Riboswitch regulation of methionine metabolism and vitamin B_{12} uptake in mycobacteria – implications for drug susceptibility and pathogenesis

Terry Kipkorir

Molecular Mycobacteriology Research Unit
Division of Medical Microbiology
Department of Pathology

A thesis submitted to the Faculty of Health Sciences, University of Cape Town, in fulfilment of the requirements for the degree of Doctor of Philosophy

January 2019

Dedication

For my mother Leah C. Ngeno, and my father, John K. Ngeno, who bequeathed me the priceless gifts of curiosity, faith and hope.

Declaration

I declare that this thesis is my own unaided work, it is being submitted for the degree of Doctor of Philosophy at the University of Cape Town. It has not been submitted for any degree of examination at any other university.

Signed by candidate

Terry Kipkorir

07-Jan-2019

Date

Abstract

Alterations in the genetic capacity for cobamide biosynthesis have been identified as potentially critical in the evolution of *Mycobacterium tuberculosis* from a putative environmental ancestor. Moreover, recent studies have implicated cobamide biosynthesis pathway genes in the adaptation of the bacillus to intracellular pathogenesis. Although mycobacteria retain essential biochemical reactions that require cobamides, the specific role of these co-factors during tuberculosis (TB) disease remains unresolved. This thesis aimed to examine the production, uptake, and utilization of cobamides in mycobacteria using *M. smegmatis* as a model. To this end, the genetic capacity for *de novo* production and uptake of cobamide in host-associated and environmental mycobacteria was assessed, followed by direct validation in *M. smegmatis*. A combination of genetics, gene expression analysis, live-cell time-lapse microscopy and targeted metabolite and protein analysis via mass spectrometry (MS) was then employed to investigate cobamide riboswitch-dependent regulation of methionine biosynthesis in *M. smegmatis*. Results indicated that, in wild-type *M. smegmatis*, *de novo* cobamide biosynthesis ensured constitutive repression of *metE*, the gene encoding the mycobacterial cobalamin-independent methionine synthase. Owing to this repression, *metH*, a gene encoding the cobalamin-dependent methionine synthase, was found to be conditionally essential for bacillary replication *in vitro*. Drug susceptibility testing to investigate the link between cobamides and the intrinsic resistance to anti-folate antibiotics confirmed novel mycobacterial vulnerabilities in cobamide-related methionine metabolism, indicating that the outcomes of cobamide-dependent regulation may have relevance to mycobacterial pathogenesis and drug discovery. In contrast to *M. tuberculosis*, which was previously shown to transport exogenous CNCbl readily, *M. smegmatis* poorly assimilated exogenous co-factor despite the presence of multiple putative cobamide transporters. However, uptake was enhanced in a mutant requiring CNCbl for growth. Elucidating the factors which regulate cobamide biosynthesis and co-factor utilization in *M. smegmatis*, an environmental mycobacterium, might provide a lens through which to consider the differential regulation and utilization of cobamides in *M. tuberculosis*, an obligate pathogen with a limited host range.

Acknowledgements

Above all, God, in whom I live, and move, and have my being. *Amen!*

I offer my gratitude to everyone who has been instrumental in making my degree successful. I am sure that without this support, it would not have been possible for me to walk see this endeavour to its end. Due to practical constraints, I am unable to list the names of all people that have supported me, nor enumerate all their specific contributions.

To Mom and Dad, I give the most thanks. You have been the principal sources of my inspiration and you never allowed me to think that I couldn't do it. *Obee kutung'ony!* To my siblings, Vivian and Tracey, your faith in me keeps me motivated. Your strength when we lost sister Dorothy was a much-needed lifebuoy. To Anna, my god-daughter, and Aislinn, her mom. You two are the right amount of cuteness one needs while pursuing a PhD degree. To Winnie, I love you! You became my unofficial supervisor. Thank you for making me feel at home with you.

To my primary supervisor, Professor Digby Warner, thank you for taking me on, believing in me, and for teaching me about mycobacteria. Your dedication and rigor are rubbing off on me, and I am glad. Thank you for your open-door policy, and for never missing our regular update meetings and brainstorming sessions.

To Professor Valerie Mizrahi, who co-supervised my doctoral work, thank you for always expecting excellence from me and for being accessible despite your heavily packed schedule. I would love to inherit your passion for the basic sciences.

To the MMRU vitamin B_{12} mini-group: Dr. Atica Moosa, who trained me in the early days of my training, when I was still wet behind the ears! To Dr. Gabriel Mashabela and Rendani Mbau, for your insights on cobamide biochemistry. Special appreciation to Gabriel, for pioneering the LC-MS/MS method for quantitatively analysing cobamide production in mycobacteria. To Dr. Gopinath Krishnamoorthy, for his valuable advice and insight into cobalamin metabolism and riboswitch dependent regulation.

To Irene Gobe, for your Tn-seq expertise; Timothy de Wet, for your helpful suggestions about CRISPRi, and for your help analysing data with "R" software; my best friend, Dr. Anastasia Koch, for your insights in mycobacterial phylogenomics, and for your critical reading of this thesis. To Vin, a special thank you for your counsel to never cut corners and to always have a backup plan, thank you, boss! Thank you, Mandy, Charles, Saber, Michael "Gonzaga!" Reiche, Tim and Irene for being respectful lab space neighbours. To Caitlin, Avuyonke, Lucas, Meagan, and Celeste, for cordially sharing the office space. To Kedebone, for the thoroughly entertaining chit-chats, which were daily reminders that there is more life beyond the walls of the university.

To the extended MMRU family, for honouring me with the royal moniker, *King*, which I shall forever carry with pride. Thank you for all the stimulating conversations we had in and out of the lab. Thank you for all the helpful troubleshooting tips and tricks.

To Dr. Jon Ambler, for your help unravelling complex whole-genome sequencing data. To Dr. Clemens Hermann and Dr. Bridget Calder and members of Professor Jonathan Blackburn laboratory at UCT, thank you for your technical advice on proteomics and MS. To Dr. Lubbe Wiesner for the access to the LC-MS/MS platform at UCT's Division of Clinical Pharmacology. To the Sarah Fortune laboratory for kindly providing the CRISPRi system, and to Dr. Stephanie Dawes for Δ*metE* and Δ*metE*Δ*cobK*::*hyg* strains.

To my IDM mentor, Professor Edward Sturrock, I appreciate your creating the time to simply chat.

To Nina, who received me on my very first day in Cape Town, and to Miriam and Stanley – you were my first friends in South Africa. To *the Family*: Joe and his wife, Beryl; Edgar and his wife Jade, and sons Gavin and Illie; Safari and his wife Dorcas; Philemon and his wife Beryl; Elzie, Karen, Joshua, Albert, Billy, Joan, Lydia, Robert, Sandra, Janet, and Aunt Chuma – thank you for the fun and wholesome fellowship. Your lives are excellent examples of ethical living. To *the Mowbray Seventh-day Adventist Church*, and *the Rondebosch cell group*, thank you for being family. To Lucy and Andy and Sharon, Shammah and Jacqueline, for your friendship.

This thesis was supported by grants from the Howard Hughes Medical Institute Senior International Research Scholar's grant (to Professor Mizrahi), the South African Medical Research Council (to Professor Mizrahi) and the National Research Foundation of South Africa (to Professor Mizrahi and Professor Warner).

Presentations

1. **Kipkorir, T.** Mashabela, G., Mizrahi, V., Warner, D.F. Conditional essentiality of the gene encoding vitamin B_{12}-dependent methionine synthase MetH in *Mycobacterium smegmatis.* Short talk given at the SASBMB-FASBMB Conference held at North West University, South Africa, 2018.

2. **Kipkorir, T.**, Mashabela, G., Mizrahi, V., Gopinath, K., Warner, D.F. Riboswitch regulation of methionine metabolism and vitamin B_{12} uptake in mycobacteria – a role in pathogenesis? Oral presentation made at the Cape Town Acid Fast Club held at Stellenbosch University, South Africa, 2017.

3. **Kipkorir, T.**, Mashabela, G., Mizrahi, V., Gopinath, K., Warner, D.F. Riboswitch regulation of methionine metabolism and vitamin B_{12} uptake in mycobacteria – a role in pathogenesis? Poster discussed at the 7th Congress of the European Microbiologists (FEMS) held at Valencia, Spain, 2017.

4. **Kipkorir, T.**, Mashabela, G., Mizrahi, V., Gopinath, K., Warner, D.F. Riboswitch regulation of methionine metabolism and vitamin B_{12} uptake in mycobacteria. Poster exhibited at the Tuberculosis: Past, Present, and Future ASM Conference held in New York, USA, 2017.

Table of Contents

List of Figures

List of Tables

Frequently used abbreviations

ACN	Acetonitrile
5-MTHF	5-methyltetrahydrofolate
ABC (buffer)	Ammonium bicarbonate buffer
AdoCba	Adenosylcobamide
AdoCbl	Adenosylcobalamin (also known as coenzyme B_{12})
ALA	5-aminolevulinic acid
Amp	Ampicilin
Amp^R	Ampicilin resistant
ATc	Anhydrotetracycline
BCG	Bacille Calmette-Guèrin
BLAST	Basic local alignment search tool
bp	base pair(s)
BSL2	Biosafety level 2
BSL3	Biosafety level 3
Cbi-P	Cobinamide-phosphate
cDNA	Complementary DNA
CNCba	Cyanocobamide
CNCbl	Cyanocobalamin (also known as vitamin B_{12})
$(CN)_2$Cbi	Dicyanobinamide
cobK	gene encoding precorrin-6x reductase
CobK	Precorrin-6x reductase
CRISPRi	Clustered regularly interspaced short palindromic repeats interference
CRISPRi-Seq	CRISPRi sequencing
CTAB	Cetyltrimethylammonium bromide
*d*Cas9$_{Sth1}$	Deactivated CRISPR-associated protein-9 nuclease from *Streptococcus thermophilus*
DCO	double crossover
ddPCR	droplet digital PCR
DMB	Dimethylbenzimidazole
DNA	Deoxyribonucleic acid
DOTS	Directly-observed therapy, short course
DTT	Dithiothreitol
dUTP	Deoxyuridine triphosphate
EC	Enzyme commission
e.g.	*exempli gratia* (for example)
EMB	Ethambutol
ESX	Early secretory antigenic target secretion system
et al.	*et alibi* (and others)
FA	Formic acid
FAM	6-carboxyfluorescein
FASP	Filter-aided sample preparation

FMN	Flavin mononucleotide
$FMNH_2$	reduced FMN
g	Gravitational force
h	Hour(s)
HGT	horizontal gene transfer
HIV	Human immunodeficiency virus
HPLC	High performance liquid chromatography
Hyg	Hygromycin
Hyg^R	Hygromycin resistant
IDM	Institute of Infectious Disease and Molecular Medicine
i.e.	*id est* (that is (to say))
INH	Isoniazid
JGI/IMG ER	Joint Genome Institute's Integrated Microbial Genomes with Expert Review
kbp	Kilo base pairs
KCN	Potassium cyanide
KD	Knockdown
KEGG	Kyoto Encyclopaedia of Genes and Genomes
Km	Kanamycin
Km^R	Kanamycin resistant
KO	Knockout
LA	Lysogeny-Bertani agar
LB	Lysogeny-Bertani broth
LC-MS/MS	Liquid chromatography tandem mass spectrometry
LOS	Lipooligosaccharides
LTBI	Latent tuberculosis infection
M. tb	*Mycobacterium tuberculosis*
MAB	*Mycobacterium abscessus* complex
MABA	Microplate alamar blue assay
MAC	*Mycobacterium avium* complex
Mb	Megabases
mc^2155	*Mycobacterium smegmatis* strain mc^2155
MCM	Methylmalonyl Co-A mutase
MDR-TB	Multi-drug resistant tuberculosis
MeCba	Methylcobamide
MeCbl	Methylcobalamin
metE	gene encoding the cobalamin-independent methionine synthase
MetE	cobalamin-independent methionine synthase
metH	gene encoding cobalamin-dependent methionine synthase
MetH	cobalamin-dependent methionine synthase
mg	Milligram
μg	Microgram
MGB	Minor groove binder

mic	Microti
MIC	Minimum inhibitory concentration
min	Minute(s)
ml	Millilitre
µl	Microliter
mM	Millimolar
µM	Micromolar
MMRU	Molecular Mycobacteriology Research Unit
MRCA	Most recent common ancestor
MRM	Multiple reaction monitoring
mRNA	Messenger RNA
ms	Millisecond
MS	Mass spectrometry
MTBC	*Mycobacterium tuberculosis* complex
NaCl	Sodium chloride
NADPH	Nicotinamide adenine dinucleotide phosphate
ng	nanogram
nM	nanomolar
nm	nanometres
ns-SNM(s)	Non-synonymous single nucleotide mutation(s)
NTM	Non-tuberculous mycobacteria
OADC	Oleic Albumin Dextrose Catalase
OD_{600nm}	Optical density (measured at 600nm wavelength)
HOCba	Hydroxycobamide
HOCbl	Hydroxycobalamin
Oligo	Oligonucleotide
PABA	para-aminobenzoic acid
PAM	protospacer adjacent motif
PAS	Para-aminosalicylic acid
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
Pfam (database)	Protein families
PRR	pattern recognition receptor
PZA	Pyrazinamide
RBS	Ribosome binding site
RCDC	Reducing agent and detergent compatible
RD	Region of difference
RIF	Rifampicin
RNA	Ribonucleic acid
RNI	Reactive nitrogen intermediates
ROS	Reactive oxygen species
rpm	revolutions per minute
rRNA	Ribosomal RNA
SAM	S-adenosyl-L-methionine

SCO	single crossover
SDS	Sodium dodecyl sulphate
SDZ	Sulfadiazine
sec	Seconds
SEM	standard error of mean
sgRNA	Short guide RNA
sigA	gene encoding RNA polymerase sigma factor A (σ^A)
SMZ	Sulfamethoxazole
SNM(s)	Single nucleotide mutation(s)
SRM	Single reaction monitoring
s-SNM(s)	Synonymous single nucleotide mutation(s)
STAGE-tips	STop And Go Extraction tips
STB	smooth tubercle bacilli
STR	Streptomycin
STZ	Sulfathiazole
Suc^S	Sucrose sensitive
SULFA	Suphonamide
TB	Tuberculosis
THF	Tetrahydrofolate
TMP	Trimethoprim
Tn	Transposon
Tn-seq	Transposon sequencing
UCT	University of Cape Town
Uro' III	Uroporphyrinogen III
USA	United States of America
UTR	Untranslated region
VIC	2'-chloro-7'-phenyl-1,4-dichloro-6-carboxyfluorescein
viz.	*videlicet* (that is (namely))
WGS	whole-genome sequencing
WHO	World Health Organization
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactoside

1 General Introduction

1.1 Tuberculosis (hereafter referred to as TB)

Formerly known as phthisis or consumption, TB claimed nearly 1.3 million lives in 2017, during which the estimated number of new cases of TB was 10.0 million people, according to the 2018 Global TB Report released by the World Health Organization (WHO) (World Health Organization, 2018). The same report further highlights that TB cases have been declining by approximately 2% since 2000, due to the use of anti-TB drugs and systematic strategies to manage the disease. This recent reversal of TB trends, however, belies an ancient relationship this disease shares with humans (Comas *et al*., 2013; Bos *et al*., 2014; Galagan, 2014; Supply & Brosch, 2017; Brynildsrud *et al*., 2018; Gagneux, 2018). Despite this shared history, the bacterium that causes TB was first described only 150 years ago by Jean-Antoine Villemin, who also demonstrated at the time that TB was a transmissible disease (Daniel, 2015). Building on this work, the Robert Koch went on to confirm in 1882 that *Mycobacterium tuberculosis*, so-named because of its fungus-like appearance in liquid culture (forming rod-shaped heaps with protective layers), was the etiological agent of TB (Cambau & Drancourt, 2014; Barberis *et al*., 2017). A century of TB studies since Koch's discovery have now revealed that this disease is caused by a group of closely related mycobacteria, collectively known as the *M. tuberculosis* complex (MTBC) (Tientcheu *et al*., 2017). In humans, TB is caused by *M. tuberculosis sensu stricto* and *M. africanum*, while animal TB-causing MTBC strains include *M. bovis*, *M. microti*, *M. pinnipedii*, *M. orygis*, *M. suricattae*, *M. mungi*, *M. caprae*, and a chimpanzee isolate (Hershberg *et al*., 2008; Boritsch & Brosch, 2016; Gagneux, 2018). For succinctness, the human TB-causing MTBC organisms are referred to as *M. tuberculosis* in this thesis, unless there is a specific need to distinguish *M. africanum*.

Koch's discoveries laid the foundation for the development of TB vaccines and four decades later, in 1921, the bacille Calmette-Guérin (BCG) vaccine, comprising a preparation of live attenuated *Mycobacterium bovis*, was pioneered. The BCG vaccine offers variable protection against TB. Nevetheless, it remains the only vaccine

currently approved by WHO, while a number of promising candidates are at various phases of clinical development (Voss *et al*., 2018).

The identification of the etiological agent of TB also paved the way for rational treatment. Initial efforts involved quarantine-based approaches, whereby patients were nursed in isolated sanatoria, some located at high altitudes to provide continuous fresh air and abundant sunshine (Greenhalgh & Butler, 2017). The first use of antibiotics against TB began in the latter half of the 1940s, followed by a rapid expansion of effective antibiotics including streptomycin (STR), para-aminosalicylic acid (PAS), isoniazid (INH), pyrazinamide (PZA), ethambutol (EMB) and rifampicin (RIF), all of which have different mechanisms of action. INH, PZA, EMB and RIF continue to be used today as first-line anti-TB drugs (World Health Organization, 2018). A significant obstacle to TB treatment is that the drugs must be taken in combination consistently over a period of six months, and preferably under direct observation by a health care provider. This TB control strategy, first promulgated by WHO in 1994, is called DOTS (directly-observed therapy, short course) (World Health Organization, 1994). Significant gains in TB diagnosis and treatment have been made using DOTS, most notably in countries where health care systems are efficient (Bloom *et al.*, 2017). Lapses in adherence to the relatively long treatment program undermines DOTS increasing the probability of undesirable outcomes such as disease recurrence, failure to control TB transmission and multi-drug resistance, the latter of which demands longer, and more complex and expensive regimens (Bloom *et al.*, 2017; Esmail *et al.*, 2018).

The emergence of multi-drug resistant *M. tuberculosis* strains and co-infection with the human immunodeficiency virus (HIV), reignited interest in anti-TB drug discovery, which had slowed in the 1970s (Vasava *et al*., 2017). The publication of the first complete whole-genome sequence of *M. tuberculosis* in 1998 (Cole *et al*., 1998) provided critical genome-based insights into the biology of the bacterium and mechanisms of drug-resistance, aiding in disease diagnosis and drug discovery efforts (Quan *et al*., 2018). Gradually, the number of potential drug candidates entering the discovery pipeline has grown, with Bedaquiline being the latest candidate to be approved and Delamanid undergoing clinical trials (https://www.newtbdrugs.org; (Esmail *et al*., 2018)). Although knowledge of TB infection, transmission, detection,

therapy and drug-resistance is accumulating, high global prevalence indicates that much work remains to be done to eradicate TB. **Figure 1.1** shows the global map of TB prevalence, with the worst affected continents including Africa and Asia.

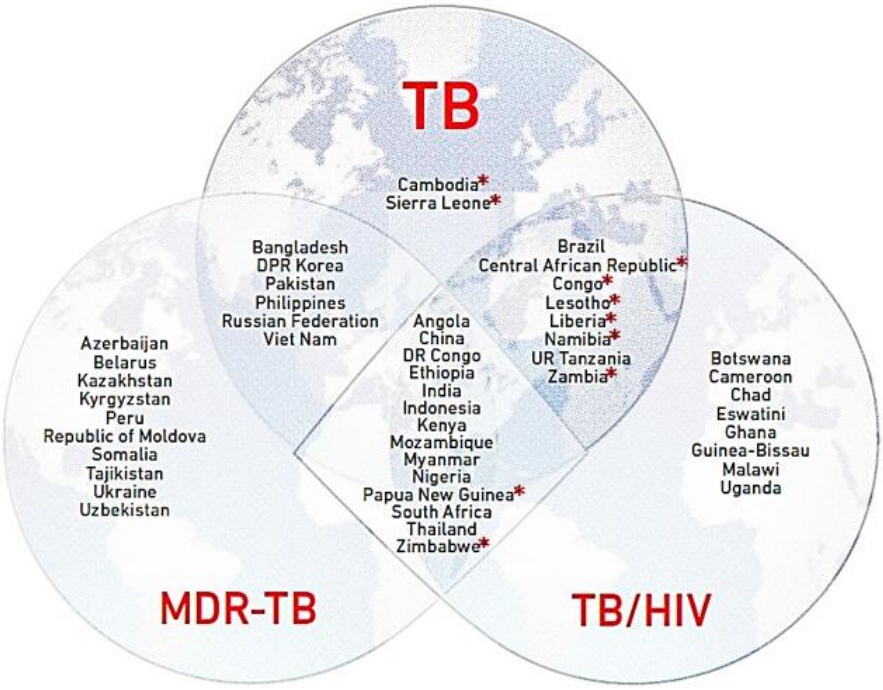


Figure 1.1: **Top 30 countries with the highest burden of TB, MDR-TB and HIV-TB co-infection in 2017.** The disease burden in the three categories is expressed as the absolute number of incident cases or in some instances, the number of incident cases per capita (asterisks). All of the worst-affected nations are found in Africa and Asia. The 14 countries in the central diamond have the highest rates in all three categories. Source: (World Health Organization, 2018).

At the time of his discovery, Robert Koch described the aerosol transmission of *M. tuberculosis* from sick to healthy pigs (reviewed in (Cambau & Drancourt, 2014)). A similar model of transmission is postulated to occur in humans, whereby the key steps involved include (1) the exit of bacilli from the respiratory tract of an individual with active TB as aerosolized droplets; (2) the survival of the bacteria in aerosol form outside of the original host; (3) inhalation of aerosols by a new host; (4) evasion of immune defences, and; (5) establishment of new infection (Turner *et al.*, 2017). Although this model assumes that humans are the exclusive reservoirs of *M. tuberculosis*, genomic studies suggest that the MTBC had a putative environmental ancestor (Supply *et al.*, 2013). Gradually, this putative ancestor gained human

virulence through a possible combination of loss or acquisition of gene function and fixation of single nucleotide mutations (SNMs) (Supply *et al*., 2013). It is believed that these events drove the evolution of *M. tuberculosis* as an obligate human pathogen (Supply *et al*., 2013; Orgeur & Brosch, 2018). In the most recent review on MTBC evolution, Sebastien Gagneux compared genomes of MTBC and non-tuberculous mycobacteria (NTM) and presented evidence of step-wise loss and/or acquisition of genomic features that have shaped the lifestyle of *M. tuberculosis* as a human pathogen (Gagneux, 2018). It is expected that further work in this area will shed more light into the evolution of mycobacteria, and the specializations that make *M. tuberculosis* a successful pathogen.

1.2 The evolution and global spread of *M. tuberculosis*

The MTBC share a hypothetical most recent common ancestor (MRCA) with *Mycobacterium canettii*, a highly recombinogenic, smooth-colony-forming mycobacterium which was isolated in the 1960s by George Canetti from lymph nodes of patients from the horn of Africa (van Soolingen *et al*., 1997; Gutierrez *et al*., 2005). Because of their unique colony morphology, *M. canettii* and *M. canettii*-like strains are also known as "smooth tubercle bacilli" (STB). Some STB strains can cause fatal disease (Koeck *et al*., 2011; Blouin *et al*., 2014), but they show reduced persistence in infection models (Supply *et al*., 2013); moreover, to date, no human-to-human transmissions of STB have been reported.

Routine applications of whole-genome sequencing (WGS) tools have now made possible genomic analyses of *M. canettii* to elucidate the molecular events that might have enabled *M. tuberculosis* to transition to humans from an environmental reservoir. MTBC and STB strains share approximately 3.9Mb of their genomes, with STB bearing on average 900 additional genes relative to *M. tuberculosis* (Supply *et al*., 2013; Boritsch *et al*., 2014). Of these extra genes, only 1% are common to all STBs, probably due to high rates of interspecies gene acquisition by horizontal gene transfer (HGT), characteristic of active association with an (as-yet unknown) environmental reservoir (Blouin *et al*., 2014; Boritsch & Brosch, 2016; Gagneux, 2018). In contrast to the high genetic diversity of STB, a pattern of clonal expansion and little to no genetic recombination is observed in clinical isolates of *M. tuberculosis* (Comas *et al*.,

2013). The extensive genetic variability of *M. canettii* strains relative to *M. tuberculosis* places STB at an earlier branching point before MTBC emergence. It is presumed that MTBC progenitor strains gradually acquired advantageous genetic features, including loss or acquisition of genes and the development of distinct host-preferences, which have contributed to their pathoadaptation (**Figure 1.2**) (Boritsch *et al*., 2014; Orgeur & Brosch, 2018).

One of the genes that has been lost in MTBC, but is present in STB, is the cobamide biosynthesis gene, *cobF*, encoding a precorrin-6a synthase homologue (**Figure 1.2**) (Supply *et al*., 2013; Blouin *et al*., 2014; Boritsch *et al*., 2014). Although there is no single organism in which the full set of cobamide biosynthesis genes is known or fully characterized, evidence put together from studies in various organisms suggest that cobamide biosynthesis is energetically costly, requiring nearly 30 enzyme-catalysed reactions (Mattes *et al*., 2017). The loss of *cobF* might explain the lack of *de novo* cobamide biosynthesis in MTBC strains (Gopinath, Venclovas, *et al*., 2013) although it has been suggested that alternative enzymes could play compensatory roles (Gopinath, Moosa, *et al*., 2013). To date, however, there is no experimental evidence that any member of the MTBC produce cobamides. It is likely, therefore, that the lack of cobamides in MTBC strains has forced these mycobacteria to depend on their host for these essential co-factors. Evidence of differential production of cobamides by STB and MTBC would strengthen this argument and, potentially, implicate cobamide metabolism in the evolution of MTBC pathogenicity.

Another important step in the evolution of MTBC virulence is speculated to have been the loss of the *pks5* gene, encoding a polyketide synthase involved in the production of lipooligosaccharides (LOS), a mycobacterial glycolipid (**Figure 1.2**). The smooth-colony morphology of STB is a result of the expression of Pks5, and experimental evidence shows that recombination events in this gene locus result in rough-colony morphology (Boritsch, Frigui, *et al*., 2016). Interestingly, Boritsch *et al.* also found that the LOS-deficient rough-colony variants had increased fitness and virulence in infection models compared to smooth-colony morphotypes. In addition, the surface modification following *pks5* recombination was shown to alter MTBC recognition by the host's immune system (Boritsch, Frigui, *et al*., 2016). Taken together, these genetic modifications appear to have contributed to MTBC virulence.

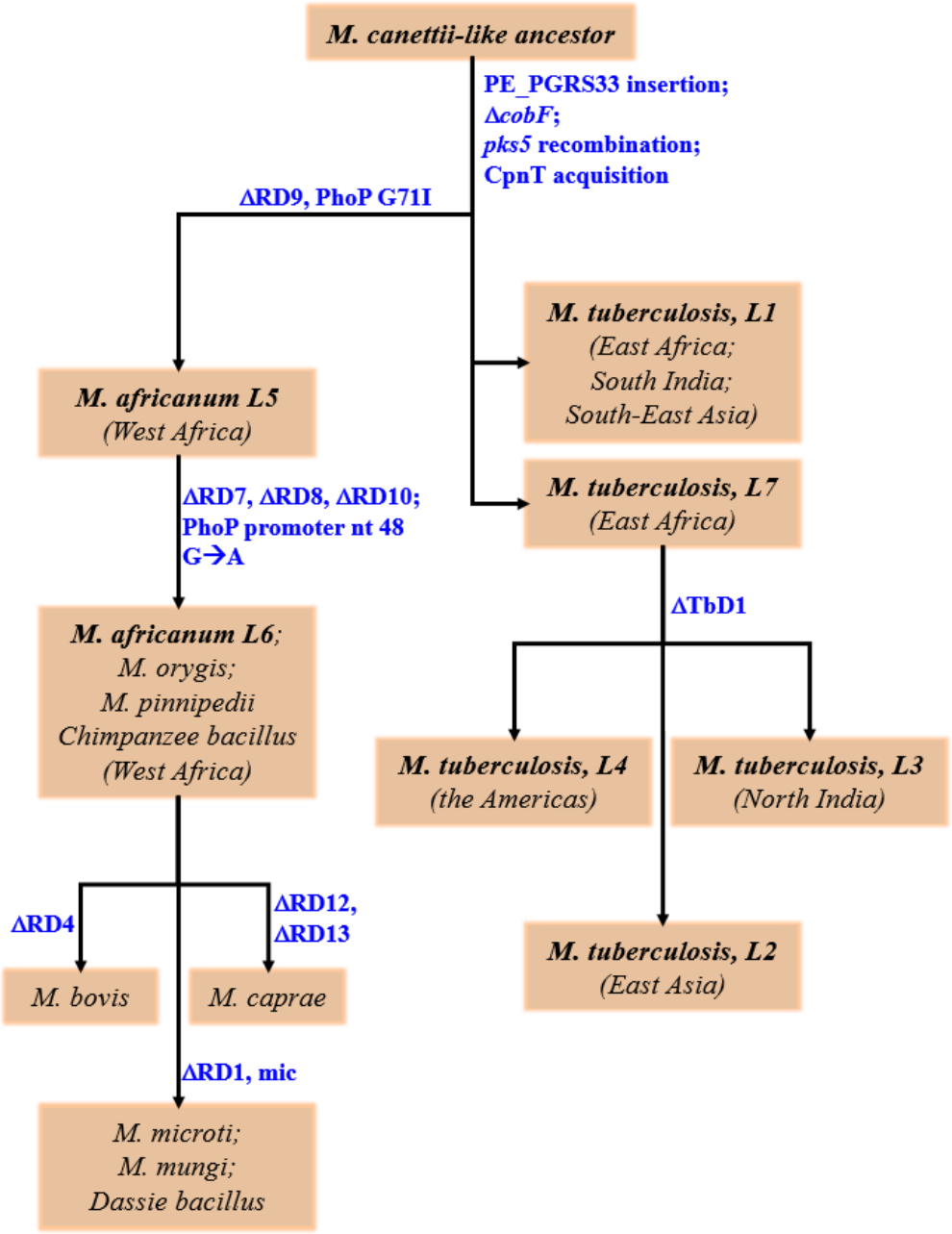


Figure 1.2: **Genetic adaptations in the evolution of virulence in the *M. tuberculosis* complex.** The MTBC mycobacteria share a hypothetical most recent common ancestor with *M. canettii*-like STB. The deletion of *cobF* (encoding a precorrin-6a synthase) and *pks5* (encoding a polyketide synthase) and insertion of *pe_pgrs33* (encoding an uncharacterized PE family protein) are some of the key molecular events in the evolution of MTBC from a recombinogenic *M. canettii*-like ancestor. The RD9 deletion gave rise to *M. africanum* L5 and subsequently, *M. africanum* L6 emerged carrying deletions in RD7, RD8 and RD10. Animal-adapted strains share a recent common ancestor with *M. africanum* L6. The RD9-deleted branch of MTBC have also accumulated mutations in *phoP* (encoding a two-component system response transcriptional positive regulator). The L1 and L7 lineages diverged before the emergence of modern lineages that have the *M. tuberculosis*-specific deletion (TbD1). The geographical locations where a strain is most prevalent are shown in brackets. Molecular events and evolutionary branching-off points are inferred from (Boritsch & Brosch, 2016; Gagneux, 2018; Orgeur & Brosch, 2018). CpnT – **c**hannel **p**rotein with **n**ecrosis-inducing **t**oxin; RD – **r**egion of **d**ifference; mic – *microti*.

The loss of genes previously encoded by the STB ancestor is not the only feature marking the transition to virulence. At the time of branching from the hypothetical most recent common ancestor (MRCA), MTBC strains possibly also acquired MTBC-specific genes such as the *PE_PGRS33* gene, encoding an uncharacterized PE family protein (**Figure 1.2**) (Boritsch & Brosch, 2016; Boritsch, Khanna, *et al*., 2016). PE_PGRS33 belongs to a subfamily of mycobacteria-specific proteins designated as PE, because of a conserved Pro-Glu (PE) motif at residues 7-8 in their N-terminal regions, and an enriched C-terminal polymorphic GC-rich sequence (PGRS) motif (Bottai & Brosch, 2009). Together with the PPE proteins, which have a characteristic Pro-Pro-Glu (PPE) motif at residues 7-9 in the N-terminal region, the PE-PPE proteins collectively encode approximately 10% of the *M. tuberculosis* genome and 4% of the total proteome (Cole *et al*., 1998; Gey van Pittius *et al*., 2006). The functions of most PE and PPE proteins remain unknown but evidence of co-evolution with the ESX secretion system and surface localization has led to notion that these proteins are involved in virulence and immunogenicity (Gey van Pittius *et al*., 2006; Delogu G, Cole ST, 2008; Sampson, 2011; Vordermeier *et al*., 2012; Gröschel *et al*., 2016; Brennan, 2017). Because of its localization at the mycobacterial outer membrane, PE_PGRS33 is reportedly involved in mycobacterial cell wall stability and in host-pathogen interactions (Zumbo *et al*., 2013).

Another MTBC-specific gene that was acquired during evolution encodes the C-terminal domain of CpnT, a mycobacterial **c**hannel **p**rotein with **n**ecrosis-inducing **t**oxin that induces necrosis in mammalian macrophages (Danilchanka *et al*., 2014, 2015; Sun *et al*., 2015). The N-terminal domain of CpnT is similar between STB and MTBC, but the C-terminal domain, which comprises the **t**uberculosis **n**ecrotizing **t**oxin (TNT) (Sun *et al*., 2015), is different. This difference implies that the acquisition of the C-terminal domain occurred after the separation of MTBC strains (**Figure 1.2**). Apart from PE_PGRS33 and CpnT, which have been linked experimentally to virulence, there are three other MTBC-specific genes (encoded by homologues of *Rv0394c*, *Rv2023c*, and *Rv3190c*) that were acquired during MTBC evolution, but which remain to be characterized (Boritsch & Brosch, 2016).

As MTBC emerged, distinct host-preferences seem to have also developed, resulting in animal-adapted strains and human-adapted lineages (**Figure 1.2**) Studies have revealed that selected deletions in regions of difference (RD) are consistent with independent evolution of animal- and human-adapted MTBC (Mostowy *et al*., 2002; Pym *et al*., 2002; Supply *et al*., 2013; Blouin *et al*., 2014). There are seven lineages of the human-adapted MTBC, which are designated as lineages 1 through 7 (L1–L7) (**Figure 1.2**). Of the 7 lineages of human-adapted MTBC strains, *M. africanum* falls into L5 and L6, which are the oldest lineages. This is followed in chronological order by L1 and L7, which subsequently gave rise to L2 (Beijing strain), L3 and L4 (Euro-American strain). These last three are considered the most modern *M. tuberculosis* lineages and have a characteristic *M. tuberculosis*-specific deletion (TbD1) (**Figure 1.2**). Lineages L2 and L4 are globally widespread, while the rest of the lineages have remained largely geographically distinct (Brosch *et al*., 2002; Merker *et al*., 2015; Stucki *et al*., 2016). L1 is most prevalent in East Africa, South India and South East Asia; L2 in East Asia; L3 in North India; L4 in the Americas; L5 & L6 in West Africa; and L7 in East Africa. The absence of RD9 is a key distinguishing feature of *M. africanum* from the other human-adapted lineages (**Figure 1.2**). The RD9-deleted branch of the MTBC also have accumulated mutations in *phoP*, encoding a two-component system response transcriptional positive regulator involved in virulence (Gonzalo-Asensio *et al*., 2014). Animal-adapted MTBC share a most recent ancestor with *M. africanum* L5 and L6, with additional deletions in RD7, RD8 and RD10 defining the emergence of *M. pinnipedii* (seal), *M. orygis* (oryx) and Chimpanzee bacillus (**Figure 1.2**). Subsequent RD4 deletion gave rise to *M. bovis* (cattle), while *M. caprae* (goat) emerged independently due to the loss of RD12 and RD13. The RD1 deletion occurred independently in *M. microti* (rodents), *M. mungi* (mongoose) and Dassie bacillus genomes (**Figure 1.2**).

According to most recent estimates, 25% of the total world population is infected with TB (Houben & Dodd, 2016). The regions between Eastern Africa and South-East Asia are often cited as the epicentre of an outward early spread of MTBC to other parts of the world, in parallel with an "out-of-Africa" human dispersal hypothesis (Hershberg *et al*., 2008; Wirth *et al*., 2008; Comas *et al*., 2013; O'Neill *et al*., 2017). A key piece of evidence supporting this view is the observation that all surviving *M. canettii* strains, and all animal- and human-adapted MTBC lineages, have been found in

several parts of Africa (Coscolla *et al*., 2013; Comas *et al*., 2015; Nebenzahl-Guimaraes *et al*., 2016). A recent analysis of lineage 4 (L4), the most widely-spread lineage worldwide, revealed an "out-of-Europe" global expansion corresponding to European colonization of Africa and the Americas (Brynildsrud *et al*., 2018). These patterns are emblematic of the intimate relationship between *M. tuberculosis* and humans. The ease of modern human movement provides opportunities for previously geographically distinct lineages to mix according to the human-to-human transmission model.

Yet, despite increased intercontinental human migration, it is still unclear why some lineages are more globally spread (*viz.*, L2 & L4) while others remain restricted. Interestingly, the most distantly-related members of the MTBC differ by roughly 2000 SNMs (Coscolla & Gagneux, 2014), which when juxtaposed with the approximately 65000-SNM difference in STB (Supply *et al*., 2013), indicate that MTBC are very closely related. Some of these SNMs, such as the *phoP* loss-of-function mutations in lineages L5, L6 and the animal-adapted strains, are associated with fitness loss (Gonzalo-Asensio *et al*., 2014). The Beijing strain (L2) has been reported to be hyper-virulent, more transmissible and more amenable to acquiring drug resistance, although the molecular events underlying this phenomena are unclear (Lopez *et al*., 2003; Rad *et al*., 2003; Mokrousov, 2013; Merker *et al*., 2015; Boritsch & Brosch, 2016).

1.3 *M. tuberculosis* pathogenesis

The human-to-human transmission model of TB is anchored by three simplistic components: (1) an infected host, (2) a transmission route (air) and (3) a new host. The outcome of infection can vary depending on the immune state of the new host and fitness of the infecting bacillus, or a combination of these factors (Drain *et al*., 2018; Lin & Flynn, 2018). Consequently, a newly-infected individual may remain asymptomatic after exposure, a (presumably) noncontagious phase of disease known as latent *M. tuberculosis* infection (LTBI), which is characterized by absence of clinical symptoms despite a positive immunologic test (Drain *et al*., 2018; Lin & Flynn, 2018). This scenario is likely what happens in most cases of exposure because only a small fraction of new infections lead to active, transmissible disease (Vanderven *et al*., 2016; Lin & Flynn, 2018). The progression to active disease can be accelerated

in immunocompromised individuals such as those co-infected with the human immunodeficiency virus (HIV) and type-2 diabetes. Alcohol abuse, cigarette smoking, malnutrition, and repeated exposure to *M. tuberculosis* in overcrowded living conditions, may also lead to poor control of infection.

Once inhaled, aerosolized bacilli are encountered by alveolar macrophages, which play a frontline role in sterilization of the invading *M. tuberculosis* (**Figure 1.3**) (Srivastava, Ernst & Desvignes, 2014). *M. tuberculosis* can also infect other cell types including dendritic cells, neutrophils, adipocytes, natural killer (NK) cells, and epithelial cells (Srivastava, Ernst & Desvignes, 2014; Randall *et al*., 2015). Furthermore, TB is not restricted to lungs; in fact, extrapulmonary TB makes up between 20–25% of notified new infections annually (World Health Organization, 2018). Therefore, differences can be expected in the metabolic interactions between bacilli and host cells in different microenvironments. However, the macrophage response has been most extensively studied and is therefore the best understood. Several mycobacterial pattern recognition receptors on the surface of macrophages bind mycobacterial lipids leading to phagocytosis, a process which captures bacilli in vacuoles called a phagosomes (Kaufmann, 2001). *M. tuberculosis* can escape or subvert host immune responses, and bacillary replication can occur within individual alveolar macrophages (**Figure 1.3**) (Sun *et al*., 2015; Uribe-Querol & Rosales, 2017). However, as more macrophages take up the mycobacteria, they become activated and a cascade of immune responses begins to slow down intracellular bacterial growth. Initially, there is an influx of trace metals that are toxic to mycobacteria into the phagosome (**Figure 1.3**). It has also been shown that trace elements that can be used by mycobacteria as micronutrients are sequestered (Neyrolles *et al*., 2015). As the macrophage matures, it fuses with a lysosome, causing the internal pH to shift from around pH 6.5 to pH 4.5 (**Figure 1.3**). *M. tuberculosis*, however, can block normal acidification of phagosomes by preventing fusion with lysosomes (Vergne *et al*., 2005), which maintains the pH at levels that are permissive to mycobacterial replication.

The macrophage maturation process is also accompanied by a superoxide burst, which further slows down *M. tuberculosis* replication (Russell, 2011). Still, bacilli can continue to replicate within the macrophage while resisting the oxidative stress (Ehrt & Schnappinger, 2009). In addition to macrophage activation, lymphocytes and lipid-

rich foamy macrophages are recruited, forming the granuloma, a mass of cells which encase the infected macrophage cell (**Figure 1.3**) (Guirado & Schlesinger, 2013). Ultimately, the granuloma develops a necrotic core and traps *M. tuberculosis* in a hypoxic, lipid-rich environment. At this point, *M. tuberculosis* can enter a metabolically "quiescent" state characterized by slowed replication (Tailleux *et al*., 2008). If *M. tuberculosis* survives this onslaught of immune responses, bacilli can ultimately break out of the caseous granuloma into the respiratory cavity, where they can be expelled and transmitted within aerosols (**Figure 1.3**) (Rhee *et al*., 2011; Gouzy *et al*., 2014). Evidently, this model oversimplifies granuloma behaviour by implying that the granulomatous response is either sterilizing or permissive, leading to LTBI. However, recent studies suggest a much more complex interaction between bacilli and macrophage, characterized by a high degree of diversity of responses depending on the lineage and sterilization efficacy of individual granulomas (Lin *et al*., 2014; Huang *et al*., 2018; Hunter, 2018).

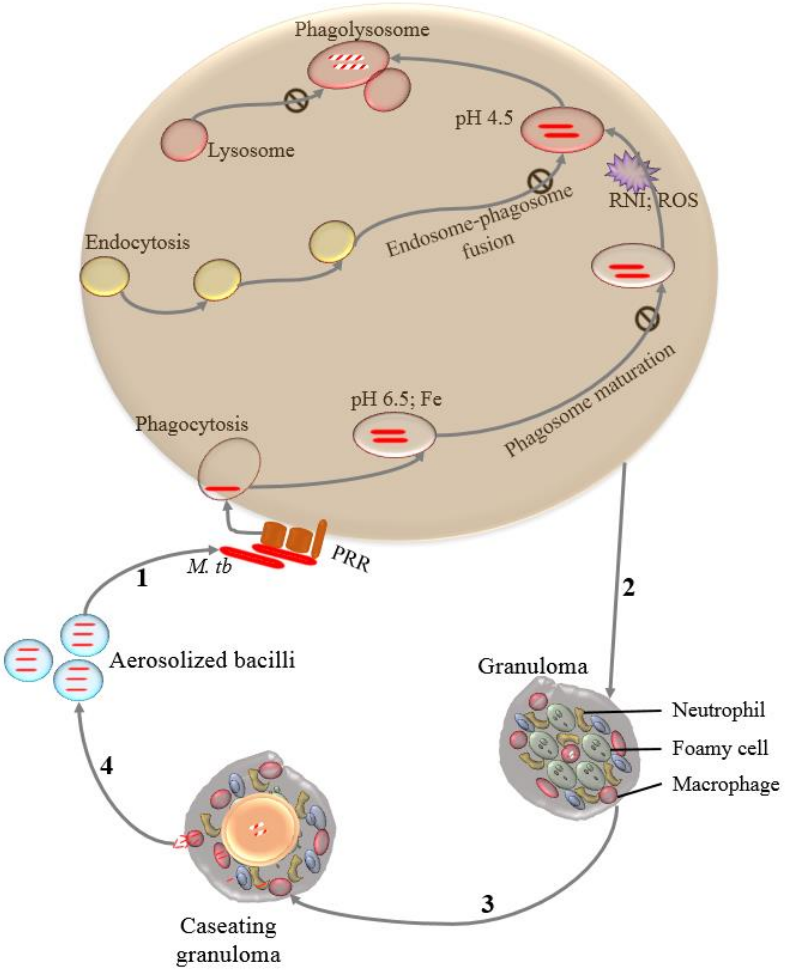


Figure 1.3: **A schematic of *M. tuberculosis* pathogenesis**. **1**. Aerosolized bacilli inhaled by a new host are recognized by pattern recognition receptors (PRR) on the surfaces of alveolar macrophages and phagocytosed. The phagosome then undergoes a process of maturation which includes acidification, fusion with lysosomes, accumulation of toxic trace metals, and release of reactive nitrogen intermediates (RNI) and reactive oxygen species (ROS) that slow down growth and kill invading bacilli. *M. tuberculosis* can counter host immune responses at various steps of this process, including blocking phagosomal maturation, endosome-phagosome fusion, and phagolysosome fusion. **2**. Granuloma formation includes the recruitment of foamy cells and neutrophils to the site of infection. **3**. Granuloma maturation leads to caseation, trapping bacilli inside a lipid-rich centre. **4**. Disease reactivation due to the breaking-out of bacilli into the respiratory passage, where they are expelled as aerosolized droplets and transmitted to new hosts.

1.4 Cobamide-related metabolism in TB disease

In the context of TB disease, bacilli and the host cells require energy for normal metabolic functions and are often pitted against one another for the same resources in

discrete niches of infection. For the purposes of this thesis, only those aspects relating to cobamide metabolism are briefly highlighted in this section.

As the primary environment for *M. tuberculosis*, the granuloma is rich in high-energy fatty acids and cholesterol, which are broken down by bacilli to generate energy and precursors for polyketides and phospholipids (Van der Geize *et al*., 2007; Peyron *et al*., 2008; Daniel *et al*., 2011; Beste *et al*., 2013; Lee *et al*., 2013; Quadri, 2014; Lovewell, Sassetti & VanderVen, 2016). *M. tuberculosis* expresses two enzymes that are required to utilize acetyl-CoA derived from fatty acids: isocitrate lyase (ICL) and phosphoenolpyruvate carboxykinase (PEPCK) (Muñoz-Elías *et al*., 2006; Marrero *et al*., 2010; Eoh & Rhee, 2014). The consumption of cholesterol yields propionyl coenzyme A (PrP-CoA), which is converted by 2-methylcitrate synthase (PrpC; EC 2.3.3.1), to 2-methylcitrate, a noxious intermediate that inhibits gluconeogenesis (Horswill, Dudding & Escalante-Semerena, 2001; Rocco & Escalante-Semerena, 2010). The toxicity of 2-methylcitrate is resolved by converting this intermediate to pyruvate and succinate, which enter the tricarboxylic acid (TCA) cycle via the glyoxylate shunt (McKinney *et al*., 2000; Eoh & Rhee, 2014). Alternatively, PrP-CoA is first converted to methylmalonyl coenzyme A (MM-CoA) , which is further converted to succinyl coenzyme A via the MM-CoA pathway. The MM-CoA route involves the cobalamin-dependent enzyme, (*R*)-methylmalonyl-CoA mutase (MCM; EC 5.4.99.2) (Savvi *et al*., 2008). Another way in which cells protect against 2-methylcitrate toxicity is by utilizing PrP-CoA and acetyl coenzyme A as building blocks for cell envelope components: phthiocerol dimycocerosates (PDIM), polyacylated trehaloses (PATS) and sulfolipids (SL), which are associated with mycobacterial virulence (Russell *et al*., 2010; Lee *et al*., 2013; Lovewell, Sassetti & VanderVen, 2016). In one study using a mouse infection model, *M. tuberculosis* mutants in which the 2-methylcitrate pathway was disabled were able to establish infection, presumably by engaging the MM-CoA pathway to overcome 2-methylcitrate toxicity (Muñoz-Elías *et al*., 2006). These findings raised the possibility that bacilli could use host-derived cobamides to enable activity of the MCM enzyme.

