Functional characterization of potential new drug target candidates in *Mycobacterium tuberculosis*

Inaugural dissertation

for the attainment of the title of doctor
in the Faculty of Mathematics and Natural Sciences
at the Heinrich Heine University Düsseldorf

presented by

Steffen Schindler
from Hamm

Düsseldorf, August 2024

from the Institute of Pharmaceutical Biology and Biotechnology
at the Heinrich Heine University Düsseldorf

Published by permission of the
Faculty of Mathematics and Natural Sciences at
Heinrich Heine University Düsseldorf

Supervisor: Prof. Dr. Rainer Kalscheuer

Co-supervisor: Prof. Dr. Karl-Erich Jaeger

Date of the oral examination

TABLE OF CONTENTS

ABBREVIATIONS

AC_1PIM_2	Acylphosphatidylinositol dimannoside
AC_2PIM_2	Diacylphosphatidylinositol dimannoside
AC_2PIM_4	Diacylphosphatidylinositol tetramannoside
AMP	Antimicrobial peptides
AMR	Antimicrobial resistance
APD	Antimicrobial peptide database
approx.	Approximately
CDC	Centers for Disease Control and Prevention
DC-SIGN	Dendritic cell-specific intercellular adhesion molecule-3-grabbing non-integrin
DNA	Deoxyribonucleic acid
DPA	Decaprenyl-monophosphoryl-D-arabinofuranose
DPM	Decaprenyl-monophosphoryl-D-mannopyranose
DPMS	Dolichyl phosphate mannose synthase
E. coli	*Escherichia Coli*
e.g.	*Exempli gratia*
Fig.	Figure
G1P	Glucose-1-phosphate
HIV	Human immunodeficiency virus
IFN-γ	Interferon-γ
IL	Interleukin
i.e.	*Id est*
LAM	Lipoarabinomannan
LM	Lipomannan
M. smegmatis	*Mycobacterium smegmatis*
M. tuberculosis	*Mycobacterium tuberculosis*
M1P	Maltose-1-phosphate
mAGP	Mycolyl-arabinogalactan-peptidoglycan complex
MDR	Multidrug-resistant
MGLP	6-O-methylglucosyl-containing lipopolysaccharides
MHC	Major histocompatibility complex
miRNA	MicroRNA
mRNA	Messenger RNA

MTBC	*Mycobacterium tuberculosis* complex
ncRNA	Non-coding RNA
nt	Nucleotide/s
OD	Optical density
PI	Phosphatidyl-myo-inositol
PIM	Phosphatidylinositol mannosides
PRR	Pattern recognition receptors
RNA	Ribonucleic acid
RR	Rifampicin-resistant
rRNA	Ribosomal RNA
SMAMP	Synthetic mimics of antimicrobial peptides
S. aureus	*Staphylococcus aureus*
S. typhimurium	*Salmonella typhimurium*
TB	Tuberculosis
TNF-α	Tumor necrosis factor α
tRNA	Transfer RNA
WHO	World Health Organization
XDR	Extensively drug-resistant

ACKNOWLEDGEMENT

First and foremost, I would like to thank Prof. Dr. Rainer Kalscheuer for giving me the opportunity to write my PhD thesis on a topic I genuinely enjoyed working on. The way you guided me through the years in your research group leaves me with nothing but gratitude. Your patience, knowledge, support, but also the trust you showed me made this part of my life a wonderful experience.

I would also like to thank Prof. Dr. Karl-Erich Jaeger for his commitment as my co-supervisor. In addition to preparing me for the tasks ahead by supervising my master's thesis, I am grateful that you have decided to continue to accompany my academic career.

I want to express my gratitude to my co-workers because it was a pleasure working with you. I would especially like to thank Heike Goldbach-Gecke for her support. I cannot imagine the working group without you, not only because of your lab management, but also because of the welcoming atmosphere you create. A big thank you to the whole crew: Lasse, Yvonne, Rosi, Rizwan, Dieter, Nidja, Viktor, Emmanuel, Tino, and Lin. I really enjoyed working with you guys, drinking coffee, chatting and goofing around. I am glad that not only scientific results were produced during this time, but also friendships were formed that I value immensely.

I am also grateful for the opportunity to participate in the MOI III Graduate School and would like to thank the entire team for scientific exchange that broadened my horizons for different research areas, and of course for letting me meet all of you.

Special thanks to Prof. Dr. Heiner Schaal and Björn Wefers for their management and help in the BSL3 lab, where I spent a lot of time.

Selbstverständlich möchte ich mich auch bei meiner Familie bedanken, ohne die diese Reise für mich nicht möglich gewesen wäre. Vielen Dank für eure ständige Unterstützung, nicht nur während meines Studiums, sondern mein ganzes Leben lang.

1. INTRODUCTION

1.1. Mycobacterium tuberculosis

In 1882, Robert Koch discovered *Mycobacterium tuberculosis*, the causative agent of tuberculosis (TB). The rod shaped, slow-growing bacterium belongs to the *Mycobacterium tuberculosis* complex (MTBC), which includes a variety of pathogens that cause infections in humans or animals. While *M. tuberculosis* primarily infects human hosts, other members of the MTBC such as *Mycobacterium pinnipedii*, *Mycobacterium microti*, and *Mycobacterium bovis* target animals such as seals, rodents, and cattle ([1–3]). However, infection of lung tissue is not limited to members of the MTBC, as non-tuberculous mycobacteria that do not belong to the MTBC group have also been associated with severe lung diseases, for example in the context of cystic fibrosis ([4]). Although TB is particularly associated with the respiratory tract, other organs such as intestine, liver, and bone marrow can also be severely affected ([5–7]). Given TB's status as a global health threat, the World Health Organization (WHO) tracks developments related to the disease, including its geographic spread, the emergence of resistant strains, and global efforts to combat it. The latter is embedded in a project called “End TB Strategy”, which aims to reduce incidence and deaths by 80% and 90%, respectively. Although there have been setbacks, tailoring milestones to specific countries is regarded as a promising approach to pave the way for achieving the goals by 2030 ([8]). Today, however, TB is detected around the world, with the highest incidence in Southeast Asia and Africa (Fig. 1). The high incidence of 7.5 million newly diagnosed TB cases in 2022 is exacerbated by the rapid evolution of drug resistance, which is independent of prevalence ([9]). Since *M. tuberculosis* is evolutionary adapted to humans, it developed mechanisms to survive in hosts without being eradicated and is therefore extremely difficult to treat. *M. tuberculosis* is capable of altering its gene expression to enter a dormant state, leaving approx. 1/3 of humanity latently infected ([10]). While the pathogen can reside in the host for years without leading to severe symptoms, reactivation of active metabolism can be triggered, leading to further spread of the disease. In 2019, the WHO tracked approx. 1.5 million TB-related deaths, making TB the leading cause of death from a single pathogen until the onset of the SARS-CoV-2 pandemic ([11]).

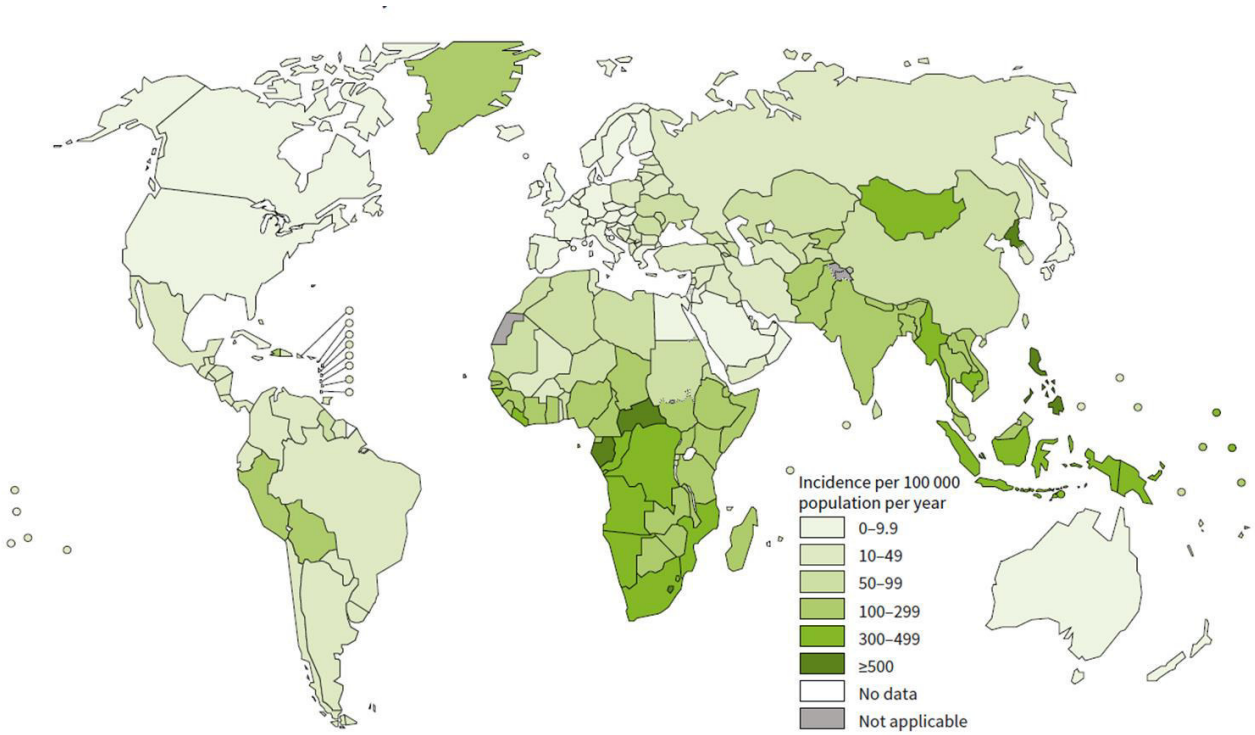


Fig. 1 Estimated TB incidence rates of 2022. Tuberculosis is a global threat to health, with the highest incidence rate in Africa and Southeast Asia, totaling 7.5 million infections in 2022 ([9]).

1.2. Infection

The high incidence rate of TB is the result of several ecological and economic factors. However, one complicating aspect is the aerosol transmission, which requires only a few bacteria to infect a new host ([12]). Upon inhalation, mycobacteria first encounter the non-professional airway epithelial cells, which recognize the pathogens *via* pattern recognition receptors (PRR) and alter the composition of the mucosa covering their surface to become more antimicrobial ([13]). The next line of defense is embodied by alveolar macrophages, which recognize the bacteria, leading to phagocytosis to eliminate the pathogens and thus terminate the infection ([14]). After internalization of a pathogen, the phagosome undergoes maturation changes and eventually fuses with the lysosome, leading to decomposition of contained microorganisms. Subsequent presentation of antigens by MHC class II molecules triggers the adaptive immune system to promote the response of the host. In contrast to other pathogens, *M. tuberculosis* has evolved a variety of virulence factors that control mechanisms within the macrophages ([15]). Utilizing the ESX-1 secretion machinery, *M. tuberculosis* gains access to the cytosol of the macrophages and employs virulence factors to evade the immune response by manipulating processes, leading to inhibition of autophagy and neutralization of reactive oxygen species or toxic metals ([16–18]). In addition, the fusion of the phagolysosome is prevented by labeling the phagosome as an early endosome, which is achieved by inhibiting the Rab

conversation ([19]). Instead of eventually killing the mycobacteria, macrophages become a unique habitat in which the bacteria can reside for decades without being affected by the human immune system. In the course of the immune response, various cytokines such as Interleukin-6 (IL-6), IL-1 and tumor necrosis factor α (TNF-α), as well as or other immunologically important compounds, particularly Interferon-γ (IFN-γ) are employed to manage the infection ([20–23]). Since IFN-γ plays a major role in this context, the link to severely affected HIV-patients becomes apparent. IFN-γ production is largely dependent on $CD4^+$ cells, a population that was observed to be decreased in HIV-positive patients, favoring TB progression ([24]).

The interplay between innate and adaptive immune system results in the detention of bacteria in macrophages and the attraction of additional immune cells that coalesce into a cluster of T-lymphocytes, neutrophils, and fibroblasts that surround the infected macrophage and form the hallmark of TB, the granuloma ([15]). The formation of granulomas is controversial because they are beneficial to the host but also shield bacteria and complicate the use of antibiotics ([25]). While preventing the spread of the pathogen, *M. tuberculosis* is able to adapt to this situation by altering its metabolism and entering a metabolically inactive, dormant state, resulting in a latent infection that can last for years. In the case of immunodepression, *M. tuberculosis* has evolved virulence factors that alter processes in the infected macrophages, causing them to no longer produce IL-6 and TNF-α but IL-10, an anti-inflammatory factor that promotes the onset of TB (summarized in Fig. 2) ([26]).

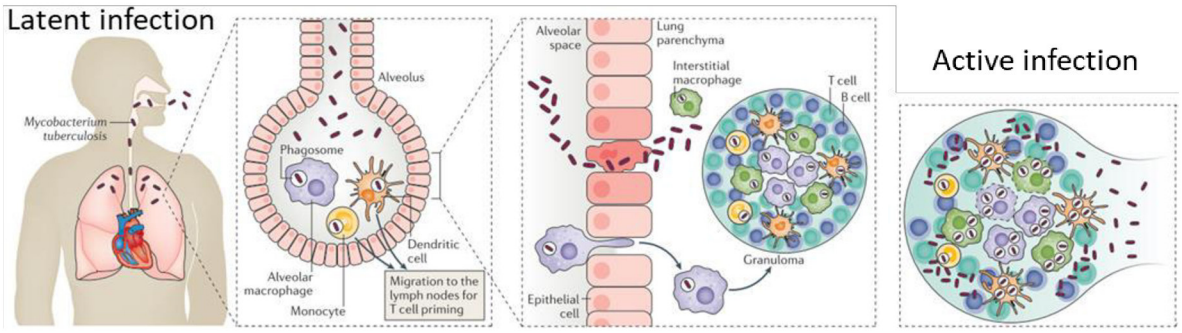


Fig. 2 Process of infection of humans by *M.* tuberculosis. Upon inhalation, mycobacteria encounter various immune cells and are eventually phagocytosed by macrophages. By modulating the host cell, *M. tuberculosis* prevents phagolysosome fusion and controls processes leading to granuloma formation. In the case of immunodepression, *M. tuberculosis* is able to alter the macrophage interleukin production to promote its escape from the cell aggregate. Modified from [27].

1.3. Therapy and resistance

According to the WHO, the standard regimen for TB is a combination therapy consisting of the four first-line antibiotics isoniazid, rifampicin, pyrazinamide, and ethambutol. Due to variations in infection, for instance, in granuloma composition, the duration of the treatment must be significantly prolonged to avoid relapse-events ([28]). Various long-term therapies have been tested and evaluated, leading to the guideline of a daily dose of all four drugs over a two-months period. Following this "Intensive phase", isoniazid and rifampicin must be taken daily in the "Continuation phase" for four months. Despite being two months shorter in duration, the 2/4-split regimen yielded fewer relapses and deaths compared to the eight-month regimen involving isoniazid and ethambutol in a six-month continuation phase ([29]). In addition to discussions about the duration of treatment, the antibiotics used have also been the subject of discussion. Since the 1960s, the continued use of the same antibiotics has led to the development of drug resistance. Of an estimated 10 million new TB cases, 4% showed rifampicin- or multidrug-resistance (RR/MDR), with MDR being defined as resistance to both rifampicin and isoniazid ([9]). The mode of resistance was elucidated for both compounds. Rifampicin, a semisynthetic product of the antimicrobial compound rifamycin B, acts on the RNA polymerase subunit β of microorganisms ([30]). Mutations in a specific section of the sequence of the target gene *rpoB* result in drug-resistance ([31]). In the case of isoniazid, resistance developed rapidly, as its discovery and also the first report of resistance date back to 1952 ([32]). Isoniazid is a pro-drug that is activated intracellularly by catalase peroxidase activity and subsequently interferes with mycobacterial cell wall formation by affecting mycolic acids ([33],[34]). The resistance mechanism against this compound is rather simple, as mutations in the KatG protein prevent the enzyme from activating the prodrug ([35]). Unfortunately, rifampicin and isoniazid are not the only first-line antibiotics for which resistance development has been reported. While the susceptibility to ethambutol, which has been shown to impair cell wall biosynthesis, is reduced by *embB* mutation, the lack of uptake of the pro-drug as well as pyrazinamidase activity results in pyrazinamide resistance ([36–39]). RR- and MDR-TB are treated differently because parts of the established regimen would not be beneficial for patients. Therefore, resistant mutants are treated with second-line drugs such as aminoglycosides (kanamycin or amikacin) or fluoroquinolones (levofloxacin or moxifloxacin). This approach, as well as procedures for treating extensively drug-resistant (XDR) strains that emerged as a result of the development of additional resistance to second-line antibiotics, is described in a 120-page guideline published by the WHO ([29]). This publication exemplifies the efforts of scientists to oppose the alarming trend of the antibiotic resistance development in times of stagnant drug development, also known as the antibiotic resistance crisis ([40]).

In order to underline that TB is a global threat, the WHO pointed out that the emergence of MDR and XDR strains is not geographically tied to high incidence locations (Fig. 3). While the incidence is particularly high in Africa, the highest proportions of drug-resistant strains have been found in Eastern Europe and Asia ([9]).

There are a number of factors that promote the development of drug resistance, which can be grouped into three main issues: misuse, overuse, and economic factors. Misuse or inappropriate use of antibiotics have demonstrated a link to the development of drug resistance, which becomes even more alarming when including findings of Klein, who discovered a 65% increase in antibiotic consumption between 2000 and 2015 ([41–43]). The overuse of antibiotics includes several aspects, such as inappropriate prescription, unrestricted access, and excessive use in agriculture. While several antibiotics are available over the counter in various countries or on the internet, 62% of antibiotics used in agriculture are considered medically important and therefore considered for human therapy ([41,44–47]). Economic factors indirectly favor the development of resistance, as, for instance, the financial barriers hamper the development of new drugs. The time-consuming research, development, and clinical trials make the production of new antibiotics less affordable or attractive to pharmaceutical companies. While antibiotics are used in a relatively short period of time yielding small profits, the duration of treatment of chronic diseases is significantly longer, guaranteeing comparatively higher revenues ([48]). For financial reasons, drug development for certain organisms is considered more profitable, neglecting research on other pathogens ([49]).

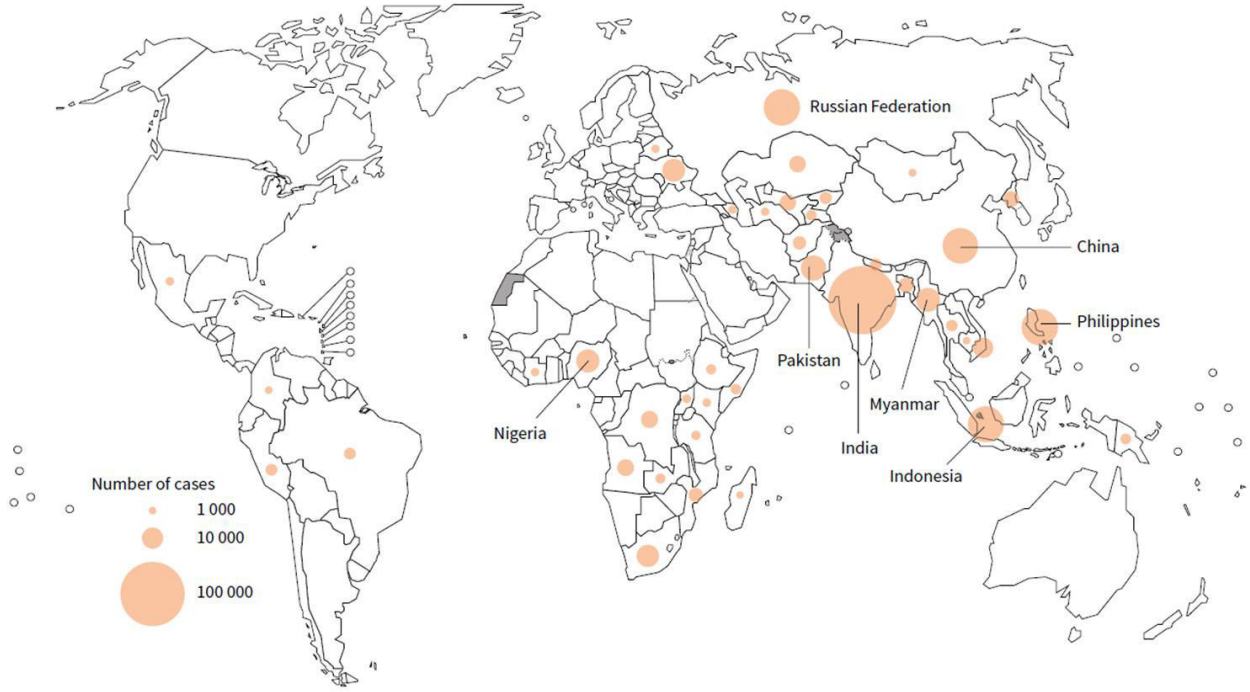


Fig. 3 Estimated number of new TB cases with multidrug-resistance (MDR) or rifampicin-resistance (RR) in 2022. The development of drug resistances is a threat around the globe that is independent from incidence rates ([9]).

1.4. Drug discovery – from Waksman to bioinformatics

Although this work is focusing on *M. tuberculosis*, it is noteworthy that the development of antimicrobial resistances (AMR) is a major threat posed by several pathogens. If the trend described in Chapter 1.3 continues, AMR are estimated to be the leading cause of death by 2050, surpassing the mortality rates of diseases such as cancer or diabetes (Fig. 4) ([50]). The urgent need for new antibiotics to counter this trend is evidenced by numerous studies calling for advances in antibiotic research.

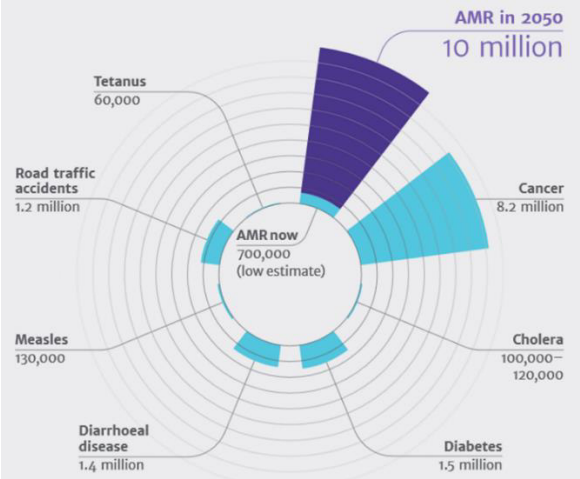


Fig. 4 Estimates of causes of death in 2050. Current antimicrobial trends predict up to 10 million deaths per year related to antimicrobial resistance, which would surpass today's leading causes of death, such as cancer or diabetes. Adapted from [50].

Historically, drug discovery has been driven by screening approaches, such as examining environmental samples for microbes that produce secondary metabolites inhibiting the growth of other microorganisms. The so-called Waksman Platform was invented by Selman Waksman and utilizes culture-based methods to evaluate the antibiotic potential of extracts derived from microorganisms ([51]). This strategy promoted the golden era of antibiotics, which lasted approx. from 1940 to 1960 and led to the discovery of several chemical classes of antibiotics ([52]). Although the next decades yielded some new antibiotic classes, such as quinolones, carbapenems, and lipopeptides, the development of innovative antibiotics began to stagnate ([53–56]). While this method has yielded a number of derivatives that also have great potential and could also circumvent resistance mechanisms, the discovery of new classes of antibiotics has been a rare event. The trend continued as no new antibiotic class has been discovered since 1987, resulting in the so-called discovery void ([57]).

Advances in digital technology have enabled more cost-effective and time-saving screening of compound libraries, limiting laboratory work to functional validation of compounds rather than identification of potential candidates. Using 3D modeling, virtual screening has identified

antibiotics that overcome resistance mechanisms by, for example, interfering with the resistance-mediating β-lactamases ([58]). Furthermore, the idea of drug repurposing became a vibrant area. For instance, halicin, originally intended to treat diabetes, was found to be a broad-spectrum antibiotic capable of killing bacteria such as *E. coli* or *M. tuberculosis* ([59]). Although *in silico* studies demand certain requirements and bioinformatics understanding, programs that are assessable and applicable to non-bioinformaticians have been developed to promote drug discovery ([60,61]). The bioinformatics approach improved drug discovery but did not lead to new mechanisms or new classes of antibiotics. However, promising advances were made that helped dealing with AMR issues.

In addition to screening compound libraries, bioinformatics facilitates other approaches to drug development, for example the rational design of antibiotics. The idea of rational design of drugs describes the development of molecules tailored to specific targets, which has been used in the past to treat cancer or fight TB ([62]). Bioinformatic features can be used to predict and optimize the physiochemical properties of molecules *in silico*, for example, to disrupt a target protein by impairing its function or its bimolecular interactions ([63,64]). Computational advances have revolutionized rational design in terms of efficacy, enabling less time-consuming and cost-intensive approaches while processing large amounts of data and reducing wet-lab work to validation of prediction rather than basic screening experiments. Although this description of bioinformatics resembles the holy grail of antibiotic research, modern prediction tools face many obstacles. The main problem is the complexity of cellular processes and the response to interventions since, for instance, the impairment of a particular function is complemented by the regulation of different elements or processes. The incomplete understanding of cellular properties and coherent networks complicates predictions and requires basic research to support bioinformatics approaches. One example of the current need for laboratory work is the antibiotic compound chlorflavonin. Although a chlorine atom has been identified experimentally to play a critical role in the antibiotic properties of this compound, *in silico* docking studies have failed to demonstrate interactions of the chlorine with the putative targets ([65]). In summary, computer-based technology offers great opportunities, but laboratory work remains essential for antibiotic development.

1.5. Drug targets

In general, the treatment of TB relies on targets that are regularly used in antibiotic treatment. In standard therapy, *M. tuberculosis* is affected at multiple targets simultaneously, e.g., DNA-dependent RNA polymerase, membrane, and cell wall, which are targeted by first-line antibiotics ([30,34,37]). Due to the emergence of drug resistance described above, therapy may depend on antibiotics targeting different mechanisms, such as DNA gyrase (fluoroquinolones), nucleic acid synthesis (sulfonamides), or protein biosynthesis (aminoglycosides) ([66–68]). While common targets like the polymerase are affected similarly to other microorganisms, the cell wall of *M. tuberculosis* possesses unique characteristics that are described in detail in this Chapter. Following, the relatively young topic of non-coding RNAs (ncRNAs) will be introduced and their potential as new drug targets will be addressed.

1.5.1. Cell wall

The cell wall of mycobacteria has a complex and unique structure that results in intrinsic resistance to a variety of antibiotics. Compounds acting on intracellular targets have to overcome a thick barrier consisting of four main layers to exert their function (Fig. 5) ([69]). To reach an intracellular location, compounds must penetrate the capsule (1), cross the mycomembrane (2), the mycolyl-arabinogalactan-peptidoglycan complex (mAGP) (3), and finally the plasma membrane (4), which is separated from the rest of the cell wall by the periplasmic space. While serving as a physical barrier against antagonistic compounds, the biosynthesis of these cell wall structures also represents an important drug target due to its crucial role in viability and cellular integrity. The following section provides insight into the different layers of the cell wall to demonstrate the potential for therapeutic intervention.

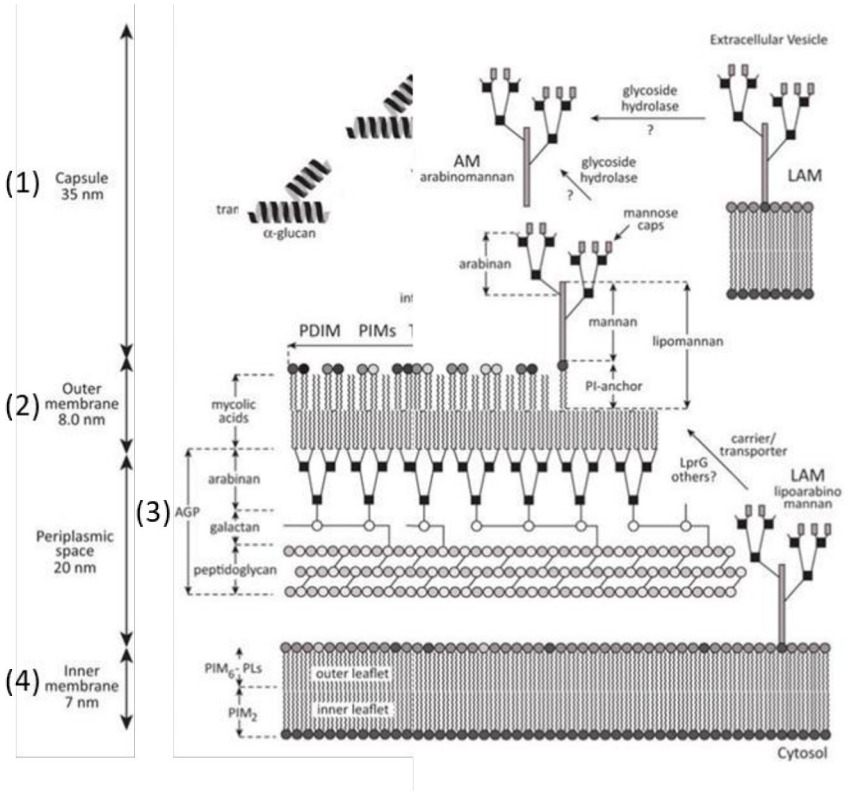


Fig. 5 The mycobacterial cell wall. The cell wall consists of four layers: the capsule (1), the outer membrane or mycomembrane (2), the mycolyl-arabinogalactan-peptidoglycan complex (mAGP) (3), and the inner or plasma membrane (4). The cell wall contains several immunoactive structures such as α-D-glucan, phosphatidylinositol mannosides (PIM), lipomannan (LM) and lipoarabinomannan (LAM). Adapted from [70].

(1) The outermost component of the envelope, the capsule, is largely composed of carbohydrates and proteins and is credited with functions such as maintaining water balance and modulating the immune response ([71–73]). The polysaccharides identified in *M. tuberculosis* are the homo- and heteropolysaccharides α-D-glucan, D-mannan and D-arabino-D-mannan ([74]). To briefly describe the formation of extracellularly localized α-glucan, there are two known pathways that provide maltose-1-phosphate (M1P) for the polymerizing maltosyltransferase activity of GlgE. This protein, analyzed in Chapter 6, forms linear α-glucan that is branched *via* GlgB and subsequently exported (Fig. 6). The process involving GlgC, GlgA, and GlgE utilizes glucose-1-phosphate (G1P) derived from glucose-6-phosphate (G6P) as substrate, whereas the pathway using trehalose relies on TreS, Pep2, and GlgE ([75]). In addition to these two metabolic pathways, Rv3032 was also found to play a role in glucan elongation, although its main function is associated with the production of 6-O-methylglucosyl-containing lipopolysaccharides (MGLP) ([76]).

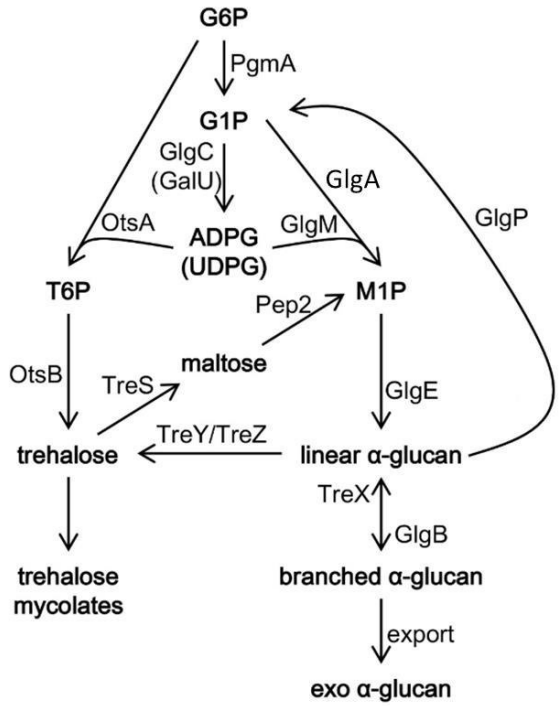


Fig. 6 The pathway of synthesizing α-glucan. There are two pathways by which α-glucan is produced starting from glucose-6-phosphate (G6P). Both pathways utilize maltose-1-phosphate (M1P), one derived from glucose-1-phosphate (G1P) and ADP- or UDP-glucose (ADPG, UDPG), the other from trehalose-6-phsphate (T6P) and trehalose. Using M1P as a substrate, α-glucan is produced by iterative action of GlgE and GlgB and finally exported. Modified from [77].

(2) The outer membrane is a lipid-rich layer composed mainly of mycolic acids covalently bound to the arabinogalactan layer ([78]). The mycolic acids of the outer or mycomembrane contain characteristic β-hydroxy-α-alkyl fatty acids and are classified as α-, methoxy-, or keto-mycolic acids based on certain structural features ([79]). The composition of mycolic acids in the cell wall varies depending on the species, which is why mycolic acids have been used for diagnostic purposes in the past ([80]). Although the ratio of α- to methoxy- and keto-mycolates varies within strains of the *M. tuberculosis* complex, α-mycolates generally remain the most abundant class ([81]). Mycolic acids contribute to the fluidity and permeability of the cell wall and have been linked to pathogen-host interactions, as an association with granuloma formation has been revealed ([82,83]). In addition, isoniazid, known as one of the most effective drugs against *M. tuberculosis*, specifically interferes with the biosynthesis of this cell wall component, which emphasizes the importance of investigating the metabolic pathways and genes involved ([34]). The biosynthesis of mycolic acids involves a complex network of different enzymes that utilize, among others, the type I and II fatty acid synthase systems, which catalyze condensation, reduction, and dehydration reactions that eventually lead to molecules containing 60 to 90 carbon atoms (reviewed in [84]). As the last step of the intricate formation of the mycolic acid trehalose monomycolate, the mycolic β-ketoester has to be reduced before it is translocated to the mycomembrane. The reductase catalyzing this reaction has first been

identified in *Corynebacterium glutamicum* and was later proposed to be encoded by the gene Rv2509 in *M. tuberculosis* ([85,86]). To shed more light on this crucial step of mycolic acid biosynthesis, we have extended the analysis of Rv2509 as described in Chapter 7.

(3) The third layer is the so-called mAGP complex, which consists of arabinogalactan and peptidoglycan. As an important structural component that accounts for approx. 35% of the mycobacterial envelope and connects individual elements, it is essential for the integrity of the mycobacterial cell wall ([87]). The arabinogalactan layer can be divided into three parts: the linker, which connects the complex to the peptidoglycan, arabinose, and galactose. The linker unit binds an unbranched chain of alternating β1→5 and β1→6 galactose units to approx. 10% of the muramic acid residues of the peptidoglycan, tethering the structure to the interior of the cell wall ([88,89]). The galactan, which is covalently bound to the linker, serves as a scaffold for highly branched arabinose chains that connect the mAGP complex with the mycolic acids ([88,90]). Due to the peptidoglycan, this complex also contributes to the shape, rigidity and osmotic stability compensating the turgor pressure ([91]). Peptidoglycan is a polymer composed of linear glycans, more specifically N-acetylglucosamine and N-acetylmuramic acid residues linked by β1→4 bonds and cross-linked by short peptides. Although composed of the same core structure, mycobacterial peptidoglycan is modified to contain unique features, such as N-glycolylmuramic acid or amidated carboxyl groups of peptides ([92,93]). As a result of the uncommon modifications, mycobacteria show increased lysozyme resistance and decreased susceptibility to β-lactam antibiotics ([94]).

(4) The innermost layer is the plasma membrane, which is separated from the rest of the cell wall by the periplasm and consists of phospholipids commonly found in bacterial membranes, such as cardiolipin, phosphatidylglycerol, phosphatidylethanolamine, phosphatidyl-*myo*-inositol (PI), and phosphatidylinositol mannosides (PIM) ([95]). The membrane is described as asymmetric because PIMs, lipomannan (LM), and lipoarabinomannan (LAM) are present in the thicker-appearing outer leaflet ([96]). However, the inner leaflet of the membrane is thought to contain excessive diacylphosphatidylinositol dimannoside (AC_2PIM_2), a precursor of LM and LAM ([97]). The lipoglycans LM, LAM, and the terminally mannosylated form ManLAM are not only bound to the plasma membrane but are also found non-covalently bound in the outer membrane, although the exact localization of these structures remains to be elucidated.

It is generally accepted that the cell wall of *M. tuberculosis* not only provides a protective barrier, particularly against hydrophilic compounds, but also contributes significantly to infection and retention in macrophages. Glycoconjugates are of particular interest because they possess affinity for several host cell receptors such as the dendritic cell-specific intercellular adhesion molecule-3-grabbing non-integrin (DC-SIGN), complement receptors, and the mannose receptor, thus modulating the immune response ([98–104]). Early research

focused on the biosynthesis of PIM, which consists of a PI backbone that acts as a mannose acceptor, and demonstrated different mannosylation levels from mono- to hexamannosylated states ([105]). PIMs have been shown not only to be involved in the interplay between *M. tuberculosis* and host cells, but also to embody the precursor structure for immunoactive LM and LAM ([106,107]). It is hypothesized that the acetylated form of PIM_4 (AC_1/AC_2PIM_4) represents the branch point that determines further modulation toward LM or LAM, as well as the structure that is translocated to the periplasmic space by a yet unknown mechanism ([108]). In addition, the further processing of PIM_4, leading to assembly of higher order PIMs, LM and LAM, was shown to rely on decaprenyl-monophosphoryl-D-mannopyranose (DPM) donors and not on nucleotide-derived elements such as GDP-Man that is used for the assembly of lower mannosylated PIMs ([105,108,109]). Analogous observations have been made with respect to decaprenyl-monophosphoryl-D-arabinofuranose (DPA) in the biosynthesis of LAM, although the process by which DPA is generated and translocated is still discussed ([110,111]).

As mentioned above, cell wall-associated components play a crucial role in the interaction between pathogen and host. These components not only promote the pathogenicity of *M. tuberculosis*, but also represent valuable drug targets due to their essential role in pathogen survival. For example, penicillin, the best-known class of antibiotics, inhibits the formation of the peptidoglycan layer by disrupting cross-linking ([112]). Accordingly, TB is treated using antibiotics that target the cell wall, including isoniazid and ethambutol ([34,37]). Based on the comprehensible success of this therapeutic regimen, the biosynthesis of various mycobacterial cell wall structures, including DPA, DPM, peptidoglycan, mycolic acids, and α-glucan, is considered a valuable drug target due to their essential role in *M. tuberculosis* ([82,113–116]). Since the density of putative drug targets in the biosynthetic pathways of glycoconjugates is remarkably high, deciphering the functions of the proteins involved seems reasonable. In this regard, PimA, which initiates PIM biosynthesis by transferring the first mannose residue to the PI backbone, MptA, which is essential for the elongation of the mannose chain in LM and LAM biosynthesis, and EmbC, which is involved in arabinosylation of LAM, are considered as valuable targets for therapeutic treatment ([117–119]). In order to expand this list, we analyzed the mycobacterial Rv3277 in Chapter 5.

1.5.2. Non-coding RNAs

The field of ncRNAs is receiving increasing attention as transcripts that do not serve as blueprints for proteins offer great potential for therapeutics. In addition to the well-characterized tRNA and rRNA, both of which play important roles in translation, ncRNA molecules are commonly used to fine-tune gene expression in bacteria. While ncRNAs are regularly used in intracellular processes that modulate, for example, metabolism or growth behavior, the expression patterns of several ncRNAs that have been analyzed so far suggest that they are also involved in pathogenicity ([120–122]). The expression of ncRNAs in response to external factors such as osmotic stress, reactive oxygen species, acidity, or nutrient limitation results in changes in gene expression to rapidly adapt to the hostile environment to which a pathogen is exposed, e.g., upon invasion of a host ([123–127]). In the last decades, knowledge about ncRNAs has increased considerably, leading to the recognition of complex ncRNA-based regulatory networks that modulate gene expression at the post-transcriptional and post-translational levels, thereby influencing, for example, the pathogenicity of various pathogens such as *Chlamydia trachomatis*, *Listeria monocytogenes*, *Staphylococcus aureus*, *Salmonella typhimurium*, or *Vibrio cholerae* ([128–132]). The list of regulatory ncRNAs under investigation is constantly expanding as the importance of these molecules in pathogens and their great potential for therapeutic purposes have been recognized. In *M. tuberculosis*, research on ncRNAs is still limited. As the existence of ncRNAs in mycobacteria has been confirmed and transcriptomics has revealed particularly interesting expression patterns, research in this area commenced ([126,133]).

As mentioned above, *M. tuberculosis* exemplifies how the innate and adaptive immune responses can not only be suppressed or evaded but also exploited to establish unique niches for residence and/or propagation (see Chapter 1.2). To achieve this, *M. tuberculosis* reprograms its transcriptional machinery to respond to threats and manipulate macrophages ([134]). Since adaption is a process which requires immediate changes, it is plausible that not only proteins but also rapidly generated RNA molecules are used to fine-tune the response of pathogens to a changing environment. While riboswitches located mainly in the 5' UTR of the target mRNA control translation in response to specific stimuli such as temperature, an increasing variety of independent ncRNA molecules are considered attractive targets for therapeutic intervention ([135], [136]). The mode of action of ncRNAs is rather simple, as the association with the target triggers several post-transcriptional and post-translational changes that mainly lead to repression or promotion of gene expression (Fig. 7). While most ncRNAs characterized to date regulate target gene expression by hindering or enabling access to the ribosome binding site of mRNAs, the formation of a ncRNA-mRNA duplex can also affect the stability of the mRNA. In general, base-pairing relies on a certain binding affinity due to

complementary sequences. Cis-acting ncRNAs, located on the opposite strand of their target gene, possess total complementary properties to the target. This characteristic offers great potential to influence the translation of the respective mRNA. Trans-acting ncRNAs are located in intergenic regions and therefore possess imperfect binding properties to their respective target mRNAs. However, trans-acting ncRNAs exhibit a wide range of targets, even proteins. By blocking the active sites of enzymes or affecting enzymes in an allosteric manner, trans-acting ncRNAs can affect intracellular processes at the post-translational level. The most prominent example of ncRNA-protein interaction was analyzed in *E.coli*, which utilizes the 6S RNA to influence the RNA polymerase to prevent transcription in stationary phase ([122]). Analysis of the association of ncRNAs with proteins revealed that the activity of trans-acting ncRNAs largely depends on the chaperone Hfq, which facilitates the interaction between ncRNA and target mRNA. Hfq was discovered independently of ncRNAs in *E. coli* and was identified to play a role in the synthesis of bacteriophages ([137]). However, 20 years later, its major function in the complex gene regulation mechanisms was deciphered ([138]). It is particularly interesting that Hfq plays an essential role in the virulence of several pathogens, for instance, *Salmonella typhimurium* ([139]). The reason for the impaired virulence in absence of this chaperone can be explained by the affinity to approx. 730 target RNAs ([140]). This large number of interaction partners is made possible by a common enzyme binding site that requires an adenine- and uracil-rich binding motif presented by the RNA molecules ([141]). Although Hfq is highly abundant in *E. coli*, Hfq or homologous proteins are not found in all pathogens known to utilize trans-acting ncRNAs, for instance, *S. aureus* or *M. tuberculosis* ([142]). It is still debated whether the protein or its homologue has not yet been discovered, or if the ncRNAs might act independently of protein guidance. Observations on *S. aureus* Rsa-family ncRNAs revealed a conserved C-rich motive that enables Hfq-independent binding of targets, thus supporting the existence of chaperone-independent ncRNA-based mechanisms ([143]). Current knowledge indicates that Hfq or homologous proteins are also absent in *M. tuberculosis* ([126]). Although no generic binding motif has yet been discovered, the unusually high GC-content of *M. tuberculosis* of approx. 66% may facilitate ncRNA function in a chaperone-independent manner similar to *S. aureus* ([144]).

Mycobacteria possess ncRNAs that are conserved from close relatives to distant mycobacteria, distributed throughout the genome, and expressed in a regulated manner, for instance, in a specific growth phase ([145]). Initial research has led to the functional characterization of ncRNAs such as F6, B55, or G2 that exhibit regulated expression at different growth phases and/or under adverse conditions, for instance, acidic or oxidative stress, to which the pathogen is exposed during infection ([146]). The use of ncRNA to regulate rapid adaptation to stress or to influence the host immune response is also known for other

pathogens, such as *S. typhimurium*, which utilizes RybB and MicA to fine-tune the expression of outer membrane proteins to alter its membrane composition ([147]). Other examples of the growing spectrum of cis- and trans-acting ncRNAs that confer virulent properties to pathogens include Sprd in *S. aureus*, SsrS in *Legionella pneumophilia*, IsrJ in *S. typhimurium*, Qrr1-4 and Vrra in *V. cholerae,* and MTS1338 in *Mycobacterium smegmatis* ([130–132,148–150]). The latter example is special as this finding is based on heterologous expression. MTS1338 is an ncRNA of *M. tuberculosis* that does not occur naturally in *M. smegmatis* but leads to increased survival rates in infection experiments upon heterologous expression ([150]). Since the function of MTS1338 in *M. tuberculosis* is not yet clear, we analyzed this ncRNA together with MTS2823, as both molecules show similar expression patterns and accumulate in stationary phase ([126,133]). Chapter 4 shows our efforts to uncover the function of these highly abundant ncRNAs, with MTS2823 being the most abundant transcript in stationary phase cells. Studying the function of ncRNAs whose roles are not yet understood could uncover direct targets for drugs or vulnerable processes they are involved in. This research might pave the way for the rational design of innovative antitubercular drugs.

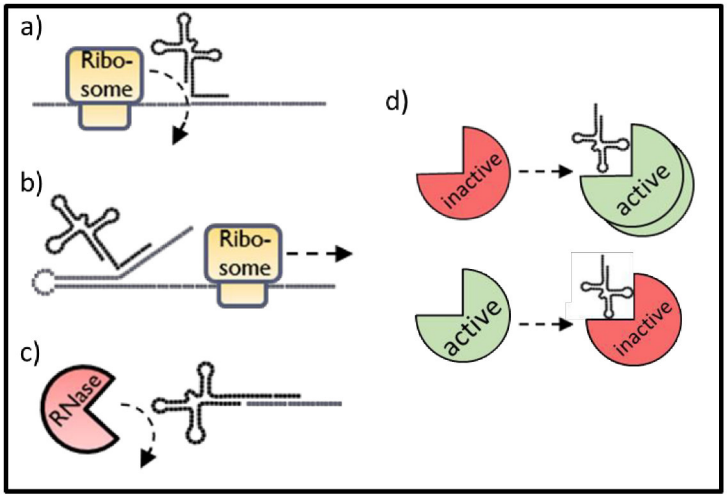


Fig. 7 Mode of action of ncRNAs. Binding to the target mRNA could affect translation by either a) blocking the ribosome binding site, thereby hindering translation, or b) enabling translation by making the ribosome binding site accessible to ribosomes. c) Stability of mRNAs may also be enhanced by ncRNAs, as RNase could be prevented from targeting the single-stranded mRNA. d) Association with enzymes could result in allosteric effects that alter the activity state of enzymes or prevent the conversion of substrates by blocking the active site of the enzyme.

2. AIM OF THIS STUDY

The objective of this study is to characterize potential drug targets in *M. tuberculosis*. Alarming trends in the emergence of drug resistances combined with the antibiotic discovery void underline the need for the identification or development of new drugs.

As with many microorganisms, the cell wall of *M. tuberculosis* represents a valuable drug target. Therefore, it is comprehensible that the standard treatment regimen for TB includes antibiotics that target components of the mycobacterial envelope. Considering that it is useful to expand the list of putative cell wall-associated drug targets to counteract possible resistance developments, we analyze proteins involved in the biosynthesis of essential cell wall components of *M. tuberculosis* and assess their potential for therapeutic advances. Moreover, we aim to complete the understanding of cell wall-associated biosynthetic processes to enable rational design of therapeutics that interfere with essential processes.

In a different approach to fight TB, we are focusing on a relatively young issue in TB research: untranslated transcripts. Over the past decade, the therapeutic potential of ncRNAs has been demonstrated in various pathogens, which has increased interest in ncRNA-associated research. As knowledge of ncRNAs involved in the regulatory machinery of *M. tuberculosis* is limited, our advances could pave the way for innovative treatments for TB.

In this study, we will use genomics, transcriptomics, proteomics, bioinformatics and techniques such as functional screening assays to analyze the individual candidates. By generating mutants of *M. tuberculosis* or *M. smegmatis* that either lack or insufficiently express target genes, we aim to decipher the function of several transcripts or proteins to evaluate their potential for therapeutic approaches.

3. SUMMARY

Although *M. tuberculosis* was replaced as the deadliest pathogen during the SARS-CoV-2 pandemic, TB remains a global threat. Millions of people are infected and killed by TB each year, and the emergence of antibiotic resistances exacerbates the current situation. Unlike assumed, antibiotic resistance is not limited to high incidence areas, highlighting its importance on a global scale. The drivers for resistance development are diverse as a number of factors ranging from misuse to economics are involved.

During the so-called golden era of antibiotics (1940 - 1960), antibiotic discovery using hit identification-based antimicrobial screening methods was successful. After this period rich in antibiotic discoveries, humanity faces the challenge of not being outcompeted by the evolution of antibiotic resistance, as the outcome is likely to be catastrophic in terms of mortality from multidrug-resistant pathogens by 2050. To counteract this trend, scientists are striving to find and develop new antibiotics and/or innovative methods to combat pathogens.

The aim of this work was to characterize new drug targets in *M. tuberculosis* to enable, for instance, the rational design of antibiotics. Rational design uses information about a target protein to engineer structure-specific molecules that interfere with the protein's function. This target-to-drug approach supports drug development based on screening approaches that analyze the effects of molecules on pathogens. Due to our efforts to generate knowledge about different genes involved in cell wall assembly, we have been able to expand the spectrum of antibiotic target candidates to include Rv3277, Rv2509 and *glgE*.

In addition to analyzing components associated with the cell wall, we also investigated ncRNAs as potential drug targets. Since the world of ncRNAs is a relatively young and underrepresented area in *M. tuberculosis* research, advances could have a great impact on drug development. We functionally characterized the two highly abundant ncRNAs MTS1338 and MTS2823 in *M. tuberculosis* and uncovered crucial regulatory functions, as both ncRNAs have been shown to be involved in the repression of gene expression during transition to and/or in stationary phase. The involvement in the regulation of essential mechanisms such as ribosomal silencing that, for instance, enables latent infections, renders these ncRNAs valuable drug targets.

4. Redundant functions of the non-coding RNAs MTS1338 and MTS2823 in *Mycobacterium tuberculosis*

Unpublished, manuscript in preparation

Overall contribution to the paper:

- Writing of the complete manuscript draft
- Generation of the *M. tuberculosis* ncRNA knockout mutants H37Rv_ΔMTS1338 and H37Rv_ΔMTS2823
- Isolation of total RNA, evaluation of transcriptome sequencing data
- Extraction of proteins, interpretation of proteome profiling data
- Analysis of expression patterns
- Evaluation of overlapping functions of MTS1338 and MTS2823
- Phenotypic characterization experiments (growth characteristics, temperature and pH sensitivity, infection of THP-1 derived macrophages, susceptibility to antibiotics)
- Detection of ncRNAs in the stationary phase culture supernatant

Redundant functions of the non-coding RNAs MTS1338 and MTS2823 in *Mycobacterium tuberculosis*

Steffen Schindler[1], David Podlesainski[2], Markus Kaiser[2], Thomas R. Ioerger[3], and Rainer Kalscheuer[1*]

[1] Institute of Pharmaceutical Biology and Biotechnology, Heinrich Heine University Düsseldorf, Germany

[2] Center of Medical Biotechnology and Hospital Hygiene, University Duisburg-Essen, Germany

[3] Department of Computer Science, Texas A&M University, College Station, Texas, USA

*Correspondence: Rainer Kalscheuer (Rainer.Kalscheuer@hhu.de)

ABSTRACT

Adaptation is a vital process for any organism, including pathogens. The moment a pathogen infects a host, it needs to a) adapt to the changed environment and b) respond to the immune defense. The key to survival under such adverse conditions is rapidity. Regarding the regulation of gene expression, non-coding RNAs in particular have gained increasing attention in the last decades. In several examples, it has been shown that RNA-based fine-tuning of gene expression drastically affects the virulence of pathogens. By enabling a rapid response to changes in, for example, temperature, pH, nutrient availability, and defense mechanisms, the probability of survival and thus the ability to infect the host is increased. Although the regulatory effects of non-coding RNAs are primarily studied in *Escherichia coli* and we are just beginning to understand the potential of gene expression regulated by non-coding RNAs in *Mycobacterium tuberculosis*, the relevance of these molecules has been suggested in several studies. The discovery of the highly abundant non-coding RNAs MTS1338 and MTS2823, which accumulate during the stationary growth phase in *M. tuberculosis*, provided the basis for the functional characterization. Based on gene-specific knockout stains that we subjected to functional screening and transcriptomic and proteomic approaches, we concluded that both non-coding RNAs have a redundant function in modulating gene expression by affecting ribosomal proteins during transition to and/or in stationary phase. Although the mechanism remains to be elucidated, the association with an essential process in the pathogen's metabolism highlights the great potential of MTS1338 and MTS2823 as potential targets for antitubercular therapy.

KEYWORDS

Mycobacterium tuberculosis, tuberculosis, non-coding RNA, drug target, infection

INTRODUCTION

Until 2019, before emergence of SARS-CoV-2, *Mycobacterium tuberculosis* was considered the world's deadliest pathogen. Although tuberculosis is now being overshadowed, it remains a global threat to human life, infecting approx. 10.6 million people and killing 1.3 million in 2022 ([1]). The effort to fight TB is impeded by the mycobacteria's ability to evade the effect of, for instance, cell wall-targeting antibiotics such as isoniazid or ethambutol, antibiotics inhibiting RNA-polymerase such as rifampicin and rifabutin, as well as other classes of antibiotics ([2,3]). All around the world, the emergence of multi and extensively drug-resistant strains of *M. tuberculosis* underlines the urgent need for new antibiotics.

In recent years, increasing knowledge about non-coding RNAs (ncRNAs) has led to the finding that RNAs are more than mere blueprints for proteins. The class of ncRNAs comprises several subclassifications that are involved in the regulation of cellular processes. Besides obvious functions of ncRNAs, such as rRNA or tRNA in translation, bacteria employ small ncRNAs in numerous processes, particularly in response to environmental changes ([4]). It is known that ncRNAs are involved in diverse processes such as catabolism of different nutrition sources, replication, secretion, modulation of the cell surface, or more general processes like metabolism, stress response, and quorum sensing ([5–14]). Particularly in pathogenic organisms, ncRNAs attracted much attention since they can confer virulence traits and/or are important for the progression of the infection in, for instance, *Pseudomonas aeruginosa*, *Staphylococcus aureus*, *Streptococcus pneumoniae*, and *Listeria monocytogenes* ([15–18]).

The general mechanism of ncRNA-based regulation relies on base-pairing with target mRNAs or on association with proteins, resulting in activation or disruption of gene expression ([4], [19]). While riboswitches regulate the expression of the particular mRNA they are part of, cis- and trans-acting small ncRNAs conduct their function as independent RNA molecules. Cis-acting ncRNAs are known to regulate mRNAs derived from the opposite DNA strain and therefore possess total complementary properties with the specific target structure. Intergenic, trans-acting ncRNAs demonstrate a broader range of targets and elicit activity despite imperfect binding to target RNAs. In this context, the protein Hfq plays an essential role, as ncRNAs largely rely on association with this protein to conduct their function ([20]). Hfq was originally discovered in 1968, but its involvement in the establishment of ncRNA-mRNA interactions was revealed only decades later ([21,22]). While this chaperone or homologous proteins are still unknown in some organisms, Hfq demonstrates essentiality for virulence traits in pathogens such as *Salmonella typhimurium* ([23]).

In 2011, Arnvig discovered ncRNAs in *M. tuberculosis* and subsequently proposed their involvement in pathogenesis ([24,25]). Studies in other bacteria revealed that certain ncRNAs are differentially expressed and/or highly abundant according to the growth phase of the cell ([26,27]).

Insight into the function and/or vulnerable processes involving ncRNAs could pave the way for the rational design of innovative antitubercular drugs ([28]). In this study, we focused on functional characterization of two small ncRNAs that we and others found to be highly abundant in exponential and/or stationary phase cells ([24,29]). The mere abundance of these small ncRNA species suggests that they might be of importance for *M. tuberculosis* during certain growth phases.

MATERIAL AND METHODS

Bacterial strains and growth conditions

The strains used in this study are listed in Tab. 1. Mycobacteria were grown in aerobically in liquid cultures (Middlebrook 7H9 broth, BD Diagnostics) or on agar plates (Middlebrook 7H10 agar, Difco) at 37°C, 5% CO2 and 80 rpm (in case of liquid cultures), supplemented with 10% ADS (0.81% NaCl; 5% BSA; 2% dextrose), 0.5% glycerol and 0.05% tyloxapol. If required, antibiotics for selection (50 µg/ml hygromycin, 20 µg/ml kanamycin, 20 µg/ml apramycin) were added.

Escherichia coli was incubated in lysogeny broth (LB) medium or on LB agar supplemented with antibiotics for selection (100 µg/ml ampicillin, 150 µg/ml hygromycin, 40 µg/ml kanamycin, 20 µg/ml apramycin) at 37°C, 5% CO2 and 80 rpm (in case of liquid cultures).

Tab. 1 Strains of microorganisms used in this study. Hyg^r, hygromycin resistant; Suc^s, sucrose sensitive.

Strain	Relevant characteristics	Source or reference
M. tuberculosis		
H37Rv		W.R. Jacobs Jr., AECOM, Bronx, NY
H37Rv_ΔMTS1338	ΔMTS1338::γδres-sacB-hyg-γδres, Knockout mutant, Hyg^r, Suc^s	This study
H37Rv_ΔMTS2823	ΔMTS2823::γδres-sacB-hyg-γδres, Knockout mutant, Hyg^r, Suc^s	This study
M. smegmatis		
mc^2155		Obtained from William R. Jacobs Jr., PhD, Albert Einstein College of Medicine, Bronx, USA
E. coli		
NEB® 5-alpha F'lq	*fhuA2 a(argF-lacZ)U169 phoA glnV44 a80a(lacZ)M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17*	New England Biolabs

Transformation of *E. coli*

Commercially available competent *E. coli* NEB5-alpha cells (New England Biolabs) were transformed according to the manufacturer's protocol. The cells were thawed on ice and mixed with 5 µl of ligation mixture or plasmid DNA and incubated on ice for 30 minutes. A heat shock was then conducted at 42°C for 30 seconds. After cooling, the cells were incubated in 500 µl of SOC-medium at 37°C. Finally, the transformed cells were spread on selective LB agar and incubated overnight at 37°C. Emerging colonies were screened for desired phenotypes.

Transformation of mycobacteria

Mycobacteria were cultivated to an $OD_{600\ nm}$ of 0.5 - 0.8, washed twice in a cold 10% (v/v) glycerol + 0.05% (v/v) tyloxapol solution and finally resuspended in 1/10 of the original culture volume. Electroporation was performed using cuvettes with 0.2 cm gap at 2500 V, 1000 Ω, and 25 µF. The transformed cells were incubated with 1 ml Middlebrook 7H9 medium and incubated at 37°C for 1-2 hours (*M. smegmatis*) or 18-20 hours (*M. tuberculosis*) and finally plated on Middlebrook 7H10 agar containing selective antibiotics.

Generation of site-specific gene deletion mutants

In order to generate site-specific mutants by allelic exchange, we used the method of specialized transduction employing temperature-sensitive mycobacteriophages as described previously ([30]). Briefly, specific allelic exchange substrates were generated by PCR and introduced into a vector containing a synthetic resistance cassette (p0004S). The sequence of allelic exchange substrates flanking the synthetic resistance cassette was introduced into a phasmid (phAE159), with which cells of *M. smegmatis* were electroporated. Subsequently, *M. smegmatis* was used to generate a high-titre phage lysate containing the phages with the desired allelic exchange sequence at the permissive temperature of 30°C. By exposing cells of *M. tuberculosis* to the phage lysate at the non-permissive temperature of 37°C, allelic exchange took place. Screening for the desired phenotypes was performed after plating the mycobacteria on Middlebrook 7H10 plates containing 50 µg/ml hygromycin for selection. The oligonucleotide primers used for generating the allelic exchange substrates as well as the resulting plasmids and phasmids are listed in Tab. 2 and 3, respectively.

Tab. 2 Oligonucleotides used in this study

Oligonucleotide	Application	Sequence 5'- 3'
SS_1338_L_fw	Synthesis of the 5' Fragment used for allelic exchange of MTS1338	GCCCAAAAATTGGGCTGGATTCAACGGCCATGC
SS_1338_L_rev	Synthesis of the 5' Fragment used for allelic exchange of MTS1338	CGCCAATTCTTGGCACCAATACTGCCACCAAAC
SS_1338_R_fw	Synthesis of the 3' Fragment used for allelic exchange of MTS1338	GCCCAAAGATTGGGATCCTCATCCTGTTACTGC
SS_1338_R_rev	Synthesis of the 3' Fragment used for allelic exchange of MTS1338	CGCCAACTTTTGGCGTTGTGGAGCGTCGTGTTC
SS_2823_L_fw	Synthesis of the 5' Fragment used for allelic exchange of MTS2823	CGGCAAAAACTGCCATTGCCAAGTCCAGCACAC
SS_2823_L_rev	Synthesis of the 5' Fragment used for allelic exchange of MTS2823	GCGCAATTCTTGCACGAAGCGGCTTTGGATTGC
SS_2823_R_fw	Synthesis of the 3' Fragment used for allelic exchange of MTS2823	GCCCAAAGATTGGTAGCCGCTCGGCTATCTAGG
SS_2823_R_rev	Synthesis of the 3' Fragment used for allelic exchange of MTS2823	CGCCAACTTTTGGTCGATGACGATCGGCTTGGG

Tab. 3 Plasmids and phasmids used in this study

Plasmid	**Application**	**Source or reference**
p0004S	Construction of knockout plasmids	
pΔ1338	Construction of ΔMTS1388 knockout phasmid	This study
pΔ2823	Construction of ΔMTS1388 knockout phasmid	This study
Phasmid	**Application**	**Source or reference**
phAE159	Construction of knockout plasmids	
phΔ1338	Generation of phages for allelic exchange of MTS1338	This study
phΔ2823	Generation of phages for allelic exchange of MTS2823	This study

Extraction of genomic DNA

Mycobacteria were grown to late log phase, pelleted by centrifugation, and resuspended in 450 µl of GTE solution (50 mM glucose; 25 mM Tris-Cl, pH 8.0; 10 mM EDTA) supplemented with 50 µl of a 10 mg/ml lysozyme solution. After overnight incubation at 37°C, 150 µl of a 2:1 solution of 10% SDS and 10 mg/ml proteinase K was added. This mixture was incubated at 55°C for 30 minutes before adding 200 µl of 5 M NaCl and 160 µl of CTAB solution (4.1 g NaCl; 10 g hexadecyltrimethylammonium bromide in 90 ml ddH_2O). After incubation for 10 minutes at 65 °C, one volume of 24:1 (v/v) chloroform : isoamyl alcohol was added. After vortexing and centrifugation, the aqueous layer was transferred for further processing. The procedure was repeated, 0.7 volumes of isopropanol were added, and the solution was centrifuged. The precipitated DNA was washed with 70% ethanol and resuspended in ddH_2O.

