchaea* in a Deep Alpine Lake”

Supervised by Prof. Jakob Pernthaler & Dr. Gianluca Corno

- Deep-water sampling, chemical water analysis, phylogenetic analysis of prokaryotes, evaluation of microbial communities

Internship

2010

Smithsonian’s tropical research institute, Panama City – Panama

- scientific assistant in drug discovery based on Panamas endemic *Streptomyces* species

Selected conferences and awards

- 9th International Conference on the Pathogenesis of Mycobacterial Infections, poster presentation, 2014, Stockholm – Sweden
- Global Health Institute retreat 2016, best poster prize, Arolla – Switzerland
- Rigi Workshop 2016 (4 day course about translational biology), best poster prize, Rigi – Switzerland
- ASM Tuberculosis 2017, oral presentation, New York City – USA
- Life Sciences Switzerland (LS²) travel grant
- ASM Microbe 2017, oral presentation, New Orleans – USA

Skills

Molecular microbiology

- DNA and RNA extraction, handling and analysis
- molecular cloning, mutant generation
- bacterial and cell cultures
- BSL 3 laboratory work
- protein expression, purification and manipulation, structural analysis
- light microscopy
- ChIP-seq, RNA-seq preparation and analysis

Communication and teaching

- public speaking and presentation
- teaching experience in classes and supervision of junior students
- report and publication writing
- science outreach

IT

- Windows and Linux operation system
- data analysis, statistics and plotting with R
- NGS analysis
- Microsoft Office, Circos, Bowtie, Image J, PyMol, Photoshop, Illustrator

Languages

	Spoken	Written
German / Swiss German	Native	Native
English	C1 – C2	C1
French	B2	B1
Italian, Spanish	A1	A1

Extra-curricular activities

- Exploring nature; walking, hiking, canoeing, enjoying the environment
- Hang gliding, active member clubs, volunteering during events
- travelling to countries far and close, experience foreign cultures
- sharing a good meal with friends and family, reading books

Selected courses

- French; several integrated skills courses at EPFL, multi-lingual meetings and tandem language learning
- Online courses: Introduction to Project Management (AdelaideX on edX), Computer Networks (University of Washington on Coursera)
- Novartis international biotechnology leadership camp (BioCamp): 60 selected students participate in a 4-day seminar and discuss burning questions about science innovation with Novartis experts
- Workshop at ASM Microbe 2017: Science outreach for the microbiologist, ideas for children and adults

Personal details

Born in Switzerland, Swiss Citizen, age 30

- Car driver's license, category B
- motor boat license, category A on Swiss inland waters, for boats with > 6 kW motor power
- PADI Open Water Diver

ACKNOWLEDGEMENTS

I would like to express my greatest thanks to my supervisor Professor Stewart Cole for giving me the opportunity to work on this exciting project in his laboratory. During the past four years, I profited from his immense knowledge about *Mycobacterium tuberculosis* and his huge experience in the field of science. Stewart also allowed me to travel to several conferences, where I could improve my presentation skills and meet other great scientists with whom I discussed my projects. A very important person during my time as a PhD student was Dr. Claudia Sala, who taught me all the necessary methods, trained me in the biosafety laboratory 3 environment, revised the manuscripts I wrote and was always available for fruitful discussions and interpretations of the latest experiments. Not only her tips and tricks how to handle *M. tuberculosis* or numerous protocols, beautifully documented or even written in Italian, but also her encouragement and advice were a great help to finish my thesis.

I also would like to express my thanks to the fellow lab members, present and former, for their technical assistance and pleasant environment. It was always interesting to share our latest results, exchange knowledge about new methods in microbiology or share a glass of good wine. Further, I would like to acknowledge my former lab, the limnologists, who prepared me during my Master thesis for this PhD thesis, and the EPFL facilities who provided an organized and well-equipped place for work.

Ich möchte mich auch ganz herzlich bei meinen Eltern, Annegret und Urs Odermatt, für die Unterstützung, moralisch sowie finanziell, bedanken; dass ihr an mich geglaubt und mich ermutigt habt, nicht nur während meiner Doktorarbeit, sondern auch die Jahre zuvor. Ein herzliches Dankeschön an meinen Bruder Alexander und an Gotti Rita, für die gemeinsame gute Zeit und herzlichen Momente mit euch, und an meine Bärenfreunde vom Studium, sowie alle Bekannten und Verwandten.

Je voudrais aussi remercier tout le monde qui m'ont reçu avec les bras ouverts dans le sport de deltaplane et m'initié aux différents sites sur la montagne ; Luftarena, Deltalab, DCV, DCR, DCS, RNVL et les bras cassés. J'ai passé des moments inoubliables en l'air avec vous !

A big thank you to everyone who supported me during my journey through this PhD!

Lausanne, 21.09.2017



Nina Theres Odermatt