1.5 Cobamide chemical structure

Cobamides are complex cyclic tetrapyrrolidine that contains a centrally-chelated cobalt ion coordinated equatorially by four nitrogens donated by the nitrogen atom of each pyrrole ring and axially by an upper and a lower nitrogenous or phenolic base (**Figure 1.4A**). The upper axial face is occupied by an alkyl ligand, represented as an "R-group," (**Figure 1.4A**) from which the corresponding cobamide form derives its common name. If "R" is a cyano group, the cobamide is referred to as cyanocobamide (CNCba); if a deoxyadenosine group, then adenosylcobamide (AdoCba); if a methyl group, then methylcobamide (MeCba); and if a hydroxyl group, then hydroxycobamide (HOCba). AdoCba is one of two biologically functional cobamide forms and is involved in radical chemistry reactions such as carbon rearrangements, dehydrations and deaminations (Halpern, 1985). The other biologically functional cobamide form is Co(I)cobamide, in which the cobalt ion is in the +1 state (Co^{+}). Co(I)cobamide is involved in methyl transfers and dehalogenation reactions (Koutmos *et al.*, 2009; Dassanayake, Cabelli & Brasch, 2013; Payne *et al.*, 2015; Dowling *et al.*, 2016). Cobalamin is the only cobamide that contains 5,6-dimethylbenzimidazole (DMB) as the lower base. The identity of the "R" group also determines the type of cobalamin form, whereby the alkyl group is a cyano in cyanocablamin (CNCbl; also known as vitamin B_{12}); an adenosyl in adenosylcobalamin (AdoCbl; also known as coenzyme B_{12}); a methyl in methylcobalamin (MeCbl) and a hydroxyl in hydroxycobalamin (HOCbl). When the nitrogen of DMB is coordinated to the cobalt ion, it is referred to as a "base-on" (*i.e.*, DMB-on) configuration (**Figure 1.4B**). When cobalamin-binding proteins interact with the co-factor, they also do so either in the "base-on" configuration or in a "base-off" conformation in which DMB has been displaced by an amino acid residue. If a histidine is the coordinating residue, the interaction is "base-off," "His-on," but if a different amino acid is involved, then the coordination "base-off," "His-off" (**Figure 1.4B**) (Drennan *et al*., 1994).

The Co–C bond between the cobalt ion and a coordinating carbon of the R-group determines the catalytic reactivity of the cobamide. The Co–C bond undergoes different types of cleavage, which alters the oxidation states of cobalt, enabling cobamides to act as co-factors in a variety of reactions. For example, in rearrangement reactions, homolytic cleavage is required to form reactive, but transient,

deoxyadenosyl radical species. In methyl transfer reactions, heterolytic cleavage cycles the oxidation state of cobalt between Co^{3+} and Co^{+} (Giedyk, Goliszewska & Gryko, 2015). The Co–C bond is usually inert in the absence of light but, when excited by light, photolysis occurs relatively easily (Weissbach *et al.*, 1960; Jones, 2017; Padmanabhan *et al.*, 2017). The versatility of the coordination states between cobalamin and proteins and the lability of the Co–C illustrate the broad spectrum of cobalamin biochemistry (Parks *et al*., 2013; Bridwell-Rabb, Grell & Drennan, 2018).

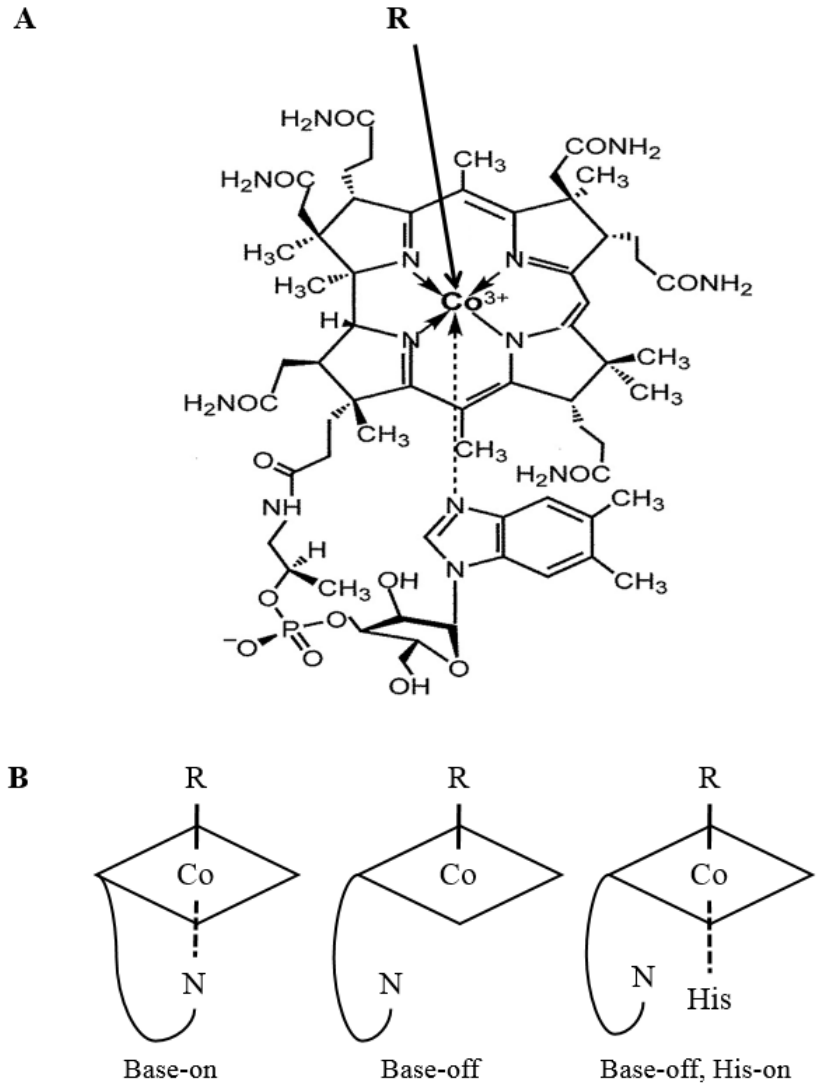


Figure 1.4: **The chemical structure of a cobamide. A**. Cobalamin structure: a corrin ring surrounds a centrally-chelated cobalt ion. The cobalt ion is coordinated by an alkyl upper ligand (R-group) and 5,6-dimethylbenzimidazole (DMB). The R-group can be occupied by cyano- (cyanocobamide), adenosyl- (adenosylcobamide), methyl- (methylcobamide) or hydroxyl- (hydroxycobamide) groups. **B**. On-base/off-base configurations of lower ligand coordination.

1.6 Cobamide biosynthesis in mycobacteria

Although cobamides are required by cells in all domains of life (Nielsen *et al*., 2012), *de novo* biosynthesis of these molecules is restricted to some bacteria and archaea (Zhang *et al*., 2009; Doxey *et al*., 2015). Mycobacteria possess most of the genes involved in *de novo* cobamide biosynthesis, except for members of the MTBC which are missing *cobF* as previously discussed in **section 1.2** of this chapter. However, the

loss of *cobF* is not the only deficiency in cobamide biosynthetic capacity in the RD9-deleted lineages (L5, L6). This deletion caused a 5' truncation of *cobL*, which occurs in an operon with *cobM* and *cobK* (Brosch *et al*., 2002), potentially disrupting function of all three genes. It has also been observed that in many clinical isolates of *M. tuberculosis*, potentially inactivating SNMs occur in *cobN* (Young, Comas & de Carvalho, 2015). Interestingly, *M. leprae*, which has a reduced genome owing to extensive decay (Cole *et al*., 2001), is predicted to either contain pseudogenes or entirely lack certain cobamide biosynthetic genes (Rodionov *et al*., 2003; Zhang *et al*., 2009; Gopinath, Moosa, *et al*., 2013). These genomic insights suggest that *M. tuberculosis* and other pathogenic mycobacteria might have lost the ability to produce cobamides and have come to rely on the host for these compounds.

Aerobic and anaerobic routes exist for the biosynthesis of cobamides, which are distinguished by their requirements for molecular oxygen, the timing of the cobalt insertion step (which occurs early in anaerobic biosynthesis, and late in aerobic biosynthesis), and the mechanism of ring contraction (Warren *et al*., 2002). As a convention, the prefix "*cob*" designates genes involved in the oxygen-dependent process, while "*cbi*" denotes genes in the oxygen-independent route. It is an acknowledged fact that there is confusion regarding the use of these gene prefixes. For example, all cobamide biosynthesis genes are prefixed "*cob*" in *P. denitrificans* whereas in *S. typhimurium* the prefix "cbi" is used to distinguish genes required to synthesise cobinamide, the precursor of cobalamin, from the genes required to convert cobinamide to cobalamin, which are prefixed "*cob*" (Blanche *et al.*, 1995; Raux *et al.*, 1998; Warren *et al.*, 2002; Mattes *et al.*, 2017). In this thesis, the "*cob*" prefix is used to designate the aerobic pathway, and to differentiate those genes used in the anaerobic pathway, the prefix "*cbi*" is used. The cobalt-containing intermediates of the anaerobic pathway are distinguished by the prefix "Co-"; for example, Co-Precorin-6A. Either route can be divided into two major steps. The first is corrin ring synthesis while the second is the assembly and attachment of the lower base.

Like siroheme, co-factor F430 and chlorophyll, cobamides are derived from the general tetrapyrrole precursor, 5-aminolevulinic acid (ALA), a product of the C-5 glutamate pathway (Oh-hama, Stolowich & Scott, 1988; Kang *et al*., 2012). ALA undergoes condensation, polymerization, rearrangement and cyclization to form

uroporphyrinogen III. Committed cobamide biosynthesis begins when uroporphyrinogen III is methylated at C-2 and C-7 to form precorrin-2 (Lawrence *et al*., 1995). This step is catalysed by CysG, a multifunctional S-adenosyl-L-methionine:uroporphyrinogen III methyltransferase (SUMT). CobA, a smaller SUMT sharing structural homology with the catalytically relevant C-terminus of CysG is encoded in *M. smegmatis*. CysG also acts as a precorrin-2 oxidase and a ferrochelatase. Another protein, HemD, can also catalyse Uroporphyrin III methylation. *M. tuberculosis* expresses *cysG* and *hemD*, both of which have been identified in transposon (Tn) mutagenesis screens as independently essential for *in vitro* growth (Griffin *et al*., 2011; Dejesus *et al*., 2017). This is likely because in addition to the cobamide pathway, the products of the reactions catalysed by the two enzymes are also precursors for siroheme and co-factor F430.

Corrin ring synthesis begins after precorrin-2, and these steps occur differently under aerobic and anaerobic conditions. In the aerobic pathway, CobG initiates the ring contraction before late-stage cobalt insertion by CobNST, a cobalt chelatase complex formed by three enzymes: CobN, CobS and CobT. In the absence of oxygen, CbiX or CbiK first catalyse the cobalt insertion step (Moore & Warren, 2012). The two routes reunite at the corrinoid adenosylation step which produces cob(II)yrinate a,c,-diamide (Warren *et al*., 2002). The second stage of cobamide biosynthesis, the assembly of DMB, can also occur either aerobically or anaerobically (Renz, 1970; Keck, Munder & Renz, 1998; Warren *et al*., 2002; Hazra *et al*., 2015). Mycobacterial cobamide production is discussed in more detail in **Chapter 3** of this thesis.

1.7 Cobamide transport in mycobacteria

Studies show that treatment outcomes worsened in TB patients suffering from pernicious anaemia when regimens that included vitamin B_{12} supplements were prescribed (Ramagopalan *et al*., 2013). These clinical cases suggest that pathogenic mycobacteria might have specialized mechanisms to access cobamides and/or precursors as a cost-effective alternative to *de novo* biosynthesis.

Accumulating evidence indicates that *M. tuberculosis* transports cobamides when cultured *in vitro* (Warner *et al*., 2007; Savvi *et al*., 2008; Griffin *et al*., 2012; Lee *et al*., 2013; Boot *et al*., 2016). This transport is mediated by BacA, an ATP-binding

cassette transporter protein encoded by the Rv1819c gene which has also been implicated in drug efflux (Gopinath, Venclovas, *et al*., 2013), hinting that cobamide uptake in mycobacteria is unlike that of Gram-negative bacteria which rely on a well-characterized system consisting of the outer membrane protein, BtuB, the transmembrane proteins TonB, ExbB and ExbD, the periplasmic protein, BtuF, and the inner membrane complex, BtuCD (Cadieux *et al.*, 2002; Locher, Lee & Rees, 2002; Noinaj *et al.*, 2010; Korkhov *et al.*, 2014). Thus, extracellular cobamides are carried as precious cargo from one protein to another, into the bacterial cytosol. Hovever, a cobamide delivery system inside the bacterial cytosol has not been described and it is unknown how cobamide-dependent enzymes find the co-factor. In mammals, an *in vitro* model for the multistep shuttling of food-derived cobamides across intestinal cell membranes and delivery to cobamide-dependent enzymes, in which at least 15 proteins carry the co-factor to cytosolic and mitochondrial sites, has been proposed (Banerjee, 2006; Padovani *et al.*, 2008; Banerjee, Gherasim & Padovani, 2009; Nielsen *et al.*, 2012).

Evidently, cobamide uptake is complex and a complete characterization of this process in mycobacteria is made all the more difficult by the fact that unlike Gram negative bacteria, mycobacteria have an unusual and notoriously complex cell envelope (Niederweis *et al*., 2010). This highly hydrophobic envelope presents a formidable barrier to passive diffusion of large water-soluble compounds such as cobamides, strongly suggesting that a distinct active transport mechanism exists for this molecule. However, given the long generation time for most pathogenic mycobacteria, the requirement for fast uptake of small molecules might not be strict, and passive transport may be all that it takes to achieve the necessary amount of the co-factor inside cells. However, the possibility of passive diffusion of cobamides is very low, as a cobamide would be fully protonated at the optimal pH for bacterial growth, which would require active transport. Therefore, the identification of cobamide chaperones could contribute significantly to understanding the role of cobamides in mycobacterial physiology and pathogenesis.

1.8 Riboswitch regulation of mycobacterial cobamide metabolism

In mycobacteria, there are only three known cobamide-related metabolic processes: biosynthesis of methionine, ribonucleotide reduction, and propionyl CoA catabolism via the methylmalonyl pathway. Interestingly, each of the three metabolic processes has cobamide-dependent and -independent enzymes (Young, Comas & de Carvalho, 2015). Of the cobamide-independent enzymes, only the cobalamin-independent methionine synthase, MetE (5-methyltetrahydropteroyltriglutamate-homocysteine methyltransferase, EC 2.1.1.13), has been experimentally shown to be under direct control by a cobalamin riboswitch (Warner *et al*., 2007). Apart from the *metE* riboswitch, comparative genomics have identified other cobalamin riboswitches in the 5' untranslated region (5' UTR) of putative operons comprising putative ABC-type cobamide transporters found only in environmental mycobacteria. In addition, these environmental organisms also have other predicted cobalamin riboswitches in front of putative operons containing predicted cobalt transporters (Novichkov *et al*., 2013). In pathogenic strains, the only other riboswitch (*i.e.*, apart from the *metE* riboswitch) occurs in the 5' UTR of a three-gene operon cosisting of a putative cobalt transport gene *ppe2*, and cobamide biosynthesis genes *cobQ1* and *cobU* (Rodionov *et al*., 2003).

Riboswitches are conserved non-coding RNA secondary structures usually located in the 5' UTRs of the messenger RNA (mRNA) of the genes under their control (Nahvi, Barrick & Breaker, 2004). These RNA elements sense and bind a variety of specific ligands including coenzymes, metal ions, nucleobases, tRNAs, sugars, amino acids, or toxic metabolites (Ignatov & Johansson, 2017). They can also act as thermo- or pH-switches to control gene expression in response to temperature or pH changes (Nechooshtan *et al*., 2009; Ignatov & Johansson, 2017). It is classically understood that ligand binding or stimulus sensing triggers conformational changes in the mRNA, which block or permit transcription and/or translation (Breaker, 2012). Transcription control is achieved by the formation of a terminator structure while the sequestration of the ribosome binding sites (RBS) leads to inhibition of translation (Lussier *et al*., 2015). Riboswitches typically comprise two parts: a highly conserved RNA aptamer and an expression platform, which usually is made of less conserved sequences (**Figure 1.5**) (Haller, Soulière & Micura, 2011). The aptamer domain tightly binds a ligand, inducing allosteric changes in the riboswitch structure, which occludes the

transcription start codon or RBS (**Figure 1.5**). These rearrangements control the expression of the adjacent gene or operon (Sherwood & Henkin, 2016).

Even before the use of the term 'riboswitch' to refer to ligand-sensing RNA, the regulation of gene and protein expression by cobalamin had been demonstrated in *E. coli* and *S. typhimurium*, in which cobalamin controlled BtuB expression by RBS sequestration in both organisms (Ravnum & Andersson, 1997; Nou & Kadner, 2000). Advances in structural biology have now made available high-resolution X-ray structures of the cobalamin riboswitches in *E. coli*, providing further insights into the riboswitch mechanism. The aptamer domain of the cobalamin riboswitch is a common core surrounded by four peripheral extensions (P), also known as stems (**Figure 1.5**) (Johnson Jr *et al*., 2012; Peselis & Serganov, 2012). The stems are connected by loops (L) and junctions (J). Another common element of this riboswitch is a tertiary structure called the "kissing loop," which is formed by an interaction of L5 and L13 (**Figure 1.5**). The kissing loop is a peculiar feature of cobalamin riboswitches which has been shown to improve ligand binding efficiency although it is not essential for ligand binding (Lussier *et al*., 2015). This tertiary structure, however, has been found to fold more slowly than secondary structures (Soulière *et al*., 2013), a mechanistic problem which is solved by the presence of three strategic regulatory pause sites that coordinate independent folding of the aptamer and expression platform domains (Perdrizet *et al*., 2012). The L13/P13 structure (**Figure 1.5**) acts as the anti-terminator or anti-sequester for transcriptional regulation or translational control, respectively.

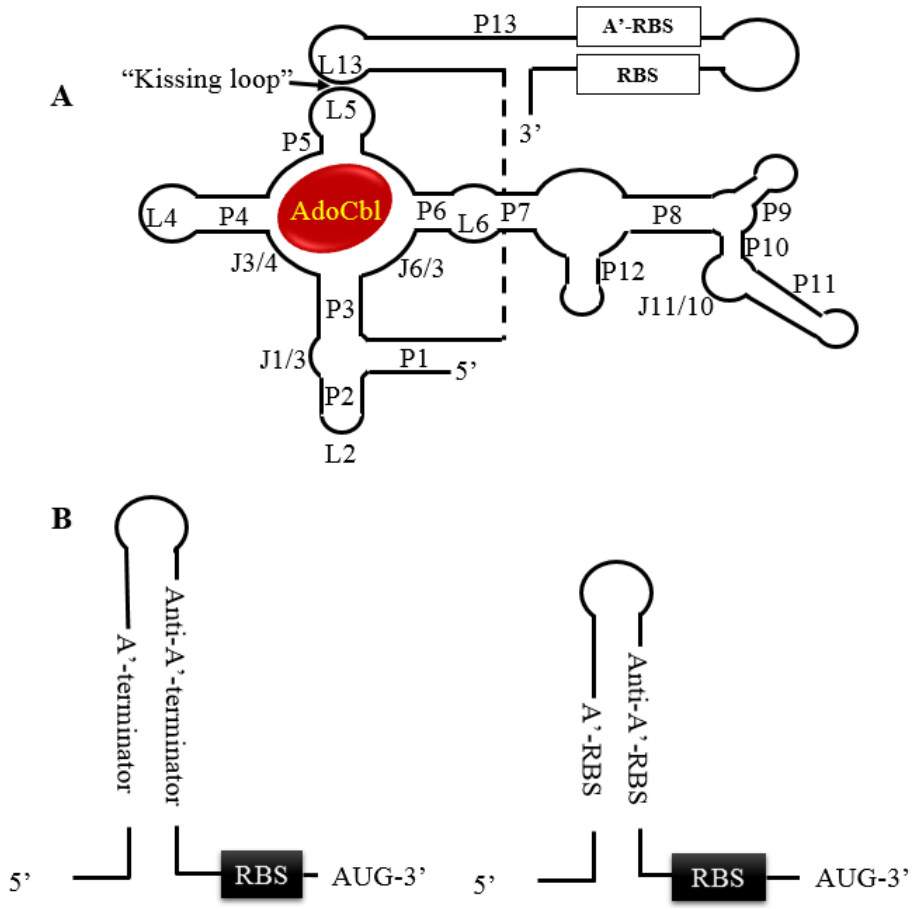


Figure 1.5: **A schematic of the consensus structure of the cobalamin riboswitch**. **A**. Ligand-bound conformation of a translationally-controlling cobalamin riboswitch. A conserved, centrally situated aptamer core lies at the intersection of a four-way junction of P3-P6. P5 and P13 interact to form the kissing loop tertiary structure. The L13/P13 structure contains the anti-terminator or anti-sequester. In this configuration, the ribosome binding site (RBS) associates with the anti-RBS sequences. In this schematic (based on (Johnson Jr *et al*., 2012; Peselis & Serganov, 2012)), AdoCbl (red oval) is bound to the aptamer. **B**. Schematic showing the expression platform in transcriptional (left schematic) and translational (right schematic) riboswitches in the absence of cobalamin. The anti-terminator and anti-RBS domains associate with P1-derived anti-anti-terminator and anti-anti-RBS sequences, respectively. AUG – start codon.

There is considerable diversity among riboswitches because they have high ligand specificity. Subtle alterations in their structure can inactivate ligand-binding or completely change the type of ligand engaged. For example, among purine riboswitches, the guanine and adenine riboswitches are distinguished by only one nucleotide difference (Serganov *et al*., 2004). Similarly, cobalamin riboswitches can discriminate between AdoCbl, MeCbl and HOCbl (Johnson Jr *et al*., 2012; Perdrizet *et al*., 2012; Peselis & Serganov, 2012; Soulière *et al*., 2013; Polaski *et al*., 2017). AdoCbl riboswitches differ from MeCbl and HOCbl riboswitches by the presence of an extension of the P6 stem that forms a peripheral extension comprising P8–P11

helices (**Figure 1.5**), which are missing in the MeCbl and HOCbl riboswitches. The specificity of riboswitches implies that, potentially, each has a unique mechanism of action. Riboswitches, therefore, represent an underexplored frontier with potential to complement current knowledge in mycobacterial biology and possibly, provide novel targets for antimycobacterial therapy.

The basic mechanism of regulation by a riboswitch involves two mutually exclusive states: an "ON" state that allows gene expression and an "OFF" state that suppresses gene expression. In general, "ON" riboswitches detect toxic or overabundant metabolites and turn on the expression of genes that eliminate those metabolites (Baker *et al*., 2012; Furukawa *et al*., 2015). Conversely, "OFF" riboswitches turn genes off in response to metabolite availability and, typically, regulate biosynthesis or transport. Experimental evidence shows, however, that the precise process by which riboswitches regulate gene expression is not simply the flipping "ON" or "OFF" of a switch, but involves complex mechanics of ligand-induced signalling, speed of transcription, and the kinetics and thermodynamics of RNA folding. (Haller, Soulière & Micura, 2011).

1.9 Methionine biosynthesis in mycobacteria

Methionine is an essential amino acid required for initiating protein biosynthesis, DNA methylation and cysteine biosynthesis. There are at least three possible pathways for the synthesis of methionine in mycobacteria. In one pathway, *O*-acylated L-homoserine undergoes trans-sulphuration with L-cysteine to form L-cystathionine, which is then cleaved to L-homocysteine. L-homocysteine is subsequently methylated to produce L-methionine (**Figure 1.6A**). Alternatively, *O*-acetyl-L-homoserine reacts with hydrogen sulphide to produce L-homocysteine, bypassing L-cystathionine (**Figure 1.6A**). In another route, L-homoserine is first esterified by succinyl-CoA to form *O*-succinyl-L-homoserine, which can be converted directly to L-homocysteine or indirectly via L-cystathionine (**Figure 1.6A**). All the genes involved in mycobacterial methionine biosynthesis are known, except those required for the esterification of L-homoserine. The methylation of L-homocysteine is catalysed either by the cobalamin-independent MetE or MetH (5-methyltetrahydrofolate-homocysteine methyltransferase, EC 2.1.1.14), which requires cobalamin as a co-factor. Both

enzymes have a mutual requirement for zinc to activate homocysteine (Matthews & Goulding, 1997). In general, organisms that are capable of synthesizing or assimilating cobamides encode MetH, while those that cannot synthesize or salvage cobamides encode only MetE.

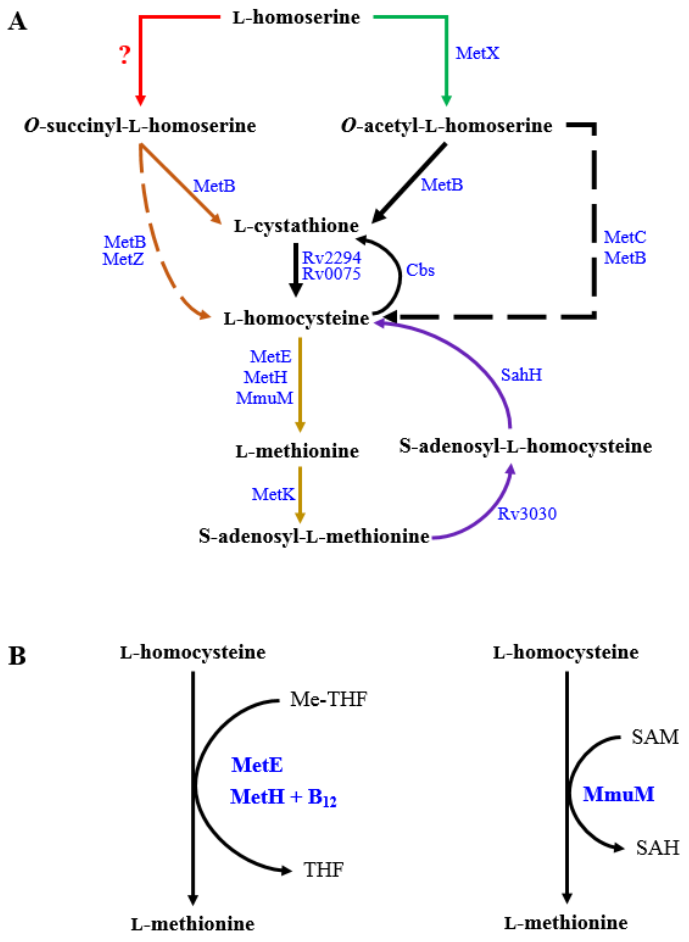


Figure 1.6: **Methionine biosynthesis and S-adenosyl methionine recycling in mycobacteria**. **A.** Possible routes to *de novo* methionine in mycobacteria. The first step is the esterification of L-homoserine by acetyl-CoA (green line) or by succinyl-CoA (red line) Trans-sulphuration of the *O*-acylated L-homoserine with L-cysteine (black solid lines) forms L-cystathionine which is reduced to L-homocysteine. *O*-acetyl-L-homoserine can undergo direct sulphuration to L-homocysteine (black dashed line). The enzyme catalysing the esterification of L-homoserine by succinyl-CoA is unknown. **B.** Methylation of homocysteine. The transfer of a methyl group from 5-methyltetrahydrofolate (Me-THF) to homocysteine is carried out by MetE or MetH. In pathogenic mycobacteria, MmuM is predicted to catalyse the transfer of a methyl group from SAM to L-homocysteine. SAH – *S*-adenosyl-L-homocysteine. Enzyme annotations are based on *M. tuberculosis* H37Rv.

Although the net reaction catalysed by MetE and MetH is essentially the same, comprising the transfer of a methyl group from 5-methyltetrahydrofolate (5-MTHF) to homocysteine, there are catalytic features unique to each of the two enzymes. For example, MetE uses the triglutamylated derivative of 5-MTHF, whereas MetH can accept monoglutamylated 5-MTHF (Gonzalez *et al*., 1992). Furthermore, while MetE catalyses the direct transfer of the methyl species from 5-MTHF, using a catalytic cysteine residue as an intermediate methyl acceptor, MetH first transfers the methyl group to Co(I)cobalamin (Gonzalez *et al*., 1992). Interestingly, the measured rate of turnover of MetH is reportedly 100-fold higher than that of MetE, implying that the direct methylation strategy employed by MetE is less efficient.

Protein structures of the mycobacterial methionine synthases have not been resolved, but useful deductions can be made by examining the crystal structures in *E. coli*, *T. maritima*, *N. crassa*, *C. albicans* and *A. thaliana* (Dixon *et al*., 1996; Ferrer *et al*., 2004; Pejchal & Ludwig, 2005; Fu *et al*., 2011; Ubhi *et al*., 2014; Wheatley, Ng & Kapoor, 2016). Sequence alignments revealed that unlike MetE, which shares homology with that of other bacteria, a MetH variant specific to actinobacteria is found in mycobacteria (Young, Comas & de Carvalho, 2015). Interestingly, a conserved hypothetical protein, Rv3015c, often annotated as a MetE homologue, is also found in *M. tuberculosis*. Rv3015c has about 44% sequence overlap with the conserved, catalytic C-terminal domain of MetE from *C. reinhardtii*, but the zinc ion-binding and homocysteine activation motifs are inactive, making it unlikely that this single-domain protein is functional (Warner *et al*., 2007).

A comparison of MetH and MetE in non-mycobacterial models reveals important structural differences between the two enzymes. For instance, MetH has four distinct domains, each performing separate catalytic functions (Goulding, Postigo & Matthews, 1997), whereas MetE has two domains linked by a catalytic groove (Ferrer *et al*., 2004; Pejchal & Ludwig, 2005). The four domains of MetH comprising the homocysteine, 5-MTHF, cobamide, and *S*-adenosylmethionine (SAM) binding and activation domains, form a modular structure orientated in the N- to C- terminal direction along the protein sequence (**Figure 1.7**) (Goulding, Postigo & Matthews, 1997). By contrast, the MetE structure is a dual $(\beta\alpha)_8$ barrel structure connected by an inter-domain linker, with N- and C-terminal domains that share substantial sequence

and structural homology (**Figure 1.7**) (Pejchal & Ludwig, 2005), possibly due to a primordial gene duplication event. The zinc binding domains of both enzymes occur in a conserved region but, in MetE, the binding site is in the C-terminal domain whereas in MetH this site is found in the N-terminal region (**Figure 1.7**). In the unbound state, the zinc ion is coordinated in a tetrahedral structure by one histidine, one glutamate and two cysteine residues in MetE. In contrast, this coordination is achieved by three cysteines and one asparagine in the unbound MetH. However, in the bound state, the zinc ligand undergoes inversion, during which the coordinating glutamate in MetE and asparagine in MetH are replaced by the sulphur of homocysteine (Koutmos *et al*., 2008; Ubhi *et al*., 2014; Ubhi & Robertus, 2015; Wheatley, Ng & Kapoor, 2016).

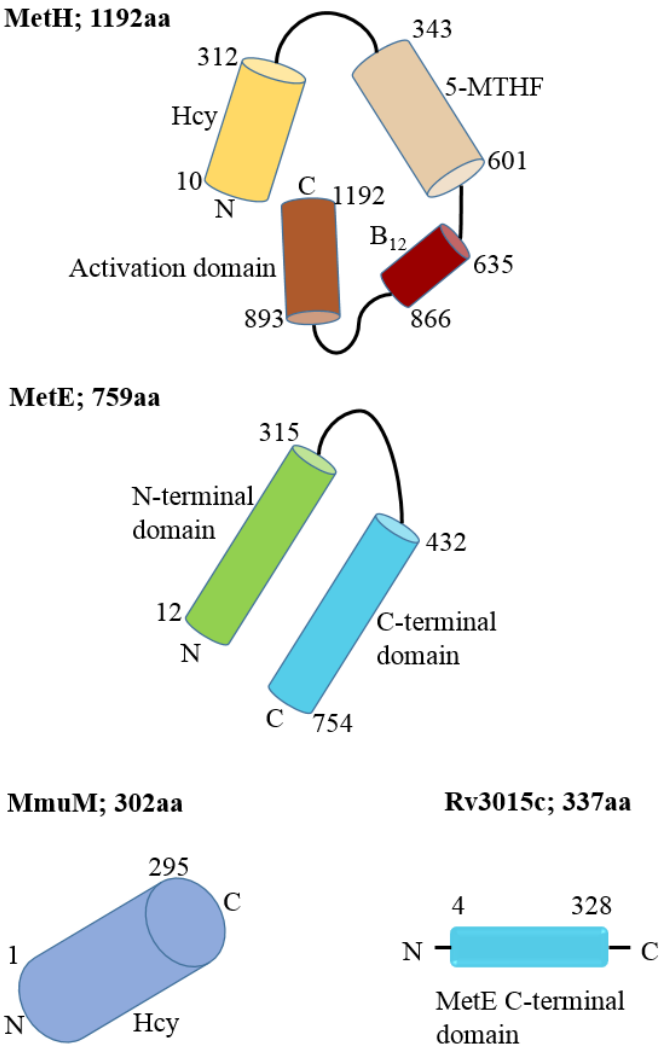


Figure 1.7: **Schematic of inferred structures of mycobacterial methionine synthases.** Protein sequences of *M. tuberculosis* MetH, MetE, MmuM and Rv3015c were scanned for conserved domains against the InterPro database (Jones *et al*., 2014) to find domain architectures. There are four domains in MetH, with distinct binding motifs for homocysteine (Hcy), the pterin group derived from 5-methyl tetrahydrofolate (5-MTHF), cobalamin (B_{12}) and SAM, which binds the C-terminal activation domain. MetE is organized into N- and C-terminal domains, while the putative methyltransferases MmuM and Rv3015c are single-domain proteins. Rv3015c shares homology with the C-terminal domain of MetE. Numbers above each module indicate the coordinates of the residues that define each domain.

In addition to MetH and MetE, several mycobacteria also encode a probable cobalamin-independent homocysteine S-methyltransferase, MmuM, which can catalyse the methylation of L-homocysteine using *S*-methylmethionine (SMM), a SAM derivative, as a methyl donor (Ying *et al*., 2015). MmuM is widespread among clinically relevant mycobacteria which encode this gene in their chromosomes or on plasmids; in contrast, *M. smegmatis* and the fish mycobacterium, *M. marinum*, lack

MmuM homologues (Ying *et al*., 2015). MmuM shares sequence and functional homology with the human betaine-homocysteine *S*-methyltransferase 2 (BHMT2; EC 2.1.1.10), which catalyses the conversion of betaine into dimethyl glycine and the methylation of homocysteine (Li *et al*., 2016). Recently, the overall structure of MmuM from *E. coli* was resolved, revealing some similarities with the homocysteine-binding domain of *T. maritima* MetH (Li *et al*., 2016). Like MetH, MmuM has an N-terminal zinc ion-binding domain of its $(\alpha\beta)_8$ TIM-barrel structure (**Figure 1.7**). Three cysteines and a tyrosine coordinate the zinc ion, and like MetE and MetH, a zinc inversion mechanism is predicted for binding with homocysteine. In *E. coli*, the absence of MetE and MetH can be compensated by MmuM, provided SAM is included in growth media (Thanbichler, Neuhierl & Böck, 1999). However, experimental evidence suggest that production of methionine via MmuM is inadequate to compensate for the lack of MetE and MetH activities in *M. tuberculosis* (Warner *et al*., 2007; Berney *et al*., 2015). Therefore, although MmuM provides a plausible way to recycle SAM and other metabolites of the methionine pathway, its role in mycobacterial metabolism remains unclear.

1.10 The intersection of cobalamin, methionine and folate metabolism

Cobalamin, methionine and folate metabolism share a unique intersection at the methylation of L-homocysteine by MetH. In the one-carbon pool by folate, 5-MTHF is generated by an irreversible conversion of 5,10-methylenetetrahydrofolate (Blanco, Coque & Martin, 1998). In the absence of cobalamin and, consequently, lack of MetH activity, cellular 5-MTHF accumulates, leading to a phenomenon called the "methyl folate trap." In humans, the methyl folate trap is associated with hyperhomocysteinemia and the onset of megaloblastic anaemia (Palmer *et al*., 2017). Originally thought to exist in eukaryotes only, the methyl folate trap was recently described in prokaryotes, including mycobacteria (Guzzo *et al*., 2016). When induced in the presence of drugs targeting the folate pathway, this phenomenon is lethal in the absence of cobalamin.

1.11 Aims and objectives

The overall goal of this thesis was to examine cobamide production, uptake and regulation of mycobacterial methionine metabolism using *M. smegmatis* as a proxy and where appropriate, making comparisons with pathogenic mycobacteria. After initiating the experimental work, it became clear that, while *M. smegmatis* was a useful model organism, a direct comparison with *M. tuberculosis* could not be made with respect to the acquisition of this co-factor. Unexpectedly, differences in cobamide-related metabolism between the two mycobacteria were revealed, as described in later chapters. These chapters have been organized to describe systematically the specific aims and objectives of this work, the research methodology, results and discussion. **Chapter 2** contains the details of the experimental methods and in **Chapter 3**, an evaluation of the genetic capacity for cobamide biosynthesis and transport in mycobacteria is presented. This is followed by an experimental validation of cobamide production and functionality of the riboswitch controlling *metE* in *M. smegmatis* in **Chapter 4**. **Chapter 5** describes how the interaction between cobalamin and the *metE* riboswitch underlies the essentiality of *metH* in *M. smegmatis*. The question of how the regulation of methionine biosynthesis by cobalamin mediate susceptibility of *M. smegmatis* to drugs targeting the folate pathway is also addressed. In **Chapter 6**, concluding remarks are made, situating this dissertation within a series of studies emanating from both the Molecular Mycobacterial Research Unit (MMRU), and other research groups, that address fundamental questions regarding cobamide-related metabolism and its inferred relevance to mycobacterial pathogenesis. It is expected that these investigations will yield key insights into the utility of targeting cobamide biosynthesis or cobamide-dependent regulation for anti-TB drug development.

2 Materials and Methods

2.1 Downloads of genomes and visualization of phylogenetic distances

A modified version of the procedure for detecting cobamide biosynthesis and salvage genes (Shelton *et al*., 2018) was followed. The names, metadata and unique identifiers for 1111 publicly available *Mycobacteriaceae* family genomes classified as either "finished" or "permanent draft" were downloaded from the Joint Genome Institute's Integrated Microbial Genomes with Expert Review (JGI/IMG ER, https://img.jgi.doe.gov/) database (Markowitz *et al*., 2012) (accessed July-Aug, 2018). The JGI/IMG ER database was selected because of its unique comparative genomics tools, including built-in annotation profiles for KEGG classifications and protein family databases such as Pfam (El-Gebali *et al*., 2018). The JGI/IMG ER database also offered user-friendly access to multiple reference datasets allowing comparative phylogenetic analyses and assessments of metabolic potential of individuals. The list of 1111 genomes was manually inspected to remove duplicates and a refined list of 1069 genomes was obtained (**Table S1, Supplementary information online**). The gene count of *Mycobacterium leprae Br4923* (gene count = 1654) was chosen as the minimum cut-off for genome size. Each unique genome name was considered a distinct strain. 16S-based phylogenetic distances between the genomes in the refined list were computed under the "Compare Genomes" functionality in JGI/IMG ER with BLAST percent identity set at 30+%. A precomputed Newick file for all genomes in the refined list was downloaded and displayed using FigTree v1.4.3 (http://tree.bio.ed.ac.uk/software/figtree). The tree was rooted on *Hoyosella sublava* as an outgroup species.

2.2 Detection of *de novo* cobamide biosynthesis and salvage genes in mycobacterial genomes

To find genes involved in the *de novo* cobamide biosynthesis and salvage pathways, the "Find Functions" tool in JGI/IMG ER was used to search for enzymes by the enzyme commission (EC) number (**Table S2, Supplementary information online**) (Ogata *et al*., 1999). The KEGG *Porphyrin and chlorophyll metabolism* and the KEGG *Riboflavin metabolism* pathways were used as references for cobamide and DMB

biosynthesis respectively. The genes were matched with genomes using the "function profile: functions vs genomes" functionality and gene hits were downloaded as a list of unique EC numbers, enzyme names and genome names. Output Excel sheets were combined into one master Excel file which was used as input for an R script that interpreted the absence or presence of genes to assess the intactness of *de novo* cobamide biosynthesis. To identify enzymes involved in the salvage pathway, the canonical *E. coli* corrinoid transporter complex BtuFCD (EC 3.6.3.33) (Cadieux *et al*., 2002) and the atypical *M. tuberculosis* cobamide transporter, which is orthologous to BacA (Gopinath, Venclovas, *et al*., 2013), were matched against the genomes using the "function profile: functions vs genomes" functionality. For convenience, the pathway was analysed in segments representing the synthesis of: (1) the 5-aminolevulinic acid (ALA) and uroporphyrinogen III biosynthesis (Uro' III), (2) aerobic and anaerobic corrin ring contraction, and (3) cobinamide-phosphate (Cbi-P) and DMB assembly. The *P. denitrificans* gene nomenclature (Blanche et al., 1995) was used.The prefix "*cob*" was used to designate genes or enzymes of the aerobic corrin ring synthesis pathway and "*cbi*" to designate those of the anaerobic corrin ring synthesis pathway.

2.3 Identification of cobamide biosynthetic genes in *M. smegmatis*

Genes encoding the cobamide biosynthetic and salvage enzymes as listed in **Table 3.2** of **Chapter 3** and *M. smegmatis* orthologs of the *M. tuberculosis* cobamide biosynthetic genes described in (Gopinath, Moosa, *et al*., 2013) were retrieved using the search function in the Mycobrowser portal (Kapopoulou, Lew & Cole, 2011).

2.4 Bacterial strains and growth conditions

The bacterial strains and plasmids used in this study are described in **Table 2.1**. Unless specified, *M. smegmatis* mc^2155 was grown in either Middlebrook (Difco) 7H9 supplemented with 10% Oleic Albumin Dextrose Catalase (OADC) (Becton Dickinson) and 0.05% Tween 80 or on Middlebrook (Difco) 7H10 supplemented with 10% OADC. When used in mycobacterial cultures, kanamycin (Km) was added at a final concentration of 25μg/ml and hygromycin (Hyg) at a final concentration of 50μg/ml. *E. coli* was grown in lysogeny broth (LB) or lysogeny agar (LA) with

50μg/ml Km or 200μg/ml Hyg, where appropriate. All cultures were incubated at 37°C unless indicated otherwise. Unless specified, growth curves were generated by seeding bacterial cells in clear round-bottom 96-well culture plates (Greiner Bio-One) and taking absorbance measurements every 1.5h, over a period of 30h, in a FLUOstar OPTIMA microplate reader (BMG Labtech).

Table 2.1: **Strains and plasmids used in this study**

Strain or plasmid	**Description**	**References**
Mycobacterium smegmatis		
mc^2155	High-frequency transformation mutant of *M. smegmatis* ATCC 607	(Snapper *et al*., 1990)
Δ*cobK*	*cobK* knockout in mc^2155	This study
Δ*cobK*Δ*metH*	*metH* knockout in Δ*cobK*	This study
Δ*metE*	*metE* knockout in mc^2155	Dawes, unpublished
Δ*metE*Δ*cobK::hyg*	*metE* knockout in Δ*cobK::hyg* background	Dawes, unpublished
metH KD	*M. smegmatis metH* knockdown strain	This study
Plasmids		
p2NIL	Suicide plasmid; Km^R	(Parish & Stoker, 2000)
pGOAL17	Cassette plasmid; Amp^R	(Parish & Stoker, 2000)
pGOAL19	Cassette plasmid; Amp^R	(Parish & Stoker, 2000)
p3875K	*M. smegmatis* Δ*cobK* vector; Km^R	This study
p3875K19	*M. smegmatis* Δ*cobK* vector containing *Pac*I cassette from pGOAL19; Km^R, Hyg^R, Suc^S	This study
p4185K	*M. smegmatis* Δ*metH* vector; Km^R	This study
p4185K19	*M. smegmatis* Δ*metH* vector containing *Pac*I cassette from pGOAL19; Km^R, Hyg^R, Suc^S	This study
p4185K17	*M. smegmatis* Δ*metH*::*hyg* vector containing *PacI* cassette from pGOAL19; Km^R, Hyg^R, Suc^S	This study
PLJR962	CRISPRi backbone for *M. smegmatis*	(Rock *et al*., 2017)
PLJR962_*metH*1	*M. smegmatis metH* knockdown construct with sgRNA1 oligo	This study
PLJR962_*metH*2	*M. smegmatis metH* knockdown construct with sgRNA2 oligo	This study
PLJR962_*metH*3	*M. smegmatis metH* knockdown construct with sgRNA3 oligo	This study
PLJR962_*metH*4	*M. smegmatis metH* knockdown construct with sgRNA4 oligo	This study

Strain or plasmid	Description	References
PLJR962_*metH*5	*M. smegmatis metH* knockdown construct with sgRNA5 oligo	This study
PLJR962_*metH*6	*M. smegmatis metH* knockdown construct with sgRNA6 oligo	This study
PLJR962_*metH*7	*M. smegmatis metH* knockdown construct with sgRNA7 oligo	This study
PLJR962_*metH*8	*M. smegmatis metH* knockdown construct with sgRNA8 oligo	This study
PLJR962_*metH*9	*M. smegmatis metH* knockdown construct with sgRNA9 oligo	This study
PLJR962_*metH*10	*M. smegmatis metH* knockdown construct with sgRNA10 oligo	This study
PLJR962_*metH*11	*M. smegmatis metH* knockdown construct with sgRNA11 oligo	This study
PLJR962_*metH*12	*M. smegmatis metH* knockdown construct with sgRNA12 oligo	This study
PLJR962_*metH*15	*M. smegmatis metH* knockdown construct with sgRNA15 oligo	This study
PLJR962_*mmpL3*	*M. smegmatis mmpL3* knockdown construct	This study

2.5 Cloning

The oligonucleotides (oligos) used for cloning and PCR are listed in **Table 2.2**. The p2NIL backbone (Addgene plasmid #20188; (Parish & Stoker, 2000)) was used to generate the suicide vector while pGOAL19 (Addgene plasmid #20190; (Parish & Stoker, 2000)) was used to generate the counter-selection cassette. An in-frame, unmarked deletion in *M. smegmatis cobK* (*MSMEG_3875*) was generated by amplifying a 912bp fragment (FR1) containing 40bp of the 5' end of *cobK* plus 872bp of upstream flanking sequence and a 923bp fragment (FR2) containing 107bp of the 3' end of *cobK* plus 816bp fragment of the downstream flanking sequence. The two fragments were joined in a three-way ligation reaction with p2NIL using *Asp718*I, *Bgl*II, and *Hin*dIII restriction, resulting in the p3875K vector carrying a deleted *cobK* allele of 120bp size. A counter-selection fragment carrying the *lacZ*, *hyg*, and *sacB* genes was excised from pGOAL19 and cloned at the *Pac*I site of p3875K to generate the suicide vector p3875K19 that was then electroporated into competent *M. smegmatis* cells.