RNA extraction

Mycobacterial cultures were grown either to the exponential phase ($OD_{600\ nm}$ of 0.5 - 0.8) or to the late log phase (cultivation for 28 days). The cultures were pelleted and resuspended in 5 ml RNA protect bacteria reagent (QIAGEN) and incubated overnight at RT. The QIAGEN RNeasy Kit (QIAGEN) was used for isolation of total RNA. The procedure was conducted following the manufacturer's protocol.

Transcriptome profiling

Cells of the ncRNA deprived *M. tuberculosis* mutants were grown from frozen stocks in 7H9 medium containing until they reached an OD_{600} of 0.5 - 0.8. Cells were harvested for RNA extraction by centrifugation and fixed overnight in 5 ml RNA Protect reagent (Qiagen). Fixed cells were centrifuged, resuspended in 1 ml RLT buffer (Qiagen) and lysed by bead beating to prepare lysates.Total RNA was extracted using the RNeasy Mini kit (Qiagen). Quality control of total RNA preparations were performed with Agilent 2100 Bioanalyzer and Fragment Analyzer, respectively. RNA was further analyzed by photometric Nanodrop measurement. Barcoded cDNA libraries were prepared according to the manufacturers protocol (Ion Total RNA-Seq Kit v2; Life Technologies). Emulsion PCR and subsequent IonProton sequencing were performed according to commercial kit protocols (Ion PI Template OT2 200 Kit v2, Ion PI Sequencing 200 Kit; Life Technologies). Demultiplexing was done using TorrentSuite software (vers. 4.0.2, Life Technologies). Raw sequencing reads were quality trimmed in CLC Genomics Workbench (vers. 6.5.2, CLCbio / Qiagen). After removal of short (< 50 nt) sequences, the remaining high-quality reads were aligned against the *M. tuberculosis* H37Rv reference sequence NC_000962.3. RPKM normalized read counts were log2 transformed for

further analysis. Differential gene expression between two experimental conditions (three biological replicates each) was statistically determined by Student´s T-test (FDR corrected). The significance threshold was set to p(corr) = 0.01.

Protein extraction

Mycobacterial cultures were grown in 20 ml Middlebrook 7H9 medium until an $OD_{600\ nm}$ of 0.5 - 0.8 was detected. The cultures were washed three times with cold PBS and finally resuspended in 2 ml of PBS. Cell lysis was achieved by bead beating using silica-zirconium beads at 50 Hz for 3 x 3 min. After addition of 200 µl of 10% SDS solution, the mixture was incubated at 4°C for 30 min. While after centrifugation the cell debris was collected together with the beads at the bottom of the vials, the protein-containing supernatant was filtered three times with 0.2 µm polyethersulfone filters and used for further applications.

LCMS/MS

LC–MS/MS assays were performed utilizing an Orbitrap Elite instrument linked to an EASY-nLC 1000 LC system (Thermo Scientific) ([31]). The LC was adjusted to one-column mode. The analytical column was a fused silica capillary (75 µm × 45 cm) with an integrated PicoFrit emitter (15 µm, New Objective), packed with Reprosil-Pur 120 C18-AQ 1.9 µm resin (Dr. Maisch). The analytical column was enclosed by a column oven (Sonation) and connected to a nanospray flex ion source (Thermo Scientifitc). The temperature of the column oven was set to 45 °C during data acquisition. The LC was operated with two mobile phases: solvent A (0.1% FA in water) and solvent B (0.1% FA in ACN), both in UPLC grade purity (Sigma-Aldrich). The peptides were loaded onto the analytical column at maximum flow rate, not reaching the set pressure limit of 980 bar (usually 0.5 - 0.6 µl/min). The peptides were then fractionated on the analytical column with a gradient of solvent A and solvent B at a flow rate of 300 nL/min for 140 minutes (gradient: start with 7% solvent B; gradient 7% to 35% solvent B for 120 min; gradient 35–100% solvent B for 10 min and 100% solvent B for 10 min). The mass spectrometer was operated with the Xcalibur software (Thermo Fischer Scientific, version 2.2 SP1.48) and set to the positive ion mode. Precursor ion scanning was performed in the Orbitrap analyzer (fourier transform mass spectrometry (FTMS)) in the scan range of m/z 300–1800 and with a resolution of 60000 with the internal lock mass option turned on (lock mass was 445.120025 m/z, polysiloxane) ([32]). The spectra obtained were recorded in data-dependent form in the ion trap (ion trap mobility spectrometry (ITMS)) in a variable scan range and with a rapid scan rate. The ionization potential (spray voltage) was set to 1.8 kV. Peptides were analyzed using a repetitive cycle that included a full precursor ion scan (1.0×10^6 ions or

50 ms) followed by 15 product ion scans (1.0×10^4 ions or 100 ms). During this process, peptides were segregated based on their intensity in the full survey scan (threshold of 500 counts) to generate tandem mass spectra (MS2), enabling peptide sequencing and identification. For MS2 spectra generation, the collision-induced dissociation (CID) energy was adjusted to 35%. For MS2 data acquisition, dynamic ion exclusion was set to 120 s with a repeat count of one and a maximum excluded ion list of 500 members. The settings enabled included monoisotopic precursor selection, FTMS (Orbitrap) preview mode, charge state screening, and ion injection time prediction. Charge states higher than one were considered for fragmentation. Subsequently, the RAW spectra were submitted to an Andromeda search in MaxQuant (version 1.5.3.30) with default settings ([33,34]). Label-free quantification was activated ([35]). MS/MS spectra data were screened against the Uniprot *Mycobacterium tuberculosis* (strain ATCC 25618 / H37Rv) reference proteome database (UP000001584; 3,993 entries). All searches included a contaminant database (as implemented in MaxQuant, 245 sequences). The contaminant database includes known MS contaminants and was added to assess the level of contamination. Andromeda searches allowed static modification of cysteine (57 Da, alkylation with iodoacetamide) as well as oxidation of methionine residues (16 Da) and acetylation of the N-terminus of the protein (42 Da) as dynamic modification. Enzyme specificity was set to “Trypsin/P”. In the Andromeda search, the instrument type was configured to Orbitrap, with the precursor mass tolerance set to ±20 ppm for the initial search and ±4.5 ppm for the main search. The MS/MS match tolerance was set to ±0.5 Da. Protein FDR and peptide spectrum match FDR were adjusted to 0.01 and a minimum peptide length of seven amino acids was required. Unique and razor peptides as well as peptides with dynamic modifications were required to have a minimum score of 40 for protein quantification. Perseus v1.5.5.3 ([36]) was used for additional sorting and evaluation of the MaxQuant output. MS/MS data were loaded into the matrix from the proteinGroups.txt file, and reverse hits and putative contaminants were excluded. The technical replicates of the samples were categorized into groups and filtered to enable statistical calculations. Protein groups with three valid values in at least one category were used in the data evaluation. For the remaining protein groups, missing values were ascribed, and a t-test was performed (number of randomizations 250; initial FDR 0.05 and S0 0.1).

Minimal inhibitory concentration (MIC) determination

A two-fold serial dilution of antibiotics was prepared in a 96-well polystyrene plate (U-bottom, Greiner). Mycobacteria were seeded at a density of $1x10^5$ cells in a total volume of 100 µl per well. After incubation at 37°C for five days, 10 µl resazurin solution (100 µg/ml) were added to each well. Following resuspension of the solution, the plates were incubated for 24 hours at RT. Finally, the mycobacteria were inactivated with 10% formalin (100 µl per well), and

fluorescence (excitation 540 nm; emission 590 nm) was quantified using a Tecan infinite F200 Pro reader (Tecan). Relative growth was calculated according to the non-treated growth control.

Phenotypic characterization assays (temperature capacity, pH tolerance)

A defined number of mycobacterial cells (10^5 cells/well) were incubated in 96-well polystyrene plates (U-bottom, Greiner) and subjected to growth conditions differing in temperature or pH-value. The relative growth of mycobacteria was determined using resazurin (see MIC determination) or, for pH-associated experiments, using the BacTiter-Glo™ Microbial Cell Viability Assay (Promega) according to the manufacturer's instructions.

Reactivation of metabolism

In order to evaluate the ability to switch between dormancy and active metabolism, cells were incubated in Middlebrook 7H9 medium until reaching an $OD_{600\ nm}$ of 1. The cultures were washed three times in phosphate-buffered saline (PBS) and further incubated as described above. After 3 months, 10^6 cells were incubated in 10 ml Middlebrook 7H9 medium to assess metabolic reactivation by monitoring the growth by measuring the $OD_{600\ nm}$ of the cultures.

Intracellular survival of mycobacteria employing THP-1 derived macrophages

A defined number of THP-1 cells (10^5 cells/well) were grown in RPMI1649 containing 10% fetal calf serum (FCS) at 37 °C and 5% CO2 in a 96 well flat-bottom microtiter plate. Differentiation was induced by adding 50 nM phorbol 12-myristate 13-acetate (PMA) and incubating the cells overnight. Before infection of the THP-1 derived macrophages with mycobacteria grown to an OD_{600} of 0.6 - 0.8, the cells were introduced to fresh medium. A ratio of 3:1 (molecules of infection = 3) was used for infection. After incubation for 3 h, the cells were washed twice with PBS before fresh RPMI1649 (supplemented with 10% FCS) was added. At selected time points (immediately and after 5 days), infected macrophages were lysed with ddH_2O. Samples of the solutions were diluted in Middlebrook 7H9 medium and plated on Middlebrook 7H10 agar. The resulting colonies were analyzed after incubation at 37 °C for 14 days.

RESULTS

The ncRNAs MTS1338 and MTS2823 are highly abundant in stationary phase cells

Using whole transcriptome sequencing in a previous study with a different context, as a side aspect, we have identified strictly regulated small ncRNAs that are highly abundant in exponential and/or stationary phase cells ([29]). To contextualize the amount of transcripts, we compared ncRNA levels of MTS1338 and MTS2823 to sigma factor F (*sigF*), which is known to be induced in stationary phase cells ([37]). The average trimmed reads per kilobase million (RPKM) of MTS1338 and MTS2823 exceeded the expression of *sigF* by 1,304- and 4,023-fold, respectively (Tab. 4). Since the expression of MTS1338 and MTS2823 correlates with cell maturation, the hypothesis of possible regulatory functions in the stationary phase becomes evident.

Tab. 4 The ncRNAs MTS1338 and MTS2823 belong to the most abundant transcripts in stationary phase cells of *M. tuberculosis*. Whole transcriptome sequencing of *M. tuberculosis* reveals the high abundance of the ncRNAs in the stationary phase. Abbreviations: nt, nucleotides; RPKM, reads per kilobase million. Data extracted from ([29]).

	Rv Identity	Description	Size (nt)	Average trimmed RPKM
Most abundant transcripts	Rvnc0036a	MTS2823	300	856824.6
	Rvnc0036	MTS1338	117	277857.0
	Rv3286c	*sigF*	786	213.4

The ncRNA-knockout mutants H37Rv_ΔMTS1338 and H37Rv_ΔMTS2823 are morphologically unobtrusive

In order to gain insight into the function of the ncRNA MTS1338 and MTS2823, previous studies by other scientists focused on the overexpression of ncRNAs in the exponential phase ([24,38–40]). Although overexpression of MTS1338, which is rarely expressed in the exponential growth phase, was a practical advance that led to interesting results, we wanted to develop an approach for functional characterization that could be applied to both ncRNAs, MTS1338 and MTS2823 ([25]). MTS2823 is one of the most abundant ncRNAs in *M. tuberculosis* even in exponential phase cells. Therefore, overexpression could miss effects, as already saturated processes would not be affected by additional expression of ncRNA ([24]). To generate knockout mutants, we utilized specialized phage transduction in which the respective ncRNA

was substituted by a synthetic cassette containing selection markers, yielding the mutants H37Rv_ΔMTS1338 and H37Rv_ΔMTS2823. The mutants growing on selective medium have been verified by diagnostic PCR (Fig. 1).

When grown on agar plates or in liquid medium, the mutants did not exhibit a specific phenotype but showed wild type-like colony formation and culture density without characteristic features, such as excessive clumping.

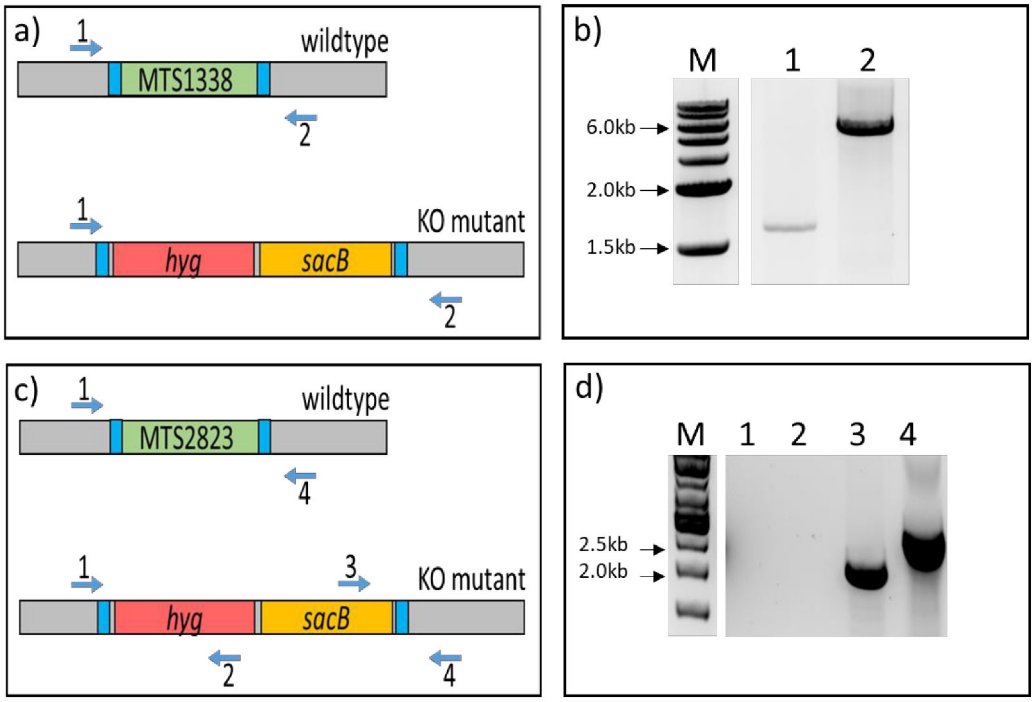


Fig. 1 Diagnostic PCR to detect allelic exchange of MTS1338 and MTS2823. a) Illustration of the allelic exchange of MTS1338. The blue area marks the flanking regions used for the allelic exchange. The upper figure represents the wild type situation, while the lower figure shows the knockout. Primers used for diagnostic PCR are indicated with arrows and numbers. **b**) Agarose gel showing the PCR products generated using primers 1 + 2 and wild type DNA as template (lane 1) or H37Rv_ΔMTS1338 DNA as template (lane 2). **c**) Illustration of the allelic exchange of MTS2823, analogous to a). **d**) Agarose gel showing the PCR products generated by using primers 1 + 2 and wild type DNA as template (lane 1) or H37Rv_ΔMTS2823 DNA as template (lane 3) as well as using primers 3 + 4 and wild type DNA as template (lane 2) or H37Rv_ΔMTS2823 DNA as template (lane 4). The depictions in a) and c) do not demonstrate actual ratios.

The absence of MTS1338 or MTS2823 reveals multiple changes in the transcriptome and suggests shared or overlapping functions

The regulated expression of the ncRNAs MTS1338 and MTS2823 indicates a potential role in the stationary phase. However, we investigated possible functions in both the exponential and the stationary phase. Whole transcriptome sequencing yielded interesting data as H37Rv_ΔMTS1338 and H37Rv_ΔMTS2823 demonstrated a variety of similar changes compared to the wild type, for example, the number of dysregulated transcripts, the ratio of up- to downregulation, and the number of affected processes.

In the exponential growth phase, H37Rv_ΔMTS1338 showed 160 differentially expressed genes, of which 59% were upregulated and 41% were downregulated. Similar differences were found when H37Rv_ΔMTS2823 was compared to the wild type, as 63% of 147 differentially expressed genes were upregulated and 37% were downregulated (Fig. 2a). In addition to the similarity in the number of regulated transcripts, we detected an overlap of 44 genes regulated by both mutants. This trend is underlined by the observation that six of the ten most up- and downregulated transcripts of H37Rv_ΔMTS1338 are also found among the most highly regulated transcripts in H37Rv_ΔMTS2823 (Fig. 2c). The transcripts of the antitoxin gene *vapB46* (Rv3385c), the gene encoding probable membrane protein Rv1004c, the glycerophosphodiester phosphodiesterase gene *glpQ1* (Rv3842), the gene encoding putative primosomal protein N' PriA (Rv1402), the gene *moeA1* (Rv0994) associated with molybdopterin biosynthesis, and Rv2257c encoding a conserved protein of unknown function are downregulated in both mutants. The observation that there was no focus on any particular process was also noted concerning the upregulated transcript as we found the transcript of the dehydratase gene *hisB* (Rv1601), the gene encoding PPE family protein PPE59 (Rv3429), Rv1871c encoding a conserved protein of unknown function, the putative acyl-CoA reductase gene Rv1543, the fatty acid-AMP ligase gene *fadD21* (Rv1185c), and the electron transport component gene *nrdH* (Rv3053) to be similarly expressed in both mutants. Overall, the similarity of the altered gene expression profile in both ncRNA mutants in the exponential phase is striking.

Comparable observations were made in the stationary phase, where both mutants showed the same transcriptome pattern as 56% of differentially expressed genes were upregulated and 44% were downregulated (Fig. 2b). As expected, the total number of differentially expressed genes was higher in the stationary phase (257 and 220 genes, respectively). We detected an overlap in gene regulation as more than 50% (127) of differentially expressed genes were dysregulated in both mutant strains. Remarkably, the regulation (positive/negative) was concordant in 98% (125 genes). As an exception to this trend, the uncharacterized genes Rv1812c and Rv2626c, which codes for the hypoxic response protein Hrp1, were slightly

downregulated in H37Rv_ΔMTS1338 and upregulated in H37Rv_ΔMTS2823. The ten most dysregulated transcripts in the stationary phase did not follow the trend of the exponential phase because in the case of downregulation we detected only the transcripts of Rv1167c, a transcriptional regulatory protein, and Rv1722, possibly encoding a carboxylase involved in the lipid mechanism, showing congruence between the mutants (Fig. 2d). Regarding the ten most upregulated genes, we observed similar regulation only in a transcript associated with amino acid biosynthesis. The high abundance of Rv1658, which encodes the argininosuccinate synthase ArgG, prompted us to investigate transcripts associated with amino acid synthesis. Apparently, arginine biosynthesis was severely affected in both mutants, as we detected upregulation of other genes belonging to the same gene cluster. Notably, in H37Rv_ΔMTS2823, the transcripts of four genes of this gene cluster were found among the most upregulated transcripts (*argB*, *argC*, *argF*, and *argG*).

Besides strikingly similar results, we discovered a high number of putative targets for MTS1338 and MTS2823 that were dysregulated in the absence of the respective ncRNA. In particular, low transcript levels of the downstream genes Rv1734c and Rv1735c were observed in the stationary phase of H37Rv_ΔMTS1338, suggesting a regulatory role for the expression of these genes. Although the investigation of the single targets is a practical advance to study the function of the ncRNAs, we subsequently focused on deciphering the reason for the vast amount of concordant changes.

At the more global level, i.e., looking for the affectedness of processes rather than individual transcripts, we recognized a remarkable similarity of regulated processes in both mutants, as the expression of genes associated with cell wall and cell processes, conserved hypotheticals, and intermediary metabolism and respiration was strongly affected (Fig. 2e-f). These results strongly suggest that MTS1338 and MTS2823 could possess similar and partially overlapping functions. This potential partial functional redundancy might explain the lack of an apparent phenotype in the background of the single gene deletion. Further supporting this hypothesis, we observed that deletion of MTS2823 resulted in a slight but significant upregulation in MTS1338 expression possibly representing a compensating response (Fig. 3). In contrast, MTS1338-deficiency did not significantly affect MTS2823 expression, suggesting unilateral complementation.

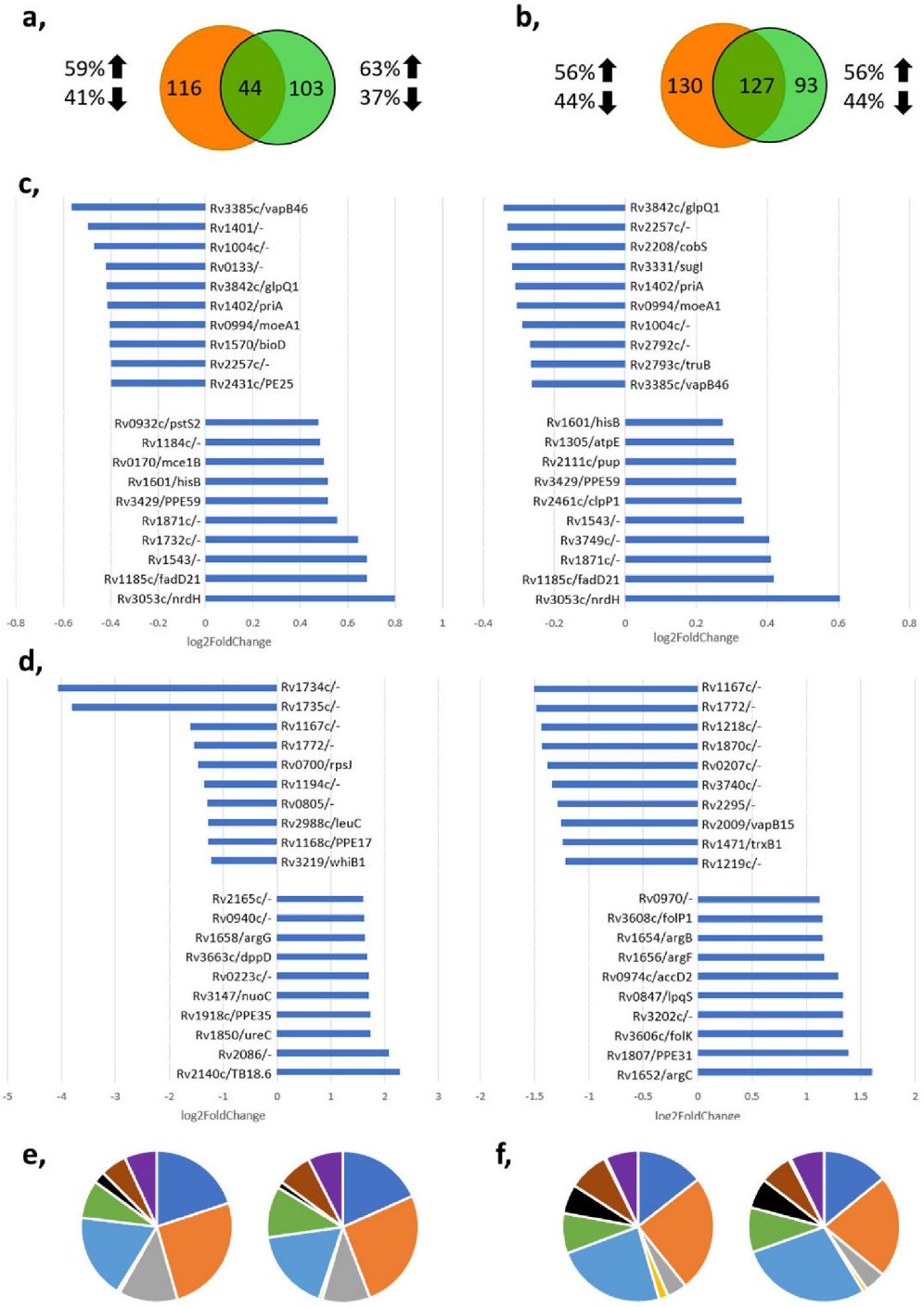


Fig. 2 Whole transcriptome sequencing reveals a similar pattern of dysregulated genes in H37Rv_ΔMTS1338 and H37Rv_ΔMTS2823. Comparative analysis of the transcriptome of H37Rv_ΔMTS1338 (left) and H37Rv_ΔMTS2823 (right) in terms of total dysregulated transcripts in exponential (**a**) and stationary phase (**b**), the ten most dysregulated transcripts in exponential (**c**) and stationary phase (**d**), and the relative distribution of regulated transcripts to different functional categories in exponential (**e**) and stationary phase (**f**). Assignments of regulated genes to functional categories was done according to Mycobrowser (https://mycobrowser.epfl.ch/): Dark blue, cell wall and cell processes; orange, conserved hypotheticals; grey, information pathways; yellow, insertion sequences and phages; light blue, intermediary metabolism and respiration; green, lipid metabolism; black, PE/PPE proteins; brown, regulatory proteins; purple, virulence, detoxification, adaptation.

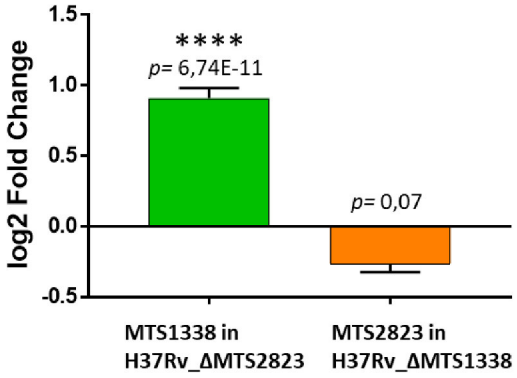


Fig. 3 MTS1338 and MTS2823 possess unilateral complementation properties. Comparison of the exponential phase expression of MTS1338 in H37Rv_ΔMTS2823 (green) and MTS2823 in H37Rv_ΔMTS1338 (orange) based on transcriptome sequencing data. Statistical significance was determined using the student's t-test. Error bars show the standard error of the mean (n = 3).

Proteome profiling reveals the potential involvement of MTS1338 and MTS2823 in ribosomal silencing

During whole transcriptome sequencing, we discovered potential homologous functions of MTS1338 and MTS2823. Accordingly, we subjected the knockout mutants H37Rv_ΔMTS1338 and H37Rv_ΔMTS2823 to mass spectrometry to understand whether the results obtained are also transferable to the proteome level.

The deprivation of a non-essential ncRNA led to widespread changes in the composition of the proteome, comparable to changes detected by transcriptome sequencing. In H37Rv_ΔMTS1338 we observed a change in the abundance of 296 proteins in the exponential phase and 748 proteins in the stationary phase (Fig. 4a-b). Comparison of the proteome of H37Rv_ΔMTS2823 with that of the wild type yielded similar results, as a total of 277 and 860 proteins belonging to diverse processes were differentially expressed in the exponential and stationary phase, respectively (Fig. 4c-d). In addition to similar quantities of dysregulated proteins, the mutants share the most dysregulated proteins in the exponential phase, as the eight most upregulated and five of the ten most downregulated proteins of H37Rv_ΔMTS1338 are also found among the most regulated proteins in H37Rv_ΔMTS2823 (Fig. 4a, c). Besides hypothetical proteins thought to be involved in transport across the membrane or having an unknown function (Rv3054c, Rv1924c, Rv0879c, Rv0849, Rv1754c), both mutants showed upregulation of proteins related to virulence. MymT, a metallothionein known to protect the cell from copper toxicity, is the most upregulated protein in both mutants and exhibited a 6-fold change compared with the wild type. Also, MalQ, a 4-alpha-glucanotransferase, and CysK2, the protein that catalyzes the synthesis of the non-proteinogenic amino acid S-sulfocysteine, were highly upregulated in both mutants. MymT and CysK2 are both associated with host-

pathogen interactions, as copper is used by the host as an intracellular defense mechanism against cells of *M. tuberculosis* residing inside macrophages, while S-sulfocysteine acts as a signaling molecule that triggers redox defense reactions in response to reactive oxygen species ([41], [42]). Among the common most downregulated proteins, we detected CrtB, which is involved in carotenoid biosynthesis of mycobacteria, CydB and CydD, both involved in energy metabolism, and proteins belonging to the PPE family (PPE18, PPE60) ([43–45]).

Concerning the proteome of stationary phase mutants, we discovered corresponding patterns, as four of the ten most upregulated and three of the ten most downregulated proteins demonstrated similar expression patterns in both mutants. Apart from proteins like MymT, Rv0849, CydD, and CrtB, which were also found among the most differentially expressed proteins in the exponential phase, PrpC, a component associated with the methyl citrate network, demonstrated high abundance in both mutants. Intriguingly, we also detected highly abundant ribosomal proteins, for example, RplP, RplX, RpmA, RpsT, and RpsZ. Since we did not expect to observe highly abundant ribosomes in the stationary phase, we analyzed the expression of each ribosomal protein that was captured by mass spectrometry. As indicated in Fig. 5a, 100% of the detected ribosomal proteins (31 and 23, respectively) showed higher abundance in the stationary phase. Further evidence for a potential redundant function of MTS1338 and MTS2823 is demonstrated in Fig. 5b, as most ribosomal proteins were equally affected in both mutants.

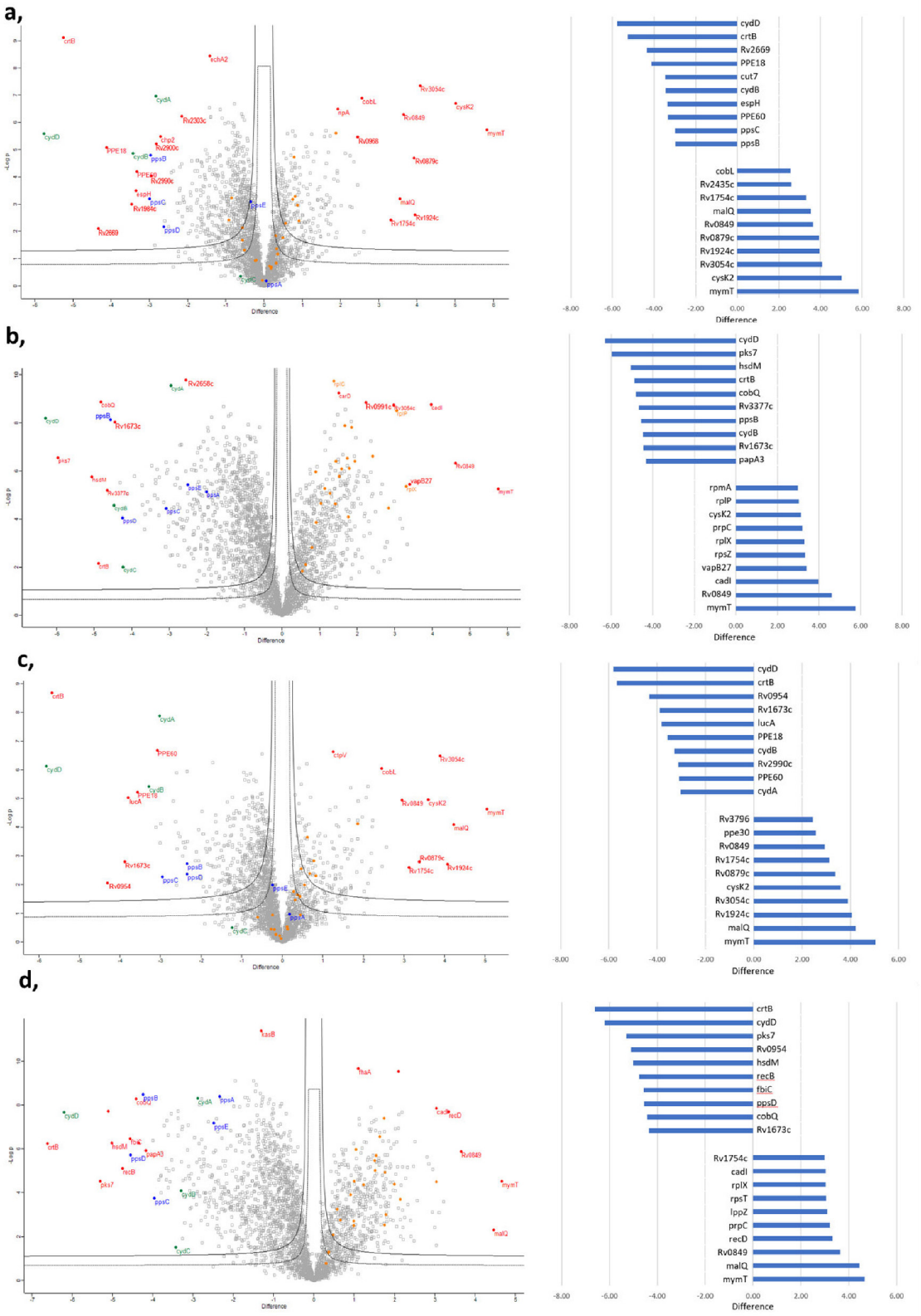


Fig. 4 Proteome profiling demonstrates the global effect of ncRNA-deficiency and a shared function of MTS1338 and MTS2823. The Volcano plots depict alterations in the protein abundance of H37Rv_ΔMTS1338 in exponential (**a**) and stationary phase (**b**) and of H37Rv_ΔMTS2823 in exponential (**c**) and stationary phase (**d**) compared to the wild type. The difference is displayed on the X-axis and the *p*-value for each protein is indicated by the Y-axis. The table on the right side of each volcano plot shows the ten most divergent protein abundances in the respective growth phase.

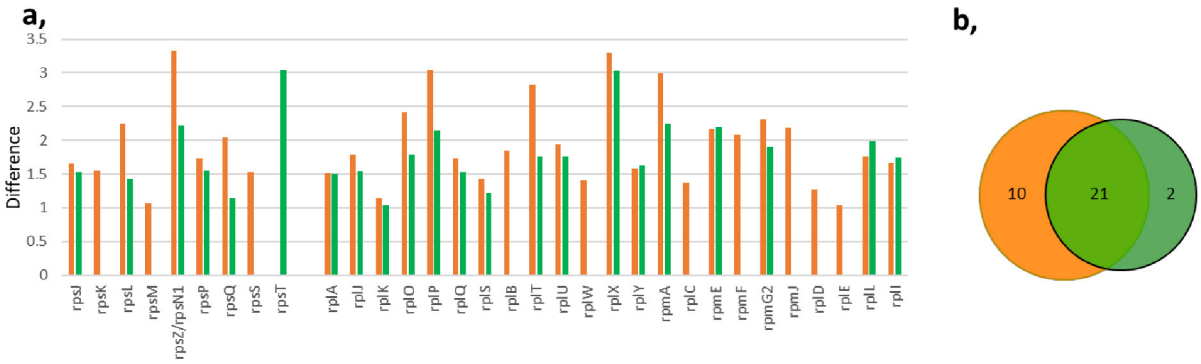


Fig. 5 MTS1338- or MTS2823-deficiency leads to high abundance of ribosomal proteins in stationary phase. Comparative analysis of H37Rv_ΔMTS1338 (orange) and H37Rv_ΔMTS2823 (green) with respect to the abundance of ribosomal proteins in stationary phase compared with wild type (**a**), and the proportion of common dysregulation of ribosomal proteins (**b**).

Phenotypic characterization by evaluation of loss/gain of function did not reveal further functions of the ncRNAs in virulence-associated experiments

Since the transcriptome and proteome of the mutants differ significantly compared with the wild type, the lack of an apparent phenotype when cultured on plate or in liquid medium is intriguing. Although we expect compensatory mechanisms, the intracellular changes could lead to altered properties that affect the fitness of the mutants. Therefore, we performed experiments specifically targeting the processes in which ncRNAs are known to be involved in other pathogens. To test the mutants in experiments associated with infection, we examined the ability to grow in diverse environments or at different temperatures, the response to nutrient deprivation, intracellular survival, and the susceptibility to the first-line antibiotics rifampicin, ethambutol, bedaquiline, and delamanid (Fig. 6).

In contrast to whole transcriptome sequencing and proteome profiling, in which both mutants exhibited numerous alterations compared with the wild type, we did not detect an obvious phenotype other than slight growth advantages of H37Rv_ΔMTS2823, as the response of the mutants was similar to that of the wild type.

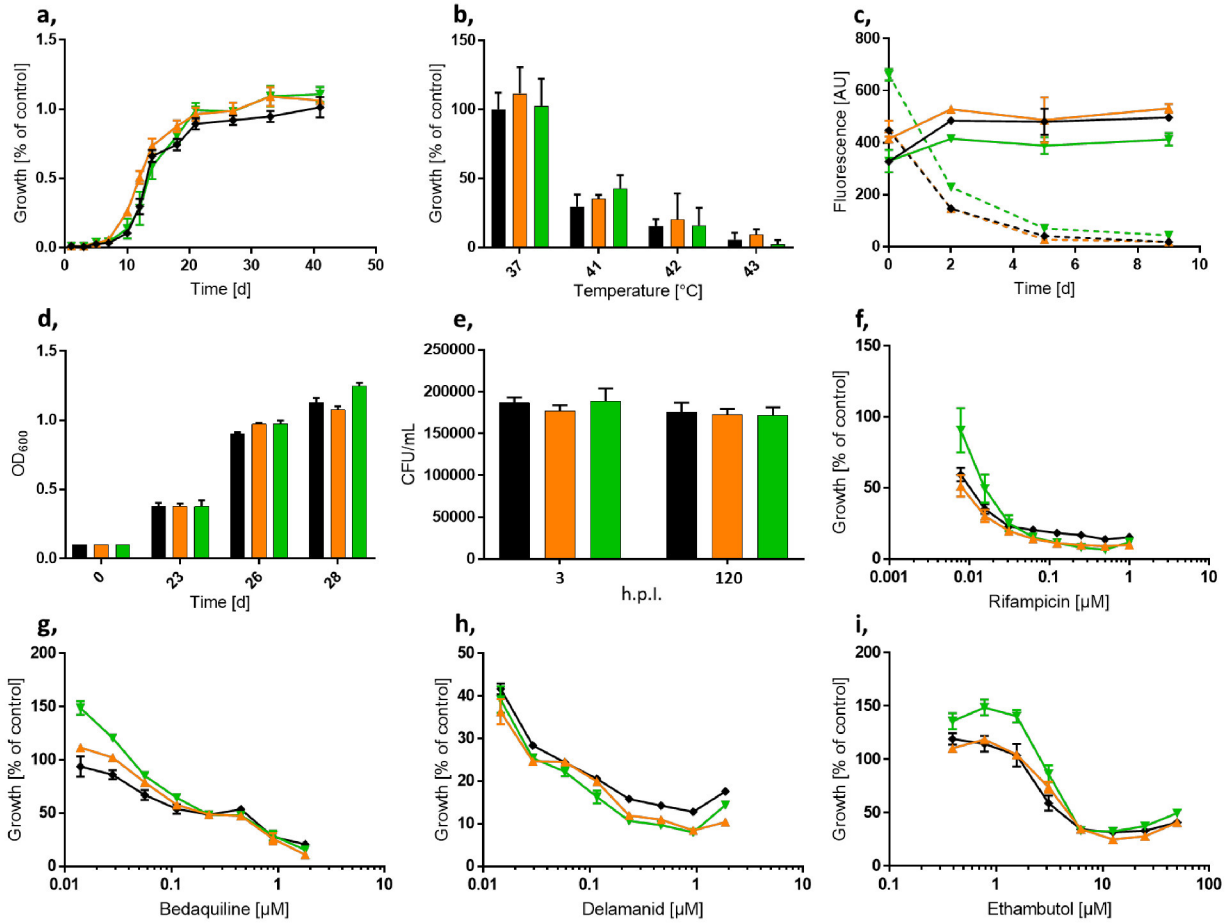


Fig. 6 Phenotypic characterization of *M. tuberculosis* H37Rv_ΔMTS1338 and H37Rv_ΔMTS2823. The strains H37Rv wild type (black), H37Rv_ΔMTS1338 (orange), and H37Rv_ΔMTS2823 (green) were tested in several *in vitro* assays to functionally characterize the ncRNAs MTS1338 and MTS2823. **a**, Growth curves in Middlebrook 7H9 broth. **b**, Viability in Middlebrook 7H9 broth incubated at different temperatures. **c**, Growth ability in acidified medium of pH 5.0 (line) and pH 4.5 (dashed line). **d**, Reactivation of metabolism after a prolonged time of nutrient starvation. **e**, Intracellular survival in THP-1 derived macrophages. **f-i**, Dose-response curves utilizing rifampicin (**f**), bedaquiline (**g**), delamanid (**h**), and ethambutol (**i**).

PERSPECTIVE

Secretion of the ncRNAs MTS1338 and MTS2823 for potential modulation of the host cell following infection?

In addition to deciphering the intracellular function of the ncRNAs, we realized that due to the vast amount of transcripts the studied ncRNAs might also be subjected to secretion and could potentially lead to manipulation of the host cell. It is known that *M. tuberculosis* modulates mechanisms within macrophages, actively hindering phagolysosome fusion, and is prone to releasing membrane vesicles, particularly during infection ([46,47]). Considering that cells of *M. tuberculosis* contain high levels of MTS1338 and MTS2823, released vesicles might contain these transcripts, which could affect gene expression in the host. In an initial test, we analyzed the secretion of the ncRNAs. Using cell-free supernatant from a stationary phase culture of *M. tuberculosis* H37Rv, we detected the transcripts MTS1338 and MTS2823 (Fig. **7**), indicating the release of the ncRNAs into the environment. In the absence of precise controls to validate our hypothesis of secretion of ncRNAs, we are aware of the need to consider the putative origin of the transcripts. The free RNA could possibly originate from lysed cells. However, with regard to the physiology of *M. tuberculosis*, dormant bacteria do not tend to lyse randomly. Considering that the disruption of the mycobacterial cell wall requires very harsh conditions such as physical pressure and the use of detergents, lysis of cells incubated for weeks appears unlikely. The stability of the ncRNA itself is also an important factor, because although we did not perform stability tests, we would not suspect the RNA would be stable over an extended period of incubation, which supports our hypothesis.

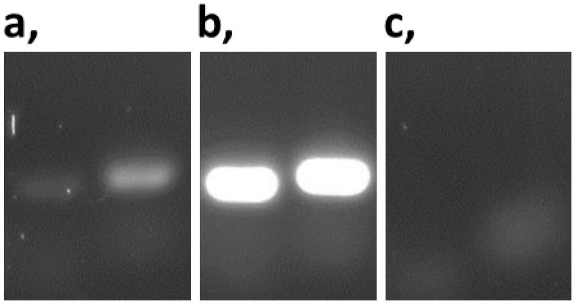


Fig. 7 RT-PCR-based detection of MTS1338 and MTS2823 in the cell-free supernatant of a stationary phase culture of *M. tuberculosis* H37Rv. Diagnostic PCR of MTS1338 (left) and MTS2823 (right) employing the supernatant of a stationary phase culture of *M. tuberculosis* wild type. For detection of the ncRNAs, cDNA synthesized from DNA-depleted culture supernatant (**a**), genomic DNA (**b**, positive control), or the DNA-depleted supernatant (**c**, negative control) was used as template.

Exploiting the highly abundant ncRNAs MTS1338 and MTS2823 as a potential diagnostic tool?

If the secreted ncRNA hypothesis was confirmed, diagnostic methods could take advantage of the large amount of MTS1338 and MTS2823 released by *M. tuberculosis*. In 2018, Demirci described an innovative approach using the 85B mRNA as a diagnostic tool to detect pulmonary *M. tuberculosis* infections ([48]). Regardless of their function, MTS1338 and MTS2823 would represent potential candidates for expanding the targets for this method, which could allow detection of the ncRNAs in patient materials such as sputum or serum.

DISCUSSION

As the treatment regimen for TB has remained unchanged for decades, the urgent need for new antibiotics and innovative approaches to fight *M. tuberculosis* becomes evident. The discrepancy between the rate of development of multi- and extensively drug-resistant strains and the progression of drug discovery is alarming ([1], [49]). While approaches based on hit identification have driven drug discovery in the past, the idea of less time-consuming rational design of drugs is gaining increasing attention ([50]). In order to enable rational design, target structures need to be discovered and validated. The concept of employing antisense approaches using RNA in therapeutic regimens dates back to 1970 and is used, for example, in the treatment of cancer and hepatitis ([51], [52]). Although the basic mode of action of ncRNAs is straightforward, the regulation in bacteria has not yet been elucidated in most cases ([53]). However, small ncRNAs are considered as potential drug targets since various pathogenic microorganisms utilize ncRNAs in, for instance, growth, virulence, and biofilm formation ([17,54–57]). Accordingly, this work aimed for functional characterization of highly abundant ncRNAs of *M. tuberculosis* in order to evaluate the ncRNAs and/or processes in which they are involved as potential drug targets.

Transcriptome sequencing conducted by our group confirmed the high abundance of the ncRNAs MTS1338 and MTS2823 in stationary phase cultures of *M. tuberculosis* H37Rv as previously reported in the literature ([24]). Surprisingly, the function of the transcripts is still not fully understood. MTS1338, a 117 nt ncRNA, is located on the forward strand and overlaps with the hypothetical protein Rv1734c. The transcription of both elements is regulated by the dormancy survival (DosR) regulon in response to environmental stress, such as hypoxia ([24,58]). The DosR-regulon consists of a two-component system that regulates the expression of more than 40 genes involved in metabolic processes, such as respiration, protein biosynthesis, and cellular growth ([59,60]). Since the DosR regulon plays an important role in bacterial virulence, associated elements are considered valid targets for therapeutic regimens ([61]). While overexpression of MTS1338 resulted in slower growth of *M. tuberculosis*, it has been directly

associated with virulence as overexpression confers pathogenic properties to the non-pathogenic *Mycobacterium smegmatis*, a close relative of *M. tuberculosis* ([38–40]). In addition to MTS1338, this study aimed for functional characterization of MTS2823. This molecule gained attention because of its high abundance, especially in the stationary phase or in infections. Arnvig suggested that MTS2823 may play an important role in regulating metabolism because overexpression in the exponential phase resulted in impaired metabolism as the methyl citrate network-associated genes *prpC* and *prpD* were downregulated, presumably in an S6 RNA-like manner ([24]). Although further studies on an *M. smegmatis* homolog of MTS2823 did not confirm this assumption, we provide evidence for the involvement of MTS2823 in the regulation of gene expression ([62]).

While studies with H37Rv_ΔMTS1338 and H37Rv_ΔMTS2823 revealed various potential target genes of both ncRNAs, such as Rv1734c and Rv1735c, whose expression was shown to be dependent on MTS1338, we were intrigued by the similarity of results we observed in both mutants. Regardless of the function of MTS1338 and MTS2823, we found largely correlative results in both ncRNA-deficient mutants in transcriptomic or proteomic approaches, indicating that the same target genes and/or processes were affected. The same trend has been reported independently in the literature, proposing a regulatory function of the ncRNAs, since overexpression of MTS1338 or MTS2823 resulted in a transcriptional shift associated with the transition to the dormant state and downregulation of genes essential for exponential growth, respectively ([24,38]). We confirmed these results by transcriptome sequencing and proteome profiling, as both mutants showed upregulation of genes associated with the methyl citrate network and amino acid biosynthesis, suggesting active metabolism rather than downregulation in knockout mutants. Since our study is based on deprivation instead of overexpression of ncRNA, observations inverse to results based on overexpression validate the roles of MTS1338 and MTS2823 in the regulatory network of *M. tuberculosis*. Following this idea, we subjected the mutants to experiments targeting processes in which ncRNAs are known or suspected to be involved. For example, in *Vibrio cholerae*, overexpression of the ncRNA VrrA plays an important role in infection by controlling the formation of membrane vesicles, a process that also occurs in *M. tuberculosis* ([63], [47]). As MTS1338 is thought to be involved in pathogen-host interactions, we used the knockout mutants H37Rv_ΔMTS1338 and H37Rv_ΔMTS2823 to evaluate potential infection-associated functions ([40]). In addition, ncRNAs in *E. coli* were found to play an important role in the development of drug resistance, as overexpression of *dsrA* led to increased production of efflux pumps, which reduced sensitivity to antibiotics ([27]). The association of MTS1338 with the expression of TatC, a transcolase linked to antibiotic resistance, suggests that ncRNA deprivation may lead to altered susceptibility to antibiotics ([64]).

Apart from the slight growth advantage of H37Rv_ΔMTS2823, which might be due to derepressed metabolism, the mutants did not exhibit any aberrant patterns compared with the wild type. It appeared that the ncRNAs do not participate in processes our experiments aimed for. However, the effect of ncRNA-deficiency could have been attenuated or compensated by other processes or by functionally homologous elements. Although we have discovered a large number of dysregulated genes that could be differentially regulated to mitigate a potential phenotype, we suggest that MTS1338 and MTS2823 could at least partially complement the absence of the respective other ncRNA. Besides correlative results, a potential argument for this hypothesis is demonstrated by the upregulation of MTS1338 in the MTS2823-deficient mutant. This effect was not observed in H37Rv_ΔMTS1338, presumably due to processes that are already saturated by the high quantity of MTS2823 exerting the specific, yet unknown, function. It remains to be elucidated how the numerous changes detected by transcriptome sequencing and proteome profiling did not alter the fitness of the mutants in any of the phenotypic characterization experiments.

Since the levels of MTS1338 and MTS2823 increase with cell maturity and are highly abundant in stationary phase, we mainly focused on a possible role in the stationary phase. The transition to the stationary phase is generally characterized by a restructuring of gene expression, limiting biosynthesis to processes necessary for survival. In this process, bacteria rely on ncRNAs to a) downregulate specific target genes, b) regulate RNA polymerase activity, and potentially c) inhibit the ribosomes to decrease translation ([65], [66]). In *E. coli*, ribosome arrest is achieved by the use of ribosome modulation factor proteins that dimerize 70S ribosome monomers, leading to a translationally inactive state known as the hibernation stage ([67]). Thus, the bacteria retain ribosomal units that allow a rapid transition to a metabolically active state without prior ribosome biosynthesis. Although protein-based approaches suggest similar processes in *M. tuberculosis* and ribosome inhibition could occur through toxin-antitoxin activity, translational inhibition is a process that is poorly characterized in *M. tuberculosis* ([68],[69]). The use of ncRNA to trigger ribosome modulation has already been suggested since a variety of ribosome-associated ncRNAs have been identified ([70]). Since we detected all ribosomal proteins exclusively in higher abundance in the stationary phase of MTS1338- and MTS2823-deficient mutants, a potential role in ribosome regulation becomes evident. The absence of components of the regulatory machinery involved in gene expression could explain the high amount of dysregulated proteins that we detected, since translation would not have been downregulated. The link to the downregulation of processes related to biosynthesis is supported by the findings of Salina and Arnvig, who both proposed ncRNA-based regulation of gene expression based on observations of changes in the transcriptome during overexpression experiments ([38], [24]).

In conclusion, this study demonstrated the involvement of MTS1338 and MTS2823 in the regulatory network of *M. tuberculosis*. Although the mode of action remains to be elucidated, both ncRNAs have been linked to the regulation of gene expression based on observations of the transcriptome and proteome, respectively. Due to similar expression patterns and the observation that deprivation of the respective ncRNA resulted in various changes in gene expression that were highly consistent in both mutants, we conclude that MTS1338 and MTS2823 may have similar functions. We can link the ncRNAs to ribosomal protein regulation, suggesting a role in impeding gene expression at the transition to and/or during the stationary phase.

Besides a role in mycobacterial metabolism, we also recognized the possibility of host cell manipulation through ncRNA secretion. Although the origin of the extracellular ncRNA is controversial, we detected both transcripts in the medium of stationary phase cultures of *M. tuberculosis*, raising the possibility of host cell interference that should be addressed in future experiments. Additionally, secretion of MTS1338 and MTS2823 could potentially be used as biomarkers of *M. tuberculosis* infection to provide diagnostic information such as quantification to classify severity or mycobacterial load.

Our study provides compelling results regarding the regulatory roles of MTS1338 and MTS2823 and underlines the value of studying ncRNAs in pathogens. Research in this area may pave the way for the rational design of therapeutics for the treatment of TB and may also hold the key to facilitating diagnostic efforts to detect infections.

AUTHOR CONTRIBUTIONS

SS and RK designed the study. SS and DP conducted experiments and acquired data. SS, RK, TRI, DP and MK interpreted the data. SS and RK wrote the manuscript.

REFERENCES

1. WHO. *Global tuberculosis report 2023*. *January* (2023).
2. Sensi, P. History of the Development of Rifampin Author (s): P . Sensi Source : Reviews of Infectious Diseases , Vol . 5 , Supplement 3 . The Use of Rifampin in the Treatment of Nontuberculous Infections (Jul . - Aug ., 1983), pp . S402-S406 Published by : Oxfo. *Rev. Infectous Dis.* **5**, S402–S406 (1983).
3. K Mikusová, R A Slayden, G S Besra, and P. J. B. Biogenesis of the mycobacterial cell wall and the site of action of ethambutol. **39**, 2484–2489 (1995).
4. Waters, L. S. & Storz, G. Regulatory RNAs in Bacteria. *Cell* **136**, 615–628 (2009).
5. Beisel, C. L. & Storz, G. Discriminating tastes: Physiological contributions of the Hfq-binding small RNA Spot 42 to catabolite repression. *RNA Biol.* **8**, 766–770 (2011).
6. Sonnleitner, E., Abdou, L. & Haas, D. Small RNA as global regulator of carbon catabolite repression in Pseudomonas aeruginosa. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 21866–21871 (2009).
7. Ge, X. Q. & Lin, H. Noncoding RNAs in the regulation of DNA replication. *Trends in Biochemical Sciences* **39**, (2014).
8. Solans, L. *et al.* The PhoP-Dependent ncRNA Mcr7 Modulates the TAT Secretion System in Mycobacterium tuberculosis. *PLoS Pathog.* **10**, (2014).
9. Tomasini, A. *et al.* The RNA targetome of Staphylococcus aureus non-coding RNA RsaA: Impact on cell surface properties and defense mechanisms. *Nucleic Acids Res.* **45**, 6746–6760 (2017).
10. Girardin, R. C. & McDonough, K. A. Small RNA Mcr11 requires the transcription factor AbmR for stable expression and regulates genes involved in the central metabolism of Mycobacterium tuberculosis. *Mol. Microbiol.* **113**, 504–520 (2020).
11. Gerrick, E. R. *et al.* Small RNA profiling in mycobacterium tuberculosis identifies mrsi as necessary for an anticipatory iron sparing response. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 6464–6469 (2018).
12. Gogol, E. B., Rhodius, V. A., Papenfort, K., Vogel, J. & Gross, C. A. Small RNAs endow a transcriptional activator with essential repressor functions for single-tier control of a global stress regulon. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 12875–12880 (2011).
13. Rutherford, S. T., Van Kessel, J. C., Shao, Y. & Bassler, B. L. AphA and LuxR/HapR reciprocally control quorum sensing in vibrios. *Genes Dev.* **25**, 397–408 (2011).
14. Boisset, S. *et al.* Staphylococcus aureus RNAIII coordinately represses the synthesis of virulence factors and the transcription regulator Rot by an antisense mechanism. *Genes Dev.* **21**, 1353–1366 (2007).
15. Wang, C., Ye, F., Kumar, V., Gao, Y. G. & Zhang, L. H. BswR controls bacterial motility and biofilm formation in Pseudomonas aeruginosa through modulation of the small RNA rsmZ. *Nucleic Acids Res.* **42**, 4563–4576 (2014).
16. Novick, R. P. *et al.* Synthesis of staphylococcal virulence factors is controlled by a regulatory RNA molecule. *EMBO J.* **12**, 3967–75 (1993).
17. Mann, B. *et al.* Control of virulence by small RNAs in Streptococcus pneumoniae. *PLoS Pathog.* **8**, 34 (2012).
18. Toledo-Arana, A. *et al.* The Listeria transcriptional landscape from saprophytism to

virulence. *Nature* **459**, 950–956 (2009).

19. Storz, G., Vogel, J. & Wassarman, K. M. Regulation by Small RNAs in Bacteria: Expanding Frontiers. *Mol. Cell* **43**, 880–891 (2011).

20. Chao, Y. & Vogel, J. The role of Hfq in bacterial pathogens. *Curr. Opin. Microbiol.* **13**, 24–33 (2010).

21. Franze De Fernandez, M. T., Eoyang, L. & August, J. T. Factor fraction required for the synthesis of bacteriophage Qβ-RNA. *Nature* **219**, 588–590 (1968).

22. Arluison, V. *et al.* Spectroscopic observation of RNA chaperone activities of Hfq in post-transcriptional regulation by a small non-coding RNA. *Nucleic Acids Res.* **35**, 999–1006 (2007).

23. Sittka, A., Pfeiffer, V., Tedin, K. & Vogel, J. The RNA chaperone Hfq is essential for the virulence of Salmonella typhimurium. *Mol. Microbiol.* **63**, 193–217 (2007).

24. Arnvig, K. B. *et al.* Sequence-based analysis uncovers an abundance of non-coding RNA in the total transcriptome of Mycobacterium tuberculosis. *PLoS Pathog.* **7**, (2011).

25. Arnvig, K. & Young, D. Non-coding RNA and its potential role in Mycobacterium tuberculosis pathogenesis. *RNA Biol.* **9**, 427–436 (2012).

26. Gryczan, T. J., Grandi, G., Hahn, J., Grandi, R. & Dubnau, D. Conformational alteration of mRNA structure and the posttranscriptional regulation of erythromycin-induced drug resistance. *Nucleic Acids Res.* **8**, 6081–6097 (1980).

27. Nishino, K., Yamasaki, S., Hayashi-Nishino, M. & Yamaguchi, A. Effect of overexpression of small non-coding DsrA RNA on multidrug efflux in Escherichia coli. *J. Antimicrob. Chemother.* **66**, 291–296 (2011).

28. Matsui, M. & Corey, D. R. Non-coding RNAs as drug targets. *Nat. Publ. Gr.* **16**, 167–179 (2016).

29. Korte, J. *et al.* Trehalose-6-Phosphate-Mediated Toxicity Determines Essentiality of OtsB2 in Mycobacterium tuberculosis In Vitro and in Mice. *PLoS Pathog.* **12**, e1006043 (2016).

30. Larsen, M. H., Biermann, K., Tandberg, S., Hsu, T. & Jacobs, W. R. Genetic Manipulation of Mycobacterium tuberculosis. *Curr. Protoc. Microbiol.* **6**, 1–21 (2007).

31. Michalski, A. *et al.* Ultra high resolution linear ion trap orbitrap mass spectrometer (orbitrap elite) facilitates top down LC MS/MS and versatile peptide fragmentation modes. *Mol. Cell. Proteomics* **11**, 1–11 (2012).

32. Olsen, J. V. *et al.* Parts per million mass accuracy on an orbitrap mass spectrometer via lock mass injection into a C-trap. *Mol. Cell. Proteomics* **4**, 2010–2021 (2005).

33. Cox, J. *et al.* Andromeda: A peptide search engine integrated into the MaxQuant environment. *J. Proteome Res.* **10**, 1794–1805 (2011).

34. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **26**, 1367–1372 (2008).

35. Cox, J. *et al.* Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol. Cell. Proteomics* **13**, 2513–2526 (2014).

36. Tyanova, S. *et al.* The Perseus computational platform for comprehensive analysis of

(prote)omics data. *Nat. Methods* **13**, 731–740 (2016).

37. Demaio, J., Zhang, Y., Ko, C., Young, D. B. & Bishai, W. R. A stationary-phase stress-response sigma factor from Mycobacterium tuberculosis. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 2790–2794 (1996).

38. Salina, E. G. *et al.* MTS1338, a small Mycobacterium tuberculosis RNA, regulates transcriptional shifts consistent with bacterial adaptation for entering into dormancy and survival within host macrophages. *bioRxiv* 474304 (2019). doi:10.1101/474304

39. Ignatov, D. V. *et al.* Dormant non-culturable Mycobacterium tuberculosis retains stable low-abundant mRNA. *BMC Genomics* **16**, 954 (2015).

40. Bychenko, O. *et al.* Mycobacterium tuberculosis Small RNA MTS1338 Confers Pathogenic Properties to Non-Pathogenic Mycobacterium smegmatis. 1–18 (2021).

41. Wolschendorf, F. *et al.* Copper resistance is essential for virulence of Mycobacterium tuberculosis. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 1621–1626 (2011).

42. Steiner, E. M. *et al.* CysK2 from Mycobacterium tuberculosis is an O-phospho-L-serine-dependent S-sulfocysteine synthase. *J. Bacteriol.* **196**, 3410–3420 (2014).

43. Ramakrishnan, L., Tran, H. T., Federspiel, N. A. & Falkow, S. A crtB homolog essential for photochromogenicity in Mycobacterium marinum: Isolation, characterization, and gene disruption via homologous recombination. *J. Bacteriol.* **179**, 5862–5868 (1997).

44. Kana, B. D. *et al.* Characterization of the cydAB-encoded cytochrome bd oxidase from Mycobacterium smegmatis. *J. Bacteriol.* **183**, 7076–7086 (2001).

45. Goldman, B. S., Gabbert, K. K. & Kranz, R. G. The temperature-sensitive growth and survival phenotypes of Escherichia coli cydDC and cydAB strains are due to deficiencies in cytochrome bd and are corrected by exogenous catalase and reducing agents. *J. Bacteriol.* **178**, 6348–6351 (1996).

46. Via, L. E. *et al.* Arrest of mycobacterial phagosome maturation is caused by a block in vesicle fusion between stages controlled by rab5 and rab7. *J. Biol. Chem.* **272**, 13326–13331 (1997).

47. Prados-Rosales, R. *et al.* Mycobacteria release active membrane vesicles that modulate immune responses in a TLR2-dependent manner in mice. *J. Clin. Invest.* **121**, 1471–1483 (2011).

48. Demirci, M. *et al.* Diagnostic performance of the RT-qPCR method targeting 85B mRNA in the diagnosis of pulmonary Mycobacterium tuberculosis infection. *J. Infect. Public Health* **11**, 662–666 (2018).

49. CDC. Antibiotic Resistance Threats in the United States, 2013. *Cent. Dis. Control Prev.* **37**, 828–831 (2013).

50. Reddy, M. R. & Parrill, A. L. Overview of rational drug design. *ACS Symp. Ser.* **719**, 1–11 (1999).

51. Zamecnik, P. C. & Stephenson, M. L. Inhibition of Rous sarcoma virus replication and cell transformation by a specific oligodeoxynucleotide. *Proc. Natl. Acad. Sci. U. S. A.* **75**, 280–284 (1978).