To generate an in-frame, unmarked deletion in *M. smegmatis metH* (*MSMEG_4185*), an amplicon (FR1) containing the 1524bp of the 5' coding sequence of *metH* and another 1480bp amplicon (FR2) containing 354bp of the 3' end of *metH* plus the downstream flanking sequence were amplified by PCR. FR1 was subjected to restriction with *Asp*718I and *Bgl*II and FR2 was digested with *Hin*dIII and *Bgl*II and the two digested fragments were then joined in a three-way ligation reaction with p2NIL to produce the p4185K vector with an 1848bp truncated *metH* allele. A counter-selection fragment carrying the *lacZ*, *hyg*, and *sacB* genes was excised from pGOAL19 and cloned at the *Pac*I site of p4185K to generate the suicide vector p4185K19 that was used in electroporation. To generate the *hyg*-marked suicide vector p4185K17, a *hyg* cassette was cloned into the *Bgl*II site of p4185K before inserting the counter-selection cassette derived from pGOAL17 (Addgene plasmid #20189) (Parish & Stoker, 2000). Constructs were validated using restriction enzyme mapping and Sanger sequencing using the primers listed in **Table 2.3**.

Table 2.2: **Oligos used for cloning, screening and gene expression analysis using ddPCR**

Oligo ID	**5' → 3' sequence**	**Description**
*3875*F1	CTCAGA**AAGCTT**GAAAGGCGGCGATT	Forward primer for FR1 Δ*cobK*; ***Hin*dIII**
*3875*R1	GCAGCAGAACTCGC**AGATCT**ATCGTC	Reverse primer for FR1 Δ*cobK*; ***Bgl*II**
*3875*F2	TGGCGG**AGATCT**TGATCATGGTGGAC	Forward primer for FR2 Δ*cobK*; ***Bgl*II**
*3875*R2	GGTCCGAGCATGC**GGTACC**GTTCTA	Reverse primer for FR2 Δ*cobK*; ***Asp*718I**
4185F1	GAGACGTT**GGTACC**GAACAT	Forward primer for FR1 Δ*metH*; ***Asp*718I**
4185R1	GCGCCCGC**AGATCT**GCTT	Reverse primer for FR1 Δ*metH*; ***Bgl*II**
4185F2	GCACCG**AGATCT**GGGCGT	Forward primer for FR2 Δ*metH*; ***Bgl*II**
4185R2	GGTGTCG**AAGCTT**ACCGGA	Reverse primer for FR2 Δ*metH*; ***Hin*dIII**
*3875*_KO_F	GGTGCGTATCGGAGGATT	

Oligo ID	**5' → 3' sequence**	**Description**
*3875*_KO_R	GTTCGCCCTCGTAGTCAT	Internal primers Δ*cobK*
*3875*_F	CAACATCGTCCCCCAACTGA	Flanking primers Δ*cobK*
*3875*_R	CGACATCACGCTCGACAAAC	
4185-KO-F	GTTCCTGTTCCACGCCAT	Internal primers Δ*metH*
4185-KO-R	CTGTCAGCCACTTCTCCT	
4185_F	CAACATGGACGAGGGCATGA	Flanking primers Δ*metH*
4185_R	GACTGCGGGTGCGAGAAATA	
6638_F	GGTTCATCGCCTCGTGGAAT	Flanking primers Δ*metE*
6638_R	GTCAACTTGTCAGGGCTGCT	
TB_*metH*_F	CGACGGTGTTGTGCAGTT	Flanking primers *PPE37-metH* locus inn *M. tb*
TB_*metH*_R	CGTCGTTGGATTCGTGCT	
*3875*_ProbeF	GGGTGCACAGCGTCACCA	Primers generating 562bp probe for southern blot on Δ*cobK*
*3875*_ProbeR	CCGACATCACGCTCGACA	
rt_*metE*_Fwd	GGAGCGCAACGACATGGT	Sequence detection primers for *metE* gene expression by ddPCR
rt_*metE*_Rev	TCTCGGTCGCGAAGAAACC	
rt_*4185*_Fwd	TCGACGAGGATTCGCTGTTC	Sequence detection primers for *metH* gene expression by ddPCR
rt_*4185*_Rev	GGCCGATCCGGTGATGT	
rt_*sigA*_Fwd	GCCCGCACCATCCGTAT	Sequence detection primer for *sigA* gene expression by ddPCR
rt_*sigA*_Rev	ATACGGCCGAGCTTGTTGAT	Sequence detection primer for *sigA* gene expression by ddPCR
rt_*metE*_probe	TATTTCGCCGAACAGC	Taqman hydrolysis probe for *metE*
rt_*4185*_probe	CATCGGTGAGCGCAC	Taqman hydrolysis probe for *metH*
rt_*sigA*_probe	CCGGTGCACATGGT	Taqman hydrolysis probe for *SigA*

Table 2.3: **Primers used for Sanger sequencing**

Primer ID	**Sequence (5' → 3')**
MSMcobK_SeqKF1	CAAGGCATGACGGTCTAC
MSMcobK_SeqKF2	ACGCCAAGGTGATCGACA
MSMcobK_SeqKF3	CAACATCGTCCCCCAACT
MSMcobK_SeqKF4	GGGGTCGAAGAGCATGTT
MSMcobK_SeqKR1	CGCCGAGTACAGGTAACT
MSMcobK_SeqKR2	ATGATGGTGGTGATGGCG
MSMcobK_SeqKR3	GAGACCTGTTGGCGCGAT
MSMcobK_SeqKR4	CGACATCACGCTCGACAA
MSMcobK_SeqKR5	CGGTGAGGGAGCAGATTT
MSMmetHKO_SeqF1	CGATCGTTGTCAGAAGTA
MSMmetHKO_SeqR1	GAAACAGCCCACCGGATA
MSMmetHKO_SeqF2	GCAGGAGAGTGATACCGAT
MSMmetHKO_SeqR2	GTCGACTCCGAAAAGCTGT
MSMmetHKO_SeqF3	CCCGACGCTTTAGTCACA
MSMmetHKO_SeqR3	CAGTACTTCGTCGTCGGCA
MSMmetHKO_SeqF4	CTTGAACAACGTGGCCTT
MSMmetHKO_SeqR4	CAAGGTCGTGGAGGGCAAA
MSMmetHKO_SeqF5	CGATCCGGTGATGTTGGT
MSMmetHKO_SeqR5	GAATACAGCACCCCGGACA
MSMmetHKO_SeqF6	GTTGAGGGTGTCGAAGAT
MSMmetHKO_SeqF7	GGGTTGAGTGTTGTTCCA
MSMmetHKO_SeqR6	GGGTTCCGATTTAGTGCT

2.6 Electroporation and screening

To make competent cells for electroporation, 50ml *M. smegmatis* cultures were grown to exponential phase cells (OD_{600nm} of 0.6–0.8) and harvested by centrifugation at 4000 × *g* for 10min. Following this, cells were subjected to three successive washes in 40ml, 30ml, and 20ml ice-cold 10% glycerol. After the third wash, cell pellets were re-suspended in 1ml ice-cold 10% glycerol and used immediately or stored at -80°C until needed. Electroporation was done in a GenePulser Xcell™ electroporator (Bio-Rad) using the time constant settings (1200V; 5ms time constant; 1mm cuvette). A total of 100µl competent cells were incubated with 1–8µg DNA for 20min on ice prior to pulsing. Single crossover (SCO) transformants were selected with Km and Hyg on 7H10-OADC plates. As colonies became visible on the plates, 50µl of 2% (w/v) 5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-gal) was underlain in each plate for the

blue/white screening of SCOs. PCR was used to confirm the SCOs, which were then cultured in antibiotic-free 7H9-OADC, followed by serial dilution and plating on 7H10-OADC containing 2% (w/v) sucrose. For the Δ*cobK* strain, double crossover recombinants (DCOs) were screened by PCR and confirmed with Southern blotting. In addition to PCR, *M. smegmatis* Δ*metH* strains were validated by Sanger sequencing using the primers in **Table 2.3**.

2.7 Extraction of *M. smegmatis* genomic DNA

Genomic DNA was isolated from 5ml of cultures at OD_{600nm} = 1 using the cetyltrimethylammonium bromide (CTAB) method (van Helden *et al.*, 2001) as follows. The cell suspension was divided into 1ml aliquots, pelleted and re-suspended in 500µl 1x TE (10mM Tris, pH 8.0; 1mM EDTA, pH 8.0). To each aliquot, 50µl of 10mg/ml lysozyme was added, followed by an overnight incubation at 37°C. Next, the cells were chemically lysed with 70µl of 10% SDS and 7µl of 10mg/ml proteinase K, with shaking at 400 rpm in a thermomixer at 65°C for 2h. This was followed by the addition of 100µl of 5M NaCl and 80µl CTAB, with further shaking at 65°C for 10min. To harvest the DNA, 700µl of 24:1 (v/v) chloroform: isoamyl alcohol was added and the tubes left to stand for 10min standing at room temperature, followed by centrifuging at 14000 × *g* for 10min. The aqueous phase was mixed with 700µl of ice-cold isopropanol and frozen for 30min at -80°C to precipitate the DNA. The DNA precipitate was pelleted by centrifugation at 14000 × *g* for 10min, washed in 70% ethanol and dried in a vacuum.

2.8 Southern blotting

For Southern blotting, 2–3µg DNA was digested overnight with *Sty*I and separated on 1% agarose gel at 80V. The DNA was then transferred onto a Hydrobond™ N+ membrane (Amersham) by overnight capillary transfer. DNA was then fixed on the membrane by microwaving for 2min on high setting. Next, the membrane was hybridized at 42°C overnight with target-specific PCR-generated probes (**Table 2.2**). Probe labelling and detection was done using the ECL Direct Nucleic Acid Labelling and Detection Systems (Amersham). The target DNA fragments were visualized on hypersensitive X-ray films.

2.9 Extraction of cobamides

Wild-type and mutant *M. smegmatis* strains were cultured until stationary phase ($OD600_{nm}$ ≥2) in 50ml 7H9-OADC supplemented with 3μg/ml (23μM) cobalt chloride. Cells were harvested by centrifugation at 4000 × *g* for 10min at 4°C. Cell pellets were then re-suspended in 8ml of 50mM sodium acetate buffer, pH4.5, and stored at -80°C until needed. Once thawed, the cells were lysed by 5min of sonication using a microtip sonicator with the following settings: 30 amplitude; 15sec pulse; 15sec cooling. Next, 16μl of a solution of 100mM KCN was added to the lysed cells. Then, the extraction tube was tightly closed and incubated in a chemical fume hood for 30min at room temperature, followed by boiling at 90°C for 45min inside the hood. The tube was then cooled on ice briefly and centrifuged at 4°C at 4000 x *g* for 10min. The supernatant was filtered through a 0.22μm filter and loaded onto a Sep-Pak C18 Plus Light Cartridge (Waters), which had been washed with 5ml 75% (v/v) ethanol and conditioned with 10ml sterile water. Following loading of the supernatant, the cartridge was washed with 10ml of water and eluted with 75% ethanol, collecting about 15 drops. The eluent was analysed immediately by liquid chromatography tandem mass spectrometry (LC-MS/MS) or stored in -20°C until needed. Prior to LC-MS/MS analysis, the cell extract was centrifuged at 14000 × *g* for 10min on a bench-top centrifuge, followed by a chloroform purification step.

2.10 Analysis of cobamides by LC-MS/MS

Eluents were analysed using LC-MS/MS in a positive ionization mode. Quantification was accomplished using the following multiple reaction monitoring (MRM) parameters: (m/z, 678→359 and m/z, 678→147). Chromatographic separation was performed through a HPLC reverse phase column (Phenomenex Synergi^{TM} Polar-RP 100Å, 50 x 2mm (Separations)) using an Agilent 1200 Rapid Resolution HPLC system equipped with a binary pump, degasser, and auto sampler, coupled to an AB Sciex 4000 QTRAP hybrid triple quadrupole linear ion-trap spectrometer. Mobile phases were 0.1% formic acid in water (A) and 0.1% formic acid in acetonitrile (B), and the following gradients were run: 0-2min, 95% A; 2-4, 5% A; 4-6min, 95% A; 6-8min, 95% A. The flow rate was maintained at 400μl/min. The MS analysis was performed

on an AB Sciex 4000 QTRAP LC mass spectrometer using the following parameters: Curtain gas (25.00); IS (5500.00); Temperature (200.00°C); GS1 (80.00); GS2 (55.00); EP (12.0). Data processing was done using the SCIEX Analyst® software.

2.11 RNA extraction and cDNA preparation

Total RNA was extracted using the FastRNA® Pro Blue kit (MP Biomedicals) from cell pellets of exponentially growing (OD_{600nm} = 0.4) or stationary phase (OD_{600nm} = 1.5) mycobacterial cultures. To remove residual contaminating genomic DNA, the isolated RNA was digested with TURBO DNAse (Ambion) after which 0.5µg of the DNase-treated RNA was used as template for cDNA synthesis, using the High Capacity RNA to cDNA kit (Thermofisher).

2.12 Quantitative gene expression analysis by ddPCR

Droplet digital PCR (ddPCR) and data analysis was performed as described (Singh *et al*., 2015). Primers and minor groove binder (MGB) Taqman probes (**Table 2.2**) were designed using Primer Express 3.0 (Applied Biosystems). TaqMan MGB probes homologous to the target genes were labelled with 2'-chloro-7'-phenyl-1, 4-dichloro-6-carboxyfluorescein (VIC) whereas the reference gene, *sigA*, was labelled with 6-carboxyfluorescein (FAM), to allow duplexing. The ddPCR reaction mixtures (20µl) containing 1x ddPCR Supermix for Probes, no dUTP (Bio-Rad), 375nM primers, 250nM VIC- or FAM-labelled probes and 1-10ng of cDNA template were emulsified in an Automated Droplet Generator (AutoDG) (Bio-Rad), and partitioned into approximately 20000 droplets. PCR was performed in a QX100 Thermal Cycler (Bio-Rad) with the following cycling conditions: 1× (95°C for 10min); 40× (94°C for 30 sec, 60°C for 1min ; 2°C/s ramp rate); 1× (98°C for 10min ; 2°C/s ramp rate); 1× (4°C infinite hold; 2°C/s ramp rate). Following amplification, droplets were quantified using the QX200 droplet reader (BioRAD). Three technical replicates were run for each target gene for each biological replicate. The absolute quantities of gene copies were determined using the QuantaSoft Analysis Pro (Bio-Rad) software. Poisson statistics were applied to the fraction of endpoint population of positive droplets to analyse the fluorescence data from each well. The absolute transcript levels of target genes were computed in copies/µl and normalized to *sigA*. The ratios of normalized

transcript count in mutant strains relative to the wild-type strain was used to assess expression differences.

2.13 Whole genome sequencing, genome assembly and variant detection

Genomic DNA was extracted using the CTAB method (van Helden *et al*., 2001) as described above, from exponential phase cultures grown from single colonies. Genomic libraries prepared using the TruSeq Nano DNA (Illumina) sample preparation kit according to the manufacturer's instructions, were sequenced using a 150bp paired-end strategy on an Illumina HiSeq 4000 instrument. Trimmomatic v0.35 (Bolger, Lohse & Usadel, 2014) was used to remove adapters, leading or trailing bases with a quality score < 3, reads shorter than 36bp in length, and bases with an average quality score of < 15 based on a 4-base sliding window. BWA v0.7.12 (Li & Durbin, 2009, 2010) was then used to map paired-end reads to the *M. smegmatis* mc^2155 reference genome (CP000480.1) (Fleischmann *et al*., 2006). SAMtools v0.1.2 (Li *et al*., 2009) was used to call bases. Sites that had Phred scores lower than 20 or coverage below 10-fold were removed from further analysis. SNPeff v4.1 (Cingolani *et al*., 2012), using the *M. smegmatis* mc^2155 (uid57701) database, was used to annotate variant positions.

2.14 Cell fractionation for protein quantification

Triplicate cultures of *M. smegmatis* were grown to OD_{600nm} ~ 1.2 in 7H9-OADC with or without 10μg/ml CNCbl. Cell lysis, fractionation and filter-assisted sample preparation (FASP) was done as described (Hermann *et al.*, 2018). Cells were harvested by centrifugation at 3500 × *g* for 10min and washed three times with phosphate-buffered saline (PBS, pH 7.4) containing 1× cOmplete™ protease inhibitor (Roche). Cells were lysed by 3min of sonication with a microtip sonicator using the settings: 30 amplitude; 15sec pulse; 15sec cooling. The cell lysate was centrifuged for 10min at 8000 × *g* to collect the cytosolic fraction (supernatant) and membrane fraction (pellet). The pellet was washed with PBS, re-suspended with 100mM Tris (pH 8.0), 4% SDS and 100mM DTT, and boiled in a water bath at 95°C for 30min. Urea was then mixed with the boiled pellet to a final concentration of 8M, with rocking for 30min at room temperature. The supernatant fraction and the dissolved pellet were

filtered through a 0.22μm filter, snap-frozen in liquid nitrogen and stored at -80°C until needed. Protein concentration was determined by the RCDC assay (BioRAD) as per manufacturer's protocol.

2.15 Filter-assisted sample preparation

Tryptic peptides were generated using the FASP protocol as described (Hermann *et al.*, 2018). Briefly, sequence-grade Trypsin (NEB) was added at 1:100 enzyme to protein ratio and incubated for 18h at 37°C in a wet chamber. Trypsin was quenched with 0.1% formic acid. Tryptic peptides were eluted with three rounds of buffer exchange with ABC buffer and desalted with homemade STAGE-tips containing Empore Octadecyl C18 solid-phase extraction discs (Supelco). C18 discs were conditioned with three washes with solvent B (80% acetonitrile (ACN); 0.1% formic acid (FA)) and three washes with solvent A (2% ACN; 0.1% FA). 20μg peptides were then eluted with solvent C (60% ACN; 0.1% FA), dried in a vacuum and re-suspended with solvent A to a concentration of 200ng/μl.

2.16 Selected Reaction Monitoring (SRM) assay development and LCMS analysis

SRM assays were developed in Skyline (version 4.1) using a spectral library generated from previous discovery MS data (Hermann *et al.*, 2018) with a cut-off score of 0.9. Skyline was set up to select two peptides per input protein, with the highest picked MS1 intensity in the discovery data, and then the top 5 most intense fragment ions for each of those peptides. A transition list was then generated for the Thermo Scientific triple stage quadrupole (TSQ) Vantage mass spectrometer. Samples were separated using a Thermo Accella LC system on a 10cm monolithic C18 column (Phenomenex) with a 4.6mm ID with a mobile phase that comprised a mixture of solvent A (water + 0.1% formic acid) and solvent B (HPLC-grade acetonitrile + 0.1% formic acid). The method was 45min in total, and the flow rate was 300μl/min. The gradient program began with 3% B, followed by a gradient from 8%–45% B from 5–25min, then an increase to 80% B at 30min for a 5-minute wash, before returning to 3% B for the remainder of the method. The LC system was run in-line into a Thermo TSQ Vantage through a heated electrospray ionisation (HESI) source. The source voltage was

+3500V, with a capillary temperature of 300°C, a vaporiser temperature of 200°C, sheath gas of 30, and aux gas of 10.

2.17 Retention time calculation and method refinement and scheduling

To determine the retention time for each peptide, methods were generated for the TSQ Vantage with a maximum of 20 transitions monitored per method. Therefore, since the original list contained 5 transitions, a total of 8 unscheduled methods was generated, with a cycle time of 5sec to maximise the amount of signal, a collision gas pressure of 1.5 mTorr, a Q1 peak width (FWHM) of 0.7, and collision energies as determined by Skyline. These unscheduled methods were then run with consecutive 2µl injections of a reference sample to further refine the list of transitions and determine the retention time for each peptide. The reference sample was generated by pooling all samples. The unscheduled runs were analysed in Skyline to determine the retention times for each peptide. Any transitions with no intensity or those with background-level intensity, interference, or ambiguous signal were removed from the method, with a minimum of 3 transitions per peptide remaining in the final list. The spectral library was used to further refine the assays, with any peptides with a *dotp* score lower than 0.7 removed from the final list.

2.18 Gene silencing of *M. smegmatis metH* by CRISPRi

Transcriptional knockdown (KD) of target genes using the clustered randomly interspaced short palindromic repeats interference (CRISPRi) approach was carried out as described (Rock *et al*., 2017). Short guide RNAs (sgRNAs) were designed as described (de Wet *et al.*, 2018), and 13 sgRNA oligo pairs– top oligo (forward primer) and bottom oligo (reverse primer) – targeting *metH* were synthesised (**Table 2.4**). The top and bottom oligos were annealed to generate an sgRNA, which was then cloned into the PLJR962 plasmid (Rock *et al*., 2017) using *Bsm*BI restriction. The ligation reaction was done overnight in a 10µl reaction using T4 DNA ligase (NEB) and the entire ligation mix from each reaction was transformed into 50µl electrocompetent *E. coli* DH5α to make thirteen independent transformations. Transformants were selected on LB plates with 50µg/ml Km. For each transformation, plasmid DNA was extracted from single colonies and sequenced with primer 1834 (**Table 2.4**) to verify the *metH*

KD constructs. As a positive control, a *mmpL3* KD construct was also generated. Competent *M. smegmatis* cells were transformed by electroporation with 200ng of the *metH* KD or *mmpL3* KD constructs. Each independent electroporation was selected on 7H10-OADC containing 25μg/ml Km. To induce gene silencing by CRISPRi, 100ng/ml anhydrotetracycline (ATc) was included in the selection plates. For uninduced controls, plates did not include ATc.

Table 2.4: **Oligos used to generate sgRNAs targeting *M. smegmatis metH***

sgRNA	**Oligo**	**5' → 3' Sequence**
1	F1	GGGAAAGTACTGCGACTGCGGGTG
	R1	AAACCACCCGCAGTCGCAGTACTT
2	F2	GGGAGTCGTCACCGACGGCGTTGGC
	R2	AAACGCCAACGCCGTCGGTGACGAC
3	F3	GGGAGTCACCGACGGCGTTGGCCGG
	R3	AAACCCGGCCAACGCCGTCGGTGAC
4	F4	GGGAGCCGGGTTGTTGAGGATGTC
	R4	AAACGACATCCTCAACAACCCGGC
5	F5	GGGAGCCCTTCATCTCCCACGCGTT
	R5	AAACAACGCGTGGGAGATGAAGGGC
6	F6	GGGAGCTTCTTCTCGGCCTCGATGT
	R6	AAACACATCGAGGCCGAGAAGAAGC
7	F7	GGGAGCGGACTTCACGACCTGCGG
	R7	AAACCCGCAGGTCGTGAAGTCCGC
8	F8	GGGAGTGCGTGATGCGTTCGCGCA
	R8	AAACTGCGCGAACGCATCACGCAC
9	F9	GGGAGTCCAGGCCGGCCTTGATGGC
	R9	AAACGCCATCAAGGCCGGCCTGGAC
10	F10	GGGAGCCGGCCTTGATGGCGTGGAA
	R10	AAACTTCCACGCCATCAAGGCCGGC
11	F11	GGGAGCCTCCCGCACCGGGTTGTTGCC
	R11	AAACGGCAACAACCCGGTGCGGGAGGC
12	F12	GGGAACCGGGTTGTTGCCGCGGAA
	R12	AAACTTCCGCGGCAACAACCCGGT
15	F15	GGGAGCTGGGCGTGCCGGATCGAGTT
	R15	AAACAACTCGATCCGGCACGCCCAGC
mmpL3	L3F	GGGAACAGACTGGCTGCCCTCGTC
	L3R	AAACGACGAGGGCAGCCAGTCTGT
P1834		TTCCTGTGAAGAGCCATTGATAATG

2.19 Single cell microfluidics and time-lapse microscopy

To perform single-cell analyses using microfluidics, a 100µl suspension of 2.0 x 10^6 bacterial cells/ml was prepared and loaded on the four-chambered CellASIC ONIX B04A-03 microfluidic platform (Merck) kept at 37°C throughout the experiment. Cells were trapped with the following pressure and flow time settings: channel A8 at 13.8 kPa for 15 seconds; channel A6 at 27.6kPa for 15 seconds. Channel A6 was then rinsed at 6.9kPa for 30 seconds. Untrapped cells were washed out by flowing inlet solution at 34.5kPa for 5min Prior to trapping, cells at OD_{600nm} = 1 was diluted to OD_{600nm} = 0.1 in 10ml 7H9-OADC with or without 100ng/ml of ATc and incubated in a shaking incubator for 6h at 37°C. ATc-containing media was perfused continuously in channels loaded with cells that had been pre-exposed to ATc and for cells that were not pre-exposed to ATc, only 7H9-OADC was perfused. The experiment was run for 43h. Live-cell imaging was performed on a Zeiss AxioObserver using a 100X, 1.4NA Objective with Phase Contrast and Colibri.7 fluorescent illumination system. Images were captured every 15min using a Zeiss Axiocam 503 and processed using FIJI software (https://fiji.sc/).

2.20 Determination of SULFAs MIC_{90} using the microplate alamar blue assays (MABA)

M. smegmatis strains were grown to OD_{600nm} 0.2–0.5 and diluted 1:1000 in 7H9-OADC. Then, 50µl of this diluted culture was added to 50µl 7H9-OADC containing 2-fold serial dilutions of sulfamethoxazole (Sigma-Aldrich®), sulfathiazole (Sigma-Aldrich®) or sulfadiazine (Sigma-Aldrich®) in clear round-bottom 96-well culture plates (Greiner Bio-One). Where applicable, 10µg/ml CNCbl was included in each well. Plates were incubated overnight at 37°C, after which 10µl of 100µg/ml resazurin was added to each well. The plates were incubated for an additional 5h at 37°C before fluorescence intensity measurements were taken using a FLUOstar OPTIMA microplate reader (BMG Labtech). The data were exported and analysed with GraphPad Prism 5.

3 Phylogenomics analysis predicts that *de novo* cobamide biosynthesis in mycobacteria is biased toward environmental species

3.1 Overview

Environmental and pathogenic species of mycobacteria colonize vastly different niches and, as a result, these microorganisms might be expected to exhibit phenotypic and metabolic diversity. For example, *M. smegmatis*, a saprophyte frequently used as a model for the general study of mycobacteria (Barry, 2001; Reyrat & Kahn, 2001), is a fast grower with doubling times of approximately 3h. On the other hand, *M. tuberculosis* is an obligate human pathogen which doubles roughly every 24h in standard liquid medium. The causative agent of leprosy, *M. leprae*, has a notoriously long generation time of two weeks, and is impossible to grow in axenic culture. Instead, *M. leprae* proliferates in footpads of mice, armadillos and macaques (Shepard, 1960; Sasaki *et al*., 2001). Various annotated mycobacterial genomes are now available in several up-to-date, publicly accessible databases such as the Mycobrowser portal (Kapopoulou, Lew & Cole, 2011), the Kyoto Encyclopaedia of Genes and Genomes (KEGG) (Ogata *et al*., 1999), and the Joint Genome Institute's Integrated Microbial Genomes with Expert Review (JGI/IMG ER) (Markowitz *et al*., 2012). These repositories enable genomic comparisons of various mycobacterial species to define key phylogenetic and metabolic differences between pathogens and non-pathogens. As reviewed in **chapter 1**, cobamide production has been implicated in the evolution of mycobacterial pathogenesis, whereby specific deletions or inactivation of genes in the *de novo* cobamide biosynthetic pathway have been observed in some pathogenic strains (Orgeur & Brosch, 2018). Consequently, these organisms might have become dependent, at least partially, on their hosts for cobamides or their precursors. In this chapter, the genetic capacity of mycobacteria for cobamide biosynthesis and salvage is analysed and discussed.

3.2 Results

3.2.1 Mycobacterial genomes dataset and phylogenetic tree

Publicly available, complete (designated as "finished") or nearly complete ("permanent draft") genome assemblies belonging to the taxonomic family *Mycobacteriaceae* were surveyed to assess the genetic potential for cobamide biosynthesis and salvage. After filtering to remove duplications or entries with fewer genes than *M. leprae*, which is known to have the most reduced mycobacterial genome (Cole *et al*., 2001), 42 entries were discarded from the original 1111 genomes, resulting in a working dataset of 1069 genomes (**Table S1, Supplementary information online**). Information about ecosystem associated with the genomes was available for 872 records (81.6%), with human hosts contributing 837 (96%) of these (**Table 3.1**). These included isolates of human and animal adapted strains of the MTBC which were sequenced at different centres around the world, the *M. abscessus* complex, and opportunistic non-tuberculous mycobacteria such as *M. fortuitum*, *M. neonaurum*, *M. xenopi*, *M. kansasii*, *M. avium*, *M. marinum* and *M. smegmatis*. There were 179 genomes that did not have associated ecosystem information, 18 genomes of environmental organisms which were not associated with animal or human hosts, and 5 genomes from engineered microbiomes (**Table 3.1**). On average, the genomes of host-associated mycobacteria were smaller (average genome size = 4.8Mb) compared to those of environmental mycobacteria (average genome size = 5.9Mb) (**Table 3.1**), consistent with host adaptation. There were no genomes associated with mycobacteria that infect small mammals such as meerkats (*M. suricattae*), voles (*M. microti*) and dassies (Dassie bacilli). It is possible their genomes might have been among those in the unclassified category (average genome size = 4.7Mb) (**Table 3.1**). *M. leprae TN* had the smallest genome (3.3Mb) and encoded 2750 genes, of which 1086 were pseudogenes (**Table S1, Supplementary information online**).

Table 3.1: **Summary of mycobacterial genomes and ecosystem metadata**

Ecosystem/Host	Number of genomes	Average genome size (Mb)
Bison	1	4.6
Water buffalo	1	5.0
Asian Elephant	1	5.3
Oryx	1	4.8
Camel	2	4.7
Goat	2	4.8
Wild boar	2	5.2
Birds	4	4.8
Deer	4	4.8
Sheep	4	4.7
Engineered microbiome	5	5.7
Cattle	9	4.1
Environmental	18	5.9
Unclassified	179	4.7
Human	836	4.5

The JGI/IMG ER database provided a functionality for computing the phylogenetic distances between genomes based on 16S rRNA sequencing (Markowitz *et al*., 2012). The tree was rooted on *H. sublava*, once grouped with the *Mycobacteriaceae* family but recently reclassified into the genus Hoyosella (Hamada *et al*., 2016). Two primary clusters of genomes were observed which, on inspection, revealed a dichotomy between host-associated genomes and environmental mycobacteria (**Figure 3.1**). Although the available data were insufficient to infer ancestry, other investigators have demonstrated that environmental species are ancestral to host-associated mycobacteria, an evolution which accompanied a transition to pathogenicity (Supply *et al*., 2013).

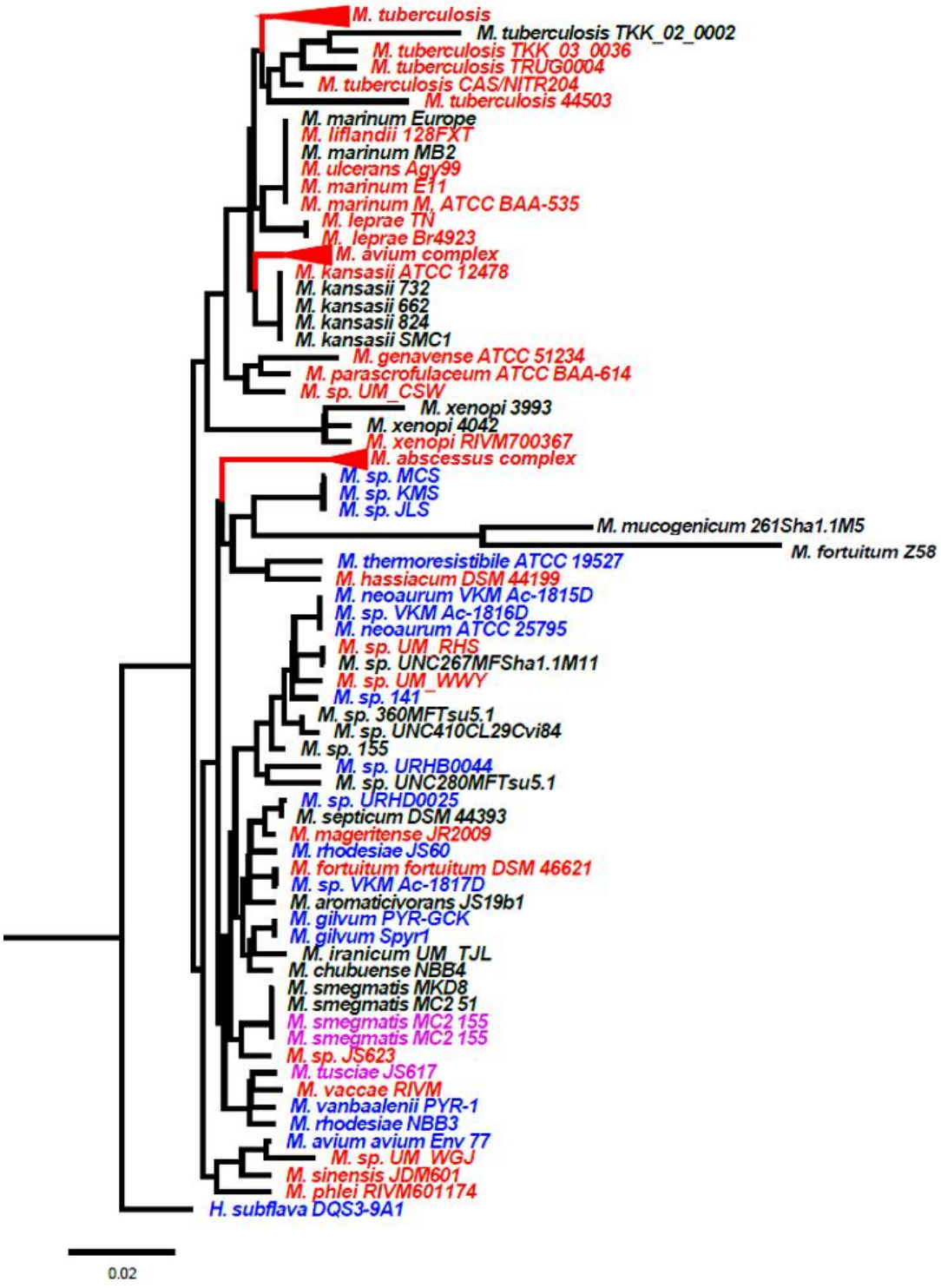


Figure 3.1: **Phylogenetic distance tree of mycobacterial genomes based on 16S rRNA alignment**. Genomes for which ecosystem and host metadata were available are highlighted in colour while those lacking these metadata are shown in black. Red indicates host-associated organisms, blue indicates environmental species and purple are genomes of engineered microbiomes. Some genomes of *M. tuberculosis* clinical isolates and of the *M. avium* and *M. abscessus* complexes are collapsed by tips (triangles) for easy visualization. *H. sublava* is the outgroup species.

3.2.2 Mycobacterial genomes encode variable capacity for corrin ring biosynthesis via the aerobic pathway

To determine the completeness of *de novo* cobamide biosynthesis in mycobacterial genomes, the pathway was segmented into three parts for analysis (**Table 3.2**) The first part grouped together enzymes involved in the synthesis of the tetrapyrrole precursors ALA and Uro' III. The second segment comprised enzymes catalysing the ring contraction steps, which were subdivided into aerobic and anaerobic pathways, and the last part grouped enzymes involved in the synthesis of Cbi-P and DMB (**Table 3.2**). More than 98% of mycobacterial genomes encoded at least one copy of the genes involved in the synthesis of ALA and Uro' III (**Table 3.2**). Except for *cobF* and *cobB*, which were found in 18% and 78% of the genomes, respectively, the genes involved in the aerobic corrin ring synthesis pathway were encoded by more than 98% of the mycobacterial genomes (**Table 3.2**). However, there were no hits for genes encoding the anaerobic corrin ring biosynthetic pathway except *cbiF*, *cbiJ*, and *cbiC*, which were contained in 99% of the genomes and *cbiA*, found in 78% of the genomes (**Table 3.2**). The annotations of these genes in representative mycobacteria were queried in the KEGG database (Ogata *et al*., 1999), and were found to correspond to CobM (CbiF), CobK (CbiJ), CobH (CbiC), and CobB (CbiA). Amino acid sequences of these "Cob" enzymes were then analysed for protein domain architecture with Pfam (El-Gebali *et al*., 2018). This analysis revealed "Cbi" family motifs in the "Cob" enzymes, which explained their being detected in the anaerobic corrin ring synthesis step.

Table3.2: **Distribution of *de novo* cobamide biosynthetic and salvage pathway enzymes in mycobacteria**.

Symbol	EC#	Synthesis step	% Genomes
GltX	EC 6.1.1.17	ALA and Uro' III synthesis	100.00%
ProS	EC 6.1.1.15	ALA and Uro' III synthesis	100.00%
HemA	EC 1.2.1.70	ALA and Uro' III synthesis	100.00%
HemL	EC 5.4.3.8	ALA and Uro' III synthesis	99.00%
HemB	EC 4.2.1.24	ALA and Uro' III synthesis	100.00%
HemC	EC 2.5.1.61	ALA and Uro' III synthesis	100.00%
HemD	EC 4.2.1.75	ALA and Uro' III synthesis	100.00%
CobA_CysG	EC 2.1.1.107	ALA and Uro' III synthesis	100.00%
CysG	EC 1.3.1.76	ALA and Uro' III synthesis	98.00%
CobI	EC 2.1.1.130	Corrin ring synthesis (aerobic)	99.00%

Symbol	EC#	Synthesis step	% Genomes
CobG	EC 1.14.13.83	Corrin ring synthesis (aerobic)	99.00%
CobJ	EC 2.1.1.131	Corrin ring synthesis (aerobic)	99.00%
CobM	EC 2.1.1.133	Corrin ring synthesis (aerobic)	99.00%
CobF	EC 2.1.1.152	Corrin ring synthesis (aerobic)	18.00%
CobK	EC 1.3.1.54	Corrin ring synthesis (aerobic)	99.00%
CobL	EC 2.1.1.132	Corrin ring synthesis (aerobic)	99.00%
CobH	EC 5.4.99.61	Corrin ring synthesis (aerobic)	98.00%
CobB	EC 6.3.5.9	Corrin ring synthesis (aerobic)	78.00%
CobNST	EC 6.6.1.2	Corrin ring synthesis (aerobic)	99.00%
CbiK_CbiX	EC 4.99.1.3	Corrin ring synthesis (anaerobic)	0.00%
CbiL	EC 2.1.1.151	Corrin ring synthesis (anaerobic)	0.00%
CbiH	EC 2.1.1.272	Corrin ring synthesis (anaerobic)	0.00%
CbiF	EC 2.1.1.271	Corrin ring synthesis (anaerobic)	99.00%
CbiG	EC 3.7.1.12	Corrin ring synthesis (anaerobic)	0.00%
CbiD	EC 2.1.1.195	Corrin ring synthesis (anaerobic)	0.00%
CbiJ	EC 1.3.1.106	Corrin ring synthesis (anaerobic)	99.00%
CbiT	EC 2.1.1.196	Corrin ring synthesis (anaerobic)	0.00%
CbiE	EC 2.1.1.289	Corrin ring synthesis (anaerobic)	0.00%
CbiC	EC 5.4.99.60	Corrin ring synthesis (anaerobic)	98.00%
CbiA	EC 6.3.5.11	Corrin ring synthesis (anaerobic)	78.00%
CobR	EC 1.16.8.1	Cbi-P synthesis	0.00%
CobO_PduO	EC 2.5.1.17	Cbi-P synthesis	100.00%
CobQ	EC 6.3.5.10	Cbi-P synthesis	99.00%
CobC_CobD_protein α	EC 6.3.1.10	Cbi-P synthesis	99.00%
CbiZ	EC 3.5.1.90	Cbi-P synthesis	0.00%
CobU_CobP	EC 2.7.1.156	Cbi-P synthesis	99.00%
CobU_CobP	EC 2.7.7.62	Cbi-P synthesis	99.00%
RibF	EC 2.7.1.26	DMB synthesis	99.00%
Fre	EC 1.5.1.38	DMB synthesis	0.03%
BluB	EC 1.13.11.79	DMB synthesis	99.00%
CobU_CobT	EC 2.4.2.21	DMB attachment	99.00%
CobC	EC 3.1.3.73	DMB attachment	0.00%
CobS_CobV	EC 2.7.8.26	DMB attachment	97.00%
BtuFCD	EC 3.6.3.33	Corrinoid uptake	0.00%
BacA	not assigned	Corrinoid uptake	99.00%

The variability in the intactness of the corrin ring synthesis step among the 1069 mycobacterial genomes was represented graphically on a binary heat map generated

in “R,” where functionality was inferred as the presence of at least one gene copy for each enzyme-catalysed step (**Figure 3.2**). This functional profile revealed that most genomes that lacked *cobF* belonged to host-associated mycobacteria. Interestingly, a large proportion of the genomes missing *cobF* also lacked *cobB* (**Figure 3.2**). The genomes of *M. leprae* (genome size = 3.3Mb) and *M. mucogenicum* (genome size = 3.4Mb) were missing all corrin ring biosynthesis enzymes (**Figure 3.2**). There was a high degree of variability in the presence of genes encoding the corrin ring synthesis step in genomes belonging to the *M. avium* complex (MAC) (**Figure 3.2**). A functional heat map for the MAC showed that this variability prevailed throughout the *de novo* cobamide biosynthesis pathway (**Figure 3.3**).

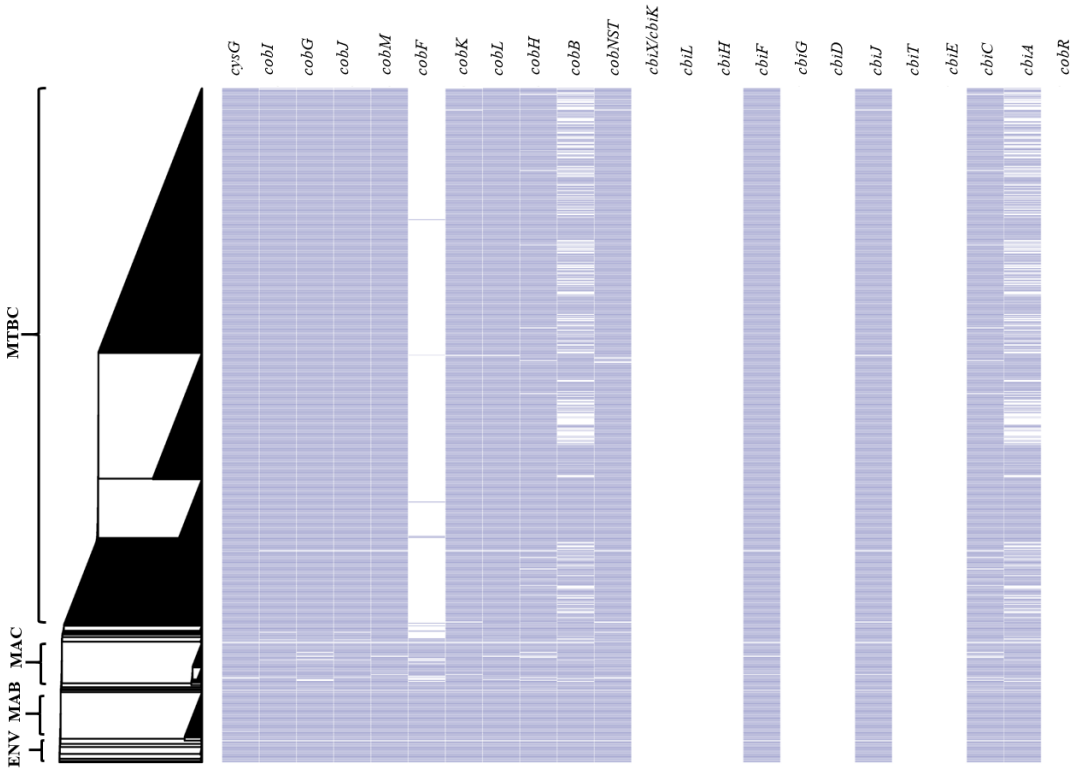


Figure 3.2: **A binary heat map showing the distribution of genes involved in corrin ring synthesis step of the *de novo* cobamide pathway in mycobacteria**. The presence or absence of at least one gene copy in the corrin ring synthesis step was assigned the value 1 or 0, respectively, and clustered by species in "R." Hits are indicated in blue while the absence of genes is displayed in white. The aerobic pathway was intact for environmental species (ENV) but species of the MTBC were missing *cobF*. Some of the MTBC genomes also lacked *cobB*. Anaerobic pathway genes were not found but cobM, *cobK*, *cobH* and *cobB* were detected as hits for *cbiF*, *cbiJ*, *cbiC* and *cbiA*, respectively. A high variability in corrin ring coding capacity was observed in the *M. avium* complex (MAC). MAB – *M. abscessus*.

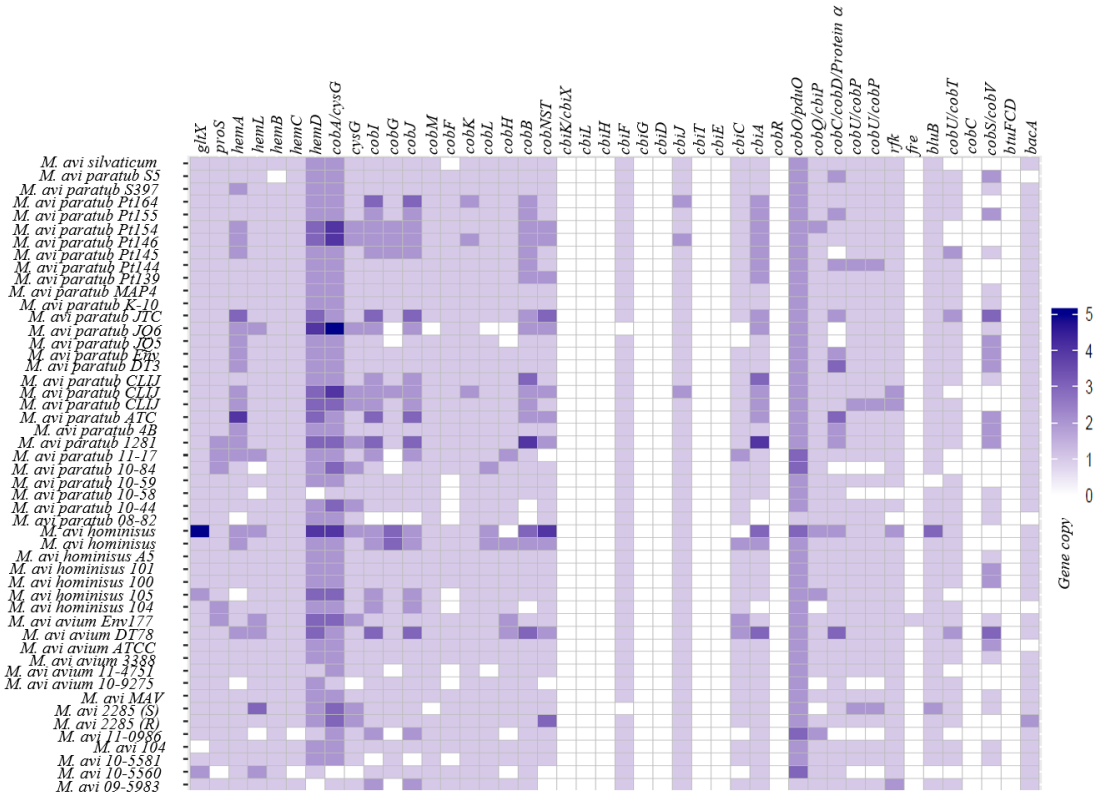


Figure 3.3: **Heat map showing the distribution of *de novo* cobamide biosynthesis genes in the *M. avium* complex genomes**. A matrix of copy numbers for genes involved in *de novo* cobamide biosynthesis and corrinoid salvage in the MAC genomes was displayed graphically in "R." The MAC genomes displayed a high degree of variation both in the presence of at least one gene copy per enzyme step, and in the number of gene copies per step.