52. Rooij, E. & Kauppinen, S. Development of micro RNA therapeutics is coming of age . *EMBO Mol. Med.* **6**, 851–864 (2014).

53. Dutta, T. & Srivastava, S. Small RNA-mediated regulation in bacteria: A growing palette of diverse mechanisms. *Gene* **656**, 60–72 (2018).

54. Yang, Y. *et al.* The Small RNA sr8384 is a crucial regulator of cell growth in solventogenic clostridia. *Appl. Environ. Microbiol.* **86**, (2020).

55. Michaux, C. *et al.* Involvement of Enterococcus faecalis small RNAs in stress response and virulence. *Infect. Immun.* **82**, 3599–3611 (2014).

56. Koo, J. T., Alleyne, T. M., Schiano, C. A., Jafari, N. & Lathem, W. W. Global discovery of small RNAs in Yersinia pseudotuberculosis identifies Yersinia-specific small, noncoding RNAs required for virulence. *Proc. Natl. Acad. Sci. U. S. A.* **108**, (2011).

57. Papenfort, K., Förstner, K. U., Cong, J.-P., Sharma, C. M. & Bassler, B. L. Differential RNA-seq of *Vibrio cholerae* identifies the VqmR small RNA as a regulator of biofilm formation. *Proc. Natl. Acad. Sci.* **112**, E766–E775 (2015).

58. Kumar, K. & Dutta, T. Transcriptional activation of the Mycobacterium tuberculosis virulence-associated small RNA MTS1338 by the response regulators DosR and PhoP. *FEBS Lett.* **598**, 1034–1044 (2024).

59. Park, H. D. *et al.* Rv3133c/dosR is a transcription factor that mediates the hypoxic response of Mycobacterium tuberculosis. *Mol. Microbiol.* **48**, 833–843 (2003).

60. Voskuil, M. I. *et al.* Inhibition of respiration by nitric oxide induces a Mycobacterium tuberculosis dormancy program. *J. Exp. Med.* **198**, 705–713 (2003).

61. Malhotra, V. *et al.* Disruption of response regulator gene, devR, leads to attenuation in virulence of Mycobacterium tuberculosis. *FEMS Microbiol. Lett.* **231**, 237–245 (2004).

62. Pánek, J. *et al.* The suboptimal structures find the optimal RNAs: Homology search for bacterial non-coding RNAs using suboptimal RNA structures. *Nucleic Acids Res.* **39**, 3418–3426 (2011).

63. Song, T. *et al.* A new Vibrio cholerae sRNA modulates colonization and affects release of outer membrane vesicles. *Mol. Microbiol.* **70**, 100–111 (2008).

64. McDonough, J. A., Hacker, K. E., Flores, A. R., Pavelka, M. S. & Braunstein, M. The twin-arginine translocation pathway of Mycobacterium smegmatis is functional and required for the export of mycobacterial β-lactamases. *Journal of Bacteriology* **187**, 7667–7679 (2005).

65. Johansen, J., Rasmussen, A. A., Overgaard, M. & Valentin-Hansen, P. Conserved Small Non-coding RNAs that belong to the σE Regulon: Role in Down-regulation of Outer Membrane Proteins. *Journal of Molecular Biology* **364**, 1–8 (2006).

66. Hnilicová, J. *et al.* Ms1, a novel sRNA interacting with the RNA polymerase core in mycobacteria. *Nucleic Acids Res.* **42**, 11763–11776 (2014).

67. Yoshida, H. *et al.* The ribosome modulation factor (RMF) binding site on the 100S ribosome of Escherichia coli. *J. Biochem.* **132**, 983–989 (2002).

68. Barth, V. C. *et al.* Toxin-mediated ribosome stalling reprograms the Mycobacterium tuberculosis proteome. *Nat. Commun.* **10**, (2019).

69. Kumar, A. *et al.* Mycobacterium tuberculosis DosR regulon gene Rv0079 encodes a putative, 'dormancy associated translation inhibitor (DATIN)'. *PLoS One* **7**, 1–7 (2012).

70. Pircher, A., Gebetsberger, J. & Polacek, N. Ribosome-associated ncRNAs: An emerging class of translation regulators. *RNA Biol.* **11**, 1335–1339 (2014).

5. The GtrA-like lipid floppase Rv3277 is essential for cell surface mannosylation in *Mycobacterium tuberculosis*

Unpublished, manuscript in preparation

Overall contribution to the paper:

- Writing the first version of the complete manuscript draft
- Generation of the conditional mutant c-*Rv3277t*_tetOFF
- Generation of the conditional mutant c-*Msmeg_1817*_tetON
- Evaluation of the conditional mutants in liquid and on solid medium
- Generation of complemented conditional mutants
- Evaluation of complementation in liquid and on solid medium
- Evaluation of 3D modeling
- Extraction of proteins, evaluation of proteome profiling data

The GtrA-like lipid floppase Rv3277 is essential for cell surface mannosylation in *Mycobacterium tuberculosis*

Kristin Vill[*], Steffen Schindler[1,*], Lasse van Geelen[1], David Podlesainski[2], Sudagar S. Gurcha[3], Milena Hänisch[1], Kathrin Lindner[1], Martine Gilleron[4], Christoph Gertzen[5], Violetta Krisilia[1], Ben J. Appelmelk[6], Holger Gohlke[5], Jérôme Nigou[4], Gurdyal S. Besra[3], Markus Kaiser[2], and Rainer Kalscheuer[1,§]

[1] Heinrich Heine University Düsseldorf, Faculty of Mathematics and Natural Sciences, Institute of Pharmaceutical Biology and Biotechnology, 40225 Düsseldorf, Germany

[2] Department of Chemical Biology, ZMB, Faculty of Biology, University of Duisburg-Essen, 45177 Essen, Germany

[3] School of Biosciences, University of Birmingham, Edgbaston, Birmingham, B15 2TT, United Kingdom

[4] Institute of Pharmacology and Structural Biology (IPBS), University of Toulouse, CNRS, University of Toulouse III-Paul Sabatier, Toulouse, France

[5] Institute of Pharmaceutical and Medicinal Chemistry, Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany

[6] Molecular Microbiology/Medical Microbiology and Infection Control, Amsterdam University Medical Centers, Amsterdam, Netherlands.

* These authors contributed equally

§ Author for correspondence: rainer.kalscheuer@hhu.de

Abstract

Tuberculosis continues to be a threat to health around the globe. The causative agent, *Mycobacterium tuberculosis*, possesses not only the ability to evade the immune system but also a complex cell wall that confers intrinsic resistance to antibiotics and is crucial for pathogenicity. Although research has made great progress on cell wall issues, some key steps in biogenesis of the unique structure remain obscure. Glycoconjugates, which have a vital function in host-pathogen interactions, are synthesized in a bipartite process in the cyto- and periplasm. Although the biosynthesis is well known, the mechanism by which the carbohydrates enter the outer leaflet of the cell membrane remains unknown. Using bioinformatics, gene inactivation and phenotypic characterization of mutant strains, we have identified Rv3277 as a GtrA-like transporter and revealed its role in the translocation of glycoconjugate building blocks. Thus, we have uncovered a missing part in the biosynthetic pathway of glycoconjugate formation and identified a potential new drug target candidate.

Keywords

Tuberculosis, mycobacteria, cell wall, glycoconjugates, floppases

Introduction

Only surpassed by Covid-19 in the last years, tuberculosis (TB) remains one of the most serious infectious diseases with approx. 10.6 million new cases and 1.6 million deaths in 2021 [1]. The increasing emergence of multidrug-resistant strains of the causative pathogen *Mycobacterium tuberculosis* underlines the urgent need for new antibiotics. Much of its pathogenicity is attributed to the unusual, complex cell wall structure (reviewed in [2]). Compared to other Gram-positive bacteria, and in addition to the characteristic mycolic acids, mycobacteria do possess various additional unique cell wall molecules that confer exceptional physicochemical resistance to various stresses, mediate high intrinsic antibiotic resistance by reducing permeability, and/or play important roles in interaction with the host´s immune system [2]. Some of the most abundant cell wall molecules are rich in mannose, arabinose, or both. These glycoconjugates comprise molecules such as phosphatidylinositol mannosides (PIMs), lipomannan (LM), arabinogalactan (AG) and lipoarabinomannan (LAM), some of which have been shown to play important roles in virulence by interaction with various host cell receptors such as the mannose receptor, complement receptors, or the dendritic cell-specific intercellular adhesion molecule-3-grabbing non-integrin (DC-SIGN) [3-8]. Since biosynthesis of these glycoconjugates is essential for viability of *M. tuberculosis*, they offer several validated or potential target structures for chemotherapy of TB, as exemplified by the first-line antibiotic ethambutol that interferes with AG biosynthesis [9, 10].

Giving their outstanding importance for viability and pathogenesis of *M. tuberculosis*, the biosynthesis of arabinosylated and mannosylated mycobacterial cell wall constituents has been extensively investigated over the last decades. However, while large aspects of the biosynthetic pathways and enzymes have been discovered, several essential steps are still unknown. Biosynthesis of the arabinosylated and mannosylated mycobacterial cell wall glycoconjugates are membrane-associated bipartite processes that are initiated at the cytoplasmic side of the plasma membrane. Subsequently, intermediates are translocated (“flopped”) to the periplasmic leaflet of the membrane, where they are extended by various arabinosyl- and/or mannosyltransferases. In the biosynthesis of AG, after synthesis at the cytosolic face of the membrane by galactofuranosyltransferases using UDP-galactofuranose as donor substrate, the galactan core C50-P-P-GlcNAc-Rha-Galf_{30} is flipped across the plasma membrane through a Wzm-Wzt-type ABC transporter comprising the proteins Rv3781 and Rv3783 [11]. The subsequent steps in AG assembly take place at the periplasmic side of the plasma membrane by arabinofuranosyltransferases that transfer arabinofuranose residues onto C50-P-P-GlcNAc-Rha-Galf_{30} utilizing decaprenyl-monophosphoryl-D-arabinofuranose (DPA) as a donor substrate [12]. Similarly, the early steps in biosynthesis of the mannosylated glycoconjugates PIM, LM, and LAM occur at the cytoplasmic side of the plasma membrane. The α-mannopyranosyltransferases PimA (Rv2610c), PimB (Rv2188c) and PimC or a yet-

unidentified mannosyltransferase mediate the stepwise mannosylation of the *myo*-inositol moiety of the phospholipid phosphatidyl-*myo*-inositol, employing GDP-mannopyranose as the donor substrate, yielding phosphatidylinositol tetramannoside (PIM_4). Following mono- and diacylation by the acyltransferase Rv2611c and a yet-unknown enzyme, the resulting Ac_1/Ac_2PIM_4 is flopped by an unknown mechanism from the cytoplasmic to the periplasmic side of the plasma membrane [13]. Here, Ac_1/Ac_2PIM_4 serves as a branch point, which can either be extended by the mannopyranosyltransferase PimE (Rv1159) to form higher-order PIMs (Ac_1/Ac_2PIM_6), or by the mannopyranosyltransferase MptA (Rv2174), MptB (Rv1459c) and MptC (Rv2181) to synthesize LM. The mentioned mannopyranosyltransferases represent GT-C type glycosyltransferases that utilize decaprenyl-monophosphoryl-D-mannopyranose (DPM) as the donor substrate [14]. For LAM biosynthesis, a yet-unidentified arabinofuranosyltransferase first primes the mannan core of LM with a few arabinofuranose residues, after which EmbC (Rv3793) extends the primed LM with the addition of 12–16 arabinofuranose residues. Subsequently, the branching enzyme AftC and the arabinofuranosyltransferase AftB, which is also involved in AG biosynthesis, complete the nonreducing arabinan domain of LAM. EmbC, AftC, and AftB all employ DPA as the donor substrate to result in addition of altogether 55–70 arabinofuranose residues to LM. Finally, the arabinan domain of LAM is modified with mannose residues by the mannopyranosyltransferases CapA (Rv1635c) and MptC (Rv2181), both employing DPM as the donor substrate, resulting in ManLAM [13, 15] (for a schematic overview see Supplementary Fig. 1).

The biosyntheses of DPA and DPM, which are required for extension of the arabinosylated and mannosylated glycoconjugates at the periplasmic face of the cytoplasm membrane, are themselves bipartite processes that are initiated at the cytoplasmic side of the plasma membrane and require subsequent translocation to the periplasmic leaflet of the membrane. DPM is synthesized at the cytoplasmic side of the membrane by transfer of mannose from the nucleotide-activated donor substrate GDP-mannopyranose to the membrane lipid carrier C50-P mediated by Ppm1 (Rv2051c), and subsequently needs to be flopped to the periplasmic side of the cytoplasm membrane (Fig. 1). In contrast, DPA requires an activated phosphosugar, phosphoribosylpyrophosphate, which is produced by phosphorylation of ribose-1-phosphate catalyzed by phosphoribosyl-1-pyrophosphate synthetase PrsA (Rv1017). Next, decaprenylphosphoryl-5-phosphoribose synthase UbiA (Rv3806c) transfers phosphoribosylpyrophosphate to C50-P, yielding decaprenyl-monophosphoryl-D-ribofuranosephosphate, which is subsequently dephosphorylated by Rv3807 to produce decaprenyl-monophosphoryl-D-ribofuranose (DPR). Decaprenylphosphoribose 2′ epimerase composed of DprE1 (Rv3790) and DprE2 (Rv3791) subunits then catalyzes the epimerization of the ribosyl unit of DPR, resulting in DPA (for a

schematic overview see Supplementary Fig. 2). Recently, epimerization mediated by DprE1-DprE2 was reported to occur extracellularly at the periplasmic face of the membrane [16]. Thus, this suggests that flopping might occur at the stage of the DPR intermediate. In summary, the biosynthesis of arabinosylated and mannosylated glycoconjugates requires several floppase reactions at the level of DPM, DPR, Ac_1/Ac_2PIM_4 and C50-P-P-GlcNAc-Rha-Gal*f*$_{30}$, of which only flopping of the galactan core C50-P-P-GlcNAc-Rha-Gal*f*$_{30}$ by the Wzm–Wzt-type ABC transporter Rv3781-Rv3783 has been clearly established [11].

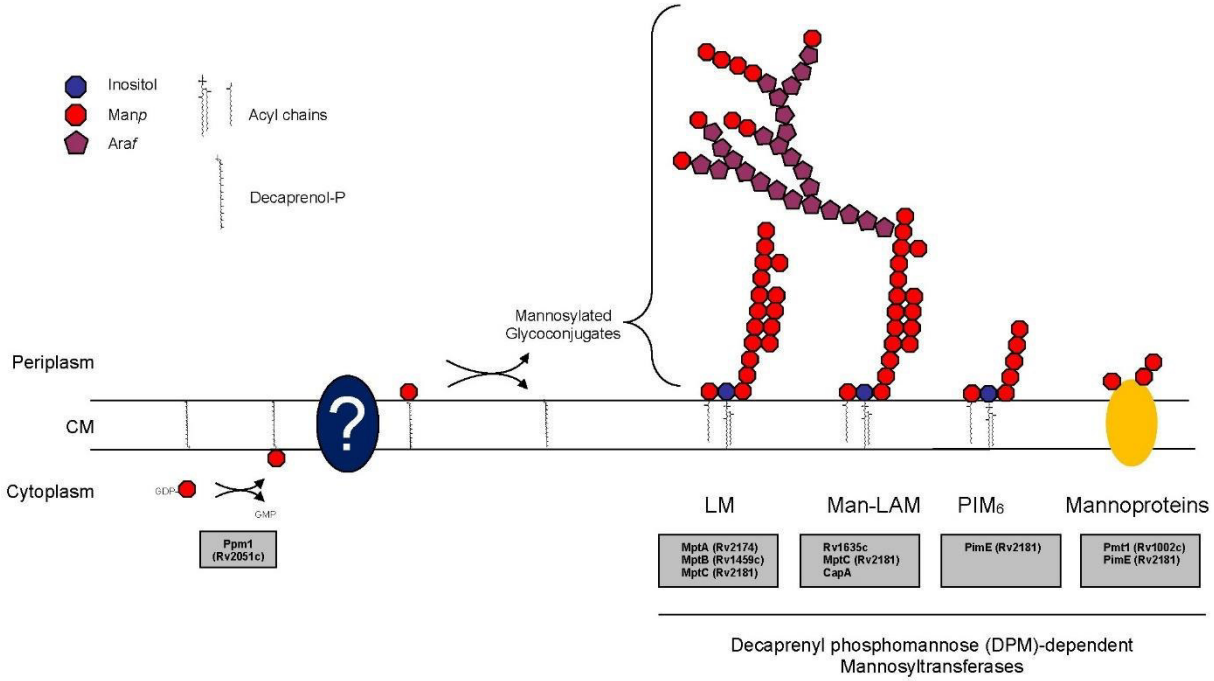


Fig. 1: Schematic representation of the bipartite biosynthesis of decaprenyl monophosphoryl-D-mannopyranose (DPM) in *M. tuberculosis*. DPM synthesis is initiated at the cytoplasmic side of the membrane by transfer of mannose from the nucleotide-activated donor substrate GDP-mannopyranose to the membrane lipid carrier C50-P mediated by Ppm1 (Rv2051c), and subsequently needs to be flopped to the periplasmic side of the cytoplasm membrane by a yet-unknown lipid translocase. At the periplasmic face of the membrane, DPM serves as the donor substrate for various mannopyranosyltransferases for the synthesis of different mannosylated cell surface glycoconjugates. LM, lipomannan; Man-LAM, mannose-capped lipoarabinomannan; PIM_6, phosphatidyl inositol hexamannoside.

In other prokaryotes, translocases that have similarities to small multidrug resistance transporters can perform lipid floppase function [17]. A well-known class of these floppases is the GtrA-like floppases. They are small, very hydrophobic enzymes of ~120 amino acids. The

best-known example is from *Shigella flexneri*, where the gene cluster *gtrA*, *gtrB*, and *gtrX* mediates the conversion of O antigen, the distal portion of LPS, by transferring a glycosyl residue [18]. GtrB catalyzes the transfer of glucose from a nucleotide-activated precursor to a lipid carrier molecule. GtrA then translocates the lipid-linked glucose across the cytoplasmic membrane, where GtrX finally mediates the actual glycosyl transfer [18-20]. In *M. tuberculosis*, Rv3789 codes for a putative GtrA-type multidrug resistance-like transporter located directly upstream of *dprE1*. It was demonstrated that the deletion of the homolog of Rv3789 in *Mycobacterium smegmatis*, Msmeg_6372, resulted in a reduction in the arabinose content of both AG (approximately -15%) and LAM (approximately -35%) and simultaneous accumulation of DPA [21], which is compatible with Rv3789 acting as a DPA floppase. However, deletion of Rv3789 in *M. tuberculosis* resulted in no discernable effects, suggesting the existence of proteins with redundant function and/or complementing mechanisms in this organism. In contrast, studies on Rv3789 by a different research group revealed a *M. tuberculosis* gene deletion mutant that exhibited a phenotype resembling that of the *M. smegmatis* Msmeg_6372 gene deletion mutant, characterized by decreased incorporation of arabinan into AG accompanied by accumulation of the precursor DPA [22]. However, it was later suggested that synthesis of DPA might take place in the periplasm via epimerization of DPR, which implies flopping occurs at the level of DPR and not DPA [16], and it was demonstrated that Rv3789 interacts with the priming arabinosyltransferase AtfA [22]. Thus, it was suggested that Rv3789 might not act as a DPA floppase but rather plays an important, albeit not strictly essential, role by recruiting AftA for AG biosynthesis. On the other hand, heterologous expression of the bona-fide *Escherichia coli* undecaprenylphosphate-4-amino-4-deoxy-L-arabinose floppase, ArnE/ArnF, functionally complemented the lack of Rv3789 and restored LAM formation, favoring the predicted function as a floppase [21]. Regardless of whether Rv3789 is a DPA floppase or just indirectly involved in AG biosynthesis by functioning as a membrane anchor for DPA-dependent arabinosyltransferases, dispensability of the encoding gene proves that one or more other yet-unknown floppases must exist in *M. tuberculosis*.

In this study, we used a bioinformatics approach to identify a second protein in *M. tuberculosis* exhibiting homology to the flippase GtrA of *S. flexneri* [18]. Considering the predicted essentiality of *Rv3277*, we investigated its potential role in the cell wall assembly and evaluated its role for the viability of *M. tuberculosis*.

Results

Identification of the essential gene *Rv3277t* with an alternative translational start site.

To identify genes encoding potential GtrA-like proteins, we performed a Domain Enhanced Lookup Time Accelerated BLAST (DELTA-BLAST) search of the *M. tuberculosis* H37Rv genome using the prototypic *S. flexneri* GtrA protein sequence (UniProt accession number Q9T1D7). In addition to Rv3789, this identified a second protein, Rv3277, exhibiting low identity (14.39% in a 139 amino acid overlap) (Fig. 2a). Rv3277 is annotated as a probable essential conserved transmembrane protein comprising 272 amino acids. It is a mycobacterial core protein conserved in many different strains [23] and contains four transmembrane helices according to Alphafold prediction [24], similar to prototypic GtrA.

To examine the predicted essential function of Rv3277, we employed a conditional silencing-approach, which enabled us to fine-tune the expression of the corresponding gene *Rv3277* in *M. tuberculosis*. For this, we inserted four *tet* operator sites upstream of the annotated start codon, allowing concentration-dependent binding of a plasmid-borne reverse tet repressor (revTetR) in presence of anhydrotetracycline (ATc). Unexpectedly however, contrasting its predicted essential function based on genome-wide analyses of transposon mutant pools [25, 26], the c-*Rv3277*_tet-OFF mutant did not show any obvious phenotype in presence of ATc to silence gene expression (Suppl. Fig 3a). When comparing Rv3277 to orthologues in other mycobacteria, we noticed that many of the encoded proteins are much shorter, revealing a substantial 49 amino acid N-terminal extension in the annotated Rv3277 protein (Suppl. Fig. 3b). We thus inspected the nucleotide sequence and identified an alternative open reading frame with a GTG translation start site 147 bp downstream of the annotated ATG translation start site, encoding a truncated version of the annotated *Rv3277* gene, which we refer to as *Rv3277t* (Suppl. Fig. 3c). We therefore presumed false annotation of Rv3277 and hypothesized that the gene is expressed recognizing the alternative GTG start codon, resulting in a protein comprising just 223 amino acids. Insertion of the controllable promoter cassette described above upstream of the alternative GTG start codon led to the generation of the conditional *M. tuberculosis* c-*Rv3277t*_tetOFF mutant (Suppl. Fig. 3d). Silencing of the target gene in this conditional mutant in the presence of ATc demonstrated the expected phenotype with lack of growth both on Middlebrook 7H10 agar and in Middlebrook 7H9 liquid medium, respectively (Fig. 2b + c). Viability dropped dose-dependently when fully induced cells were plated on ATc-containing solid medium, indicating that loss of Rv3277t function is bactericidal (Fig. 2d). The 147-bp region upstream of the GTG start codon likely includes the native promoter, explaining why our initial silencing approach with the controllable promoter cassette inserted upstream of the annotated ATG start codon was unsuccessful since the gene was still expressed from its endogenous promoter. In addition to our gene silencing experiments, next-generation sequence analysis of a saturated Himar1 transposon

library in *M. tuberculosis* supported the assumption that the annotated N-terminal extension of Rv3277 is incorrect as the absence of transposon insertions showed gene essentiality only for potential TA dinucleotide insertion sites downstream of the mentioned GTG codon [26]. The use of GTG as a translation start site is common in *M. tuberculosis* as it is used in approx. 35% of tubercular genes [27].

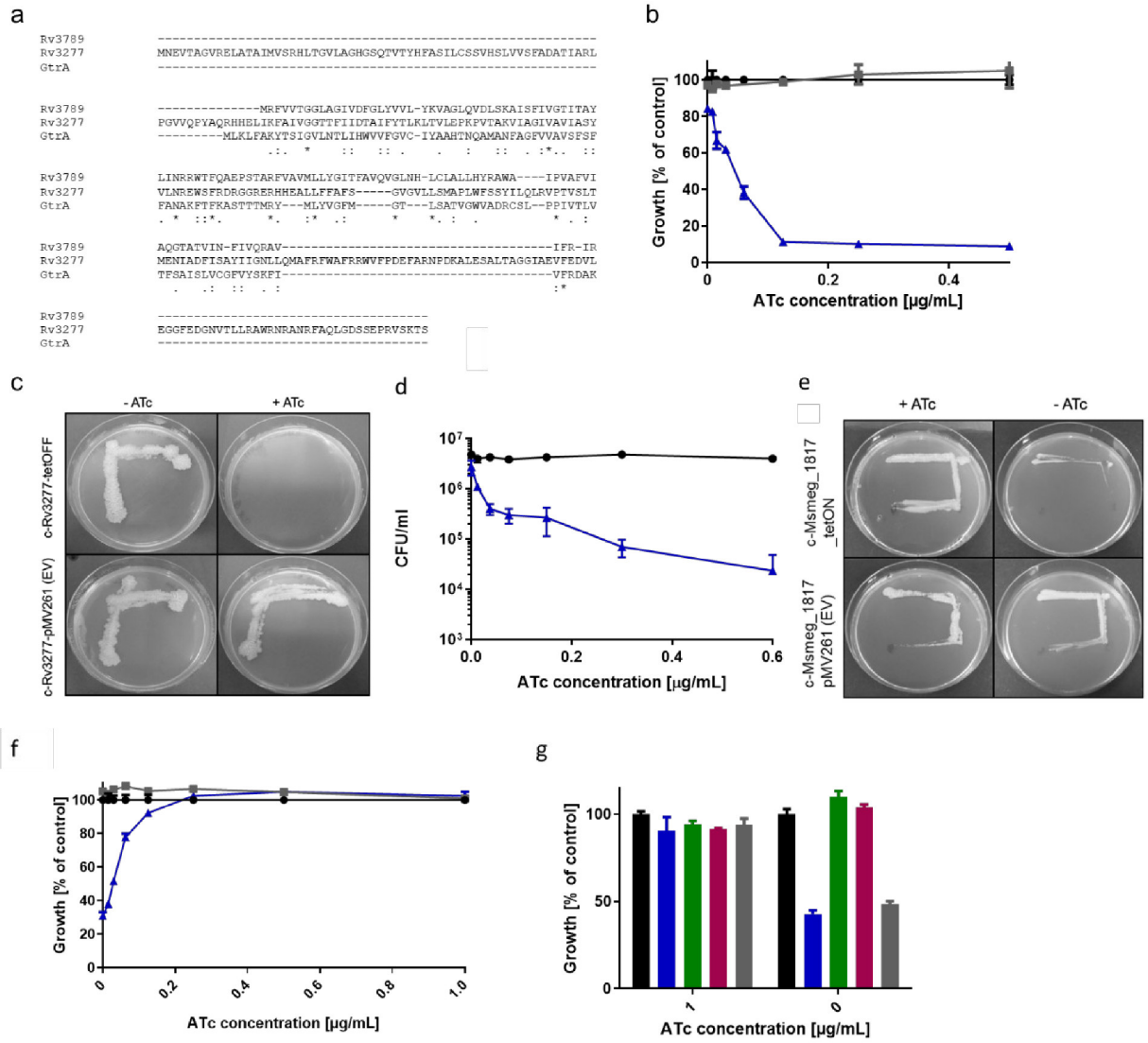


Fig. 2: Demonstration of essentiality of the gene *Rv3277t* in *M. tuberculosis* and its orthologue *MSMEG_1817* in *M. smegmatis*. **a**, Multiple sequence alignment of *Rv3789* (*M. tuberculosis*), *Rv3277* (*M. tuberculosis*), and *gtrA* (*S. flexneri*) using CLUSTAL by Muscle (3.8). Asterisks show positions that are highly identical, whereas colon and period show positions with lower similar residues. The N-terminus of Rv3277 is not available in the compared proteins. **b – c,** ATc-dependent growth inhibition of the conditional *M. tuberculosis* c-*Rv3277t*_tetOFF mutant (blue), the empty vector control strain (c-*Rv3277t* pMV261, grey) and the *M. tuberculosis* H37Rv wild type (black) in Middlebrook 7H9 liquid medium (**b**) and on

Middlebrook 7H10 agar (**c**). Growth in **c** was documented after 14 days of incubation. **d,** Dose-response curves of c-*Rv3277t*_tetOFF (blue) and *M. tuberculosis* wild type H37Rv (black) against increasing concentrations of ATc on 7H10 Middlebrook agar. A fully induced liquid culture was grown to $OD_{600\,nm}$ = 0.3, washed, and serial dilutions were plated on a solid medium containing the indicated ATc concentrations. Colony-forming units were counted after 14 days. A decrease in viable cell counts indicates a bactericidal effect of loss of Rv3277t functionality. **e - f,** ATc-dependent growth of the conditional *M. smegmatis* c-*Msmeg_1817*_tetON mutant (blue) and of the empty vector control strain (c-*Msmeg_1817* pMV261, grey) on Middlebrook 7H10 agar (**e**) and in Middlebrook 7H9 liquid medium (**f**). Growth in **e** was documented after 3 days of incubation. **g,** Complementation of the conditional *M. smegmatis* c-*Msmeg_1817*_tetON mutant (blue) *via* constitutive expression of a plasmid-borne copy of *Msmeg_1817* (green) or *Rv3277t* (pink). The empty vector control (c-*Msmeg_1817*_tetON pMV361, grey) and the wild type *M. smegmatis* mc^2155 (black) are shown as a reference. Constitutive recombinant expression of both *Msmeg_1817* and *Rv3277t* rescues growth inhibition during silencing of the endogenous *Msmeg_1817* gene in the absence of ATc in Middlebrook 7H9 liquid medium. Data in **b**, **d, f** and **g** are shown as mean values (in = 3) ± SEM.

Rv3277t from *M. tuberculosis* and its orthologue MSMEG_1817 from *M. smegmatis* are functionally homologous proteins.

Next, to study the effect of gene silencing in non-tuberculous mycobacteria, we analyzed the orthologue of *Rv3277t* in *M. smegmatis*, which is a fast-growing, non-pathogenic bacterium commonly used as a surrogate to study aspects of biochemistry and physiology of *M. tuberculosis*. The annotated protein MSMEG_1817 exhibits 80% identity on a 223 amino acid overlap to Rv3277t. To fine-tune gene expression in *M. smegmatis*, we generated the c-*Msmeg_1817*_tetON mutant analogous to the c-*Rv3277t*_tetOFF mutant by inserting a promoter cassette comprising four *tet* operator sites upstream of the annotated start codon (Suppl. Fig. 4) and expressing a plasmid-borne tet repressor (TetR) that binds to the operator site in absence of the inducer ATc. The conditional *M. smegmatis* c-*Msmeg_1817*_tetON mutant showed strict ATc-dependent growth, indicating that the studied gene is also essential for in vitro growth of *M. smegmatis* (Fig. 2e + f). In order to unambiguously attribute the observed phenotype to the regulation of the target gene and to rule out possible polar effects that might result from integration of a synthetic promotor cassette, the c-*Msmeg_1817*_tetON mutant was genetically complemented. ATc-dependency of growth of the c-*Msmeg_1817*_tetON mutant was rescued by constitutive expression of both *MSMEG_1817*

and *Rv3277t* (Fig. 2g). This demonstrated the absence of any polar effects relevant to the observed silencing phenotype and the functional homology between MSMEG_1817 and Rv3277t, allowing inferences to *M. tuberculosis* from studying the protein in *M. smegmatis* [27].

Silencing of *Rv3277t* in *M. tuberculosis* results in altered cell morphology.

In order to functionally characterize Rv3277t, we established conditions establishing partial silencing of *Rv3277t* in *M. tuberculosis* H37Rv during growth in liquid culture to allow comparative phenotypic characterization of wild type-like fully induced and of stressed cells (Fig. 3a). Scanning electron microscopy (SEM) revealed a slightly altered morphology in partially silenced cells with rougher and more irregular surfaces and less rounded, flattened poles (Fig. 3b). Occasionally, one to two blebs were observed emanating from the cell surface in partially silenced cells (Fig 3c), which is somewhat reminiscent to morphological alterations occurring in mycobacterial cells treated with isoniazid [28]. Although the observed alterations in partially silenced cells were rather subtle and no substantial cell lysis was observed, these findings support a potential role for Rv3277t in mycobacterial cell wall biogenesis.

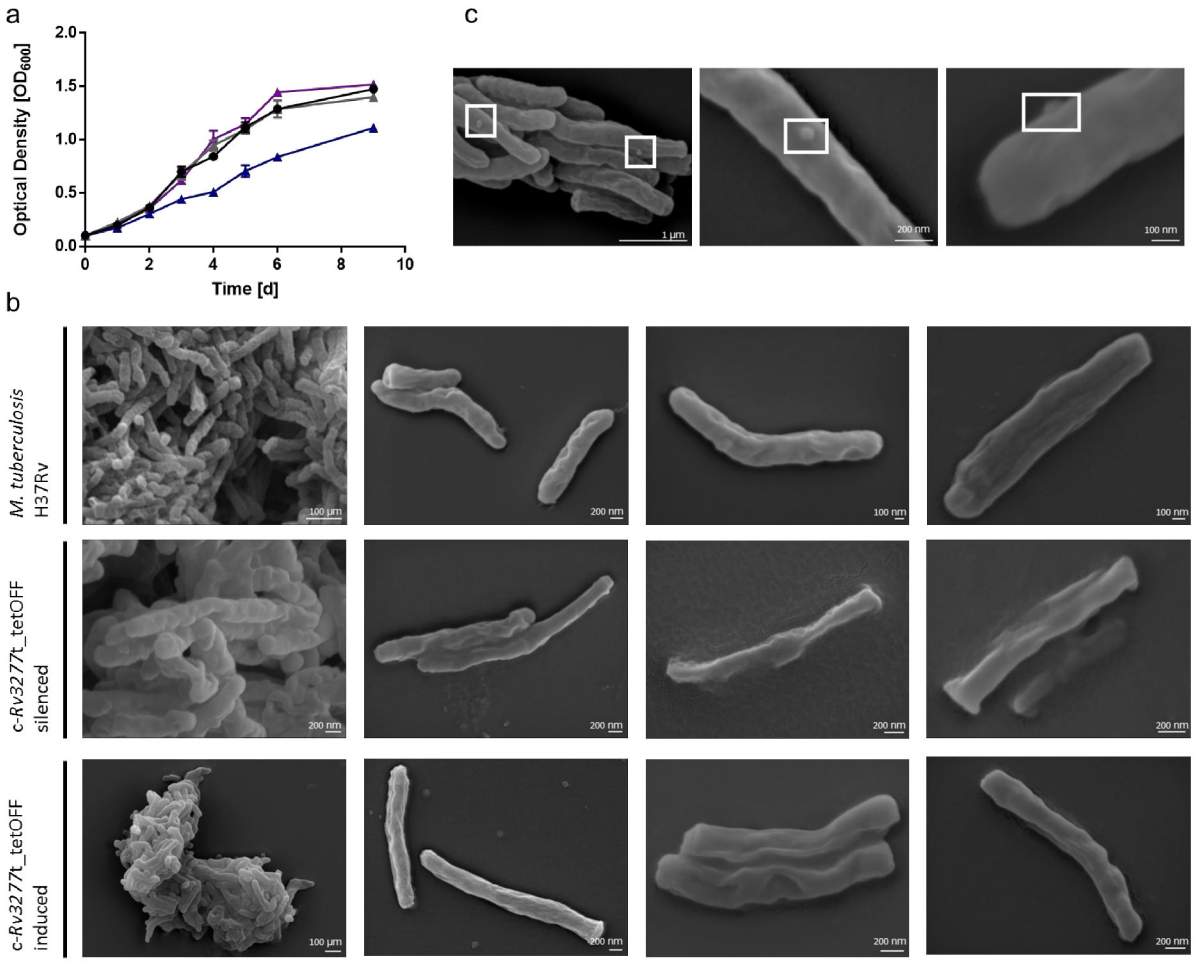


Fig. 3: Partial silencing of *Rv3277t* in *M. tuberculosis* results in altered cellular morphology. a, Growth curves in Middlebrook 7H9 medium of partially silenced (blue) and induced (purple) cells of the *M. tuberculosis* c-*Rv3277t*_tetOFF mutant (pink) compared to the empty vector control strain (c-*Rv3277t* pMV261, grey) and to the *M. tuberculosis* H37Rv wild type (black) strain. For full induction of the target gene, cells of the *M. tuberculosis* c-*Rv3277t*_tetOFF mutant were grown in medium without ATc, whereas 0.3 µg/mL ATc was added for partial silencing. Data represent mean values (n = 3) ± SEM. **b,** Scanning electron microscopy images of partially silenced and induced cells of the *M. tuberculosis* c-*Rv3277t*_tetOFF mutant cultivated in liquid culture compared to wild type. Wild type cells and fully induced cells of the conditional mutant were cultivated for 14 d in medium without ATc, whereas 0.3 µg/mL ATc was added for partial silencing of the conditional mutant. Size of scale bars as indicated. **c,** Scanning electron microscopy images showing occasional blebs forming at the surface of partially silenced cells of the *M. tuberculosis* c-*Rv3277t*_tetOFF mutant**.** Size of scale bars as indicated.

Silencing of *Rv3277t* in *M. tuberculosis* and *MSMEG_1817* in *M. smegmatis* leads to global decrease in mannosylated cell-surface glycoconjugates.

Based on the demonstrated impact on mycobacterial cell morphology and its conserved essential function in diverse mycobacteria, we hypothesized that Rv3277t is involved in biosynthesis of one or more core mycobacterial cell wall components that are crucial for viability in vitro. Previous studies implicated a role of the related GtrA-like protein Rv3789 in arabinogalactan formation and suggested it might work as a non-essential, redundant floppase involved in translocation of DPA or DPR [16]. Thus, we speculated that Rv3277t could potentially function as a floppase for DPR/DPA or for other decaprenol phosphate-linked carbohydrates as a precursor for glycoconjugate formation. In addition to arabinosylated cell wall polymers, the mycobacterial cell wall is characterized by high abundance of various mannosylated glycoconjugates such as PIM, LM, and LAM, the biosynthesis of which involves mannosyltransferases that require DPM as an activated donor substrate. Since several genes involved in biosynthesis of PIM, LM, and LAM are essential [25, 26], we concluded that the unknown translocase mediating flopping of DPM needs to be likewise crucial for the viability of mycobacteria, and we hypothesized that Rv3277t could have this essential flopping activity.

To study the potential role in mannosylated glycoconjugate biosynthesis, we first analyzed lysates generated from partially silenced cells of the *M. smegmatis* c-*Msmeg_1817*_tetON mutant. Glycoconjugates were separated by SDS-PAGE and probed employing the monoclonal antibodies F30-5, which is specific for the Ara_6 terminal motif of LAM [29, 30], and F183-24, which recognizes alpha-1,2-linked mannosyl residues as present

in the mannose cap on higher-order PIMs (Ac_1/Ac_2PIM_6) [30, 31]. Immunoblots employing the anti-LAM and anti-PIM_6 antibodies revealed that cellular levels of both LAM (Fig. 4a, top panel) and PIM_6 (Fig. 4a, bottom panel) were drastically reduced in partially silenced cells of the *M. smegmatis* c-*Msmeg_1817*_tetON mutant compared to fully induced cells, which exhibited wild type levels of these glycoconjugates. Genetic complementation of the c-*Msmeg_1817*_tetON mutant by constitutive expression of either *MSMEG_1817* or *Rv3277t* completely reversed the ATc-dependent silencing effect, unambiguously confirming that the observed phenotype was only attributed to the regulation of the target gene and ruling out relevant polar effects (Fig. 4c). Next, we confirmed the observed phenotypes using complementary methods. Separation by SDS-PAGE of polar lipids extracted from ^{14}C-glucose labeled partially silenced and fully induced cells of the c-*Msmeg*_1817_tetON mutant followed by autoradiographic analysis not only confirmed strongly downregulated levels of LAM in partially silenced cells but also demonstrated a concomitant dramatic decrease of the precursor LM (Fig. 4b). Furthermore, 2D-TLC analysis of the polar lipids fraction extracted from ^{14}C-glucose labeled partially silenced and fully induced cells of the c-*Msmeg*_1817_tetON mutant followed by autoradiographic analysis corroborated the depletion of higher-order PIMs (Ac_1/Ac_2PIM_6), while levels of lower-order PIMs (Ac_1/Ac_2PIM_2) as well as of the precursor phosphatidyl-*myo*-inositol remained unaltered (Fig. 4d).

To corroborate congruent phenotypes in slow-growing mycobacteria, we assessed levels of LAM and PIM_6 in fully induced and partially silenced cells of the *M. tuberculosis* c-*Rv3277t*_tetOFF mutant by immunoblotting. Similar to *M. smegmatis*, we found strongly reduced levels of both mannosylated glycoconjugates in partially silenced cells. (Fig. 4e). To confirm these findings, we performed MALDI-MS analyses of polar lipid fractions isolated from fully induced and partially silenced cells of the *M. tuberculosis* c-*Rv3277t*_tetOFF mutant. In agreement with our observations in *M. smegmatis*, silencing of *Rv3277t* resulted in substantial reduction of mono- and diacylated PIM_6, while levels of lower-order PIM_2 species (PIM_2, Ac_1PIM_2, Ac_2PIM_2) as well as PI remained unaltered (Fig. 4e). Accumulation of the precursors for all higher forms of mannosylated cell wall glycoconjugates, Ac_1PIM_4 and Ac_2PIM_4, was not detectable. However, we observed some accumulation of Ac_1PIM_5 in partially silenced cells, which was not present in fully induced cells (Fig. 4f).

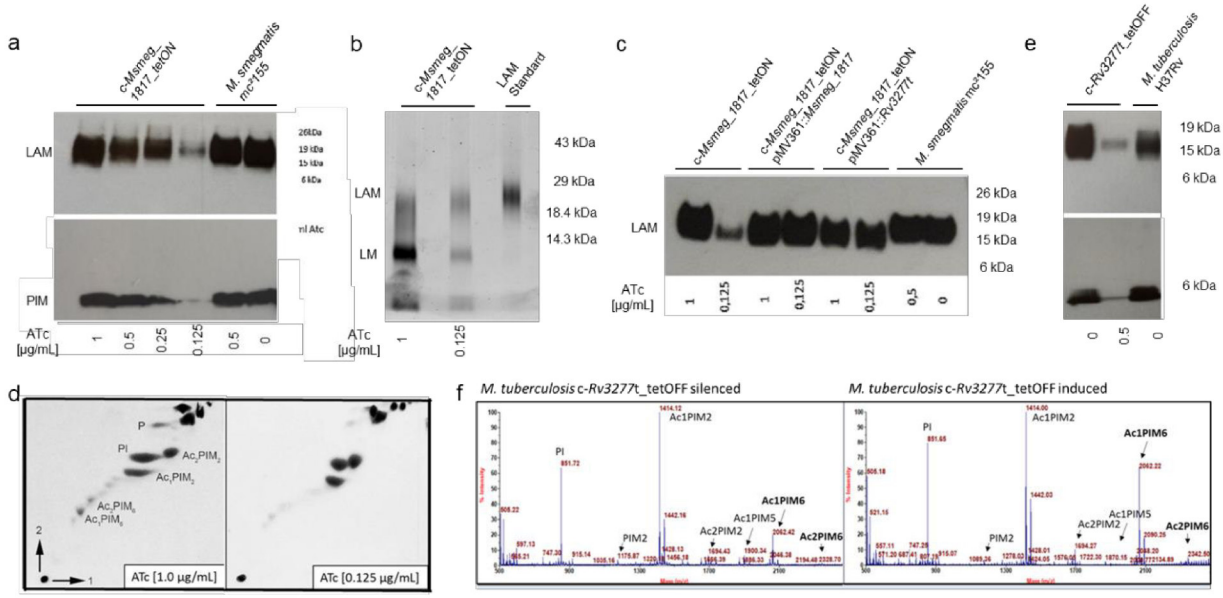


Fig. 4: Silencing of *Rv3277t* in *M. tuberculosis* and of *MSMEG_1817* in *M. smegmatis* blocks biosynthesis of the higher order mannosylated cell wall glycoconjugates PIM_6, LM and LAM. a, SDS-PAGE immunoblot analyses of lysates of fully induced (at 1 µg/mL ATc) and partially silenced (at 0.5, 0.25 and 0.125 µg/mL ATc) cells of the c-*Msmeg_1817*_tetON mutant compared to wild type *M. smegmatis* mc^2155 (left panel), and fully induced (at 0 µg/mL ATc) and partially silenced (at 0.5 µg/mL ATc) cells of the *M. tuberculosis* c-*Rv3277t*_tetOFF mutant compared to wild type *M. tuberculosis* H37Rv (right panel), employing the LAM-specific monoclonal antibody F30-5 (upper panel) and the PIM_6-specific antibody F183-24 (lower panel). Experiments were repeated once with similar results (Suppl. Fig. 5a-c). **b**, SDS-PAGE autoradiographic analysis of the LM/LAM fraction prepared from ^{14}C-glucose labeled fully induced (at 1 µg/mL ATc) and partially silenced (at 0.125 µg/mL ATc) cells of the *M. smegmatis* c-*Msmeg_1817*_tetON mutant. A full, uncut blot photo can be found in Suppl. Fig. 5d. **c,** Genetic complementation of the *M. smegmatis* c-*Msmeg_1817*_tetON mutant by constitutive expression of a plasmid-borne copy of *Msmeg_1817* or *Rv3277t* restores LAM biosynthesis. SDS-PAGE immunoblot analyses of lysates of fully induced (at 1 µg/mL ATc) and partially silenced (at 0.125 µg/mL ATc) cells of the c-*Msmeg_1817*_tetON mutant compared to wild type *M. smegmatis* mc^2155, employing the LAM-specific monoclonal antibody F30-5. **d,** 2D-TLC autoradiographic analysis of polar lipid fractions isolated from ^{14}C-glucose labeled fully induced (at 1 µg/mL ATc) and partially silenced (at 0.125 µg/mL ATc) cells of the *M. smegmatis* c-*Msmeg_1817*_tetON mutant. **e,** SDS-PAGE immunoblot analyses of lysates of fully induced (at 0 µg/mL ATc) and partially silenced (at 0.5 µg/mL ATc) cells of the *M. tuberculosis* c-*Rv3277t*_tetOFF mutant compared to wild type *M. tuberculosis* H37Rv employing the LAM-specific monoclonal antibody F30-5 (upper panel) and the PIM_6-specific antibody F183-24 (lower panel). **f,** MALDI-ToF MS analysis of lipids extracted from fully

induced (at 0 µg/mL ATc) and partially silenced (at 0.5 µg/mL ATc) cells of the *M. tuberculosis* c-*Rv3277t*_tetOFF mutant. The experiment was repeated once with similar results (Suppl. Fig. 5e). P, phospholipids; PI, phosphatidyl-*myo*-inositol; PIM_2, phosphatidylinositol dimannoside; Ac_1PIM_2, monoacylphosphatidylinositol dimannoside; Ac_2PIM_2, diacylphosphatidylinositol dimannoside; Ac_1PIM_6, monoacylphosphatidylinositol hexamannoside; Ac_2PIM_6, diacylphosphatidylinositol hexamannoside.

Impact of MSMEG_1817 on the arabinogalactan composition in *M. smegmatis*

The phenotypes described above resulting from silencing of *Rv3277t* and *Msmeg_1817* in *M. tuberculosis* and *M. smegmatis*, respectively (i.e., global downregulation of all higher forms of mannosylated cell wall glycoconjugates (PIM_6, LM and LAM), which all arise from Ac_1PIM_4 or Ac_2PIM_4,), are compatible with two alternative scenarios. Since all these glycoconjugates are formed via extension of the precursors Ac_1PIM_4 or Ac_2PIM_4 by various mannosyltransferases at the periplasmic site of the cytoplasm membrane using DPM as a substrate, the GtrA-like proteins could either mediate flopping of DPM from the cytoplasmic to the periplasmic face of the membrane or might be somehow required for formation of DPM at the cytoplasmic leaflet. Alternatively, the GtrA-like proteins might rather be required for flopping of the precursors Ac_1PIM_4 / Ac_2PIM_4 from the cytoplasmic to the periplasmic face of the membrane or might be involved in their synthesis from Ac_1PIM_2, / Ac_2PIM_2, at the cytoplasmic leaflet.

To gain hints regarding the specificity of the GtrA-like proteins Rv3277t and MSMEG_1817 in cell wall glycoconjugate formation, the LM/LAM as well as the arabinogalactan fractions were purified from fully induced and partially silenced cells of the *M. smegmatis* c-*Msmeg_1817*_tetON mutant and analyzed for their sugar monomer composition by GC. Consisting with its role in mannosylation, silencing of *MSMEG_1817* led to a gradual increase in the arabinose-to-mannose ratio of LM/LAM (Fig. 5). Surprisingly, however, silencing of *Msmeg_1817* concomitantly also led to a gradual decrease in the arabinose-to-galactose ratio of arabinogalactan (Fig. 5). Collectively, while predominantly controlling mannose incorporation, these findings show that the GtrA-like proteins *Rv3277t* and *MSMEG_1817* also contribute to the incorporation of arabinose into arabinogalactan. The most plausible explanation for the observed phenotypes is the function of Rv3277t and MSMEG_1817 as a floppase that preferentially translocates DPM but to some extent can also translocate DPR or DPA.

	ATc [µg/mL]	Ara : Man ratio of LM/LAM fraction	Ara : Gal ratio of AG fraction
M. smegmatis c-*Msmeg_1817*_tetON	1	2.5 : 1	2.2 : 1
	0.5	3.0 : 1	1.75 : 1
	0.125	3.4 : 1	1.5 : 1

Fig. 5: Impact of MSMEG_1817 on glycoconjugate monomer composition in *M. smegmatis*. Relative composition of LM/LAM and arabinogalactan purified from fully induced and partially silenced cells of the *M. smegmatis* c-*Msmeg_1817*_tetON mutant. Sugar composition was analyzed by GC and ratios were normalized to mannose or galactose content, respectively. Data represent single measurements. Ara, arabinose; Man, mannose; Gal, galactose; LM, lipomannan; LAM, lipoarabinomannan; AG, arabinogalactose.

Characterization of the floppase properties of Rv3277t

To gain insights into the possible flopping activity of Rv3277t, we performed 3D modeling employing the TopModel suite, revealing secondary structures that allow for transmembrane localization [32]. The secondary structure of Rv3277t was modeled based on homology with the template protein dolichyl phosphate mannose synthase (DPMS) of *Pyrococcus furiosus*, which is involved in N-glycosylation of proteins in this archaeon [33]. The cellular location was predicted by the PPM server [34]. Rv3277t is divided into an N-terminal transmembrane domain (TM) consisting of four α-helices and an unfolded C-terminus with periplasmic orientation (Fig. 6a). The TM domain resembles the C-terminal TM domain of the template DPMS in that the four α-helices are arranged as two dimers [33]. It is worth noting that we detected a change from hydrophobic to hydrophilic amino acids (shown as orange sticks) between the two dimeric structures, indicating altered physicochemical properties of the channel-like structure in Rv3277t (Fig. 6b, c). By transferring the co-crystallized detergent lauryl dimethylamine-N-oxide of the DPMS of *P. furiosus* to our model, we demonstrated a potential interaction between the hydrophilic lipid heads and the hydroxyl group of threonine, a hydrophilic amino acid not found in the TM of DPMS (Fig. 6c). In contrast to similar TM domains, the alignment of Rv3277t and the X-ray crystal structure of the DPMS template showed a significant difference between the proteins, as Rv3277t has a shorter N-terminus (Fig. 6d).

To assess whether the identified distinct hydrophilic amino acids in the channel-like structure (T35, K59, N141; Fig. 6c) are critical for the function of Rv3277t, a variant of the Rv3277t protein, where all three positions were mutated to alanine, was constitutively expressed from an integrative plasmid in the *M. tuberculosis* c-*Rv3277t*_tetOFF mutant. Functionality was then evaluated by the capability to suppress growth inhibition of the

M. tuberculosis c-*Rv3277t*_tetOFF mutant during ATc-induced silencing of the endogenous *Rv3277t* gene. While ectopic constitutive expression of wild type *Rv3277t* fully suppressed the growth defect during ATc-induced silencing of the endogenous gene, expression of the *Rv3277t*-T35A-K59A-N141A variant resulted in only partial suppression (Fig. 6e), suggesting that hydrophilic amino acids in the channel promote translocation of the probable substrates DPM and DPR/DPA likely by facilitating interaction with the polar head groups.

In order to further characterize the suggested floppase activity of Rv3277t, we sought to test whether heterologous expression of *bona-fide* lipid flippases can phenotypically complement the *M. tuberculosis* c-*Rv3277t*_tetOFF mutant under conditions when *Rv3277t* is silenced. Several studies demonstrated that flippases can exhibit relaxed substrate specificity, so that upon heterologous expression they can restore activity in absence of certain target flippases despite low sequence similarity between these translocases [21, 35-38]. Previously, it was shown that loss of Rv3789 function could be compensated for by heterologous expression of ArnE/F [21], which encode a flippase translocating undecaprenol phosphate-linked 4-amino-4-deoxy-L-arabinose involved in lipid A modification in *E. coli* [39]. Integrative plasmids for constitutive expression of the following genes were generated and electroporated into the *M. tuberculosis* c-*Rv3277t*_tetOFF mutant: the *arnE/F* genes; the *wzxE* gene from *E. coli* encoding an undecaprenol phosphate-linked oligosaccharide translocase involved in O-antigen assembly [40], and *ltaA* from *Staphylococcus aureus*, which represents a member of the major facilitator superfamily (MFS) and is a proton-coupled antiporter flippase for the anchor lipid-linked-disaccharide gentiobiosyl-diacylglycerol [41, 42]. Additionally, the *M. tuberculosis* gene *Rv3789*, encoding a putative GtrA-type multidrug resistance-like transporter, presumably acting as DPA-floppase was used for constitutive expression [21]. Phenotypic functional complementation was evaluated by the capability to suppress growth inhibition of the *M. tuberculosis* c-*Rv3277t*_tetOFF mutant during ATc-induced silencing of *Rv3277t* (Fig. 6f). While constitutive expression of the wild type *Rv3277t* completely suppressed the growth defect of the *M. tuberculosis* c-*Rv3277t*_tetOFF mutant, neither the expression of *Rv3789* nor of the other tested flippases significantly affected the ATc-dependent growth behavior of the mutant cells, resulting in phenotypes comparable to cells containing the empty vector. This strongly suggests that Rv3277c exhibits a narrow substrate specificity that does not overlap with the tested lipid translocases.

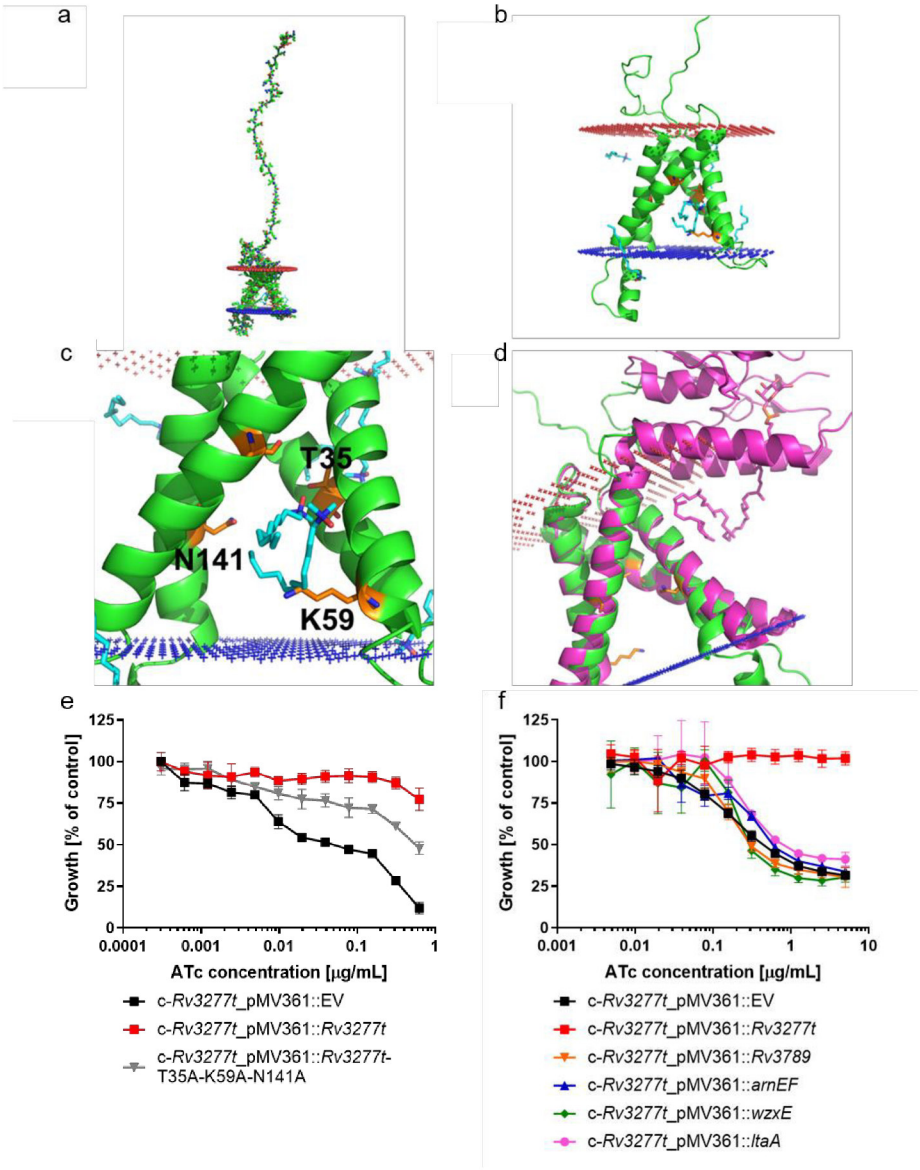


Fig. 6: Homology modeling of Rv3277t exhibits significant similarities to the DPMS of *P. furiosus*. Homology model of Rv3277t (green) made with the TopModel suite [32]. The red and blue dots represent the outer and inner membrane boundaries as predicted by the PPM server, respectively. The orange residues shown as sticks are residues that have undergone a change from hydrophobic to hydrophilic compared to the dolichyl phosphate mannose synthase (DPMS) template (PDB-ID: 5MLZ) **a**, Overview of the Rv3277t homology model. The C-terminus, pointing away from the membrane, is unfolded. **b**, The transmembrane part of the Rv3277t homology model. The co-crystallized detergent lauryl dimethylamine-*N*-oxide of the template DPMS (PDB-ID: 5MLZ) was transferred to the homology model and is shown as blue sticks. **c**, Interactions of the template lipids with the residues changed from hydrophobic to hydrophilic in Rv3277t. Here, the hydroxyl-group of T35 interacts directly with the hydrophilic lipid head. **d**, Alignment of the Rv3277t homology model with the X-ray crystal structure of the template DPMS (PDB-ID: 5MM1, magenta) with co-crystallized dolichyl phosphate mannose as shown in sticks interacting with the transmembrane region. The synthase part is not present

in Rv3277t as it possesses a short N-terminus. **e**, ATc-dependent growth inhibition of the *M. tuberculosis* c-*Rv3277t*_tetOFF mutant complemented with integrative plasmids constitutively expressing either the wild type version (pMV361::*Rv3277t*) or the mutated version of *Rv3277t* (pMV361::*Rv3277t*_T35A_K59A_N141A) compared to the empty vector control (pMV361::EV). Data are shown as mean values (in = 3) ± SEM. **f**, ATc-dependent growth inhibition of the *M. tuberculosis* c-*Rv3277t*_tetOFF mutant complemented with integrative plasmids constitutively expressing the *bona-fide* lipid flippases *wzxE* from *E. coli*, *arnE/F* from *E. coli*, *ltaA* from *Staphylococcus aureus* and the putative floppase *Rv3789* (from *M. tuberculosis*. Growth dependency was compared to the empty vector control (pMV361::EV) and the *M. tuberculosis* c-*Rv3277t*_tetOFF mutant complemented with the wild type version of *Rv3277t* (pMV361::*Rv3277t*). The average of two independent clones is shown ± SEM, with each measuring point of the two clones determined in triplicates.

Proteome stress response of *M. tuberculosis* during silencing of *Rv3277t*

To understand the importance of Rv3277t for viability of *M. tuberculosis* and to obtain insights into the cellular processes that are affected by silencing of *Rv3277t*, whole protein cell lysates were prepared from fully induced and partially silenced cells of the *M. tuberculosis* c-*Rv3277t*_tetOFF mutant to perform global proteome analysis and comparison of the elicited stress responses. Compared to fully induced conditions, partially silenced cells exhibited a significant change in the abundance of 242 proteins (Fig. 4a). As expected by the role of Rv3277t in cell wall glycoconjugate biosynthesis, gene ontology analysis using Fisher´s exact test revealed that mainly proteins related to lipid metabolism, plasma membrane and cell wall formation were affected in partially silenced cells (Fig. 4b). Within the overall global differential proteome profile, several overlapping, coordinated distinct stress responses could be identified (Supplementary Table 1). Strikingly, we found increased abundance of proteins involved in glycoconjugate mannosylation: the polyprenol-monophosphomannose synthase Ppm1 (Rv2051c) that transfers mannose from GDP-mannose to polyprenol phosphates at the cytosolic face of the membrane to yield the precursor DPM; the mannosyltransferase PimB (Rv2188c) that transfers mannose from GDP-mannose to PIM_1 to yield PIM_2; and the alpha-1-,6-mannopyranosyltransferase MptA (Rv2174) that mediates elongation of mono- and diacylated PIM_4 involved in biosynthesis of the alpha-1-,6-mannan core of LM biosynthesis employing DPM as donor substrate [20,21,36]. This upregulation of enzymes mediating early steps in mannosylation pathways might represent a directed compensatory stress response to counteract the observed global decrease in mannosylated cell surface glycoconjugates when Rv3277t is silenced. In agreement with adaptations aiming at compensating disturbed cell wall architecture, we also recognized increased protein abundances related to biosynthesis and

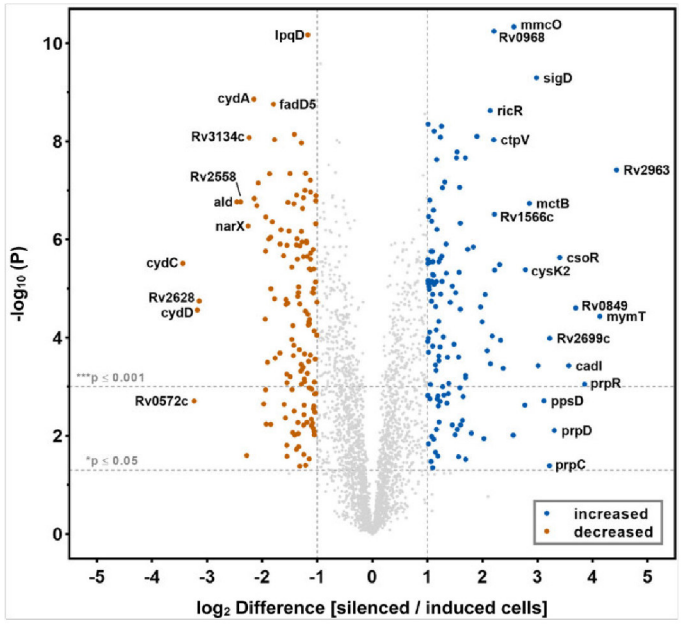
assembly of other cell wall constituents, such as phthioceroldimycocerosates (PDIM) and related glycolipids (PpsA – PpsD, FadD26, Mas, Pks15, Rv2958c), peptidoglycan (Rv3717, MurG, MurC, RipA) and mycolates (FbpC). Further, upregulation of cardiolipin synthase PgsA2 suggested stimulated phospholipid production. Another distinct response was differential abundance of a large group of proteins belonging to a regulon that is controlled by the copper-sensing transcriptional repressor CsoR and that responds to stress conferred by elevated copper concentrations: CadI (Rv2641), CsoR (Rv0967), CtpV (Rv0969), CysK2 (Rv0848), Rv0430, Rv0500A, Rv0849, Rv0968, Rv2963, RicR (Rv0190) [43]. Downregulation of Rv0500A and strong upregulation of all other mentioned proteins is congruent with cells experiencing copper stress. Additionally, we also found upregulation of proteins that do not belong to the copper stress regulon but are involved in copper homeostasis: MymT (Rv0186A), which is a copper-binding metallothionein, as well as the multicopper oxidase MmcO (Rv0846c) and the outer membrane copper transporter MctB (Rv1698) that are both required for copper resistance [44]. Thus, intriguingly, despite growing bacteria in regular culture medium, cells appear to experience copper stress when Rv3277t is silenced. It is not known whether there is indeed an increased intracellular concentration of copper in silenced cells, and if so, how the impaired cell surface mannosylation is mechanistically connected to this effect. It cannot be ruled out that the copper stress regulon is just misleadingly triggered at normal copper concentration, for example by an aberrant activation of the copper-sensing transcriptional repressor CsoR.