3.2.3 *De novo* Cbi-P and DMB synthesis steps are conserved in mycobacteria

The step immediately following corrin ring synthesis is the reduction of cob(II)yrinate a,c diamide by CobR, a predicted cob(II)yrinate a,c diamide reductase (Lawrence *et al.*, 2008). In this analysis, no homologues of the gene encoding this enzyme were found in any genome (**Table 3.2**). However, BluB, which also catalyses the formation of DMB, has been proposed as an alternative cob(II)yrinate a,c diamide reductase based on sequence homology to CobR (Gopinath, Moosa, *et al.*, 2013). The BluB enzyme was present in more than 99% of the genomes analysed (**Table 3.2**). The enzymes which catalyse the subsequent reactions leading up to Cbi-P were also found in more than 99% of the genomes (**Table 3.2**). The gene encoding CobO (EC 2.5.1.17), involved in the adenosylation of the cobalt ion to form adenosylcobyrinate a,c-diamide were also found in 99% of the genomes in this analysis (**Table 3.2**). Moreover, genes

encoding aminopropanol biosynthesis enzymes CobC/CobD (EC 6.3.1.10), which, together with the as-yet-unidentified protein α, form an enzyme complex that catalayses the attachment of aminopropanol to adenosylcobyrinate, were also detected in 99% of genomes analysed (**Table 3.2**). Genes encoding CobU/CobT (EC 2.4.2.21) and CobS/CobV (EC 2.7.8.26), involved in the attachment of DMB to Cbi-P were also found in 97% of the genomes in this analysis (**Table 3.2**). However, no homologue of the gene encoding CobC (EC 3.1.3.73), an adenosylcobalamin 5'-phosphatase involved in the final step of adenosylcobalamin biosynthesis, was detected in any genome (**Table 3.2**). In addition to BluB, the riboflavin kinase RibF, and a flavin mononucleotide (FMN) reductase are also required for DMB synthesis in prokaryotes that use the aerobic pathway. More than 99% of the genomes encoded RibF but the FMN reductase was found in only 31 genomes (0.03%) (**Table 3.2**), of which 28 were genomes belonging to environmental organisms and 3 were genomes belonging to host-associated mycobacteria (*M. xenopi*; 2 genomes, and *M. avium*; 1 genome). Apart from participating in the *de novo* cobamide biosynthesis pathway, CobO and PduO are also predicted to be involved in the salvage of cobamides (Gopinath, Venclovas, *et al*., 2013). However, previous genetic screens conducted in the MMRU identified an atypical transporter, BacA, which was consistent with the fact that no mycobacterial genomes encoded homologues of the canonical BtuFCD transporter complex found in Gram negative bacteria (Gopinath, Venclovas, *et al*., 2013). In the present analysis, none of the genomes encoded the BtuFCD transporter complex but more than 99% of the genomes encoded at least one copy of BacA or CobO/PduO enzymes (**Table 3.2**).

3.3 Discussion

Cobamides are required by cells in all domains of life (Nielsen *et al*., 2012) yet *de novo* biosynthesis of this co-factor is restricted to some bacteria and archaea (Zhang *et al*., 2009; Doxey *et al*., 2015). In this chapter, a comparative analysis of over 1000 mycobacterial genomes from the publicly accessible JGI/IMG ER database (Markowitz *et al*., 2012) is presented. Phylogenetic clustering based on 16S alignments separated the genomes into two primary groups of environmental and host-associated mycobacteria. This analysis revealed that environmental mycobacteria encode a complete pathway for *de novo* cobamide biosynthesis via the aerobic route while pathogenic strains lack at least one enzyme in the pathway, CobF being the most

notably absent among the MTBC strains. The activity of CobF is presumed to have been lost during evolution of MTBC from a non-pathogenic environmental ancestor (Orgeur & Brosch, 2018). The detection of CobO/PduO enzyme classes and BacA in more than 99% the genomes further suggested that mycobacteria have a conserved genetic capacity to bypass *de novo* biosynthesis if exogenous cobamides available. Interestingly, genomes of the *M. avium* complex, which cause opportunistic diseases in humans, showed highly variable genetic capacity for *de novo* cobamide synthesis. This was not surprising, however, given that high frequency of genetic variability is associated with this complex (Rindi & Garzelli, 2014; Imperiale *et al*., 2017). The overall preservation of genes involved in either *de novo* cobamide biosynthesis or salvage among mycobacteria signal the indispensability of this co-factor in these organisms irrespective of their habitat and pathogenicity.

In *M. leprae*, only those genes related to the synthesis of the general tetrapyrrole precursors appeared intact. However, it was also observed that the draft genome of *M. mucogenicum*, a rapidly growing NTM associated with opportunistic infections related to contaminated water systems and hospital equipment (Springer *et al*., 1995), was also unexpectedly small. However, recent whole-genome resequencing of *M. mucogenicum* estimated that the number of genes it encoded was 6003 (Asmar *et al*., 2015) as opposed to 3125 reported in the JGI/IMG ER database, implying that occasional errors may be found with genome drafts in this database, necessitating manual curation of data.

Although the absence of CobF is assumed to have disrupted *de novo* cobamide production, a putative replacement for this enzyme (Rv2067c) has been suggested, implying a possibility of recovering cobamide from *M. tuberculosis* (Gopinath, Moosa, *et al*., 2013). Interestingly, most organisms that lacked CobF also lacked CobB, which catalyses the amidation of the corrin ring prior to the cobalt chelation step. In *M. leprae*, *cobB* is a pseudogene, and in the analysis conducted here, it appeared as "missing." It is likely, therefore, that this was also the case in those genomes where *cobB* was marked as absent, instead of its being completely "missing." Ongoing studies in the MMRU have failed to provide direct proof of cobamide production in *M. tuberculosis* (Dr. Gabriel Mashabela, personal communication), despite the presence of putative *cobF* substitutes, encoded by *Rv2067c* and *Rv039*1

(Gopinath, Moosa, *et al*., 2013). It is unclear how loss-of-function CobB mutations alone (assuming putative replacement enzymes for CobF are present in some mycobacteria) would affect cobamide biosynthesis in those organisms.

The natural loss of the *cobB* gene has not been reported in the literature for any other organism. Recently, however, relatively frequent frameshift mutations affecting the 3' region of *cobB* were discovered in clinical isolates of *M. tuberculosis*, a finding that suggests that this gene might also be involved in adapting to pressures within the host (Minias, Minias & Dziadek, 2018). In transposon mutagenesis screens, it has been shown that the loss of cobamide biosynthesis genes including *cobB* is associated with decreased production of PDIM, a mycobacterial lipid associated with virulence (Russell *et al*., 2010; Lee *et al*., 2013; Lovewell, Sassetti & VanderVen, 2016) and with increased fitness in *Acanthamoeba* (Weerdenburg *et al*., 2015; Boot *et al*., 2016). In light of these reports, it is tempting to speculate that the seemingly parallel loss of *cobF* and *cobB* represents an ongoing evolutionary event in pathogenic mycobacteria, in agreement with the theory that the evolution of mycobacterial pathogenesis is associated with genome decay (Gagneux, 2018; Orgeur & Brosch, 2018). Minias and colleagues also found that several other *de novo* cobamide biosynthesis genes in *M. tuberculosis* clinical isolates appear to be undergoing purifying selection (Minias, Minias & Dziadek, 2018). While their findings pointed to the possibility of cobamide production in *M. tuberculosis*, the lack of experimental evidence of cobalamin production in this organism suggests that the negative selection pressure acting on these genes might be an adaptation to host-induced stressors, antimycobacterial treatment or a combination of these pressures rather than in response to the limited access to cobamides in the host.

Among the mycobacterial genomes studied, none expressed any anaerobic "Cbi" enzymes, which was expected as mycobacteria typically inhabit oxygenated environments. However, four "Cob" enzymes (CobB/H/K/M) were detected as "Cbi" enzymes (CbiA/C/J/F), catalysing the corresponding corrin ring synthesis steps in the anaerobic pathway. This observation suggested possible enzymatic dexterity in this part of the cobamide biosynthetic pathway, whereby reaction steps of the aerobic and anaerobic routes can be catalysed by enzymes of either pathway. This view is supported by observations that *P. aeruginosa*, an aerobic bacterium which lacks an

identifiable CobF but expresses the corresponding enzymes CbiG and CbiD, is capable of *de novo* cobamide production (Crespo, Blanco-Cabra & Torrents, 2018). The discovery of Rv0259c as a possible homologue of the sirohydrochlorin cobaltochelatase CbiX/CbiK in *M. tuberculosis* implies that mycobacteria might also be able, at least partially, to use the anaerobic pathway (Gopinath, Moosa, *et al*., 2013).

In a previous study in the MMRU, two non-homologous *de novo* cobamide biosynthesis pathway enzymes, CobO (a homologue of the Salmonella CobA, the most extensively studied adenosyltransferase (Escalante-Semerena, Suh & Roth, 1990)) and PduO (a homologue of the Salmonella PduO (Johnson *et al.*, 2001)), were implicated in corrinoid assimilation, although only the latter was shown to be required *in vitro* for the assimilation of dicyanocobinamide $(CN)_2$Cbi (Gopinath, Venclovas, *et al*., 2013). There is ample experimental evidence (reviewed in-depth in (Mattes *et al.*, 2017)) that the Salmonella CobA is also involved in *de novo* synthesis of the corrin ring and the nucleotide loop. Given that CobO was widespread in mycobacterial genomes analysed in this study it would not be surprising that in mycobacteria, this protein functions in the same manner as the Salmonella CobA. Other bacteria and some archaea encode an adenosylcobinamide hydrolase, named CbiZ (EC 3.5.1.90), that is used to remodel adenosylcobamide by hydrolyzing the amide bond between aminopropanol and the ring, yielding adenosylcobyrinate, which can then be converted to a cobamide that the bacterium can use through the incorporation of a base other than DMB (Woodson & Escalante-Semerena, 2004). A homologue of this enzyme was not found in any mycobacterial genome analysed in this study (**Table 3.2**), hinting that this remodelling might not exist and that the final product in the pathway is an adenosylated cobamide.

The genes involved in Cbi-P synthesis, and in the assembly and attachment of DMB were highly conserved in the mycobacterial genomes used in this analysis. The gene encoding BluB, which catalyses the oxygen-dependent conversion of reduced FMN ($FMNH_2$) to DMB (Campbell *et al*., 2006; Gray & Escalante-Semerena, 2007), was widespread in mycobacterial genomes. $FMNH_2$ is the product of NADPH-dependent FMN reductase, but the enzyme catalysing this reaction was detected in only a few genomes of environmental organisms. Recently, one research group presented experimental evidence that anaerobic bacteria such as *Eubacterium limosum* encode *bzaABCDE* genes required for a four-step reaction leading to DMB, starting with 5-

aminoimidazole ribotide (AIR), a precursor for thiamine (vitamin B_6) biosynthesis (Hazra *et al*., 2015). In their paper, these authors hypothesized that DMB could be assembled anaerobically using methionine, glycine, formic acid and erythrose as building blocks (Renz, 1970). It is unknown, however, whether mycobacteria encode the enzymes that would be involved in recycling these metabolites. Subsequent comparative genomics by the same investigators did not find identifiable homologs of any *bza* genes in mycobacteria, suggesting that anaerobic DMB biosynthesis in mycobacteria is unlikely (Shelton *et al*., 2018).

The question of how mycobacteria accomplish cobamide transport continues to be a challenging puzzle. Several lines of indirect evidence suggest that, although mycobacteria lack the classical bacterial cobamide transporter complex BtuFCD, they can salvage cobamides using the BacA, an efflux protein implicated in drug resistance (Warner *et al*., 2007; Savvi *et al*., 2008; Gopinath, Venclovas, *et al*., 2013; Guzzo *et al*., 2016; Lawrence *et al*., 2018). The fact that the BtuFCD complex was absent in mycobacteria implies that the canonical mechanism of cobamide transport found in Gram negative bacteria (Lewinson *et al*., 2010) might not be at play in mycobacteria. The reason behind this could be that, unlike Gram negative bacteria, mycobacteria are characterized by a highly complex cell envelope (Niederweis *et al*., 2010), which provides a unique barrier to small molecule transport in these organisms. The current model of mycobacterial cobamide transport situates BacA in the cytoplasmic membrane (Gopinath, Venclovas, *et al*., 2013). Therefore, the array of proteins that would deliver cobamide through the mycobacterial outer cell membrane and periplasmic space remain unknown. In summary, mycobacteria encode a variable capacity for cobamide biosynthesis. Moreover, there is a strong impetus to develop mycobacteria as distinct models these organisms as they do not conform to classic models of cobamide biosynthesis and metabolism.

4 Production of cobamides by *M. smegmatis* strain mc^2 155

4.1 Overview

Cobamides are fascinating co-factors that have a rich scientific history marked by intensive efforts to characterise their biosynthesis. Because of their complex chemical structure, sophisticated machinery is required to shuttle cobamides between cellular microenvironments or across the interfaces of pathogens or symbionts and their hosts (Croft *et al*., 2005; Nielsen *et al*., 2012; Fang, Kang & Zhang, 2017). Despite the fact that cells in all domains of life require this co-factor, only a small group of prokaryotes can synthesise cobamides via an energetically-expensive process requiring nearly 30 enzyme-catalysed steps (Smith, Warren & Refsum, 2018). The catalytic efficiencies of cobamide-dependent enzymes seemingly subsidize the costs of *de novo* production, implying there is strong evolutionary pressure to keep these enzymes in all living systems. Only plants appear to have dispensed with their dependence on cobamides, although recent studies reveal that certain herbs can accumulate exogenous co-factor (Lawrence *et al*., 2018).

In mycobacteria, cobamide-dependent enzymes catalyse the biosynthesis of methionine and deoxyribonucleotides and the catabolism of propionate and ethanolamine (Young, Comas & de Carvalho, 2015). Knowledge about the roles played by cobamide in mycobacterial physiology and during disease has remained limited, partly because mycobacterial pathogens are challenging to study. For instance, the risk of exposure to *M. tuberculosis* is a serious threat to researchers and experiments requiring the use of this organism must be handled in costly, specialized biosafety level 3 (BSL3) facilities. Moreover, *M. tuberculosis* has a long generation time of 24h in nutrient-rich broth and takes more than 3 weeks to form colonies on solid media (James, Williams & Marsh, 2000). To circumvent these challenges, various aspects of mycobacterial physiology, metabolism and pathogenesis have been studied using non-pathogenic mycobacterial models such as *M. smegmatis* and attenuated strains such as engineered auxotrophs and *M. bovis BCG* (Shiloh & DiGiuseppe Champion, 2010). *M. smegmatis* is a particularly convenient model because it is amenable to tractable genetic manipulation, safe to handle in biosafety level 2 (BSL2) laboratories and has a relatively fast doubling time of only 3–4h.

Previous studies in the MMRU established that CNCbl uptake pathways are intact *M. tuberculosis* but *de novo* cobamide biosynthesis appears to be absent (Warner *et al*., 2007). Subsequently, CNCbl transport in *M. tuberculosis* was validated experimentally and BacA (Rv1819c) was discovered as a non-canonical cobamide transporter (Gopinath, Venclovas, *et al*., 2013). In this chapter, direct evidence of *de novo* cobalamin production in *M. smegmatis* is presented. The regulation of methionine biosynthesis by cobalamin in this bacterium is also described. Lastly, the uptake of exogenous CNCbl and dicyanocobanimade ($(CN)_2$Cbi), is examined and findings are compared with current knowledge about corrinoid transport in *M. tuberculosis*.

4.2 Results

4.2.1 The *de novo* cobamide biosynthetic pathway of *M. smegmatis* is functional

Multiple comparative genomics studies have predicted that mycobacteria can make or utilize cobamides (Rodionov *et al*., 2003; Vitreschak *et al*., 2003; Zhang *et al*., 2009; Gopinath, Moosa, *et al*., 2013; Minias, Minias & Dziadek, 2018; Shelton *et al*., 2018). Here, the cobamide biosynthetic genes in *M. smegmatis* were identified by querying the Mycobrowser portal (Kapopoulou, Lew & Cole, 2011) for orthologs of the identified *M. tuberculosis* genes. Hits from this ortholog search indicated that *M. smegmatis* encodes all enzymes required for *de novo* cobamide biosynthesis (**Table 4.1**). Enzymes catalysing the anaerobic pathway were missing, although CobB/H/K/M were found to be hypothetical replacements for CbiA/C/J/F, respectively (**Table 4.1**). Two paralogous genes, *MSMEG_0431* and *MSMEG_4529*, encoded putative CbiX replacements (**Table4.1**). Genes encoding corrinoid salvage and DMB synthesis and attachment were also found.

Besides these, *M. smegmatis* encoded four other proteins (MSMEG_1123, MSMEG_6048, MSMEG_6069, and MSMEG_2607) for which enzymatic functions have not been assigned, although their predicted gene annotations implicated them in cobamide biosynthesis. MSMEG_1123/6048/6069 share homology with CobW protein of proteobacteria, which has been suggested to participate in cobalt chelation due to its proximity to CobNST cobalt-chelatase complex (Rodionov *et al*., 2003). In *M. smegmatis*, the putative CobW proteins were not proximal to CobNST.

MSMEG_2607 was annotated as CbiM, which is a predicted cobalt transporter (Rodionov *et al*., 2003). In sum, this analysis revealed that *M. smegmatis* has the full genetic capacity for the *de novo* biosynthesis of cobamides via the aerobic pathway.

Table 4.1: ***De novo* cobamide biosynthetic genes in *M. smegmatis***

Symbol	EC#	Synthesis step	*M. smegmatis* gene number
GltX	EC 6.1.1.17	ALA and Uro' III	*MSMEG_2383/MSMEG_6306*
ProS	EC 6.1.1.15	ALA and Uro' III	*MSMEG_2621*
HemA	EC 1.2.1.70	ALA and Uro' III	*MSMEG_0952/MSMEG_0661*
HemL	EC 5.4.3.8	ALA and Uro' III	*MSMEG_0969/MSMEG_5970*
HemB	EC 4.2.1.24	ALA and Uro' III	*MSMEG_0956*
HemC	EC 2.5.1.61	ALA and Uro' III	*MSMEG_0953*
HemD	EC 4.2.1.75	ALA and Uro' III	*MSMEG_0432/MSMEG_0954*
CobA_CysG	EC 2.1.1.107	ALA and Uro' III	*MSMEG_2618*
CysG	EC 1.3.1.76	ALA and Uro' III	*MSMEG_2618*
CobI	EC 2.1.1.130	Corrin ring (aerobic)	*MSMEG_3873*
CobG	EC 1.14.13.83	Corrin ring (aerobic)	*MSMEG_3871*
CobJ	EC 2.1.1.131	Corrin ring (aerobic)	*MSMEG_3873*
CobM	EC 2.1.1.133	Corrin ring (aerobic)	*MSMEG_3877*
CobF	EC 2.1.1.152	Corrin ring (aerobic)	*MSMEG_5548*
CobK	EC 1.3.1.54	Corrin ring (aerobic)	*MSMEG_3875*
CobL	EC 2.1.1.132	Corrin ring (aerobic)	*MSMEG_3878*
CobH	EC 5.4.99.61	Corrin ring (aerobic)	*MSMEG_3872*
CobB	EC 6.3.5.9	Corrin ring (aerobic)	*MSMEG_2617*
CobNST	EC 6.6.1.2	Corrin ring (aerobic)	*MSMEG_3864 (cobN); MSMEG_4277 (cobS); MSMEG_4275 (cobT)*

Symbol	EC#	Synthesis step	*M. smegmatis* gene number
CbiK_CbiX	EC 4.99.1.3	Corrin ring (anaerobic)	*MSMEG_0431/MSMEG_4529*
CbiL	EC 2.1.1.151	Corrin ring (anaerobic)	unidentified
CbiH	EC 2.1.1.272	Corrin ring (anaerobic)	unidentified
CbiF	EC 2.1.1.271	Corrin ring (anaerobic)	*MSMEG_3877 (cobM)*
CbiG	EC 3.7.1.12	Corrin ring (anaerobic)	unidentified
CbiD	EC 2.1.1.195	Corrin ring (anaerobic)	unidentified
CbiJ	EC 1.3.1.106	Corrin ring (anaerobic)	*MSMEG_3875 (cobK)*
CbiT	EC 2.1.1.196	Corrin ring (anaerobic)	unidentified
CbiE	EC 2.1.1.289	Corrin ring (anaerobic)	unidentified
CbiC	EC 5.4.99.60	Corrin ring (anaerobic)	*MSMEG_3872 (cobH)*
CbiA	EC 6.3.5.11	Corrin ring (anaerobic)	*MSMEG_2617 (cobB)*
CobR	EC 1.16.8.1	Cbi-P	*MSMEG_6035*
CobO_PduO	EC 2.5.1.17	Cbi-P	*MSMEG_2616 (cobO); MSMEG_4934 (pduO); MSMEG_1544 (pduO)*
CobQ	EC 6.3.5.10	Cbi-P	*MSMEG_2588*
CobC_CobD Protein α	EC 6.3.1.10	Cbi-P	*MSMEG_4305 (cobC); MSMEG_4310 (cobD)*
CobU_CobP	EC 2.7.1.156	Cbi-P	*MSMEG_4274 (cobU)*
CobU_CobP	EC 2.7.7.62	Cbi-P	*MSMEG_4274 (cobU)*
RibF	EC 2.7.1.26	DMB	*MSMEG_2653*
Fre	EC 1.5.1.38	DMB	*MSMEG_2800*
BluB	EC 1.13.11.79	DMB	*MSMEG_6053*
CobU_CobT	EC 2.4.2.21	DMB attachment	*MSMEG_4275 (cobT)*
CobC	EC 3.1.3.73	DMB attachment	*MSMEG_4305 (cobC)*
CobS_CobV	EC 2.7.8.26	DMB attachment	*MSMEG_4277 (cobS)*

4.2.2 Validation of *de novo* cobalamin biosynthesis in *M. smegmatis* by LC-MS/MS

Previous observations using an *M. smegmatis* Δ*metE*Δ*cobK*::*hyg* strain had suggested that *M. smegmatis* could produce and use cobalamin as a co-factor for MetH (Dr. Stephanie Dawes, unpublished). This genotype of this mutant was validated by PCR, confirming an in-frame deletion of the entire open reading frame (ORF) of *metE* and a disruption of *cobK* generated by the insertion of a *hyg* fragment (**Figure 4.1A**). To examine the growth phenotype of the Δ*metE*Δ*cobK*::*hyg* strain, 1×10^6 cells/ml were seeded in a 96-well plate and grown at 37°C overnight in 7H9-OADC medium containing 10-fold dilutions of CNCbl or methionine. The highest concentration of CNCbl tested was 10μg/ml and the highest methionine concentration was 1mM. Growth was inferred by the formation of visible colonies after 3 days' incubation. There were no colonies in wells containing the Δ*metE*Δ*cobK*::*hyg* strain without supplement but in 7H9-OADC medium supplemented with either 1mM methionine or 10μg/ml CNCbl, similar colony sizes were observed between this mutant and a wild-type control (**Figure 4.1B**). Below 10ng/ml CNCbl or 1μM methionine (*i.e.*, 1000-fold dilutions of either supplement), the Δ*metE*Δ*cobK*::*hyg* strain did not form visible colonies in wells (**Figure 4.1B**). There were no differences between the growth curve of the Δ*metE*Δ*cobK*::*hyg* strain grown in 7H9-OADC medium supplemented with 10μg/ml CNCbl or in medium supplemented with 1mM methionine (**Figure 4.1C**). These data demonstrated that the Δ*metE*Δ*cobK*::*hyg* strain required methionine or CNCbl for growth in 7H9-OADC medium due to the absence of both MetE and MetH activity (due to the absence of *de novo* cobamide biosynthesis).

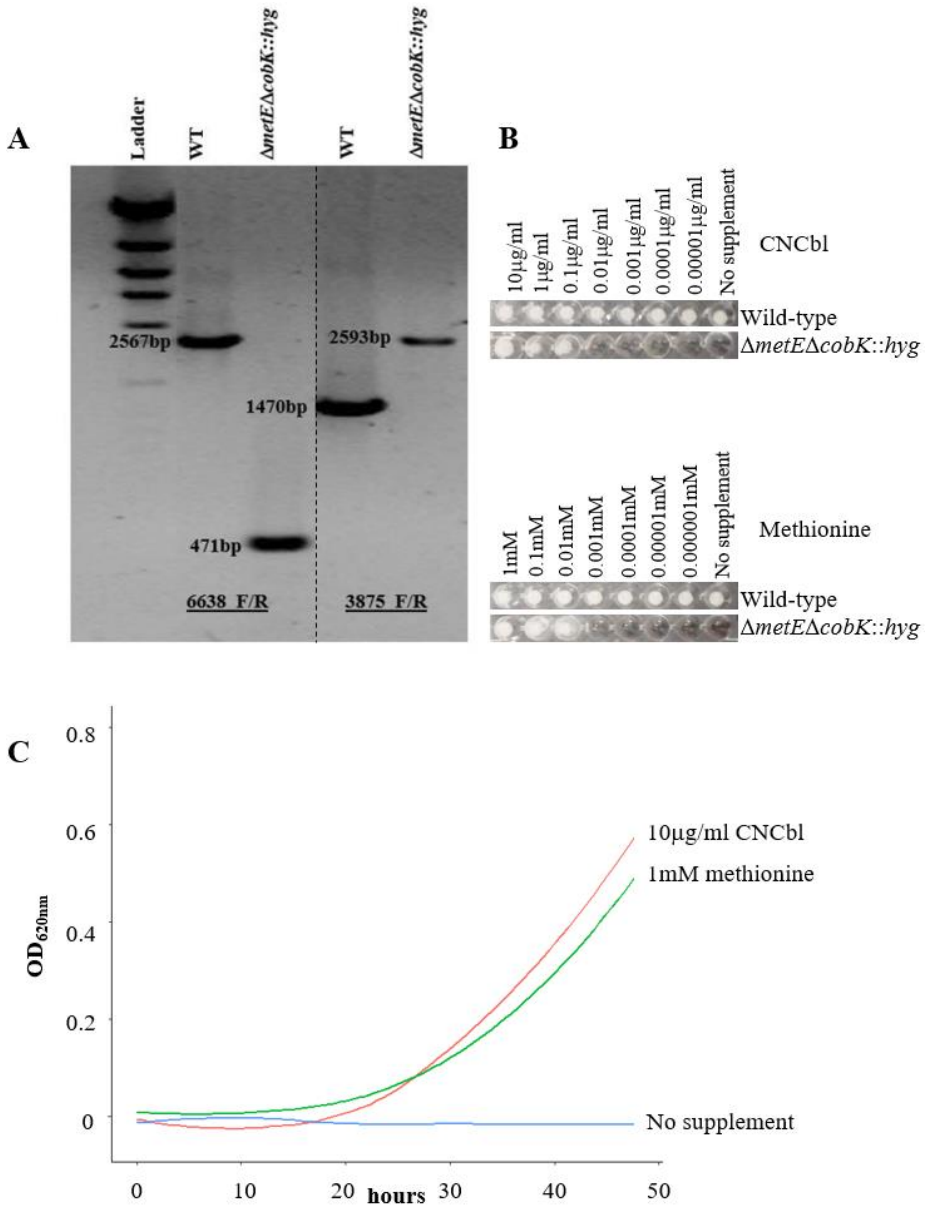


Figure 4.1: **Genotypic and phenotypic characterization of the Δ*metE*Δ*cobK*::*hyg* strain**. **A**. The genotype of Δ*metE*Δ*cobK*::*hyg* received from Dr. Stephanie Dawes (unpublished) was verified by PCR using primers 3875_F/R flanking the *cobK* allele and 6638_F/R flanking the *metE* allele. PCR confirmed the deletion of *metE* gene, resulting in a 471bp PCR product and the insertion of an 1124bp *hyg* fragment into the *cobK* allele, resulting in a 2593bp PCR fragment. The wild-type PCR products for *metE* and *cobK* were 2567bp and 1470bp, respectively. The dashed line shows the interface of images merged from two distict gels. **B**. Rescue of the Δ*metE*Δ*cobK*::*hyg* strain by exogenous CNCbl (CNCbl) or methionine. **C**. Growth curve of the Δ*metE*Δ*cobK*::*hyg* strain in 7H9-OADC medium in the presence or absence of 10µg/ml exogenous CNCbl or 1mM methionine, showing failure of the strain to grow in unsupplemented medium. Absorbance measurements (OD_{620nm}) were taken every 15min on a microplate reader over a time course of 48h. The growth curve was plotted using "R."

Previous work had also shown that when the medium contained exogenous CNCbl, the expression of MetE was repressed by a riboswitch (Warner *et al*., 2007). To investigate cobalamin production in *M. smegmatis*, a sensitive LC-MS/MS method was developed in the MMRU to directly quantify internal cobalamin concentrations (Dr. Gabriel Mashabela, unpublished). In this method, extracted cobalamin was derivatized to the cyanated form by boiling cell lysates in solution containing potassium cyanide. To examine cobalamin-dependent regulation of methionine biosynthesis, a strain which was cobalamin-deficient but expressing both MetE and MetH was needed. To generate this mutant, allelic exchange mutagenesis was used to replace the wild-type allele with deleted *MSMEG_3875* allele, creating an unmarked in-frame deletion of 80% of the *cobK* ORF (**Figure 4.2A**). PCR and Southern blotting confirmed the successful isolation of Δ*cobK* (**Figure 4.2B-C**), showing that this gene was nonessential for *in vitro* growth of *M. smegmatis*. The Δ*cobK* strain exhibited no morphological differences compared to the parental wild-type and grew at the same rate as the wild-type strain (**Figure 4.2D**).

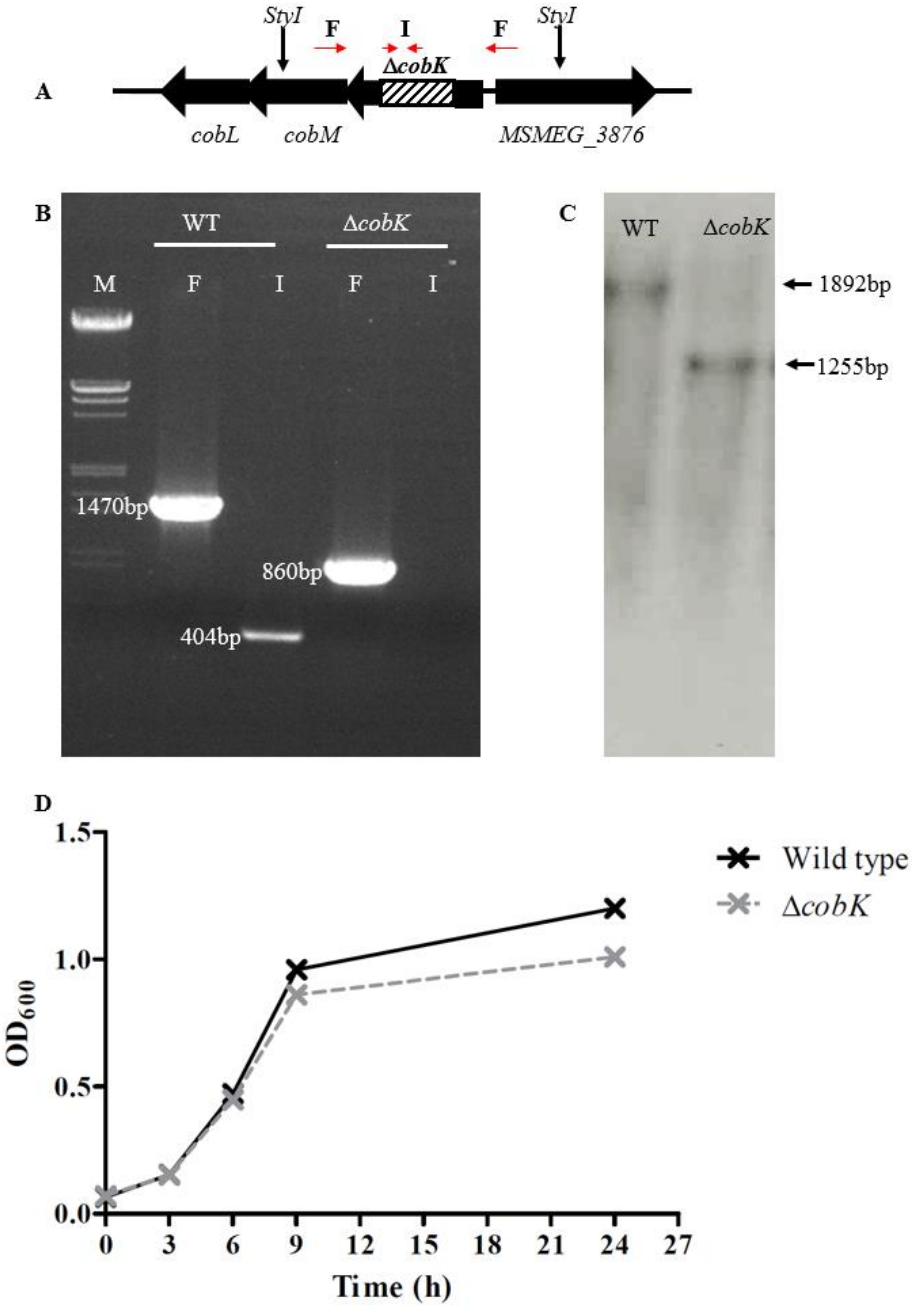


Figure 4.2: **Generation and confirmation of *M. smegmatis* Δ*cobK* strain**. **A.** The deletion construct eliminated 636bp of *cobK* (diagonally patterned rectangle). Primers amplifying the region flanking (F) or internal (I) to the deleted portion of *cobK* were designed for PCR screening. **B.** The deletion was confirmed with PCR. Expected product sizes were observed for wild-type and Δ*cobK* when flanking primers were used, but internal primers yielded a product in wild-type only. M – DNA molecular weight marker. **C.** Southern blotting confirmation of ΔcobK. A PCR-generated probe for *cobK* was used to detect a fragment between two naturally occurring *Sty*I restriction sites (down-facing arrows in **A**). The specific bands for wild-type (1892bp) and mutant (1255bp) were observed. **D.** Growth curves of wild-type (solid black line) and Δ*cobK* strain (grey dashed line) were generated by plotting $OD600_{nm}$ values of a shaking culture every 3h for 9h and after 24h.

To test whether the *de novo* cobamide biosynthetic pathway in *M. smegmatis* was compromised by deleting *cobK*, whole-cell lysates of the Δ*cobK* strain and the parental wild-type strain were analysed by LC-MS/MS. Cobalamin was positively identified using an MRM method that detected two transitions corresponding to α-ribazole 5-phosphate (m/z, 678→359) and DMB (m/z, 678→147) (**Figure 4.3A**). If cobalamin were present, these two transitions would be observed as co-eluting peaks (**Figure 4.3B**).

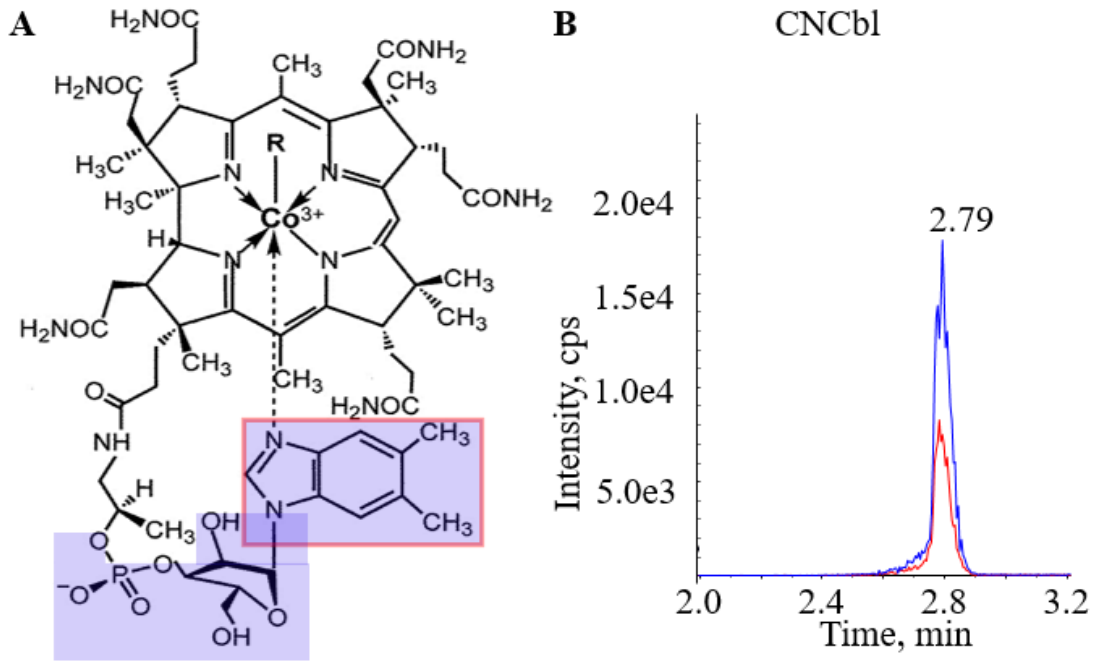


Figure 4.3: **Detection of cobalamin in *M. smegmatis* using LC-MS/MS. A.** Cobalamin chemical structure. A corrin ring surrounds a centrally-chelated cobalt ion, coordinated by an alkyl upper ligand (R-group) and a lower base, DMB. Several chemical groups can occupy the R position, giving a unique identity to the cobalamin form: cyano- (cyanocobalamin), adeno- (adenosylcobalamin), methyl- (methylcobalamin) or hydroxyl- (aquacobalamin). **B.** Traces corresponding to a cyanocobalamin standard (CNCbl) showing the co-elution of α-ribazole 5-phosphate (blue trace) and DMB (red trace) at 2.79min. The LC-MS/MS method was optimized to detect α-ribazole 5-phosphate (MW = 359g.mol^{-1}; blue background in ***A***) and DMB (MW = 147g.mol^{-1}; red box in ***A***).

As expected, cell extracts of wild-type *M. smegmatis* culture produced high intensity cobalamin peaks (**Figure 4.4**). These findings provided direct proof of the functionality of the aerobic cobamide pathway in this organism. By contrast, cobalamin was not detected in the Δ*cobK* strain (**Figure 4.4**). The deletion of *cobK*, therefore, abolished *de novo* cobalamin biosynthesis in *M. smegmatis*. That this mutant

grew normally without exogenously provided CNCbl indicated that MetE was functional in *M. smegmatis* and substituted for MetH in the absence of the co-factor. Because the LC-MS/MS method was optimized to specifically identify the cobamide bearing DMB as the lower ligand, the detection of cobalamin in *M. smegmatis* confirmed also DMB biosynthesis in this organism was also functional, via an as-yet-unidentified pathway.

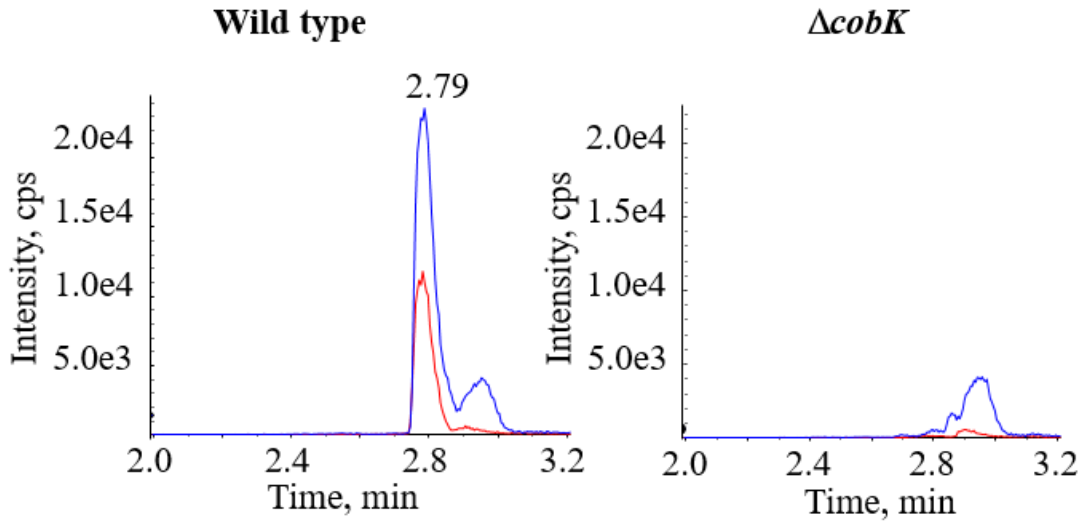


Figure 4.4: **LC-MS detection of *de novo* cobalamin biosynthesis in *M. smegmatis***. Cell extracts of wild type Δ*cobK* analysed by LC-MS/MS to detect the two cobalamin transitions co-eluting peaks at 2.79min. The peaks corresponding to α-ribazole 5-phosphate (blue trace) and DMB (red trace) were detected in extracts of the wild type strain but not in the Δ*cobK* strain. Peak intensities are expressed as counts per second (cps).

4.2.3 Transcriptional repression of *metE* by endogenous cobalamin in *M. smegmatis*

Besides its role as a co-factor for MetH, cobalamin potentially regulates methionine biosynthesis in *M. smegmatis* as a ligand of a predicted riboswitch upstream of *metE*. In *M. tuberculosis*, cobalamin induces transcriptional repression of *metE* via this riboswitch (Warner *et al*., 2007). To test whether this was also true in *M. smegmatis*, *metE* gene expression levels in wild-type and Δ*cobK* strains were compared using ddPCR. This quantitation revealed that the levels of *metE* in Δ*cobK* were elevated by 12.5× relative to the parental wild-type strain (**Figure 4.5**). This result indicated that the predicted riboswitch upstream of *metE* in *M. smegmatis* was functional and

sensitive to endogenous cobalamin. However, some *metE* transcripts were still detectable in the wild-type strain, implying that repression by internal co-factor was not absolute (**Figure 4.5**). The increased transcript count following the elimination of cobalamin in Δ*cobK* was nonetheless consistent with predicted de-repression of the *metE* riboswitch. Cobalamin deficiency in Δ*cobK* did not affect *metH* transcript levels (**Figure 4.5**).

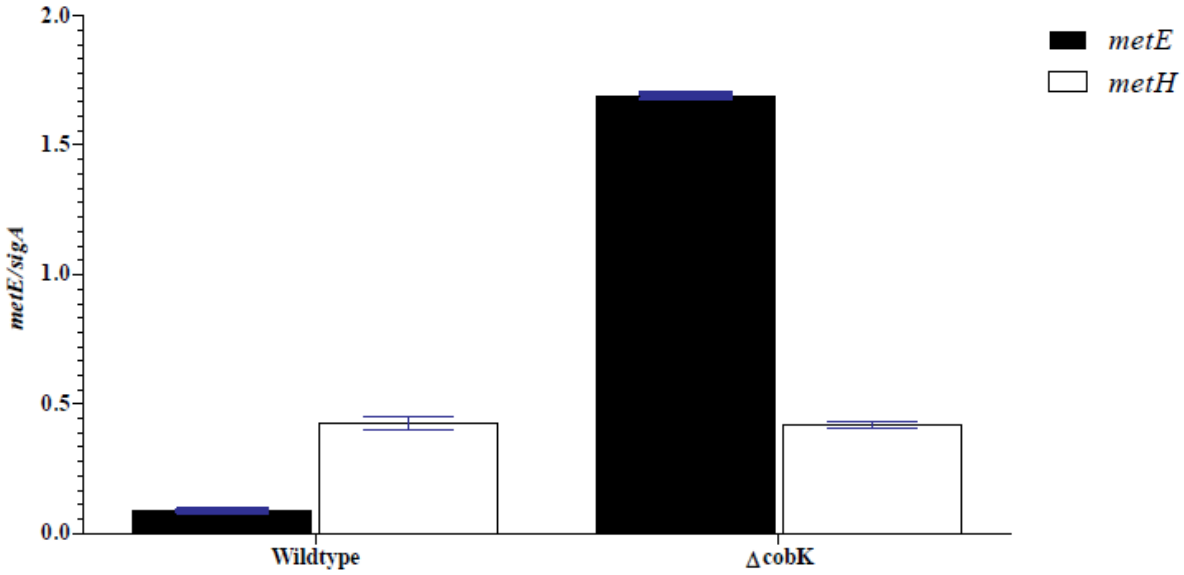


Figure 4.5: **Transcriptional repression of *metE* by endogenous cobalamin in *M. smegmatis*. A.** Absolute copy numbers of *metE* transcripts in *M. smegmatis* cultures were determined using ddPCR and normalized to the housekeeping gene, *sigA*. The cobalamin-deficient Δ*cobK* strain had over 10-fold increase in *metE* gene expression relative to wild-type. There was no difference in relative expression of *metH* between wild-type and Δ*cobK* strains. The data are representative of two independent experiments. The error bars show the standard error of the mean for three technical replicates.

4.2.4 Marginal control of *metE* gene expression by exogenous CNCbl in *M. smegmatis*

The impact of exogenous CNCbl on the *metE* riboswitch was also evaluated by comparing gene expression levels in wild-type and Δ*cobK* strains grown in 7H9 medium supplemented with CNCbl. It was expected that, in the presence of exogenous CNCbl, the Δ*cobK* strain would phenocopy the wild-type strain with respect to *metE* transcription. However, ddPCR showed that exogenous CNCbl only slightly repressed

metE in the mutant but did not cause detectable changes in *metE* expression in the wild-type strain (**Figure 4.6**).