Among the strongest upregulated proteins were 2-methylcitrate synthase PrpC, 2-methylcitrate dehydratase PrpD and the cognate transcriptional regulator PrpR. Accumulation of methylcitrate cycle intermediates has recently been linked to drug tolerance [45]. Therefore, potential stimulation of the methylcitrate cycle by upregulation of PrpC and PrpD might represent an unspecific general stress adaptation.

Regarding processes that are negatively affected by silencing of Rv3277t, we observed lower abundance of several ribosomal proteins (RplQ, RpmC, RpsO, RplU, RpmA, RpmD, RpmF), suggesting global downregulation of protein biosynthesis. Likewise, downregulation of numerous transcriptional regulatory proteins and putative DNA-binding proteins suggests massive perturbation of transcriptional control (CspA, Rv3295, Rv0232, Rv3716c, Rv0081, DosR, CarD, Rv0500A, Rv3788, ArgR, Rv0386, Rv1404). A related coordinated stress response that could additionally affect the transcriptome is the down-regulation of several antitoxins (VapB29, MazE3, VapB38, VapB33, VapB47) that likely release their corresponding toxins, which putatively exhibit ribonuclease activity. The reduced level of the RNA polymerase-binding transcription factor CarD is particularly noteworthy since this protein is necessary for survival during exposure to several stressors such as oxidative stress, DNA damage, and nutrient limitation [46]. In the same direction, downregulation of several universal

stress proteins (Rv1996, Rv2005c, Rv2623, Rv2624c and Rv3134c) and of the chaperone HspX might indicate an impaired ability to adequately respond to various endogenous and exogenous noxes. Therefore, curtailing of different stress adaptation mechanisms may partly explain the lethal effect of Rv3277t silencing. Downregulation of the ribonucleoside-diphosphate reductase subunits NrdF2 and NrdB implies that DNA replication was dampened, while concurrently cell division appeared to be stimulated as presumed by upregulation of FtsK, SepF and FtsQ, which should actually increase the demand of deoxynucleotide triphosphates. Finally, among the strongest downregulated proteins were components of the electron transport chain or factors required for their assembly (NADH-quinone oxidoreductase subunit NuoE, integral membrane cytochrome D ubiquinol oxidase CydA, nitrate reductase-like protein NarX, and cytochrome assembly ABC transporter CydC-CydD), pointing toward impaired aerobic respiration and a curtailed capacity to generate ATP in silenced cells.

a



b

Type	Category	Size	Intersec.	Enrichment factor	FDR	p-value
Pathway	Lipid metabolism	61	48	1.668	1.67E-04	2.78E-07
GO	Peptidoglycan-based cell wall [GO:0009274]	354	209	1.251	6.94E-04	5.00E-07
GO	Cytosol [0005829]	343	192	1.186	5.48E-03	1.09E-04
GO	Plasma membrane [0005886]	821	453	1.169	3.17E-05	1.14E-08

c

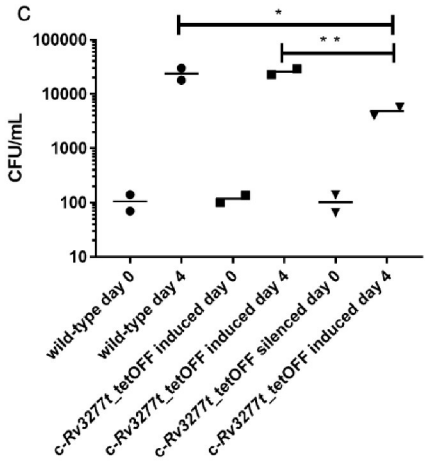


Fig. 7: Silencing *Rv3277t* leads to a drastic change in the proteome composition and impairs intracellular growth in *M. tuberculosis*. a, LC-MS/MS-based whole protein analysis of silenced cells of the *M. tuberculosis* H37Rv c-*Rv3277t*_tetOFF mutant compared to the corresponding fully induced cells. The volcano plot illustrates the log_2-fold change in abundance in silenced vs. induced cells (X-axis) and corresponding -log10 *p* values (Y-axis). Proteins complying with the chosen threshold of significance and showing a log_2-fold

change ≥ 1 or ≤ -1 are marked in blue or orange, respectively. For silencing conditions, cells of the *M. tuberculosis* H37Rv c-*Rv3277t*_tetOFF mutant were cultivated in presence of 0.3 µg/mL ATc until they reached the mid-logarithmic phase ($OD_{600\ nm}$ of 0.6 - 0.8). Quantification was done via label free quantification (LFQ) of five replicates per sample group. To identify statistically significant hits from the analysis, $P \leq 0.05$ (Student's T-test; permutation-based FDR with 250 randomizations and FDR = 0.01) was applied. See Supplementary Fig. 6 for LC-MS/MS-based control proteomic profiles. **b**, Gene ontology analysis reveals an enrichment of proteins related to plasma membrane and cell envelope formation in silenced cells of the *M. tuberculosis* H37Rv c-*Rv3277t*_tetOFF mutant. Fischer Exact Test was performed with significant protein groups from silenced vs induced cells t-test as input (Benjamini-Hochberg FDR, Threshold value 0.05). **c**, THP-1 derived macrophages were infected with *M. tuberculosis* H37Rv wild type and fully induced (at 0 µg/mL ATc) cells of the *M. tuberculosis* c-*Rv3277t*_tetOFF mutant at a multiplicity of infection = 3. After washing to remove non-internalized cells, mutant cells were incubated either in the presence of 0.3 µg/mL ATc (silenced) or in absence of ATc (induced). Macrophages were lysed after 4 days of incubation, and lysates were subjected to CFU plating to count viable cells. CFU/mL were calculated after 2 weeks of incubation (n=2). *, $p < 0.05$; **, $p < 0.01$.

Silencing of *Rv3277t* strongly impairs intracellular growth in macrophages

The data so far have demonstrated the essentiality of *Rv3277t* for in vitro growth of *M. tuberculosis* in axenic culture. During infection, however, *M. tuberculosis* largely resides and replicates inside macrophages. Thus, we were interested in studying the effect of silencing on intracellular growth capacity in a human macrophage infection model. For this, the *Rv3277t* gene was silenced by addition of ATc after infection of differentiated human THP-1 cells with the *M. tuberculosis* c-*Rv3277t*_tetOFF mutant. Compared to wild type and fully induced cells of the *M. tuberculosis* c-*Rv3277t*_tetOFF mutant, silenced cells showed a significantly reduced intracellular growth (Fig. 7c), demonstrating that Rv3277t is essential in this infection model.

Discussion

For decades, the cell wall has been a useful target for antitubercular therapy. Remarkably, the biosynthetic pathways responsible for the assembly of various cell wall constituents have not been completely deciphered, leaving several potential drug targets undiscovered. In particular, it remained unknown until now how the undecaprenyl phosphate-linked molecules DPM and DPA are translocated from the cytosolic to the periplasmic face of the membrane to serve as activated substrates for membrane-bound mannosyl- and arabinosyltransferases. This study provides evidence for the role of Rv3277t in the translocation of DPM, a crucial step in the biosynthesis of cell wall-associated mannosylated glycoconjugates. Although direct proof of the enzymatic activity of Rv3277t is pending, several lines of evidence strongly support the conclusion that Rv3277t is a DPM floppase and might also contribute to flopping of DPA or DPR. The global impact on cell surface mannosylation affecting glycoconjugates all relying on DPM as the substrate is in agreement with Rv3277t representing a DPM floppase. In contrast, the DPM-independent formation of lower-order PIM species continued normally during silencing of Rv3277t. The inability of *bona fide* lipid translocases exhibiting relaxed substrate specificity to compensate for the loss of Rv3277t suggests a distinct substrate specificity.

In eukaryotes, a similar bipartite mechanism is known in the context of N-glycosylation of proteins, which occurs in the endoplasmic reticulum (ER) [47]. Here, the dolichylphosphate β-D-mannose synthase (DPMS) utilizes GDP-Man to transfer α-D-mannose to dolichylphosphate, resulting in dolichylphosphate β-D-mannose (Dol-*P*-Man), the eukaryotic pendant to DPM located in the cytoplasm [48]. After translocation of Dol-*P*-Man from the cytosol to the lumen of the ER, the glycosylation of proteins can occur [49]. Interestingly, the state of knowledge is also incomplete in eukaryotes because it is unknown which proteins conduct the translocation. However, it has been shown that in vitro DPMS is required for the transfer through membranes [50]. Intriguingly, our molecular modeling revealed a high similarity between Rv3277t and the transmembrane domain of DPMS from *P. furiosus*. Since Rv3277t lacks the cytosolic N-terminal synthase part of the DPMS, we assumed that Rv3277t is not involved in the synthesis of DPM (which is known to be mediated by Ppm1 in *M. tuberculosis*) but rather allows translocation due to the hydrophilic amino acids within the TM domain. This hypothesis is supported by Reichenbach, who suggested that the transmembrane domain of *P. furiosus* DPMS may have a function independent of the N-terminal catalytic domain, since a truncated version lacking the transmembrane domain was reported enzymatically functional [33]. We revealed potential interactions of co-crystallized template lipids with certain hydrophilic amino acids that resemble a “polar pocket” between α-helices in the transmembrane domain of Rv3277t, indicating a binding site for the mannosyl head group [51]. Mutating this polar pocket to a hydrophobic character resulted in strong

impairment of Rv3277t function, implying this feature is important in facilitating reorientation of the polar mannosyl head group across the membrane.

Based on findings with other floppases [52-54], inhibition of Rv3277t activity was expected to result in some accumulation of DPM accumulation at the cytosolic face of the membrane. However, none was detectable in partially silenced cells of the *M. tuberculosis* c-*Rv3277t*_tetOFF mutant or the *M. smegmatis* c-*Msmeg_1817*_tetON mutant. We recognized upregulation of the DPM-producing enzyme Ppm1 and of the mannosyltransferases PimB and MptA as part of the stress response in partially silenced cells of the *M. tuberculosis* c-*Rv3277t*_tetOFF mutant, implying potential regulatory feedback mechanisms controlling flux of shared intermediates through the mannosylated glycoconjugate biosynthetic pathways. It can be speculated that DPM sequestration is prevented by hydrolysis to make the lipid anchor undecaprenyl phosphate available for other pathways. Accumulation of Ac_1/Ac_2PIM_4 has been reported in *M. tuberculosis* to result from inactivation of mannosyltransferase PimE [31]. We did not observe Ac_1/Ac_2PIM_4 accumulation in partially silenced cells of the *M. tuberculosis* c-*Rv3277t*_tetOFF mutant. Instead, we detected some accumulation of Ac_1PIM_5. Ac_1PIM_5 is an intermediate of the extension of Ac_1PIM_4 to Ac_1PIM_6 by the mannopyranosyltransferase PimE (Rv1159) at the periplasmic side of the membrane employing DPM as the donor substrate and does not accumulate in wild type cells due to its rapid further extension [55]. Possibly, when DPM-flopping is blocked by silencing of Rv3277t, there is a build-up of the precursor Ac_1PIM_4, which might be further extended by an unknown mannosyltransferase that uses GDP-mannose as the donor substrate at the cytosolic side of the membrane, resulting in Ac_1PIM_5 as a shunt product.

Our data support a floppase activity of Rv3277t not only with respect to DPM. The specific reduction in the arabinose content of mycobacterial AG in partially silenced conditional mutants provides evidence for an additional role in DPA reorientation. There are conflicting reports concerning the possible function of the GtrA-like protein Rv3789 in AG biosynthesis, either as a translocase mediating flopping of DPA or rather indirectly as a membrane anchor for recruitment of DPA-dependent arabinosyltransferases [22]. Further, it has not been clearly established yet whether precursor translocation occurs at the level of DPA or DPR [16]. The presence of Rv3277t as a floppase contributing to reorientation of DPR or DPA could explain why Rv3789 is important but not strictly required for AG production. However, overall similarity between Rv3789 and Rv3277t is low, and while both proteins harbor a GtrA-typic transmembrane domain comprising four helices, Rv3277t is substantially larger than Rv3789 (223 vs. 121 amino acids) and contains a short N-terminal as well as a larger C-terminal extension. This implies that Rv3277t is not simply redundant to Rv3789 but has additional essential functions.

Consistent with its central role in mannosylated cell wall glycoconjugate biosynthesis, partial silencing of Rv3277t resulted in a compensatory upregulation of Ppm1 that produces the central precursor DPM and of enzymes mediating early steps in PIM and LM synthesis. In addition, upregulation of pathways for the unrelated cell wall components PDIM, mycolates and peptidoglycan potentially also aim at compensating disturbed cell wall architecture. In addition, unexpected pleiotropic effects were observed, including inhibition of protein biosynthesis and massive disturbance of transcriptional control. Some of the responses appear to represent misguided mechanisms that likely potentiate stress to the cells, such as downregulation of different general stress adaptation mechanisms and downregulation of dNTP biosynthesis while concomitantly forcing cell division. One of the most surprising stress responses was induction of the copper responsive regulon and further proteins involved in copper homeostasis. It can be speculated that the impaired cell wall structure might promote uptake of copper from the medium leading to toxic concentrations. Downregulation of central components of the electron transport chain might promote this process. At least, a curtailed capacity to generate ATP via aerobic respiration likely contributes to an inability to properly cope with the multitude of stressors emerging in silenced cells. The parallel corruption of several elementary cellular processes plausibly explains the bactericidal effect of Rv3277t silencing. In addition to their crucial role for viability of mycobacteria, however, the mannosylated glycoconjugates that are produced in a Rv3277t-dependent manner are also involved in M. tuberculosis pathogenicity by interaction with various host cell receptors or by controlling cellular innate immune responses. Thus, the observed intracellular growth defect in infected human macrophages is likely driven both by reduced viability and reduced ability to manipulate the host cell. Further work is necessary to dissect to which extent each aspect contributes to the in vivo phenotype. Regardless, its dual influence on bacterial viability and on pathogen-host interaction suggests that Rv3277t is a particularly interesting target for the development of novel antitubercular chemotherapeutics.

Methods

Bacterial strains and growth conditions

Strains used in this study are listed in Supplementary Table 2. Mycobacterial cultures were grown aerobically at 37 °C shaking in Middlebrook 7H9 broth or statically on Middlebrook 7H10 agar each supplemented with 10% ADS (0.8% NaCl, 5% BSA, 2% dextrose), 0.5% glycerol, 0.05% tyloxapol, and appropriate antibiotics (50 µg/mL hygromycin, 20 µg/mL kanamycin, 10 µg/mL apramycin). *E. coli* was grown in lysogeny broth (LB)-medium or LB agar containing appropriate antibiotics (150 µg/mL hygromycin, 40 µg/mL kanamycin, 100 µg/mL ampicillin 20 µg/mL apramycin).

Generation of conditional mutants

The generation of conditional mycobacterial mutants was conducted employing a modified specialized transduction protocol [56, 57]. The flanking regions of the gene of interest were amplified *via* PCR employing oligonucleotide primers listed in Supplementary Table 3. Subsequently, the flanking regions were treated with *Van91*I and ligated with *Van91*I-digested pcRv1327c-4XtetO vector arms, resulting in integration of a synthetic gene cassette (hyg-Pmyc1-4XtetO) comprising a hygromycin resistance gene and the strong P*myc1* promoter from *M. smegmatis* engineered to contain four *tetO* operator sites immediately upstream of the start codon of the target gene. This yielded the allelic exchange knock-in plasmids listed in Supplementary Table 4. The resulting knock-in plasmids were linearized with *Pac*I and cloned into the temperature-sensitive TM4-derived phasmid phAE159, yielding the corresponding knock-in phasmids (Supplementary Table 3). Knock-in phages were produced by electroporating the phasmids into *M. smegmatis* at the permissive temperature of 30°C which was propagated in *M. smegmatis*. Targeted gene knock-in in *M. tuberculosis* and *M. smegmatis* was achieved by specialized transduction employing high-titer lysates of the temperature-sensitive mycobacteriophages at the non-permissive temperature of 37°C. Regulated gene expression was established by heterologous expression in the generated knock-in mutants of a plasmid-borne repressor (revTetR or TetR) that binds to the tetO sites either in presence (tetOFF) or absence (tetON) of anhydrotetracyclin (ATc), respectively.

Genetic complementation of conditional mycobacteria mutants

For complementation of the conditional *M. tuberculosis* c-*Rv3277t*_tetOFF mutant and the *M. smegmatis* c-*Msmeg_1817*_tetON mutant, various genes were amplified by PCR using oligonucleotide primers listed in Supplementary Table 3. PCR products were cloned using the restriction enzymes *Pac*I and *Hin*dIII into the single-copy integrative plasmid pMV361(Apra)-PacI, which is a derivative of pMV361(Kan) engineered to contain a unique *Pac*I restriction site and an apramycin resistance gene [57]. The resulting plasmids providing constitutive gene

expression from the HSP60 promoter were transformed by electroporation into the conditional mutants and selected on solid medium containing 10 µg/mL apramycin.

Growth determination

To evaluate growth of conditional mutants in liquid medium in dependence of ATc, mycobacteria were grown to $OD_{600\ nm}$ of approx. 0.6 - 0.8 and diluted to 10^6 CFU/mL in Middlebrook 7H9 medium. Subsequently, 50 µL of this suspension was added to each well of a 96-well round-bottom microtiter plate containing a twofold serial dilution of ATc in 50 µl Middlebrook 7H9 medium per well. The microtiter plate was incubated statically at 37 °C for 5 days. Growth was determined via resazurin reduction assay. For this, 10 µl of a 100 µg/mL resazurin solution was added to each well of the microtiter plate and incubated at RT for 24 h. After fixation of cells by adding 100 µL of a 10% formalin solution and incubation at RT for 30 min, the reduction of resazurin was detected by measuring fluorescence (535 nm excitation, 590 nm emission) employing a Tecan Infinite F200 Pro (TECAN, Männedorf, Switzerland). In another experiment, growth in liquid medium was measured in a kinetic manner by growing cultures until reaching mid-logarithmic phase ($OD_{600\ nm}$ of 0.6 – 0.8), after which they were diluted to an $OD_{600\ nm}$ of 0.01 using Middlebrook 7H9 medium. To induce sub-lethal silencing of the cells of the *M. tuberculosis* c-*Rv3277t*_tetOFF mutant, 0.3 µg/mL ATc was added to the culture media. Subsequently, the $OD_{600\ nm}$ of cultures were recorded daily over several days to obtain growth curves.

For quantification of ATc-dependent growth on solid medium, a culture of mycobacteria was grown to $OD_{600\ nm}$ of approx. 0.6 - 0.8 before 10 µl of serially diluted aliquots of the cell suspension was plated out on Middlebrook 7H10 agar supplemented with increasing concentrations of ATc. Colony forming units were counted after 3 or 14 days for *M. smegmatis* and *M. tuberculosis*, respectively.

Scanning electron microscopy

Liquid cultures of *M. tuberculosis* H37Rv wild type and of the c-*Rv3277t*_tetOFF mutant were grown to an $OD_{600\ nm}$ of 0.5. The conditional c-*Rv3277t*_tetOFF mutant was silenced by addition of 0.3 µg/mL ATc, whereas the induced cells were grown without the addition of ATc. After 14 days of incubation, the cells were centrifuged at 4,000 rpm for 5 minutes and fixed at a final concentration of 2% paraformaldehyde and 2.5% glutaraldehyde for 1 hour. The fixed cells were then washed twice with PBS followed by washing with PBS containing 0.025% tyloxapol. Cells were dehydrated stepwise with a range of ethanol concentrations (50%, 70%, 80%, 90%, 96%, and 100%) for 5 minutes each. Images were captured using a scanning electron microscope ZEISS Supra 55VP.

Preparation of protein lysates

Mycobacteria were grown in 20 mL Middlebrook 7H9 medium supplemented with the indicated concentrations of ATc. After reaching an $OD_{600\,nm}$ of approx. 0.6 - 0.8, the cultures were washed three times with ice-cold PBS and resuspended in 2 mL PBS. Subsequently, the cells were transferred to screw-cap tubes containing silica-zirconium beads and lysed by bead beating at 50 Hz for 3 x 3 min. Next, 200 µl of 10% SDS were added followed by 30 min incubation at 4 °C. Finally, the samples were centrifuged and the supernatant was sterile filtered three times employing 0.2 µm polyethersulfone filters (Sarstedt AG & Co.KG, Nümbrecht, Germany).

Global proteome profiling by LC-MS/MS analysis

Following in-solution digestion (ISD) of protein lysates, peptides were subjected to desalting using homemade C18 StageTips [58]. In brief, the peptide solution was passed through a StageTip that was pre-conditioned with methanol (MeOH) and equilibrated with 0.5% formic acid (FA). The immobilized peptides were subsequently washed twice with 0.5% (v/v) FA and then eluted from the StageTip using a solution containing 80% (v/v) acetonitrile (ACN) and 0.5% (v/v) FA. The eluted peptides were dried using a vacuum concentrator (Eppendorf) before being resuspended in 10 µl of 0.1% (v/v) FA in preparation for LC-MS/MS analysis.

LC–MS/MS experiments were conducted on an Orbitrap Elite instrument from Fischer Scientific, coupled with an EASY-nLC 1000 liquid chromatography system also from Fischer Scientific, operating in one-column mode [59]. The analytical column was a fused silica capillary with an integrated PicoFrit emitter, packed in-house with Reprosil-Pur 120 C18-AQ 1.9 µm resin from Dr. Maisch. The analytical column was enclosed by a Sonation column oven and connected to a nanospray flex ion source from Thermo. During data acquisition, the column oven temperature was set at 45 °C. The LC was equipped with two mobile phases: solvent A (0.1% formic acid in water) and solvent B (0.1% formic acid in acetonitrile). Both solvents were of UPLC grade from Sigma. Peptides were loaded directly onto the analytical column at a maximum flow rate that did not exceed the set pressure limit of 980 bar, usually around 0.5–0.6 µl/min. Peptides were separated on the analytical column using a 140 min gradient of solvent A and solvent B at a flow rate of 300 nL/min, with a gradient starting at 7% B, followed by 7 to 35% B for 120 min, 35 to 100% B for 10 min, and 100% B for 10 min. The mass spectrometer was operated using Xcalibur software, version 2.2 SP1.48 from Thermo Fischer Scientific in positive ion mode. Precursor ion scanning was performed in the Orbitrap analyzer in the scan range of m/z 300–1,800, with a resolution of 60,000 and the internal lock mass option turned on. The lock mass used was 445.120025 m/z, polysiloxane [60]. Product ion spectra were recorded in a data-dependent fashion in the ion trap in a variable scan range and rapid scan rate. The spray voltage was set at 1.8 kV. Peptides were analyzed using a repeating cycle consisting of a full precursor ion scan followed by 15 product ion scans, where

peptides are isolated based on their intensity in the full survey scan for tandem mass spectrum generation that permits peptide sequencing and identification. Collision-induced dissociation energy was set at 35% for MS2 spectra generation. During MS2 data acquisition, dynamic ion exclusion was set to 120 s, with a maximum list of excluded ions consisting of 500 members and a repeat count of one. Ion injection time prediction, preview mode for the orbitrap, monoisotopic precursor selection, and charge state screening were enabled, with only charge states higher than 1 considered for fragmentation.

Immunoblot analysis of glycoconjugates

For the specific detection of the cell wall components, protein lysates were separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and transferred to a PVDF membrane by Western blot transfer according to standard protocols [61, 62]. The membrane was blocked with 10% skim milk for 30 min, washed once with TBS (50 mM Tris, 150 mM NaCl, pH 7.6) containing 0,05% Tween 80 (TBST), and then incubated with the murine monoclonal antibodies F30-5 or 183-24 in 2% skim milk at 4°C overnight as primary antibodies for detection of LAM and PIM_6, respectively [30]. Following washing of the membrane three times with TBST, an anti-mouse antibody conjugated to horseradish peroxidase (HRP) in 2% skim milk was used as secondary antibody. After 1 h incubation at room temperature, the membrane was washed three times with TBST, three times with TBS and once with H_2O. By adding the substrate solution (Lumiglo®Reagent, Cell signaling Technology, Denver, USA), the antibody-linked HRP allowed detection *via* the generation of luminescence. The luminescence was visualized on Amersham Hyperfilm™ MP using the Cawomat 2000 IR developer.

SDS-PAGE analysis of lipoglycans

Lipoglycans were extracted from 10 mL of cultures of fully induced or partially silenced cells of the *M. smegmatis* c-*Msmeg_1817*_tetON mutant grown with 10 µCi/mL of radiolabeled ^{14}C-D-glucose. The bacteria were harvested, dried and resuspended in water. Five consecutive extractions were performed with an equal volume of 50% (v/v) ethanol in water at 85°C for 6 hours. The supernatants were pooled, dried, and subjected to a phenol-H_2O treatment at 65°C. The aqueous phase containing the crude lipoglycans was dialyzed against water, dried, and the incorporation of ^{14}C-glucose was quantified by liquid scintillation. Equal amounts (50,000 cpm) were then loaded onto a 15% SDS-PAGE gel and separated electrophoretically. The visualization was performed autoradiographically by 48 h exposure of X-ray films.

Two-dimensional thin-layer chromatography

Cells of the *M. smegmatis* c-*Msmeg_1817*_tetON mutant were grown to an $OD_{600\ nm}$ of 0.8, and radiolabeled with 10 µCi/mL of ^{14}C-D-glucose during incubation for 4 hours at 37°C. For extraction of polar lipids, 10 mL of culture was harvested, washed twice with PBS and then extracted with 2 mL CH_3OH : 0.3% NaCl [63] (100:10, v/v) and 2 mL petroleum ether for 30 min. After centrifugation, the pellet was resuspended in 2.3 mL $CHCl_3$: CH_3OH : 0.3% NaCl [63] (90:100:30, v/v/v), stirred for one hour and centrifuged again. The pellet was mixed with 750 µl $CHCl_3$: CH_3OH : 0.3% NaCl [63] (50:100:40, v/v/v) for 30 min, centrifuged, and the supernatant was combined with the previous fraction. After repeating the last step, the combined supernatants were mixed with 1.3 mL of $CHCl_3$ and 1.3 mL of 0.3% NaCl [63], centrifuged, and the lower organic phase was isolated and dried. The dried extract of polar lipids was then resuspended in $CHCl_3$: CH_3OH (2:1, v/v), and the incorporation of ^{14}C-glucose was quantified by liquid scintillation using a 5% aliquot of the lipid extract in 5 mL EcoScint A (National Diagnostics). Equal amounts of the polar lipid extracts (50,000 cpm) were applied to Silica Gel 60 F254 (Merck) aluminum-backed TLC plates and run in the solvent system E for polar lipids: $CHCl_3$: CH_3OH : H_2O (60:30:6, v/v/v) in the first dimension and $CHCl_3$: CH_3CO_2H : CH_3OH : H_2O (40:25:3:6, v/v/v/v) in the second dimension. The visualization was performed autoradiographically by 48 h exposure of X-ray films.

MALDI-ToF mass spectroscopy

Cells of the conditional *M. tuberculosis* c-*Rv3277t*_tetOFF mutant were grown in 100 mL 7H9 medium in duplicates. Cultures were grown first under fully induced conditions to ca. late log phase to obtain biomass before splitting and continue culturing for further 4 days either with or without addition of 1 µg/mL ATc to silence gene expression. Next, the $OD_{600\ nm}$ was measured to allow normalization of the amount of lipid extract used for subsequent analyses. The cells were then harvested and washed twice with water, resuspended in 1 mL water and heat-inactivated in glass tubes by boiling for 30 min.

MALDI-ToF analyses were performed on an AB Sciex TOF/TOF 5800 mass spectrometer using the reflectron mode, as described elsewhere [64, 65]. Briefly, ionisation was accomplished through irradiation using an Nd:YAG laser (349 nm) operating with a pulse rate of 400 Hz. The laser intensity was set at 3500 with continuous stage motion utilised at a velocity of 600 µm/s. The matrix used was 2,5-dihydroxybenzoic acid (Sigma) at a concentration of 10 mg/mL in chloroform/methanol 9:1 or 1:1.

Extraction of AG and lipoglycan fractions and monomeric compositional analysis *via* GC/MS

Cells of the *M. smegmatis* c-*Msmeg_1817*_tetON mutant were grown at the indicated ATc concentrations, harvested, and frozen at -20°C. The thawed cells were then resuspended in PBS containing 2% Triton X-100 (pH 7.2), sonicated, and pelleted at 27,000 *g*. The cell wall was extracted three times with PBS containing 2% SDS at 95°C for one hour, then washed with water, 80% (v/v) acetone in water, and finally with acetone before lyophilization. Cells were hydrolyzed with 2 M trifluoroacetate, reduced with NaB_2H_4, and the alditols resulting from this treatment were per-*O*-acetylated. The per-*O*-acetylation was performed after mixing the lipoglycans with 0.5 mL DMSO containing 4.8 M dimethylsulfinyl carbanion. The mixture was stirred three times for one hour, mixed with 50 µL of iodomethane and stirred again for one hour. The mixture was diluted with one volume of water, dialyzed against water, and analyzed by GC/MS for alditol acetate sugar composition.

Molecular modeling

The protein sequence of Rv3277t was modeled using TopModel [32] and TopScore [66]. Using default program parameters, the comparative modeling includes template identification, sequence alignment, modeling, refinement, and scoring. This resulted in a homology model with a good TopScore of 0.307, where the unstructured C-terminus has a low confidence and the transmembrane region a high confidence.

Subsequently the orientation in the membrane and the membrane boundaries were predicted using the PPM server [34]. This resulted in a predicted membrane thickness of 29.4 ± 1.3 Å with a transfer free energy ΔG = -45.2 kcal mol^{-1} and a tilt angle of 2° ± 1°.

Macrophage infection assay

THP-1 cells were cultured in RPMI medium supplemented with 10% FBS in a humidified atmosphere with 5% CO_2 at 37 °C. Cell counting was performed using a haemocytometer, and 10^5 cells were added into each well of a 96-well flat-bottom polystyrene microtiter plate (Greiner) in a total volume of 100 µL. To differentiate THP-1 into adherent cells with macrophage-like characteristics, the medium was supplemented with 50 nM phorbol-12-myristate-13-acetate (PMA). Following overnight differentiation, the THP-1 derived macrophages were washed twice with PBS. For infection, cells of the conditional *M. tuberculosis* c-*Rv3277t*_tetOFF mutant grown under fully induced conditions were harvested, washed, and resuspended in RPMI supplemented with 10% FBS to a density of $3x10^6$ CFU/mL. 100 µL of this cell suspension was added to each well containing the differentiated THP-1 cells, resulting in a multiplicity of infection = 3. After 3 hours, the cells were washed twice with PBS to remove non-phagocytized bacteria, and 100 µL RPMI

containing 10% FBS was added. To silence target gene expression, the medium contained 0.3 µg/mL ATc. The THP-1 derived macrophages were incubated at 37 °C, 5% CO_2, and 85% humidity for 4 days. After incubation, cells were fixed with formalin (final concentration of 5%) and lyzed with deionized H_2O for 30 minutes. Dilutions of the suspensions from each well were plated onto Middlebrook 7H10 plates. Colonies were counted after 3 weeks of incubation at 37 °C.

References

1. World Health, O., *Global tuberculosis report 2022. .* licence: cc bY-Nc-sa 3.0 iGo, 2022.
2. Batt, S.M., D.E. Minnikin, and G.S. Besra, *The thick waxy coat of mycobacteria, a protective layer against antibiotics and the host's immune system.* Biochem J, 2020. **477**(10): p. 1983-2006.
3. Torrelles, J.B., A.K. Azad, and L.S. Schlesinger, *Fine Discrimination in the Recognition of Individual Species of Phosphatidyl- myo -Inositol Mannosides from Mycobacterium tuberculosis by C-Type Lectin Pattern Recognition Receptors.* The Journal of Immunology, 2006. **177**(3): p. 1805-1816.
4. Kang, P.B., et al., *The human macrophage mannose receptor directs Mycobacterium tuberculosis lipoarabinomannan-mediated phagosome biogenesis.* Journal of Experimental Medicine, 2005. **202**(7): p. 987-999.
5. Schlesinger, L.S., *Macrophage phagocytosis of virulent but not attenuated strains of Mycobacterium tuberculosis is mediated by mannose receptors in addition to complement receptors.* Journal of immunology (Baltimore, Md. : 1950), 1993. **150**(7): p. 2920-30.
6. Villeneuve, C., et al., *Mycobacteria use their surface-exposed glycolipids to infect human macrophages through a receptor-dependent process.* Journal of Lipid Research, 2005. **46**(3): p. 475-483.
7. Tallieux, L., et al., *DC-SIGN is the major Mycobacterium tuberculosis receptor on human dendritic cells.* Journal of Experimental Medicine, 2003. **197**(1): p. 121-127.
8. Armstrong, J.A. and D.a. Hart, *Response of cultured macrophages to Mycobacterium Tuberculosis, with observations on fusion of lysosomes with phagosomes.* Journal of Experimental Medicine, 1971. **134**(3): p. 713-740.
9. Zhang, Y., et al., *Mode of action of pyrazinamide: Disruption of Mycobacterium tuberculosis membrane transport and energetics by pyrazinoic acid.* Journal of Antimicrobial Chemotherapy, 2003. **52**(5): p. 790-795.
10. Mikusová, K., et al., *Biogenesis of the mycobacterial cell wall and the site of action of ethambutol.* Antimicrob Agents Chemother, 1995. **39**(11): p. 2484-9.
11. Savková, K., et al., *An ABC transporter Wzm-Wzt catalyzes translocation of lipid-linked galactan across the plasma membrane in mycobacteria.* Proc Natl Acad Sci U S A, 2021. **118**(17).
12. Wolucka, B.A., et al., *Recognition of the lipid intermediate for arabinogalactan/arabinomannan biosynthesis and its relation to the mode of action of ethambutol on mycobacteria.* J Biol Chem, 1994. **269**(37): p. 23328-35.
13. Jankute, M., et al., *Assembly of the Mycobacterial Cell Wall.* Annu Rev Microbiol, 2015. **69**: p. 405-23.
14. Mishra, A.K., et al., *Lipoarabinomannan and related glycoconjugates: structure, biogenesis and role in Mycobacterium tuberculosis physiology and host–pathogen interaction.* FEMS Microbiology Reviews, 2011. **35**(6): p. 1126-1157.

15. Mishra, A.K., et al., *Lipoarabinomannan and related glycoconjugates: Structure, biogenesis and role in Mycobacterium tuberculosis physiology and host-pathogen interaction.* FEMS Microbiology Reviews, 2011. **35**(6): p. 1126-1157.
16. Brecik, M., et al., *DprE1 Is a Vulnerable Tuberculosis Drug Target Due to Its Cell Wall Localization.* ACS Chemical Biology, 2015. **10**(7): p. 1631-1636.
17. Higgins, C.F. and M.M. Gottesman, *Is the multidrug transporter a flippase?* Trends Biochem Sci, 1992. **17**(1): p. 18-21.
18. Guan, S. and N.K. Verma, *Functional analysis of the O antigen glucosylation gene cluster of Shigella flexneri.* Microbiology, 1999. **145**(5).
19. Allison, G.E. and N.K. Verma, *Serotype-converting bacteriophages and O-antigen modification in Shigella flexneri.* Trends Microbiol, 2000. **8**(1): p. 17-23.
20. Korres, H., et al., *Topological analysis of GtrA and GtrB proteins encoded by the serotype-converting cassette of Shigella flexneri.* Biochem Biophys Res Commun, 2005. **328**(4): p. 1252-60.
21. Larrouy-Maumus, G., et al., *A small multidrug resistance-like transporter involved in the arabinosylation of arabinogalactan and lipoarabinomannan in mycobacteria.* Journal of Biological Chemistry, 2012. **287**(47): p. 39933-39941.
22. Kolly, G.S., et al., *GtrA protein Rv3789 is required for arabinosylation of arabinogalactan in Mycobacterium tuberculosis.* Journal of Bacteriology, 2015. **197**(23): p. 3686-3697.
23. Marmiesse, M., et al., *Macro-array and bioinformatic analyses reveal mycobacterial 'core' genes, variation in the ESAT-6 gene family and new phylogenetic markers for the Mycobacterium tuberculosis complex.* Microbiology (Reading), 2004. **150**(Pt 2): p. 483-496.
24. Jumper, J., et al., *Highly accurate protein structure prediction with AlphaFold.* Nature, 2021. **596**(7873): p. 583-589.
25. Griffin, J.E., et al., *High-resolution phenotypic profiling defines genes essential for mycobacterial growth and cholesterol catabolism.* PLoS Pathog, 2011. **7**(9): p. e1002251.
26. Dejesus, M.A., et al., *Comprehensive essentiality analysis of the Mycobacterium tuberculosis genome via saturating transposon mutagenesis.* mBio, 2017. **8**(1): p. 1-17.
27. Cole, S.T., et al., *Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence.* Nature, 1998. **393**(6685): p. 537-544.
28. Vilchèze, C., et al., *Inactivation of the inhA-encoded fatty acid synthase II (FASII) enoyl-acyl carrier protein reductase induces accumulation of the FASI end products and cell lysis of Mycobacterium smegmatis.* J Bacteriol, 2000. **182**(14): p. 4059-67.
29. Appelmelk, B.J., et al., *The mannose cap of mycobacterial lipoarabinomannan does not dominate the Mycobacterium-host interaction.* Cell Microbiol, 2008. **10**(4): p. 930-44.
30. Kolk, A.H.J., et al., *Production and characterization of monoclonal antibodies to Mycobacterium tuberculosis, M. bovis (BCG) and M. leprae.* Clinical and Experimental Immunology, 1984. **58**(3): p. 511-521.
31. Driessen, N.N., et al., *Role of phosphatidylinositol mannosides in the interaction between mycobacteria and DC-SIGN.* Infect Immun, 2009. **77**(10): p. 4538-47.
32. Mulnaes, D., et al., *TopModel: Template-Based Protein Structure Prediction at Low Sequence Identity Using Top-Down Consensus and Deep Neural Networks.* Journal of Chemical Theory and Computation, 2020. **16**(3): p. 1953-1967.
33. Gandini, R., et al., *Structural basis for dolichylphosphate mannose biosynthesis.* Nat Commun, 2017. **8**(1): p. 120.
34. Lomize, M.A., et al., *OPM database and PPM web server: Resources for positioning of proteins in membranes.* Nucleic Acids Research, 2012. **40**(D1): p. 370-376.
35. Feldman, M.F., et al., *The activity of a putative polyisoprenol-linked sugar translocase (Wzx) involved in Escherichia coli O antigen assembly is independent of the chemical structure of the O repeat.* J Biol Chem, 1999. **274**(49): p. 35129-38.

36. Marolda, C.L., J. Vicarioli, and M.A. Valvano, *Wzx proteins involved in biosynthesis of O antigen function in association with the first sugar of the O-specific lipopolysaccharide subunit.* Microbiology (Reading), 2004. **150**(Pt 12): p. 4095-105.
37. Alaimo, C., et al., *Two distinct but interchangeable mechanisms for flipping of lipid-linked oligosaccharides.* Embo j, 2006. **25**(5): p. 967-76.
38. Chin, C.Y., et al., *Francisella FlmX broadly affects lipopolysaccharide modification and virulence.* Cell Rep, 2021. **35**(11): p. 109247.
39. Yan, A., Z. Guan, and C.R. Raetz, *An undecaprenyl phosphate-aminoarabinose flippase required for polymyxin resistance in Escherichia coli.* J Biol Chem, 2007. **282**(49): p. 36077-89.
40. Hong, Y., M.A. Liu, and P.R. Reeves, *Progress in Our Understanding of Wzx Flippase for Translocation of Bacterial Membrane Lipid-Linked Oligosaccharide.* J Bacteriol, 2018. **200**(1).
41. Lambert, E., et al., *Evidence for a trap-and-flip mechanism in a proton-dependent lipid transporter.* Nat Commun, 2022. **13**(1): p. 1022.
42. Zhang, B., et al., *Structure of a proton-dependent lipid transporter involved in lipoteichoic acids biosynthesis.* Nat Struct Mol Biol, 2020. **27**(6): p. 561-569.
43. Ward, S.K., E.A. Hoye, and A.M. Talaat, *The global responses of Mycobacterium tuberculosis to physiological levels of copper.* J Bacteriol, 2008. **190**(8): p. 2939-46.
44. Rowland, J.L. and M. Niederweis, *A multicopper oxidase is required for copper resistance in Mycobacterium tuberculosis.* J Bacteriol, 2013. **195**(16): p. 3724-33.
45. Quinonez, C.G., et al., *The Role of Fatty Acid Metabolism in Drug Tolerance of Mycobacterium tuberculosis.* mBio, 2022. **13**(1): p. e0355921.
46. Stallings, C.L., et al., *CarD is an essential regulator of rRNA transcription required for Mycobacterium tuberculosis persistence.* Cell, 2009. **138**(1): p. 146-59.
47. Snider, M.D., *Topography Of Glycosylation In The Rouch Endoplasmatic Reticulum And Golgy Apparatus.* 1987.
48. Orlean, P., C. Albright, and P.W. Robbins, *Cloning and sequencing of the yeast gene for dolichol phosphate mannose synthase, an essential protein.* The Journal of biological chemistry, 1988. **263**(33): p. 17499-17507.
49. Sanyal, S. and A.K. Menon, *Stereoselective transbilayer translocation of mannosyl phosphoryl dolichol by an endoplasmic reticulum flippase.* Proceedings of the National Academy of Sciences of the United States of America, 2010. **107**(25): p. 11289-11294.
50. Haselbeck, A. and W. Tanner, *Dolichyl phosphate-mediated mannosyl transfer through liposomal membranes.* Proceedings of the National Academy of Sciences of the United States of America, 1982. **79**(5): p. 1520-1524.
51. Reichenbach, T., *Structural and biochemical insights into biosynthesis and degradation of N -glycans Structural and biochemical insights into biosynthesis and degradation of N -glycans*. 2020.
52. Zhou, Z., et al., *Function of Escherichia coli MsbA, an essential ABC family transporter, in lipid A and phospholipid biosynthesis.* J Biol Chem, 1998. **273**(20): p. 12466-75.
53. Margolles, A., et al., *The Purified and Functionally Reconstituted Multidrug Transporter LmrA of Lactococcus lactis Mediates the Transbilayer Movement of Specific Fluorescent Phospholipids.* Biochemistry, 1999. **38**(49): p. 16298-16306.
54. Sanyal, S. and A.K. Menon, *Flipping lipids: why an' what's the reason for?* ACS Chem Biol, 2009. **4**(11): p. 895-909.
55. Morita, Y.S., et al., *PimE is a polyprenol-phosphate-mannose-dependent mannosyltransferase that transfers the fifth mannose of phosphatidylinositol mannoside in mycobacteria.* Journal of Biological Chemistry, 2006. **281**(35): p. 25143-25155.
56. Jain, P., et al., *Specialized transduction designed for precise high-throughput unmarked deletions in Mycobacterium tuberculosis.* mBio, 2014. **5**(3): p. e01245-14.
57. Korte, J., et al., *Trehalose-6-Phosphate-Mediated Toxicity Determines Essentiality of OtsB2 in Mycobacterium tuberculosis In Vitro and in Mice.* PLoS pathogens, 2016. **12**(12): p. e1006043-e1006043.

58. Rappsilber, J., M. Mann, and Y. Ishihama, *Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips.* Nature Protocols, 2007. **2**(8): p. 1896-1906.
59. Michalski, A., et al., *Ultra high resolution linear ion trap orbitrap mass spectrometer (orbitrap elite) facilitates top down LC MS/MS and versatile peptide fragmentation modes.* Molecular and Cellular Proteomics, 2012. **11**(3): p. 1-11.
60. Olsen, J.V., et al., *Parts per million mass accuracy on an orbitrap mass spectrometer via lock mass injection into a C-trap.* Molecular and Cellular Proteomics, 2005. **4**(12): p. 2010-2021.
61. Laemmli, U.K., *Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4.* Nature, 1970. **227**: p. 680-685.
62. Mahmood, T. and P.C. Yang, *Western blot: technique, theory, and trouble shooting.* N Am J Med Sci, 2012. **4**(9): p. 429-34.
63. Trias, J. and R. Benz, *Permeability of the cell wall of Mycobacterium smegmatis.* Molecular Microbiology, 1994. **14**(2): p. 283-290.
64. Gilleron, M., V.F. Quesniaux, and G. Puzo, *Acylation state of the phosphatidylinositol hexamannosides from Mycobacterium bovis bacillus Calmette Guerin and mycobacterium tuberculosis H37Rv and its implication in Toll-like receptor response.* J Biol Chem, 2003. **278**(32): p. 29880-9.
65. Mosquera-Restrepo, S.F., et al., *A Mycobacterium tuberculosis fingerprint in human breath allows tuberculosis detection.* Nat Commun, 2022. **13**(1): p. 7751.
66. Mulnaes, D. and H. Gohlke, *TopScore: Using Deep Neural Networks and Large Diverse Data Sets for Accurate Protein Model Quality Assessment.* J Chem Theory Comput, 2018. **14**(11): p. 6117-6126.

Acknowledgement

Financial support for this study was provided to R.K. by the Jürgen Manchot Stiftung (graduate schools MOI III and IV) and the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) by project number 270650915 / GRK 2158 and by award number KA 2259/2-1.

Author contribution

Conceptualization, funding acquisition, and supervision, R.K; microbiological investigation, K.V., S.S., L.v.G., M.H., K.L., V.K.; lipid analyses, S.S.G., M.G.; molecular modelling, C.G.; proteome analysis, D.P., data analysis, H.G., J.N., G.B., M.K.; providing reagents, B.J.A.; writing – original draft, K.V., S.S:, R.K..

Ethics declarations

Competing interests

All authors declare no competing interests.

Supplementary Information

GtrA-like lipid floppase Rv3277 is essential for cell surface mannosylation in *Mycobacterium tuberculosis*

Kristin Vill[*], Steffen Schindler[1,*], Lasse van Geelen[1], David Podlesainski[2], Sudagar S. Gurcha[3], Milena Hänisch[1], Kathrin Lindner[1], Martine Gilleron[4], Christoph Gertzen[5], Violetta Krisilia[1], Ben J. Appelmelk[6], Holger Gohlke[5], Jérôme Nigou[4], Gurdyal S. Besra[3], Markus Kaiser[2], and Rainer Kalscheuer[1,§]

[1] Heinrich Heine University Düsseldorf, Faculty of Mathematics and Natural Sciences, Institute of Pharmaceutical Biology and Biotechnology, 40225 Düsseldorf, Germany

[2] Department of Chemical Biology, ZMB, Faculty of Biology, University of Duisburg-Essen, 45177 Essen, Germany

[3] School of Biosciences, University of Birmingham, Edgbaston, Birmingham, B15 2TT, United Kingdom

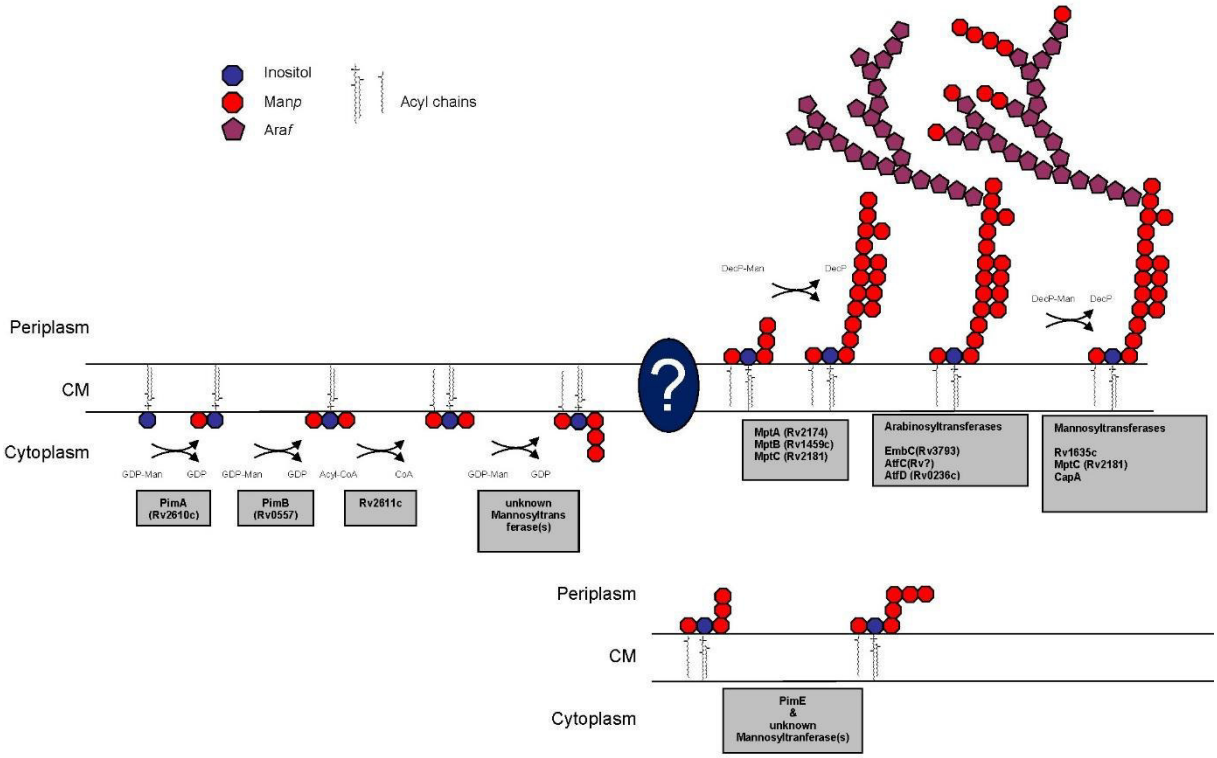
[4] Institute of Pharmacology and Structural Biology (IPBS), University of Toulouse, CNRS, University of Toulouse III-Paul Sabatier, Toulouse, France

[5] Institute of Pharmaceutical and Medicinal Chemistry, Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany

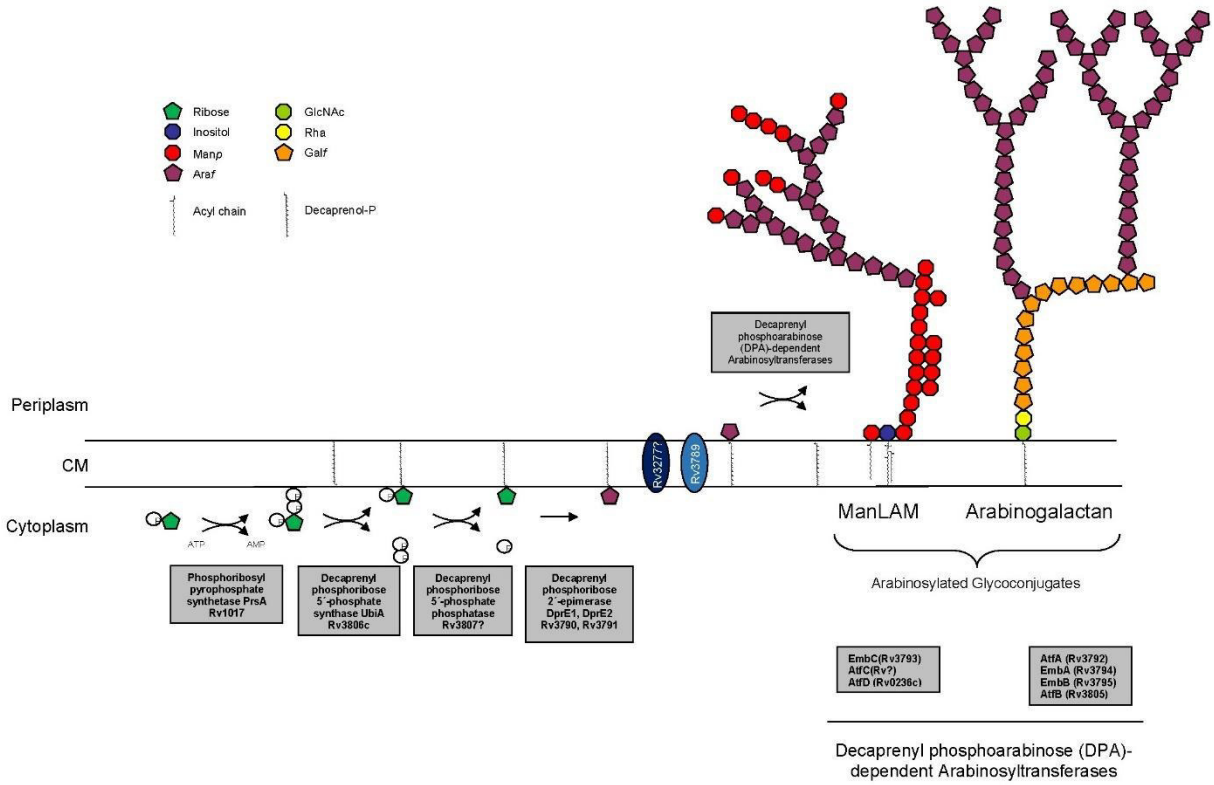
[6] Molecular Microbiology/Medical Microbiology and Infection Control, Amsterdam University Medical Centers, Amsterdam, Netherlands.

* These authors contributed equally

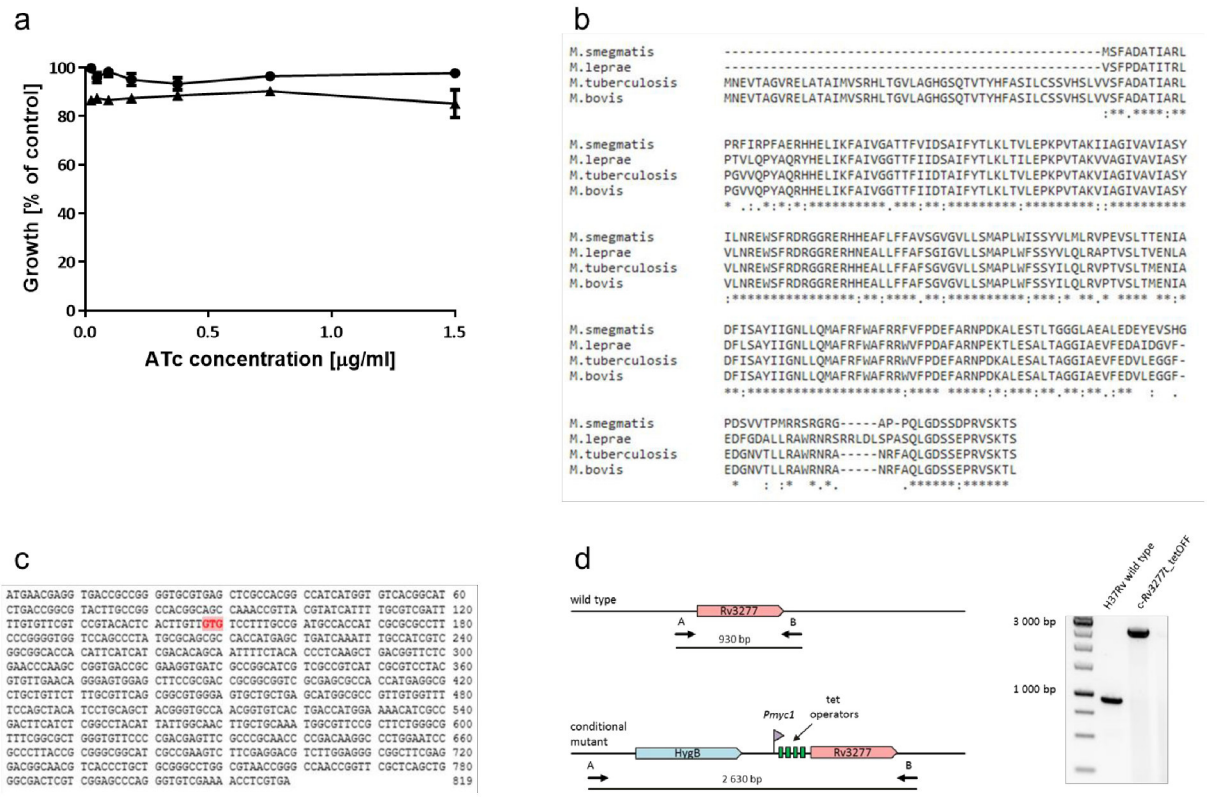
§ Author for correspondence: rainer.kalscheuer@hhu.de



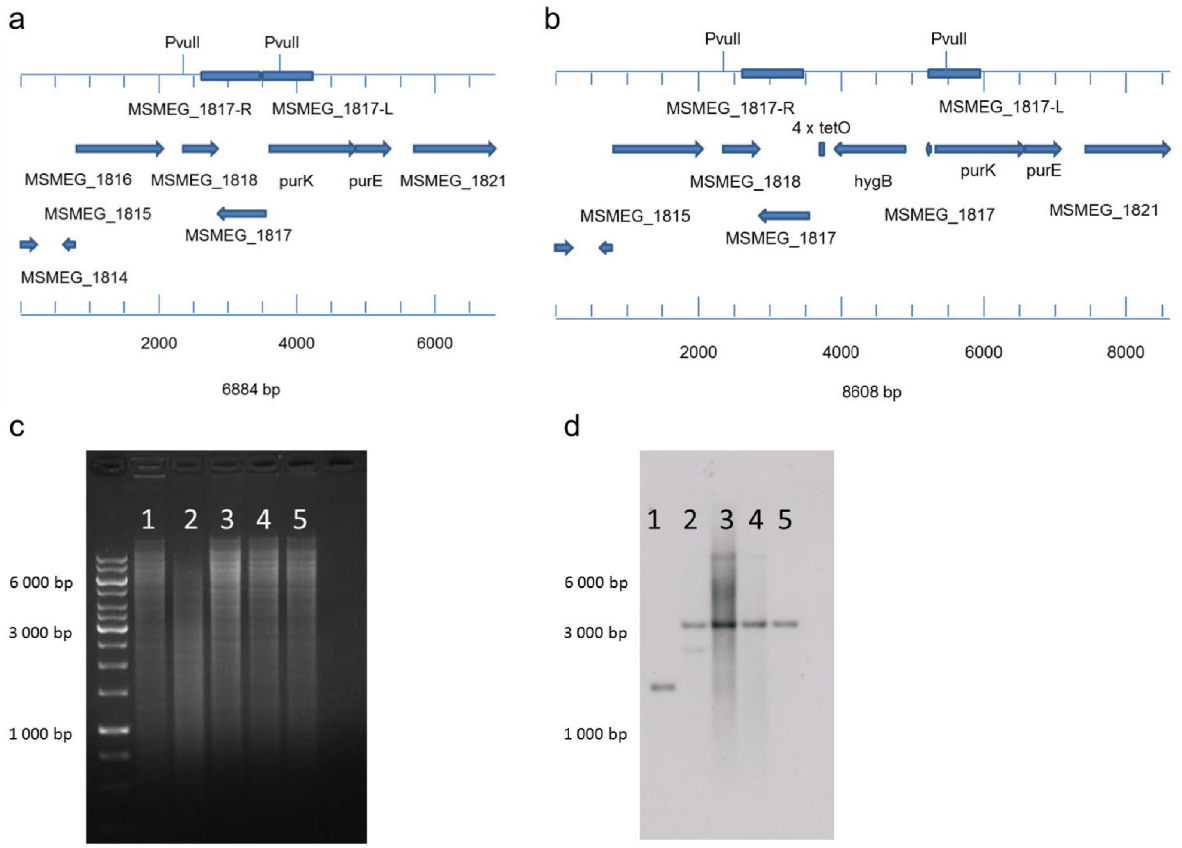
Supplementary Fig. 1: Scheme of the bipartite biosynthesis of phosphatidyl inositol-based mannosylated cell surface glycoconjugates in *M. tuberculosis*. The α-mannopyranosyltransferases PimA (Rv2610c), PimB′ (Rv2188c) and PimC or a yet-unidentified mannosyltransferase mediate the stepwise mannosylation of the *myo*-inositol moiety of the phospholipid phosphatidyl-*myo*-inositol, employing GDP-mannopyranose as the donor substrate, yielding phosphatidylinositol tetramannoside (PIM_4). Following mono- and diacylation by the acyltransferase Rv2611c and a yet-unknown enzyme, the resulting Ac_1/Ac_2PIM_4 is flopped by an unknown mechanism from the cytoplasmic to the periplasmic side of the plasma membrane. Ac_1/Ac_2PIM_4 serves as a branch point, which can either be extended by the mannopyranosyltransferase PimE (Rv1159) to form higher-order PIMs (Ac_1/Ac_2PIM_6), or by the mannopyranosyltransferase MptA (Rv2174), MptB (Rv1459c) and MptC (Rv2181) to synthesize LM. The involved mannopyranosyltransferases represent GT-C type glycosyltransferases that utilize decaprenyl-monophosphoryl-D-mannopyranose (DPM) as the donor substrate. A yet-unidentified arabinofuranosyltransferase first primes the mannan core of LM with a few arabinofuranose residues, after which EmbC (Rv3793) extends the primed LM with the addition of 12–16 arabinofuranose residues. Subsequently, the branching enzyme AftC and the arabinofuranosyltransferase AftB complete the nonreducing arabinan domain of LAM. EmbC, AftC and AftB all employ DPA as the donor substrate to result in addition of altogether 55–70 arabinofuranose residues to LM. Finally, the arabinan domain of LAM is modified with mannose residues by the DPM-dependent mannopyranosyltransferases CapA (Rv1635c) and MptC (Rv2181), resulting in ManLAM.