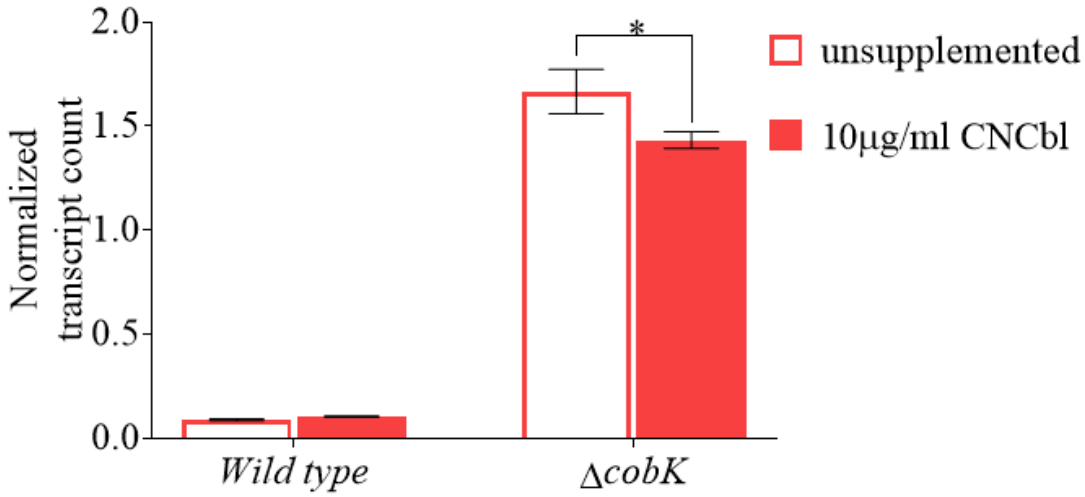
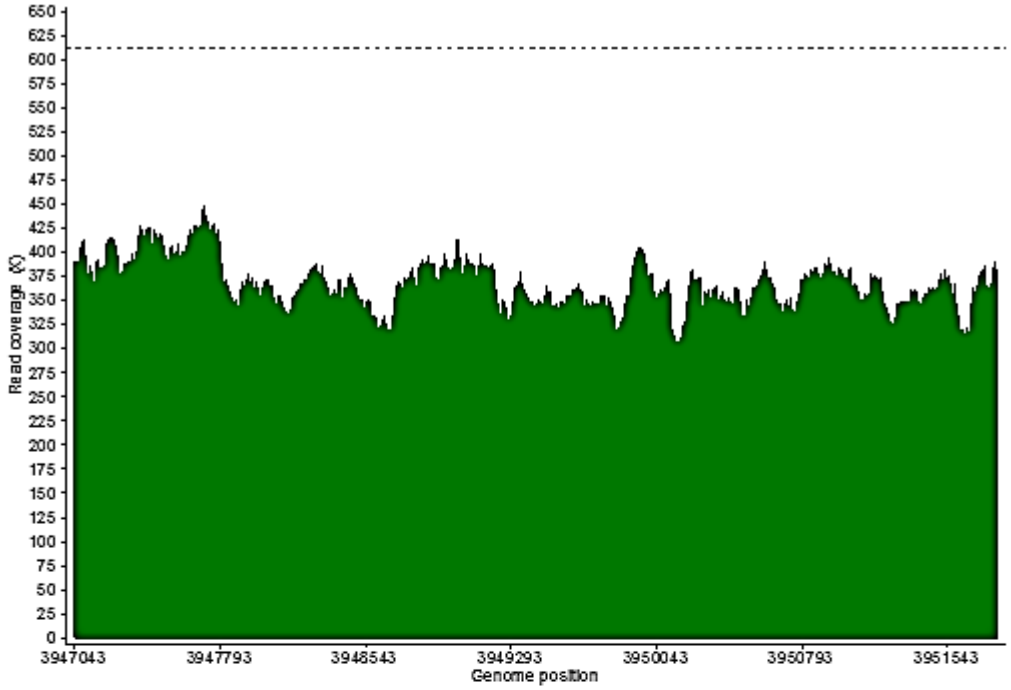
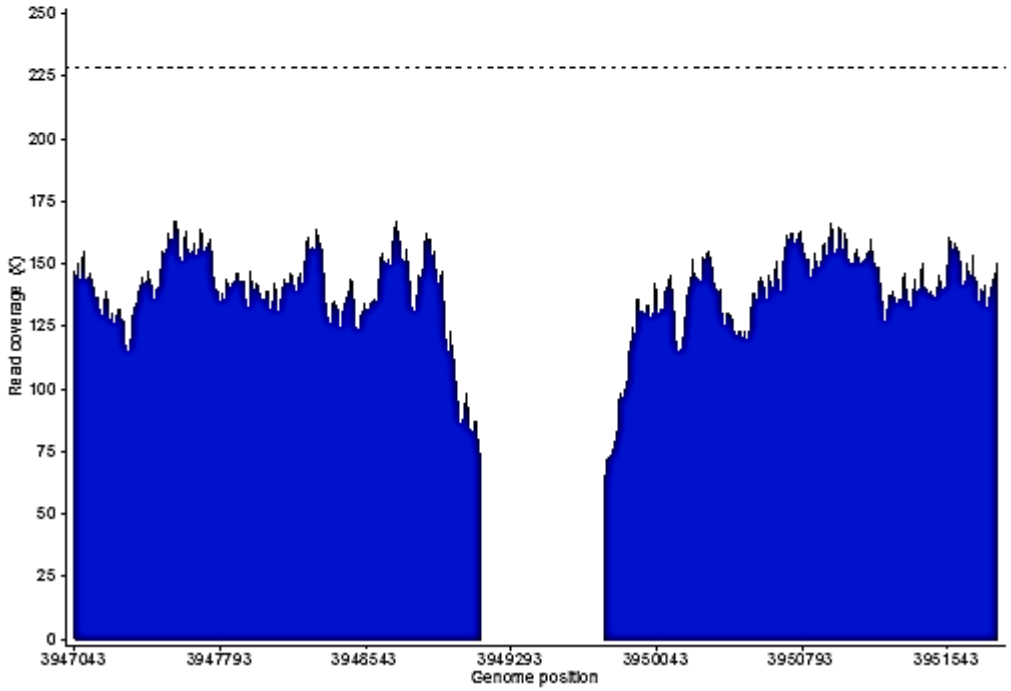


Figure 4.6: **Relative *metE* gene expression in the presence of exogenous CNCbl**. Quantification of *metE* transcription in wild type and Δ*cobK* in the presence or absence of exogenous CNCbl using ddPCR. In unsupplemented cultures, the Δ*cobK* mutant shows ≥10-fold increase in *metE* gene expression relative to wild type. Providing B_{12} in the growth media produced a small but significant decrease in *metE* levels in Δ*cobK* (p=0.0118; (*); ordinary two-way ANOVA) but not in the wild type, relative to unsupplemented cultures.

As previously noted, the Δ*metE*Δ*cobK*::*hyg* strain could be rescued by methionine or CNCbl, which implied that uptake systems were intact in *M. smegmatis*. Therefore, it was surprising that exogenous CNCbl repressed *metE* only slightly in the Δ*cobK* strain. One possibility was that the mutant might have harboured secondary mutations affecting the binding of cobalamin to the *metE* riboswitch. To test this assumption, the Δ*cobK* strain and the parental wild-type strain were whole-genome sequenced to screen for such mutations. WGS data were also used as an additional screen for the deletion of *cobK*. To do this, a plot of read coverage along the genomic locus of *cobK* was made (**Figure 4.7**).



Wildtype



ΔcobK

Figure 4.7: **A plot of read coverage along the genomic coordinates of *cobK* in *M. smegmatis***. The wild-type strain is plotted in green and Δ*cobK* in blue. The average coverage for each strain is indicated by a dashed line above the plot. The drop in read depth in Δ*cobK* strain corresponded to the deleted portion of the gene.

Analysis of the WGS data revealed no SNMs in the 5' upstream region or in the open reading frame of *metE* in either wild-type or Δ*cobK*. In addition, no mutations were detected in any of the other genes involved in *de novo* cobalamin biosynthesis, confirming that the deletion of *cobK* was solely responsible for the abrogation of cobalamin production in this mutant. However, SNMs were found in six genes in the Δ*cobK* strain but not in the parental wild-type strain (**Table 4.2**). Out of the six SNMs, three were non-synonymous (ns-SNM) in *MSMEG_2148* (putative HNH endonuclease protein), *MSMEG_3876* (putative phosphotransferase) and *MSMEG_6127* (putative anti-anti-sigma factor). The other three were SNMs found upstream of the start codons of *MSMEG_0691* (putative transcriptional regulatory protein), *MSMEG_6270* (hypothetical protein) and *MSMEG_6423* (a putative phosphodiesterase).

Table 4.2: **Unique SNMs detected in *M. smegmatis* Δ*cobK* but not in the parental wild-type strain**

Gene	**Annotation** (Kapopoulou, Lew & Cole, 2011)	**SNM**	**Type of SNM**	**Amino Acid Change**
MSMEG_0691	putative transcriptional regulatory protein	1T>G	upstream of start codon	
MSMEG_2148	HNH endonuclease domain protein	1135T>C	ns-SNM	Trp379Arg
MSMEG_3876	putative phosphotransferase enzyme family protein	887G>A	ns-SNM	Arg296His
MSMEG_6127	anti-anti-sigma factor	107T>G	ns-SNM	Leu36Arg
MSMEG_6270	hypothetical protein	1A>C	upstream of start codon	

Gene	Annotation (Kapopoulou, Lew & Cole, 2011)	SNM	Type of SNM	Amino Acid Change
		1G>A		
MSMEG_6423	Glycerophosphoryl diester phosphodiesterase family protein	1T>C 1T>G	upstream of start codon	

4.2.5 Uptake and assimilation of dicyanocobinamide in *M. smegmatis*

Because WGS did not find any obvious mutations responsible for the weak repression of *metE* by exogenous CNCbl, it was then postulated that CNCbl uptake was so poor that only marginal control of *metE* expression was achieved. A second intriguing possibility was that the *metE* riboswitch discriminated between *de novo* synthesised cobalamin and exogenous CNCbl. To interrogate this notion, the cobalamin-deficient mutants were fed $(CN)_2$Cbi; (procured from Sigma-Aldrich®) as a precursor for cobamide production. Previous studies had demonstrated that *M. tuberculosis* could assimilate this precursor into a physiologically functional cobamide (Gopinath, Venclovas, *et al*., 2013). Therefore, it was reasoned that *M. smegmatis* would similarly transport and assimilate $(CN)_2$Cbi.

It was first necessary to determine whether *M. smegmatis* could assimilate $(CN)_2$Cbi. This was tested by finding out if $(CN)_2$Cbi could rescue the Δ*metE*Δ*cobK*::*hyg* strain. To do this, the mutant was first grown in 7H9 medium to exponential phase with 1mM methionine to support its growth, and then serially diluted and spotted on agar containing 10μM $(CN)_2$Cbi. This assay confirmed that $(CN)_2$Cbi supported the growth of Δ*metE*Δ*cobK*::*hyg* (**Figure 4.8**). However, at this concentration of $(CN)_2$Cbi, Δ*metE*Δ*cobK*::*hyg* took longer to form colonies than when a similar concentration of CNCbl was used (**Figure 4.8**). In the CNCbl-containing plates, only three days of

incubation were required for distinctly visible colonies to emerge. By contrast, it took 13 days to form colonies of comparable sizes in the $(CN)_2$Cbi plates (**Figure 4.8**). This was not surprising, as it has been reported that the assimilation of $(CN)_2$Cbi in mycobacteria is very poor (Gopinath, Venclovas, *et al.*, 2013). Nevertheless, these observations clearly demonstrated that *M. smegmatis* could utilize $(CN)_2$Cbi.

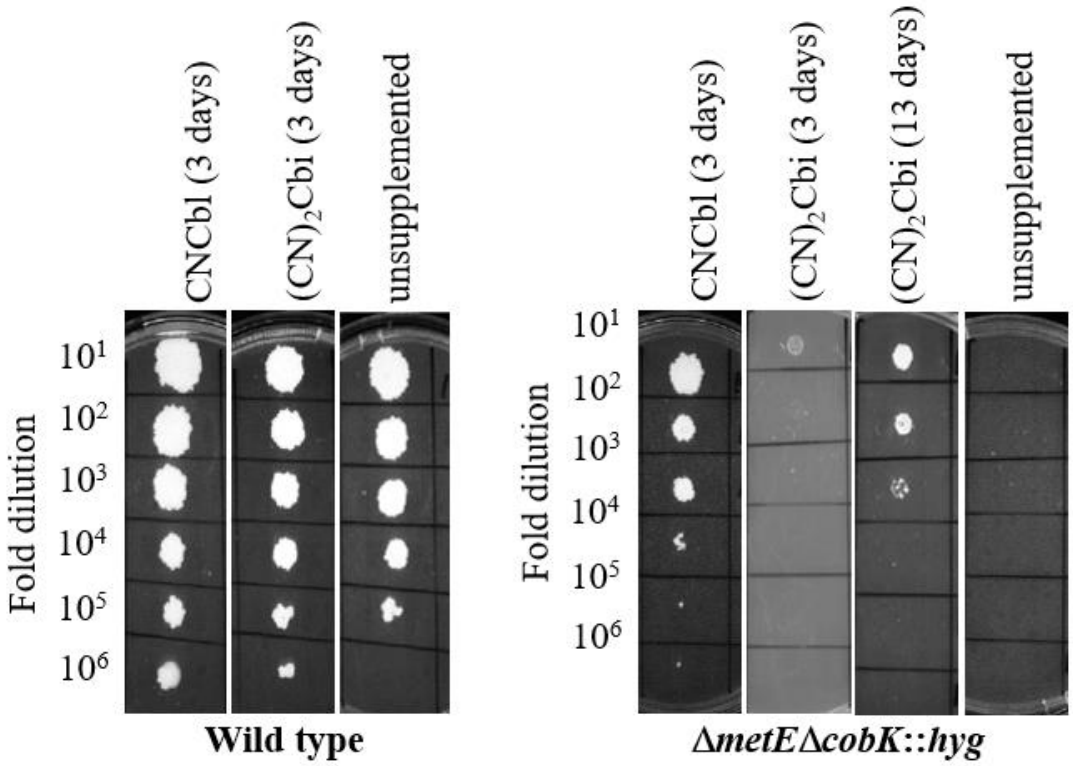


Figure 4.8: **Uptake and assimilation of dicyanocobinamide by *M. smegmatis*.** Spotting assays of exponentially-growing cultures of wild-type and Δ*metE*Δ*cobK*::*hyg* strains on 7H10-OADC agar containing either 10μg/ml CNCbl or 10μM $(CN)_2$Cbi, or without supplementation. Plates were imaged at 3 days and the $(CN)_2$Cbi plate was reimaged after 13 days.

Interestingly, there was no difference in the growth rate of Δ*metE*Δ*cobK*::*hyg* in liquid culture when either $(CN)_2$Cbi or CNCbl was used; *i.e.*, in both conditions, similar $OD600_{nm}$ values were obtained at any given time point. However, the uptake of $(CN)_2$Cbi in liquid culture was accompanied by a remarkable change in colour of the spent media (**Figure 4.9**). In the Δ*metE*Δ*cobK*::*hyg* strain, the colour of the media changed from the initial purple colour of $(CN)_2$Cbi to a final pale-yellow colour, in contrast to that of wild-type and Δ*cobK* spent media, which changed to a rusty colour.

These observations hinted that, although the uptake of $(CN)_2$Cbi was poor in *M. smegmatis*, metabolic necessity could trigger increased salvage. In this case, the absolute requirement for cobalamin in Δ*metE*Δ*cobK*::*hyg* might have presented a stronger impetus for $(CN)_2$Cbi assimilation, unlike in the wild-type and Δ*cobK* strains, which could grow without supplementation.

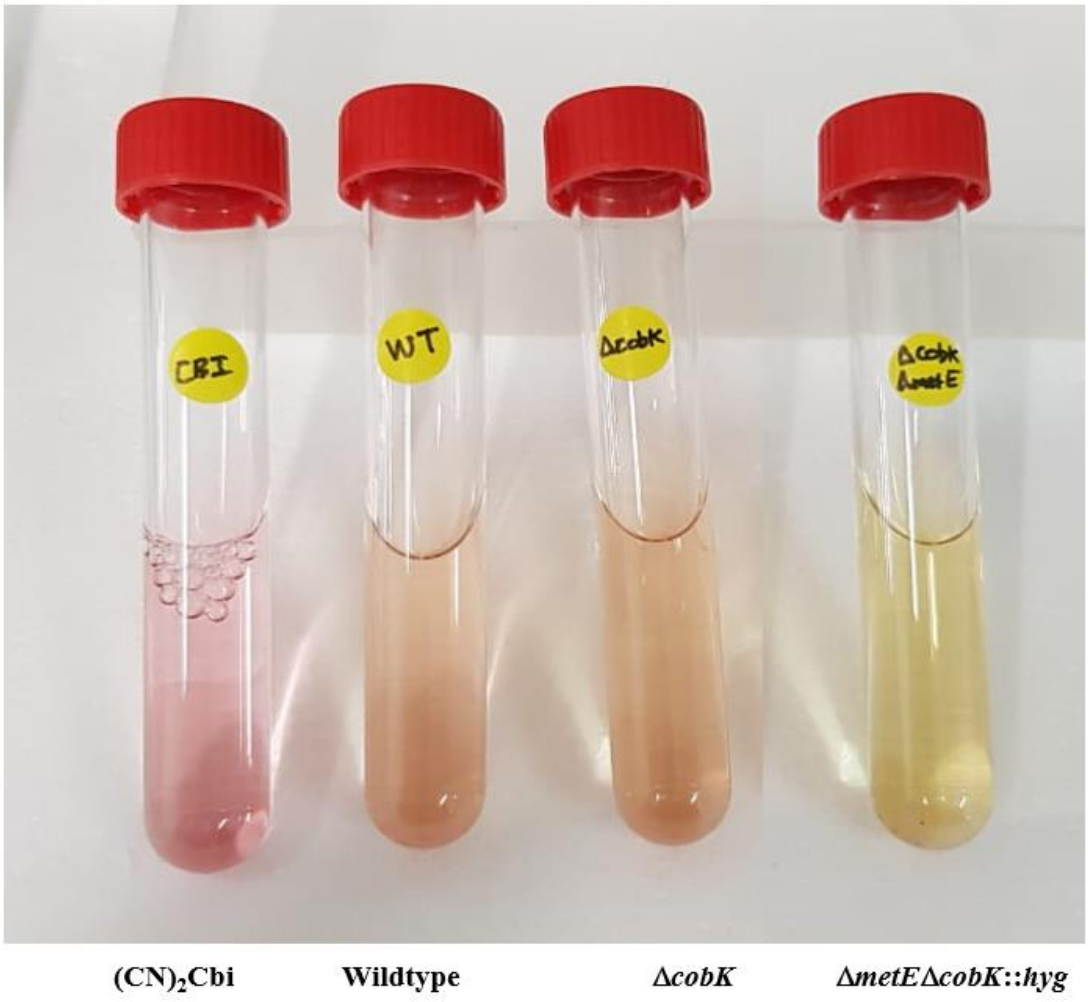


Figure 4.9: **Dicyanocobinamide uptake in *M. smegmatis* in liquid cultures.** *M. smegmatis* wild-type or mutant strains were grown to stationary phase (OD_{600nm} > 2) with 30μM $(CN)_2$Cbi overnight in a shaking incubator. The spent media was harvested by centrifugation and filtered through a 0.22μm syringe filter. $(CN)_2$Cbi uptake was accompanied by distinctive spent media colour changes. The purple colour of $(CN)_2$Cbi faded more strikingly in the Δ*metE*Δ*cobK*::*hyg* strain than in the other strains.

4.2.6 LC-MS/MS confirmation of *de novo* cobalamin biosynthesis via dicyanocobinamide assimilation

To confirm that *M. smegmatis* used exogenous $(CN)_2$Cbi as a precursor for *de novo* cobalamin biosynthesis, wild-type, Δ*cobK* and Δ*metE*Δ*cobK*::*hyg* strains were grown in liquid cultures containing $(CN)_2$Cbi and cell extracts were analysed by LC-MS/MS for *de novo* synthesised cobalamin, which was derivatized to the cyanated form by boiling cell lysates in solution containing potassium cyanide. To circumvent poor uptake, $(CN)_2$Cbi was used at a high concentration (30μM). Since $(CN)_2$Cbi was 93% pure and contained trace amounts (≤ 0.1%) of CNCbl as per the manufacturer's specifications, a 1μM control $(CN)_2$Cbi sample was analysed for background cobalamin peaks, expecting that at this concenctration, possible CNCbl contamination (≤ 1nM) was high enough to detect via LC-MS/MS. LC-MS/MS data showed that the cobalamin signal in the $(CN)_2$Cbi control was below the limit of detection (**Figure 4.10**), confirming that any cobalamin found in $(CN)_2$Cbi-supplemented cultures would be a true reflection of $(CN)_2$Cbi conversion. Consistent with phenotypic data, cobalamin peaks were detected in cell extracts of Δ*metE*Δ*cobK*::*hyg* supplemented with $(CN)_2$Cbi, although the intensity of the cobalamin peaks was approximately 3-fold lower that of unsupplemented wild-type strain (**Figure 4.10**). In contrast, cobalamin was not detected in the Δ*cobK* strain despite supplementation with excess $(CN)_2$Cbi (**Figure 4.10**), which reinforced the presumption that measurable $(CN)_2$Cbi uptake occurred only in cells in which the co-factor was essential for growth. In sum, the LC-MS/MS method validated the intactness of the $(CN)_2$Cbi salvage pathway; *i.e.*, *de novo* cobalamin biosynthesis in *M. smegmatis* can proceed via the $(CN)_2$Cbi salvage pathway.

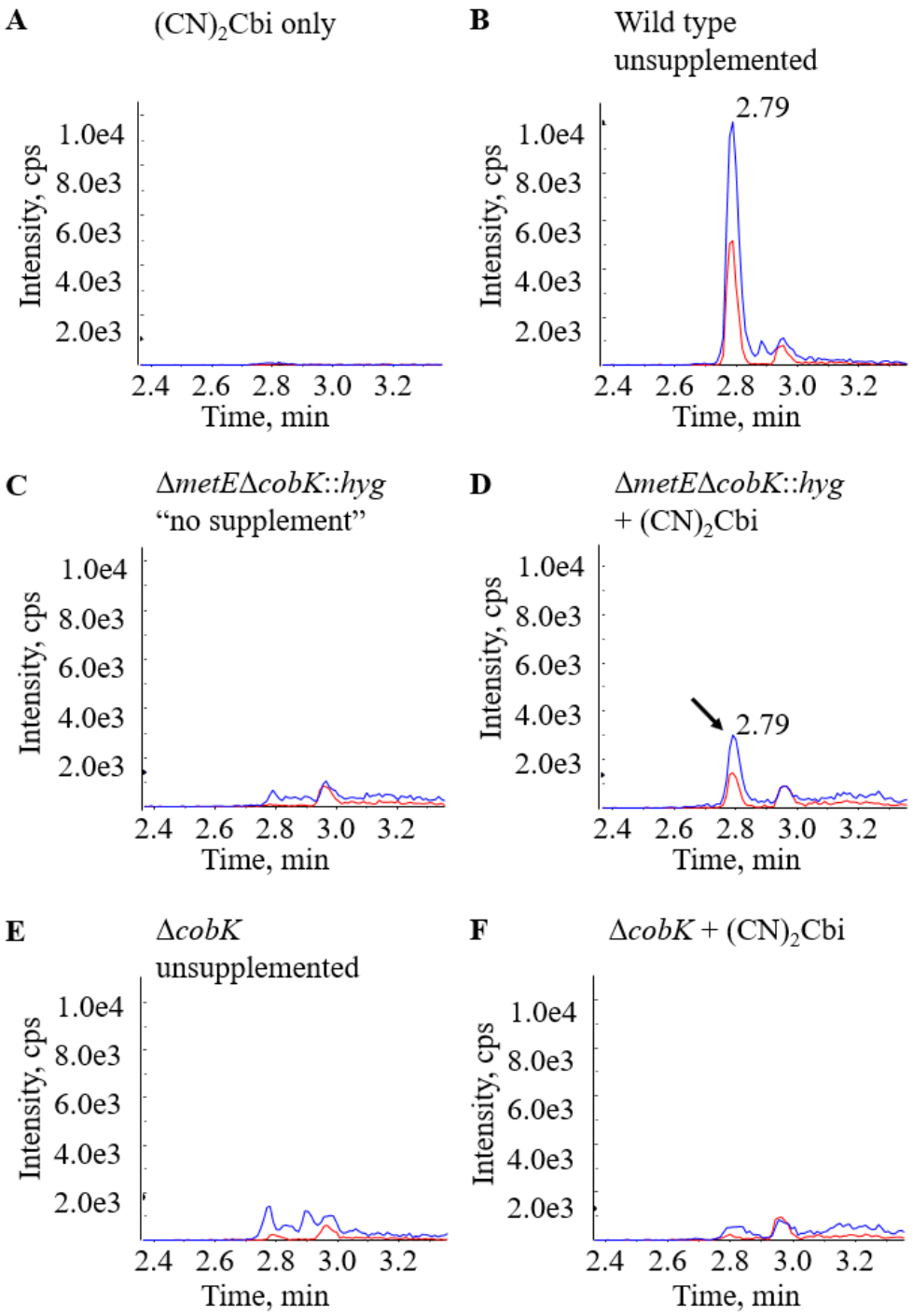


Figure 4.10: **LC-MS confirmation of dicyanocobinamide-assimilation in *M. smegmatis*. A.** A sample of commercial $(CN)_2$Cbi showing that potentially-contaminating CNCbl is undetectable. **B.** *De novo* cobalamin production in wild type bacilli cultured in unsupplemented 7H9-OADC medium. **C-F.** Detection of cobalamin in Δ*metE*Δ*cobK*::*hyg* **(C-D)** or Δ*cobK* **(E-F)** strains grown with or without $(CN)_2$Cbi. Peaks corresponding to α-ribazole 5-phosphate (blue trace) and DMB (red trace) were detected in the $(CN)_2$Cbi-supplemented Δ*metE*Δ*cobK*::*hyg* strain only (arrow). The "no supplement" Δ*metE*Δ*cobK*::*hyg* strain was grown in medium containing 1mM methionine.

4.2.7 Control of *metE* repression by *de novo* synthesised cobalamin derived from assimilated dicyanocobinamide

Because LC-MS/MS showed that the amount of cobalamin derived from $(CN)_2$Cbi assimilation was too low to be detected in wild-type and the Δ*cobK* strains, it was interesting to explore whether these ultralow concentrations were still sufficient to measurably repress *metE*. To investigate this possibility, cultures of wild-type and Δ*cobK* strains were grown in $(CN)_2$Cbi-containing media, followed by quantitation of *metE* transcripts using ddPCR. Again, a high concentration of $(CN)_2$Cbi was used to offset poor uptake. Quantitation analysis showed that $(CN)_2$Cbi supplementation had no repressive effect on *metE* (**Figure 4.11**), which was consistent with poor uptake. Additionally, these findings suggested that on its own, $(CN)_2$Cbi might not have been a compatible ligand for the *metE* riboswitch. However, since the LC-MS/MS method specifically detected DMB as a marker of *de novo* synthesised cobalamin, it was not possible to measure intracellular accumulation of $(CN)_2$Cbi itself.

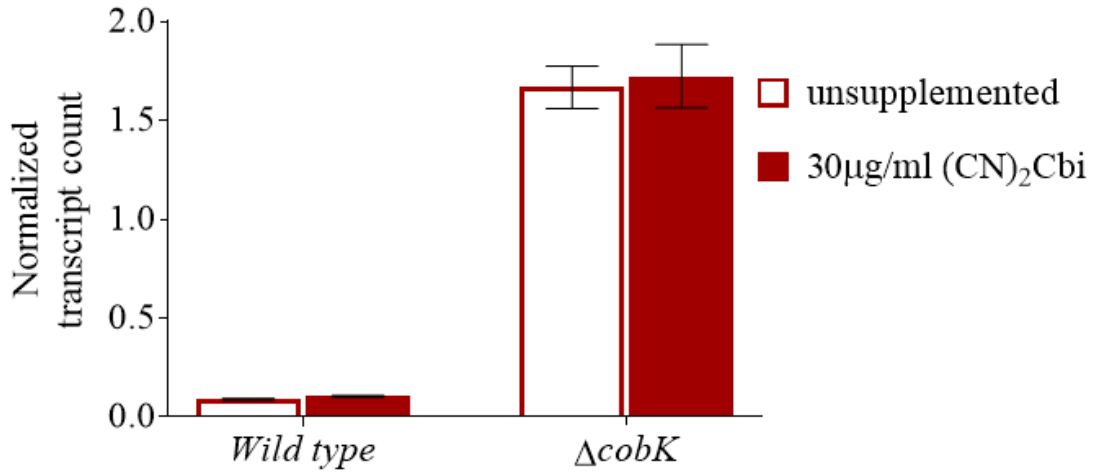


Figure 4.11: **Relative expression of *metE* in wild-type and Δ*cobK* cultures supplemented with dicyanocobinamide.** Absolute quantification of *metE* transcripts in the presence or absence of $(CN)_2$Cbi using ddPCR. $(CN)_2$Cbi supplementation did not change *metE* transcript level in either wild type or Δ*cobK*, relative to unsupplemented cultures. The graphed data are representative of two independent experiments. The error bars show the standard error of the mean.

4.2.8 Putative cobamide transporters in *M. smegmatis*

The genomics, gene expression and LC-MS/MS data supported the functionality of CNCbl and (CN)2Cbi uptake and assimilation systems in *M. smegmatis*, even though the proteins and mechanisms involved are uncharacterized. In Gram negative bacteria, extracellular cobamides are trafficked into cells by the BtuB-TonB-ExbBD-BtuF-BtuCD system (Noinaj *et al*., 2010). Unlike *M. tuberculosis*, *M. smegmatis* is one of the few mycobacteria that encode homologues of *btuF*, *btuC*, and *btuD*. In *M. smegmatis*, these genes occur as co-operonic partners: *MSMEG_6064* (*btuF*), *MSMEG_6063* (*btuC*), and *MSMEG_6062* (*btuD*) (**Figure 4.12A**; **Table 4.3**) (Novichkov *et al*., 2013). The *M. smegmatis* genome also encodes paralogs of these three genes in a separate operon which comprises two copies of *btuF2* (*MSMEG_4561* and *MSMEG_4560*), *btuC2* (*MSMEG_4559*) and *btuD2* (*MSMEG_4557*). Cobalamin riboswitches are predicted upstream of *MSMEG_6064* and *MSMEG_4561*, which are the first genes in their respective operons (**Figure 4.12A**; **Table 4.3**). In addition to these proteins, two paralogous genes (*MSMEG_3655/MSMEG_4380*) encoding homologues of the *M. tuberculosis* BacA are also found in the *M. smegmatis* genome (**Figure 4.12A**; **Table 4.3**). These proteins have been predicted to associate with the inner membrane of the mycobacterial cell envelope (Hermann *et al.*, 2018).

Table 4.3: **Putative cobamide transporters in *M. smegmatis***

Symbol	**Annotation** (Kapopoulou, Lew & Cole, 2011)	***M. smegmatis* gene number**
bacA	ABC transporter, permease/ATP-binding protein	*MSMEG_3655; MSMEG_4380*
btuF*	Vitamin B_{12} ABC transporter, B_{12}-binding component	*MSMEG_6064*
btuC	Vitamin B_{12} ABC transporter, permease component	*MSMEG_6063*
btuD	Vitamin B_{12} ABC transporter, ATPase component	*MSMEG_6062*
btuF2* (*MSMEG_4561*)	Vitamin B_{12} ABC transporter, B_{12}-binding component	*MSMEG_4561; MSMEG_4560*
btuC2	Vitamin B_{12} ABC transporter, permease component	*MSMEG_4559*
btuD2	Vitamin B_{12} ABC transporter, ATPase component	*MSMEG_4557*

*** Cobalamin riboswitches are predicted upstream of these genes (Novichkov *et al*., 2013).**

4.2.9 Riboswitch control of putative cobamide transport in *M. smegmatis*

Because putative cobalamin riboswitches occurred upstream of the *btuFCD* and *btuF2C2D2* operons, the expression levels of these proteins in the presence or absence of CNCbl were assessed using targeted MS. To this end, membrane fractions of wild-type and Δ*metE*Δ*cobK*::*hyg* strains grown in methionine-supplemented media were prepared and digested with trypsin to generate unique peptides. Tryptic peptides corresponding to BtuF and BtuF2 (MSMEG_4561) were then detected and quantitated by MS and analysed with the Skyline software. As a control for a non-riboswitch associated transporter, the expression profile of one of the two BacA paralogs (MSMEG_3655) was also assessed. MSMEG_4380 was not analysed since the reference spectral library did not contain any peptides for this protein. Analysis showed that all three proteins had higher peak intensities in the Δ*metE*Δ*cobK*::*hyg* strain than in the wild-type (**Figure 4.12B**). The peak intensities of BacA, BtuF and BtuF2 (MSMEG_4561) were 1.4×, 3.5×, and 9.4× higher in Δ*metE*Δ*cobK*::*hyg* than in wild-type, respectively (**Figure 4.12B**). The significantly lower levels of BtuF and BtuF2 (MSMEG_4561) in the wild-type strain – a *de novo* cobamide producer – were suggestive of riboswitch-based downregulation of these proteins. That is, the inferred overexpression of these putative transporters in Δ*metE*Δ*cobK*::*hyg* – which lacks *de novo* cobamide – indicated de-repression and reflected the demand for cobalamin in this mutant.

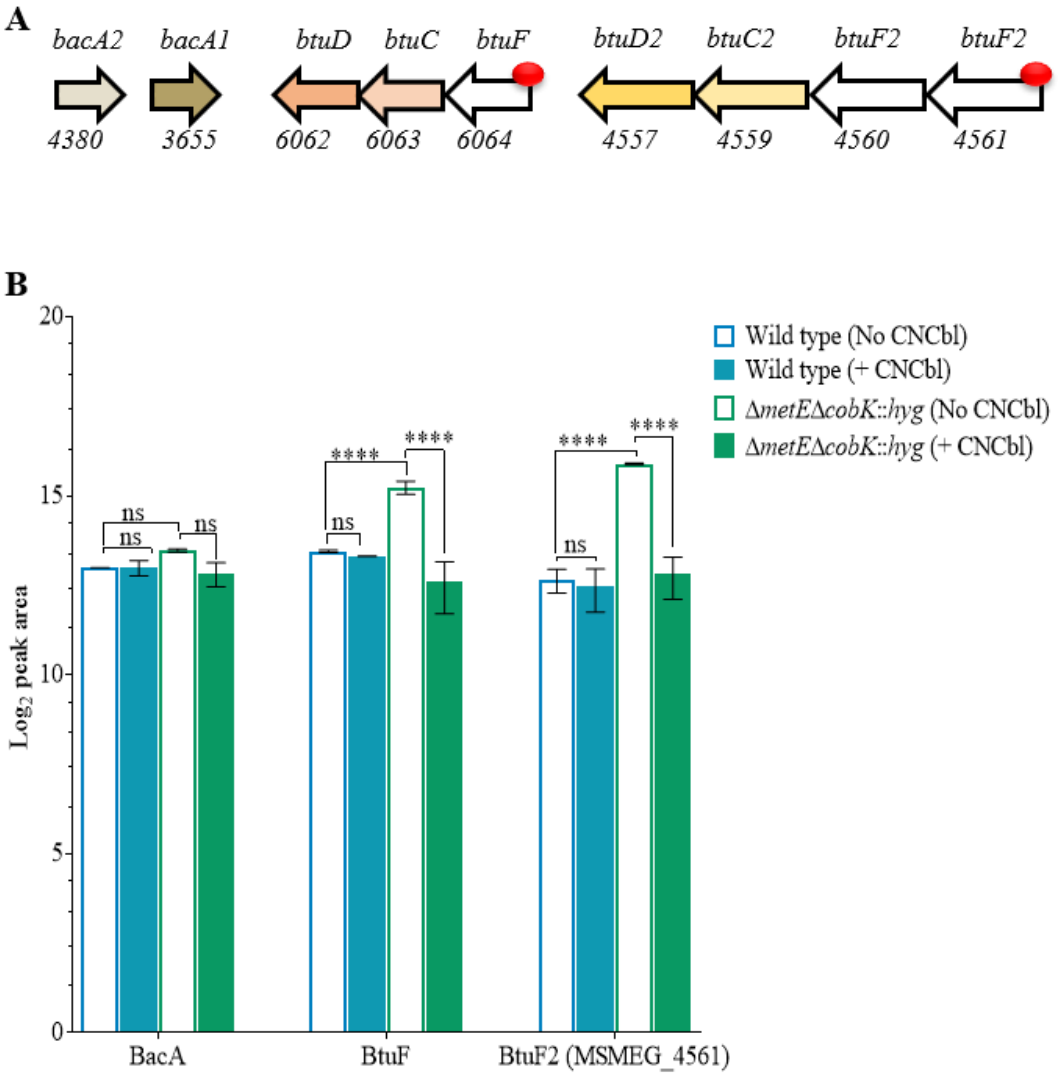


Figure 4.12: **Protein expression profiles of BacA, BtuF and** BtuF2 (MSMEG_4561) **in wild-type or Δ*metE*Δ*cobK*::*hyg***. **A**. A schematic representation of *bacA* genes and the *btuFCD* and *btuF2C2D2* operons in *M. smegmatis*. Cobalamin riboswitches (red spheres) are predicted upstream of *btuF* and *btuF2* (*MSMEG_4561*). Gene numbers in *M. smegmatis* are shown below each gene name. **B**. Targeted MS analysis of BacA, BtuF and BtuF2 (MSMEG_4561) protein expression. Log_2 peak intensities of BacA, BtuF and BtuF2 (MSMEG_4561) in cell wall fractions of wild type and Δ*metE cobK*::*hyg* strains cultured in media containing 10μg/ml CNCbl or without CNCbl (No CNCbl). Similar levels of BacA were observed in the presence or absence of CNCbl in both wild type and Δ*metE cobK*::*hyg* strains. In the absence of CNCbl, BtuF and BtuF2 (MSMEG_4561) expression was higher in the Δ*metE cobK*::*hyg* strain than in the wild type (p<0.0001; (****); ordinary two-way ANOVA). CNCbl strongly repressed BtuF and BtuF2 (MSMEG_4561) in the Δ*metE cobK*::*hyg* strain (p<0.0001; (****); ordinary two-way ANOVA) but did not affect the levels of both proteins in the wild type strain. The data shown represent two independent experiments. The error bars show the standard error of the mean. ns – not significant.

When CNCbl was used as a supplement instead of methionine, the peak intensities of the three proteins were unchanged in the wild-type strain (**Figure 4.12B**). The observation that BtuF and BtuF2 (MSMEG_4561) peak intensities were not affected implied that exogenous CNCbl had not resulted in more repression of these two riboswitch-associated proteins; *i.e.*, in addition to repression by endogenously produced cobalamin. Another explanation for this result was that further repression had occurred, but the change was probably too small to be detected by this analysis. As expected, protein levels of BacA, which does not have a predicted riboswitch, were unaffected by CNCbl (**Figure 4.12B**).

In contrast, BtuF and BtuF2 (MSMEG_4561) levels were reduced in Δ*metE*Δ*cobK*::*hyg* to levels similar to those observed in wild-type (**Figure 4.12B**). The ability of exogenous CNCbl to repress the BtuF/ BtuF2 (MSMEG_4561) proteins indicated the presence of functional cobalamin riboswitches controlling their respective operons. However, it seemed that cobamide uptake was not affected despite the observed riboswitch-based repression of the BtuFCD/BtuF2C2D2 operons because Δ*metE*Δ*cobK*::*hyg* was still able to grow on exogenous CNCbl, possibly because the level of uptake via BtuFCD/BtuF2C2D2 was still sufficient to support growth even under repressed conditions. In *M. smegmatis*, there are two homologues of BacA sharing strong amino acid sequence identities between each other as evaluated by standard BLASTP alignment (Altschul *et al.*, 1997). Both putative *M. smegmatis* BacA have been reported to associate with the inner membrane of the cell envelope (Hermann *et al.*, 2018), supporting their potential roles as cobamide transporters. Therefore, the ability of Δ*metE*Δ*cobK*::*hyg* to grow in exogenous CNCbl, despite the repression of BtuF and BtuF2 (MSMEG_4561), might have also been because CNCbl uptake was unaffected due to the presence of the two copies of *M. smegmatis* BacA.

4.3 Discussion

Genomic analyses revealed that *M. smegmatis* has the genetic potential for *de novo* cobamide biosynthesis, encoding nearly 30 enzyme-catalysed steps of the the *de novo* cobamide biosynthetic pathway. Prior to this study, *de novo* production of cobalamin in *M. smegmatis* had been inferred from indirect evidence only, notwithstanding the widespread use of this organism as a mycobacterial model. Several decades ago, *de*

novo cobamide production in *M. smegmatis* was suggested when whole cell lysates and culture filtrates of this organism supported the growth of *Lactobacillus leichmannii*, a cobamide auxotroph (Karasseva, Weiszfeiler & Lengyel, 1977). Another piece of indirect evidence pointing to the presence of endogenous cobamides in *M. smegmatis* came from an experiment showing that mutants lacking genes involved in the methylcitrate pathway, which is essential for catabolizing propionate, used the cobalamin-dependent methylmalonyl pathway to deal with the toxicity of the stereoisomers of 2-methylcitrate, the by-product of propionate produced by 2-methylcitrate synthase (PrpC; EC 2.3.3.1) (Upton & McKinney, 2007). This chapter also presented direct evidence for *de novo* biosynthesis of cobalamin and cobalamin-dependent regulation of methionine biosynthesis in *M. smegmatis*. The experimental approach involved a combination of genomics, genetics, analytical chemistry, and targeted profiling of gene and protein expression. The deletion of CobK, a precorrin-6x reductase, eliminated *de novo* cobamide production, validating the functionality of this gene in *M. smegmatis*. Further experiments are necessary to to verify the functions of other putative proteins in the *de novo* cobamide biosynthesis and salvage pathways in this organism.

The functionality of $(CN)_2$Cbi salvage in *M. smegmatis* was confirmed by the data showing that cobalamin could also be produced using $(CN)_2$Cbi as a precursor. Although the metabolite analysis described in this chapter was insufficient to fully characterize the chemical structure of the *de novo* synthesised cobamide in *M. smegmatis*, the LC-MS/MS method provided clues about the structure of the cobamide molecule produced by this organism, identifying cobalamin as the likely cobamide form in this organism and DMB as the likely lower base. Whether *M. smegmatis* produces other cobamide forms or whether DMB is the only lower base or ligand, was beyond the scope of the work described in this thesis and remains the subject of future investigation. The LC-MS/MS data presented herein showed that deleting 80% of the coding sequence of *cobK* (**Figure 4.2**), encoding a precorrin-6a reductase involved in corrin ring synthesis, completely abolished *de novo* cobalamin biosynthesis. Similar LC-MS/MS analysis of Δ*metE*Δ*cobK*::*hyg*, in which the *cobK* gene was interrupted by a *hyg* fragment, also confirmed that the inactivation of *cobK* was sufficient to eliminate *de novo* cobalamin biosynthesis. However, when $(CN)_2$Cbi was provided exogenously, the corrin ring step could be bypassed, restoring cobalamin biosynthesis

in this bacterium. The mechanism by which corrinoids are assimilated in mycobacteria is still unknown, and mycobacterial homologues of BtuM, a CNCbl decyanase described in other organisms (Rempel *et al*., 2018) are as-yet unidentified.

The utilization of exogenous cobamides by mycobacteria is widely reported in the literature (Gopinath, Venclovas, *et al*., 2013; Boot *et al*., 2016; Guzzo *et al*., 2016; Lawrence *et al*., 2018). Targeted proteomics confirmed that *M. smegmatis* expresses homologues of classic BtuFCD-type cobamide transporters (Noinaj *et al*., 2010) as well as the atypical transport protein BacA (Gopinath, Venclovas, *et al*., 2013). The presence of these proteins in *M. smegmatis* raises the possibility of the existence of an active transport system for cobamides akin to the Ton-dependent system found in Gram negative bacteria. However, the *M. smegmatis* BtuFCD proteins have been found to associate with the inner membrane in this organism (Hermann *et al.*, 2018). The proteins involved in ferrying cobamides through the mycobacterial outer membrane and the periplasmic space still await discovery.

It was clear that corrinoid uptake in *M. smegmatis* was poor except under strict methionine auxotrophy. Cobalamin riboswitches are predicted upstream of BtuF and BtuF2 (MSMEG_4561), each of which occur in the 5' end of three-gene operons with BtuCD and BtuC2D2, respectively. Proteomics data established that these riboswitches are functional, suppressing their cognate proteins in the presence of cobalamin. This riboswitch-based control might explain, albeit partially, the poor uptake of CNCbl although it could not explain the apparently poor transport of $(CN)_2$Cbi, which would not be a natural ligand for these riboswitches. It is unclear why this organism also expressed BacA alongside BtuFCD/BtuF2C2D2 transporters, although the fact BacA has been linked to drug resistance (Domenech *et al*., 2009) supports the view that cobamide transport might be a secondary role for this protein.

Due to the presence of multiple candidate cobamide transporters in *M. smegmatis*, complex genetic manipulations that were beyond the scope of this thesis would be required to selectively or collectively inactivate them for more in-depth characterization. However, the preliminary data presented herein signify important first steps to understanding mycobacterial acquisition of cobamides. Follow-up work is planned to further explore putative cobamide transporters in *M. smegmatis* using

inducible knockdown platforms that are amenable to multiplexing such as the recently described mycobacterial CRISPRi system (Rock *et al*., 2017).

In this chapter, the notion that the risk of methionine starvation was a major driver of *in vitro* uptake and assimilation of CNCbl and $(CN)_2$Cbi was proposed. The data supporting this hypothesis stemmed from observations that the Δ*metE*Δ*cobK*::*hyg* strain, which required CNCbl for growth, more efficiently converted $(CN)_2$Cbi to cobalamin than the Δ*cobK* strain. Because MetE was active in the Δ*cobK* strain, the requirement for CNCbl was obviated. However, when *metE* was also deleted in the background of a disrupted *cobK*, generating Δ*metE*Δ*cobK*::*hyg*, methionine auxotrophy occurred because the absence of cobalamin deactivated MetH. Consequently, a need for externally-sourced co-factor was created, which subsequently triggered corrinoid salvage.

The cobalamin riboswitch upstream of *metE* was shown to be functional, and transcriptionally controlled this gene. It was clear that although the suppression of *metE* by endogenous cobalamin was substantial, it was incomplete, and some gene expression could still be detected. Furthermore, exogenous CNCbl only marginally repressed *metE* presumably due to poor corrinoid uptake in *M. smegmatis*. This phenotype was in sharp contrast to that described in *M. tuberculosis*, in which CNCbl has been shown to be readily transported *in vitro* (Warner *et al*., 2007). This difference between *M. smegmatis* and *M. tuberculosis* could be related to their unique environmental adaptations. Presumed to be cobamide-deficient, *M. tuberculosis* may have evolved to fully rely on exogenous cobamides, possibly via a constitutive uptake machinery. That *M. tuberculosis* appears to have dispensed with the riboswitch-controlled BtuFCD protein complex supports this theory. In this regard, it will be instructive to fully elucidate the proteins and mechanisms involved in regulating cobamide utilization during infection.

5 Conditional essentiality of the gene encoding the cobalamin-dependent methionine synthase, MetH, in *M. smegmatis*

5.1 Overview

In the previous chapter, it was demonstrated that the absence of cobalamin resulted in higher *metE* gene expression due to the de-repression of the *metE* riboswitch whereas in the wild-type *M. smegmatis*, constitutive cobalamin biosynthesis kept *metE* levels low. Accordingly, it was hypothesized that the repression of *metE* caused *M. smegmatis* to rely solely on MetH for methionine biosynthesis. In this chapter, the conditional essentiality of *metH* is defined based on loss of cell viability following the inactivation of the *metH* gene. In light of recent findings linking MetH and cobalamin to the intrinsic resistance of *M. tuberculosis* to sulphonamides (Guzzo *et al*., 2016), the implications of *metH* essentiality for antimycobacterial drug discovery research are highlighted and discussed.

5.2 Results

5.2.1 Allelic exchange mutagenesis fails to produce a *metH* knockout in cobalamin–replete *M. smegmatis*

Methionine is essential for bacterial viability owing to its multiple metabolic roles, including the initiation of translation and the synthesis of S-adenosylmethionine (SAM), a cellular donor of methyl groups (Ferla & Patrick, 2014; Ferguson, 2016). In *M. smegmatis*, the last step in the biosynthesis of methionine is catalysed by two methionine synthases, MetE and MetH (**Figure 5.1**). In the preceding chapter, it was shown that the elimination of cobamide biosynthesis resulted in the upregulation of MetE, suggesting that *metE* must be under constitutive repression by *de novo* synthesised cobalamin in wild-type *M. smegmatis*. However, because Co(I)cobalamin is a co-factor for MetH (Gonzalez *et al.*, 1992), the observed deficiency in MetE activity in the wild-type strain would not impede methionine biosynthesis. This observation led to the hypothesis that if *M. smegmatis metH* were deleted, it would

effectively deplete *de novo* methionine biosynthesis because of the cobalamin-mediated suppression of *metE* expression.