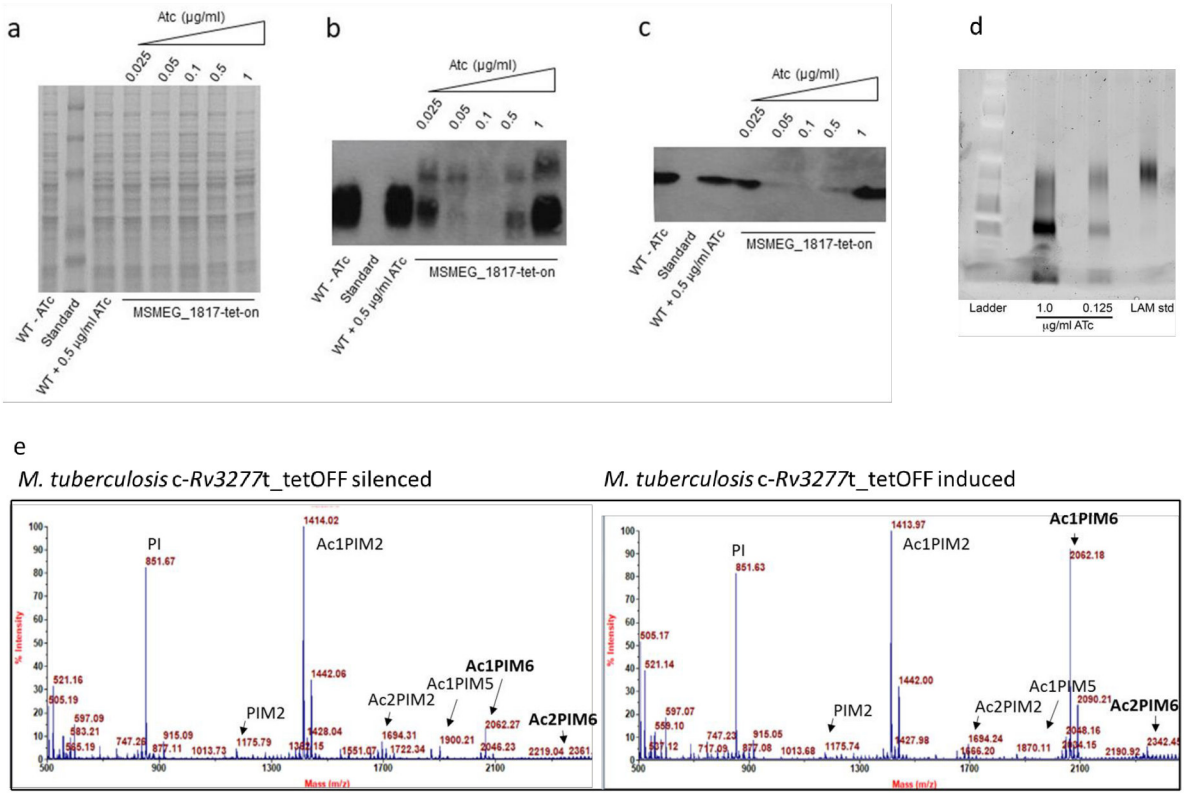
Supplementary Fig. 2: Scheme of the bipartite biosynthesis of decaprenyl monophosporyl-D-arabinofuranose (DPA). Phosphoribosyl pyrophosphate is produced by phosphorylation of ribose-1-phosphate catalyzed by phosphoribosyl-1-pyrophosphate synthetase PrsA (Rv1017). Next, decaprenylphosphoryl-5-phosphoribose synthase UbiA (Rv3806c) transfers phosphoribosyl pyrophosphate to C50-P, yielding decaprenyl-monophosphoryl-D-ribofuranosephosphate, which is subsequently dephosphorylated by Rv3807 to produce decaprenyl-monophosphoryl-D-ribofuranose (DPR). Decaprenylphosphoribose 2′ epimerase composed of DprE1 (Rv3790) and DprE2 (Rv3791) subunits then catalyzes the epimerization of the ribosyl unit of DPR, resulting in DPA. DPA is then translocated from the cytosolic to the periplasmic face of the membrane by the GtrA-like floppase Rv3789 and/or other yet-unknown lipid floppases potentially including Rv3277t. At the periplasmic face of the membrane, DPA serves as the substrate for various arabinofuranosyltransferases involved in arabinosylation of lipomannan or the galactan core C50-P-P-GlcNAc-Rha-Galf_{30}, resulting in formation of lipoarabinomannan (LAM) and arabinogalactan, respectively. LAM is subsequently subject to terminal mannosylation steps, yielding mannose-capped LAM (ManLAM).



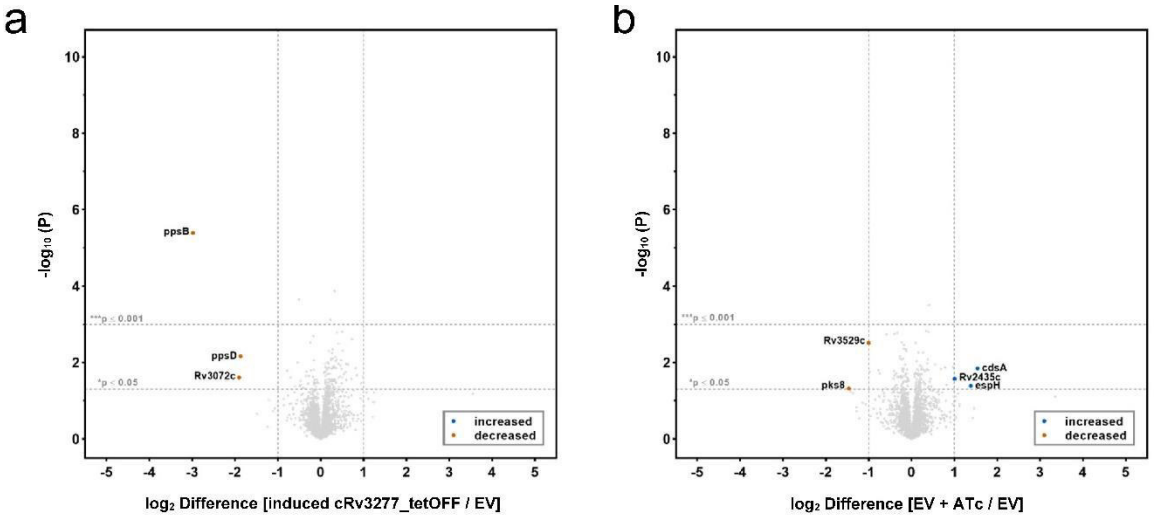
Supplementary Fig. 3: Reannotation of the essential gene *Rv3277* with an alternative starting codon. a, ATc-dependent growth of the conditional *M. tuberculosis* c-*Rv3277*_tetOFF mutant (triangle) and of the *M. tuberculosis* H37Rv wild type strain (circle) indicating no essentiality of the annotated full-length *Rv3277* gene. **b,** Multiple sequence alignment of the *M. tuberculosis* Rv3277 orthologues in *M. smegmatis*, *M. leprae*, and *M. bovis* using CLUSTAL by Muscle (3.8). Asterisks show positions that are highly identical, whereas colon and period show positions with lower similar residues. **c,** Sequence of *Rv3277* as annotated in Mycobrowser. The 5' ATG codon is annotated as the start codon, although the highlighted GTG codon 147 bp downstream was experimentally shown to be the probable start site for translation. **d,** Generation of the M. tuberculosis Rv3277t conditional mutant. Organization of the Rv3277 locus in M. tuberculosis H37Rv wild type and the c-Rv3277t_tetOFF conditional mutant with the four tet operator sites upstream of the annotated start codon (left panel). The relevant primers A and B are indicated. The sizes of the relevant diagnostic PCR products for verification are shown. Diagnostic PCR to verify the generation of the conditional c-Rv3277t_tetOFF mutant (right panel). The observed PCR products contained 930 bp for the M. tuberculosis H37Rv wild type strain and 2630 bp for the M. tuberculosis c-Rv3277t_tetOFF mutant strain. Diagnostic PCRs were carried out on genomic DNA of the respective strains.



Supplementary Fig. 4: Generation of the c-*Msmeg_1817*_tetON mutant analogous to the c-*Rv3277t*_tetOFF mutant. a + b, *MSMEG_1817* loci with the restriction sites for the Southern blot. Both Sequences were cut with *Pvu*II. The restriction sites are indicated in each map (*Pvu*II). The sequences were hybridized with *MSMEG_1817* -R as probe. *MSMEG_1817* -R and *MSMEG_1817* -L are indicating the flanking arms. **a,** The wild type *MSMEG_1817* locus is shown as a structure with an expected signal of 1321 bp. **b,** The locus of the knock-in mutant *MSMEG_1817* with the additional hygromycin resistance (*hygB*) and the promoter including four operator sites is shown with an expected signal at 3045 bp. **c,** Gel electrophoresis of RNase treated and *Pvu*II digested DNA of *M. smegmatis* mc^{2}155 (1), and 4 putative mutants (2-5). **d,** Verification of mutants by Southern blot. RNase treated and *Pvu*II digested DNA of *M. smegmatis* mc^{2}155 (1), and 4 putative mutants (2-5) were electrophoretically separated, transferred to a nylon membrane, and finally hybridized with a probe. Binding was visualized by chemiluminescence using luminol.



Supplementary Fig. 5: Silencing of *MSMEG_1817* blocks biosynthesis of the higher order mannosylated glycoconjugates LM, LAM and PIM. a, Coomassie-stained SDS-PAGE showing loading controls. **b - c,** SDS-PAGE immunoblots employing the LAM-specific monoclonal antibody F30-5 (b) and the PIM6-specific monoclonal antibody F183-24 (c). The rebounding levels of LAM and PIM6 at the lowest ATc concentration of 0.025 µg/ml probably results from the enrichment of non-regulated suppressor mutants that constitutively express *MSMEG_1817*, explainable by the strong selective pressure that occurs when silencing this essential gene. **d,** SDS-PAGE autoradiographic analysis of the LM/LAM fraction prepared from ^{14}C-glucose labeled fully induced (at 1 µg/ml ATc) and partially silenced (at 0.125 µg/ml ATc) cells of the *M. smegmatis* c-*Msmeg_1817*_tetON mutant. e) MALDI-ToF MS analysis of lipids extracted from fully induced (at 0 µg/ml ATc) and partially silenced (at 0.5 µg/ml ATc) cells of the *M. tuberculosis* c-*Rv3277t*_tetOFF mutant. PI, phosphatidylinositol; PIM_2, phosphatidylinositol dimannoside; Ac_1PIM_2, monoacylphosphatidylinositol dimannoside; Ac_2PIM_2, diacylphosphatidylinositol dimannoside; Ac_1PIM_6, monoacylphosphatidylinositol hexamannoside; Ac_2PIM_6, diacylphosphatidylinositol hexamannoside.



Supplementary Fig. 6: Comparative control LC-MS/MS-based proteomic profilings, related to Fig. 7. **a**, LC-MS/MS-based whole protein analysis of fully induced cells of the *M. tuberculosis* H37Rv c-*Rv3277t*_tetOFF mutant compared to isogenic cells carrying an empty vector control (EV), illustrating largely unchanged protein abundances. **b**, LC-MS/MS-based whole protein analysis of cells of the *M. tuberculosis* H37Rv c-*Rv3277t* knock-in mutant carrying an empty vector control (EV) incubated in presence or absence of ATc, demonstrating minimal effect of ATc on protein abundances. The volcano plots illustrate the log_2- fold change in abundance (X-axis) and corresponding -log10 *p* values (Y-axis). Proteins complying with the chosen threshold of significance and showing a log_2-fold change ≥ 1 or ≤ -1 are marked in blue or orange, respectively. Quantification was done via label free quantification (LFQ) of five replicates per sample group. To identify statistically significant hits from the analysis, P ≤ 0.05 (Student's T-test; permutation-based FDR with 250 randomizations and FDR = 0.01) was applied.

Supplementary Table 1: Identified response clusters in the proteomic stress profile of silenced vs. induced cells of the *M. tuberculosis* H37Rv c-*Rv3277t*_tetOFF mutant. Chosen threshold: log2 Difference ≤ -1 and ≥ 1; P ≤ 0.05.

Pathway	-Log$_{10}$ (P-value)	Log$_2$ Difference	Gene names	Protein names
Copper stress response	6.396	4.529	Rv2963	Putative permease Rv2963
	6.854	4.105	cysK2 (Rv0848)	O-phospho-L-serine-dependent S-sulfocysteine synthase
	4.600	3.123	csoR (Rv0967)	Copper-sensing transcriptional repressor CsoR
	3.069	2.997	Rv0849	Uncharacterized MFS-type transporter Rv0849
	6.474	2.924	mctB (Rv1698)	Copper transporter MctB
	9.024	2.665	mmcO (Rv0846c)	Multicopper oxidase MmcO
	1.405	2.280	mymT (Rv0186A)	Copper-binding metallothionein
	7.754	2.245	Rv0968	Uncharacterized protein Rv0968
	8.739	2.236	ricR (Rv0190)	Copper-sensing transcriptional repressor RicR
	9.344	2.226	ctpV (Rv0969)	Probable copper-exporting P-type ATPase V
	3.089	2.180	cadI (Rv2641)	Cadmium-induced protein CadI
	1.937	1.723	Rv0430	Uncharacterized protein
	6.608	-1.788	Rv0500A	Putative DNA-binding protein Rv0500A
Electron transport chain	6.203	-1.295	nuoE (Rv3149)	NADH-quinone oxidoreductase subunit E
	8.873	-2.036	cydA (Rv1623c)	Probable integral membrane cytochrome D ubiquinol oxidase (Subunit I)
	6.374	-2.241	narX (Rv1736c)	Nitrate reductase-like protein NarX
	5.751	-2.514	cydC (Rv1620c)	Probable 'component linked with the assembly of cytochrome' transport transmembrane ATP-binding protein ABC transporter CydC
	5.906	-2.684	cydD (Rv1621c)	Probable 'component linked with the assembly of cytochrome' transport transmembrane ATP-binding protein ABC transporter CydD
Cell devision	5.167	1.577	ftsK (Rv2748c)	DNA translocase FtsK
	2.864	1.329	sepF (Rv2147c)	Cell division protein SepF
	6.014	1.166	ftsQ (Rv2151c)	Cell division protein FtsQ
Cell wall	6.268	2.430	pgsA2 (Rv1822)	Putative cardiolipin synthase
	3.264	2.160	ppsB (Rv2932)	Phenolphthiocerol/phthiocerol polyketide synthase subunit B
	3.127	2.156	ppsD (Rv2934)	Phenolphthiocerol/phthiocerol polyketide synthase subunit D
	2.637	2.048	ppsC (Rv2933)	Phenolphthiocerol/phthiocerol polyketide synthase subunit C
	5.663	1.812	Rv3717	N-acetylmuramoyl-L-alanine amidase
	5.261	1.630	fadD26 (Rv2930)	Long-chain-fatty-acid--AMP ligase FadD26
	7.592	1.576	murG (Rv2153c)	Undecaprenyl-PP-MurNAc-pentapeptide-UDPGlcNAc GlcNAc transferase
	5.297	1.416	ppsA (Rv2931)	Phenolphthiocerol/phthiocerol polyketide synthase subunit A
	1.539	1.288	ripA (Rv1477)	Peptidoglycan endopeptidase RipA
	3.161	1.282	pks15 (Rv2947c)	Putative inactive phenolphthiocerol synthesis polyketide synthase type I Pks15
	5.465	1.180	fbpC (Rv0129c)	Diacylglycerol acyltransferase/mycolyltransferase Ag85C (Fibronectin-binding protein C)
	5.525	1.146	murC (Rv2152c)	UDP-N-acetylmuramate--L-alanine ligase
	5.031	1.084	mas (Rv2940c)	Probable multifunctional mycocerosic acid synthase Mas
	5.566	1.047	mmpL10 (Rv1183)	Acyltrehalose exporter MmpL10

Supplementary Table 1 (continued)

Pathway	-Log10 (P-value)	Log2 Difference	Gene names	Protein names
Ribosomal proteins	3,637	-1,000	rplQ (Rv3456c)	50S ribosomal protein L17
	4,125	-1,076	rpmC (Rv0709)	50S ribosomal protein L29
	4,276	-1,077	rpsO (Rv2785c)	30S ribosomal protein S15
	4,644	-1,240	rplU (Rv2442c)	50S ribosomal protein L21
	2,801	-1,365	rpmA (Rv2441c)	50S ribosomal protein L27
	3,210	-1,432	rpmD (Rv0722)	50S ribosomal protein L30
	1,534	-1,490	rpmF (Rv0979A)	50S ribosomal protein L32
Stress proteins	6,962	-1,177	Rv2624c	Universal stress protein Rv2624c
	6,039	-1,199	Rv2623	Universal stress protein Rv2623
	3,942	-1,343	Rv2005c	Universal stress protein Rv2005c
	5,242	-1,416	Rv1996	Universal stress protein Rv1996
	6,031	-1,622	hspX (Rv2031c)	Alpha-crystallin (Acr) (16 kDa antigen)
	7,934	-2,141	Rv3134c	Universal stress protein Rv3134c
Antitoxins	1,983	-1,083	vapB29 (Rv0616A)	Putative antitoxin VapB29
	5,637	-1,351	mazE3 (Rv1103c)	Antitoxin MazE3
	7,263	-1,551	vapB38 (Rv2493)	Putative antitoxin VapB38
	1,345	-1,572	vapB33 (Rv1241)	Antitoxin VapB33
Transcription factors / DNA binding proteins	2,748	-1,050	cspA (Rv3648c)	Probable cold shock protein A
	6,257	-1,054	Rv1404	Probable transcriptional regulatory protein
	5,851	-1,126	Rv3295	Probable transcriptional regulatory protein (Probably TetR-family)
	3,007	-1,148	Rv0232	Probable transcriptional regulatory protein (Probably TetR/AcrR-family)
	4,906	-1,163	Rv3716c	Nucleoid-associated protein Rv3716c
	5,072	-1,262	Rv0081	Uncharacterized HTH-type transcriptional regulator Rv0081
	7,298	-1,427	dosR (Rv3133c)	DNA-binding transcriptional activator DevR/DosR
	6,525	-1,496	carD (Rv3583c)	RNA polymerase-binding transcription factor CarD
	4,449	-1,545	argR (Rv1657)	Arginine repressor
	7,259	-1,628	Rv0386	Probable transcriptional regulatory protein (Probably LuxR/UhpA-family)
	6,608	-1,788	Rv0500A	Putative DNA-binding protein Rv0500A
	6,272	-1,800	Rv3788	Uncharacterized protein Rv3788
dNDP synthesis	4,791	-1,151	nrdF2 (Rv3048c)	Ribonucleoside-diphosphate reductase subunit beta NrdF2
	2,307	-1,326	nrdB (Rv0233)	Ribonucleotide reductase R2-like ligand binding oxidase

Supplementary Table 2: Strains used in this study

Strain	Relevant properties	Origin
Mycobacterium tuberculosis H37Rv wild type	Wild type	
M. tuberculosis c-*Rv3277*_tetOFF	*Rv3277* silenced, annotated start codon ATG, kan^R, hyg^R	This study
M. tuberculosis c-*Rv3277t*_tetOFF	*Rv3277t* silenced, reannotated start codon GTG, 147 bp downstream, kan^R, hyg^R	This study
M. tuberculosis c-*Rv3277t*_tetOFF pMV361::EV	*Rv3277t* silenced, complemented with an empty vector	This study
M. tuberculosis c-*Rv3277t*_tetOFF pMV361::*Rv3277t*	*Rv3277t* silenced, complemented with a wild type copy of *Rv3277t* constitutively expressed from a single-copy integrative plasmid	This study
M. tuberculosis c-*Rv3277t*_tetOFF pMV361::*Rv3277t* T35A_K59A_N141A	*Rv3277t* silenced, complemented with a mutated copy of *Rv3277t*, harboring the mutations T35A, K59A, and N141A constitutively expressed from a single-copy integrative plasmid	This study
M. tuberculosis c-*Rv3277t*_tetOFF pMV361::*Rv3789*	*Rv3277t* silenced, complemented with a wild type copy of *Rv3789,* constitutively expressed from a single-copy integrative plasmid	This study
M. tuberculosis c-*Rv3277t*_tetOFF pMV361::*arnEF*	*Rv3277t* silenced, complemented with a wild type copy of *arnEF*, isolated from *E. coli,* constitutively expressed from a single-copy integrative plasmid	This study
M. tuberculosis c-*Rv3277t*_tetOFF pMV361::*wzxE*	*Rv3277t* silenced, complemented with a wild type copy of *wzxE*, isolated from *E. coli,* constitutively expressed from a single-copy integrative plasmid	This study
M. tuberculosis c-*Rv3277t*_tetOFF pMV361::*ltaA*	*Rv3277t* silenced, complemented with a wild type copy of *ltaA,* isolated from *S. aureus*, constitutively expressed from a single-copy integrative plasmid	This study

Mycobacterium smegmatis mc^{2}155 wild type	Wild type	
M. smegmatis c-*Msmeg_1817*_tetON	*MSMEG_1817* silenced, annotated start codon ATG, kan^R, hyg^R	This study
Escherichia coli NEB5α	Cloning strain for plasmids	New England Biolabs (Cat.-No. C2987I)
E. coli HB101	Cloning strain for phasmids	

Supplementary Table 3: Oligonucleotides used in this study. Restriction sites used for cloning purposes are underlined.

Oligonucleotide	Sequence [5‘ – 3‘]
c-*Rv3277*-RR-*Van*91I	TTTTTCCATCTTTTGGGGTGCATGCGTTGCCATC
c-*Rv3277*-RL-*Van*91I	TTTTTCCATAGATTGGATGAACGAGGTGTCAGTG
c-*Rv3277*-LR-*Van*91I	TTTTTCCATTTCTTGGGGCTGGACCACCAGAGG
c-*Rv3277*-LL-*Van*91I	TTTTTTCCATAAATTGGGGAATTATGTGGCCGGAG
c-*Rv3277t*-RR-*Van*91I	TTTTTCCATAGATTGGATGTCCTTTGCCGATGCCACCATC
c-*Rv3277t*-RL-*Van*91I	TTTTTCCATCTTTTGGCCGCTCGAGTTCTACAACATTCC
c-*Msmeg_1817*-RR-Van91I	TTTTTCCATCTTTTGGGGCCAGGATCAACAGTGTGGA
c-*Msmeg_1817*-RL-Van91I	TTTTTCCATAGATTGGATGTCCTTCGCTGATGCAACG
c-*Msmeg_1817*-LR-Van91I	TTTTTCCATTTCTTGGAGCTCATGGTGCCGTTCGGCG
c-*Msmeg_1817*-LL-Van91I	TTTTTCCATAAATTGGTCAGATACTTCGCGGCCACCT
Msmeg_1817-3-HindIII	TTTTTAAGCTTCACGAAGTCTTCGACACCCTGG
Msmeg_1817-5-PacI	TTTTTAATTAATGTCCTTCGCTGATGCAACGATC
c-Rv3277-3-HindIII	TTTTTAAGCTTCACGAGCTTTTCGACACCCTGG
c-Rv3277-5-PacI	TTTTTAATTAATGAACGAGGTGACCGCCGGGGTG
c-*Rv3277t*-PacI	TTTTTAATTAATGTCCTTTGCCGATGCCACCATC
3' *Rv3789* HindIII	AAGCTTGCCCGATTCGACCCTGACC
5' *Rv3789* PacI	GGGCGTTAATTAATGCACCTGTCGGCACAGGTA
5' PacI *arnEF*	GCGTTAATTAATCTGGTGCTGATTCAGTATCGTCC
3' HindIII *arnEF*	GCGAAGCTTGTGTAAAGTGGCAGGGAAAAC
5 PacI *wzxE*	CCGGCGTTAATTAAGCGACTTTGTTGAACTACTTTTCCTGAT
3 wzxE *HindIII*	CGCGGAAGCTTCGATCCCAGTACGTGAATCAGTACAG
5' PacI *ltaA*	GCGGTTAATTAATTTATGGAAAGGTTCCTTTATAT
3' *ltaA* HindIII	GCGGAAGCTTTCTTTATTTTAAAATACGTTTTAACC

Supplementary Table 4: Plasmids and Phasmids used in this study

Plasmid (p) or Phasmid (ph)	Relevant properties
p0004S	hyg^R
phAE159	Temperature sensitive shuttle phasmid, derivative of mycobacteriophage TM4, amp^R
pc*Rv3277*	Knock-in cassette, hyg^R, sacB, flanking regions of *Rv3277*
phc*Rv3277*	Knock-in cassette, hyg^R, sacB, flanking regions of *Rv3277*
pc*Rv3277t*	Knock-in cassette, hyg^R, sacB, flanking regions of *Rv3277t*
phc*Rv3277t*	Knock-in cassette, hyg^R, sacB, flanking regions of *Rv3277t*
pc*Msmeg_1817*	Knock-in cassette, hyg^R, sacB, flanking regions of *Msmeg_1817*
phc*Msmeg_1817*	Knock-in cassette, hyg^R, sacB, flanking regions of *Msmeg_1817*
pMV261	kan^R
pMV261_revtetR_RBS-Mut.E	kan^R
pMV361	$apra^R$
pMV361_*Msmeg1817*	$apra^R$, complementation plasmid
pMV361_*Rv3277t*	$apra^R$, complementation plasmid
pMV361::*Rv3277t*-T35A-K59A-N141A	$apra^R$, complementation plasmid
pMV361::*Rv3789*	$apra^R$, complementation plasmid
pMV361::*arnE/F*	$apra^R$, complementation plasmid
pMV361::*wzxE*	$apra^R$, complementation plasmid
pMV361::*ltaA*	$apra^R$, complementation plasmid

6. A temperature-sensitive *Mycobacterium smegmatis glgE* mutation leads to a loss of GlgE enzyme activity and thermostability and the accumulation of α-maltose-1-phosphate

Published in: Biochimica et Biophysica Acta – General Subjects

Impact Factor: 3.422 (2019)

DOI: 10.1016/j.bbagen.2020.129783

Overall contribution to the paper:

- Construction of microbial strains
- Generation and evaluation of growth curves
- Interpretation of the data obtained
- Feedback on the manuscript draft

BBA - General Subjects 1865 (2021) 129783

Contents lists available at ScienceDirect

BBA - General Subjects

journal homepage: www.elsevier.com/locate/bbagen

A temperature-sensitive *Mycobacterium smegmatis glgE* mutation leads to a loss of GlgE enzyme activity and thermostability and the accumulation of α-maltose-1-phosphate

Karl Syson [a], Sibyl F.D. Batey [a], Steffen Schindler [b], Rainer Kalscheuer [b], Stephen Bornemann [a,*]

[a] Biological Chemistry Department, John Innes Centre, Norwich Research Park, Norwich NR4 7UH, United Kingdom
[b] Institute of Pharmaceutical Biology and Biotechnology, Heinrich Heine University, 40225 Düsseldorf, Germany

ARTICLE INFO

Keywords:
Glycogen
Temperature-sensitive mutation
Mycobacterium smegmatis
GlgE
α-Maltose-1-phosphate
α-Glucan

ABSTRACT

Background: The bacterial GlgE pathway is the third known route to glycogen and is the only one present in mycobacteria. It contributes to the virulence of *Mycobacterium tuberculosis*. The involvement of GlgE in glycogen biosynthesis was discovered twenty years ago when the phenotype of a temperature-sensitive *Mycobacterium smegmatis* mutation was rescued by the *glgE* gene. The evidence at the time suggested *glgE* coded for a glucanase responsible for the hydrolysis of glycogen, in stark contrast with recent evidence showing GlgE to be a polymerase responsible for its biosynthesis.

Methods: We reconstructed and examined the temperature-sensitive mutant and characterised the mutated GlgE enzyme.

Results: The mutant strain accumulated the substrate for GlgE, α-maltose-1-phosphate, at the non-permissive temperature. The glycogen assay used in the original study was shown to give a false positive result with α-maltose-1-phosphate. The accumulation of α-maltose-1-phosphate was due to the lowering of the k_{cat} of GlgE as well as a loss of stability 42 °C. The reported rescue of the phenotype by GarA could potentially involve an interaction with GlgE, but none was detected.

Conclusions: We have been able to reconcile apparently contradictory observations and shed light on the basis for the phenotype of the temperature-sensitive mutation.

General significance: This study highlights how the lowering of flux through the GlgE pathway can slow the growth mycobacteria.

1. Introduction

A third biosynthetic pathway to produce glycogen has been elucidated relatively recently [1,2]. This involves the enzyme GlgE (Fig. 1) acting as the polymerase [1,3]. GlgE was first identified as an enzyme involved in glycogen metabolism in 1999 [4]. Intriguingly, it was then reported that a temperature-sensitive mutant appeared to accumulate glycogen, which is the opposite of what would be expected given our current understanding of the GlgE pathway. This study addresses this apparent discrepancy.

Glycogen is well known as a carbon and energy storage molecule [6]. When a temperature-sensitive *Mycobacterium smegmatis* mutant that slowed growth and appeared to accumulate glycogen was isolated (SMEG53), the phenotype was complemented by the *glgE* gene [4]. The sequence of the genomic copy of the *glgE* gene revealed a His349Tyr amino acid substitution was present in the GlgE protein of the temperature-sensitive mutant. There appeared to be a correlation between the accumulation of glycogen with the cessation of exponential growth at the non-permissive temperature of 42 °C. Furthermore, there also appeared to be a correlation between the accumulation of glycogen and growth rate. Suppression of the temperature-sensitive phenotype was observed in certain growth conditions and by the over-expression of *garA*, which codes for a forkhead-associated domain protein consequently referred to as a glycogen accumulation regulator [7].

GlgE is a GH13_3 CAZy family member [8]. The GH13 family comprises proteins with several enzyme activities, most notably α-amylases

* Corresponding author at: The Sainsbury Laboratory, Norwich Research Park, Norwich NR4 7UH, United Kingdom.
E-mail addresses: karl.syson@astrazeneca.com (K. Syson), Sibyl.batey@jic.ac.uk (S.F.D. Batey), Steffen.schindler@hhu.de (S. Schindler), Rainer.Kalscheuer@hhu.de (R. Kalscheuer), stephen.bornemann@tsl.ac.uk (S. Bornemann).

https://doi.org/10.1016/j.bbagen.2020.129783
Received 21 April 2020; Received in revised form 19 October 2020; Accepted 4 November 2020
Available online 7 November 2020

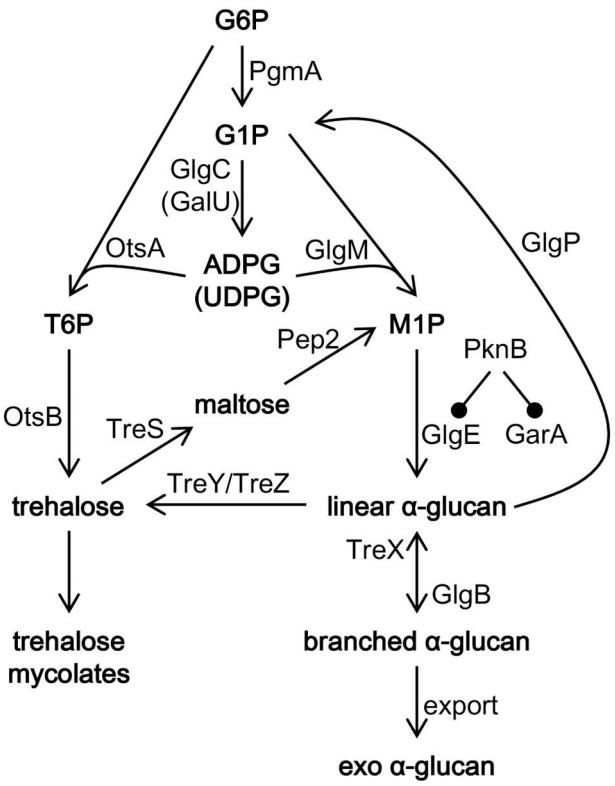


Fig. 1. Glycogen (α-glucan) pathways in mycobacteria. The scheme shows how glucose-6-phosphate (G6P) can be converted to α-glucan via the following intermediates; glucose-1-phosphate (G1P), ADP-glucose (ADPG) or UDP-glucose (UDPG), trehalose-6-phosphate (T6P), and α-maltose-1-phosphate (M1P). The regulation of GlgE and GarA by PknB is indicated. The figure is adapted from [5].

[9]. On this basis, GlgE was initially predicted to be a hydrolase capable of degrading the α-1,4 linkages of glycogen (glucanase) [4]. Given the interpretation of the results of the above study and the bioinformatic prediction, it was concluded that GlgE is a glucanase and that it is involved in glycogen recycling that is essential for exponential growth.

However, it has subsequently and unequivocally been shown that GlgE possesses α-maltose-1-phosphate:(1 → 4)-α-D-glucan 4-α-D-maltosyltransferase polymerase-type activity (EC 2.4.99.16) and some disproportionation activity, but no detectable hydrolase activity [1,10]. The polymerases for each of the known glycogen biosynthetic pathways extend existing oligomers and polymers to generate linear α-1,4-linked chains [3,11]. In all three pathways, a branching enzyme GlgB then introduces α-1,6 branchpoints. The GlgE pathway uses α-maltose-1-phosphate as the building block for the biosynthesis of an α-glucan that is similar to classical glycogen except for its fine structure [3]. There are two routes to the formation of α-maltose-1-phosphate in mycobacteria (Fig. 1). One is via trehalose and maltose plus ATP, involving trehalose synthase and maltose kinase [1,2], while the other is via glucose-1-phosphate plus ADP-glucose involving α-maltose-1-phosphate synthase [5,12]. Importantly, the loss of the *glgE* gene in *M. smegmatis* is known to lead to the accumulation of α-maltose-1-phosphate and the slowing of bacterial growth [1,5]. The GlgE-derived glycogen is partially exported by *Mycobacterium tuberculosis* to form a loosely attached capsule that is associated with increased virulence [5,13]. Notably, the two classical glycogen biosynthetic pathways that use ADP-glucose or UDP-glucose [6] are now known to be absent in mycobacteria [5].

To address an apparent inconsistency between the initial study suggesting GlgE is a hydrolase and subsequent work showing that it is a polymerase, several hypotheses were tested. Since GlgE is a polymerase that uses α-maltose-1-phosphate, a variant of GlgE that is somehow compromised would be expected to accumulate this building block leading to a lower bacterial rate of growth, as has been observed in other studies [1]. It is therefore possible that the glycogen assay used in the 1999 study, based on an amyloglucosidase and the quantification of released glucose [4], was also capable of detecting α-maltose-1-phosphate. Interestingly, the GarA protein is known to regulate carbon and nitrogen metabolism through binding to other proteins [7]. Perhaps GarA binds to GlgE and rescues the temperature-sensitive version. However, it is also a substrate of the regulatory Ser/Thr protein kinase PknB, which in turn has also been shown to negatively regulate GlgE [14]. Perhaps GarA therefore rescues the GlgE mutation through interfering with the regulatory function of PknB. The present study aims to test these hypotheses and resolve the apparent inconsistencies. This was done by reconstructing and phenotyping the mutant strain together with testing the hydrolysis of α-maltose-1-phosphate by amyloglucosidase and characterising the H349Y variant of the GlgE enzyme.

2. Materials and methods

2.1. Recombinant M. smegmatis GlgE and GarA

The *M. smegmatis glgE* gene was synthesized with optimized codon usage for expression in *Escherichia coli* (Genscript Corporation, Piscataway, NJ), allowing the production of GlgE with a 6 × His tag and a TEV cleavage site at its *N*-terminus. The gene coding for the H349Y GlgE variant was generated from the synthesized wild-type *glgE* gene using a QuikChange Lightening kit (Agilent). The constructs were ligated into pET21a expression vectors (Novagen, Darmstadt, Germany) using *Nde*I and *Bam*HI restriction sites. *E. coli* BL21 (DE3) Star (Novagen) bearing each plasmid were grown at 25 °C to an optical density of 0.6 at 600 nm in Lysogeny Broth (LB) and expression was induced with 0.5 mM isopropyl β-D-thiogalactopyranoside (IPTG). Bacteria were harvested and lyzed after a further 16 h incubation. The enzyme was purified using a 1 ml HisTrap FF column (GE Healthcare, Amersham, United Kingdom) with imidazole gradient elution and an S200 16/60 gel filtration column (Pharmacia Biotech, Amersham, United Kingdom) eluted with 20 mM Tris buffer, pH 8.5, containing 100 mM NaCl. GlgE-containing fractions were pooled and concentrated to ~1.5 mg/ml and aliquots were stored at −80 °C. An *M. smegmatis garA* construct was ligated into a pETPhos expression vector using *NdeI* and *BamHI* restriction sites [14]. *E. coli* BL21 (DE3) Star (Novagen) bearing this plasmid was grown at 30 °C to an optical density of 0.6 at 600 nm in LB and expression was induced with 0.2 mM IPTG. Bacteria were harvested and lyzed after a further 3 h incubation. The enzyme was purified using a 1 ml HisTrap FF column (GE Healthcare, Amersham, United Kingdom). The enzyme was dialysed into 20 mM Bis-Tris propane, pH 7.0, containing 150 mM NaCl and concentrated to 9 mg/ml and aliquots were stored at −80 °C.

2.2. Enzyme assays

Unless otherwise stated, all enzyme assays were carried out in 100 mM Bis-Tris propane, pH 7.0, containing 50 mM NaCl. GlgE activity was monitored at 21 °C using an end-point assay monitoring the quantitative release of inorganic phosphate with malachite green [15]. The concentration of free inorganic phosphate was estimated from a standard curve. α-Maltose-1-phosphate was synthesized as described earlier [10]. Reaction mixtures of 25 μl comprised 1 mM maltohexaose and 0.25 mM α-maltose-1-phosphate. Enzyme concentrations were such to allow reactions to progress linearly for 10 min with total donor consumption being <40%. Reactions were quenched with 175 μl of malachite green and incubated for 20 min at 21 °C before the absorbance at 630 nm was measured on a SpectraMAX Plus microplate spectrophotometer using SOFTmax PRO 3.1.1 software. The effect of pH was determined using 100 mM MES (pH 6.0), Bis-Tris (pH 6.5), Bis-Tris propane (pH 7.0), HEPES (pH 7.5) and Tris (pH 8.0) buffers. Acceptor preference used maltose, maltotriose, maltotetroase, maltopentaose, maltohexaose and

maltoheptaose at 1 mM each. The effect of salt was determined from 0 to 350 mM NaCl. Temperature de-activation of GlgE and GlgE H349Y was performed by pre-incubating protein at 21, 35, 40, 45 and 50 °C for 20 min and assaying at 21 °C. Temperature de-activation rescue experiments were performed as above, but in the presence of 50 μM *M. smegmatis* GarA. The effect of temperature on enzyme activity was determined by measuring initial rates (v_0/[E]) over a temperature range of 25 to 55 °C in 5 °C increments, with reaction mixtures of 40 μl comprising 1 mM maltohexaose and 1 mM α-maltose-1-phosphate. Enzyme kinetics at permissive (30 °C) and non-permissive (42 °C) temperatures were performed with maltohexoase concentrations between 1 and 150 mM and 1 mM α-maltose-1-phosphate. Initial rates were measured by quenching 6 μl reaction aliquots in 94 μl of 1 M HCl at time points from 1 to 8 min. The quenched reactions were incubated with 700 μl of malachite green assay solution for 20 min at 21 °C, and the absorbance at 630 nm was measured on a Perkin Elmer Lambda 18 or 25 spectrophotometer. Reaction rates were linear over at least the first 4 min. Enzyme kinetics for maltohexaose were fitted to a substrate inhibition model, as defined by equation, 1 using GraphPad Prism v8.

$$v_0/[E] = \frac{k_{cat}^{app}[S]}{K_m^{app} + [S]\left(1 + \frac{[S]}{K_i}\right)} \quad (1)$$

Eq. 1 is a simplified form derived of Eq. 2 [16], where A would be α-maltose-1-phosphate and B the acceptor substrate in the case of GlgE [17].

$$v_0/[E] = \frac{k_{cat}[A][B]}{K_{mB}[A] + K_{mA}[B]\left(1 + \frac{[B]}{K_{iB}}\right) + [A][B]} \quad (2)$$

Eq. 2 can rearranged to give Eq. 3.

$$v_0/[E] = \frac{\frac{k_{cat}[A][B]}{K_{mA}}}{\frac{K_{mB}[A]}{K_{mA}} + [B]\left(1 + \frac{[B]}{K_{iB}} + \frac{[A]}{K_{mA}}\right)} \quad (3)$$

When the initial concentration of A (α-maltose-1-phosphate) is fixed and the initial concentration of B (acceptor) is varied, Eq. 3 appears in the form of Eq. 1 such that k_{cat}^{app} and K_m^{app} are defined by Eqs. 4 and 5, respectively.

$$k_{cat}^{app} = \frac{k_{cat}[A]}{K_{mA}} \quad (4)$$

$$K_m^{app} = \frac{K_{mB}[A]}{K_{mA}} \quad (5)$$

The hydrolysis of α-maltose-1-phosphate was tested using *Aspergillus niger* amyloglucosidase from Megazyme (high purity for use in Megazyme Total Starch method) in 100 mM sodium acetate buffer, pH 4.0, at 25 °C.

2.3. Thermofluor assay

Assay mixtures of 50 μl containing 0.6 mg/ml protein, 20 mM Bis-Tris propane, pH 7.0 and SYPRO Orange protein stain dye (Sigma). Melting curves were recorded with a temperature range of 20 to 95 °C, using a DNA Engine Opticon 3 real time PCR system (MJ Research) with Opticon Monitor 3.1 analysis software (BioRad).

2.4. Circular dichroism

Experiments were performed using a Chirascan-Plus CD spectrophotometer (Applied Photo-physics). Purified proteins were dialysed into 10 mM sodium phosphate, pH 7.0. CD measurements were carried out in a quartz glass cell with a 0.5 mm path length at a ~ 0.3 mg/ml protein concentration. To obtain overall CD spectra, wavelength scans between 180 and 260 nm were collected at 20, 30 and 42 °C using a 1.0 nm bandwidth, a 0.5 nm step size, and a time per point of 1 s. Triplicate spectra were averaged, followed by the subtraction of buffer only control spectra. Time course spectra were recorded when the sample reached 42 °C and again two hours later. The raw data in millidegree units were corrected for background and converted to mean residue molar ellipticity. Secondary structure assignments were made using the DichroWeb server [18], using the CDSSTR algorithm and reference set 7 [19]. Thermal melt curves were determined with wavelength scans between 195 and 260 nm using 1 nm bandwidth, 1.0 nm step size and time per point of 0.7 s. The temperature was increased from 20 to 90 °C at a rate of 1 °C per min and a complete scan was collected at 1 °C intervals. The raw data in millidegree units, from 201 to 260 nm at each temperature measured, were analysed using the Global 3 software package (Applied Photophysics). The Global 3 package fitted the full-spectrum data using the nonlinear regression method of Marquardt and Levenberg [20] to generate a global analysis of unfolding. This procedure determines fitted and optimized temperatures of transition (melting points: T_m) and their associated Van't Hoff enthalpies (ΔH).

2.5. Protein-protein interaction studies

Surface plasmon resonance experiments were carried out on a Biacore T200 system (GE Healthcare). Immobilisation of either GarA or GlgE was attempted using both amine coupling (CM5 chip) and nickel affinity (NTA chip) approaches as per the manufacturer's instructions.

Fluorescence was recorded on a Perkin-Elmer LS55 fluorescence spectrometer connected to a PTP1 Peltier system. All assays were recorded using 100 mM Bis-Tris propane, pH 7.0, containing 50 mM NaCl. Anisotropy experiments used GlgE and its variants labelled with dansyl chloride in 50 mM HEPES, pH 9.0, containing 50 mM NaCl over 45 mins at 21 °C [21], in a manner used by us in other systems [22]. Labelled protein was then dialysed into 20 mM Tris, pH 8.5. Excitation and emission wavelengths were 335 and 515 nm, respectively. Experiments were carried out by titrating GarA into 1 μM labelled protein, up to a concentration of 630 μM at 25 and 42 °C. Separately, tryptophan fluorescence was measured with 1 μM H349Y and a GarA concentration up to 100 μM.

2.6. Bacterial strains, growth curves and metabolite analyses

Two strains of *M. smegmatis* mc^2155 were used. The parent strain was c-*glgE*-tet-off [5], which allowed the *glgE* gene to be silenced in the presence of anhydrotetracycline (ATc). A plasmid bearing the H349Y variant of the *glgE* gene, pMV361(Apra)::*glgE*-Ts, was introduced into this strain. Cells were grown aerobically at 30 or 42 °C in Middlebrook 7H9 medium supplemented with 0.5% *v*/v glycerol and 0.05% v/v tyloxapol and containing 10% v/v ADS enrichment (5% *w/v* bovine serum albumin fraction V (BSA), 2% w/v glucose, 0.85% w/v sodium chloride). Hygromycin (50 mg/l), kanamycin (20 mg/l) and apramycin (10 mg/l) were added for the selection of appropriate strains. Starter cultures of *M. smegmatis* strains with the appropriate antibiotic selection were grown at 30 °C in the presence or absence of ATc until they reached logarithmic phase. They were then diluted to an optical density of 0.3 at 600 nm and divided into two equal aliquots, one of which was grown at 30 °C and the other at 42 °C. When one of the cultures reached an optical density of ≥1.5 at 600 nm, they were each diluted to an OD of 1.5 (equating to 4.5 10^8 colony forming units per ml) and harvested by centrifugation, washed twice with phosphate buffered saline and resuspended in 2 ml of water. Soluble metabolites were released from cells by heating at 95 °C for 4 h, and cell debris was removed by centrifugation. The resulting supernatant was filtered, freeze-dried and re-suspended in 500 μl of D_2O. ^1H nuclear magnetic resonance spectra were recorded on a Bruker Avance III 400 MHz spectrometer at 22 °C and data were analysed using Topspin 3.5 software (Bruker Biospin Ltd).

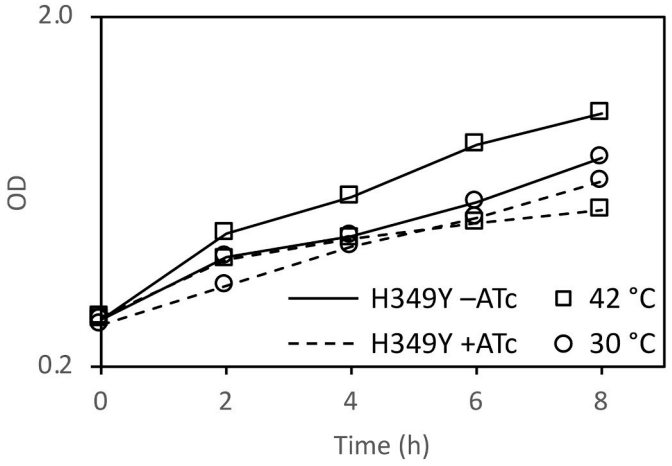


Fig. 2. The growth of *M. smegmatis* is compromised by the H349Y mutation of GlgE, particularly at 42 °C. The optical density of cultures was monitored at 600 nm in two independent experiments and the mean values were plotted on a log scale. The c-*glgE*-tet-off pMV361(Apra)::*glgE*-Ts strain grew more rapidly at 42 than 30 °C in the absence of ATc (annotated as H349Y –ATC), as expected for a wild-type-like strain with a functional GlgE. This strain in the presence of ATc (annotated as H349Y + ATC) grew more slowly than the wild-type strain at each respective temperature. The temperature-sensitive strain initially grew more rapidly at 42 °C than at 30 °C, but growth was arrested after 2 h of growth. These data with the reconstructed mutant faithfully recapitulate those reported for the isolated MSEG53 mutant [4].

NMR signals were quantified by comparing integrals with sodium 3-(trimethylsilyl)propionate-2,2,3,3-d_4 as an internal standard, which was also consistent with a standard curve generated with authentic trehalose.

3. Results

3.1. The temperature sensitive mutant accumulates α-maltose-1-phosphate at the non-permissive temperature

If the H349Y mutation in GlgE compromises enzyme activity, it would be expected that this would lead to the accumulation of α-maltose-1-phosphate [1]. To determine which small water-soluble metabolites accumulate in *M. smegmatis* as a result of the presence of the H349Y variant of GlgE at permissive and non-permissive temperatures, a gene silencing approach was used. The *glgE* gene has previously been brought under the control of a repressible promotor [5], allowing it to be robustly silenced in the presence of ATc. This was important because mutations in *glgE* are known to cause genetic instability over time because of toxicity associated with the accumulation of α-maltose-1-phosphate, particularly during growth in the presence of this metabolite's precursor trehalose. To generate the temperature-sensitive mutant, the H349Y variant of the *glgE* was introduced on a plasmid into the conditionally silenced strain. The addition of ATc to the growth medium would silence the wild-type copy making the H349Y variant the only GlgE present. Thus, it was possible to generate the temperature sensitive mutant synthetically, avoiding the possibility of additional mutations being introduced elsewhere in the genome during storage or manipulation. This would allow any phenotypes to be attributable to the single amino acid substitution in GlgE with confidence.

The two bacterial strains were then grown in the presence or absence of ATc at either the permissive or non-permissive temperatures of 30 and 42 °C, respectively. Growth kinetics confirmed that the c-glgE-tet-off strain expressing the H349Y variant of the *glgE* gene in the presence of ATc (Fig. 2) faithfully reproduced the temperature-sensitive growth phenotype reported for the SMEG53 strain [4]. This shows that the H349Y mutation of the *glgE* gene is sufficient to result in temperature-sensitivity at the cellular level.

Other phenotypes associated with SMEG53, such as altered colony morphology and partial rescue by osmolytes [4], were not observed in the reconstructed mutant strain in the presence of ATc. Since SMEG53 was isolated after treatment with the mutagen nitrosoguanidine and mutations in *glgE* are known to lead to genetic instability, it is likely that mutations in genes other than *glgE* might have contributed to the other phenotypes.

The soluble metabolites were then extracted from the strains grown under the different conditions and analysed using ^{1}H NMR spectroscopy. Whenever ATc was absent, the wild-type GlgE would be expected to be present. It would therefore be expected that no α-maltose-1-phosphate would accumulate [1] at any temperature, whether or not the H349Y

Table 1
The temperature sensitive strain accumulates α-maltose-1-phosphate at the non-permissive temperature. The *glgE* gene in *M. smegmatis* was silenced with ATc using a tet-off system. The temperature sensitive H349Y variant of *glgE* was introduced into the parent strain on a plasmid. Each of the two strains were grown at either 30 or 42 °C in either the absence or presence of ATc. The expected presence of each variant of the GlgE enzyme is indicated in each case. The level of α-maltose-1-phosphate (M1P) was determined using ^{1}H NMR spectroscopy. The values represent means of two independent experiments (with a limit of detection of 2.5 μg/10^9 colony forming units (CFU)). See also Fig. 3.

Strain	T	ATc	GlgE present	M1P μg/10^9 CFU
	°C			
c-*glgE*-tet-off	30	–	wild-type	< 2.5
c-*glgE*-tet-off	30	+	none	20.2
c-*glgE*-tet-off	42	–	wild-type	< 2.5
c-*glgE*-tet-off	42	+	none	25.5
c-*glgE*-tet-off pMV361(Apra)::*glgE*-Ts	30	–	wild-type + H349Y	< 2.5
c-*glgE*-tet-off pMV361(Apra)::*glgE*-Ts	30	+	H349Y	< 2.5
c-*glgE*-tet-off pMV361(Apra)::*glgE*-Ts	42	–	wild-type + H349Y	< 2.5
c-*glgE*-tet-off pMV361(Apra)::*glgE*-Ts	42	+	H349Y	8.2

Table 2
Kinetics of wild-type and H349Y GlgE. Enzyme activity was monitored by detecting phosphate release. Assays were performed in duplicate and fitted values are expressed with standard errors. Quoted constants are apparent because the enzyme obeys ping-pong (substituted enzyme) kinetics [1]. See also Fig. 8.

GlgE	T	k_{cat}^{app}	K_m^{app}	K_i	k_{cat}^{app}/K_m^{app}
	°C	s^{-1}	mM	mM	M^{-1} s^{-1}
wild-type	30	78.9 ± 6.0	9.7 ± 1.5	126 ± 24	8100 ± 1400
wild-type	42	112 ± 14	14.4 ± 3.5	187 ± 67	7800 ± 2100
H349Y	30	15.0 ± 3.0	10.0 ± 3.8	95 ± 43	1500 ± 640
H349Y	42	14.5 ± 1.3	12.5 ± 2.3	366 ± 140	1160 ± 240

Table 3
Net secondary structures of wild-type and H349Y *M. smegmatis* GlgE. The values were determined using circular dichroism spectroscopy (see Fig. 9). Values are means of triplicate measurements. Note, the elements add up to either 99, 100 or 101 depending on rounding introduced by the software.

GlgE	T	t	% helix	% sheet	% turns	% unstructured
	°C	h				
wild-type	25	0	22	30	21	28
wild-type	42	0	21	28	20	30
wild-type	42	2	20	32	20	29
H349Y	25	0	23	26	20	31
H349Y	42	0	20	31	20	29
H349Y	42	2	18	30	21	30

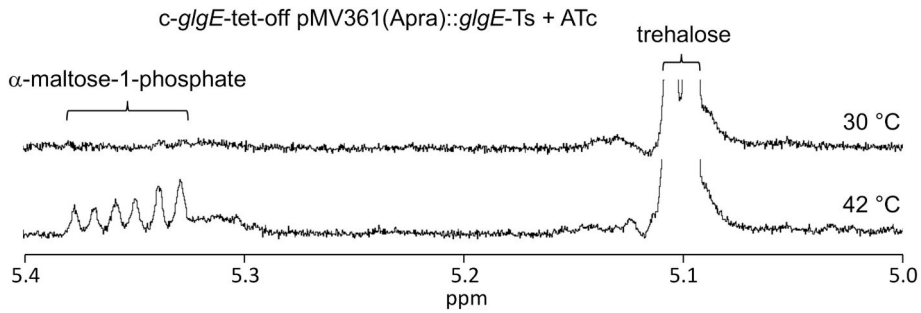


Fig. 3. The H349Y mutation of GlgE leads to the accumulation of α-maltose-1-phosphate when *M. smegmatis* is grown at 42 °C. Representative ^{1}H NMR spectra out of two biological replicates are shown of extracts of the c-*glgE*-tet-off pMV361(Apra)::*glgE*-Ts strain grown in the presence of ATc at either 30 or 42 °C. See also Table 1.

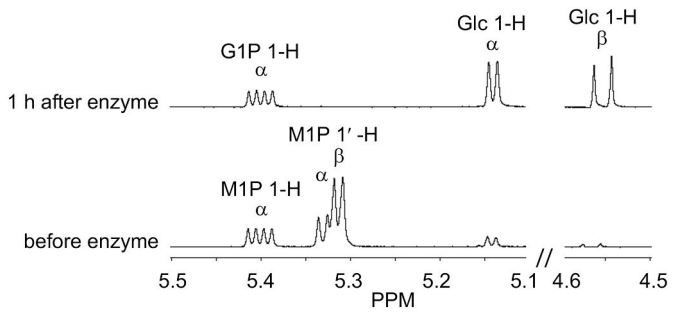


Fig. 4. *A. niger* amyloglucosidase cleaves the α-1,4-linked non-reducing end glucose from both α and β-maltose-1-phosphate. An anomeric mixture of α/β-maltose-1-phosphate was incubated with amyloglucosidase. The reaction was monitored using ^{1}H nuclear magnetic resonance spectroscopy and representative spectra are shown. The peaks for maltose-1-phosphate (M1P), glucose-1-phosphate (G1P) and glucose (Glc) were assigned as previously reported [10,23]. The 1′-H resonances of α/β-maltose-1-phosphate associated with the α-glucosyl-(1 → 4)-glucose linkage are lost at the same time as those of the 1-H reducing end of glucose appear. The 1-H resonances of α-maltose-1-phosphate and α-glucose-1-phosphate are essentially coincident. They are also slightly down-field compared with Fig. 3 because the solution was buffered at pH 4.0. Note that the 1-H resonances from β-maltose-1-phosphate and β-glucose-1-phosphate are not shown for clarity.

variant of GlgE was present. This was indeed the case (Table 1). The only other small soluble carbohydrate metabolite that accumulated was trehalose, which was present in all strains including the wild-type, as reported previously. Interestingly, the level of trehalose increased at the higher temperature (e.g. from typically 29 to 46 μg/10^{9} colony forming units). When ATc was added in the absence of the H349Y variant, no GlgE would be expected to be present. This would lead to the accumulation of α-maltose-1-phosphate (to non-lethal levels because no trehalose was added to the grown medium [1]). Accumulation indeed occurred, particularly at the higher temperature when more of the precursor, trehalose, was available. Finally, when ATc was added in the presence of the H349 variant, the mutated H349Y version of GlgE would be expected to be the only form present. When this strain was grown at 30 °C, no significant amount of α-maltose-1-phosphate was detected. However, a moderate amount of α-maltose-1-phosphate had clearly accumulated at 42 °C. One can therefore conclude that the temperature sensitivity of the H349Y GlgE strain led to the accumulation of α-maltose-1-phosphate implying GlgE was compromised by the mutation at the higher temperature, but not completely absent.

3.2. α-Maltose-1-phosphate is hydrolysed by amyloglucosidase

An anomeric mixture of α/β-maltose-1-phosphate was incubated with *A. niger* amyloglucosidase. Both anomers were efficiently degraded by the enzyme to yield glucose-1-phosphate and glucose (Fig. 4). As expected for this type of enzyme (EC 3.2.1.3), the β anomer of glucose was liberated initially, which subsequently reaches an equilibrium with the α anomer within an hour or so. It is therefore clear that any α-maltose-1-phosphate that accumulates in *M. smegmatis* would be degraded by this enzyme to yield glucose.

3.3. The H349Y GlgE mutation leads to a loss of enzyme activity and an increased temperature sensitivity

If the activity of GlgE is compromised by the H349Y mutation, it could manifest itself through a loss of either activity or stability. First, the kinetics of the extension of maltohexaose with α-maltose-1-phosphate by wild-type and H349Y GlgE were each monitored by the release of inorganic phosphate. It was immediately apparent that the activity of H349Y GlgE was lower than that of the wild-type enzyme. Nevertheless, the activity profiles of the two enzymes as a function of either pH or NaCl concentration were similar (Fig. 5A and B). Specifically, both enzymes showed optimal activity at pH 7.0, and NaCl up to 350 mM had little effect on either enzyme. Each enzyme had an acceptor preference for maltohexaose or maltoheptaose (Fig. 5C) with little or no difference between their acceptor profiles. The mutation therefore had little or no effect on these properties of GlgE despite a reduction in overall activity.

To ascertain what effect temperature had on the wild-type and H349Y GlgE, their temperature stabilities were determined by pre-incubating each at a range of temperatures before assaying them at 21 °C. The wild-type GlgE enzyme showed a 24 to 23% drop in activity after exposure to temperatures of 40 and 45 °C, respectively (Fig. 6). By contrast, the activity of H349Y GlgE was more severely affected with drops in activity of 34 and 71% at 40 and 45 °C, respectively. From this trend, it is possible to estimate a 50% drop would have occurred at 42 °C. One can conclude that the H349Y mutation roughly doubled the rate of deactivation at 42 °C from ~24% to ~50% over a period of 20 min.

Over-expression of GarA in vivo resulted in a partial rescue of the temperature sensitive phenotype of SMEG53 cells [4] and GarA has been demonstrated to bind to and regulate not only phosphorylated but also non-phosphorylated targets [24]. To test whether GarA can directly protect H349Y GlgE from temperature de-activation, H349Y was pre-incubated with GarA. However, GarA showed no significant protective effect (Fig. 6). Indeed, it seemed to somewhat further destabilise H349Y GlgE.

The effect of temperature on GlgE activity was also determined. Enzyme activity was linearly dependent upon time over the 10 min period of the assay in all cases except for assays with the H349Y variant at 45 °C and above when initial rates had to be estimated. The wild-type enzyme showed an increase in activity between 25 and 45 °C, followed by a modest decrease at 50 and 55 °C (Fig. 7). By contrast, H349Y GlgE was considerably less active, and showed only a modest increase in activity between 25 and 40 °C, followed by a complete loss of activity at 50 °C and above, consistent with the temperature stability experiments. In contrast to the wild-type enzyme, there was no significant difference

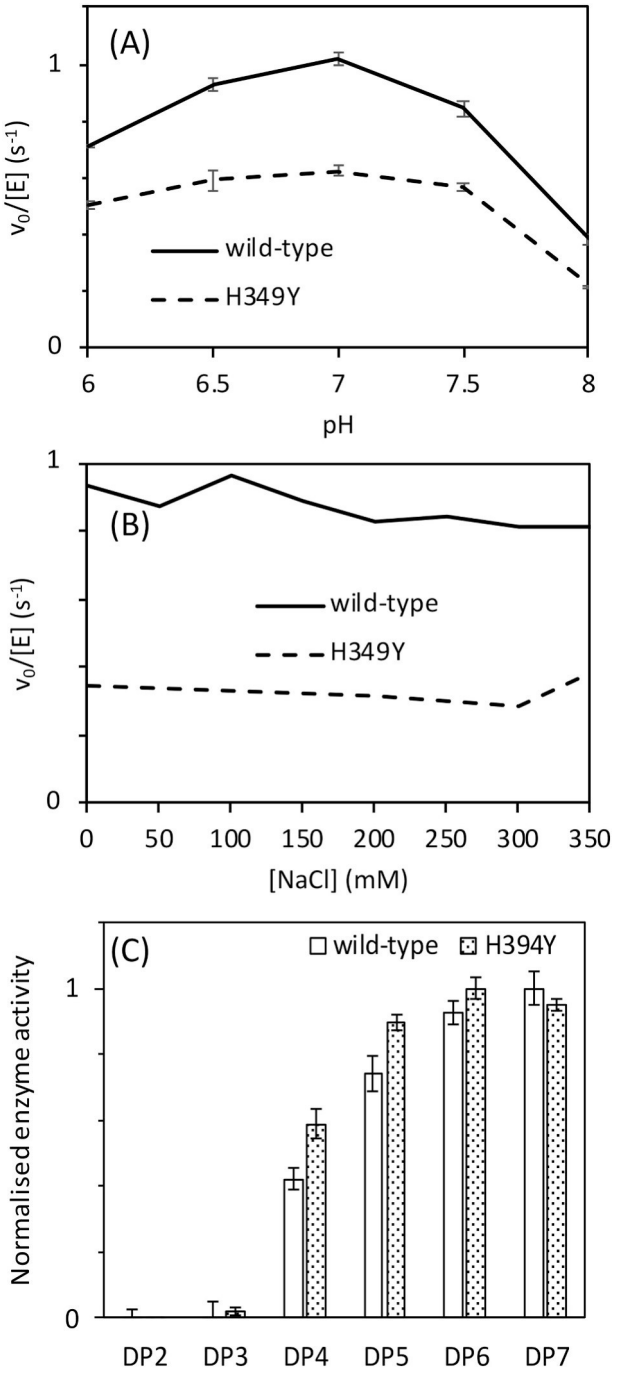


Fig. 5. The H349Y mutation changes neither the pH optimum, salt-sensitivity nor acceptor specificity of *M. smegmatis* GlgE at 21 °C. Experiments were carried out using 0.25 mM α-maltose-1-phosphate and 1 mM acceptor. (A) shows the pH optima of wild-type (solid line) and H349Y (broken line) GlgE to be 7. Error bars represent standard errors (which are very small) from triplicate experiments. (B) shows the tolerance of wild-type (solid line) and H349Y (broken line) GlgE to NaCl up to 350 mM. (C) shows the acceptor preferences of wild-type (plain bars) and H349Y (shaded bars) GlgE are very similar. The acceptors tested ranged from maltose (degree of polymerisation (DP) of 2) up to maltoheptaose (DP of 7). Means are normalised to the highest value for each enzyme to aid comparison. The error bars represent standard errors from triplicate experiments.

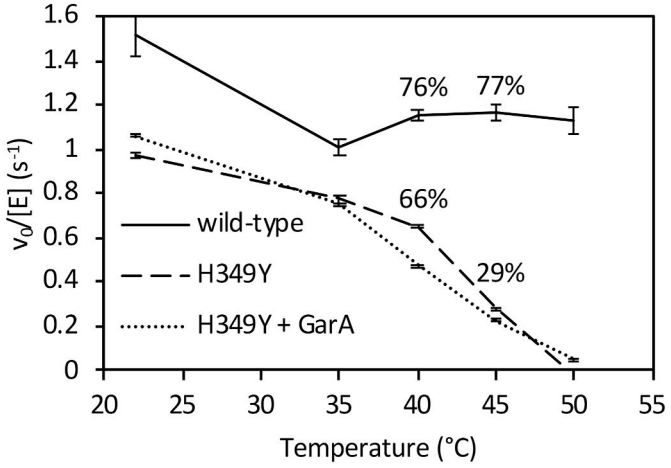


Fig. 6. Temperature deactivation of wild-type and H349Y GlgE enzyme activity. Residual activity of wild-type (solid line) and H349Y GlgE in the absence (broken) or presence of GarA (dotted line) after pre-incubation at an elevated temperature for 20 min before assaying at 21 °C. Error bars represent standard errors of triplicate experiments. Percentage deactivation at 40 and 45 °C are indicated for wild-type and H349Y enzymes compared with their respective activities at 22 °C.

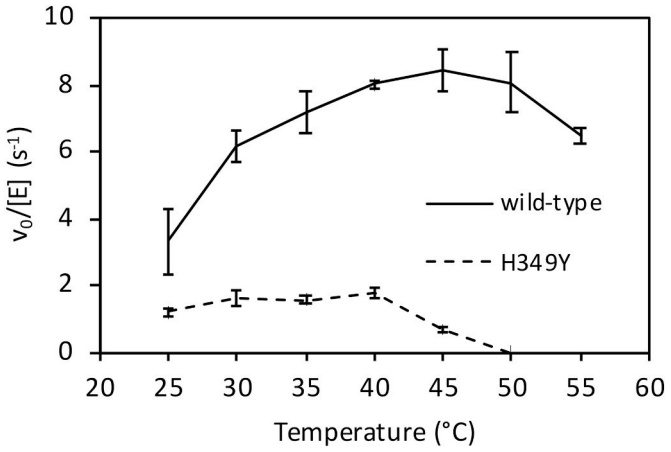


Fig. 7. The effect of temperature on the enzyme activities of wild-type and H349Y GlgE. The activity of wild-type (solid line) and H349Y (broken) GlgE as a function of temperature was determined by monitoring phosphate release. Error bars represent standard errors of triplicate experiments.

in the activity of the H349Y GlgE at 30 and 40 °C. Under these conditions, there were therefore ~3.5-fold and ~ 4-fold differences in turnover rates between the two enzymes at 30 and 40 °C, respectively.

To address whether the impaired activity of H349Y GlgE was due to a k_{cat} or K_m effect, the kinetic parameters with maltohexaose were measured for each enzyme at permissive and non-permissive temperatures for bacterial growth. Assays were conducted over a maximum of 10 min, minimising any impact of thermoinstability. The data exhibited substrate inhibition, so were fitted to a substrate inhibition model (Fig. 8 and Table 2). Substrate inhibition can be rationalised by high concentrations of maltohexaose promoting a competing disproportion reaction [1,10]. With the wild-type enzyme, both the k_{cat}^{app} and K_m^{app} for maltohexaose increased slightly at the higher temperature, resulting in a similar k_{cat}^{app}/K_m^{app} value (Table 2). The values of K_m^{app} of H349Y GlgE for maltohexaose at each temperature were very similar to those of the wild-type enzyme. Given that K_m^{app} is a function of the true K_m values for both substrates (Eq. 5), the affinity of neither substrate appears to have changed. By contrast, the values of k_{cat}^{app}, and therefore k_{cat}^{app}/K_m^{app}, were at

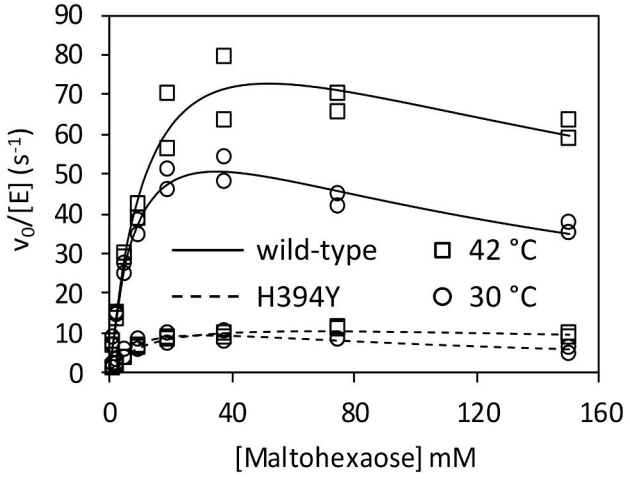


Fig. 8. Kinetics of maltohexaose extension with α-maltose-1-phosphate by wild-type and H349Y GlgE. The kinetics were determined with enzyme concentrations between 12.5 and 100 nM and 1 mM α-maltose-1-phosphate at either 42 (squares) or 30 °C (circles) with wild-type (solid line) or H349Y (broken line). Experiments were carried out in duplicate. The lines of best fit adhere to a substrate inhibition model. See Table 2 for the derived kinetic constants.

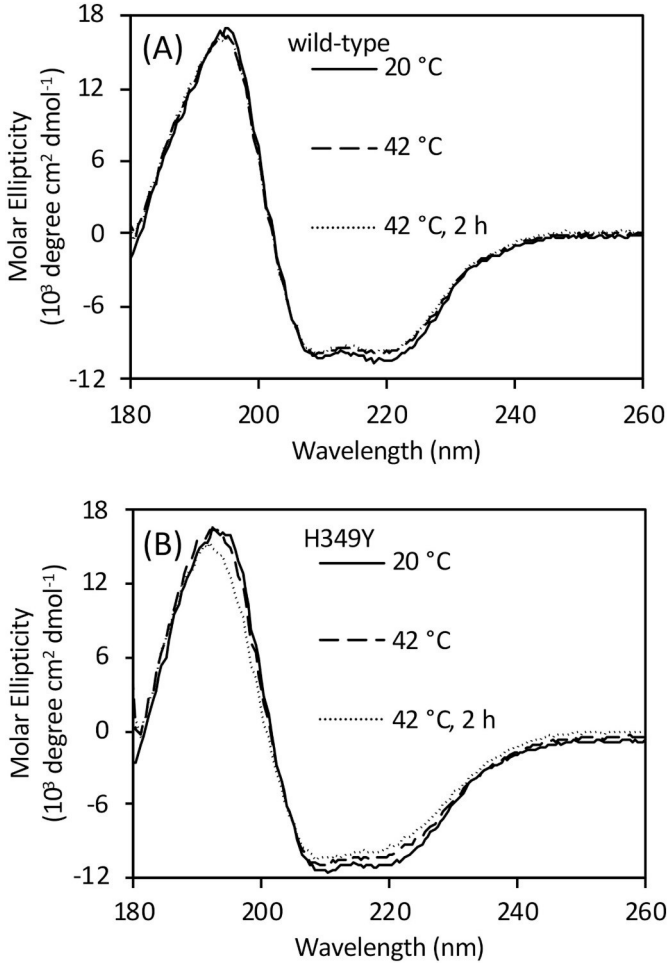


Fig. 9. Circular dichroism spectra of wild-type (panel A) and H349Y (panel B) GlgE. Spectra were recorded at either 20 or 42 °C, as indicated. An additional spectrum was recorded for each enzyme after 2 h of incubation at 42 °C. See Table 3 for the corresponding analysis of net secondary structures.

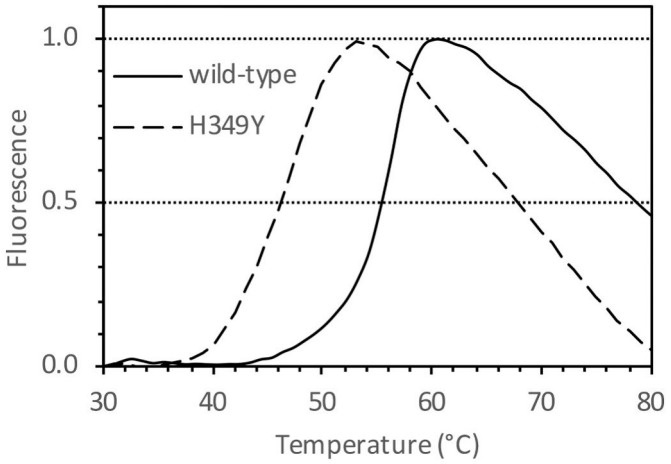


Fig. 10. Thermofluor analysis of wild-type and H349Y GlgE. The proteins were each heated in the presence of CYPRO Orange. Fluorescence (arbitrary units) is shown normalised to the maximum signal of each sample. The melting temperatures for wild-type (solid line) and H349Y (broken line) GlgE were 56 and 47 °C, respectively, as highlighted by the horizontal guideline at 0.5. The normalised fluorescence values at 42 °C for wild-type and H349Y GlgE were 0.01 and 0.16, respectively.

five to seven-fold lower, consistent with the above experiments. Since k_{cat}^{app} is a function of the true value of k_{cat} (Eq 4), the H349Y mutation affects k_{cat} rather than the affinities of either substrate. Importantly, the k_{cat}^{app}/K_m^{app} value for H349Y GlgE was about a third lower at the non-permissive vs the permissive temperature.

3.4. The H349Y mutation reduces thermal stability but has limited impact at the non-permissive temperature

Circular dichroism spectroscopy was used to probe the secondary structure and temperature stability of wild-type and H349Y GlgE. The two forms of GlgE had very similar net secondary structure at 25 and 42 °C (Fig. 9 and Table 3). After 2 h of incubation at 42 °C, the H349Y enzyme showed a shift in the spectrum compared with the wild-type enzyme (Fig. 9). However, this shift was very small, equating to changes in secondary structural elements of ≤2% each (Table 3). Melting temperatures were determined by monitoring the loss of secondary structure as a function of temperature. The H349Y mutation led to a 7.6 °C destabilisation of the GlgE protein from 55.7 ± 0.2 to 48.1 ± 0.3 °C. Consistent with this observation, a Thermofluor assay that follows the thermal denaturation of proteins via binding of a fluorescent dye was in close agreement, showing a 9 °C destabilisation from 56 to 47 °C (Fig. 10). At 42 °C, the amount of melting of the wild-type and H349Y enzymes was 1 and 16%, respectively. Therefore, no gross denaturation of the protein occurred at the non-permissive temperature for bacterial growth, consistent with only partial loss of activity for the H349Y enzyme at 42 °C (Fig. 6).

3.5. Neither GlgE nor phospho-GlgE bind to GarA

Although GarA was shown above not to rescue the temperature sensitivity of the H349Y mutation, it was not possible to rule out a protein-protein interaction between the two. We therefore employed several methods to explore whether any interaction could occur. A surface plasmon resonance approach was not possible because neither protein bound successfully to a CM5 sensor chip at all values of pH tested (4.0 and 4.5). Similarly, only non-specific binding was observed when attempting to use an NTA sensor chip to immobilise each His-tagged protein. We therefore took another approach. GarA has no tryptophan residues. By contrast, GlgE has intrinsic fluorescence with its 17 tryptophan residues, which could change when another protein binds to

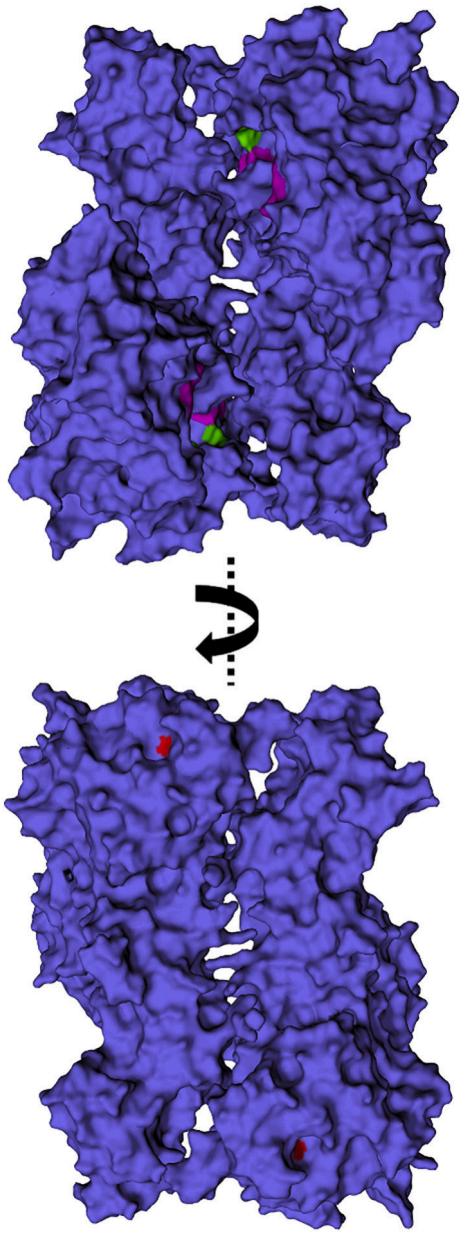


Fig. 11. The location of the substituted His residue on GlgE. The structure of *Streptomyces coelicolor* GlgE isoform I is shown with surface representation (PDB entry 5LGV) [25]. This structure is of a Glu423Ala variant with a maltooligosaccharide bound to both the donor and acceptors sites (with this ligand omitted for clarity). The equivalent residue to His349 in the *M. smegmatis* enzyme is His329 in the *S. coelicolor* enzyme, which is highlighted in red and is visible in the lower panel. Representative residues associated with the binding site of α-maltose-1-phosphate (catalytic residues Glu423Ala and Asp394, together with Gln324 and Asn352; *S. coelicolor* numbering) are highlighted in green, while some of those associated with the binding site of the acceptor (His397, Trp448 and Arg449) are highlighted in magenta. All of these residues associated with the binding of substrates are visible in the upper panel.

GlgE. However, the intrinsic fluorescence of GlgE (at 1 μM) did not change in the presence of GarA (up to 100 μM). In separate experiments, the labelling of GlgE with dansyl groups allowed the measurement of fluorescence anisotropy. This signal would change when the tumbling rate of GlgE slows on binding another protein. However, no such change was detected in the presence of GarA. It was not possible to repeat the experiment with unlabelled GlgE and labelled GarA due to protein solubility limitations. Nevertheless, given the sensitivity of fluorescence anisotropy, one can conclude that GlgE and GarA do not interact in the conditions tested.

4. Discussion

We have shown that the temperature-sensitive mutation in the *glgE* gene leads to the accumulation of α-maltose-1-phosphate at the non-permissive temperature (Table 1 and Fig. 3). This is entirely consistent with the phenotype of other mutations in mycobacterial *glgE* reported recently [1,5], where a loss of GlgE activity leads to the accumulation of α-maltose-1-phosphate and the slowing of bacterial growth. However, this contrasts with the 1999 study [4], where it was concluded that the accumulation of glycogen led to the cessation of bacterial growth. It is easy to reinterpret the 1999 study with hindsight, but the conclusions drawn at the time were entirely reasonable. First, GlgE was predicted to be a glucanase and second, the only known potential precursors of glycogen at that time were UDP-glucose or ADP-glucose, which would not have been susceptible to hydrolysis by amyloglucosidase. By contrast, the recently discovered α-maltose-1-phosphate precursor possesses a glucosyl-(1 → 4)-glucose linkage that is susceptible to hydrolysis, leading to the liberation of glucose. Since the standard glycogen assay involves the detection of glucose liberated by amyloglucosidase, α-maltose-1-phosphate would have given a false positive result. It therefore follows that the interpretation of the observations in 1999 that led to the conclusion that glycogen recycling was essential for growth were based on an assumption that we now know to be incorrect. It therefore remains to be seen whether glycogen recycling is indeed essential for growth.

The basis for the temperature-sensitivity of the H349Y mutation appears not to be a simple case of protein instability at the non-permissive temperature. Relative to the wild-type enzyme, the H349Y enzyme lost only ~half of its activity after 20 min of incubation at 42 °C (Fig. 6) with only minimal changes in net secondary structure after 2 h (Fig. 9 and Table 3) and only 15% more melting (Fig. 10). Rather, it appears to be also down to the loss of activity of the enzyme at the non-permissive temperature caused by a ~ 7-fold reduction in the k_{cat}^{app} relative to the wild-type enzyme (Fig. 8 and Table 2). This was due to a much smaller intrinsic value of k_{cat}^{app} (Fig. 8 and Table 2) coupled with a much smaller apparent temperature coefficient than the wild-type enzyme (Fig. 7) leading to a smaller k_{cat}^{app} for the H349Y variant at the higher vs lower temperature. Since there was a ~ 5-fold reduction in activity caused by the mutation at the permissive temperature (Fig. 8), this implies that adequate flux through the GlgE pathway is on a knife-edge somewhere between the permissive and non-permissive temperatures. At the higher temperature, the accumulation of α-maltose-1-phosphate is evidently correlated with reduced bacterial growth-rate, consistent with all previous studies [1,5].

The H349Y mutation of GlgE maps onto the surface of the protein (Fig. 11) that is on the opposite side of the molecule to the substrate binding sites [10,25]. The closest residue that forms part of the substrate binding sites is within five amino acids within the sequence. Note that this His residue is highly conserved in GlgE enzyme across mycobacteria, including *M. tuberculosis,* and streptomycetes. Quite how the substitution affects enzyme activity is therefore not clear, which mirrors our lack of understanding of how phosphorylation lowers activity [14].

GarA has been reported to rescue the temperature-sensitive mutation by reducing the accumulation of carbohydrate [4]. All attempts to identify an interaction between GarA and GlgE suggested no such interaction occurs. In addition, GarA did not strongly influence the temperature stability of GlgE (Fig. 6). Both GarA and GlgE are substrates of the phosphorylase PknB. Phosphorylation of GarA precludes its binding to other proteins. Phosphorylation of GlgE leads to its negative regulation, primarily through a lowering of k_{cat}^{app} by up to two orders of magnitude [14]. Perhaps GarA binds to phosphorylated GlgE and stabilises it. However, preliminary anisotropy experiments with GarA and GlgE phosphorylated by PknB indicated no interaction (data not shown).

In any case, phosphorylated GlgE has very low activity, so its stabilisation would not recover flux through the enzyme. The most likely explanation for the ability of GarA to rescue the mutation is therefore that it acts as a decoy at high concentrations, preventing the negative regulation of GlgE. It therefore follows that GlgE would normally be somewhat negatively regulated by PknB at the non-permissive temperature, perhaps to counter its temperature-dependent increase in activity (Fig. 7). It could therefore be the combination of negative regulation, thermo-instability and a reduced k_{cat}^{app} that breaks the threshold for tolerance of α-maltose-1-phosphate toxicity in the mutant at the higher temperature leading to a lower rate of bacterial growth.

5. Conclusions

In this study, we have shown that it is possible to reconcile the original report that GlgE is a glucanase [4] with what is now known about the biosynthesis of glycogen in mycobacteria [5]. When GlgE is compromised by the H349Y amino acid substitution, α-maltose-1-phosphate accumulates to toxic levels due to a partial loss of polymerase activity, particularly at an elevated temperature, coupled with a loss of stability. This implies the regulation of GlgE is delicately balanced to control flux through the glycogen biosynthetic pathway without compromising growth. This increases the attractiveness of the GlgE pathway as a potential drug target [26].

Author contributions

KS carried out the enzymology. SFDB carried out the bacterial growths and metabolite analyses. SS and RK constructed the microbial strains and also carried out bacterial growths. SB and RK conceived the project and supervised the work. SB wrote the manuscript from a draft written by KS with subsequent feedback from all authors.

Funding

This work was supported by the United Kingdom Biotechnology and Biological Sciences Research Council (Doctoral Training Partnership [BB/J014524/1] and Institute Strategic Programme [BB/J004561/1] grants) and the John Innes Foundation. R.K acknowledges support from the Jürgen Manchot Foundation.

Declaration of Competing Interest

The authors declare that no conflict of interest exists.

Acknowledgements

We thank Virginie Molle for the *garA* and *pknB* expression vectors and for feedback on the manuscript.

References

[1] R. Kalscheuer, K. Syson, U. Veeraraghavan, B. Weinrick, K.E. Biermann, Z. Lui, J. C. Sacchettini, G. Besra, S. Bornemann, W.R. Jacobs, Self-poisoning of *Mycobacterium tuberculosis* by targeting GlgE in an a-glucan pathway, Nat. Chem. Biol. 6 (2010) 376–384, https://doi.org/10.1038/nchembio.340.

[2] A.D. Elbein, I. Pastuszak, A.J. Tackett, T. Wilson, Y.T. Pan, Last step in the conversion of trehalose to glycogen: a mycobacterial enzyme that transfers maltose from maltose 1-phosphate to glycogen, J. Biol. Chem. 285 (2010) 9803–9812, https://doi.org/10.1074/jbc.M109.033944.

[3] A. Rashid, S.F.D. Batey, K. Syson, H. Koliwer-Brandl, F. Miah, E.J. Barclay, K. C. Findlay, K.P. Nartowski, Y.Z. Khimyak, R. Kalscheuer, S. Bornemann, The assembly of a-glucan by GlgE and GlgB in mycobacteria and streptomyces, Biochemistry 55 (2016) 3270–3284.

[4] A.E. Belanger, G.F. Hatfull, Exponential-phase glycogen recycling is essential for growth of *Mycobacterium smegmatis*, J. Bacteriol. 181 (1999) 6670–6678.

[5] H. Koliwer-Brandl, K. Syson, R. van de Weerd, G. Chandra, B. Appelmelk, M. Alber, T.R. Ioerger, W.R. Jacobs, J. Geurtsen, S. Bornemann, R. Kalscheuer, Metabolic network for the biosynthesis of intra- and extracellular α-glucans required for virulence of *Mycobacterium tuberculosis*, PLoS Pathog. 12 (2016), e1005768.

[6] J. Preiss, T. Romeo, Physiology, biochemistry and genetics of bacterial glycogen synthesis, Adv. Microb. Physiol. 30 (1989) 183–238.

[7] P. England, A. Wehenkel, S. Martins, S. Hoos, G. André-Leroux, A. Villarino, P. M. Alzari, The FHA-containing protein GarA acts as a phosphorylation-dependent molecular switch in mycobacterial signaling, FEBS Lett. 583 (2009) 301–307, https://doi.org/10.1016/j.febslet.2008.12.036.

[8] V. Lombard, H. Golaconda Ramulu, E. Drula, P.M. Coutinho, B. Henrissat, The carbohydrate-active enzymes database (CAZy) in 2013, Nucleic Acids Res. 42 (2014) D490–D495, https://doi.org/10.1093/nar/gkt1178.

[9] M.R. Stam, E.G.J. Danchin, C. Rancurel, P.M. Coutinho, B. Henrissat, Dividing the large glycoside hydrolase family 13 into subfamilies: towards improved functional annotations of a-amylase-related proteins, Protein Eng. Des. Sel. 19 (2006) 555–562, https://doi.org/10.1093/protein/gzl044.

[10] K. Syson, C.E.M. Stevenson, M. Rejzek, S.A. Fairhurst, A. Nair, C.J. Bruton, R. A. Field, K.F. Chater, D.M. Lawson, S. Bornemann, Structure of a *Streptomyces* maltosyltransferase GlgE: a homologue of a genetically validated anti-tuberculosis target, J. Biol. Chem. 286 (2011) 38298–38310, https://doi.org/10.1074/jbc.M111.279315.

[11] J. Preiss, Glycogen biosynthesis, in: M. Schaechter (Ed.), The Encyclopedia of Microbiology Vol. 5, Elsevier, Oxford, U. K., 2009, pp. 145–158.

[12] K. Syson, C.E.M. Stevenson, D.M. Lawson, S. Bornemann, Structure of the *Mycobacterium smegmatis* a-maltose-1-phosphate synthase GlgM, Acta Crystallogr. F 76 (2020) 175–181, https://doi.org/10.1107/S2053230X20004343.

[13] T. Sambou, P. Dinadayala, G. Stadthagen, N. Barilone, Y. Bordat, P. Constant, F. Levillain, O. Neyrolles, B. Gicquel, A. Lemassu, M. Daffé, M. Jackson, Capsular glucan and intracellular glycogen of *Mycobacterium tuberculosis*: biosynthesis and impact on the persistence in mice, Mol. Microbiol. 70 (2008) 762–774, https://doi.org/10.1111/j.1365-2958.2008.06445.x.

[14] J. Leiba, K. Syson, G. Baronian, I. Zanella-Cleon, R. Kalscheuer, L. Kremer, S. Bornemann, V. Molle, *Mycobacterium tuberculosis* maltosyltransferase GlgE, a genetically validated anti-tuberculosis target, is negatively regulated by Ser/Thr phosphorylation, J. Biol. Chem. 288 (2013) 16546–16556, https://doi.org/10.1074/jbc.M112.398503.

[15] P.A. Lanzetta, L.J. Alvarez, P.S. Remack, O.A. Candia, An improved assay for nanomole amounts of inorganic phosphate, Anal. Biochem. 100 (1979) 95–97.

[16] A. Cornish-Bowden, Fundamentals of Enzyme Kinetics, Third ed., Portland Press, London, 2004.

[17] K. Syson, C.E.M. Stevenson, A.M. Rashid, G. Saalbach, M. Tang, A. Tuukkanen, D. I. Svergun, S.G. Withers, D.M. Lawson, S. Bornemann, Structural insight into how *Streptomyces coelicolor* maltosyl transferase GlgE binds α-maltose 1-phosphate and forms a maltosyl-enzyme intermediate, Biochemistry 53 (2014) 2494–2504, https://doi.org/10.1021/bi500183c.

[18] L. Whitmore, B.A. Wallace, DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data, Nucleic Acids Res. 32 (2004) W668–W673, https://doi.org/10.1093/nar/gkh371.

[19] N. Sreerama, R.W. Woody, Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set, Anal. Biochem. 287 (2000) 252–260, https://doi.org/10.1006/abio.2000.4880.

[20] D.W. Marquardt, An algorithm for least-squares estimation of nonlinear parameters, J. Soc. Ind. Appl. Maths. 11 (1963) 431–441, https://doi.org/10.1137/0111030.

[21] T. Meyer, P.I. Hanson, L. Stryer, H. Schulman, Calmodulin trapping by calcium-calmodulin-dependent protein kinase, Science 256 (1992) 1199–1202, https://doi.org/10.1126/science.256.5060.1199.

[22] J.B. Miller, A. Pratap, A. Miyahara, L. Zhou, S. Bornemann, R. Morris, G.E. D. Oldroyd, CCaMK is negatively and positively regulated by calcium, providing a novel mechanism for decoding calcium responses, Plant Cell 25 (2013) 5053–5066, https://doi.org/10.1105/tpc.113.116921.

[23] F. Miah, M.J. Bibb, J.E. Barclay, K.C. Findlay, S. Bornemann, Developmental delay in a *Streptomyces venezuelae glgE* null mutant is associated with the accumulation of α-maltose 1-phosphate, Microbiology 162 (2016) 1208–1219.

[24] T. Alber, Signaling mechanisms of the *Mycobacterium tuberculosis* receptor Ser/Thr protein kinases, Curr. Opin. Struct. Biol. 19 (2009) 650–657.

[25] K. Syson, C.E. Stevenson, F. Miah, J.E. Barclay, M. Tang, A. Gorelik, A.M. Rashid, D.M. Lawson, S. Bornemann, Ligand-bound structures and site-directed mutagenesis identify the acceptor and secondary binding sites of *Streptomyces coelicolor* maltosyltransferase GlgE, J. Biol. Chem. 291 (2016) 21531–21540, https://doi.org/10.1074/jbc.M116.748160.

[26] R. Kalscheuer, W.R. Jacobs, The significance of GlgE as a new target for tuberculosis, Drug News Perspect. 23 (2010) 619–624, https://doi.org/10.1358/dnp.2010.23.10.1534855.

7. The mycolic acid reductase Rv2509 has distinct structural motifs and is essential for growth in slow-growing mycobacteria

Published in: Molecular Microbiology

Impact Factor: 3.816 (2017)

DOI: 10.1111/mmi.14437

Overall contribution to the paper:

- Generation of *Mycobacterium bovis* mutants
- Conduction of growth experiments
- Preparation of Supplementary Material (Suppl. Fig.6)
- Interpretation of the data obtained
- Feedback on the manuscript draft

Received: 8 October 2018 | Revised: 11 November 2019

DOI: 10.1111/mmi.14437

RESEARCH ARTICLE

WILEY

The mycolic acid reductase Rv2509 has distinct structural motifs and is essential for growth in slow-growing mycobacteria

Asma Javid[1] | Charlotte Cooper[1] | Albel Singh[1] | Steffen Schindler[2] | Milena Hänisch[2] | Robert L. Marshall[1] | Rainer Kalscheuer[2] | Vassiliy N. Bavro[3] | Apoorva Bhatt[1]

[1]School of Biosciences and Institute of Microbiology and Infection, University of Birmingham, Birmingham, UK

[2]Institute of Pharmaceutical Biology and Biotechnology, Heinrich Heine University Düsseldorf, Düsseldorf, Germany

[3]School of Life Sciences, University of Essex, Colchester, UK

Correspondence
Vassiliy N. Bavro, School of Life Sciences, University of Essex, Colchester CO43SQ, UK.
Email: v.bavro@essex.ac.uk

Apoorva Bhatt, School of Biosciences and Institute of Microbiology and Infection, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK.
Email: a.bhatt@bham.ac.uk

Funding information
Medical Research Council, Grant/Award Number: MRC Career Development Award

Abstract

The final step in mycolic acid biosynthesis in *Mycobacterium tuberculosis* is catalysed by mycolyl reductase encoded by the *Rv2509* gene. Sequence analysis and homology modelling indicate that Rv2509 belongs to the short-chain fatty acid dehydrogenase/reductase (SDR) family, but with some distinct features that warrant its classification as belonging to a novel family of short-chain dehydrogenases. In particular, the predicted structure revealed a unique α-helical C-terminal region which we demonstrated to be essential for Rv2509 function, though this region did not seem to play any role in protein stabilisation or oligomerisation. We also show that unlike the *M. smegmatis* homologue which was not essential for growth, *Rv2509* was an essential gene in slow-growing mycobacteria. A knockdown strain of the *BCG2529* gene, the *Rv2509* homologue in *Mycobacterium bovis* BCG, was unable to grow following the conditional depletion of BCG2529. This conditional depletion also led to a reduction of mature mycolic acid production and accumulation of intermediates derived from 3-oxo-mycolate precursors. Our studies demonstrate novel features of the mycolyl reductase *Rv2509* and outline its role in mycobacterial growth, highlighting its potential as a new target for therapies.

KEYWORDS
dehydrogenase, *Mycobacterium*, mycolic acid, reductase, tuberculosis

1 | INTRODUCTION

The cell walls of mycobacteria including the tuberculosis-causing *Mycobacterium tuberculosis* contain distinct long-chain fatty acids termed mycolic acids. These α-alkyl, β-hydroxy fatty acids form an integral part of the cell wall of mycobacteria, and of other related bacteria of the suborder Corynebacterineae, including the genera *Nocardia*, *Corynebacterium* and *Rhodococcus* (Marrakchi, Laneelle, & Daffe, 2014). Mycolic acids can be found covalently linked to cell wall arabinogalactan and as part of the glycolipids' trehalose monomycolate (TMM), trehalose dimycolate and glucose monomycolate. They are essential for mycobacterial viability (Marrakchi et al., 2014; Nataraj et al., 2015) and virulence and are synthesised by a complex

Asma Javid and Charlotte Cooper joint first authors.

array of enzymes that includes a mammalian-like Fatty Acid Synthase I (FAS-I), and a multienzyme complex of Fatty Acid Synthase-II (FAS-II) (Marrakchi et al., 2014). One of the late stages of mycolic acid biosynthesis involves the Claisen condensation of a FAS-II-derived long meromycolate chain (C_{42}-C_{62}) with a short FAS-I-derived fatty acid (C_{24}-C_{26}) to yield α-alkyl, β-keto fatty acid intermediate, catalysed by the polyketide synthase Pks13 (Gande et al., 2004; Portevin et al., 2004). The final step in the formation of a mature mycolic acid involves the enzymatic reduction of the β-keto group of the product of Pks13 to produce α-alkyl, β-hydroxy fatty acids (mature mycolic acid). In *M. tuberculosis*, the reductase required for this final step is encoded by *Rv2509* (Bhatt, Brown, Singh, Minnikin, & Besra, 2008; Lea-Smith et al., 2007). Initial studies in *Corynebacterium glutamicum*, a species that can survive the loss of mycolic acids, indicated that deletion of *cmrA* (*NCgl2385*), the *Rv2509* orthologue, resulted in a strain that accumulated the α-alkyl, β-keto fatty acid precursor instead of mature mycolic acids (Lea-Smith et al., 2007). Mycolic acid biosynthesis genes are essential in mycobacteria; however, surprisingly, we were able to generate a viable mutant of *MSMEG4722*, the homologue of *Rv2509* in the fast-growing *Mycobacterium smegmatis* (Bhatt et al., 2008). Similar to the *C. glutamicum NCgl2385* mutant, the *M. smegmatis MSMEG4722* mutant produced α-alkyl, β-keto fatty acyl precursors of mycolic acids (3-oxo-mycolic acid precursors) which were transported and esterified to arabinogalactan in the cell wall (Bhatt et al., 2008). While the precise knowledge of the substrate for *Rv2509* is lacking, functional analysis of the preceding enzyme in the biosynthetic pathway, Pks13, indicated that, while this polyketide synthase primarily catalyses the formation of the α-alkyl, β-keto fatty acyl precursor, it also contains enzymatic motifs that facilitate the release of the nascent fatty acyl chain and its subsequent transfer to a trehalose residue to produce a trehalose residue esterified with the α-alkyl, β-keto fatty acyl precursor of mycolic acid (Gavalda et al., 2014). These findings suggest that the trehalose-bound mono 3-oxo-mycolic acid precursor may likely be the substrate for Rv2509, which catalyses its conversion to TMM (Figure 1). In this study, we set out to do a detailed study of the amino acid sequence of Rv2509 with the aim of deducing its structure and functional connections, which led to the identification of series of unique sequences, and by proxy, novel structural features. We then probed the role of one distinct structural feature of the mycobacterial mycolyl reductase in vitro by deletion analysis. We also probed the essentiality of *Rv2509* for the growth and viability of the slow-growing *M. tuberculosis* complex using *Mycobacterium bovis* BCG.

2 | RESULTS

2.1 | Homologues of *M. tuberculosis Rv2509* are found across mycolic acid-producing species

While mycolic acids are produced by several species belonging to the suborder Corynebacterineae, the biosynthesis of mycolates has been studied in detail exclusively in the genera *Mycobacterium* and *Corynebacterium*. While the two genera share a common set of enzymes required for mycolic acid biosynthesis, select differences also exist. For example, while the meromycolate chains are synthesised by a Type-II fatty acid synthase complex in species of *Mycobacterium*, the same function is carried out by a mammalian-like Type-I fatty acid synthase in *Corynebacterium* species (Radmacher et al., 2005). Similarly, differences exist in the transport of mycolic acids between the two genera (Varela et al., 2012; Yang et al., 2014). To investigate whether the reduction of premycolates by an exclusive mycolyl reductase, was an enzymatic process conserved across other mycolic acid-producing species, we used the amino acid sequence of Rv2509 as a source template for a BLASTp search of genome sequences of other known mycolic acid-producing bacteria. Homologues of *Rv2509* were found across the mycolate-producing genera (Figure 2). A phylogenetic analysis of *Rv2509*, its homologues in other mycobacteria and in select mycolic acid-producing species revealed that mycolyl reductases from Corynebacterial species branched early (Supporting information Figure S1). Corynebacteria produce relatively short chain mycolic acids (Radmacher et al., 2005). To test whether mycolyl reductases across the mycolata had evolved specificities to accommodate differing chain lengths, we transformed the *M. smegmatis* Δ*MSMEG4722* strain with a plasmid-containing *C. glutamicum* mycolyl reductase gene *NCgl2385* to generate the strain Δ*MSMEG4722*-C*NCgl2385*. Mature mycolic acid production was restored in the transformed strain, indicating that *NCgl2385* could reduce *M. smegmatis* 3-oxo-mycolic acid precursors and consequently fully complement the mutant *M. smegmatis* strain. This outcome suggested that mycolyl reductases across the mycolata likely did not evolve specificities for differing chain lengths (Figure 3).

2.2 | Predicted structure of *M. tuberculosis* Rv2509

We attempted to deduce possible functional properties of Rv2509 using a comparative sequence and structural analysis with proteins of known function. Sequence analysis of Rv2509 using the NCBI conserved domain database (Marchler-Bauer et al., 2015) categorises the protein as a part of the DltE family of short-chain dehydrogenases (cluster COG0300), which are in turn members of the cl27753 short-chain dehydrogenase/reductase (SDR) superfamily of NAD(P)-dependent oxidoreductases. This is a diverse group, containing over 80,000 unique sequences (Fujisawa, Nagata, & Misono, 2003). However, despite this diversity, all SDR-superfamilies display the classical Rossmann fold typical of NAD^+/$NADP^+$-binding proteins and hence can be classified as belonging to the cl21454 family of topologically and structurally related proteins.

Next, we screened the peptide sequence of *Rv2509* against the PDB database to find structural templates for homology modelling. The serine dehydrogenase YdfG (UniProt ID P39831; PDB ID 3ASU/3ASV) (Yamazawa, Nakajima, Mushiake, Yoshimoto, & Ito, 2011) had the highest identity score; however, structural analysis revealed several gaps in the pairwise alignment and the C-terminal region was shorter than that observed in *Rv2509*, which necessitated

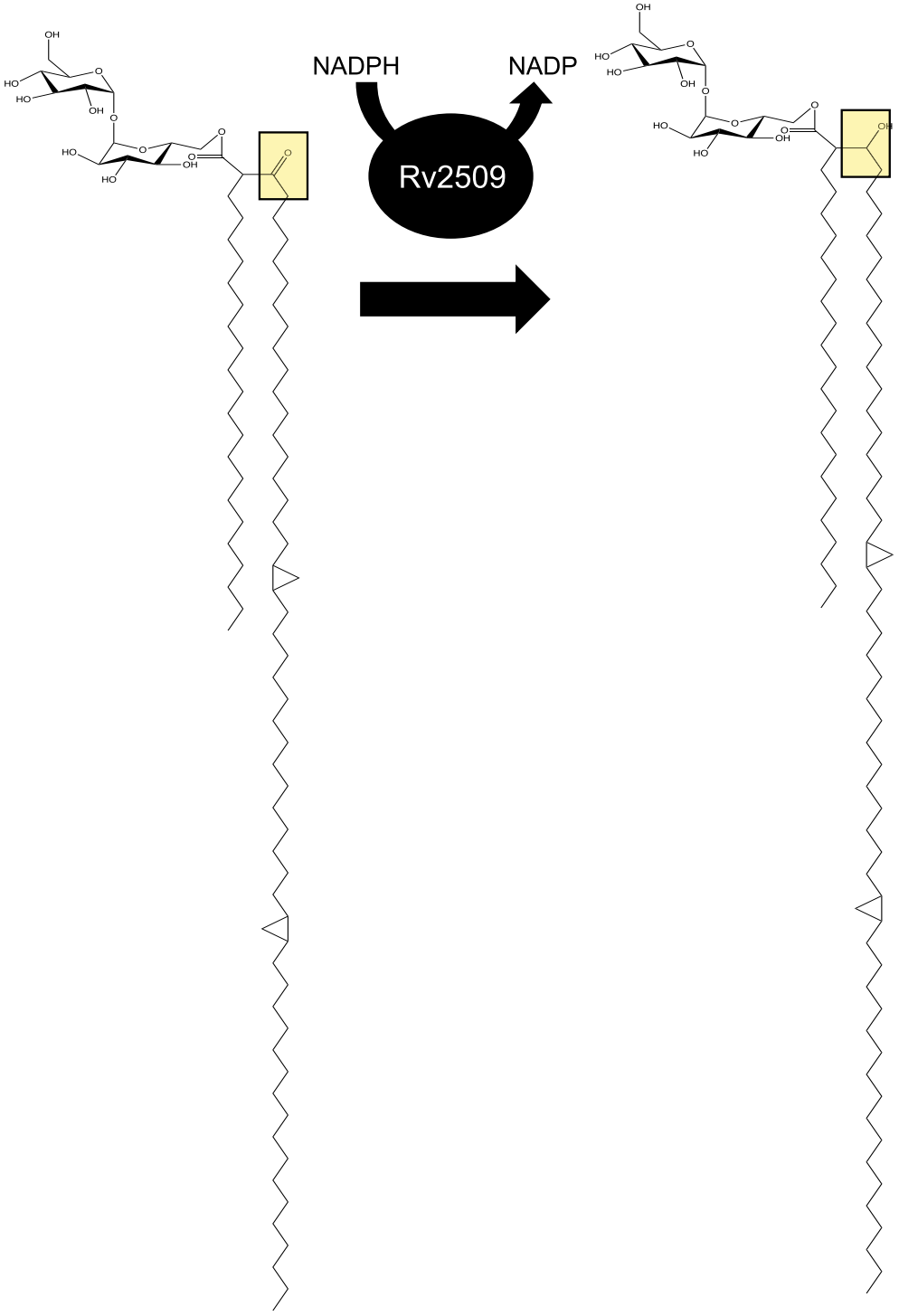


FIGURE 1 Proposed reaction catalysed by Rv2509. The putative substrate is trehalose mono-premycolate and the product is trehalose monomycolate. The examples shown here is that of a α-mycolic acid from *Mycobacterium tuberculosis* [Colour figure can be viewed at wileyonlinelibrary.com]

additional template input. Therefore, the structures of the four highest-scoring experimental SDRs (3ASU/3ASV, 1XG5, 4X54 and 4BMV) were used in parallel as a basis for comparative homology modelling of *Rv2509* (Figure 4) using the I-TASSER package (Yang et al., 2015) with specific template selection enabled. Notably, while, three of the templates yielded models with roughly equivalent (confidence) C-scores, the one based on the SDR from *Sphingobium yanoikuyae* (4BMV.pdb; GenBank ID ACB78183.1) stood out with the highest C-score of 0.6, indicating a high quality of prediction.

As visualised by the structural alignment (Figure 2) which shows the predicted secondary structure elements of Rv2509 aligned to its Corynebacterial homologues, the core structure of the Rv2509 model presents a typical NADP-binding protein Rossmann fold (residues 1–235) (Persson, Kallberg, Oppermann, & Jornvall, 2003)

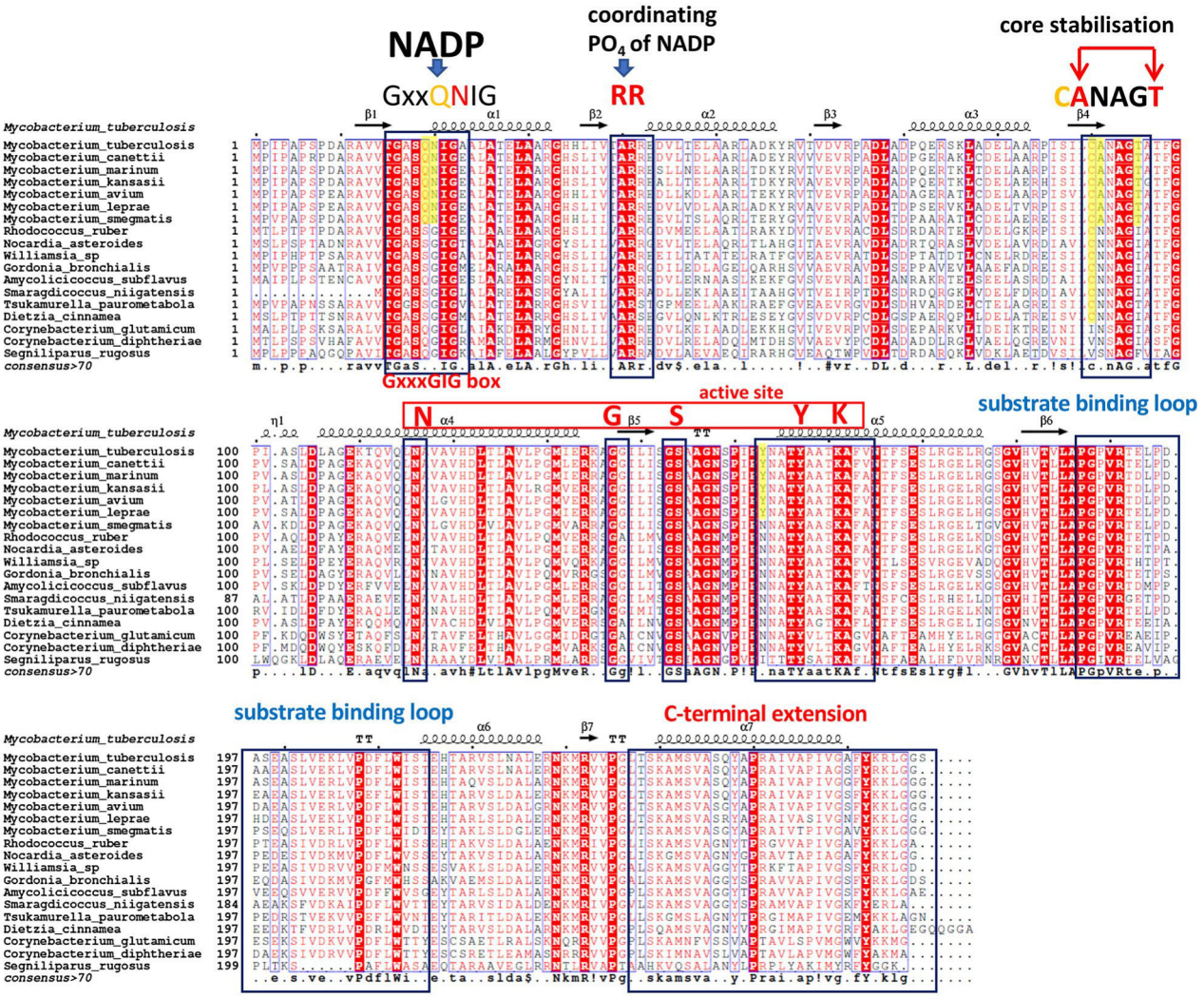


FIGURE 2 A structural alignment of the predicted Rv2509 structure from *Mycobacterium tuberculosis*, with annotated secondary structure elements (top) with the sequences of the Mycobacterial and Corynebacterial homologues highlighting the unique sequence motifs of the *Rv2509* subfamily of SDRs. The GxxQNIG motif along with RR pair identifies a novel NADP(H) binding consensus, limited to *Mycobacteriaceae* (highlighted in yellow). GxxSGIG is the wider consensus seen in the more remotely related Corynebacteria. Similarly, the unique Rossmann core-stabilisation signature CANAGT reported here is also restricted to the *Mycobacteriaceae* (highlighted in yellow) and reverts to the more common (C)NNAG(I/F) in the more remote homologues. Active site residues N-G-S-Y-K in alpha-helix 4 and alpha-helix 5 are preserved throughout the family; however, a unique additional tyrosine residue can be seen to be present in *Mycobacteriaceae* (highlighted in yellow). The C-terminal extension is also uniquely conserved across the wider Corynebacterial species

with seven parallel β-strands forming a central β-sheet, which is sandwiched between two layers of α-helices—three on each side (Figure 4). Although an earlier study has correctly classified Rv2509 within the short-chain reductase (SDR) superfamily and identified crucial sequence signatures such as the potential NAD(H)/NADP(H)-binding motif and active site in residues 157–161 (Lea-Smith et al., 2007), it did not include further structural prediction and analysis, and was limited in terms of sequences used. We, therefore, sought to expand this initial analysis by performing a systematic analysis of the Rv2509 orthologues from a wide range of *Corynebacterineae* to represent all known mycolate-producing genera.

We initiated these studies by analysing the cofactor binding site. SDRs can be divided into two large families, "Classical" SDRs which are around 250 amino acid residues long and "Extended" SDRs which contain around 350 residues (Jornvall et al., 1995), exhibiting differences in the arrangement of glycine residues within the coenzyme binding sites. Coenzyme preference-predictions based on the strong conservation of the signature residues (Persson et al., 2003) classified the Rv2509 into the group of the NADP(H)-binding SDRs, consistent with its mycolyl reductase function. However, it was difficult to categorically allocate Rv2509 into a specific group of SDRs as the predicted NADP-binding site of *Rv2509* ($G_{16}XXQ_{19}N_{20}I_{21}G_{22}$) displays a unique and novel variation of the (T)-G-X-X-X-G-I-G motif described by Brakoulias and Jackson (2004) (Figure 2). The variation seems to be restricted to the long-chain mycolates-producing species of *Mycobacteriaceae*, as short-chain mycolates-producing Corynebacteria retain the more established G-X-X-X-G-I-G motif. (Figure 2). Key residue

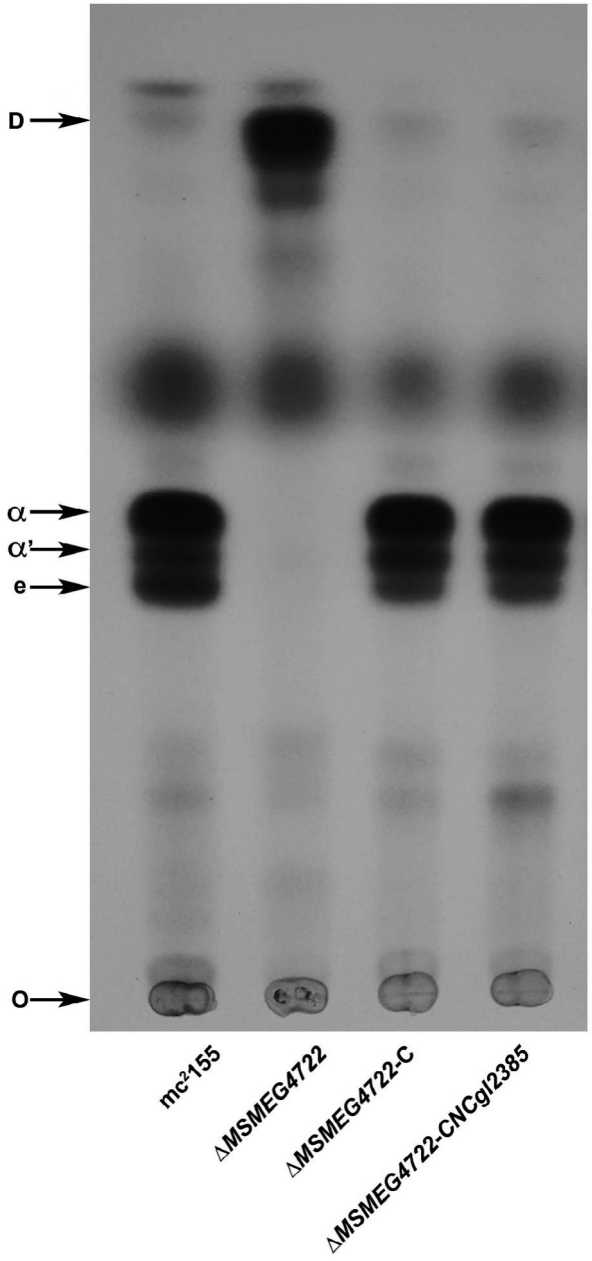


FIGURE 3 Complementation of the *M. smegmatis* Δ*MSMEG4722* mutant with the Corynebacterial mycolyl reductase. TLC analysis of ^{14}C-labelled mycolic acid methyl esters (MAMEs) isolated from different strains. α; α MAMEs, α′; α′ MAMEs, e; epoxy MAMEs, O; origin, D, degradation products derived from 3-oxo-mycolate precursors. Solvent system; petroleum ether:acetone (95:5, v:v)

positions in the NADP(H) site suggest that *Rv2509* is a cP2 member of the "Classic" SDR family, according to the classification system of Persson et al., (Brakoulias & Jackson, 2004), which is further reinforced by the presence of two specific discriminator residues, R-R, found in positions 41–42 in both *Rv2509* and YdfG from *Escherichia coli* (3ASV.pdb), as these residues are associated with coordination of the phosphate group of the NADP(H) cofactor. Further analysis of *Rv2509* suggests that it has a Class IV NADP(H)-binding site (with consensus [AVIC]-[LIV]-[VIL]-T-G-[ASGC]-X_2-[GR]-[ILF]-G-X_6-[LFY] (Hua, Wu, Sargsyan, & Lim, 2014). Thus, our analysis firmly places the Rv2509 family within the NADP(H)-binding subfamily of SDRs.

The active site in SDRs is formed by a tight cleft, lined by α-helix 4, β-strand 4 and α-helix 5 (Filling et al., 2002; Oppermann et al., 1997). Rv2509 exhibits the "classical" SDR motif Yx[AS][ST]K in α-helix 5 (Y157 and K161 being the central catalytic residues), while the GxggxgSS/T motif in β-strand 4 (G137, S144) and the central residue of α-helix 4, N116 (Figure 2) are also instantly recognisable. In addition, it is possible that some contact with the substrate may be provided by the β-strand 4.

Rather intriguingly, the consensus residues "NNAG" involved in the stabilisation of the central β-sheet of the Rossmann fold in the "Classical" SDRs and located at the core of the protein are not fully conserved in *Rv2509* but take the form of a novel "ANAG" sequence (see Figure 2). Furthermore, this novel sequence is paired with a preceding Cys residue, which is also correlated with the presence of a C-terminal Thr (in position 91 in *M. tuberculosis*). This unique "CANAGT" signature makes *Rv2509* orthologues in *Mycobacteriaceae* instantly recognisable (Figure 2).

In SDR proteins, the substrate-binding loop links the sixth β-sheet with the seventh α-helix and is one of the larger loops. In the predicted structure of Rv2509 this loop spans roughly from P_{187} to S_{214} and appears largely conserved within the mycolic acid-producing species (Figure 2). While the loop's length and position corresponds closely to the length of the substrate-binding loop in both 4BMV (P183-E206) and 3ASV (P178-T208), there was little conservation of the loop's sequence relative to Rv2509 in these and other divergent SDRs (see ConSurf analysis below (Ashkenazy et al., 2016); Supporting information Figure S2), suggesting, as expected, major differences in the structure of the substrate for Rv2509.

The most intriguing finding upon the analysis of the *Rv2509* orthologues was that all of them seem to present a C-terminal domain which seems to be highly specific to the *Corynebacterineae*. It proved also the most difficult to model reliably, as in the majority of the available SDR structural templates, the C-terminal region, appears to be either truncated or disorganized. ConSurf mapping using the multiple sequence alignment shown in Figure 2 and multiple sequence alignment based on 150 nonredundant sequences of remote SDR homologues revealed that the C-terminus of Rv2509 is uniquely a feature of the subfamily of Corynebacteriales (Supporting information Figure S3). Both pairwise and multiple sequence analyses indicated that the C-terminal region of the 4BMV.pdb is closely related to Rv2509 and thus can be tentatively used as a template for that region. This was further corroborated that by comparative modelling (Song et al., 2013) and pure *ab-initio* modelling (Raman et al., 2009), as implemented in the NewRobetta server (http://new.robetta.org). To validate this hypothesis, we conducted additional ab initio secondary structure predictions of *Rv2509* (236–268) using JPred4 (Drozdetskiy, Cole, Procter, & Barton, 2015) and PSIPRED (Buchan & Jones, 2019; Jones, 1999). Both of these bioinformatics tools supported a high alpha-helical propensity of the C-terminal region of *Rv2509*, suggesting that

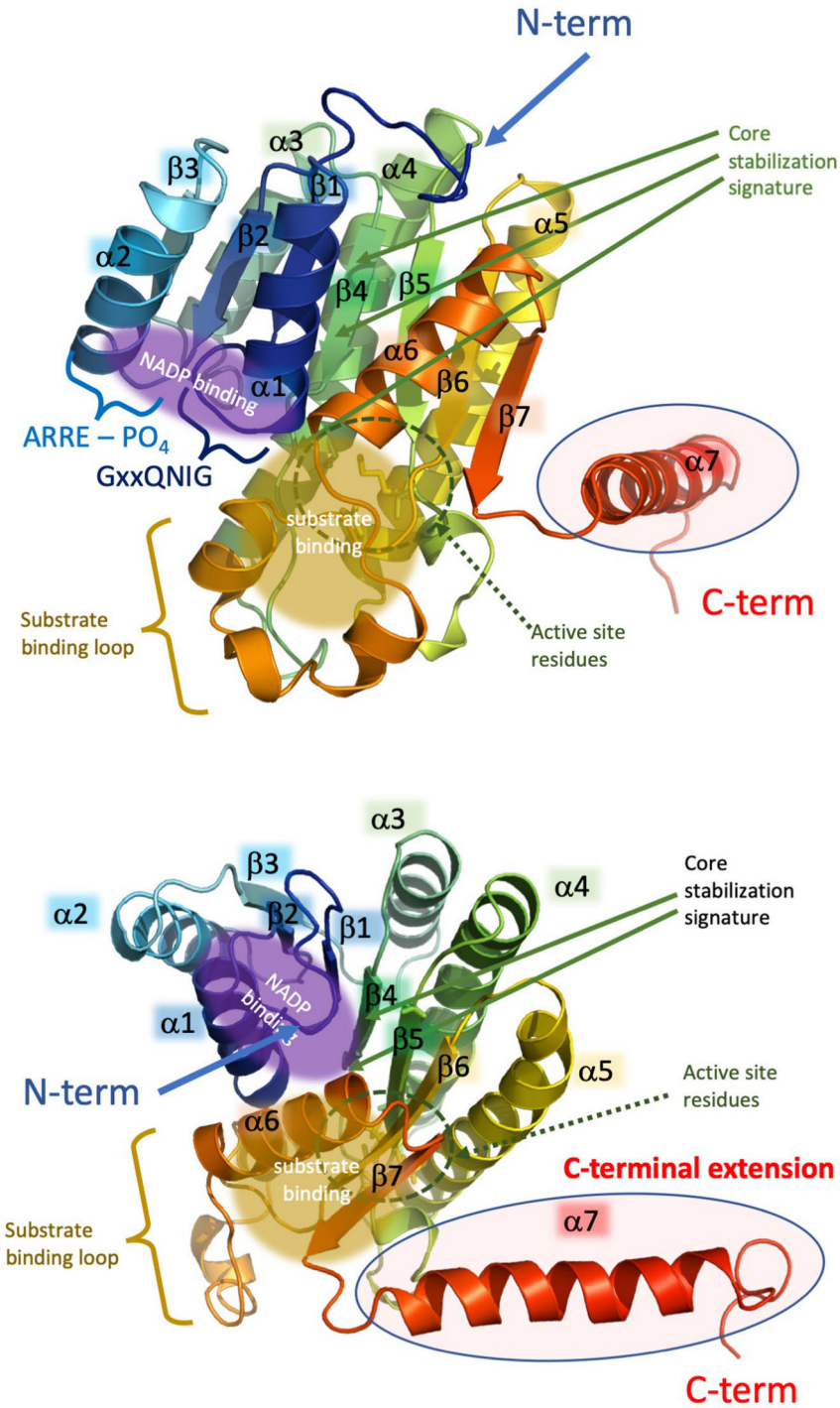


FIGURE 4 Visualisation of the predicted structure of *Mycobacterium tuberculosis* Rv2509 monomer. Two different orientations are given; side-view and top-down view towards the coenzyme binding site. The protein is coloured like a rainbow from blue (N-terminus) to red (C-terminus) and the main secondary structure elements are labelled, as well as the predicted C-terminal extension domain is circled in a red oval. The location of the key sequence and structural motifs discussed in the text is also provided

4BMV as a reliable template for the region (Supporting information Figure S3). Indeed, the C-terminus of the 4BMV.pdb structure, similarly to *Rv2509*, appears to be extended and is well-structured, presenting an α-helical organisation.

2.3 | The extended C-terminal of Rv2509 is essential for the function

Mature mycolic acid production could be restored in the *M. smegmatis* Δ*MSMEG4722* mutant on complementation with the *M. tuberculosis* orthologue *Rv2509* (Bhatt et al., 2008). The mutant strain thus provided us with a means of identifying residues and domains required for Rv2509 function, by assessing the ability of mutated constructs of plasmid-borne *Rv2509* to complement the Δ*MSMEG4722* strain by restoring mature mycolic acid biosynthesis. To test if the extended C-terminal of Rv2509 was required for function, we generated truncated versions of *Rv2509* cloned in the replicative shuttle plasmid pMV261 (Stover et al., 1991). Individual plasmids were then introduced by electroporation into the *M. smegmatis* Δ*MSMEG4722* mutant and the transformed strain analysed for the restoration of mature mycolic acid biosynthesis. Deletion of 14 amino acid residues (Ala_{255}-Ser_{268}) from

the C-terminal of Rv2509 resulted in a complete loss of function as the strains transformed with the truncated construct failed to restore mature mycolic acid biosynthesis in the Δ*MSMEG4722* strain (Figure 5; Δ*MSMEG4722*-CRvD1). A construct with a shorter truncation (Tyr_{262}-Ser_{268} deleted) also failed to complement the Δ*MSMEG4722* strain (Figure 5; Δ*MSMEG4722*-CRvD2). Expression of *Rv2509* and truncated constructs could be detected by RT-PCR confirming that the loss of Rv2509 function was linked to the deletion of the C-terminal residues (Supporting information Figure S4). These results demonstrated that the C-terminus of Rv2509 spanning terminal residues Ala_{255}-Ser_{268} was critical for function. The findings suggested an essential role for the unique, C-terminal extension in the activity of Rv2509.