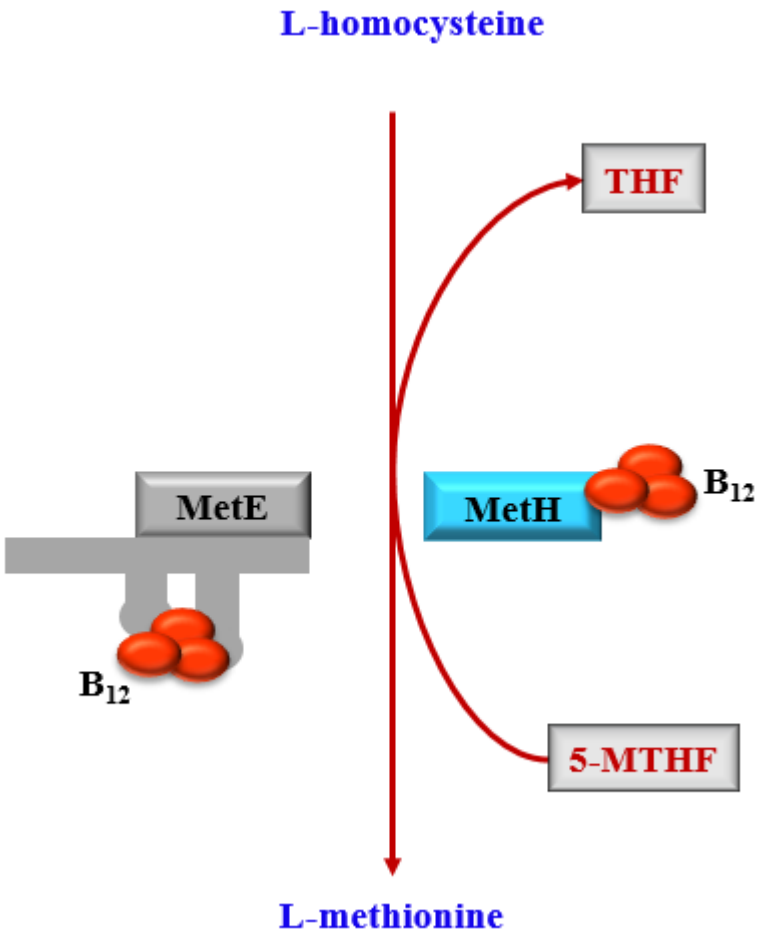


Figure 5.1: **The final step in the *de novo* methionine pathway in *M. smegmatis***. The transfer of a methyl group from 5-methyl-THF (5-MTHF) to homocysteine resulting in methionine and tetrahydrofolate (THF) as a by-product is catalysed by either MetE, the cobalamin-independent methionine synthase, or MetH, with Co(I)cobalamin as a co-factor. MetE is under constitutive repression by a cobalamin riboswitch.

To test this premise, an attempt was made to generate an unmarked *metH* knockout (KO) mutant of *M. smegmatis* in either a cobalamin-replete (wild-type) or –deficient (Δ*cobK*) background using the standard two-step allelic exchange mutagenesis (Parish & Stoker, 2000; Gopinath, Warner & Mizrahi, 2015). First, a suicide vector carrying the *metH* deletion construct (**Figure 5.2**) was inserted into wild-type *M. smegmatis* by electroporation. It was feared that the strong repression of *metE* by *de novo* synthesised cobalamin as demonstrated in the previous chapter, would counter-select against *metH*

KO mutants because *metH* deletion would essentially eliminate all *M. smegmatis* methionine synthases, depleting the bacilli of methionine. Therefore, to increase the likelihood of recovering *metH* KOs, selection plates were supplemented with exogenous methionine.

Out of 154 putative double crossover (DCO) colonies that were screened by PCR, none bore the Δ*metH* allele. It was possible that this was a result of the inherently low recombination efficiency of the allelic exchange method (Gopinath, Warner & Mizrahi, 2015). Therefore, it was thought that by using a *hyg*-marked Δ*metH*::*hyg* construct (**Figure 5.2**), recombination efficiency would be improved by essentially "forcing" the recovery of colonies bearing the Δ*metH*::*hyg* allele. Again, selection plates were supplemented with methionine. With the construct carrying the deletion-insertion allele, 60 sucrose resistant DCOs were screened, all having the wild-type allele and were presumed to have acquired resistance to hygromycin spontaneously or via rearrangement of the allelic exchange vector. The failure to delete *metH* despite methionine supplementation hinted that this gene was essential for growth in the wild-type *M. smegmatis* background due to the repression of MetE in the presence of *de novo* synthesised cobalamin in this background. In the previous chapter, it was observed that the minimal concentration of methionine that supported the growth of the Δ*metE*Δ*cobK*::*hyg* strain – a cobalamin/methionine auxotroph – was 1μM (**Figure 4.1B**), 1000-fold less than the concentration provided exogenously in this experiment. However, as the growth rate of the the Δ*metE*Δ*cobK*::*hyg* strain was the same as that of the wild type strain supplemented with 1mM methionine (**Figure 4.1C**), this concentration was not harmful to the bacilli. It was assumed, therefore, that although exogenous methionine was taken up by the bacilli, this uptake could not keep up with the demand for methionine and hence Δ*metH* or Δ*metH*::*hyg* colonies could not be recovered in a cobalamin-replete strain. To confirm this possibility, an attempt was made to delete this gene in the Δ*cobK* strain, a cobalamin-deficient background.

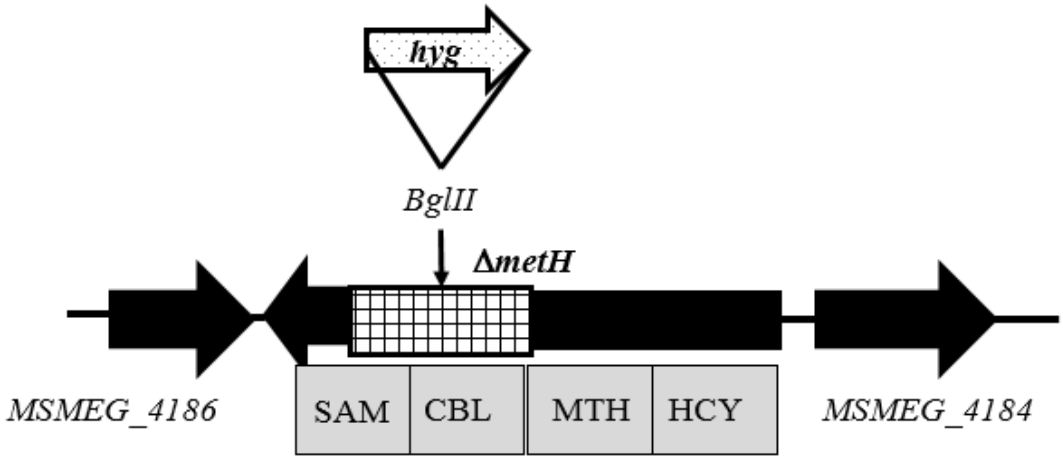


Figure 5.2: **The deletion strategy for generating the *metH* knockout**. The deleted region (chequered pattern) eliminated the entire cobalamin binding domain (CBL) and partially disrupted the SAM binding domain (SAM). A *hyg* fragment was inserted into a *Bgl*II site in the original Δ*metH* plasmid to generate a marked construct. MTH – methyltetrahydrofolate binding domain; HCY – homocysteine binding domain

Following the failure to isolate either Δ*metH* or Δ*metH*::*hyg* strains in the wild-type background, the deletion construct was electroporated into the Δ*cobK* strain and resulting transformants were screened for the Δ*metH* allele. In this case, exogenous methionine was not provided in selection plates because it was anticipated that the lack of cobalamin in Δ*cobK* would alleviate *metE* repression and provide an alternative route (via MetE) to *de novo* methionine biosynthesis. In the Δ*cobK* background, Δ*metH* strains were readily isolated: PCR screening verified the Δ*cobK*Δ*metH* genotype in 9 of 34 colonies tested (**Figure 5.3**). In all cases, the Δ*metH* genotype was validated by Sanger sequencing.

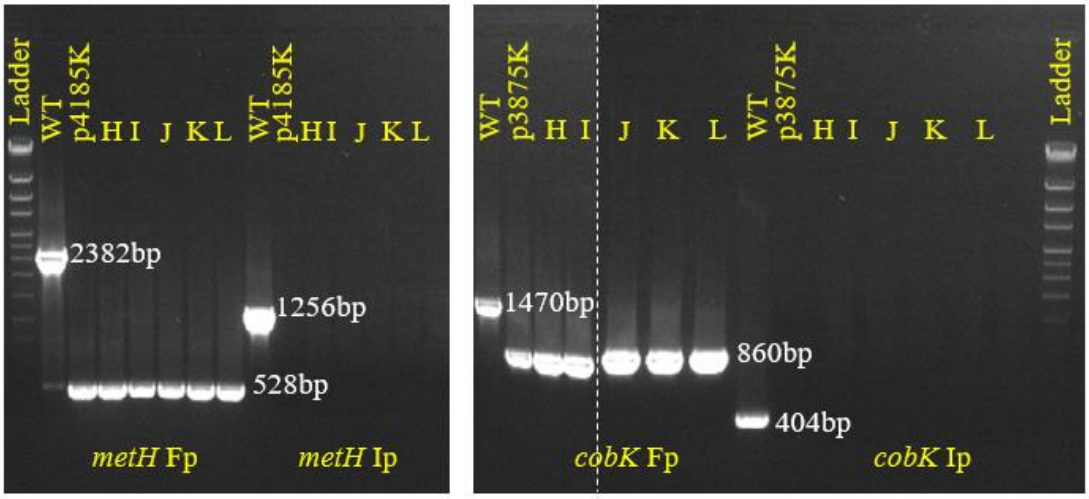


Figure 5.3: **PCR confirmation of *metH* deletion in the Δ*cobK* background**. Flanking (Fp) and internal (Ip) primers were designed to amplify outside or inside the deleted genes. Six out of 9 positive Δ*cobK*Δ*metH* clones (H-L) are shown. Flanking primers yielded PCR products of expected sizes (*metH* – 2382bp; *cobK* – 1470bp). When using internal primers, PCR products of expected sizes were observed in wild-type (*metH* – 1256bp; *cobK* – 404bp) but were missing in the plasmids or clones. The dashed line shows the break between images merged from two distict gels.

5.2.2 Lack of *de novo* cobalamin biosynthesis in the Δ*cobK*Δ*metH* strain

De novo production of cobalamin and the degree to which $(CN)_2$Cbi was assimilated by the Δ*cobK*Δ*metH* strain was assayed by LC-MS/MS analysis as described (chapter 2, *Materials and Methods*). As expected, cobalamin could not be detected in extracts of the Δ*cobK*Δ*metH* strain (**Figure 5.4A**). As was the case with the parental Δ*cobK* strain, *de novo* synthesised cobalamin was also not detectable in extracts of the Δ*cobK*Δ*metH* strain when the organism was cultured in the presence of $(CN)_2$Cbi (**Figure 5.4B**). As described in chapter 3, exogenous CNCbl was shown to result in modest repression of *metE* transcription. It was therefore hypothesised that, in the context of *metH* deletion, this partial repression would reduce internal methionine levels and possibly retard *M. smegmatis* growth. To investigate whether this was the case, the Δ*cobK*Δ*metH* strain was cultured in media containing CNCbl. However, exogenous CNCbl did not affect growth of the Δ*cobK*Δ*metH* strain in batch culture when compared to an unsupplemented control. Similarly, there were no differences in growth on solid media between CNCbl-supplemented and controls cells (**Figure**

5.4C). Furthermore, supplementation with $(CN)_2$Cbi had no effect on growth (**Figure 5.4C**), consistent with the observations reported in the previous chapter that the uptake of CNCbl and $(CN)_2$Cbi in *M. smegmatis* was poor.

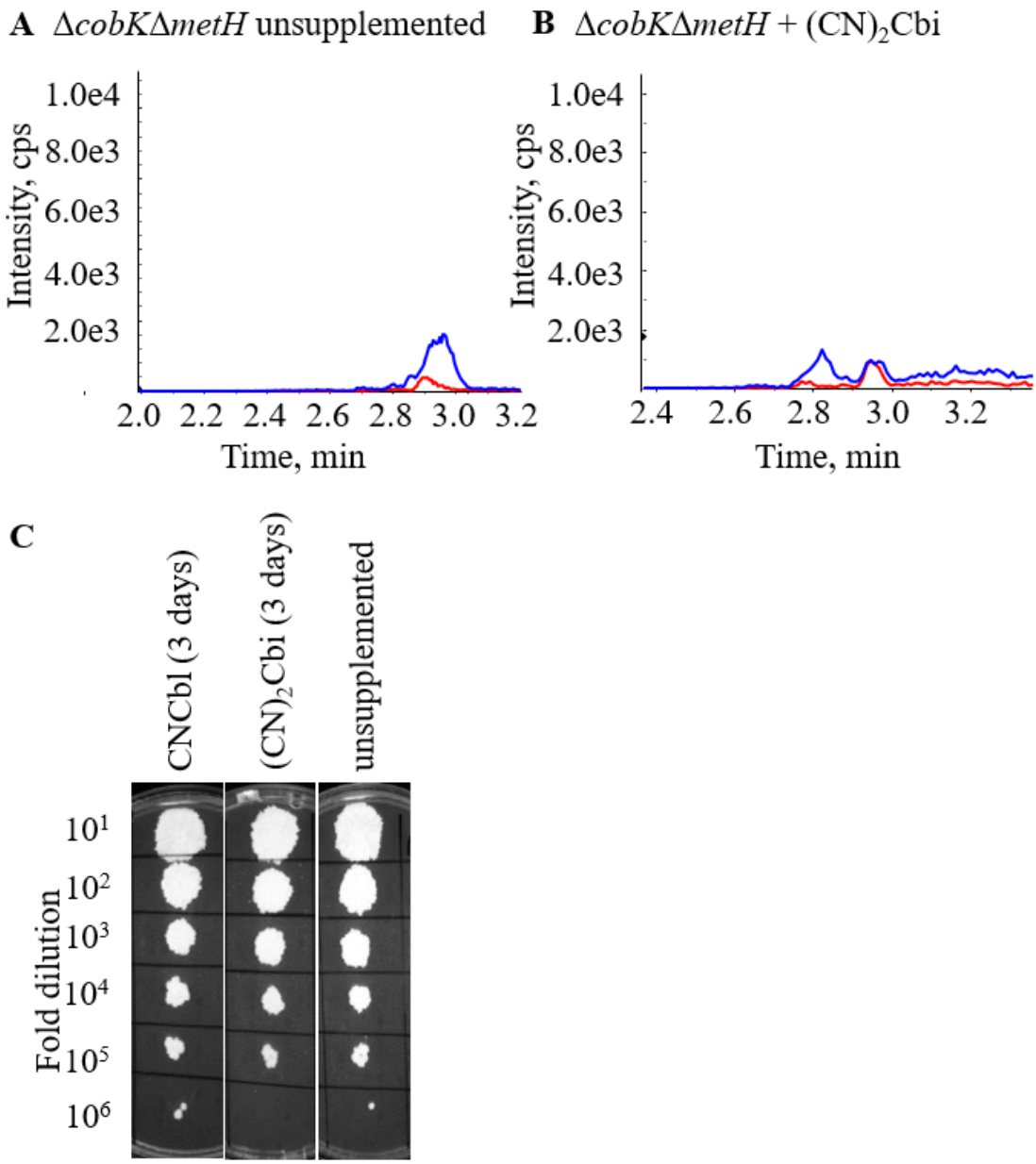


Figure 5.4: **Cyanocobalamin and dicyanocobinamide uptake by Δ*cobK*Δ*metH*. A-B.** LC - MS/MS traces of Δ*cobK*Δ*metH* cell extracts showing the lack of detectable cobalamin production despite Cbi supplementation. **C.** CNCbl and $(CN)_2$Cbi uptake in the Δ*cobK*Δ*metH* strain determined by a spotting assay on 7H10 agar containing CNCbl or $(CN)_2$Cbi. Prior to spotting, cells were grown in 7H9-OADC medium to exponential phase ($OD_{600} \approx 0.6$) followed by 10-fold serial dilutions. Plates were imaged after 3 days' incubation at 37°C.

5.2.3 Riboswitch-mediated regulation of *metE* in the Δ*cobK*Δ*metH* strain

The levels of *metE* transcript in the Δ*cobK*Δ*metH* strain in the presence of *de novo* synthesised cobalamin or exogenous CNCbl were also quantified by ddPCR and compared to gene expression levels in the parental Δ*cobK* strain. Similar to the Δ*cobK* strain, the Δ*cobK*Δ*metH* strain had elevated *metE* gene expression compared to wild-type strain (**Figure 5.5A**). As in the Δ*cobK* parental strain, the levels of *metE* in Δ*cobK*Δ*metH* were also slightly, but not significantly, decreased by exogenous CNCbl and unaffected by exogenous $(CN)_2$Cbi (**Figure 5.5B**).

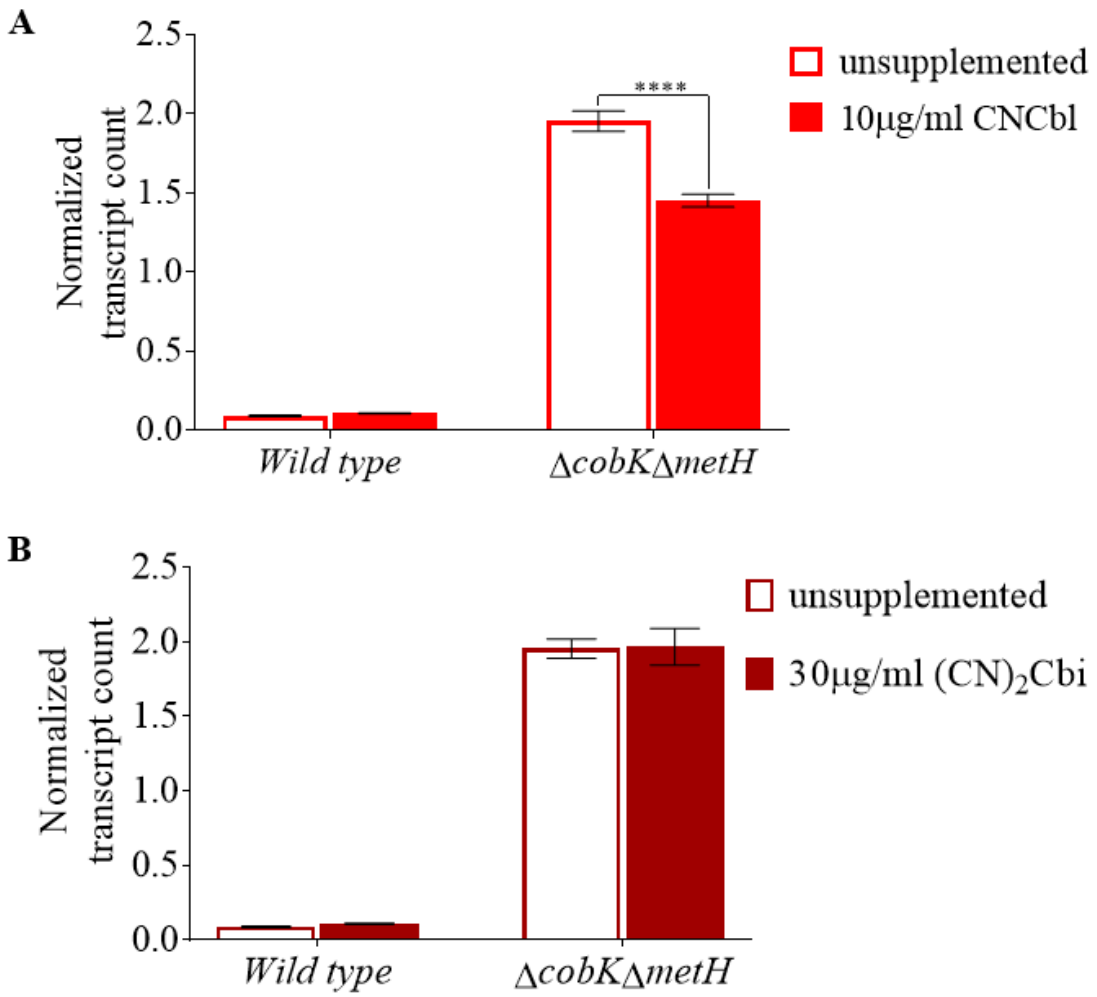


Figure 5.5: **Analysis of *metE* transcription in Δ*cobK*Δ*metH* relative to wild type in the presence of or absence of exogenous CNCbl**. **A.** Transcript copy numbers of *metE* in CNCbl-supplemented wild-type, Δ*cobK* and Δ*cobK*Δ*metH* strains were determined by ddPCR Exogenous CNCbl caused a 1.3× reduction in *metE* transcript levels in Δ*cobK*Δ*metH* (p<0.0001; (****); two-way ordinary ANOVA). **B.** *metE* transcription in Δ*cobK*Δ*metH* relative to wild type in the presence or absence of exogenous $(CN)_2$Cbi. The data shown are representative of two independent experiments. Error bars represent the standard error of the mean.

To find out whether *de novo* synthesised cobalamin had any effect on *metH* expression, levels of *metH* transcript were compared between the wild-type and Δ*cobK* strains. To examine the *metH* expression profile in the context of *metE* deficiency, *metH* transcript levels were analysed in a Δ*metE* strain, which had an in-frame deletion of the entire *metE* ORF (Dr. Stephanie Dawes, unpublished) (**Figure 5.6A**). Across these three strains, similar levels of *metH* transcripts were observed (**Figure 5.6B**), implying that *metH* gene expression was independent of cobalamin availability and *metE* levels.

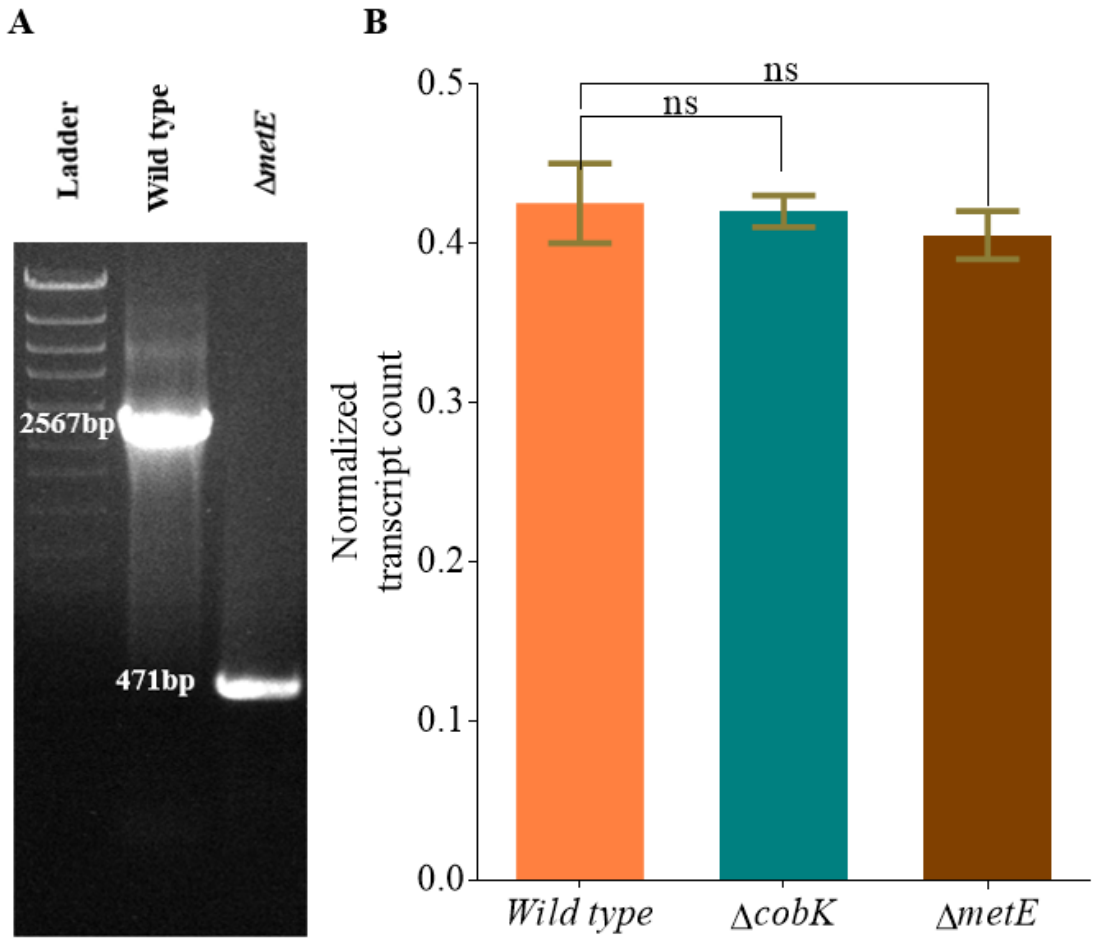


Figure 5.6: **Analysis of *metH* expression profile in *M. smegmatis*. A**. PCR genotyping of the wild type and Δ*metE* strains using the 6638_F/R primer pair targeting the flanking regions of the *metE* gene confirmed the deletion of the *metE* ORF in the Δ*metE* strain, resulting in a 471bp PCR product. The wild-type PCR product was 2567bp. **B**. Relative *metH* gene expression in wild-type, Δ*cobK* and Δ*metE* strains. Absolute copy numbers of *metH* transcript were determined by ddPCR and normalized to the copy number of *sigA* transcript. There was no difference in *metH* expression between the mutants and the wild type strain. PCR image courtesy of Rendani Mbau. ns – not significant (ordinary one-way ANOVA).

It was suspected that the small (but statistically significant) reduction in *metE* transcript due to exogenous CNCbl (**Figure 5.5A**), was related to poor CNCbl uptake. This finding also explained why the Δ*cobK*Δ*metH* strain, which encodes the cobalamin sensitive *metE* gene, grew in culture containing CNCbl. To exclude the possibility that this unexpected lack of sensitivity to CNCbl was due to suppressor mutations in the riboswitch, WGS analysis was performed on this mutant. WGS revealed no mutations in the *metE* riboswitch in the Δ*cobK*Δ*metH* strain, implying that the lack of sensitivity to the exogenous cofactor in this strain was not due to any inactivating mutations in

the *metE* riboswitch. However, one synonymous SNM (s-SNM) and three ns-SNMs that were missing in the parental Δ*cobK* strain were identified in Δ*cobK*Δ*metH* (**Table 5.1**). These mutations were found in three genes (*MSMEG_0619*, *MSMEG_4183* and *MSMEG_5492*) of unknown functions. Based on the genomic context of these genes, their functions were unlikely to be related to cobamide or methionine metabolism. The *M. tuberculosis* homologues encoding proteins with assigned functions were found for *MSMEG_0619* (PPE family protein PPE4) and *MSMEG_5492* (acetyl-/PrP-CoA carboxylase (beta subunit) AccD2) (**Table 5.1**). *MSMEG_4183* did not have any homologues in *M. tuberculosis* but the Mycobrowser portal (Kapopoulou, Lew & Cole, 2011) predicted it to encode a putative phosphoglycolate phosphatase respectively (**Table 5.1**). Because the cloned Δ*metH* fragment (**Figure 5.2**) is immediately adjacent to *MSMEG_4183*, it was assumed that the ns-SNM detected in this gene was likely due to mismatches that might have happened during the repair of double stranded DNA breaks which sometimes occur during homologous recombination.

Table 5.1: **Unique SNMs detected in Δ*cobK*Δ*metH* relative to the parental Δ*cobK* strain**

Gene ID	Product	SNM	Type of SNM	Amino Acid Change
MSMEG_0619	PE/PPE family protein	271G>T	ns-SNM	Val91Phe
MSMEG_4183	phosphoglycolate phosphatase, chromosomal	34C>A	ns-SNM	Leu12Met
MSMEG_5492	acetyl-CoA carboxylase carboxyltransferase (AccD2)	480G>A	s-SNM	Gln160Gln

5.2.4 Mass spectrometric profiling of MetE protein expression

To test the effect of endogenous and exogenous CNCbl on MetE protein expression, MS was used to detect two unique MetE peptides in the wild-type and Δ*cobK* strains grown with or without CNCbl. Consistent with incomplete *metE* repression in the presence of cobalamin (as reported in the previous chapter), MetE peptides were detected in the wild-type strain (**Figure 5.7**). As expected, analysis of absolute peak

intensities showed that MetE protein levels were 3.5× higher in Δ*cobK* than in wild-type because of the riboswitch-based repression by endogenous cobalamin. (**Figure 5.7**). When CNCbl was added to growth medium, MetE peptides could no longer be detected in the wild-type strain, possibly due to an additive repressive effect by exogenous CNCbl. However, in the Δ*cobK* strain, MetE peak intensity was halved in the presence of exogenous CNCbl (**Figure 5.7**). Based on these findings and the gene expression analysis described above, it is plausible that, in *M. smegmatis*, the *metE* riboswitch is not in a natural stoichiometric equilibrium with *de novo* synthesised cobalamin ligand and that increasing the concentration of internal cobalamin (*i.e.*, through supplementation with exogenous CNCbl) led to stronger repression.

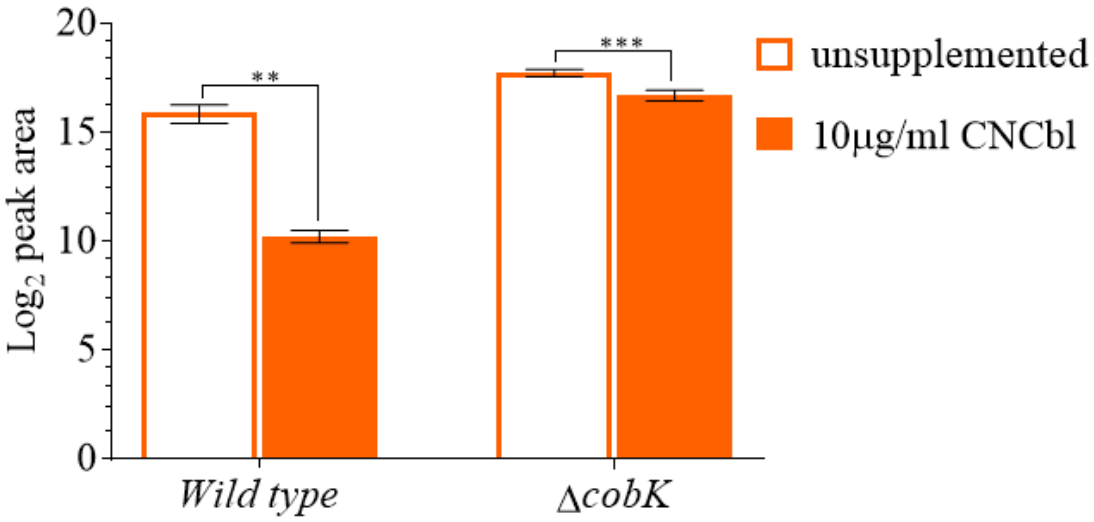


Figure 5.7: **MetE protein expression analysis using mass spectrometry**. Targeted MS analysis of MetE protein expression. Log_2 peak intensities of MetE tryptic peptides in wild type and Δ*cobK* grown in the presence or absence of exogenous CNCbl (CNCbl). MetE peptide levels were more decreased in CNCbl-supplemented wild type (p=0.0071; (**) ordinary two-way ANOVA) than in CNCbl-supplemented Δ*cobK* (p=0.0001; (***); ordinary two-way ANOVA).

5.2.5 Loss of cell viability following transcriptional knockdown of *metH* in *M. smegmatis*

Since it was not possible to generate allelic exchange deletions in *metH* in a wild-type *M. smegmatis* background, a conditional knockdown of this gene was made using the CRISPRi strategy (Rock *et al.*, 2017). For this approach, PLJR962, a Km-selectable,

chromosomally integrating CRISPRi plasmid optimized for *M. smegmatis* (Rock *et al*., 2017) was used. This vector contains an anhydrotetracycline (ATc)-inducible deactivated CRISPR-associated protein-9 nuclease from *Streptococcus thermophilus* (*d*Cas9$_{Sth1}$) as well as an ATc-inducible, target-specific sgRNA handle (**Figure 5.8**). Upon induction by ATc, *d*Cas9$_{Sth1}$ is directed to the target gene by the sgRNA, where the *d*Cas9$_{Sth1}$-sgRNA complex binds a protospacer adjacent motif (PAM) on the DNA. This binding physically blocks the RNA polymerase preventing transcription initiation or elongation (**Figure 5.8**). Transcriptional interference in a gene required for growth would lead to loss of cell viability, which was inferred from the loss of colony formation ability.

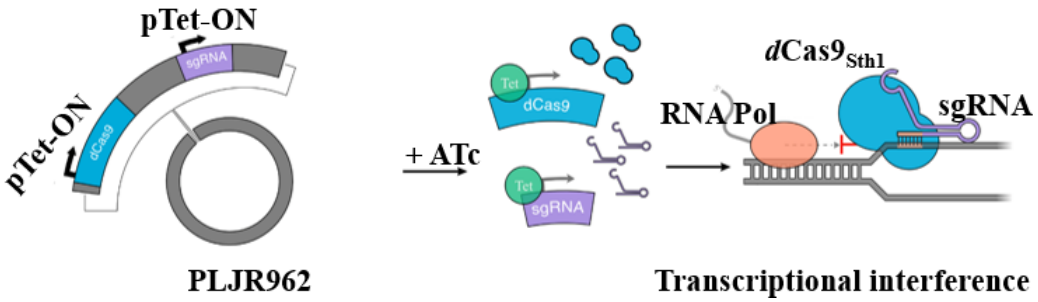


Figure 5.8: **Schematic of the mycobacterial CRISPRi system** (Rock *et al*., 2017). The PLJR962 CRISPRi vector is a kanamycin selectable and contains tetracycline-inducible promoters for the catalytically-dead Cas9 protein from *S. thermophilus* (*d*Cas9$_{Sth1}$) (blue) and a scaffold of short guide RNAs (sgRNA) (purple) specific to target DNA. Anhydrotetracycline (ATc) (green) induces the expression of *d*Cas9$_{Sth1}$ and the sgRNA scaffold, which form a complex and bind target DNA blocking the RNA polymerase (orange) and causing transcriptional interference (indicated by the red ⊢ symbol). Figure adapted with permission from (de Wet *et al.*, 2018)

Rock *et al.* reported that some leakiness in the CRISPRi system could be expected (Rock *et al*., 2017). However, the authors also recommended that, to improve target specificity and minimize leaky gene expression, high concentrations of ATc (100ng/ml) and sgRNAs with appropriate PAM sequences should be used. Therefore, to assess which sgRNAs targeting *M. smegmatis metH* would optimally induce gene silencing, a panel of 13 pairs of sgRNA oligos was tested (**Table 5.2**). Each sgRNA had an assigned PAM score between 1 and 25 in descending order of predicted strength (Rock *et al*., 2017). Among the 13 sgRNAs, the highest PAM score was 11 and the lowest was 1 (**Table 5.2**). The sgRNAs were designed to target random positions from the transcription start site (TSS) of *metH* (**Table 5.2**).

Table 5.2: **sgRNAs targeting the *metH* ORF in *M. smegmatis***

***metH* construct**	**sgRNA ID**	**Distance from TSS**	**PAM Score**
PLJR962_*metH*1	sgRNA1	3603	3
PLJR962_*metH*2	sgRNA2	3060	4
PLJR962_*metH*3	sgRNA3	3057	1
PLJR962_*metH*4	sgRNA4	2934	11
PLJR962_*metH*5	sgRNA5	2904	5
PLJR962_*metH*6	sgRNA6	2242	4
PLJR962_*metH*7	sgRNA7	2184	11
PLJR962_*metH*8	sgRNA8	2002	4
PLJR962_*metH*9	sgRNA9	1791	11
PLJR962_*metH*10	sgRNA10	1785	11
PLJR962_*metH*11	sgRNA11	1743	11
PLJR962_*metH*12	sgRNA12	1737	5
PLJR962_*metH*15	sgRNA15	729	5

The success of the transformation and gene silencing were simultaneously determined by including ATc in the selection plates. An sgRNA for *mmpL3*, a known essential gene involved in mycolic acid biosynthesis, was used as a positive control. As expected, induction of *mmpL3* KD by ATc was lethal (**Figure 5.9**). Interestingly, varying degrees of inhibition of colony formation were observed among the 13 *metH* KD constructs (**Figure 5.9**). Six constructs fully inhibited colony growth, two partially suppressed colony formation, and five showed no differences in colony numbers between induced and uninduced plates (**Figure 5.9**). These variations could have been due to leakiness in the CRISPRi system.

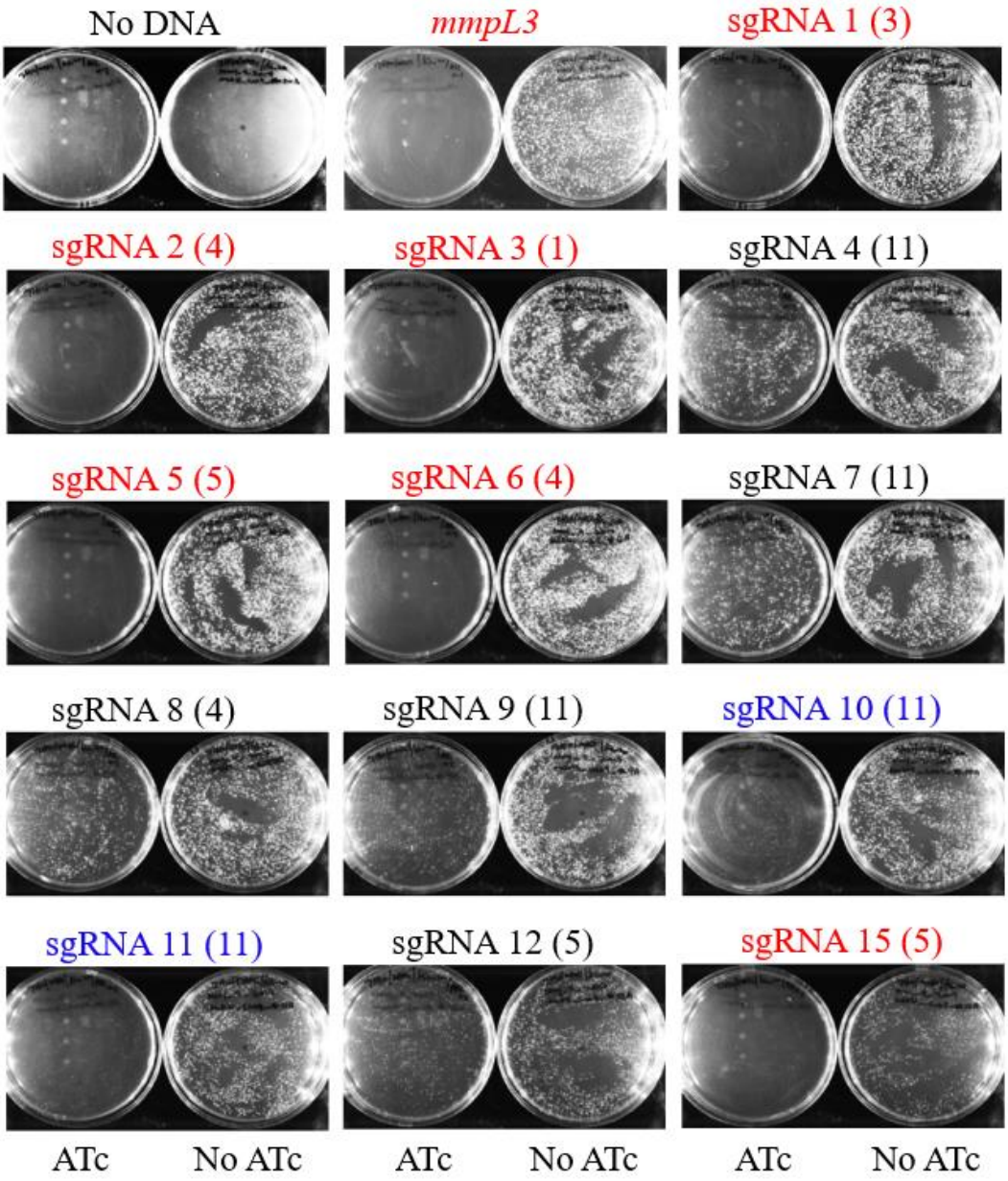


Figure 5.9: **Optimization of CRISPRi knockdown of *metH* in *M. smegmatis*.** The *mmpL3* or *metH* conditional knockdown constructs were electroporated in the wild type *M. smegmatis* background and selected on Km-containing 7H10-OADC agar with (ATc) or without (No ATc) 100ng/ml ATc. ATc-induced CRISPRi gene silencing of *mmpL3* completely inhibited colony formation but the thirteen conditional knockdown constructs of *metH* suppressed colony formation to varying degrees. The sgRNAs that produced complete inhibition of colony formation are highlighted in red, those with partial inhibition in blue, and those with no inhibition in black. The PAM score for each sgRNA is indicated in brackets.

From these results, it was clear that stronger gene silencing was associated with PAM score values closer to 1, *i.e.*, the six constructs that completely inhibited growth (sgRNA1, 2, 3, 5, 6, 15) had PAM scores ranging from 1 to 5 (**Table 5.2**). Out of the

five constructs that failed to suppress growth, three had a PAM score of 11 (sgRNA4, 7, 9), and two had a score of 4 (sgRNA8) and 5 (sgRNA12) (**Table 5.2**). The observed differences in knockdown efficiencies were unlikely to be related to the genomic coordinates of the sgRNA. There was no relationship between the degree of growth suppression and the distance of the sgRNA from transcription start site of *metH*. Using CRISPRi-seq, however, it was observed that higher knockdown efficiency was associated with sgRNAs targeting the first 5% sequences of a gene (*i.e.*, closer to the 5' end) (de Wet *et al.*, 2018).

Notwithstanding the variations in the degree of growth inhibition, these data validated the essentiality of *metH* for the growth of *M. smegmatis*, in line with predictions made independently by CRISPRi-seq, a high throughput knockdown screen developed in the MMRU to target *M. smegmatis* genes which have known *M. tuberculosis* homologs and by Tn-seq analysis (**Figure 5.10**) (de Wet *et al.*, 2018).

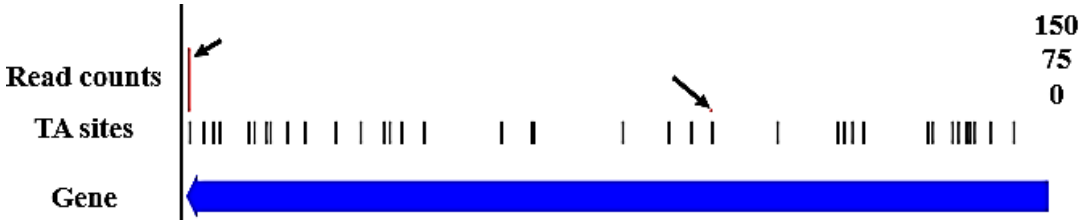


Figure 5.10: **Tn-seq analysis of the essentiality of *M. smegmatis metH*.** Sequenced data were analysed using the Hidden Markov Model (HMM) which is embedded in the TRANSIT software for Himar1 Tn-seq analysis (DeJesus *et al.*, 2015). Read counts for *metH* gene in wild-type *M. smegmatis* were obtained from deep sequencing of Tn insertion libraries. Only 2 insertions (shown in red, indicated by arrowheads) out of the 37 TA sites in this gene (black vertical lines) were found in the *metH* ORF. The data shown represent three biological replicates. Analysis courtesy of Irene Gobe.

5.2.6 Transcriptional interference of the *M. smegmatis metH* causes methionine deficiency

Since *metE* was under cobalamin-dependent repression in *M. smegmatis*, the knockdown of *metH* was expected to effectively inactivate both methionine synthases, resulting in lethal methionine depletion. Given that attempts to generate a Δ*metH* strain

had been unsuccessful in spite of methionine supplementation, the question of whether exogenous methionine would restore colony formation in the *metH* KD was tested. To do this, knockdowns were selected on methionine-containing plates using sgRNA2 and sgRNA3, which reproducibly suppressed colony growth upon CRISPRi induction with ATc. Results showed that, when methionine was included in selection plates, gene silencing with CRISPRi failed to suppress growth in contrast to when there was no methionine (**Figure 5.11**).

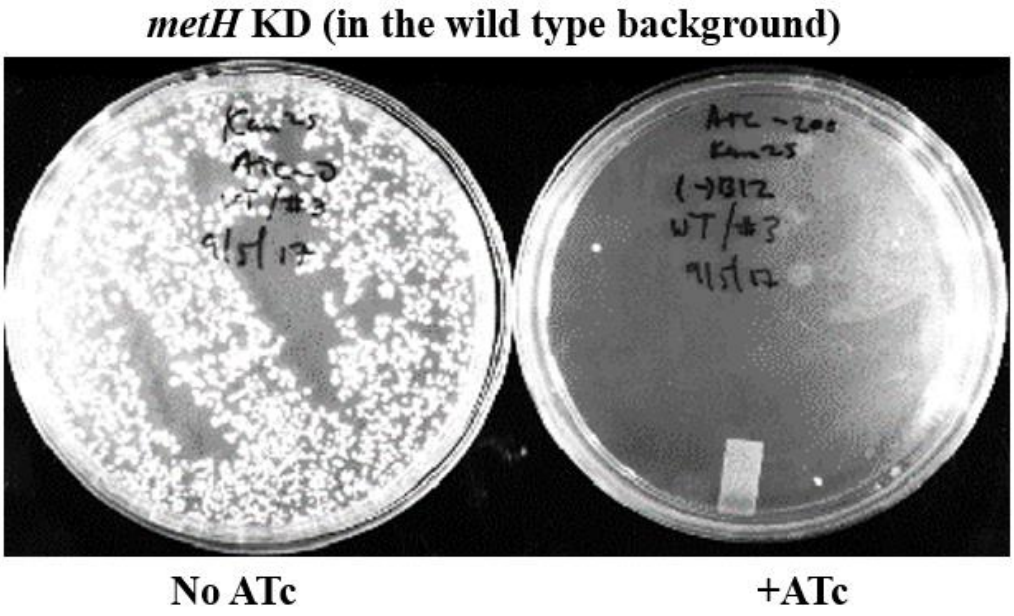


Figure 5.11: **Rescue of the *metH* knockdown in the wild-type background by methionine**. *M. smegmatis* wild-type cells were transformed with *metH* knockdown constructs and selected on plates containing 1mM methionine with ATc (+ATc) or without ATc (No ATc). CRISPRi knockdown of *metH* was growth-inhibitory in *M. smegmatis* but growth could be restored by exogenous methionine.

To determine whether the restoration of viability in *metH* knockdowns was possible at 1μM – the lowest concentration of exogenous methionine that supported the growth of the Δ*metE*Δ*cobK*::*hyg* strain (a methionine auxtroph) (**Figure 5.12A**) – a *metH*

knockdown strain was generated by inserting the *metH* CRISPRi vector in the Δ*metE* backround. It was expected that in the Δ*metE* background, ATc-mediated silencing of *metH* would eliminate the only remaining methionine synthase, resulting in the complete shutdown of methionine biosynthesis. Transformants carrying the *metH* KD construct were plated on agar containing Km, without ATc, and a single colony was then cultured to exponential phase and serially diluted and spotted on media containing ATc and 1μM or no methionine. For comparison, another spotting assay was set up with 10μM methionine. Knocking down *metH* in the Δ*metE* background resulted in growth inhibition in the absence of methionine (**Figure 5.12B**), establishing methionine deficiency as the reason for the loss of viability in *M. smegmatis metH* knockdown strains. Both 1μM and 10μM methionine concentrations restored colony growth but, at the lower concentration, colony sizes were markedly smaller (**Figure 5.12B**).