To further query whether this C-terminal extension played a role in either protein stability, protecting Rv2509 from degradation, or in oligomerisation, we expressed N-terminal His-tagged versions of Rv2509 and Rv2509-D1 (Rv2509 with Ala_{255}-Ser_{268} deleted) in *E. coli*. Western blot analysis of native polyacrylamide gels showed that there were no patterns of degradation detected in Rv2509-D1, indicating that the loss of function due to the deletion of this region was not related to protein stability (Supporting information Figure S5).

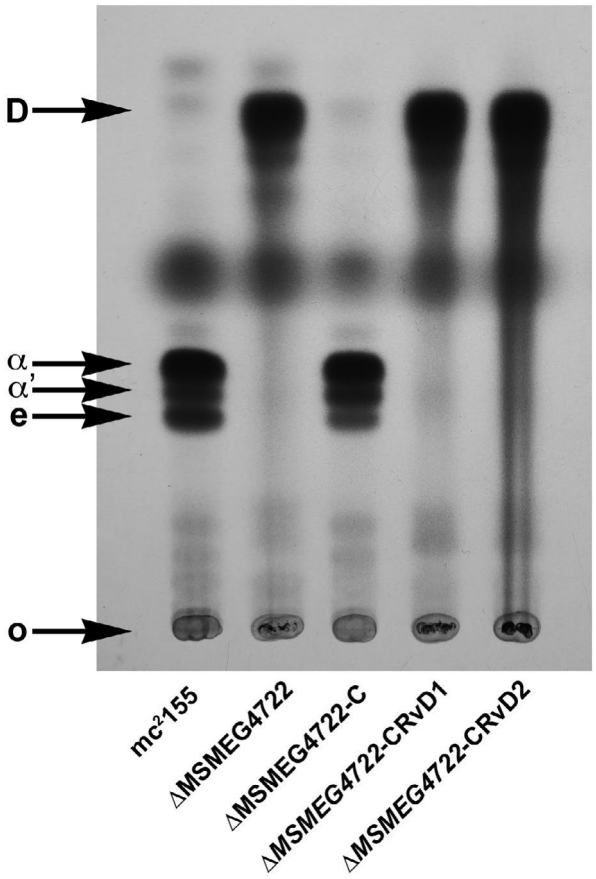


FIGURE 5 TLC analysis of ^{14}C-labelled MAMEs isolated from the strains of *M. smegmatis* Δ*MSMEG4722* mutant transformed with the different deletion constructs of *Rv2509*. α; α MAMEs, α′; α′ MAMEs, e; epoxy MAMEs, O; origin, D, degradation products derived from 3-oxo-mycolate precursors. Solvent system; petroleum ether:acetone (95:5, v:v)

Furthermore, we did not observe additional slower migrating bands, indicating dimerisation (or oligomerisation) of full-length or truncated Rv2509. Furthermore, cross-linking Rv2509 with glutaraldehyde did not reveal additional slower migrating bands, suggesting that Rv2509 did not oligomerise in solution (Supporting information Figure S5).

2.4 | *Rv2509* is an essential mycobacterial gene

Given the nonessentiality of the mycolyl reductase gene in *M. smegmatis* and prediction of transposon-site hybridisation screens of a slow-growing *M. tuberculosis Rv2509* mutant, we aimed to generate a null mutant of *Rv2509* in *M. tuberculosis* using Specialised Transduction (Bardarov et al., 2002). Such a mutant would also allow us to test the role of the mycolyl reductase in virulence using macrophage and mouse models of infection. However, repeated attempts failed to generate transductants, suggesting that *Rv2509* was essential for growth in *M. tuberculosis*. Indeed, subsequent transposon mutagenesis screens utilising deep sequencing predicted *Rv2509* to be essential for in vitro growth of *M. tuberculosis* (Griffin et al., 2011). We then utilised a promoter replacement strategy to demonstrate the essentiality of *Rv2509* in slow-growing mycobacterial vaccine strain *Mycobacterium bovis* BCG. The P*myc1* promoter from *M. smegmatis* engineered to contain four *tetO* operator sites (Korte et al., 2016) was inserted immediately upstream of the start codon of *BCG2529*, the *Rv2509* homologue in *M. bovis* BCG-Pasteur, using Specialised Transduction generating the strain BCG::P_{Tet}-*BCG2529*. Controlled gene expression of the *BCG2529* gene was achieved using a plasmid-borne synthetic gene (*rev-tetR*) derived from Tn10 *tetR* encoding a mutated TetR protein with reversed binding affinity to *tetO* sites upon the binding of tetracycline. The addition of anhydrotetracycline (ATc) results in the loss of expression.

The addition of ATc resulted in the loss of growth of BCG::P_{Tet}-*BCG2529* in broth, indicating that the expression of the gene encoding the mycolyl reductase was essential for the in vitro growth of slow-growing mycobacteria (Figure 6a). Furthermore, while concentrated cultures of BCG::P_{Tet}-*BCG2529* showed confluent growth on 7H10 agar plates, no confluent growth was observed on plates containing ATc (Figure 6b). We did observe a few scattered colonies on the ATc-containing plates and these are likely to be suppressors allowing leaky expression of P_{Tet} driven *BCG2529*. These results demonstrated that the mycolyl reductase was essential for the growth and viability of slow-growing mycobacteria.

2.5 | Conditional depletion of the mycolyl reductase results in the loss of mature mycolic acids

The *M. smegmatis* mycolyl reductase mutant was unable to synthesise fully mature mycolic acids and instead produced 3-oxo-mycolate precursors that were found esterified to the cell wall arabinogalactan and to trehalose (Bhatt et al., 2008). These labile precursors gave rise to palmitone derivatives following alkali treatment, which were visualised on Thin Layer Chromatography (TLC) plates migrating close to the

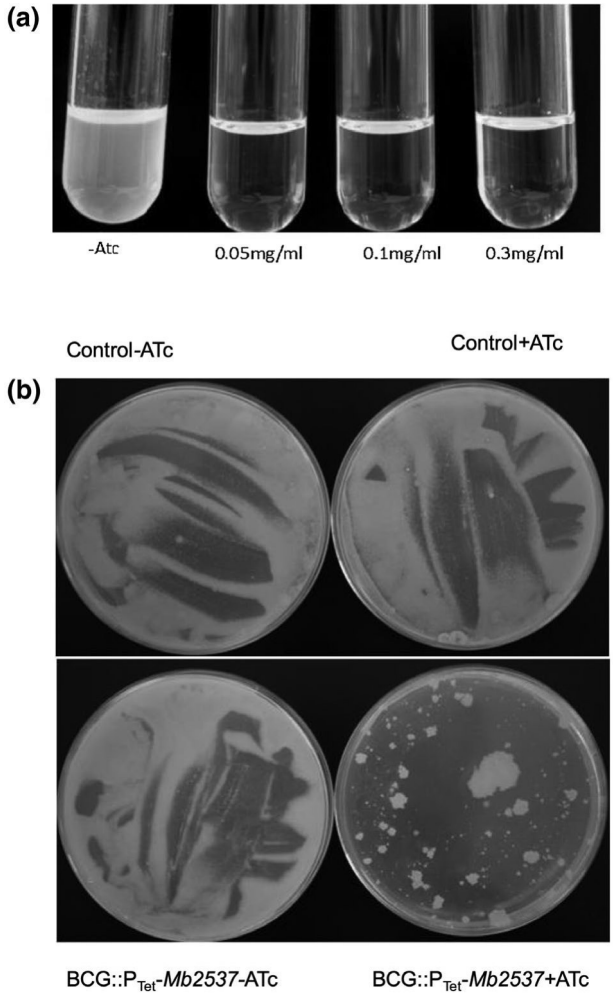


FIGURE 6 Conditional depletion of the mycolyl reductase in BCG::P$_{Tet}$-*BCG2529* leads to the loss of growth. (a) growth in different concentrations of ATc in 7H9 broth, (b) growth on 7H10 agar containing 5 μg/ml of ATc. Control strain; BCG transformed with pMV261::*rev-tetR*-RBS-D

solvent front in a solvent system used for separating methyl esters of mycolic acids and fatty acids (FAMEs and MAMEs) (Bhatt et al., 2008). To test the effects of conditional depletion of the mycolyl reductase we labelled the cultures of BCG::P$_{Tet}$-*BCG2529* with ^{14}[C]-acetic acid following grown in the presence and absence of ATc, and extracted MAMEs from the cultures. TLC analysis of the extracted MAMES showed a decrease in the levels of α and keto MAMES and an accumulation of the degradation products derived from 3-oxo-mycolate precursors close to the solvent front (Figure 7). These results showed that, as in *M. smegmatis*, loss of the mycolyl reductase function led to the loss of mature mycolic acid production and accumulation of the premature 3-oxo-mycolate precursors in slow grow mycobacteria.

3 | DISCUSSION

As summarised above, the predicted structure of the Rv2509 displays a Rossmann fold with seven parallel β-strands forming a central beta-sheet, which is sandwiched between two layers of alpha-helices—three on each side (Figure 4). While largely consistent with the "classical" family of SDRs (Persson et al., 2003), in *Mycobacteriaceae*, the NADP(H)-binding motif (G-x-x-Q-N-I-G), alongside the core stabilisation consensus (C-A-N-A-G-T) is significantly divergent from the consensus, to instantly identify this group of SDRs, and in our view is distinct enough to be separated into a distinct family.

While we currently lack knowledge about the exact mode of cofactor and ligand binding within the Rv2509 family, some information could be inferred from the structurally related FabG family of β-ketoacyl reductases which are classified as "Complex family SDRs" (Persson et al., 2003). Given that the structural overlay of Rv2509 models with existing ligand-bound FabG members yields an RMSD of around 1 Å over the C-alpha backbone atoms, and that FabG share common active site residues with *Rv2509* it is tempting to suggest that these structures share common ligand-binding and NADP(H) orientations towards the active site residues with that observed in the aceto-acyl-CoAs. Furthermore, the predicted 3D structure of Rv2509 also revealed a unique α-helical C-terminal extension, which we subsequently showed to be essential for function—constructs with deletions in the C-terminal failed to complement (restore mycolate biosynthesis) in the *M. smegmatis* mutant.

The role of this C-terminal domain is difficult to assign with certainty; however, in related SDRs similar C-terminal extensions play a number of functions. In the NADP(H)-dependent tetrameric serine-dehydrogenase YdfG from *E. coli* (3ASV.pdb) (Yamazawa et al., 2011), the C-terminal region protrudes from the main core of the protein and is required for tetramer formation, while it also plays a role in the substrate binding *via* stabilisation of the substrate loop. It is also notable that in FabG family of related SDRs discussed above also organise as tetrameric assemblies with the participation of their C-terminal domains (Javidpour et al., 2014). Furthermore, in the mycobacterial FabG4 dimer, Arg146 and Arg445 of one protomer interact with the C-terminus of the second protomer and play an essential role in the substrate association and catalysis (Dutta, Bhattacharyya, Roychowdhury, Biswas, & Das, 2013). However, our experiments with purified full-length and truncated Rv2509 do not seem to suggest that the mycolyl reductase forms oligomers.

While it isn't in itself surprising that Rv2509 presents a conserved SDR fold, our analysis has identified it for the first time decisively as a member of the NADPH-binding family and indicated that it possesses a number of unique features. Taken together with the unique signatures of the NADPH-binding site, core stabilisation domains, as well as C-terminal domain extension, the unique features of Rv2509 possibly merit its inclusion into a distinct family.

Our previous studies with *M. smegmatis* suggested that the mycolyl reductase would likely not be essential for the growth in *M. tuberculosis* and related slow-growing mycobacteria. However, this work demonstrated that gene encoding the reductase, *Rv2509*, was an essential gene required for the growth of slow-growing mycobacteria in laboratory media. This, along with the identification of unique domains in the predicted structure of Rv2509 highlights its potential as a novel anti-TB drug target.

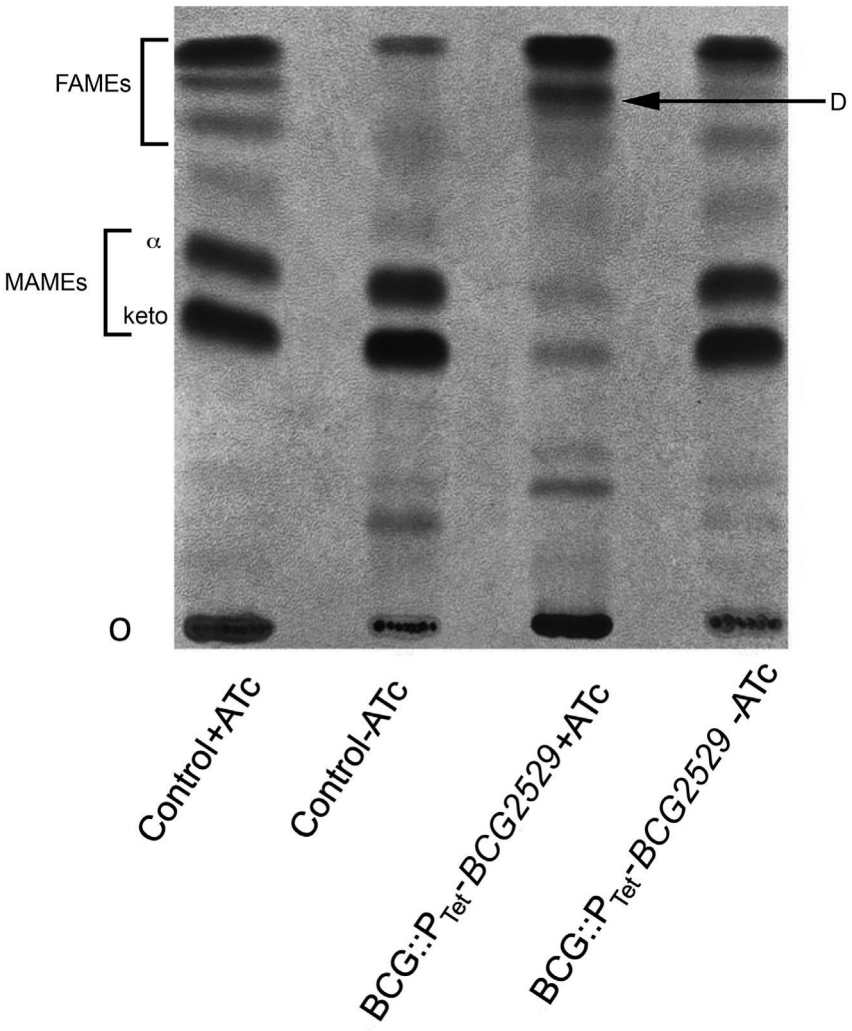


FIGURE 7 TLC analysis of ^{14}C-labelled MAMEs isolated from BCG::P_{Tet}-*BCG2529* strain. α; α MAMEs, keto; keto MAMEs, O; origin, D, degradation products derived from 3-oxo-mycolate precursors. Solvent system; petroleum ether:acetone (95:5, v:v)

4 | EXPERIMENTAL PROCEDURES

4.1 | Growth conditions, strains, phages and plasmids

M. smegmatis mc^{2}155 and derived strains was cultured in Tryptic Soy Broth (TSB) or on TSB-agar, *M. bovis* BCG strains were grown in 7H9 broth or on 7H10 agar plates. *E. coli* strains were grown in LB broth or LB agar. Hygromycin was used at a concentration of 100 mg/ml for *E. coli* and mycobacterial strains. Kanamycin was used at a concentration of 25 mg/ml and 50 mg/ml for mycobacteria and *E. coli* respectively. Recombinant mycobacterial strains, phages and plasmid constructs used in this study are listed and described in Table 1.

4.2 | Sequence analysis and homology modelling

The multiple sequence alignments (MSA) were prepared using MAFFT and NJ/UPGMA phylogeny algorithms as implemented in MAFFT v.7 server (https://mafft.cbrc.jp/ (Katoh, Misawa, Kuma, & Miyata, 2002)) and the resulting phylogenetic trees were visualised using Archaeopteryx (Han & Zmasek, 2009). Multiple sequence alignment visualisation and structural annotations were performed using ESPript 3 (Gouet, Robert, & Courcelle, 2003). For the purposes of homology modelling, we employed I-TASSER (Yang et al., 2015) with the assignment of templates. Outputs were independently corroborated by ab-initio modelling using the New Robetta server (http://new.robetta.org). Sequence conservation analysis was performed using ConSurf (Ashkenazy et al., 2016) and visualised using PyMOL (The PyMOL Molecular Graphics System, Version 1.71 Schrödinger, LLC) with additional local manual refinement and structural superposition performed in Coot (Emsley, Lohkamp, Scott, & Cowtan, 2010). Secondary structure predictions for the C-terminal domain were performed with JPred4 (Cole et al., 2008) and PSIPRED (Buchan et al., 2013).

4.3 | Expression, purification and analysis of Rv2509 and its C-terminal truncated derivative

Rv2509 was cloned in the expression vector pET28a by PCR from *M. tuberculosis* H37Rv genomic DNA using the primer pair Pet28a_Rv2509 _F2 (5′-TTTACATATGCCGATA CCCGCGCCC-3′) and Pet28a_Rv2509_R1 (5′-ATTAAAGCTTCTAGCTGCCCCCAA

TABLE 1 Plasmids and bacterial strains used in this study

	Description	Reference/Source
Plasmids		
pMV261	*E. coli-Mycobacterium* shuttle plasmid vector with *hsp60* promoter and Kan^R cassette (*aph*)	Stover et al. (1991)
pMV261-*Rv2509*	*Rv2509* cloned in pMV261	Bhatt et al. (2008)
pMV261-*NCgl2385*	*NCgl2385* cloned in pMV261	This work
pMV261-*Rv2509D1*	Deletion construct of *Rv2509* cloned in pMV261 encodes a truncated protein missing the C-terminal residues Ala_{255}-Ser_{268}	This work
pMV261-*Rv2509D2*	Deletion construct of *Rv2509* cloned in pMV261 encodes a truncated protein missing the C-terminal residues Tyr_{262}-Ser_{268}	This work
pMV261::*rev-tetR*-RBS-D	Episomal *E. coli*-mycobacterium shuttle plasmid pMV261::*rev-tetR*-RBS-D providing constitutive *rev-tetR* gene expression from the HSP60 promoter in mycobacteria	Korte et al. (2016)
pET28a-*Rv2509*	Plasmid for expressing *Rv2509* in *E. coli*	This work
pET28a-*Rv2509D1*	Plasmid for expressing C-terminal truncated *Rv2509* in *E. coli*	This work
Bacterial strains		
mc^2155	Electroporation-proficient *ept* mutant of *M. smegmatis* strain mc^26	Snapper, Melton, Mustafa, Kieser, and Jacobs (1990)
Δ*MSMEG4722*	Deletion mutant of mc^2155 in which *MSMEG4722* is replaced by *hyg*	Bhatt et al. (2008)
Δ*MSMEG4722-CRv*	Δ*MSMEG4722*-containing pMV261-*Rv2509*	Bhatt et al. (2008)
Δ*MSMEG4722-CNCgl2385*	Δ*MSMEG4722*-containing pMV261-*NCgl2385*	This work
Δ*MSMEG4722*-CRvD1	Δ*MSMEG4722*-containing pMV261-*Rv2509D1*	This work
Δ*MSMEG4722*-CRvD2	Δ*MSMEG4722*-containing pMV261-*Rv2509D2*	This work
BCG::P_{Tet}-*BCG2529*	BCG Pasteur strain containing the P*myc1* promoter from *M. smegmatis* engineered to contain four *tetO* operator sites, inserted immediately upstream of the start codon of *BCG2529*, the BCG homologue of *Rv2509*	This work

GCCTCTTG-3′). Similarly, the C-terminal truncated version of Rv2509 (missing the last 14 amino acids) was PCR amplified using the primer pair Pet28a_Rv2509 _F2 (5′-TTTACATA TGCCGATACCCGCGCCC-3′) and Pet28a_Rv2509_delR1 (5′-ATTAAAGCTTCTACA CGATGGCGCGCGGAGC-3′). The constructs were cotransformed into BL21 cells with *M. tuberculosis* chaperone GroES 60.2. Cultures were seeded 1:100 with an overnight starter culture and grown at 37°C in terrific broth supplemented with ampicillin (100 μg/ml) and kanamycin (50 μg/ml) to OD600 = 0.6 and induced with 0.5 mM IPTG overnight at 37°C or 16°C respectively. Cells were harvested at 4,000 rpm at 4°C, washed with PBS and frozen at −80°C until further use. Full-length protein was purified on cobalt-IMAC in buffer 50 mM NAH2PO4 pH 7.5, 500 mM NaCl, 10% glycerol and truncated in the same buffer composition at pH 7.0. The eluted fraction of 200 mM imidazole was taken for native-PAGE analysis.

A native polyacrylamide gel with the separated samples was transferred to a Hybond™ nitrocellulose membrane (GE Healthcare Life Sciences) which was then blocked in 5% milk 20 mM Tris pH 7.5, 150 mM NaCl TBS 0.05% Tween for 1h at room temperature followed by incubation with 0.1% milk TBS-tween with penta-his antibody BSA free (Qiagen). Following washing with TBS-tween the membrane was incubated with 0.1% milk TBS-tween with anti-mouse IgG-Alkaline Phosphatase (Sigma Aldrich) for 30 min at room temperature, washed with TBS-tween and a final wash with TBS. The membrane was stained for 5 min with BCIP/NBT (SIGMAFAST™) in water for visualisation.

Cross-linking of full-length Rv2509 was carried out in 50 mM NaH_2PO_4 pH 7.5, 10% glycerol at a final concentration of 0.1% glutaraldehyde and allowed to progress at room temperature for 2 min. The reaction was halted with 1M Tris pH 8 and subsequently run on SDS-PAGE for visualisation using SYPRO-ruby (Molecular Probes).

4.4 | Conditional depletion of the Rv2509 homologue *BCG2529* in *M. bovis* BCG-Pasteur

For establishing the regulated expression of the *BCG2529* gene, a synthetic gene cassette (*hyg*-P*myc1*-4X*tetO*) comprising a hygromycin-resistance gene and the P*myc1* promoter from *M. smegmatis* engineered to contain four *tetO* operator sites (Korte et al., 2016) was inserted immediately upstream of the *BCG2529* start codon in *M. bovis* BCG-Pasteur. Targeted gene knock-in was achieved by specialised transduction employing temperature-sensitive mycobacteriophages essentially as described previously (Korte et al., 2016). Briefly, for the generation of an allelic exchange construct for site-specific insertion

of the *hyg*-P*myc1*-4X*tetO* cassette in *M. bovis* BCG-Pasteur, upstream- and downstream DNA regions flanking the *BCG2529* start codon were amplified by PCR employing the oligonucleotide pairs *Mb2537*_LL 5′-TTTTTCCATAAATTGGAACCGCTACCTGACATGAAACCC-3′ and *Mb2537*_LR 5′-TTTTTCCATTTCTTGGGCCGATGTTCTGCGAAGCCCCGG-3′ as well as *Mb2537*_RL 5′-TTTTTCCATAGATTGGATGCCGATACCCGCGCCCAGCCC-3′ and *Mb2537*_RR 5′-TTTTTCCATCTTTTGGCGGTGTGGGAGGAGATACTCAAG-3′ respectively. The gene *BCG2528c* is localised in antilinear orientation upstream of *BCG2529* with only 84 bp between the start codons of both genes. Since it could not be excluded that the *Mb2536c* promoter region might overlap with the *BCG2529* coding region, the upstream flanking region contained a duplication of 66 bp of the 5′-end of *BCG2529* to conserve 150 bp in front of *BCG2529* that likely comprised its promoter. Subsequently, the upstream and downstream flanks were digested with *Van*91I (restriction sites underlined) and ligated with *Van*91I-digested pcRv1327c-4XtetO vector arms (Korte et al., 2016). The resulting knock-in plasmid was then linearised with *Pac*I and cloned and packaged into the temperature-sensitive phage phAE159 (Jain et al., 2014), yielding a knock-in phage which was propagated in *M. smegmatis* at 30°C. Allelic exchange in *M. bovis* BCG-Pasteur using the knock-in phage at the nonpermissive temperature of 37°C was achieved by specialised transduction using hygromycin (50 mg/L) for selection, resulting in the site-specific insertion of the *hyg*-P*myc1*-4X*tetO* cassette. The obtained *M. bovis* c-BCG_2529-4XtetO knock-in mutant was verified by the diagnostic PCR of genomic DNA using the oligonucleotide pair BCG_2529_L_fw 5′ GTCAGGTAGACGGAGAACAC-3′ and BCG_2529_L_rev 5′ AGCTCACCGCGCAGAGATTC-3′ binding outside the allelic exchange substrates used to generate this mutant and subsequent sequencing of PCR products (Supporting information Figure S6). For achieving the controlled gene expression of the *Mb2537* gene, a synthetic gene (*rev-tetR*) derived from Tn10 *tetR* encoding a mutated TetR protein with reversed binding affinity to *tetO* sites upon the binding of tetracycline (Klotzsche, Ehrt, & Schnappinger, 2009) was heterologously expressed in the knock-in mutant by electroporation of the episomal *E. coli*-mycobacterium shuttle plasmid pMV261::*rev-tetR*-RBS-D providing constitutive *rev-tetR* gene expression from the HSP60 promoter in mycobacteria using solid medium containing 50 mg/L of hygromycin and 20 mg/L of kanamycin for selection (Famulla et al., 2016). This yielded the conditional mutant BCG-Pasteur c-*BCG2529*-4X*tetO* pMV261::*rev-tetR*-RBS-D (referred to as BCG::P_{Tet}-*BCG2529*) allowing silencing of the *BCG2529* gene in the presence of anhydrotetracycline (ATc).

For silencing experiments, precultures of the BCG::P_{Tet}-*BCG2529* strain were first grown in liquid medium containing 0.3 µg/ml of ATc, before these precultures were used to inoculate the test cultures containing 0–5 µg/ml of ATc. Growth in 96-well microtitre plates was quantified after incubation at 37°C for five days by adding 10% of 100 µg/ml of resazurin solution. After further incubation at ambient temperature for 16 hr, cells were fixed for 30 min by formalin addition (5%, v/v, final concentration) and fluorescence was quantified using a microplate reader (excitation 540 nm, emission 590 nm). For the growth on 7H10 agar, nonpermissive conditions were obtained by adding ATc to a final concentration of 5 µg/ml.

4.5 | Functional complementation studies

Deletion constructs for *Rv2509* were generated by PCR amplification of selected regions of *Rv2509* using pMV261-*Rv2509* as a template, incorporating a premature stop codon in the reverse primer. The plasmid constructs generated in this manner were called pMV261-Rv2509D1 and pMV261-Rv2509D2, encoding a truncated protein missing the C-terminal residues Ala_{255}-Ser_{268} and Tyr_{262}-Ser_{268} respectively. The *C. glutamicum Rv2509* orthologue, *NCgl2385* was PCR amplified using *C. glutamicum* genomic DNA as a template and the PCR product was cloned into pMV261 using primer incorporated BamHI and EcoRI sites generating the plasmid pMV261-*NCgl2385*. Plasmid clones were electroporated into the *M. smegmatis* Δ*MSMEG4722* mutant strain using previously described protocols (Snapper et al., 1990). Kanamycin-resistant transformants were cultured in TSB and FAMEs and MAMEs were extracted from the cell pellets and separated by TLC as previously described (Vilcheze & Jacobs, 2007).

4.6 | RT-PCR of *Rv2509* transcripts

M. smegmatis strains were grown to mid-log phase and RNA was extracted by acid phenol and chloroform extraction. The purified RNA was treated with DNase and converted to cDNA, which was used as a template for quantitative PCR. PCR products were run on a 1% agarose gel to detect the expression of *Rv2509*. The gene *sigA* was monitored as a constitutively expressed gene. The primers used for PCR amplification were Rv2509RT-f (5′-ACAAGTACCGCGTCACGGTC-3′) and Rv2509RT-r (5′-AAGTCCGGCACC AGCTTCTC-3′) for the amplification of nondeleted sections of *Rv2509*.

ACKNOWLEDGMENTS

CC is funded by a BBSRC PhD studentship provided by the BBSRC Midlands Integrative Biosciences Training Partnership (MIBTP). AB acknowledges previous grant support from the Medical Research Council (UK). RK acknowledges financial support from the Jürgen Manchot Foundation.

AUTHOR CONTRIBUTIONS

AB, RK and VB designed the study. AJ, CC, AS, SS, MH and RM conducted experiments and acquired data. AJ, CC, AS, SS, MH, RM, AB, RK and VB interpreted the data. AB, RK and VB wrote the manuscript.

ORCID

Vassiliy N. Bavro https://orcid.org/0000-0003-2330-8924
Apoorva Bhatt https://orcid.org/0000-0002-6655-1636

REFERENCES

Ashkenazy, H., Abadi, S., Martz, E., Chay, O., Mayrose, I., Pupko, T., & Ben-Tal, N. (2016). ConSurf 2016: An improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Research*, *44*, W344–W350. https://doi.org/10.1093/nar/gkw408

Bardarov, S., Bardarov, S., Jr., Pavelka, M. S., Jr., Sambandamurthy, V., Larsen, M., Tufariello, J., … Jacobs, W. R., Jr. (2002). Specialized transduction: An efficient method for generating marked and unmarked targeted gene disruptions in *Mycobacterium tuberculosis*, *M. bovis* BCG and *M. smegmatis*. *Microbiology*, *148*, 3007–3017. https://doi.org/10.1099/00221287-148-10-3007

Bhatt, A., Brown, A. K., Singh, A., Minnikin, D. E., & Besra, G. S. (2008). Loss of a mycobacterial gene encoding a reductase leads to an altered cell wall containing beta-oxo-mycolic acid analogs and accumulation of ketones. *Chemistry & Biology*, *15*, 930–939.

Brakoulias, A., & Jackson, R. M. (2004). Towards a structural classification of phosphate binding sites in protein-nucleotide complexes: An automated all-against-all structural comparison using geometric matching. *Proteins*, *56*, 250–260. https://doi.org/10.1002/prot.20123

Buchan, D. W. A., & Jones, D. T. (2019). The PSIPRED protein analysis workbench: 20 years on. *Nucleic Acids Research*, *47*(W1), W402–W407.

Buchan, D. W., Minneci, F., Nugent, T. C., Bryson, K., & Jones, D. T. (2013). Scalable web services for the PSIPRED Protein Analysis Workbench. *Nucleic Acids Research*, *41*(Web Server issue), W349–W357. https://doi.org/10.1093/nar/gkt381

Cole C., Barber, J. D., & Barton, G. J. (2008). Jpred 3 secondary structure prediction server. *Nucleic Acids Research*, *36*(Web Server issue), W197–W201. https://doi.org/10.1093/nar/gkn238T

Drozdetskiy, A., Cole, C., Procter, J., & Barton, G. J. (2015). JPred4: A protein secondary structure prediction server. *Nucleic Acids Research*, *43*, 389–394. https://doi.org/10.1093/nar/gkv332

Dutta, D., Bhattacharyya, S., Roychowdhury, A., Biswas, R., & Das, A. K. (2013). Crystal structure of hexanoyl-CoA bound to beta-ketoacyl reductase FabG4 of *Mycobacterium tuberculosis*. *The Biochemical Journal*, *450*, 127–139.

Emsley, P., Lohkamp, B., Scott, W. G., & Cowtan, K. (2010). Features and development of Coot. *Acta Crystallographica Section D: Biological Crystallography*, *66*, 486–501.

Famulla, K., Sass, P., Malik, I., Akopian, T., Kandror, O., Alber, M., … Brotz-Oesterhelt, H. (2016). Acyldepsipeptide antibiotics kill mycobacteria by preventing the physiological functions of the ClpP1P2 protease. *Molecular Microbiology*, *101*, 194–209. https://doi.org/10.1111/mmi.13362

Filling, C., Berndt, K. D., Benach, J., Knapp, S., Prozorovski, T., Nordling, E., … Oppermann, U. (2002). Critical residues for structure and catalysis in short-chain dehydrogenases/reductases. *Journal of Biological Chemistry*, *277*, 25677–25684. https://doi.org/10.1074/jbc.M202160200

Fujisawa, H., Nagata, S., & Misono, H. (2003). Characterization of short-chain dehydrogenase/reductase homologues of *Escherichia coli* (YdfG) and *Saccharomyces cerevisiae* (YMR226C). *Biochimica et Biophysica Acta*, *1645*, 89–94. https://doi.org/10.1016/S1570-9639(02)00533-2

Gande, R., Gibson, K. J., Brown, A. K., Krumbach, K., Dover, L. G., Sahm, H., … Eggeling, L. (2004). Acyl-CoA carboxylases (accD2 and accD3), together with a unique polyketide synthase (Cg-pks), are key to mycolic acid biosynthesis in *Corynebacterianeae* such as *Corynebacterium glutamicum* and *Mycobacterium tuberculosis*. *Journal of Biological Chemistry*, *279*, 44847–44857.

Gavalda, S., Bardou, F., Laval, F., Bon, C., Malaga, W., Chalut, C., … Quemard, A. (2014). The polyketide synthase Pks13 catalyzes a novel mechanism of lipid transfer in mycobacteria. *Chemistry & Biology*, *21*, 1660–1669. https://doi.org/10.1016/j.chembiol.2014.10.011

Gouet, P., Robert, X., & Courcelle, E. (2003). ESPript/ENDscript: Extracting and rendering sequence and 3D information from atomic structures of proteins. *Nucleic Acids Research*, *31*, 3320–3323. https://doi.org/10.1093/nar/gkg556

Griffin, J. E., Gawronski, J. D., Dejesus, M. A., Ioerger, T. R., Akerley, B. J., & Sassetti, C. M. (2011). High-resolution phenotypic profiling defines genes essential for mycobacterial growth and cholesterol catabolism. *PLoS Pathogens*, *7*, e1002251. https://doi.org/10.1371/journal.ppat.1002251

Han, M. V., & Zmasek, C. M. (2009). phyloXML: XML for evolutionary biology and comparative genomics. *BMC Bioinformatics*, *10*, 356. https://doi.org/10.1186/1471-2105-10-356

Hua, Y. H., Wu, C. Y., Sargsyan, K., & Lim, C. (2014). Sequence-motif detection of NAD(P)-binding proteins: Discovery of a unique antibacterial drug target. *Scientific Reports*, *4*, 6471. https://doi.org/10.1038/srep06471

Jain, P., Hsu, T., Arai, M., Biermann, K., Thaler, D. S., Nguyen, A., … Jacobs, W. R., Jr. (2014). Specialized transduction designed for precise high-throughput unmarked deletions in *Mycobacterium tuberculosis*. *MBio*, *5*, e01245-14. https://doi.org/10.1128/mBio.01245-14

Javidpour, P., Pereira, J. H., Goh, E. B., McAndrew, R. P., Ma, S. M., Friedland, G. D., … Beller, H. R. (2014). Biochemical and structural studies of NADH-dependent FabG used to increase the bacterial production of fatty acids under anaerobic conditions. *Applied and Environment Microbiology*, *80*, 497–505. https://doi.org/10.1128/AEM.03194-13

Jones, D. T. (1999). Protein secondary structure prediction based on position-specific scoring matrices. *Journal of Molecular Biology*, *292*, 195–202.

Jornvall, H., Persson, B., Krook, M., Atrian, S., Gonzalez-Duarte, R., Jeffery, J., & Ghosh, D. (1995). Short-chain dehydrogenases/reductases (SDR). *Biochemistry*, *34*, 6003–6013.

Katoh, K., Misawa, K., Kuma, K., & Miyata, T. (2002). MAFFT: A novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research*, *30*, 3059–3066. https://doi.org/10.1093/nar/gkf436

Klotzsche, M., Ehrt, S., & Schnappinger, D. (2009). Improved tetracycline repressors for gene silencing in mycobacteria. *Nucleic Acids Research*, *37*, 1778–1788. https://doi.org/10.1093/nar/gkp015

Korte, J., Alber, M., Trujillo, C. M., Syson, K., Koliwer-Brandl, H., Deenen, R., … Kalscheuer, R. (2016). Trehalose-6-phosphate-mediated toxicity determines essentiality of OtsB2 in *Mycobacterium tuberculosis* in vitro and in mice. *PLoS Pathogens*, *12*, e1006043. https://doi.org/10.1371/journal.ppat.1006043

Lea-Smith, D. J., Pyke, J. S., Tull, D., McConville, M. J., Coppel, R. L., & Crellin, P. K. (2007). The reductase that catalyzes mycolic motif synthesis is required for efficient attachment of mycolic acids to arabinogalactan. *Journal of Biological Chemistry*, *282*, 11000–11008. https://doi.org/10.1074/jbc.M608686200

Marchler-Bauer, A., Derbyshire, M. K., Gonzales, N. R., Lu, S., Chitsaz, F., Geer, L. Y., … Bryant, S. H. (2015). CDD: NCBI's conserved domain database. *Nucleic Acids Research*, *43*, D222–D226. https://doi.org/10.1093/nar/gku1221

Marrakchi, H., Laneelle, M. A., & Daffe, M. (2014). Mycolic acids: Structures, biosynthesis, and beyond. *Chemistry & Biology*, *21*, 67–85. https://doi.org/10.1016/j.chembiol.2013.11.011

Nataraj, V., Varela, C., Javid, A., Singh, A., Besra, G. S., & Bhatt, A. (2015). Mycolic acids: Deciphering and targeting the Achilles' heel of the tubercle bacillus. *Molecular Microbiology*, *98*, 7–16. https://doi.org/10.1111/mmi.13101

Oppermann, U. C., Filling, C., Berndt, K. D., Persson, B., Benach, J., Ladenstein, R., & Jornvall, H. (1997). Active site directed mutagenesis of 3 beta/17 beta-hydroxysteroid dehydrogenase establishes differential effects on short-chain dehydrogenase/reductase reactions. *Biochemistry*, *36*, 34–40.

Persson, B., Kallberg, Y., Oppermann, U., & Jornvall, H. (2003). Coenzyme-based functional assignments of short-chain dehydrogenases/reductases (SDRs). *Chemico-Biological Interactions*, *143–144*, 271–278. https://doi.org/10.1016/S0009-2797(02)00223-5

Portevin, D., De Sousa-D'Auria, C., Houssin, C., Grimaldi, C., Chami, M., Daffe, M., & Guilhot, C. (2004). A polyketide synthase catalyzes the last condensation step of mycolic acid biosynthesis in mycobacteria and related organisms. *Proceedings of the National Academy of Sciences of the United States of America*, *101*, 314–319. https://doi.org/10.1073/pnas.0305439101

Radmacher, E., Alderwick, L. J., Besra, G. S., Brown, A. K., Gibson, K. J., Sahm, H., & Eggeling, L. (2005). Two functional FAS-I type fatty acid synthases in *Corynebacterium glutamicum*. *Microbiology*, *151*, 2421–2427. https://doi.org/10.1099/mic.0.28012-0

Raman, S., Vernon, R., Thompson, J., Tyka, M., Sadreyev, R., Pei, J., … Baker, D. (2009). Structure prediction for CASP8 with all-atom refinement using Rosetta. *Proteins*, *9*, 89–99.

Snapper, S. B., Melton, R. E., Mustafa, S., Kieser, T., & Jacobs, W. R., Jr. (1990). Isolation and characterization of efficient plasmid transformation mutants of *Mycobacterium smegmatis*. *Molecular Microbiology*, *4*, 1911–1919.

Song, Y., DiMaio, F., Wang, R. Y., Kim, D., Miles, C., Brunette, T., … Baker, D. (2013). High-resolution comparative modeling with RosettaCM. *Structure*, *21*, 1735–1742. https://doi.org/10.1016/j.str.2013.08.005

Stover, C. K., de la Cruz, V. F., Fuerst, T. R., Burlein, J. E., Benson, L. A., Bennett, L. T., … Bloom, B. R. (1991). New use of BCG for recombinant vaccines. *Nature*, *351*, 456–460. https://doi.org/10.1038/351456a0

Varela, C., Rittmann, D., Singh, A., Krumbach, K., Bhatt, K., Eggeling, L., … Bhatt, A. (2012). MmpL genes are associated with mycolic acid metabolism in mycobacteria and corynebacteria. *Chemistry & Biology*, *19*, 498–506. https://doi.org/10.1016/j.chembiol.2012.03.006

Vilcheze, C., & Jacobs, W. R. (2007) Isolation and analysis of *Mycobacterium tuberculosis* mycolic acids. *Current Protocols in Microbiology*, *Chapter 10*, Unit 10A.3.

Yamazawa, R., Nakajima, Y., Mushiake, K., Yoshimoto, T., & Ito, K. (2011). Crystal structure of serine dehydrogenase from *Escherichia coli*: Important role of the C-terminal region for closed-complex formation. *Journal of Biochemistry*, *149*, 701–712. https://doi.org/10.1093/jb/mvr024

Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J., & Zhang, Y. (2015). The I-TASSER Suite: Protein structure and function prediction. *Nature Methods*, *12*, 7–8. https://doi.org/10.1038/nmeth.3213

Yang, L., Lu, S., Belardinelli, J., Huc-Claustre, E., Jones, V., Jackson, M., & Zgurskaya, H. I. (2014). RND transporters protect *Corynebacterium glutamicum* from antibiotics by assembling the outer membrane. *Microbiologyopen*, *3*, 484–496.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Javid A, Cooper C, Singh A, et al. The mycolic acid reductase Rv2509 has distinct structural motifs and is essential for growth in slow-growing mycobacteria. *Mol Microbiol*. 2020;113:521–533. https://doi.org/10.1111/mmi.14437

Supplementary Information

The mycolic acid reductase Rv2509 has distinct structural motifs and is essential for growth in slow-growing mycobacteria

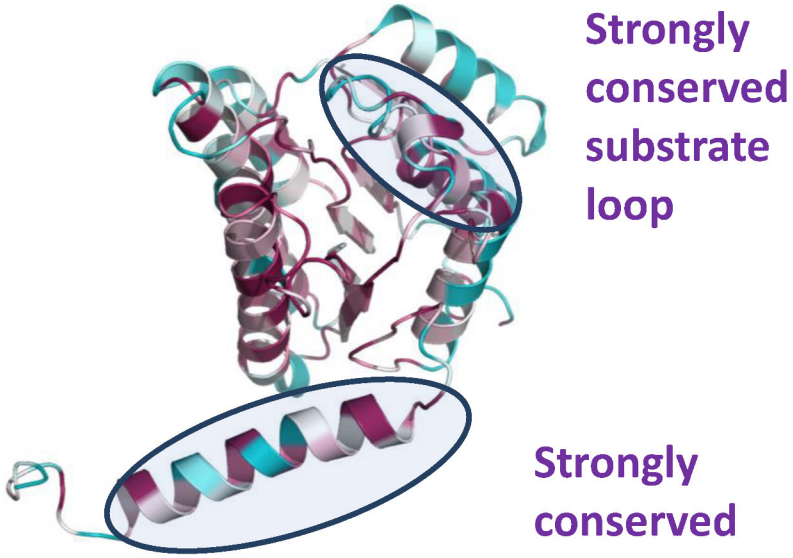
Asma Javid[1], Charlotte Cooper[1], Albel Singh[1], Steffen Schindler[2], Milena Hänisch[2], Robert L. Marshall[1], Rainer Kalscheuer[2], Vassiliy N. Bavro[3], Apoorva Bhatt[1]

[1]School of Biosciences and Institute of Microbiology and Infection, University of Birmingham, Birmingham, UK
[2]Institute of Pharmaceutical Biology and Biotechnology, Heinrich Heine University Düsseldorf, Düsseldorf, Germany
[3]School of Life Sciences, University of Essex, Colchester, UK

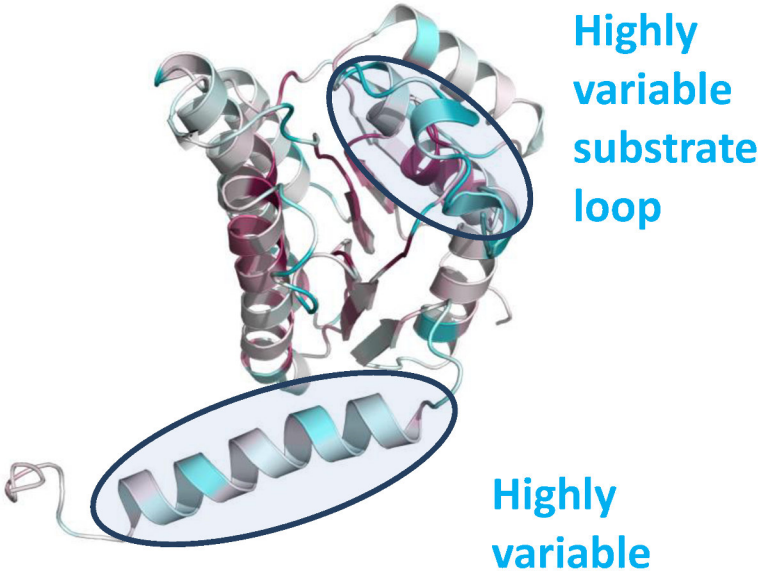
Supporting Information Fig. 1

Supporting Information Fig. 2

(A) – Consurf mapping of conservation from multiple sequence alignment based on the corynebacterial species represented in Fig1.



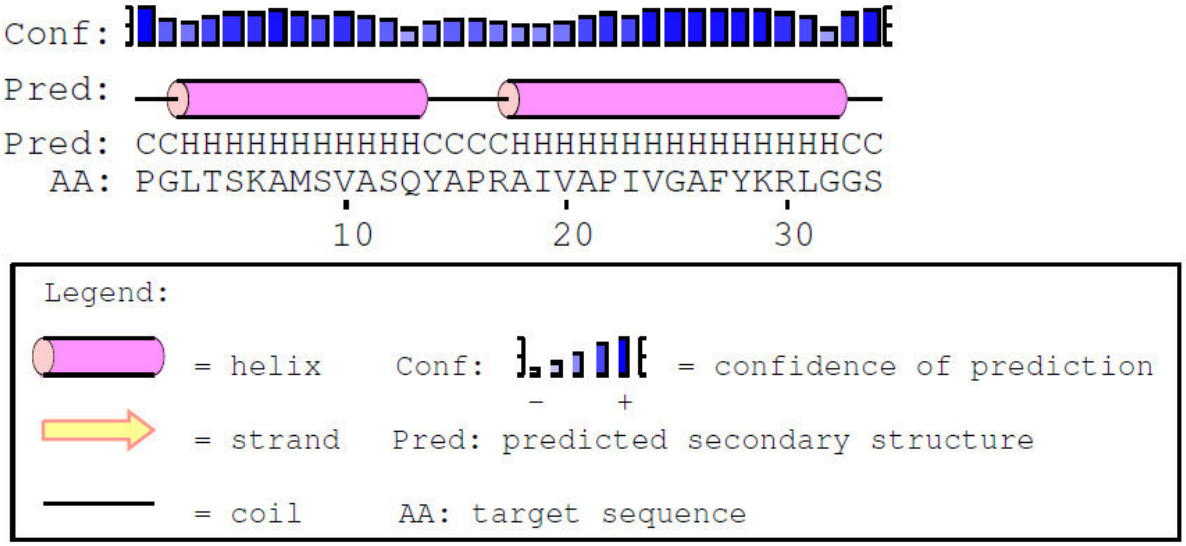
(B) Consurf mapping of conservation from multiple sequence alignment based on **expanded SDR homologues** shows conservation of the core, but not of the C-terminus and the substrate binding loop.



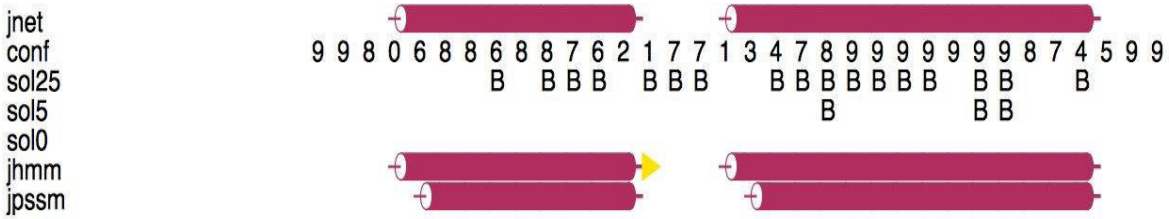
Supporting Information Fig. 3

Secondary structure prediction for the C-terminal domain of Rv2509 suggests strongly alpha-helical propensity (A) Output from PSIPRED and (B) JPRED4 output incorporating JNET.

(A)

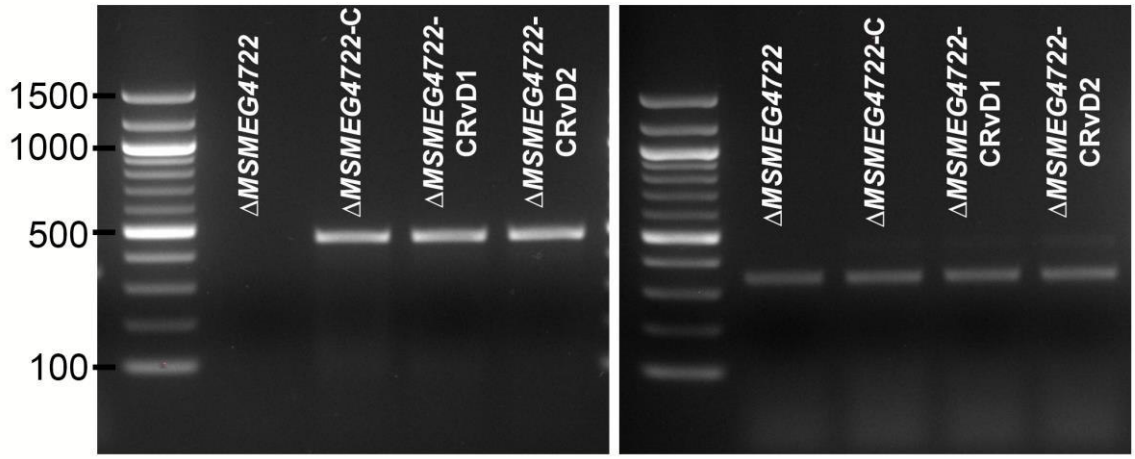


(B)



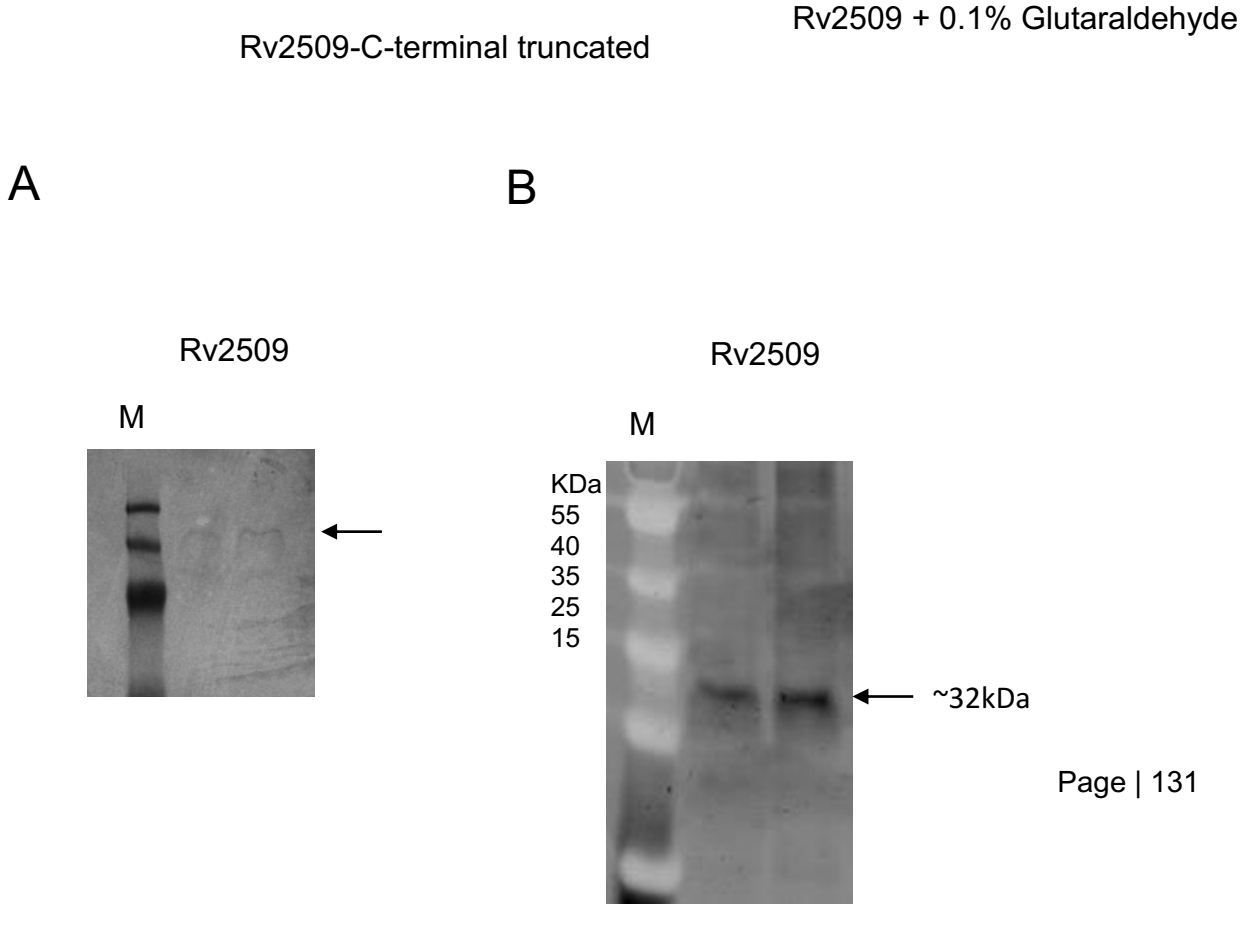
Supporting Information Fig. 4

Detection of *Rv2509* transcripts in *M. smegmatis* strains transformed with C-terminal deletion constructs by RT-PCR. (A) shows products for *Rv2509* while (B) shows products obtained for control (*sigA*).



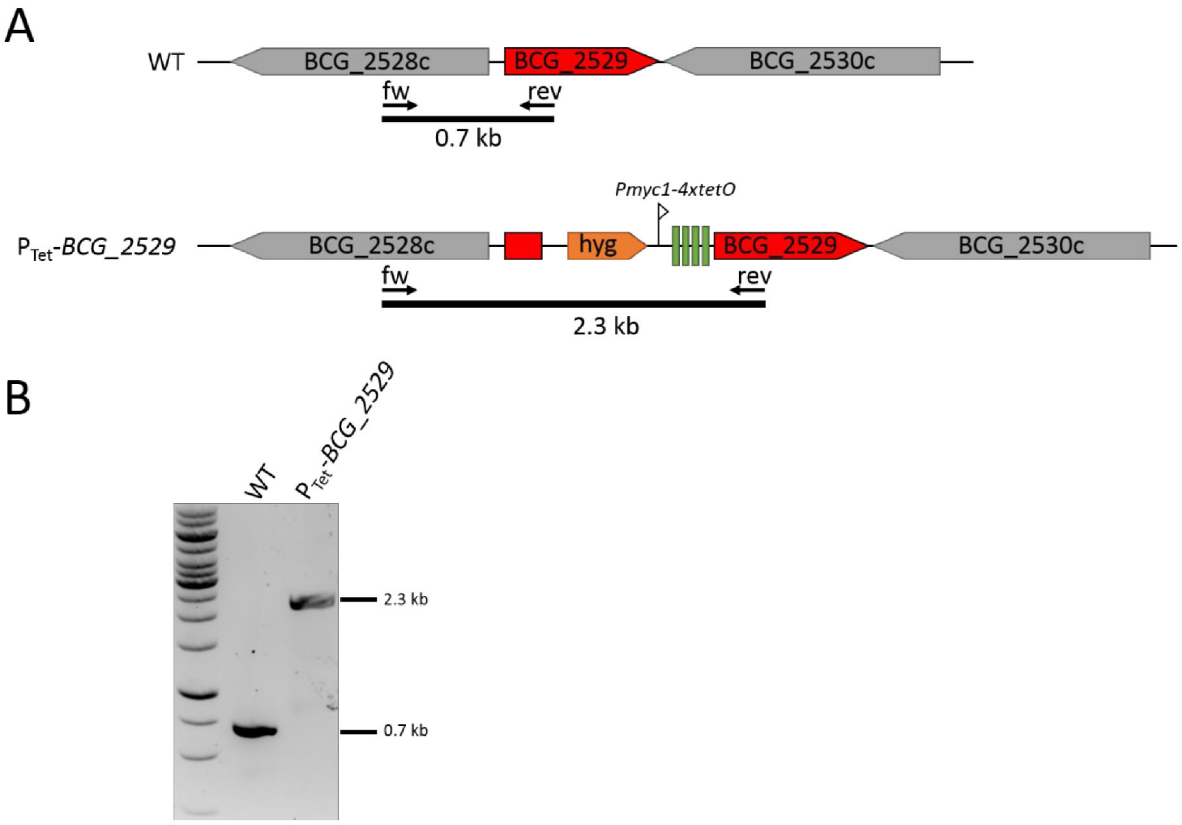
Supporting Information Fig. 5

Analysis of full length and C-terminal truncated Rv2509 by Native PAGE (A) and glutaraldehyde crosslinking (B). (A) shows a western blot of the Native gel and (B) shows SDS PAGE of untreated and glutaraldehyde treated Rv2509. M; PageRuler™ Prestained Protein Ladder (Thermo Fisher Scientific). Sizes are indicated in (B).



Supporting Information Fig. 6. Generation of a conditional *M. bovis* BCG-Pasteur BCG_2529 mutant.

(A) Organization of the *BCG_2529* locus in *M. bovis* BCG-Pasteur wild-type as well as in a *BCG_2529* conditional mutant. The sizes of relevant fragments as well as the location of the oligonucleotides used for diagnostic PCR are indicated. WT, wild-type; fw, oligonucleotide *BCG_2529*_L_fw; rev, oligonucleotid *BCG_2529*_L_rev; hyg, hygromycin resistance gene. (B) Diagnostic PCR of genomic DNA using oligonucleotides binding to the position indicated in A, showing the insertion of a synthetic gene cassette (*hyg*-P*myc1*-4X*tetO*) that contained a duplication of 66 bp of the 5´-end of *BCG_2529*.



8. DISCUSSION AND PERSPECTIVES

In this section, we will discuss the findings of the publications and place them in the current picture of AMR research. Therefore, we will first outline the results of the publications, followed by a summary of antibiotic research, particularly on chemical compound screening and rational design of drugs. In order to capture the complexity of the topic, we briefly introduce other fields of antimicrobial research before assessing the current situation and future prospects.

In order to find valuable drug targets to fight *M. tuberculosis*, our studies focus on identifying and evaluating several elements involved in both the biosynthesis and assembly of the mycobacterial cell wall, as well as in essential metabolic processes. Although there is a reliable regimen to fight TB, further research on additional drug targets is urgently needed, as TB continues to be a global threat, killing millions of people each year ([9]). In general, the treatment of TB is complicated by the adaptation of the mycobacteria to the human host, as the innate immune cells, normally responsible for eradicating pathogens, instead provide a unique niche to reside and propagate ([15]). In addition to the challenges posed by *M. tuberculosis* evading the immune system, there is an alarming increase of MDR and XDR strains that underlines the pathogen's ability to overcome countermeasures that have historically been effective in treating TB. Despite diverse advances in drug development, whether through screening approaches or rational design, many newly discovered antitubercular agents have unfortunately proven ineffective in the fight against TB. While public perception of threats to human survival mainly revolves around global warming, scientists warn that antibiotic resistances could become the greatest threat to humanity, potentially resulting in millions of deaths every year ([50]).

To counter the alarming trend of accelerating resistance development in the context of stagnant drug discovery, our efforts aimed on introducing innovative strategies and/or adding new candidates to the list of valuable drug targets in *M. tuberculosis*. Therefore, we focused on a relatively young subject in *M. tuberculosis*: ncRNA. Although it has been emphasized in recent years that ncRNAs are more than mere blueprints for proteins, research on tubercular ncRNAs is limited. Although RNA-based mechanisms such as riboswitches have long been known, the list of independent ncRNAs that exert their regulatory function through base-pairing or association with proteins is constantly increasing. The genomic location of ncRNA provides information about the putative target spectrum. While cis-acting ncRNAs possess complete complementary properties to the gene they oppose, trans-acting, intergenic ncRNAs interact with targets despite imperfect affinity and therefore demonstrate a broader target spectrum. To facilitate trans-acting ncRNA-mRNA connection, many organisms rely on the chaperone Hfq, which is either not existing or has not yet been found in *M. tuberculosis* ([141]). However, there are examples of organisms using ncRNAs independent from chaperones, such as

S. aureus ([143]). Given the unusually high GC content of *M. tuberculosis* DNA, a chaperone-independent ncRNA-mRNA duplex may be possible as hydrogen bonds between guanine and cytosine thermodynamically stabilize duplex formations more effectively than those between adenine and thymine. The use of ncRNAs is particularly interesting concerning pathogens, as several examples corroborate involvement in virulence. Potential drug targets include ncRNAs that are involved in intracellular processes or associated with pathogenic traits, as influencing these molecules could immensely affect the pathogen.

In Chapter 4: "Redundant functions of the non-coding RNAs MTS1338 and MTS2823 in *Mycobacterium tuberculosis*", we analyzed two highly abundant ncRNAs that accumulate in the stationary phase and dramatically affect the gene expression machinery. Although the absence of MTS1338 or MTS2823 in single-knockout approaches exhibited the individual non-essentiality of both ncRNAs and resulted in no apparent growth phenotype, transcriptome sequencing and proteome profiling demonstrated tremendous changes in response to ncRNA-deprivation. The effect on ribosomal proteins was particularly interesting, as a lack of MTS1338 or MTS2823 resulted in higher abundance, suggesting a possible function in silencing the gene expression during the transition to and/or in stationary phase. Another intriguing finding was that the mutants deprived of MTS1338 or MTS2823 revealed unexpected similar gene expression patterns and protein abundances, suggesting at least partial functional homology. Since silencing of gene expression is a vital process in microorganisms, and the mere abundance of the studied ncRNAs indicates certain importance, it was remarkable to find that these genes did not demonstrate essentiality. According to our studies, we concluded that MTS1338 and MTS2823 could exert redundant functions in silencing expression of a broad range of target genes and thus could complement for the absence of the other ncRNA in single-knockout approaches. The conditional essentiality of MTS1338 in a mutant lacking MTS2823 and *vice versa* should to be addressed in the future to evaluate the importance of the respective ncRNAs, potentially by constructing double-knockout or conditional silenced mutants. However, the general use of redundant ncRNAs that elicit important functions is already known in *E. coli*. The carbon storage regulator (Csr) system is a prominent example of the use of the functionally similar ncRNAs CsrB and CsrC, both of which bind the protein CsrA ([151]). CsrA conducts wide-spread functions in modulating, for instance, intracellular carbon flux, as it is able to exert positive and negative regulation of essential carbon-associated processes such as gluconeogenesis, glycolysis and glycogen biosynthesis, the latter being a process that we also studied in *M. tuberculosis* in Chapter 6 ([152]). Involvement in an essential and virulence-associated process qualifies MTS1338 and MTS2823 as attractive drug targets. Although the mode of action by which the ncRNAs influence ribosomal proteins is yet unknown, the ncRNAs could be targeted by small molecules or antisense RNAs. Deprivation of both

ncRNAs could interfere with the transition to stationary phase, creating a completely new situation concerning *M. tuberculosis* infections. In this hypothetical model, the immune system could cope with the mycobacterial infection because although the much-discussed granuloma would still shield the bacteria from antibiotics, the bacteria would be locked in a nutrient-poor niche and starve to death as a result of being unable to slow down their metabolism and switch to persistence. Further research on the mode of action of MTS1338 and MTS2823 may enable the design of small drug-like molecules that can substantially impair the silencing of the translation machinery, a crucial process for the establishment of a persistence phenotype.

While the ncRNAs analyzed above show great potential for use in antitubercular therapy, we also turned our attention to one of the most prominent targets for antibacterial agents: the bacterial cell wall.

Since Alexander Fleming discovered penicillin in 1928, the cell wall of pathogens has been considered an attractive drug target ([153]). This is also true for *M. tuberculosis*, as, for instance, the first-line antibiotics isoniazid and ethambutol target the mycobacterial cell wall ([34,37]). The unusual structure and composition of the tubercular cell wall not only reduces susceptibility to a variety of drugs, but also provides a reservoir of susceptible targets for therapy. In Chapter 5: "The GtrA-like lipid floppase Rv3277 is essential for cell surface mannosylation in *Mycobacterium tuberculosis*", we studied a protein involved in the biosynthesis of glycolipids, which play an important role in the pathogenicity of *M. tuberculosis*. While the immunomodulatory roles of, for instance, PIM, LM, and LAM have been revealed in various studies and their biosynthesis has already been studied in detail (see Chapter 1.5), the biosynthetic pathway of these components was still missing a crucial step. The biosynthesis of glycolipids starts in the cytoplasm with processing of a precursor molecule, which is subsequently translocated to the periplasm, where further components are added to form complete structures ([118]). The intracellular biosynthesis of lower-order PIMs and the process by which they are extracellularly expanded with additional carbohydrates have been deciphered in detail. However, the process by which the PIMs, mannose, and arabinose required for extracellular expansion are delivered to the periplasm has not yet been identified.

In this context, we analyzed Rv3277, a protein that showed a close connection to the abundance of several prominent cell wall glycolipids. In addition to *M. tuberculosis*, we conducted several experiments with a close relative, *M. smegmatis*, in which a functional homolog of Rv3277, Msmeg_1817, was identified and proven to be a functional homolog by complementation experiments. This allowed conclusions to be drawn for both genes and organisms, respectively. Immunodetection experiments and GC analyses demonstrated that periplasmic formation of glycolipids depends on function of Rv3277, as the amounts of mannose and arabinose incorporated into the cell wall, and ultimately the levels of PIM, LM,

and LAM, correlated with Msmeg_1817 expression. 3D modeling employing the TopModel suite suggested that Rv3277 is a transmembrane protein that may transfer DPM and DPA from the cytosol to the periplasm and thus may be the protein that is missing in the current understanding of glycolipid formation. According to the 3D model, Rv3277 demonstrates a striking homology to the TM domain of the dolichyl phosphate mannose synthase (DPMS) of *Pyrococcus furiosus*, a protein that conducts the transfer of carbohydrates to a lipid backbone and its subsequent translocation through the membrane. Based on the absence of a DPMS-like synthase domain in Rv3277, we suggested the use of different enzymes for processing of DPM or DPA prior to translocation. The identification of the Ppm1, a protein that performs the transfer of mannose to the PI backbone, supports our hypothesis ([114]). Moreover, proteome profiling indicated the involvement of Rv3277 in the biosynthetic pathway leading to the formation of glycolipids, as insufficient expression of Rv3277 resulted in immense cell wall stress. Among several differently expressed proteins, we detected the upregulation of, for instance, Ppm1, PimB, and MptA; proteins which are directly involved in the same biosynthetic pathway that we hypothesize Rv3277 to be involved in ([114,118,154]). The increased expression of these proteins may represent adaptive stress responses to compensate for the lack of function of Rv3277. In conclusion, the unique cell wall protects mycobacteria from certain stresses and renders the organism less susceptible to antibiotics, but also provides vulnerable targets for therapy. Rv3277 joins recently studied antibiotic targets such as DprE1, which are also associated with the cell wall and represent valuable drug targets due to their essentiality and localization ([111]). The discovery or development of inhibitors for these proteins could facilitate the treatment of TB.

In order to gain further insights into essential elements of the cell wall assembly that could represent valuable drug targets, we analyzed *glgE* in Chapter 6: “A temperature-sensitive *Mycobacterium smegmatis glgE* mutation leads to a loss of GlgE enzyme activity and thermostability and the accumulation of α-maltose-1-phosphate”. In previous studies, *glgE* was classified as being involved in glycogen degradation. Given the state of knowledge at the time, categorizing GlgE as a glycogen-degrading enzyme was comprehensible. However, in context of the present thesis, we deciphered that the proposed glucanase function of GlgE was based on a misinterpretation of observed phenotypes ([155]). Conditional silencing of wild type *glgE* with concomitant expression of mutated *glgE* produced a temperature-sensitive phenotype resembling the mutant that was analyzed two decades ago. In the context of low activity of GlgE, accumulation of M1P, the building blocks of glycogen, pointed towards a polymerizing function we attribute to GlgE. In contrast to previous findings, we suggest that GlgE is involved in the construction of glycogen rather than recycling it for energy purposes. Since glycogen is a crucial part of the cell envelope and is involved in mechanisms that, for instance, modulate

the hosts immune response, the potential of interfering with the function of the GlgE is thought to be exploitable for therapeutic advances ([73]). The involvement in an essential process directly associated with virulent traits and the observation that a loss of function of GlgE resulted in a toxic accumulation of M1P that eventually slows bacterial growth highlights *glgE* as a new drug target in *M. tuberculosis*.

In Chapter 7: “The mycolic acid reductase Rv2509 has distinct structural motifs and is essential for growth in slow-growing mycobacteria”, we analyzed the gene Rv2509, which encodes a protein important for the biosynthesis of mycolic acids ([86]). The biosynthetic pathway of mycolic acids is a complex process including various enzymes that modify molecules leading to the formation of large lipid structures that are eventually located in the mycobacterial cell wall (for further information, see Chapter 1.5). The intracellular biosynthesis of mycolic acids requires a final reduction of a mycolic β-ketoester prior to translocation *via* MmpL3, which is conducted by Rv2509. In contrast to the non-essential homologous genes in fast-growing mycobacteria such as *M. smegmatis*, we demonstrated the essentiality of Rv2509 in slow-growing *M. bovis* BCG. This publication also sheds light on the unique features of Rv2509 that could lead to categorization into a new family of short-chain dehydrogenases as the NADP(H)-binding motif and also the core stabilization sequence differ significantly from the classical mycobacterial standards. Using 3D modeling, the C-terminus of Rv2509 demonstrated an α-helical extension that is diverse from the classical structure of SDR family members and essential for the function of the protein. It has been suggested that the C-terminal region may facilitate oligomerization, although we did not find evidence to support this hypothesis. The important role of Rv2509 in mycobacterial growth underlines its potential as a valuable drug target in *M. tuberculosis*.

In summary, we analyzed five potential drug targets of *M. tuberculosis* and highlighted their potential regarding antitubercular therapy. We analyzed the two ncRNAs MTS1338 and MTS2823, revealing their involvement in regulating the gene expression, as well as the three proteins Rv3277, Rv2509 and GlgE, which demonstrated essentiality for cell wall assembly or intracellular and capsular glucan formation. Interfering with the function of the respective target could facilitate the treatment of TB. While each manuscript outlines our rationale for proposing the analyzed structures as valuable drug targets, the following section situates these discoveries in the broader context of current antimicrobial research.

In order to eradicate TB, the WHO monitors and guides progress (see Chapter 1.1). Due to these efforts, approx. 43 million lives were saved between 2000 and 2014, and the TB-associated mortality rate decreased by 47% ([8]). The current objective is to reduce TB-related deaths by 90% by 2030, compared to 2015 ([8]). While this thesis focuses on *M. tuberculosis*, AMR is a threat posed by a variety of microorganisms. For this reason, the WHO and institutions such as the Centers for Disease Control and Prevention (CDC) publish lists of MDR

pathogens that categorize and evaluate the urgency of drug development measures ([156,157]). In addition to *M. tuberculosis*, the ESKAPE pathogens (*Enterococcus faecium*, *S. aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa*, and *Enterobacter species*) are of particular interest as they tend to develop resistance against antimicrobial agents.

There are several key components in fighting AMR, one of which is diagnostics. Although the WHO publishes guidelines for MDR-TB treatment tailored to the respective resistance, pathogens are spread, for instance, due to inadequate or insufficient diagnostics ([158]). As a consequence of a lack of diagnosis or misdiagnosis, broad-spectrum antibiotics are increasingly applied. The excessive use of broad-spectrum antibiotics accelerates the proliferation of resistant pathogens, as they unintentionally destroy the gut microbiome and thereby create free niches that can be colonized ([159]). But even if screening for AMR were to occur more regularly, there are significant barriers to appropriate treatment, such as the distribution of antibiotics to economically underdeveloped countries ([160]). The general issues associated with the handling of antibiotics, which are discussed in more detail in Chapter 1.3, promote the emergence of resistance development.

For such sensitive issues, it seems appropriate that different higher-level institutions such as WHO, CDC, or NIH (National Institutes of Health) monitor and/or manage progress. Considering the antibiotic discovery void and the prediction that completely resistant bacteria will become a major challenge for treatments, it is important to approach antibiotic resistance from various perspectives. As the versatility of intervention methods could be purposeful for the treatment of resistant pathogens, the following sections provide a brief overview of the vibrant areas of AMR research to contextualize the discoveries of this study. Although host-directed therapy is a growing field in antitubercular therapy and would be worth exploring in more detail, the following sections focus on therapeutics that specifically target the pathogen rather than supporting the host cell.

In the past, screening of chemical substances derived from environmental samples for antimicrobial activity has yielded a diverse reservoir of molecules. However, innovative compounds have not been found for a long time, leading to the so-called antibiotic discovery void ([57]). Activity-based screening of antimicrobial compounds is expensive and time-consuming, as it requires the mode of action of the identified compounds to be elucidated and validated *in vivo*, along with their cytotoxicity. This complex process makes it unattractive for pharmaceutical companies to invest, as the relatively low profit margin for short-term therapies does not compensate for the time and expense involved. Despite the approx. $1.4 billion funding gap identified by the WHO, this approach continues to be pursued, but declining funding is causing future research efforts to shift to other areas of antimicrobial research ([8]).

In rational design, pathogens are addressed with a divergent approach. Since screening compounds for antimicrobial properties is based on hit identification, further research is required to validate the target and the mode of action. In contrast, rational design of antibiotics describes the process of developing antimicrobial agents tailored to a given target, such as interfering with bacterial gene expression. Designing molecules that act specifically on the target requires in-depth knowledge of the potential targets, including their structure and involvement in processes or networks. In addition, as with any clinically used antibiotic, therapeutics must have a certain specificity, since killing bacteria is of no use if the cytotoxic side effects also affect eukaryotic cells. The analysis of targets and design of suitable inhibitors require a) basic laboratory research, b) bioinformatics knowledge and tools, c) genetic engineering or chemical synthesis, and d) validation studies.

In conclusion, the use of molecules as antimicrobial therapeutics, based on either hit identification or rational design, requires a comprehensive understanding of inter- and intracellular dependencies, as well as bioinformatics and laboratory skills, tools, time, and financial resources. However, even if new compounds were found or developed, one of the obstacles to chemical therapeutics is the application. In TB, the defense mechanisms of the human immune system complicate the use of molecules that directly target the pathogen. Antibiotics must navigate through a cluster of human cells, the granuloma, before reaching the thick cell wall that must be penetrated to achieve intracellular localization (see chapter 1.2 and 1.5.1). Despite this physical barrier, it is difficult to find appropriate antibiotics to treat MDR pathogens, especially when dealing with multidrug efflux systems that confer resistance against multiple antibiotics by exporting the compounds ([161]).

As the golden era of antibiotics appears to be coming to an end, alternative therapeutic approaches, including some previously utilized methods, are regaining attraction in managing hard-to-treat cases. The following sections briefly describe the vibrant areas of antimicrobial research that do not revolve around conventional chemical compounds. Advances in therapies using A) small interfering RNA (siRNA), B) phages, C) humanized antibodies, or D) antimicrobial peptides (AMPs) are expected to facilitate the treatment of pathogenic infections as a stand-alone therapy or antibiotic adjuvant.

A) When considering influencing the expression of pathogenic genes, RNA molecules are perceived as valuable targets. In particular, 3D structures such as loops or hairpins provide potential binding sites for molecules that interfere with the translation machinery and/or association with nucleic acids or proteins. In the context of ncRNAs, which exert their regulatory function mainly *via* binding to mRNA targets, it seems reasonable to disrupt ncRNA-mRNA duplex formation by the use of "decoy-ncRNAs". Recently, the use of RNA-based therapeutics has gained increasing attention as bioinformatics tools such as LncTar

enable the prediction of RNA targets, facilitating research approaches ([162]). So-called siRNAs bind to target RNAs and interfere with the execution of the intended function. One example is the microRNA (miRNA)-targeted drug miravirsen, which is used in human cells to selectively sequester miR-122 to prevent the spread of hepatitis C ([163,164]). The liver-expressed miR-122 is an ncRNA that promotes viral replication by binding the 5'-UTR of the viral genome, protecting HCV RNA from degradation ([165,166]). To date, siRNA approaches have yielded more than 20 therapeutics for the treatment of cancer or infectious diseases, although the route of application is still an obstacle that several studies are addressing ([167]). As the example of miravirsen shows, antiviral siRNAs need to be introduced into human cells, which could allow administration by injection or viral carriers. With respect to *M. tuberculosis*, the strategy of applying siRNA allows for two concepts, one being host cell-directed and the other targeting the pathogen. Since *M. tuberculosis* releases membrane vesicles containing molecules that modulate immune responses, and possibly also nucleic acids that help the bacterium evade the immune system's detection, these nucleic acids could potentially be targeted for neutralization by siRNA molecules. This interference could help suppress the bacterium's ability to disrupt immune defenses, such as the fusion of phagosomes and lysosomes, which are critical for the immune system to effectively fight the pathogen ([168]). The other hypothetical model involves the use of siRNA to modulate essential processes in mycobacteria, although the route of administration would face the same challenges as conventional antibiotics. However, progress with other pathogens such as *S. aureus* demonstrates the potential of siRNA approaches for the treatment of pathogens that have become problematic in terms of resistance development ([169]).

B) Phages are bacteria-specific viral particles consisting of a protein capsule and genetic material, either RNA or DNA. The reproduction cycle of phages involves two distinct routes that revolve around bacterial hosts and leave eukaryotic cells unaffected. Upon introduction of their genome into a bacterial host cell, the genetic material can either be introduced into the bacterial DNA, leading to the lysogenic life cycle, or used to replicate within the bacteria, eventually lysing the cell and releasing phages into the environment *via* the lytic pathway (reviewed in [170]). Although phage therapy has been widely used in Eastern Europe, the lack of strict guidelines has not yielded reliable clinical trials, a circumstance that the EMA and FDA are addressing in recent approaches ([171]). Studies using mice or wax moths as model organisms have demonstrated the efficacy of phage treatment, as biofilms of *P. aeruginosa* were effectively disrupted ([172]). Since biofilms contribute to pathogens becoming less susceptible to antibiotics, the observation of re-sensitization of MDR

pathogens promotes the idea of using combination therapies of phages and antibiotics ([173]). In addition to model organisms, phage therapy has been shown to be effective in treating human patients infected with, for instance, *A. baumanii*, a pathogen that belongs to the group of ESKAPE pathogens ([174]). Since the development of resistance is also an issue in phage therapy, phages are used in the form of cocktails, i.e., the simultaneous application of different phages, to reduce the risk of AMR (analogous to TB treatment with a mixture of antibiotics, see 1.3). With respect to antibiotic compounds, the basic biology behind this approach is a major advantage ([175]). Particularly in TB, coevolution of *M. tuberculosis* with the human host poses a problem for treatment. However, the same idea can be exploited when realizing that phages also evolve with bacterial hosts, leading to mechanisms that overcome defense mechanisms ([176]). While phage therapy has advantages over other forms of antimicrobial treatment, it also faces obstacles. In addition to the lack of clinical studies, the route of administration limited the use of phages in the past, as it has been restricted to topical application. Advances in intranasal, intravenous, and percutaneous administration have shown desired effects, but the efficacy still needs to be optimized to facilitate the treatment of MDR pathogens in the future ([172,174]).

C) Humanized antibodies are optimized immunoglobulins of other species that are used for the treatment of human patients. Antibodies have long been used to treat infections, although the era of antibiotics rendered this expensive treatment almost obsolete. Palivizumab, the only monoclonal antibody approved for treatment of an infectious disease, is administered primarily to children prone to respiratory infections ([177]). However, several studies have demonstrated how antibodies intervene in bacterial infections, for instance by disrupting biofilms or by bactericidal activities ([178,179]). In addition, cytotoxicity and off-target effects, such as those affecting the gut microbiome, are minimized when using pathogen-specific antibodies, circumventing a major difficulty associated with the use of chemical agents in the treatment of human patients ([180]). While the use of antibodies is advantageous in the treatment of extracellular pathogens, intracellular organisms such as *M. tuberculosis* are difficult to target with antibodies due to the poor internalization of these macromolecules. Another disadvantage of using monoclonal antibodies is that minor amino acid changes in the antigen can reduce efficacy, making antibody production less attractive for therapeutic treatments. In this regard, using polyclonal antibodies targeting multiple epitopes of the antigen is a promising but expensive attempt to target pathogens. This approach is currently being tested in phase II clinical trials for the treatment of SARS-CoV-2 ([181]). While discussing the potential of antibodies in TB therapy, the

emergence of MDR and XDR pathogens has reinvigorated efforts in the field of monoclonal antibodies, leading to advances in the treatment of, for example, *S. aureus*, *Clostridium difficile*, or HIV ([182–184]).

D) The innate immune system includes substantial elements known as AMPs. Providing an overview of this diverse field of AMPs is difficult, as their occurrence and functions are manifold, a circumstance that led to the assumption of a low probability of resistance development ([185]). Currently, more than 3,400 AMP are registered in the antimicrobial peptide database (APD), originating from bacteria, archaea, protists, fungi, plants, animals, and synthetic origin ([186]). As functions have been demonstrated to range from antimicrobial to immune response-stimulating properties, AMPs are promising an immense potential for the use in therapeutic regimens (reviewed in [187–189]). Due to the variety of structures or chemical compositions of the different AMPs, it is impossible to describe an unambiguous mode of action of the AMPs. However, several publications underline the antibiotic potential of AMP based on the disruption of microbial membranes, as exemplified by lactoferrin B ([190]). There are different theories about how the disruption occurs, but models based on the cationic properties of AMPs are mentioned frequently, although the efficacy in physiological environments containing cations such as Mg^{2+}, Ca^{2+}, or Na^{+} is questioned ([191,192]). In addition to membrane disruption, AMPs are known to affect intracellular processes such as gene expression or replication. Although several obstacles impede the clinical use of AMPs, such as low efficacy, high cost, toxicity, and protease degradation, scientists recognized the potential of AMPs for the treatment of infections, leading to efforts to artificially optimize AMPs to develop so-called synthetic mimics of antimicrobial peptides (SMAMP, reviewed in [193]). Concerning the organism investigated in context of this thesis, the APD contains 14 AMPs that show promising activity against *M. tuberculosis*. For example, RNase7 has been shown to be expressed in airway epithelial cells in response to *M. tuberculosis* infection, and teixobactin, a cell wall biosynthesis-inhibiting AMP, demonstrates great potential for therapeutic use in infections with *S. aureus* or *M. tuberculosis* ([185,186,194]).

In this last section, we summarize the key messages of this thesis to recapitulate the main idea of this work. In the publications stated in Chapters 4 to 7, we identified several elements of *M. tuberculosis* as potential drug targets and provided scientific evidence to support this assessment. Therefore, we investigated metabolically relevant ncRNAs and essential elements for the establishment of the mycobacterial cell wall and eventually demonstrated a

connection to pathogenic properties. In total, five potential drug targets were identified that could be therapeutically exploited using the approaches explained above. To raise awareness and to remind us that the problem of antibiotic resistance is an ever-increasing threat not only for *M. tuberculosis*, we reviewed the world of antimicrobial research, highlighting historical and modern approaches, including those aside from antibiotic compounds. Although antibiotics are our best defense against pathogens and we have been involved in expanding the diverse portfolio of potential drug targets, it is important to recognize that there are other promising approaches to facilitate the fight against antibiotic resistance. To this end, we briefly described four areas of AMR research beyond the conventional antibiotics in the form of chemical compounds that have been used in combination therapies to treat TB since the 1960s. However, the term “combination” referred to different chemical compounds and not the additional use of, for instance, phages. Targeting pathogens from different angles, e.g., by combining antibiotics with the simultaneous application of phages, humanized antibodies, or AMPs could promote disease treatment while reducing the risk of resistance development.

The motivation behind the projects of this thesis is clear to anyone involved in healthcare or disease control. However, the public may not be as aware as the scale of the problem actually necessitates. While global warming is considered the greatest threat to humanity, people are only beginning to understand the menace that is growing within and around us. Efforts to develop antimicrobial therapies resemble an arms race between resistance development and drug discovery, with one competitor finding ways to circumvent the strategies of the other. In order to witness the outcome of threats such as global warming, humanity must prevail in the fight against microorganisms, as resistant pathogens are predicted to be the leading cause of death by 2050 ([50]). In this study, we investigated treatment approaches with respect to TB, although the underlying principle is applicable to different pathogens as well. By analyzing several putative drug targets, we have gained insights into *M. tuberculosis* that will enable rational design of drugs and support the fight against tuberculosis.

9. REFERENCES

1. Cousins, D. V. *et al.* Tuberculosis in seals caused by a novel member of the Mycobacterium tuberculosis complex: Mycobacterium pinnipedii sp. nov. *Int. J. Syst. Evol. Microbiol.* **53**, 1305–1314 (2003).
2. Wells, A. Q. Tuberculosis in Wild Voles. *Nature* **139**, 917–917 (1937).
3. Phillips, C. J. C., Foster, C. R. W., Morris, P. A. & Teverson, R. The transmission of Mycobacterium bovis infection to cattle. *Res. Vet. Sci.* **74**, 1–15 (2003).
4. Cañdido, P. H. C. *et al.* Multidrug-resistant nontuberculous mycobacteria isolated from cystic fibrosis patients. *J. Clin. Microbiol.* **52**, 2990–2997 (2014).
5. Kentley, J. *et al.* Intestinal tuberculosis: a diagnostic challenge. *Trop. Med. Int. Heal.* **22**, 994–999 (2017).
6. Hickey, A. J., Gounder, L., Moosa, M. Y. S. & Drain, P. K. A systematic review of hepatic tuberculosis with considerations in human immunodeficiency virus co-infection. *BMC Infect. Dis.* **15**, 1–11 (2015).
7. Tahir, M., Nida, A. & Qamar, S. Hiding in the bone: a case of miliary tuberculosis with bone marrow involvement. *AME Case Reports* **2**, 36–36 (2018).
8. World Health Organization. The End Strategy TB. *End TB Strateg.* **53**, 1689–1699 (2015).
9. WHO. *Global tuberculosis report 2023*. *January* (2023).
10. Houben, R. M. G. J. & Dodd, P. J. The Global Burden of Latent Tuberculosis Infection: A Re-estimation Using Mathematical Modelling. *PLoS Med.* **13**, 1–13 (2016).
11. World Health Organization. *Global tuberculosis report 2020*. (2020).
12. RILEY, R. L. Aerial dissemination of pulmonary tuberculosis. *Am. Rev. Tuberc.* **76**, 931–941 (1957).
13. Li, Y., Wang, Y. & Liu, X. The role of airway epithelial cells in response to mycobacteria infection. *Clin. Dev. Immunol.* **2012**, (2012).
14. Armstrong, J. A. & Hart, D. Response of cultured macrophages to Mycobacterium Tuberculosis, with observations on fusion of lysosomes with phagosomes. *J. Exp. Med.* **134**, 713–740 (1971).
15. Queval, C. J., Brosch, R. & Simeone, R. The macrophage: A disputed fortress in the battle against Mycobacterium tuberculosis. *Front. Microbiol.* **8**, 1–11 (2017).
16. Simeone, R. *et al.* Phagosomal rupture by Mycobacterium tuberculosis results in toxicity and host cell death. *PLoS Pathog.* **8**, (2012).
17. Shin, D. M. *et al.* Mycobacterium tuberculosis eis regulates autophagy, inflammation, and cell death through redox-dependent signaling. *PLoS Pathog.* **6**, (2010).
18. Miller, J. L., Velmurugan, K., Cowan, M. J. & Briken, V. The type I NADH dehydrogenase of Mycobacterium tuberculosis counters phagosomal NOX2 activity to inhibit TNF-α-mediated host cell apoptosis. *PLoS Pathog.* **6**, 1–14 (2010).
19. Via, L. E. *et al.* Arrest of mycobacterial phagosome maturation is caused by a block in vesicle fusion between stages controlled by rab5 and rab7. *J. Biol. Chem.* **272**, 13326–13331 (1997).
20. Ladel, C. H. *et al.* Lethal tuberculosis in interleukin-6-deficient mutant mice. *Infect.*

Immun. **65**, 4843–4849 (1997).

21. Di Paolo, N. C. *et al.* Interdependence between Interleukin-1 and Tumor Necrosis Factor Regulates TNF-Dependent Control of Mycobacterium tuberculosis Infection. *Immunity* **43**, 1125–1136 (2015).

22. Flynn, J. A. L. *et al.* Tumor necrosis factor-α is required in the protective immune response against mycobacterium tuberculosis in mice. *Immunity* **2**, 561–572 (1995).

23. Zhang, S. Y. *et al.* Inborn errors of interferon (IFN)-mediated immunity in humans: Insights into the respective roles of IFN-α/β, IFN-γ, and IFN-λ in host defense. *Immunol. Rev.* **226**, 29–40 (2008).

24. Okoye, A. A. & Picker, L. J. CD4+ T-Cell Depletion In Hiv Infection: Mechanisms Of Immunological Failure. *Immunol. Rev.* **254**, 54–64 (2013).

25. Cicchese, J. M., Dartois, V., Kirschner, D. E. & Linderman, J. J. Both Pharmacokinetic Variability and Granuloma Heterogeneity Impact the Ability of the First-Line Antibiotics to Sterilize Tuberculosis Granulomas. *Front. Pharmacol.* **11**, 1–15 (2020).

26. Refai, A., Gritli, S., Barbouche, M. R. & Essafi, M. Mycobacterium tuberculosis virulent factor ESAT-6 drives macrophage differentiation toward the pro-inflammatory M1 phenotype and subsequently switches it to the anti-inflammatory M2 phenotype. *Front. Cell. Infect. Microbiol.* **8**, 1–14 (2018).

27. Pai, M. *et al.* Tuberculosis. *Nat. Rev. Dis. Prim.* **2**, (2016).

28. Jo, K. W. *et al.* Risk factors for 1-year relapse of pulmonary tuberculosis treated with a 6-month daily regimen. *Respir. Med.* **108**, 654–659 (2014).

29. World Health Organization. Treatment of Tuberculosis - Guideline, Fourth Edition. **249**, 344–345 (2010).

30. Wehrli, W. & Staehelin, M. Actions of the rifamycins. *Bacteriol. Rev.* **35**, 290–309 (1971).

31. Telenti, A. *et al.* Detection of rifampicin-resistance mutations in Mycobacterium tuberculosis. *Lancet* **341**, 647–651 (1993).

32. Isoniazid in Pulmonary Tuberculosis. *Br. Med. J.* **2**, 443 (1952).

33. Johnsson, K. & Schultz, P. G. Mechanistic Studies of the Oxidation of Isoniazid by the Catalase Peroxidase from Mycobacterium tuberculosis. *J. Am. Chem. Soc.* **116**, 7425–7426 (1994).

34. Winder, F. G., Collins, P. & Rooney, S. A. Effects of isoniazid on mycolic acid synthesis in Mycobacterium tuberculosis and on its cell envelope. *Biochem. J.* **117**, 1959 (1970).

35. Barry, C. E., Slayden, R. A. & Mdluli, K. Mechanisms of isoniazid resistance in Mycobacterium tuberculosis. *Drug Resist. Updat.* **1**, 128–134 (1998).

36. Silve, G. *et al.* Ethambutol inhibition of glucose metabolism in mycobacteria: A possible target of the drug. *Antimicrob. Agents Chemother.* **37**, 1536–1538 (1993).

37. Takayama, K. & Kilburn, J. O. Inhibition of synthesis of arabinogalactan by ethambutol in Mycobacterium smegmatis. *Antimicrob. Agents Chemother.* **33**, 1493–1499 (1989).

38. Sreevatsan, S. *et al.* Ethambutol resistance in Mycobacterium tuberculosis: Critical role of embB mutations. *Antimicrob. Agents Chemother.* **41**, 1677–1681 (1997).

39. Raynard, C. *et al.* Mechanisms of pyrazinamide resistance in mycobactena:

Importance of lack of uptake in addition to lack of pyrazinamidase activity. *Int. J. Lepr. Other Mycobact. Dis.* **67**, 516–517 (1999).

40. Ventola, C. L. The Antibiotic Resistance Crisis. *P T a peer-reviewed J. Formul. Manag.* **40**, 277–283 (2015).

41. Charles-Edouard Luyt*, Nicolas Bréchot, J.-L. T. and J. C. Antibiotic stewardship in the intensive care unit. *Antibiot. Pharmacokinet. Considerations Crit. Ill* 265–275 (2014). doi:10.1007/978-981-10-5336-8_13

42. The antibiotic alarm. *Nature* 141 (2013).

43. Klein, E. Y. *et al.* Global increase and geographic convergence in antibiotic consumption between 2000 and 2015. *Proc. Natl. Acad. Sci. U. S. A.* **115**, E3463–E3470 (2018).

44. Mainous, A. G., Everett, C. J., Post, R. E., Diaz, V. A. & Hueston, W. J. Availability of antibiotics for purchase without a prescription on the internet. *Ann. Fam. Med.* **7**, 431–435 (2009).

45. Khachatourians, G. G. Agricultural use of antibiotics and the evolution and transfer of antibiotic-resistant bacteria. *Cmaj* **159**, 1129–1136 (1998).

46. Manyi-Loh, C., Mamphweli, S., Meyer, E. & Okoh, A. *Antibiotic use in agriculture and its consequential resistance in environmental sources: Potential public health implications*. *Molecules* **23**, (2018).

47. Food and Drug Administration. Summary Report on Antimicrobials Sold or Distributed for Use in Food Producing Animals. *Food Drug Adm.* **9**, 76–99 (2010).

48. Van Der Gronde, T., Uyl-De Groot, C. A. & Pieters, T. *Addressing the challenge of high-priced prescription drugs in the era of precision medicine: A systematic review of drug life cycles, therapeutic drug markets and regulatory frameworks*. *PLoS ONE* **12**, (2017).

49. Talbot, G. H. *et al.* Bad bugs need drugs: An update on the development pipeline from the Antimicrobial Availability Task Force of the Infectious Diseases Society of America. *Chemother. J.* **15**, 97–105 (2006).

50. Premanandh, J., Samara, B. S. & Mazen, A. N. Race Against Antimicrobial Resistance Requires Coordinated Action – An Overview. *Front. Microbiol.* **6**, 1–6 (2016).

51. Wright, G. D. Opportunities for natural products in 21st century antibiotic discovery. *Nat. Prod. Rep.* **34**, 694–701 (2017).

52. Brown, E. D. & Wright, G. D. Antibacterial drug discovery in the resistance era. *Nature* **529**, 336–343 (2016).

53. Emmerson, A. M. & Jones, A. M. The quinolones: Decades of development and use. *J. Antimicrob. Chemother.* **51**, 13–20 (2003).

54. Papp-Wallace, K. M., Endimiani, A., Taracila, M. A. & Bonomo, R. A. Carbapenems: Past, present, and future. *Antimicrob. Agents Chemother.* **55**, 4943–4960 (2011).

55. Karas, J. A. *et al.* Structure-Activity Relationships of Daptomycin Lipopeptides. *J. Med. Chem.* **63**, 13266–13290 (2020).

56. Debono, M. *et al.* A21978C, a complex of new acidic peptide antibiotics: Isolation, chemistry, and mass spectral structure elucidation. *J. Antibiot. (Tokyo).* **40**, 761–777 (1987).

57. Silver, L. L. Challenges of antibacterial discovery. *Clin. Microbiol. Rev.* **24**, 71–109 (2011).

58. Shakil, S., Danish Rizvi, S. M. & Greig, N. H. High throughput virtual screening and molecular dynamics simulation for identifying a putative inhibitor of bacterial CTX-M-15. *Antibiotics* **10**, (2021).

59. Stokes, J. M. *et al.* A Deep Learning Approach to Antibiotic Discovery. *Cell* **180**, 688-702.e13 (2020).

60. Morris, G. M. *et al.* AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J. Comput. Chem.* **30**, 2785–2791 (2009).

61. Rizvi, S. M., Shazi, S. & Mohd., H. A simple click by click protocol to perform docking : *EXCLI J.* **12**, 831–857 (2013).

62. Somu, R. V. *et al.* Rationally-designed nucleoside antibiotics that inhibit siderophore biosynthesis of Mycobacterium tuberculosis. *J. Med. Chem.* **49**, 31–34 (2006).

63. Timofeeva, O. A. *et al.* Rationally Designed Inhibitors Identify STAT3 N-Domain as a Promising Anticancer Drug Target. *ACS Chem. Biol.* **2**, 799–809 (2007).

64. Mandal, S., Moudgil, M. & Mandal, S. K. Rational drug design. *Eur. J. Pharmacol.* **625**, 90–100 (2009).

65. Rehberg, N. *et al.* Chlorflavonin Targets Acetohydroxyacid Synthase Catalytic Subunit IlvB1 for Synergistic Killing of Mycobacterium tuberculosis. *ACS Infect. Dis.* **4**, 123–134 (2018).

66. Khodursky, A. B. & Cozzarelli, N. R. The mechanism of inhibition of topoisomerase IV by quinolone antibacterials. *J. Biol. Chem.* **273**, 27668–27677 (1998).

67. Pradhan, S. & Sinha, C. Sulfonamide derivatives as Mycobacterium tuberculosis inhibitors: in silico approach. *Silico Pharmacol.* **6**, 1–17 (2018).

68. Hancock, R. E. W. Aminoglycoside uptake and mode of action-with special reference to streptomycin and gentamicin: II. Effects of aminoglycosides on cells. *J. Antimicrob. Chemother.* **8**, 429–445 (1981).

69. Jarlier, V. & Nikaido, H. Mycobacterial cell wall: Structure and role in natural resistance to antibiotics. *FEMS Microbiol. Lett.* **123**, 11–18 (1994).

70. Kalscheuer, R. *et al.* The Mycobacterium tuberculosis capsule: A cell structure with key implications in pathogenesis. *Biochem. J.* **476**, 1995–2016 (2019).

71. Ortalo-Magne, A. *et al.* Molecular composition of the outermost capsular material of the tubercle bacillus. *Microbiology* **141**, 1609–1620 (1995).

72. Ophir, T. & Gutnick, D. L. A role for exopolysaccharides in the protection of microorganisms from desiccation. *Appl. Environ. Microbiol.* **60**, 740–745 (1994).

73. Yoshida, K. *et al.* Role of bacterial capsule in local and systemic inflammatory responses of mice during pulmonary infection with Klebsiella pneumoniae. *J. Med. Microbiol.* **49**, 1003–1010 (2000).

74. Lemassu, A. & Daffe, M. Structural features of the exocellular polysaccharides of Mycobacterium tuberculosis. *Biochem. J.* **297**, 351–357 (1994).

75. Kalscheuer, R. *et al.* Self-poisoning of Mycobacterium tuberculosis by targeting GlgE in an α-glucan pathway. *Nat. Chem. Biol.* **6**, 376–384 (2010).

76. Stadthagen, G. *et al.* Genetic basis for the biosynthesis of methylglucose

lipopolysaccharides in Mycobacterium tuberculosis. *J. Biol. Chem.* **282**, 27270–27276 (2007).

77. Syson, K., Batey, S. F. D., Schindler, S., Kalscheuer, R. & Bornemann, S. A temperature-sensitive Mycobacterium smegmatis glgE mutation leads to a loss of GlgE enzyme activity and thermostability and the accumulation of α-maltose-1-phosphate. *Biochim. Biophys. Acta - Gen. Subj.* **1865**, 129783 (2021).

78. McNeil, M., Daffe, M. & Brennan, P. J. Location of the mycolyl ester substituents in the cell walls of mycobacteria. *J. Biol. Chem.* **266**, 13217–13223 (1991).

79. Takayama, K., Wang, C. & Besra, G. S. Pathway to synthesis and processing of mycolic acids in Mycobacterium tuberculosis. *Clin. Microbiol. Rev.* **18**, 81–101 (2005).

80. Butler, W. R., Kilburn, J. O. & Kubica, G. P. High-performance liquid chromatography analysis of mycolic acids as an aid in laboratory identification of Rhodococcus and Nocardia species. *J. Clin. Microbiol.* **25**, 2126–2131 (1987).

81. Qureshi, N., Takayama, K., Jordi, H. C. & Schnoes, H. K. Characterization of the purified components of a new homologous series of α-mycolic acids from Mycobacterium tuberculosis H37Ra. *J. Biol. Chem.* **253**, 5411–5417 (1978).

82. Liu, J., Barry, C. E., Besra, G. S. & Nikaido, H. Mycolic acid structure determines the fluidity of the mycobacterial cell wall. *J. Biol. Chem.* **271**, 29545–29551 (1996).

83. Ribi, E. *et al.* Induction of resistance to tuberculosis in mice with defined components of mycobacteria and with some unrelated materials. *Zentralblatt fur Bakteriol. Mikrobiol. und Hyg. - Abt. 1 Orig. A* **251**, 345–356 (1982).

84. Barry, C. E. *et al.* Mycolic acids: Structure, biosynthesis and physiological functions. *Prog. Lipid Res.* **37**, 143–179 (1998).

85. Lea-Smith, D. J. *et al.* The reductase that catalyzes mycolic motif synthesis is required for efficient attachment of mycolic acids to arabinogalactan. *J. Biol. Chem.* **282**, 11000–11008 (2007).

86. Bhatt, A., Brown, A. K., Singh, A., Minnikin, D. E. & Besra, G. S. Loss of a Mycobacterial Gene Encoding a Reductase Leads to an Altered Cell Wall Containing β-oxo- Mycolic Acid Analogs and Accumulation of Ketones. *Chem. Biol.* **15**, 930–939 (2008).

87. McNeil, M., Wallner, S. J., Hunter, S. W. & Brennan, P. J. Demonstration that the galactosyl and arabinosyl residues in the cell-wall arabinogalactan of Mycobacterium leprae and Myobacterium tuberculosis are furanoid. *Carbohydr. Res.* **166**, 299–308 (1987).

88. Amar, C. & Vilkas, E. [Isolation of arabinose phosphate from the walls of Mycobacterium tuberculosis H 37 Ra]. *C. R. Acad. Sci. Hebd. Seances Acad. Sci. D.* **277**, 1949–51 (1973).

89. McNeil, M., Daffe, M. & Brennan, P. J. Evidence for the nature of the link between the arabinogalactan and peptidoglycan of mycobacterial cell walls. *J. Biol. Chem.* **265**, 18200–18206 (1990).

90. Besra, G. S. *et al.* A New Interpretation of the Structure of the Mycolyl—Arabinogalactan Complex of Mycobacterium Tuberculosis as Revealed Through Characterization of Oligoglycosylalditol Fragments by Fast-Atom Bombardment Mass Spectrometry and 1H Nuclear Magnetic Resonance . *Biochemistry* **34**, 4257–4266 (1995).

91. Vollmer, W., Blanot, D. & De Pedro, M. A. Peptidoglycan structure and architecture. *FEMS Microbiol. Rev.* **32**, 149–167 (2008).

92. Azuma, I. *et al.* Occurence of N-glucosylmuramic acid in bacterial cell walls. *Biochim Biophys Acta* **20**, 444–451 (1970).

93. Mahapatra, S. *et al.* Mycobacterial lipid II is composed of a complex mixture of modified muramyl and peptide moieties linked to decaprenyl phosphate. *J. Bacteriol.* **187**, 2747–2757 (2005).

94. Raymond, J. B., Mahapatra, S., Crick, D. C. & Pavelka, M. S. Identification of the namH gene, encoding the hydroxylase responsible for the N-glycolylation of the mycobacterial peptidoglycan. *J. Biol. Chem.* **280**, 326–333 (2005).

95. Shiva Kumar Angala, Juan Manuel Belardinelli, Emilie Huc-Claustre, William H. Wheat, and M. J. *The cell envelope glycoconjugates of Mycobacterium tuberculosis*. *Critical Reviews in Biochemistry and Molecular Biology* **49**, (2014).

96. Mamadou, D., Marie-Ange, D. & Nicole, G. The cell envelope of Mycobacterium smegmatis: cytochemistry and architectural implications. *FEMS Microbiol. Lett.* **61**, 89–94 (1989).

97. Bansal-Mutalik, R. & Nikaido, H. Mycobacterial outer membrane is a lipid bilayer and the inner membrane is unusually rich in diacyl phosphatidylinositol dimannosides. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 4958–4963 (2014).

98. Tallieux, L. *et al.* DC-SIGN is the major Mycobacterium tuberculosis receptor on human dendritic cells. *J. Exp. Med.* **197**, 121–127 (2003).

99. Schlesinger, L. S. Macrophage phagocytosis of virulent but not attenuated strains of Mycobacterium tuberculosis is mediated by mannose receptors in addition to complement receptors. *J. Immunol.* **150**, 2920–30 (1993).

100. Villeneuve, C. *et al.* Mycobacteria use their surface-exposed glycolipids to infect human macrophages through a receptor-dependent process. *J. Lipid Res.* **46**, 475–483 (2005).

101. Schlesinger, L. S., Hull, S. R. & Kaufman, T. M. Binding of the terminal mannosyl units of lipoarabinomannan from a virulent strain of Mycobacterium tuberculosis to human macrophages. *J. Immunol.* **152**, 4070–9 (1994).

102. Nigou, J., Zelle-Rieser, C., Gilleron, M., Thurnher, M. & Puzo, G. Mannosylated Lipoarabinomannans Inhibit IL-12 Production by Human Dendritic Cells: Evidence for a Negative Signal Delivered Through the Mannose Receptor. *J. Immunol.* **166**, 7477–7485 (2001).

103. Astarie-Dequeker, C. *et al.* The mannose receptor mediates uptake of pathogenic and nonpathogenic mycobacteria and bypasses bactericidal responses in human macrophages. *Infect. Immun.* **67**, 469–477 (1999).

104. Kang, P. B. *et al.* The human macrophage mannose receptor directs Mycobacterium tuberculosis lipoarabinomannan-mediated phagosome biogenesis. *J. Exp. Med.* **202**, 987–999 (2005).

105. Brennan, P. & Ballou, C. E. Biosynthesis of mannophosphoinositides by Mycobacterium phlei. The family of dimannophosphoinositides. *J. Biol. Chem.* **242**, 3046–3056 (1967).

106. Chatterjee, D., Hunter, S. W., McNeil, M. & Brennan, P. J. Lipoarabinomannan. Multiglycosylated form of the mycobacterial mannosylphosphatidylinositols. *J. Biol.*

Chem. **267**, 6228–6233 (1992).

107. Torrelles, J. B., Azad, A. K. & Schlesinger, L. S. Fine Discrimination in the Recognition of Individual Species of Phosphatidyl- myo -Inositol Mannosides from Mycobacterium tuberculosis by C-Type Lectin Pattern Recognition Receptors. *J. Immunol.* **177**, 1805–1816 (2006).

108. Morita, Y. S. *et al.* PimE is a polyprenol-phosphate-mannose-dependent mannosyltransferase that transfers the fifth mannose of phosphatidylinositol mannoside in mycobacteria. *J. Biol. Chem.* **281**, 25143–25155 (2006).

109. Besra, G. S., Morehouse, C. B., Rittner, C. M., Waechter, C. J. & Brennan, P. J. Biosynthesis of mycobacterial lipoarabinomannan. *J. Biol. Chem.* **272**, 18460–18466 (1997).

110. Wolucka, B. A., McNeil, M. R., De Hoffmann, E., Chojnacki, T. & Brennan, P. J. Recognition of the lipid intermediate for arabinogalactan/arabinomannan biosynthesis and its relation to the mode of action of ethambutol on mycobacteria. *J. Biol. Chem.* **269**, 23328–23335 (1994).

111. Brecik, M. *et al.* DprE1 Is a Vulnerable Tuberculosis Drug Target Due to Its Cell Wall Localization. *ACS Chem. Biol.* **10**, 1631–1636 (2015).

112. Tipper, D. J. and J. L. S. Mechanism of action of penicillins: a proposal based on their structural similarity to acyl-D-alanyl-D-alanine. *Proc Natl Acad Sci U S A* **54**, 1133–41 (1965).

113. Christophe, T. *et al.* High content screening identifies decaprenyl-phosphoribose 2′ epimerase as a target for intracellular antimycobacterial inhibitors. *PLoS Pathog.* **5**, (2009).

114. Gurcha, S. S. *et al.* Ppm1, a novel polyprenol monophosphomannose synthase from Mycobacterium tuberculosis. *Biochem. J.* **365**, 441–450 (2002).

115. Agrawal, P., Gupta, P., Swaminathan, K. & Parkesh, R. α-Glucan Pathway as a Novel Mtb Drug Target: Structural Insights and Cues for Polypharmcological Targeting of GlgB and GlgE. *Curr. Med. Chem.* **21**, 4074–4084 (2014).

116. Tran, A. T. *et al.* Sansanmycin natural product analogues as potent and selective anti-mycobacterials that inhibit lipid i biosynthesis. *Nat. Commun.* **8**, 1–9 (2017).

117. Korduláková, J. *et al.* Definition of the first mannosylation step in phosphatidylinositol mannoside synthesis. PimA is essential for growth of mycobacteria. *J. Biol. Chem.* **277**, 31335–31344 (2002).

118. Mishra, A. K. *et al.* Identification of an α(1→6) mannopyranosyltransferase (MptA), involved in Corynebacterium glutamicum lipomanann biosynthesis, and identification of its orthologue in Mycobacterium tuberculosis. *Molecular Microbiology* **65**, 1503–1517 (2007).

119. Zhang, N. *et al.* The Emb proteins of mycobacteria direct arabinosylation of lipoarabinomannan and arabinogalactan via an N-terminal recognition region and a C-terminal synthetic region. *Mol. Microbiol.* **50**, 69–76 (2003).

120. Miyakoshi, M., Matera, G., Maki, K., Sone, Y. & Vogel, J. Functional expansion of a TCA cycle operon mRNA by a 3 end-derived small RNA. *Nucleic Acids Res.* **47**, 2075–2088 (2019).

121. Papenfort, K., Sun, Y., Miyakoshi, M., Vanderpool, C. K. & Vogel, J. Small RNA-mediated activation of sugar phosphatase mRNA regulates glucose homeostasis. *Cell*

153, 426–437 (2013).

122. Trotochaud, A. E. & Wassarman, K. M. A highly conserved 6S RNA structure is required for regulation of transcription. *Nat. Struct. Mol. Biol.* **12**, 313–319 (2005).

123. Bessaiah, H. *et al. The RyfA small RNA regulates oxidative and osmotic stress responses and virulence in uropathogenic Escherichia coli*. *PLoS Pathogens* **17**, (2021).

124. Schachterle, J. K., Onsay, D. M. & Sundin, G. W. Small RNA ArcZ Regulates Oxidative Stress Response Genes and Regulons in Erwinia amylovora. *Front. Microbiol.* **10**, 1–17 (2019).

125. Lease, R. A., Smith, D., McDonough, K. & Belfort, M. The small noncoding DsrA RNA is an acid resistance regulator in Escherichia coli. *J. Bacteriol.* **186**, 6179–6185 (2004).

126. Arnvig, K. B. *et al.* Sequence-based analysis uncovers an abundance of non-coding RNA in the total transcriptome of Mycobacterium tuberculosis. *PLoS Pathog.* **7**, (2011).

127. Gerrick, E. R. *et al.* Small RNA profiling in mycobacterium tuberculosis identifies mrsi as necessary for an anticipatory iron sparing response. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 6464–6469 (2018).

128. Grieshaber, N. A., Grieshaber, S. S., Fischer, E. R. & Hackstadt, T. A small RNA inhibits translation of the histone-like protein Hc1 in Chlamydia trachomatis. *Molecular Microbiology* **59**, 541–550 (2006).

129. Toledo-Arana, A. *et al.* The Listeria transcriptional landscape from saprophytism to virulence. *Nature* **459**, 950–956 (2009).

130. Chabelskaya, S., Gaillot, O. & Felden, B. A Staphylococcus aureus small RNA is required for bacterial virulence and regulates the expression of an immune-evasion molecule. *PLoS Pathog.* **6**, 1–11 (2010).

131. Padalon-Brauch, G. *et al.* Small RNAs encoded within genetic islands of Salmonella typhimurium show host-induced expression and role in virulence. *Nucleic Acids Res.* **36**, 1913–1927 (2008).

132. Song, T. *et al.* A new Vibrio cholerae sRNA modulates colonization and affects release of outer membrane vesicles. *Mol. Microbiol.* **70**, 100–111 (2008).

133. Korte, J. *et al.* Trehalose-6-Phosphate-Mediated Toxicity Determines Essentiality of OtsB2 in Mycobacterium tuberculosis In Vitro and in Mice. *PLoS Pathog.* **12**, e1006043 (2016).

134. Schnappinger, D. *et al.* Transcriptional Adaptation of *Mycobacterium tuberculosis* within Macrophages. *J. Exp. Med.* **198**, 693–704 (2003).

135. Altuvia, S., Kornitzer, D., Teff, D. & Oppenheim, A. B. Alternative mRNA structures of the cIII gene of bacteriophage λ determine the rate of its translation initiation. *Journal of Molecular Biology* **210**, 265–280 (1989).

136. Loh, E. *et al.* A trans-Acting Riboswitch Controls Expression of the Virulence Regulator PrfA in Listeria monocytogenes. *Cell* **139**, 770–779 (2009).

137. Franze De Fernandez, M. T., Eoyang, L. & August, J. T. Factor fraction required for the synthesis of bacteriophage Qβ-RNA. *Nature* **219**, 588–590 (1968).

138. Arluison, V. *et al.* Spectroscopic observation of RNA chaperone activities of Hfq in post-transcriptional regulation by a small non-coding RNA. *Nucleic Acids Res.* **35**,

999–1006 (2007).

139. Sittka, A., Pfeiffer, V., Tedin, K. & Vogel, J. The RNA chaperone Hfq is essential for the virulence of Salmonella typhimurium. *Mol. Microbiol.* **63**, 193–217 (2007).

140. Sittka, A. *et al.* Deep sequencing analysis of small noncoding RNA and mRNA targets of the global post-transcriptional regulator, Hfq. *PLoS Genet.* **4**, (2008).

141. Moller, T. *et al.* Hfq: A Bacterial Sm-like protein that mediates RNA-RNA interaction. *Mol. Cell* **9**, 23–30 (2002).

142. Kajitani, M., Kato, A., Wada, A., Inokuchi, Y. & Ishihama, A. Regulation of the Escherichia coli hfq gene encoding the host factor for phage Q(β). *J. Bacteriol.* **176**, 531–534 (1994).

143. Geissmann, T. *et al.* A search for small noncoding RNAs in Staphylococcus aureus reveals a conserved sequence motif for regulation. *Nucleic Acids Res.* **37**, 7239–7257 (2009).

144. Cole, S. T. *et al.* Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence. *Nature* **393**, 537–544 (1998).

145. DiChiara, J. M. *et al.* Multiple small RNAs identified in Mycobacterium bovis BCG are also expressed in Mycobacterium tuberculosis and Mycobacterium smegmatis. *Nucleic Acids Res.* **38**, 4067–4078 (2010).

146. Arnvig, K. B. & Young, D. B. Identification of small RNAs in Mycobacterium tuberculosis. *Mol. Microbiol.* **73**, 397–408 (2009).

147. Papenfort, K. *et al.* σE-dependent small RNAs of Salmonella respond to membrane stress by accelerating global omp mRNA decay. *Molecular Microbiology* **62**, 1674–1688 (2006).

148. Faucher, S. P., Friedlander, G., Livny, J., Margalit, H. & Shuman, H. A. Legionella pneumophila 6S RNA optimizes intracellular multiplication. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 7533–7538 (2010).

149. Hammer, B. K. & Bassler, B. L. Regulatory small RNAs circumvent the conventional quorum sensing pathway in pandemic Vibrio cholerae. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 11145–11149 (2007).

150. Bychenko, O. *et al.* Mycobacterium tuberculosis Small RNA MTS1338 Confers Pathogenic Properties to Non-Pathogenic Mycobacterium smegmatis. 1–18 (2021).

151. Liu, M. Y. *et al.* The RNA molecule CsrB binds to the global regulatory protein CsrA and antagonizes its activity in Escherichia coli. *J. Biol. Chem.* **272**, 17502–17510 (1997).

152. Sabnis, N. A., Yang, H. & Romeo, T. Pleiotropic regulation of central carbohydrate metabolism in Escherichia coli via the gene csrA. *J. Biol. Chem.* **270**, 29096–29104 (1995).

153. Fleming, A. On the antibacterial action of cultures of a penicillium, with special reference to their use in the isolation of B.influenzae. *Br. J. Exp. Pathol.* **10**, 226–236 (1929).

154. Schaeffer, M. L. *et al.* The pimB gene of Mycobacterium tuberculosis encodes a mannosyltransferase involved in lipoarabinomannan biosynthesis. *J. Biol. Chem.* **274**, 31625–31631 (1999).

155. Belanger, A. E. & Hatfull, G. F. Exponential-phase glycogen recycling is essential for

growth of Mycobacterium smegmatis. *J. Bacteriol.* **181**, 6670–6678 (1999).

156. Tacconelli, E. *et al.* Discovery, research, and development of new antibiotics: the WHO priority list of antibiotic-resistant bacteria and tuberculosis. *Lancet Infect. Dis.* **18**, 318–327 (2018).

157. CDC. Antibiotic Resistance Threats in the United States, 2013. *Cent. Dis. Control Prev.* **37**, 828–831 (2013).

158. World Health Organization. *WHO consolidated guidelines on tuberculosis. Module 4: treatment, drug-resistant tuberculosis treatment. Who* (2020).

159. Singh, S. B., Young, K. & Silver, L. L. What is an “ideal” antibiotic? Discovery challenges and path forward. *Biochem. Pharmacol.* **133**, 63–73 (2017).

160. Kallberg, C. *et al.* Introduction and geographic availability of new antibiotics approved between 1999 and 2014. *PLoS One* **13**, 1–19 (2018).

161. Poole, K. Mechanisms of bacterial biocide and antibiotic resistance. *J. Appl. Microbiol. Symp. Suppl.* **92**, 55–64 (2002).

162. Li, J. *et al.* LncTar: A tool for predicting the RNA targets of long noncoding RNAs. *Brief. Bioinform.* **16**, 806–812 (2014).

163. Lindow, M. & Kauppinen, S. Discovering the first microrna-targeted drug. *J. Cell Biol.* **199**, 407–412 (2012).

164. van der Ree, M. H. *et al.* Miravirsen dosing in chronic hepatitis C patients results in decreased microRNA-122 levels without affecting other microRNAs in plasma. *Aliment. Pharmacol. Ther.* **43**, 102–113 (2016).

165. Jopling, C. L., Yi, M. K., Lancaster, A. M., Lemon, S. M. & Sarnow, P. Modulation of hepatitis C virus RNA abundance by a liver-specific MicroRNA. *Science (80-.).* **309**, 1577–1581 (2005).

166. Shimakami, T. *et al.* Base Pairing between Hepatitis C Virus RNA and MicroRNA 122 3′ of Its Seed Sequence Is Essential for Genome Stabilization and Production of Infectious Virus. *J. Virol.* **86**, 7372–7383 (2012).

167. Burnett, J. C., Rossi, J. J. & Tiemann, K. Current progress of siRNA/shRNA therapeutics in clinical trials. *Biotechnol. J.* **6**, 1130–1146 (2011).

168. Prados-Rosales, R. *et al.* Mycobacteria release active membrane vesicles that modulate immune responses in a TLR2-dependent manner in mice. *J. Clin. Invest.* **121**, 1471–1483 (2011).

169. Yanagihara, K. *et al.* Effects of short interfering RNA against methicillin-resistant Staphylococcus aureus coagulase in vitro and in vivo. *J. Antimicrob. Chemother.* **57**, 122–126 (2006).

170. Salmond, G. P. C. & Fineran, P. C. A century of the phage: Past, present and future. *Nat. Rev. Microbiol.* **13**, 777–786 (2015).

171. Wright, A., Hawkins, C. H., Änggård, E. E. & Harper, D. R. A controlled clinical trial of a therapeutic bacteriophage preparation in chronic otitis due to antibiotic-resistant Pseudomonas aeruginosa; A preliminary report of efficacy. *Clin. Otolaryngol.* **34**, 349–357 (2009).

172. Forti, F. *et al.* Design of a broad-range bacteriophage cocktail that reduces pseudomonas aeruginosa biofilms and treats acute infections in two animal models. *Antimicrob. Agents Chemother.* **62**, 1–13 (2018).

173. Chan, B. K. *et al.* Phage selection restores antibiotic sensitivity in MDR Pseudomonas aeruginosa. *Sci. Rep.* **6**, (2016).

174. Schooley, R. T. *et al.* Development and Use of Personalized Bacteriophage-Based Therapeutic Cocktails To Treat a Patient with a Disseminated Resistant Acinetobacter baumannii Infection. *Antimicrob. Agents Chemother.* **61**, 1–14 (2017).

175. Oechslin, F. Resistance development to bacteriophages occurring during bacteriophage therapy. *Viruses* **10**, (2018).

176. Burrowes, B., Harper, D. R., Anderson, J., McConville, M. & Enright, M. C. Bacteriophage therapy: Potential uses in the control of antibiotic-resistant pathogens. *Expert Rev. Anti. Infect. Ther.* **9**, 775–785 (2011).

177. Olchanski, N. *et al.* Palivizumab prophylaxis for respiratory syncytial virus: Examining the evidence around value. *Open Forum Infect. Dis.* **5**, 1–9 (2018).

178. Tursi, S. A. *et al.* Salmonella Typhimurium biofilm disruption by a human antibody that binds a pan-amyloid epitope on curli. *Nat. Commun.* **11**, 1–13 (2020).

179. Storek, K. M. *et al.* Monoclonal antibody targeting the β-barrel assembly machine of Escherichia coli is bactericidal. *Proceedings of the National Academy of Sciences of the United States of America* **115**, 3692–3697 (2018).

180. Riechmann, L., Clark, M., Waldmann, H. & Winter, G. Reshaping human antibodies for therapy. *Nature* **332**, 323–327 (1988).

181. Vanhove, B. *et al.* XAV-19, a Swine Glyco-Humanized Polyclonal Antibody Against SARS-CoV-2 Spike Receptor-Binding Domain, Targets Multiple Epitopes and Broadly Neutralizes Variants. *Front. Immunol.* **12**, 1–11 (2021).

182. Weisman, L. E. *et al.* A randomized study of a monoclonal antibody (pagibaximab) to prevent staphylococcal sepsis. *Pediatrics* **128**, 271–279 (2011).

183. Lowy, I. *et al.* Treatment with Monoclonal Antibodies against Clostridium difficile Toxins. *N. Engl. J. Med.* **362**, 197–205 (2010).

184. Bruno, C. J. & Jacobson, J. M. Ibalizumab: An anti-CD4 monoclonal antibody for the treatment of HIV-1 infection. *J. Antimicrob. Chemother.* **65**, 1839–1841 (2010).

185. Ling, L. L. *et al.* A new antibiotic kills pathogens without detectable resistance. *Nature* **517**, 455–459 (2015).

186. Antimicrobial Peptide Database. *https://aps.unmc.edu/* (2022).

187. Brogden, K. A. Antimicrobial peptides: Pore formers or metabolic inhibitors in bacteria? *Nat. Rev. Microbiol.* **3**, 238–250 (2005).

188. Aerts, A. M., François, I. E. J. A., Cammue, B. P. A. & Thevissen, K. The mode of antifungal action of plant, insect and human defensins. *Cell. Mol. Life Sci.* **65**, 2069–2079 (2008).

189. Klotman, M. E. & Chang, T. L. Defensins in innate antiviral immunity. *Nat. Rev. Immunol.* **6**, 447–456 (2006).

190. Wakabayashi, H., Hiratani, T., Uchida, K. & Yamaguchi, H. Antifungal spectrum and fungicidal mechanism of an N-terminal peptide of bovine lactoferrin. *J. Infect. Chemother.* **1**, 185–189 (1996).

191. Oren, Z. & Shai, Y. Mode of action of linear amphipathic α-helical antimicrobial peptides. *Biopolymers* **47**, 451–463 (1998).

192. Hancock, R. E. W. & Sahl, H. G. Antimicrobial and host-defense peptides as new anti-infective therapeutic strategies. *Nat. Biotechnol.* **24**, 1551–1557 (2006).

193. Som, A., Vemparala, S., Ivanov, I. & Tew, G. N. Synthetic mimics of antimicrobial peptides. *Biopolym. - Pept. Sci. Sect.* **90**, 83–93 (2008).

194. Torres-Juarez, F. *et al.* RNase 7 but not psoriasin nor sPLA2-IIA associates with Mycobacterium tuberculosis during airway epithelial cell infection. *Pathog. Dis.* **76**, 1–8 (2018).

Eidesstattliche Erklärung

Hiermit versichere ich eidesstattlich, dass die vorliegende Dissertation eigenständig und ohne unzulässige Hilfe Fremder unter Beachtung der „Grundsätze zur Sicherung guter wissenschaftlicher Praxis an der Heinrich-Heine-Universität Düsseldorf“ verfasst wurde.

Hamm, _____

Steffen Schindler