Considering the inability of exogenous methionine to rescue *metH* KOs, this was also a curious finding. In principle, by knocking down *metH,* the activity of the protein would be gradually reduced. As a result, *metH* KDs in the wild-type background would have had the chance to accumulate methionine to overcome the inactivation of MetH and MetE (due to the repression by *de novo* synthesised cobalamin). In contrast, a *metH* deletion would have eliminated the ORF of the gene, completely abrogating MetH activity. Because of this, the *metH* KOs would survive in the wild-type background, since the demand for methionine could not be met and hence Δ*metH* colonies could not form on selection plates. To verify that there was no residual MetH activity in the *metH* knockdown, one of the spotting assay plates had CNCbl. It was expected that if the CRISPRi system had partially reduced MetH activity, then CNCbl would restore growth of the KD strain in the Δ*metE* background. Exogenous CNCbl was unable to rescue the *metH* knockdown in the Δ*metE* background (**Figure 5.12B**), showing that there was no residual MetH activity. However, the possibility that this result was due to inefficient CNCbl uptake could not be ruled out in this assay.

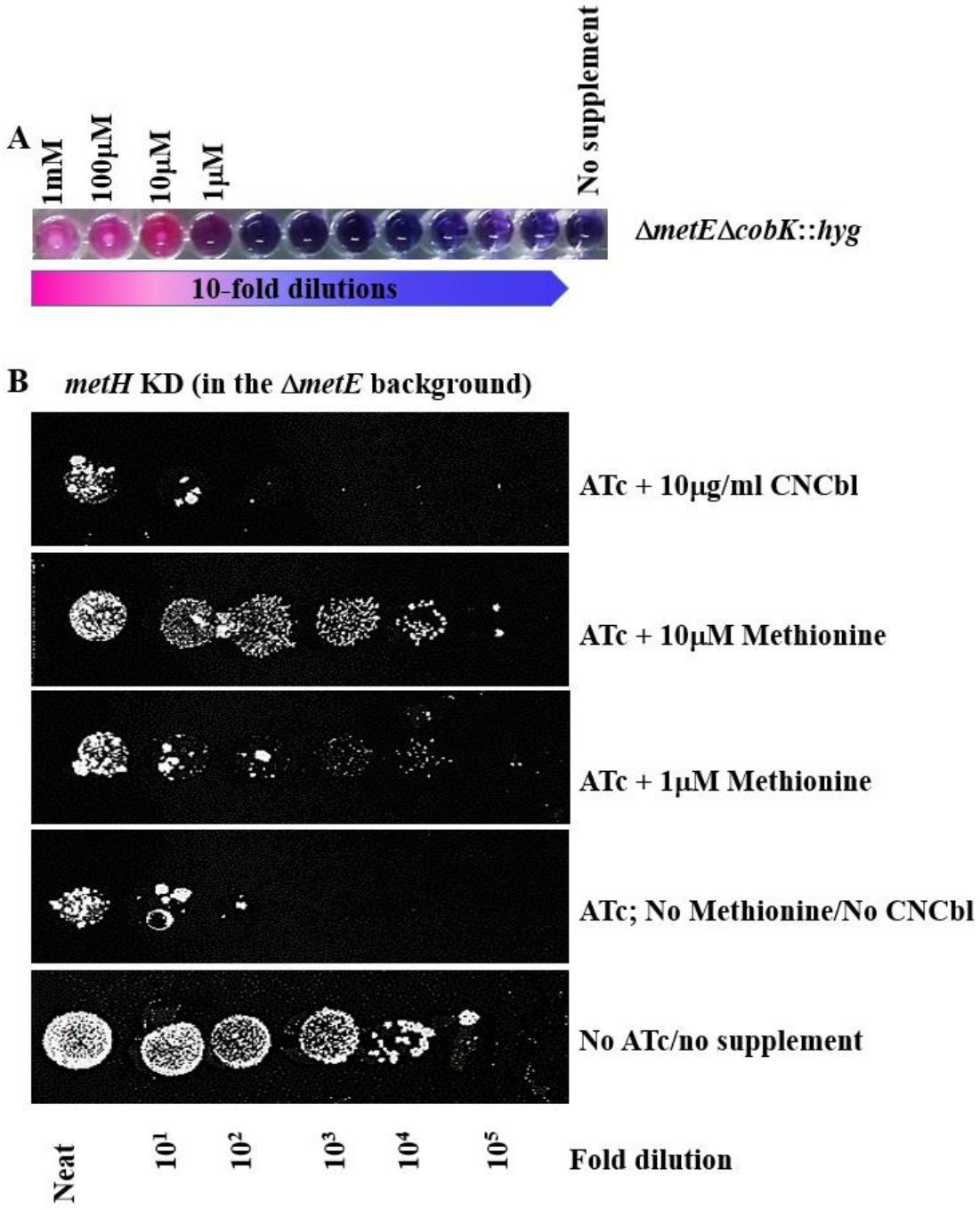


Figure 5.12: **The rescue of *metH* knockdown in the Δ*metE* background by methionine**. **A**. Cell viability assay of the Δ*metE*Δ*cobK*::*hyg* strain, a methionine auxotroph, in the presence or absence of 10-fold dilutions exogenous methionine, as determined using the MABA. The transition from blue to pink indicates cell growth. **B**. Spotting assay of a 10-fold serially diluted exponential-phase culture of the Δ*metE* strain carrying the *metH* knockdown construct. The agar plates contained either methionine or CNCbl and either ATc to induce *metH* silencing or without ATc as a control. The colony clearance in the *metH* KD in the Δ*metE* strain was rescued by 1µM and 10µM methionine.

5.2.7 Essentiality of *M. smegmatis metH* is conditional on cobalamin-mediated *metE* repression

The CRISPRi system was also used to test the hypothesis that the absence of MetE activity and the availability cobalamin were both prerequisites for *metH* to become essential. To achieve this aim, the effect of *metH* silencing was compared between the cobalamin replete (wild-type) and the cobalamin deplete (Δ*cobK*) backgrounds. Since *metE* was de-repressed in the Δ*cobK* strain, allowing the expression of MetE as an alternative methionine synthase, the induction of *metH* knockdown did not suppress growth (**Figure 5.13**), as predicted. However, when CNCbl was included in selection plates, colonies still formed even after inducing the *metH* knockdown (**Figure 5.13**), a phenotype which was reminiscent of the survival of the Δ*metH*Δ*cobK* strain in 7H10-OADC agar containing CNCbl (**Figure 5.4C**), and supported the observations made previously that CNCbl uptake in the Δ*cobK* strain was poor.

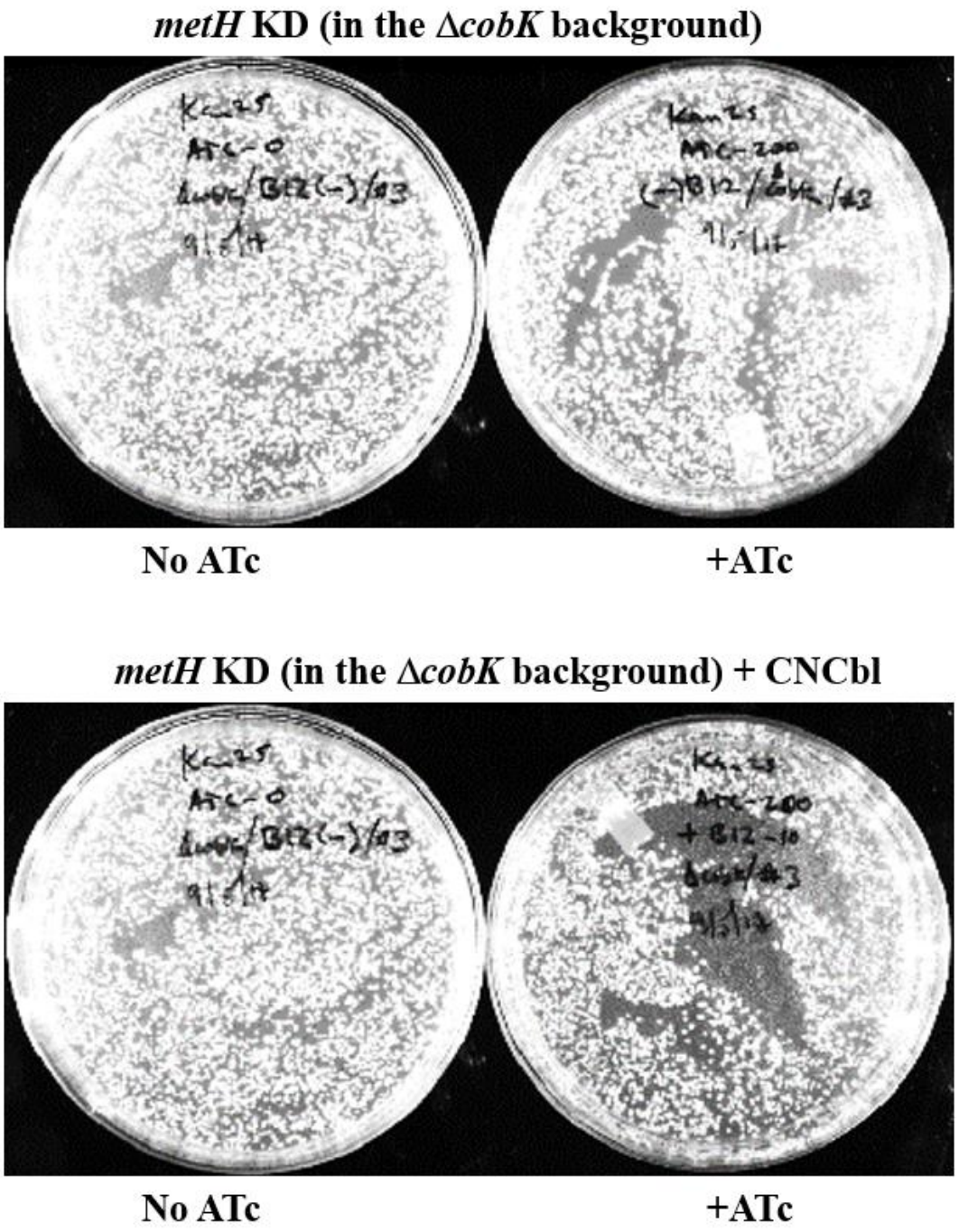


Figure 5.13: **The essentiality of *metH* in *M. smegmatis* is dependent on intact *de novo* cobalamin biosynthesis**. The *metH* CRISPRi construct was transformed into Δ*cobK* strains and selected on kanamycin with ATc (+ATc) or without ATc (No ATc) and with or without CNCbl. Induction of *metH* knockdown in ATc-containing plates did not inhibit colony formation in Δ*cobK*. Addition of CNCbl to ATc plates did not affect colony formation.

The ability of Δ*cobK metH* KDs to survive exogenous CNCbl also implied that the amount of externally-sourced cobalamin in this mutant was too low to repress *metE*

enough to cause methionine depletion. To test the theory, a spotting assay of the *metH* knockdown in the wild-type background was done, in which ATc-containing plates were supplemented with methionine or CNCbl. In the ATc only plates, barely discernible colonies were detected at 100-, 1000- and 10000-fold dilutions (Figure 4.13). However, these faint colonies disappeared completely in plates containing CNCbl (**Figure 5.14**). The disappearance of these faint colonies in the wild-type strain after exposure to exogenous CNCbl endorsed the previously stated idea of incomplete repression of *metE* by *de novo* synthesised cobalamin. This apparently additive effect of exogenous CNCbl on colony suppression also confirmed that CRISPRi-mediated silencing of the *metH* gene was robust; *i.e.*, the faint colonies were not a result of residual MetH activity. Overall, these findings established that the two conditions that made *metH* essential in *M. smegmatis* were the availability of enough internal cobalamin and a strong repression of *metE*.

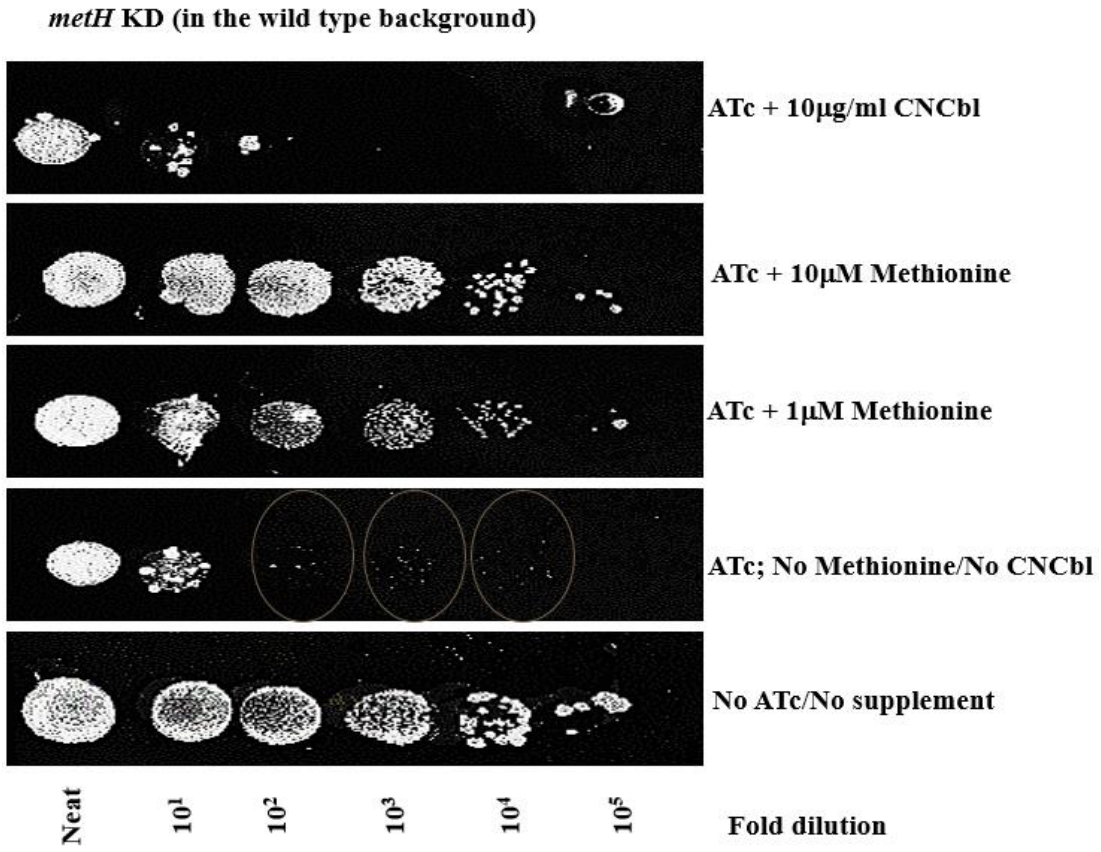


Figure 5.14: **Spotting assays of *metH* knockdowns in the wild-type strain**. An exponential culture of *M. smegmatis* wild-type strain carrying the *metH* knockdown construct was serially diluted and spotted on agar containing methionine or CNCbl with ATc (ATc) or without ATc (No ATc). In the wild-type strain, ATc-induced *metH* knockdown was growth-inhibitory but faint colonies (circled) were visible in the ATc only plates which disappeared in plates containing ATc and CNCbl. Methionine complemented the knockdown phenotype.

5.2.8 Single cell analysis with time-lapse microscopy reveals a growth arrest phenotype in the *metH* knockdowns

The *metH* KD phenotype was analysed at the single-cell level with microfluidics and live-cell imaging using time-lapse phase contrast microscopy. To induce gene silencing, cells were exposed to ATc for 6h (two doublings) prior to running the microfluidics protocol for a total of 43h. For comparison, a control experiment using the uninduced strain (*i.e.* a wild-type background carrying the *metH* KD construct without exposure to ATc) was set up in parallel. Compared to the uninduced control, the *metH* KD proliferated more slowly and cell division stagnated after 24h (**Figure 5.15**; **Supplementary Movies 1-5, supplementary information online**). By 24h,

colonies of the uninduced control had expanded into a lawn of cells occupying the entire field of view (**Figure 5.15**; **Supplementary Movies 1-5, supplementary information online**). In contrast, *metH* KD colony sizes were smaller. The limited expansion of *metH* KD colonies was consistent with methionine depletion and correlated with the previous observations showing that the *metH* KD in uninduced control was unable to form colonies on agar. These observations provided compelling evidence of the essentiality of *metH* for growth in *M. smegmatis*.

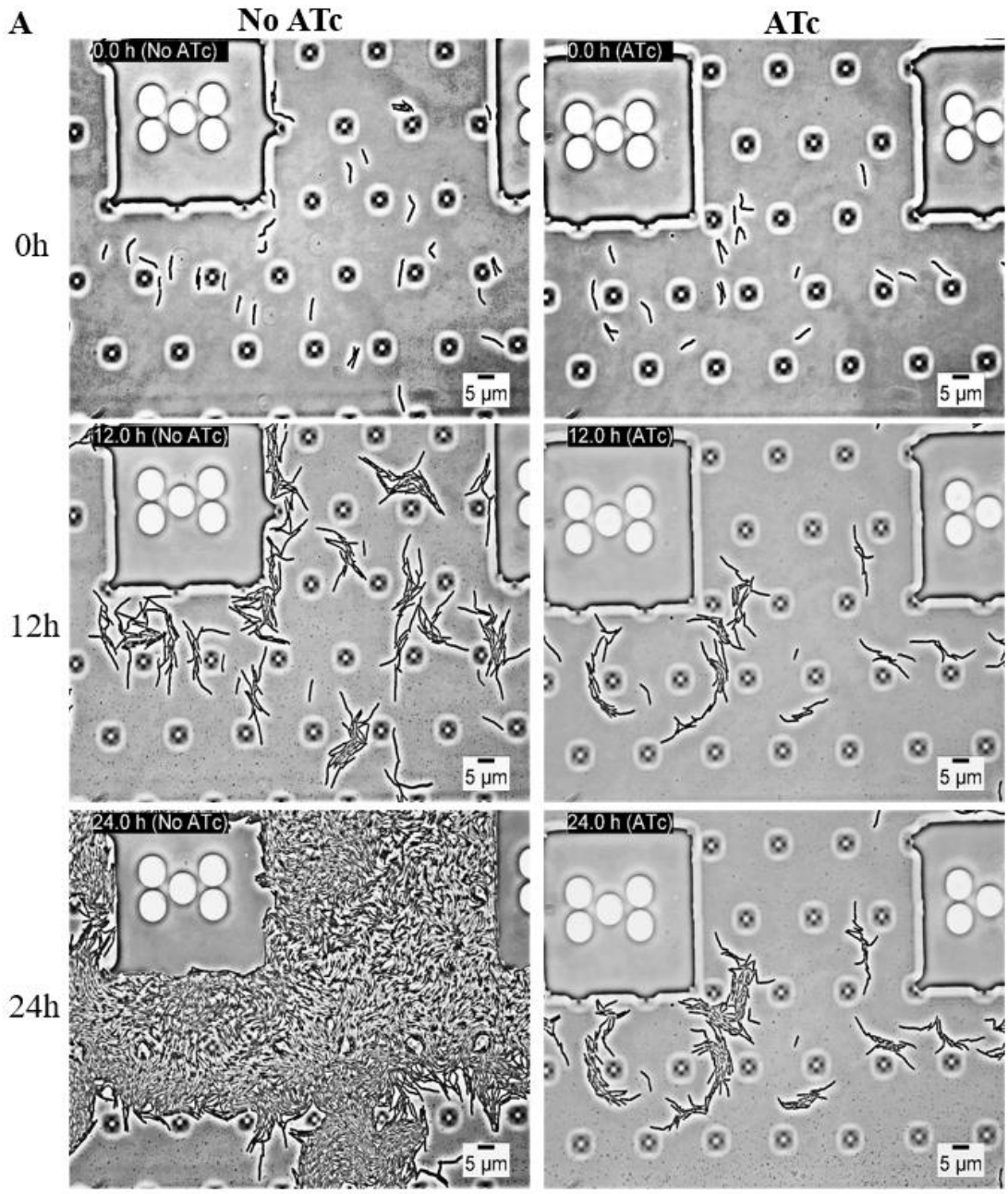


Figure 5.15: **Live-cell imaging of the *metH* knockdown using time-lapse phase-contrast microscopy**. Real-time visualization of methionine starvation in the *metH* conditional knockdowns using phase-contrast time-lapse microscopy. Compared to the uninduced control (No ATc) the *metH* conditional knockdown (ATc) had severe growth stunting. The images shown are still-frames captured at 0, 12, and 24h time points. Images were taken from Supplementary Movie 2. Scale bars represent 5µm.

5.2.9 Whole-genome sequencing analysis of *M. tuberculosis* clinical isolates reveals naturally occurring *metH* deletions

The *metH* deletion construct described in this chapter was made to mimic a natural large sequence deletion in the *ppe37-metH* genomic locus which eliminates the cobalamin binding domain and partially disrupts the SAM binding domain in *M. tuberculosis* CDC1551 (Warner *et al*., 2007). To determine if this deletion was unique to CDC1551, which is a lineage 4 (L4) strain, an analysis of WGS data for 169 *M. tuberculosis* clinical isolates from a high burden disease setting in Cape Town, South Africa (Koch *et al*., 2017), was conducted. This analysis identified 12 clinical isolates belonging to a subclade of L4 strains with the same large sequence deletion in MetH as the CDC1551 strain (**Figure 5.16**). Genomic DNA samples isolated from these 12 *M. tuberculosis* clinical strains was screened by PCR to validate the predicted deletions. In 10 of the 12 *M. tuberculosis* clinical strains, PCR confirmed the *metH* deletions (**Figure 5.16**). The deletions in the two other strains could not be ascertained by PCR because their genomic DNA had degraded while in storage.

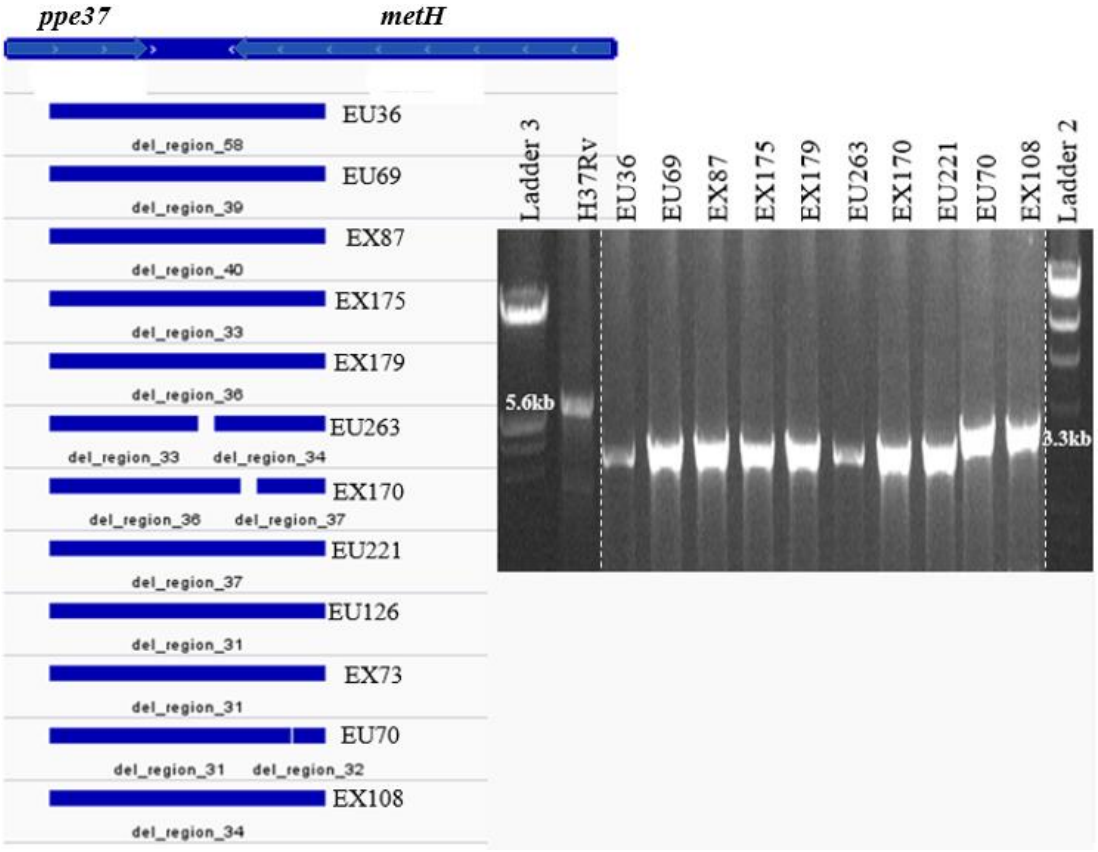


Figure 5.16: **Whole genome analysis and PCR verification of naturally-occurring *metH* deletions in *M. tuberculosis* clinical isolates**. Whole-genome sequence data were analysed for large deletions in genomes of 169 circulating clinical strains from Cape Town, South Africa. A large deletion in the *ppe37-metH* genomic locus was identified in 12 isolates (indicated by blue bar). The deletions were verified in 10 of the 12 isolates by PCR using the high fidelity Phusion polymerase and primers TB_*metH*_F/R, which target flanking the *PPE37-metH* region of *M. tuberculosis H37Rv*. In H37Rv, this region was intact, resulting in a 5.6kbp PCR product, whereas in the clinical isolates, the deletion produced a 3.3kbp PCR product. The dashed lines show break between images merged from distinct gels. WGS analysis courtesy of Dr. Jon Ambler.

These findings suggested that potentially inactivating *metH* mutations were widespread globally. Other genomic screens of clinical isolates from around the world have identified 25 ns-SNMs spread throughout the MetH protein (Comas *et al*., 2013). In the analysis conducted here, 10 s-SNMs and 13 ns-SNMs in *metH* were also found distributed among 39 of the 169 *M. tuberculosis* clinical isolates. (**Table 5.3**).

Table 5.3: **SNMs identified at *metH* locus in *M. tuberculosis* clinical strains**

Rv number	SNM	Amino Acid Change	Type of SNM	Number of strains with SNM at site
Rv2124c	G1A	Val1Met	ns-SNM	3
Rv2124c	C40T	Leu14Phe	ns-SNM	2
Rv2124c	C69T	Val23Val	s-SNM	2
Rv2124c	A96G	Leu32Leu	s-SNM	1
Rv2124c	T275G	Iso92Ser	ns-SNM	1
Rv2124c	T294C	Asp98Asp	s-SNM	1
Rv2124c	G373A	Gly125Arg	ns-SNM	1
Rv2124c	T405C	Thr125Thr	s-SNM	6
Rv2124c	C416T	Thr139Iso	ns-SNM	3
Rv2124c	G660T	Glu220Asp	ns-SNM	2
Rv2124c	C723T	His241His	s-SNM	2
Rv2124c	G887A	Gly296Asp	ns-SNM	2
Rv2124c	A1002G	Ala334Ala	s-SNM	1
Rv2124c	A1447C	Lys483Gln	ns-SNM	1
Rv2124c	G1616A	Arg539His	ns-SNM	1
Rv2124c	A2254G	Iso752Val	ns-SNM	1
Rv2124c	C2952T	Asp984Asp	s-SNM	1
Rv2124c	T3292G	Tyr1098Asp	ns-SNM	1
Rv2124c	G3397A	Gly1133Ser	ns-SNM	2
Rv2124c	C3423T	Gly1141Gly	s-SNM	2
Rv2124c	C3445G	Arg1149Gly	s-SNM	1
Rv2124c	G3537C	Ala1179Ala	s-SNM	1
Rv2124c	C3540T	Phe1180Phe	ns-SNM	1

5.2.10 The role of MetH in anti-folate resistance in *M. smegmatis*

The methionine and folate biosynthesis pathways are tightly coupled. 5-MTHF, the donor of the methyl group used in the methylation of homocysteine by MetH or MetE is a product of the folate biosynthesis pathway. Blanco *et al*. showed that the disruption of the 5,10-methylenetetrahydrofolate reductase, MetF, which catalyses the biosynthesis of 5-MTHF, led to methionine auxotrophy in *Streptomyces lividans* (Blanco, Coque & Martin, 1998). In humans, an imbalance in folate biosynthesis is associated with cobalamin deficiency (Palmer *et al*., 2017). The biochemical

anomalies involved in humans are well characterized, and they involve the inability of MetH – the sole methionine synthase in humans – to transfer methyl groups from 5-MTHF in the absence of cobalamin (Krebs, Hems & Tyler, 1976). This causes a phenomenon termed the "methyl folate trap," which is the harmful accumulation of 5-MTHF, depleting thymine. Because cobalamin deficiency in humans also impairs the biosynthesis of SAM, a repressor of MetF, the reduced SAM levels ultimately also increase the cellular pool of 5-MTHF (Krebs, Hems & Tyler, 1976).

Recently, methionine biosynthesis in *M. smegmatis* was implicated in intrinsic mycobacterial resistance to sulphur-containing antibiotics known as sulphonamides or SULFAs (Guzzo *et al*., 2016). SULFAs were the first drugs used for TB treatment, but their use gradually declined due to toxicity and were replaced with the safer INH-RIF combination therapy (Libecco & Powell, 2004; Forgacs *et al*., 2009). Currently, SULFAs are clinically indicated for urinary tract infections, bacterial meningitis, *Salmonella* infections, and chronic bronchitis (Libecco & Powell, 2004).

SULFAs target *de novo* folate biosynthesis by competing with para-aminobenzoic acid (PABA), a substrate of dihydropteroate synthase (DHPS) (**Figure 5.17**) (Libecco & Powell, 2004). The enzymatic reaction catalysed by DHPS generates THF, which is a precursor for the one-carbon pool by folate where it is terminally converted to 5-MTHF. The transfer of the methyl group from 5-MTHF to homocysteine by MetH recycles 5-MTHF back to THF (**Figure 5.17**). Interestingly, Guzzo *et al*. showed that either the inactivation of *metH*, but not of *metE*, or a deletion of the gene encoding the cobalamin biosynthesis fusion protein CobIJ, led to the accumulation of 5-MTHF in *M. smegmatis* (Guzzo *et al*., 2016). Furthermore, these mutants (with the exception of the Δ*metE* strain) showed increased susceptibilities to SULFAs relative to the wild-type strain (Guzzo *et al*., 2016). These tantalizing results prompted an examination of the role of MetH in anti-folate resistance.

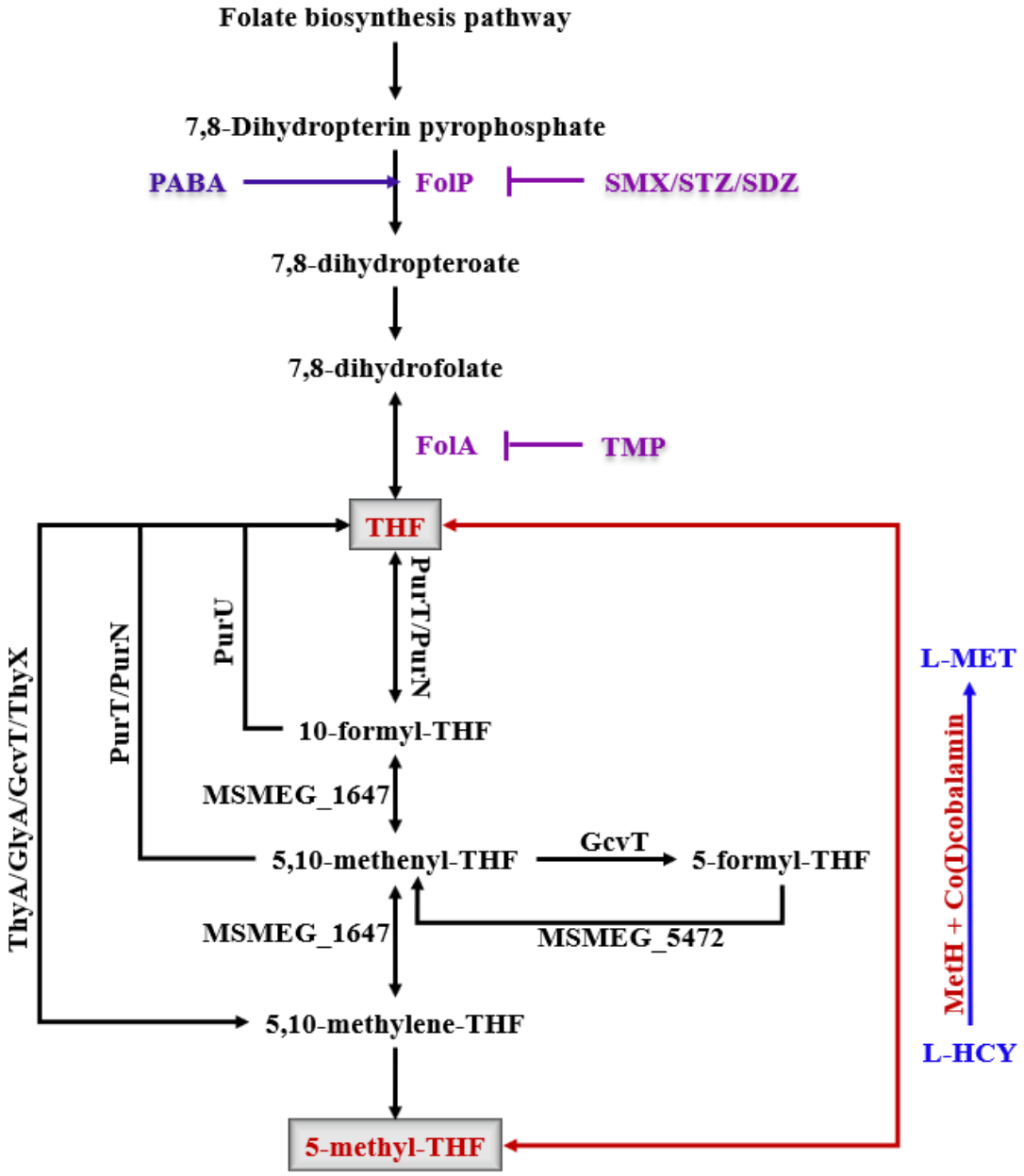


Figure 5.17: **The link between methionine biosynthesis and the one carbon pool by folate in *M. smegmatis***. Tetrahydrofolate (THF) is a derivative of the folate biosynthesis pathway and is a precursor of 5-methyl THF, the methyl donor in the conversion of L-homocysteine (HCY) to L-methionine (MET) by MetH using Co(I)cobalamin as a co-factor. SULFAs (SMZ – sulfamethoxazole; STZ – sulfathiazole; SDZ – sulfadiazine) target folate biosynthesis by competing with para-aminobenzoic acid (PABA) for binding to FolP (dihydropteroate synthase). Trimethoprim (TMP) targets FolA (dihydrofolate reductase).

5.2.11 Cobalamin deficiency in *M. smegmatis* enhances susceptibility to sulphonamides

As described in the previous sections, attempts to delete *metH* in *M. smegmatis* were unsuccessful due to the demonstrated essentiality of this gene in this organism. However, with the cobalamin deficient – and consequently MetH deficient – mutants (Δ*cobK* and Δ*cobK*Δ*metH*) and the MetE null mutant (Δ*metE*) in hand, the predicted role of cobalamin deficiency in boosting SULFA sensitivity in *M. smegmatis* could be examined. For this objective, sulfamethoxazole (SMZ), sulfathiazole (STZ) and sulfadiazine (SDZ) were used as representative SULFAs.

First, the minimum inhibitory concentration, defined as the concentration of each SULFA required to inhibit 90% (MIC_{90}) of the growth of wild-type *M. smegmatis* culture from a starting inoculum of 2.5–5.0 × 10^3 cfu/ml was determined using the microplate alamar blue (resazurin) assay (MABA). The observed MIC_{90} value for all three SULFAs was 1.56μg/ml (**Table 5.4**), which corresponded to reported values (Guzzo *et al*., 2016).

Table 5.4: **MIC_{90} of SULFAs used in this study**

SULFA Drug	WT	Δ*cobK*	Δ*metE*	Δ*cobK*Δ*metH*
SMZ	1.56μg/ml	0.2μg/ml	1.56μg/ml	0.2μg/ml
STZ	1.56μg/ml	0.2μg/ml	1.56μg/ml	0.2μg/ml
SDZ	1.56μg/ml	0.2μg/ml	1.56μg/ml	0.2μg/ml

The susceptibilities of strains lacking *de novo* synthesised cobalamin (Δ*cobK* and Δ*cobK*Δ*metH*) or strains capable of *de novo* cobalamin biosynthesis (wild-type and Δ*metE*) to SULFAs were then compared. For all three SULFAs, the MIC_{90} value in strains lacking *de novo* synthesised cobalamin was 8-fold lower than in those with functional *de novo* cobalamin biosynthesis (**Table 5.4**). There were no differences in MIC_{90} between the wild-type and Δ*metE* strains (**Table 5.4**). These results directly implicated cobalamin deficiency and, consequently, MetH inactivation, in enhancing sensitivity to SULFAs. Therefore, it was thought that exogenous CNCbl might reverse this sensitivity. To explore this possibility, the MABA was repeated, including

10µg/ml CNCbl in the wells. Again, MIC_{90} values of SMZ, STZ and SDZ were compared between strains lacking *de novo* synthesised cobalamin-deficient and those capable of *de novo* cobalamin biosynthesis.

Exogenous CNCbl did not change the MIC_{90} values of all three SULFAs in wild-type, Δ*cobK*, and Δ*metE* strains (**Table 5.5**), which was not surprising in the context of poor CNCbl uptake as already discussed. Unexpectedly, however, the Δ*cobK*Δ*metH* strain was hypersensitive to SULFAs in the presence of exogenous CNCbl. The SULFA MIC_{90} values of this mutant in the presence of CNCbl was 0.02µg/ml (**Table 5.5**)

Table 5.5: **MIC_{90} of SULFAs in the presence of exogenous CNCbl**

SULFA + CNCbl	WT	Δ*cobK*	Δ*metE*	Δ*cobK*Δ*metH*
SMZ + CNCbl	1.56µg/ml	0.2µg/ml	1.56µg/ml	0.02µg/ml
STZ + CNCbl	1.56µg/ml	0.2µg/ml	1.56µg/ml	0.02µg/ml
SDZ + CNCbl	1.56µg/ml	0.2µg/ml	1.56µg/ml	0.02µg/ml

Although the sensitivity of the Δ*cobK*Δ*metH* strain had been predicted to occur due to the cobalamin-mediated suppression of MetE, this phenotype had not been observed in shaking cultures and in agar as discussed in previous sections. Serendipitously, however, it appeared that this sensitivity phenotype was unmasked by the lower inoculum concentration in the MABA (2.5–5 $\times 10^3$ cfu/ml) compared to the higher concentration of the inoculum typically used to start batch cultures (1 $\times 10^6$ cfu/ml). Compounded with poor CNCbl uptake, this inoculum effect explained how an outgrowth of cells that inefficiently transported CNCbl would have eclipsed the CNCbl sensitivity phenotype in batch cultures of the the Δ*cobK*Δ*metH* strain. By contrast, the lower inoculum in the MABA effectively decreased the probability of an outgrowth. Therefore, when SULFAs were used in the assay in combination with CNCbl, the growth of Δ*cobK*Δ*metH* became severely inhibited (**Figure 5.18**; **Figure 5.19**; **Figure 5.20**). The severe retardation of the growth of Δ*cobK*Δ*metH* by CNCbl

in the MABA proved that *M. smegmatis* could take up small amounts of exogenous co-factor.

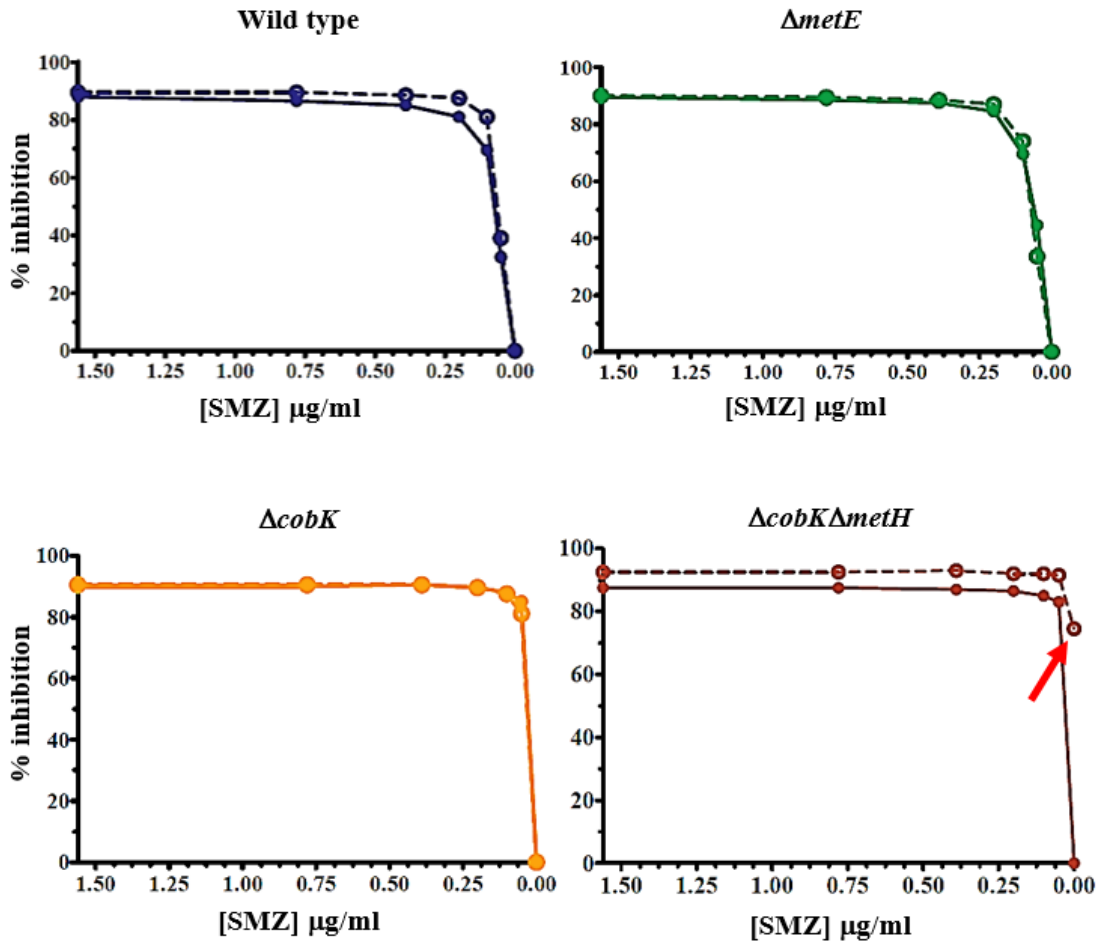


Figure 5.18: **Inhibition of *M. smegmatis* growth by sulfamethoxazole.** The percentage inhibition at different concentrations of SMZ in the presence (solid lines) or absence (dashed lines) of CNCbl was determined using the MABA. Even without SMZ, CNCbl inhibited 70% of cell growth of Δ*cobK*Δ*metH* (red arrowhead).

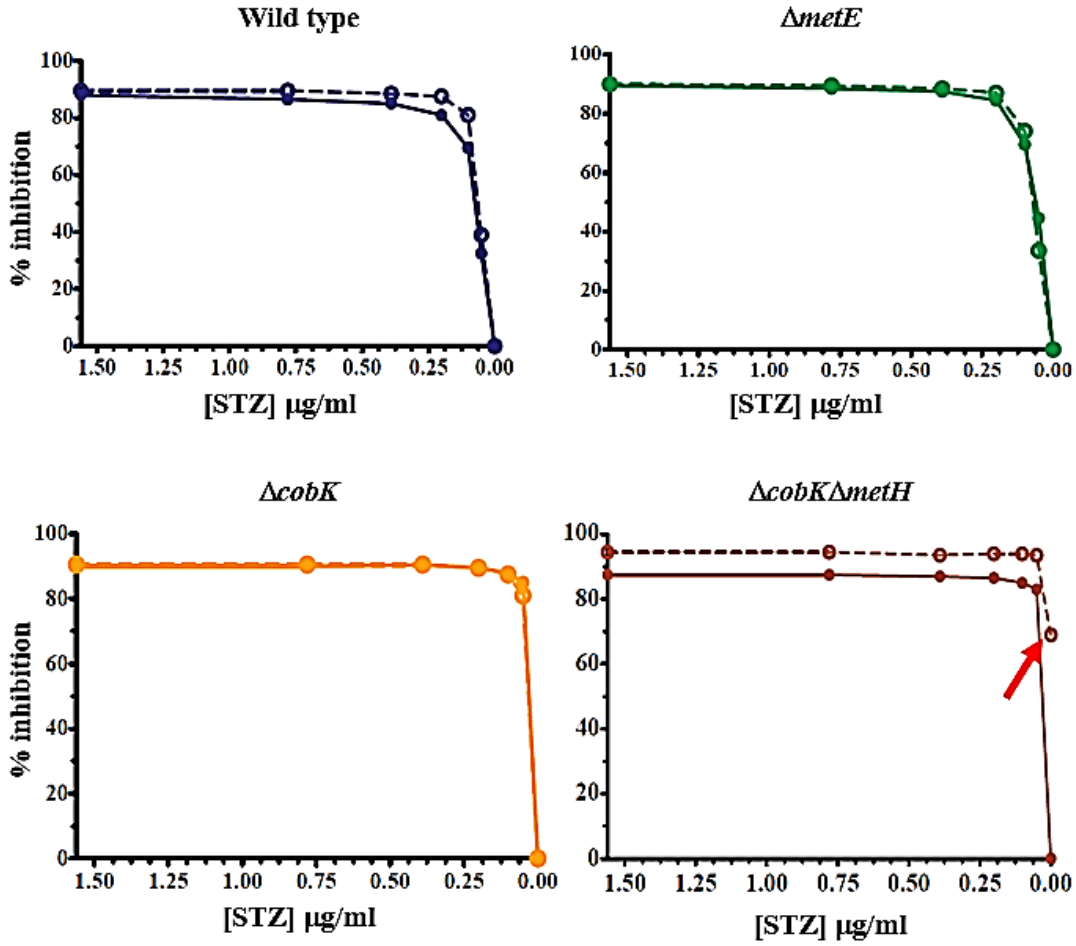


Figure 5.19: **Inhibition of *M. smegmatis* growth by sulfathiazole.** The percentage inhibition at different concentrations of STZ in the presence (solid lines) or absence (dashed lines) of CNCbl was determined using the MABA. Even without STZ, CNCbl inhibited 70% of cell growth of Δ*cobK*Δ*metH* (red arrowhead).

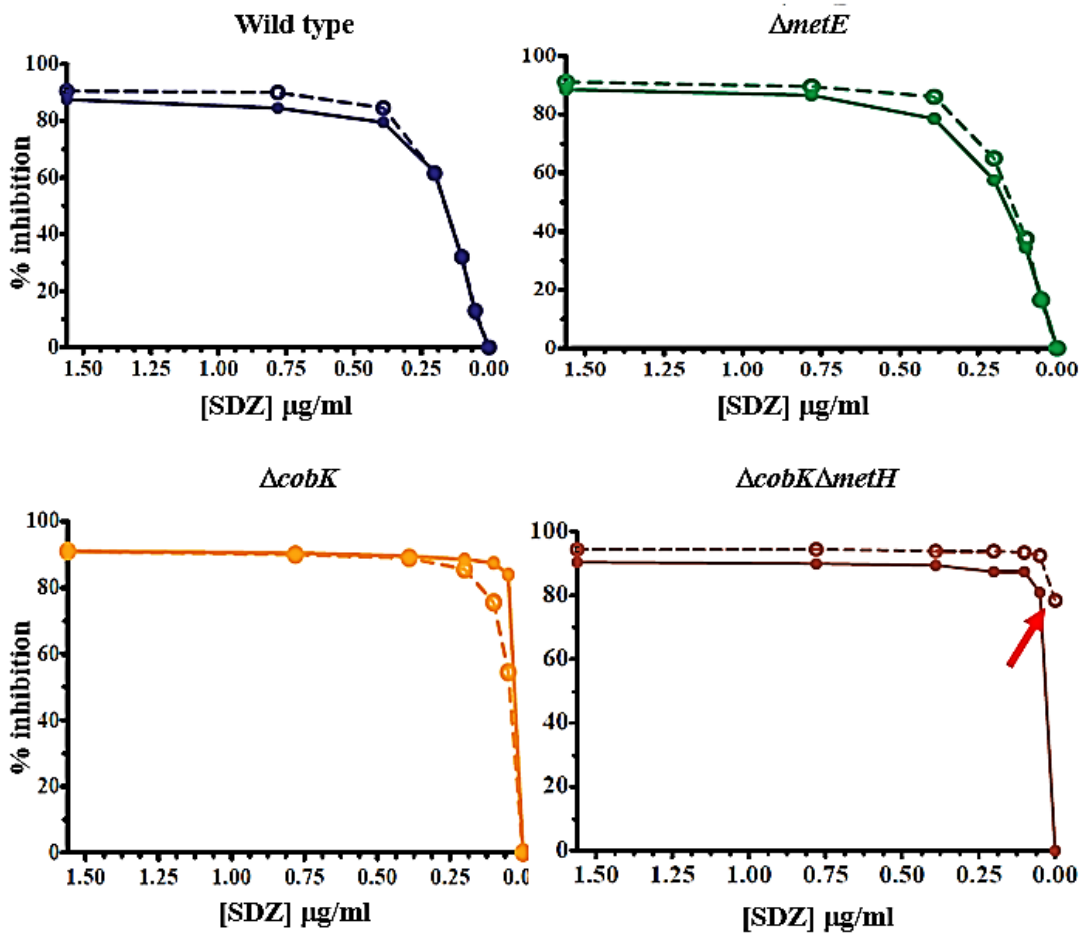


Figure 5.20: **Inhibition of *M. smegmatis* growth by sulfadiazine.** The percentage inhibition at different concentrations of SDZ in the presence (solid lines) or absence (dashed lines) of CNCbl was determined using the MABA. Even without SDZ, CNCbl inhibited 80% of cell growth of Δ*cobK*Δ*metH* (red arrowhead).

Due to the observed limited CNCbl uptake, it was thought that preincubation of *M. smegmatis* with the co-factor before setting up the MABA would allow cobalamin to accumulate inside cells, thereby enhancing the resistance of these cells to SULFAs. To test this idea, an experiment was set up whereby cells were seeded on the microplate and preincubated with CNCbl for one (3h) and two doubling times (6h) prior to introducing SULFAs. Alternatively, the assay was set up with cells that had not been preincubated with CNCbl (0h); *i.e.*, the SULFAs and CNCbl were added simultaneously at the beginning of the assay. The proportion of viable cells at the endpoint of each assay (0h, 3h, and 6h) – as indicated by fluorescence intensity following resazurin reduction – was plotted against the concentration of SULFA drug. Overall, longer preincubation times with CNCbl were associated with improved ability

to tolerate SULFAs. However, the Δ*cobK*Δ*metH* strain was severely inhibited by SULFAs (**Figure 5.21**; **Figure 5.22**; **Figure 5.23**).

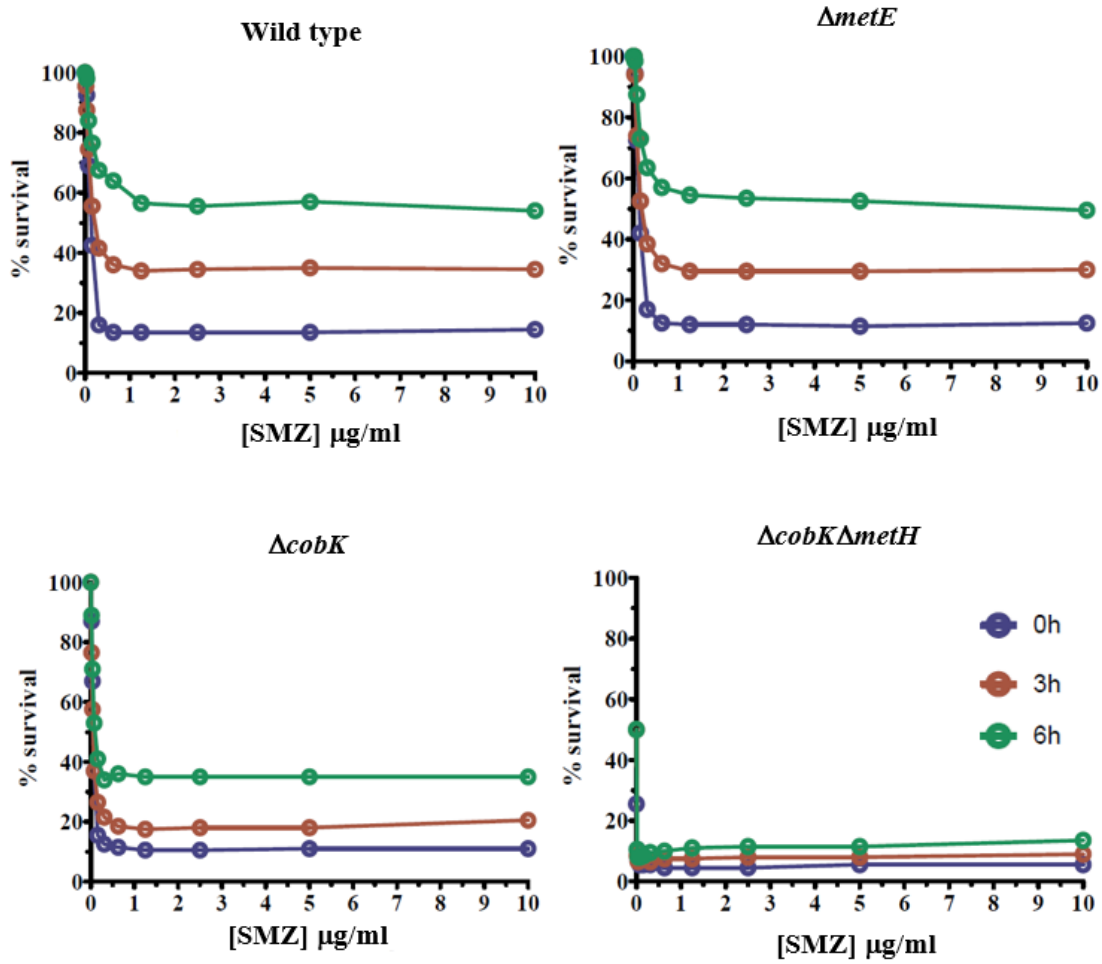


Figure 5.21: **Extended preincubation with CNCbl enhances resistance to sulfamethoxazole**. Cells were preincubated with CNCbl for 0h (blue), 3h (orange) and 6h (green) prior to SMZ exposure. The percentage of cell survival was calculated using the MABA. Longer duration of pre-exposure to CNCbl was associated with resistance to SMZ.

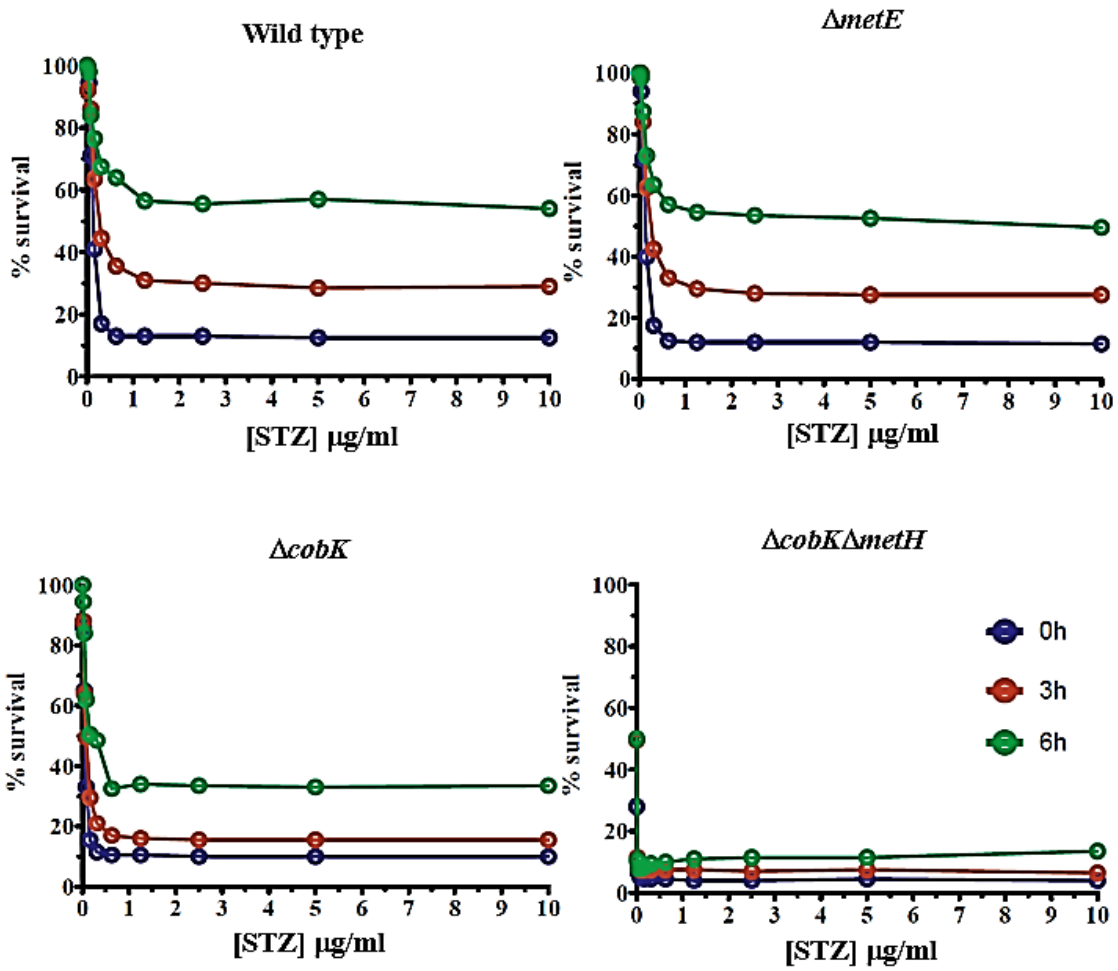


Figure 5.22: **Extended preincubation with CNCbl enhances resistance to sulfathiazole**. Cells were preincubated with CNCbl for 0h (blue), 3h (orange) and 6h (green) prior to STZ exposure. The percentage of cell survival was calculated using the MABA. Longer duration of pre-exposure to CNCbl was associated with resistance to STZ

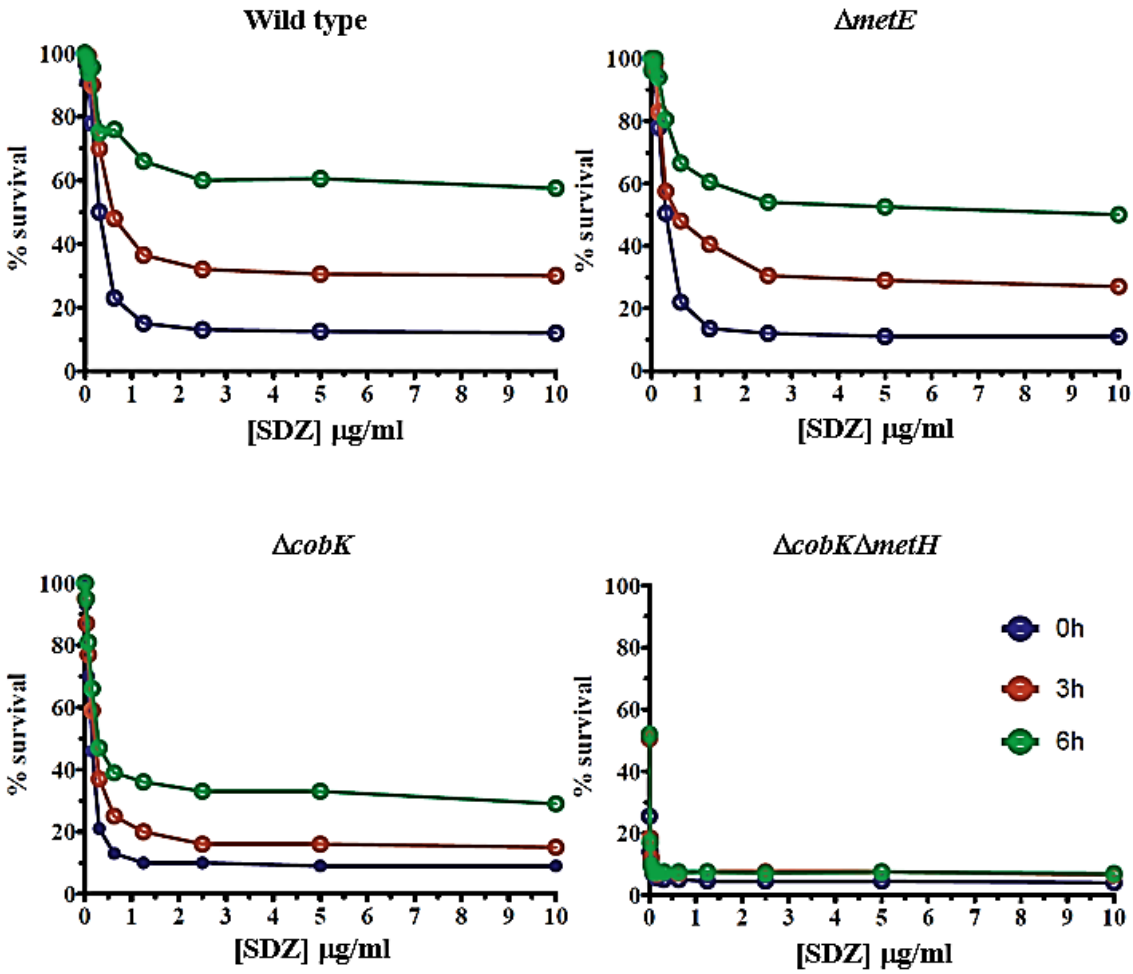


Figure 5.23: **Extended preincubation with CNCbl enhances resistance to sulfadiazine.** Cells were preincubated with CNCbl for 0h (blue), 3h (orange) and 6h (green) prior to SDZ exposure. The percentage of cell survival was calculated using the MABA. Longer duration of pre-exposure to CNCbl was associated with resistance to SDZ.

At 0h preincubation, the concentration of STZ and SMZ that was needed to inhibit 50% of growth (MIC_{50}) of wild-type and Δ*metE* in the presence of CNCbl was roughly 0.15µg/ml while that of SDZ was 0.3µg/ml. By contrast, 0.05µg/ml of STZ and SMZ and 0.15µg/ml of SDZ were required to inhibit 50% growth of CNCbl-supplemented Δ*cobK* (**Table 5.6**). Pre-exposing the cultures to CNCbl for 3h doubled the MIC_{50} of all the three SULFAs for wild-type and Δ*metE* only, but the MIC_{50} for Δ*cobK* remained constant. At 6h preincubation, the proportion of wild-type and Δ*metE* cells surviving was above 50% across the range of concentrations tested (0–10µg/ml) (**Table 5.6**; **Figure 5.21**; **Figure 5.22**; **Figure 5.23**). However, at 6h pre-exposure, the survival rate of Δ*cobK* strain was 30–40%, and in this case, the MIC_{50} values for STZ, SMZ

and SDZ in this mutant were 0.1µg/ml, 0.19µg/ml, and 0.29µg/ml respectively (**Table 5.6**). The MIC_{50} of Δ*cobK*Δ*metH* could not be determined, as the growth of Δ*cobK*Δ*metH* was inhibited by 50% by CNCbl alone (**Figure 5.21**; **Figure 5.22**; **Figure 5.23**).

Table 5.6: **MIC_{50} of SULFAs used in the preincubation experiment**

0h	**WT**	**Δ*cobK***	**Δ*metE***	**Δ*cobK*Δ*metH***
SMZ + CNCbl	0.15µg/ml	0.05µg/ml	0.14µg/ml	undetermined
STZ + CNCbl	0.13µg/ml	0.06µg/ml	0.13µg/ml	undetermined
SDZ + CNCbl	0.3µg/ml	0.15µg/ml	0.31µg/ml	undetermined
3h	**WT**	**Δ*cobK***	**Δ*metE***	**Δ*cobK*Δ*metH***
SMZ + CNCbl	0.22µg/ml	0.05µg/ml	0.19µg/ml	undetermined
STZ + CNCbl	0.27µg/ml	0.08µg/ml	0.26µg/ml	undetermined
SDZ + CNCbl	0.63µg/ml	0.22µg/ml	0.56µg/ml	undetermined
6h	**WT**	**Δ*cobK***	**Δ*metE***	**Δ*cobK*Δ*metH***
SMZ + CNCbl	>10µg/ml	0.10µg/ml	>10µg/ml	undetermined
STZ + CNCbl	>10µg/ml	0.19µg/ml	>10µg/ml	undetermined
SDZ + CNCbl	>10µg/ml	0.29µg/ml	>10µg/ml	undetermined

5.2.12 MetH activity is a primary determinant of resistance to SULFA in *M. smegmatis*

The observation that the presence of CNCbl correlated with increased ability to survive SULFAs supported the notion that SULFA resistance was mediated primarily by MetH (Guzzo *et al*., 2016). In the wild-type and Δ*metE* strains, which had better survival profiles than Δ*cobK*, MetH was the primary methionine synthase. However, in Δ*cobK*, MetH was inactive due to the absence of *de novo* synthesised cobalamin. Also, because of poor CNCbl uptake, the exogenous co-factor marginally improved survivability of Δ*cobK* in the presence of SULFAs. Thus, these data also supported the notion that resistance to SULFAs required MetH activity.

To further explore this idea, a SULFA sensitivity assay was developed using the *metH* CRISPRi knockdown strain. To allow cells carrying the *metH* CRISPRi construct to

survive *metH* silencing, 1mM methionine was added to ATc-containing wells. Starting with an inoculum of 1 x 10^3 cfu/ml, the assay was set up in a 96-well plate, using 0.4µg/ml SMZ (equivalent to 0.25× MIC) and bacterial growth was monitored by taking periodic OD_{620nm} measurements over the course of 15 doubling times (45h). As expected, *metH* knockdown led to growth inhibition but was rescued by supplementing with methionine (**Figure 5.24A**). When SMZ was used at 0.25× MIC, there was reduced growth in uninduced cells (*i.e.*, wild-type cells) and a clear separation in growth curves between *metH* knockdowns and wild-type cells could be observed (**Figure 5.24B**). The survival of *metH* knockdowns in the presence of SMZ was improved when methionine was included, and the separation of growth curves became clearer when the experiment duration was by another 5 doubling times to 60h (**Figure 5.24B**). From these findings, it was concluded that MetH activity was needed to resist SMZ and that exogenous methionine somewhat antagonized SMZ in the absence of MetH.

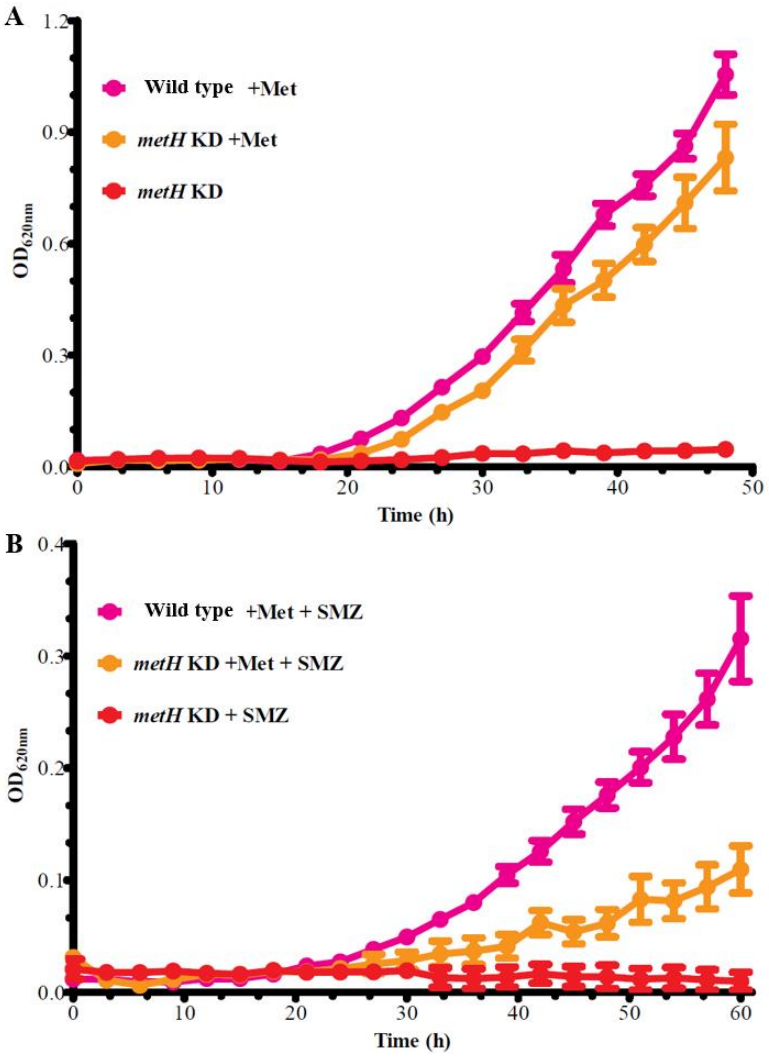


Figure 5.24: **MetH is a driver of sulfamethoxazole resistance in *M. smegmatis***. **A**. The growth of the *metH* knockdown or wild-type cells grown with or without methionine was monitored by taking periodic OD_{620nm} measurements. The *metH* knockdown failed to grow over the course of 15 doubling times but was rescued by exogenous methionine. **B**. Susceptibility of *metH* knockdowns to SMZ in the presence of exogenous methionine. Wild-type cells had better survival compared to the *metH* knockdown. Exogenous methionine slightly improved survivability of the knockdown.

5.3 Discussion

In *M. tuberculosis*, the gene encoding the cobalamin-independent methionine synthase, MetE, is annotated as essential but the gene encoding the cobalamin-dependent methionine synthase, MetH, is non-essential (Kapopoulou, Lew & Cole, 2011). The essentiality of *metE* assumes that *M. tuberculosis* is incapable of *de novo* cobalamin biosynthesis, a deficiency which has a two-fold impact on methionine

biosynthesis. First, the absence of cobalamin inactivates MetH by depriving it of its co-factor. Secondly, lack of cobalamin releases MetE from a riboswitch-mediated repression (Warner *et al*., 2007). Therefore, MetE is indispensable in *M. tuberculosis*. On the other hand, the cobalamin-producing *M. smegmatis* can use either methionine synthase. In theory, MetE is expected to take over from MetH if *de novo* cobalamin production were compromised and alternatively, MetH would compensate if MetE activity were eliminated. However, due to the presence of a cobalamin riboswitch controlling MetE in *M. smegmatis*, the loss of MetH in a background of *de novo* cobalamin production would be growth inhibitory. In the present work, the essentiality of this gene was confirmed using the mycobacterial CRISPRi system (Rock *et al*., 2017). The abrogation of *de novo* cobalamin biosynthesis sufficiently offset the growth inhibition phenotype that was observed following the loss of *metH*.

Consequently, a simple model for the essentiality of *metH* in *M. smegmatis* is proposed. In this mechanism, a functional *metE* riboswitch and enough concentrations of internal cobalamin are required (**Figure 5. 25A**). Because endogenous cobalamin is available in *M. smegmatis*, the co-factor facilitates methionine biosynthesis via MetH, while simultaneously repressing *metE* via the riboswitch. However, when *de novo* cobalamin biosynthesis is disrupted, MetE is available whereas MetH becomes inactive (**Figure 5.25B**). In this manner, *M. smegmatis* is protected from methionine starvation. However, when MetH is deleted while *de novo* cobalamin biosynthesis is still intact, cells effectively lose both methionine synthases, resulting in methionine depletion and eventual growth inhibition (**Figure 5.25C**). This also creates the methyl folate trap, which has been reported to underlie increased susceptibility of. *smegmatis* and *M. tuberculosis* to SULFAs (Guzzo *et al*., 2016).

A formal validation of the functionality of the *M. smegmatis metE* riboswitch functionality has not been presented because attempts to introduce mutations in the aptamer region of this riboswitch were not successful despite repeated efforts to do so. However, in spite of this missing piece of data, the evidence presented in this thesis is compelling to support the assertion that the availability of cobalamin and the presence of a functional *metE* riboswitch underlies the essentiality of *metH*.

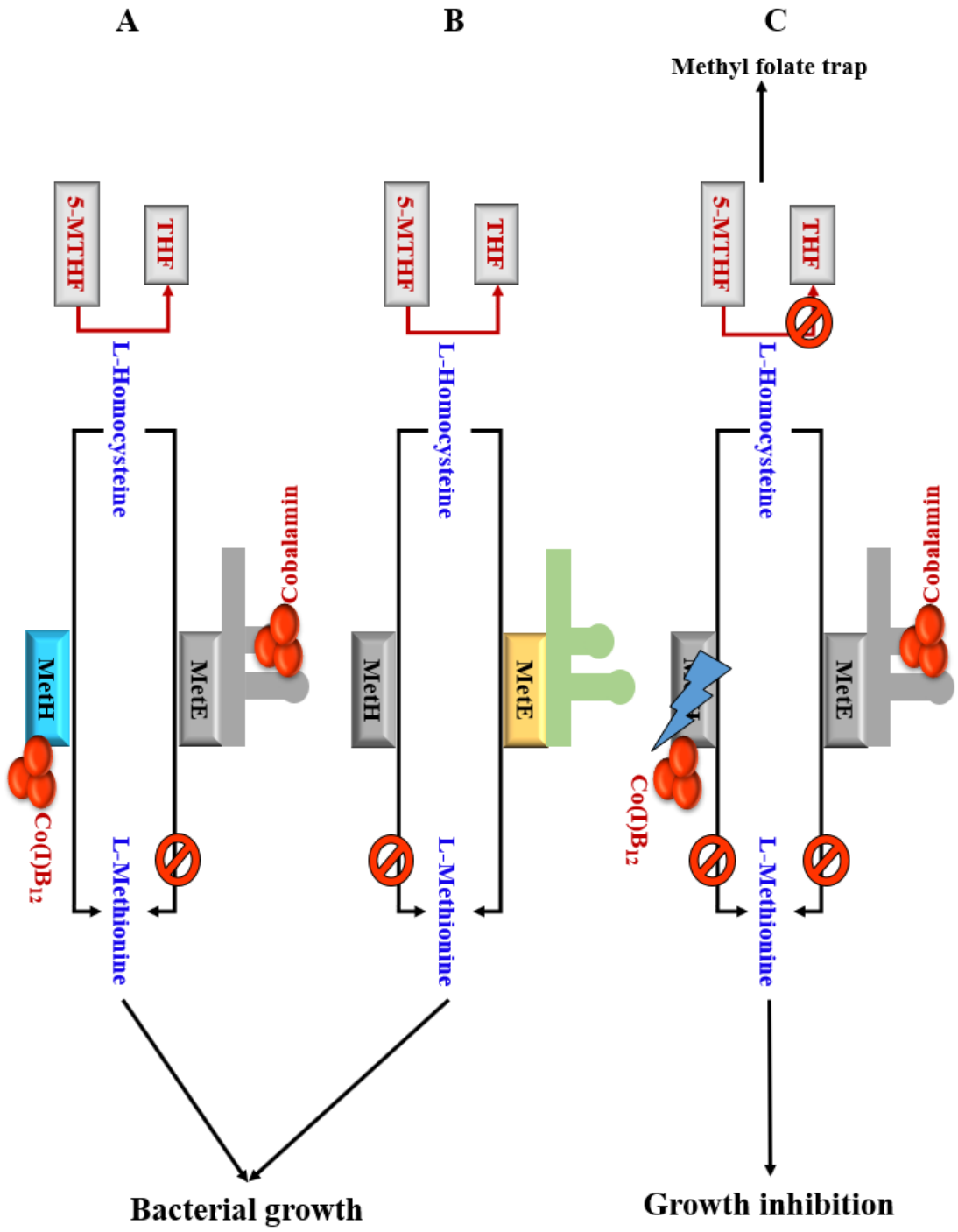


Figure 5.25: **A proposed mechanism for the essentiality of *metH* in *M. smegmatis*. A-B.** *De novo* methionine biosynthesis in *M. smegmatis* can occur via MetH using Co(I)cobalamin ($Co(I)B_{12}$) as a co-factor, or via MetE. Cobalamin represses *metE* via a riboswitch whereas in the absence of cobalamin, only MetE is viable. MetH and MetE transfer a methyl group from 5-methyl tetrahydrofolate (5-MTHF) to homocysteine, producing tetrahydrofolate (THF) as an intermediate. **C.** If MetH were deleted in the presence of cobalamin, the cell would lose both methionine synthases, succumbing to methionine starvation. Loss of MetH may also lead to the accumulation of 5-MTHF, creating the methyl folate trap.

Surprisingly, Guzzo *et al*. (2016) reported the use of a standard recombineering method to successfully delete the *M. smegmatis metH*. The evidence presented herein, which has experimentally validated *metH* as an essential gene in *M. smegmatis* and is supported by independent Tn-seq and CRISPRi-seq analysis (de Wet *et al.*, 2018), therefore contradicts the findings of these authors. In the present study, *metH* could not be deleted by allelic exchange mutagenesis which takes advantage of endogenous recombination pathways in mycobacteria to replace a wild-type with a mutant allele. Since MetE is under cobalamin-mediated repression in *M. smegmatis*, the deletion of *metH* would render the organism unable to produce methionine, making the *metH* knockout mutants non-viable. One possible explanation for the discrepancy between the results presented here and those in Guzzo *et al*. (2016) is the likely presence of spontaneous suppressor mutations affecting the function of the *metE* riboswitch in the *M. smegmatis* laboratory strain used by those authors. It is well known that random mutations accumulate due to serial passaging of bacterial culture (Kucukyildirim *et al*., 2016; Schroeder *et al*., 2018), a caveat that the paper did not address. To exclude the possibility of inadvertent second-site mutations as potential confounding factors when drawing conclusions from genetic studies, all the mutant strains generated in this study were subjected to WGS analysis.

There are efforts to engineer cobalamin derivatives as delivery vehicles for drugs and biomarkers (Romine *et al*., 2017; Gryko *et al*., 2018; Lawrence *et al*., 2018). Studies have presented clear evidence of cobamide uptake in *M. tuberculosis* via a non-canonical transporter (Warner *et al*., 2007; Gopinath, Moosa, *et al*., 2013; Lawrence *et al*., 2018). However, the findings presented here suggest that the transport of cobamides might be tightly controlled in *M. smegmatis*. Possibly due to this regulation, the quantities of the *metE* transcript were only slightly decreased the presence of CNCbl, even though the exogenous co-factor was provided at high concentration (**Figure 5.5**). Similarly, the suppression of the growth of a double mutant deficient in cobalamin and lacking MetH (Δ*cobK*Δ*metH*) by exogenous CNCbl, was attained only when a very dilute inoculum of this strain was used in culture (**Figures 5.16-21**). Technical limitations undetermined the ability to measure the minimum internal concentration of cobalamin that would be required to completely shut down MetE activity. Furthermore, attempts to do titrate *metE* repression using exogenous CNCbl were hindered by poor CNCbl uptake in *M. smegmatis*.

The combination of SULFAs and loss of MetH activity was lethal in *M. smegmatis*. Both *in vitro* and clinical evidence indicate the susceptibility of *M. tuberculosis*, including drug resistant strains, to cotrimoxazole, a combination of sulfamethoxazole and TMP (Forgacs *et al*., 2009; Huang *et al*., 2012; Ameen & Drancourt, 2013). Given that there are circulating clinical strains that lack *metH*, it will be worthwhile to investigate whether this susceptibility is related to the presence of an intact *metH* allele or if wild-type (*e.g.*, *H37Rv*) and derivative *metH*-null strains have differential susceptibilities to SULFA drugs. In addition, if *M. tuberculosis* can access cobalamin *in vivo*, SULFA drug testing in animal models might provide critical insight that will inform the feasibility of repurposing and reintroducing these drugs into current antimycobacterial regimens.

Naturally occurring mutations affecting MetH activity have been reported in clinical *M. tuberculosis* isolates, but the implications of these alleles remain unclear. The ns-SNMs in the *M. tuberculosis metH* identified in this analysis and in other genomic screens (Comas *et al*., 2013) appear to occur in the tips of a phylogenetic tree, which implies their origins are recent. The fact that *M. tuberculosis* strains lacking MetH are circulating in populations presents an exciting opportunity for further research on the global prevalence of these strains and their relevance for *M. tuberculosis* physiology, TB epidemiology and treatment. Further WGS analyses are planned to confirm whether these mutations confer any selective advantages.

6 Concluding remarks

Cobamide acquisition and related metabolism in mycobacteria has been a subject of study for many years, especially in light of the obligate lifestyle adopted by *M. tuberculosis*, which is incapable of *de novo* cobmide biosynthesis despite requiring the co-factor for three known cobamide-dependent metabolic pathways (Gopinath, Moosa, *et al*., 2013; Young, Comas & de Carvalho, 2015). Like other mycobacteria, *M. tuberculosis* utilizes cobamide in several unrelated metabolic pathways including the biosynthesis of methionine and deoxyribonucleotides and the metabolism of propionate, a by-product of cholesterol and fatty acid metabolism (Young, Comas & de Carvalho, 2015). This thesis focused only on the methionine biosynthesis pathway. The overall goal of the thesis was to describe cobamide biosynthesis and cobamide-dependent regulation of methionine biosynthesis in mycobacteria. As the best characterised saprophytic mycobacterium, *M. smegmatis* was chosen as the model for cobamide biosynthesis, transport, assimilation and function in this study. The data described in this thesis show that the cobamide form produced by this organism is cobalamin, having DMB as the lower base. The data described in this thesis also show that methionine biosynthesis in *M. smegmatis* is under the control of a cobalamin riboswitch upstream of the *metE* gene. Since this riboswitch contains the "B_{12} element," an evolutionarily conserved RNA structure typical of AdoCbl-sensing riboswitches (Vitreschak *et al.*, 2003), the functionality of this riboswitch suggests that *M. smegmatis* produces AdoCbl. More work is needed to verify whether this organism produces other cobamide forms. The data further demonstrate that both the *de novo* and salvage pathways to cobalamin biosynthesis are functional in *M. smegmatis*.

The phylogenomic approach used to evaluate the genetic capacity for *de novo* cobamide production in mycobacteria indicated that environmental mycobacteria mostly have the potential for cobamide biosynthesis via the *de novo* aerobic pathway or via corrinoid salvage. However, host-associated mycobacterial species appeared to have partial genetic capacity for *de novo* cobamide biosynthesis but encoded intact pathways for corrinoid salvage. The differences in cobamide biosynthesis between pathogenic and non-pathogenic mycobacteria correspond to a major point of departure in their evolution.

The genetic capacity of mycobacteria to acquire exogenous cobamides via the ABC-type transporter, BacA, was reinforced in the work conducted here. This protein, which is also involved in antibiotic resistance (Domenech *et al*., 2009), was found in all 1069 mycobacterial genomes that were analysed in this thesis (chapter 3). In addition to encoding two paralogous genes encoding a putative BacA, *M. smegmatis* also contains two operons containing paralogous genes encoding putative homologues of BtuFCD, which are involved in the canonical cobamide transporter complex with BtuB and TonB-ExBD (Noinaj *et al.*, 2010). It was found that BtuF and BtuF2 (MSMEG_4561) proteins were sensitive to exogenous CNCbl, presumably due to repression via putative riboswitches in the 5' UTRs of their genes. As these genes are the first in the *btuFCD* and *btuF2C2D2* operons, respectively, their repression likely have knock-on effects on the expression of their downstream operonic partners. Since all the work conducted in this thesis was performed in *M. smegmatis* and *in vitro*, the question of whether pathogenic mycobacteria can access cobamides *in vivo* remains unresolved. The organisation of cobamide transport in mycobacteria is the subject of an ongoing study in the MMRU. Toward this end, cobalamin-based chemical probes are being utilized to track cobamide uptake in both *M. smegmatis* and *M. tuberculosis* and identify cobamide binding proteins in these two mycobacteria. A surprising finding in this thesis was that, despite the presence of these multiple putative transporters, the uptake exogenous CNCbl was poor in *M. smegmatis*. However, corrinoid salvage was demonstrably more efficient in a cobalamin/methionine auxotroph (**Figure 4.9**; **Figure 4.10**), which was dependent on exogenous CNCbl for growth. It is hereby proposed that cobamide transport and assimilation in *M. smegmatis* are modulated by a fine-tuning process whereby enhanced corrinoid salvage is triggered when endogenous cobalamin is missing; *i.e.*, when cells have an absolute requirement for the co-factor.

The control of mycobacterial methionine biosynthesis via a cobalamin riboswitch was first demonstrated in *M. tuberculosis* more than a decade ago (Warner *et al*., 2007). This early work demonstrated that in *M. tuberculosis*, the gene encoding MetE was transcriptionally repressed by cobalamin. In the present work, *metE* was found to be under constitutive transcriptional repression in *M. smegmatis* via the cobalamin riboswitch. The riboswitch-mediated regulation of *metE* was also identified as being responsible for the essentiality of the *M. smegmatis metH*, a finding that mirrored Tn-

seq and CRISPRi-seq predictions of the essentiality of this gene (de Wet *et al.*, 2018). Given the lack of cobamide biosynthesis in *M. tuberculosis*, the demonstration of constitutive riboswitch regulation of *metE* in *M. smegmatis* lays the foundation for planned future work to determine the structure and riboswitch mechanism of the ligand-on or ligand-off riboswitch states represented by *M. smegmatis* and *M. tuberculosis*, respectively.

The role of the mycobacterial MetH in mediating intrinsic resistance to SULFAs as recently proposed by Guzzo *et al*. (Guzzo *et al*., 2016) was also confirmed in this work. Specifically, the activity of SULFAs was potentiated when MetH activity, but not MetE activity, was missing (**Table 5.4**; **Table 5.5**). Considering potentially inactivating *metH* mutations discovered in several *M. tuberculosis* clinical isolates have uncovered, the potential for SULFAs to effectively target *metH*-null bacilli should be investigated. Advancements in sequencing technologies and increased application in routine TB diagnosis (Quan *et al*., 2018; Zignol *et al*., 2018) should make the rapid detection of *metH* mutations in the clinical context feasible, enabling targeted therapies.

In summary, this thesis provided direct evidence that mycobacteria can synthesise cobalamin either via the *de novo* pathway or via the salvage pathway, if corrinoid precursors can be accessed. Moreover, it appears that environmental and pathogenic mycobacteria differentially regulate cobamide uptake. The precise factors underlying these differences remain uncharacterized, as are many other aspects of *M. tuberculosis* physiology and metabolism, which poses significant barriers to the development of improved drugs and effective vaccines. New insights into the metabolic mechanisms that underlie pathogenesis and long-term survival of *M. tuberculosis* in infected individuals are required. Mycobacterial RNA-regulators (*e.g.*, cobalamin riboswitches), represent an emerging research focus with the potential to uncover novel drug targets. This thesis presents the idea that the outcomes of cobamide-dependent regulation might have implications for mycobacterial evolution, physiology, pathogenesis and drug discovery. Drawing from this idea, three key focus areas for future studies are proposed: (1) the complete characterization of mycobacterial cobamide biosynthesis and transport; (2) the determination of the structure and mechanism of action of the mycobacterial cobalamin riboswitch; and (3)

the evaluation of the prevalence of mutations affecting genes involved in cobamide-related metabolism in *M. tuberculosis* clinical isolates.

7 Supplementary information online

Table S1: https://uct.figshare.com/s/3881705f0809898c96e0

Table S2: https://uct.figshare.com/s/76e15301fe6dec848c00

Supplementary Movies 1–5: https://uct.figshare.com/s/65105b9914196c4b4654

8 References

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D.J. 1997. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. Nucleic Acids Research. 25(17):3389–3402. DOI: 10.1093/nar/25.17.3389.

Ameen, S.M. & Drancourt, M. 2013. *In Vitro* Susceptibility of *Mycobacterium tuberculosis* to Trimethoprim and Sulfonamides in France. *Antimicrobial Agents and Chemotherapy*. 57(12):6370–6371. DOI: 10.1128/AAC.01683-13.

Asmar, S., Rascovan, N., Robert, C. & Drancourt, M. 2015. Draft Genome Sequence of *Mycobacterium mucogenicum* Strain CSUR P2099. *Genome Announcement*. 3(6):10–11. DOI: 10.1128/genomeA.01369-15.Copyright.

Baker, J.L., Sudarsan, N., Weinberg, Z., Roth, A., Stockbridge, R.B. & Breaker, R.R. 2012. Widespread Genetic Switches and Toxicity Resistance Proteins for Fluoride. *Science*. 335(6065):233–235. DOI: 10.1126/science.1215063.

Banerjee, R. 2006. B_{12} trafficking in mammals: A case for coenzyme escort service. *ACS Chemical Biology*. 1(3):149–159. doi:10.1021/cb6001174

Banerjee, R., Gherasim, C., Padovani, D. 2009. The tinker, tailor, soldier in intracellular B_{12} trafficking. *Current Opinion in Chemical Biology*. 13:484–491

Barberis, I., Bragazzi, N.L., Galluzzo, L. & Martini, M. 2017. The History of Tuberculosis: from the First Historical Records to the Isolation of Koch's bacillus. *Journal of Preventive Medicine and Hygiene*. 58(1): E9–E12. DOI: 10.15167/2421-4248/jpmh2017.58.1.728.

Barry, C.E. 2001. *Mycobacterium smegmatis*: An Absurd Model for Tuberculosis? Response from Barry, III. *Trends in Microbiology*. 9(10):473–474. DOI: 10.1016/S0966-842X(01)02169-2.

Berney, M., Berney-Meyer, L., Wong, K.-W., Chen, B., Chen, M., Kim, J., Wang, J., Harris, D., *et al*. 2015. Essential Roles of Methionine and S-Adenosylmethionine in the Autarkic Lifestyle of *Mycobacterium tuberculosis*. *Proceedings of the National Academy of Sciences of the United States of America*. 112(32):10008–13. DOI: 10.1073/pnas.1513033112.

Beste, D.J. V, Nöh, K., Niedenführ, S., Mendum, T.A., Hawkins, N.D., Ward, J.L., Beale, M.H., Wiechert, W., *et al*. 2013. 13C-Flux Spectral Analysis of Host-Pathogen Metabolism Reveals A Mixed Diet for Intracellular *Mycobacterium tuberculosis*. *Chemistry and Biology*. 20(8):1012–1021. DOI: 10.1016/j.chembiol.2013.06.012.

Blanche, F., Cameron, B., Crouzet, J., Debussche, L., Thibaut, D., Vuilhorgne, M., Leeper, F. J., Battersby, A. R. 1995. Vitamin B_{12}: How the Problem of Its Biosynthesis Was Solved. *Angewandte Chemie International Edition in English*. 34(4): 383-411. DOI: 10.1002/anie.199503831

Blanco, J., Coque, J.J.R. & Martin, J.F. 1998. The Folate Branch of the Methionine Biosynthesis Pathway in *Streptomyces lividans*: Disruption of the 5,10-Methylenetetrahydrofolate Reductase Gene Leads to Methionine Auxotrophy. *Journal of Bacteriology*. 180(6):1586–1591.

Bloom, B.R., Atun, R., Cohen, T., Dye, C., Fraser, H., Gomez, G.B., Knight, G., Murray, M., *et al*. 2017. Tuberculosis. *In*: *Major Infectious Diseases*. 3rd ed. K.K. Holmes, S. Bertozzi, B.R. Bloom, & P. Jha, Eds. Washington (DC): *The International Bank for Recons*. 233–313. DOI: 10.1596/978-1-4648-0524-0/ch11.

Blouin, Y., Cazajous, G., Dehan, C., Soler, C., Vong, R., Hassan, M.O., Hauck, Y., Boulais, C., *et al*. 2014. Progenitor "*Mycobacterium canettii*" Clone Responsible for Lymph Node Tuberculosis Epidemic, Djibouti. *Emerging Infectious Diseases*. 20(1):21–28. DOI: 10.3201/eid2001.130652.

Bolger, A.M., Lohse, M. & Usadel, B. 2014. Trimmomatic: A Flexible Trimmer for Illumina Sequence Data. *Bioinformatics*. 30(15):2114–2120. DOI: 10.1093/bioinformatics/btu170.

Boot, M., Sparrius, M., Jim, K.K., Commandeur, S., Speer, A., Van De Weerd, R. & Bitter, W. 2016. *iniBAC* Induction is Vitamin B_{12}- and MutAB-Dependent in *Mycobacterium marinum*. *Journal of Biological Chemistry*. 291(38):19800–19812. DOI: 10.1074/jbc.M116.724088.

Boritsch, E.V.A.C. & Brosch, R. 2016. Evolution of *Mycobacterium tuberculosis*: New Insights into Pathogenicity and Drug Resistance. *Microbiology Spectrum*. 4(5):1–20. DOI: 10.1128/microbiolspec.TBTB2-0020-2016.

Boritsch, E.C., Supply, P., Honoré, N., Seeman, T., Stinear, T.P. & Brosch, R. 2014. A Glimpse into the Past and Predictions for the Future: The Molecular Evolution of the Tuberculosis Agent. *Molecular Microbiology*. 93(5):835–852. DOI: 10.1111/mmi.12720.

Boritsch, E.C., Frigui, W., Cascioferro, A., Malaga, W., Etienne, G., Laval, F., Pawlik, A., Le Chevalier, F., *et al*. 2016. Pks5-Recombination-Mediated Surface Remodelling in *Mycobacterium tuberculosis* Emergence. *Nature Microbiology*. 1(2). DOI: 10.1038/nmicrobiol.2015.19.

Boritsch, E.C., Khanna, V., Pawlik, A., Honoré, N., Navas, V.H., Ma, L., Bouchier, C., Seemann, T., *et al*. 2016. Key Experimental Evidence of Chromosomal DNA Transfer Among Selected Tuberculosis-Causing Mycobacteria. *Proceedings of the National Academy of Sciences*. 113(35):9876–9881. DOI: 10.1073/pnas.1604921113.

Bos, K.I., Harkins, K.M., Herbig, A., Coscolla, M., Weber, N., Comas, I., Forrest, S.A., Bryant, J.M., *et al*. 2014. Pre-Columbian Mycobacterial Genomes Reveal Seals

as A Source of New World Human Tuberculosis. *Nature*. 514(7253):494–497. DOI: 10.1038/nature13591.

Bottai, D. & Brosch, R. 2009. Mycobacterial PE, PPE And ESX Clusters: Novel Insights into The Secretion of These Most Unusual Protein Families. *Molecular Microbiology*. 73(3):325–328. DOI: 10.1111/j.1365-2958.2009.06784.x.

Breaker, R.R. 2012. Riboswitches and The RNA World. *Cold Spring Harbor Perspectives in Biology*. 4(2):1–15. DOI: 10.1101/cshperspect.a003566.

Brennan, M.J. 2017. The Enigmatic PE/PPE Multigene Family of Mycobacteria and Tuberculosis Vaccination. *Infection and Immunity*. 85(6):1–8. DOI: 10.1128/IAI.00969-16.

Bridwell-Rabb, J., Grell, T.A.J. & Drennan, C.L. 2018. A Rich Man, Poor Man Story of S -Adenosylmethionine and Cobalamin Revisited. *Annual Review of Biochemistry*. 87:555–584. DOI: 10.1146/annurev-biochem-062917-012500.

Brosch, R., Gordon, S. V., Marmiesse, M., Brodin, P., Buchrieser, C., Eiglmeier, K., Garnier, T., Gutierrez, C., *et al*. 2002. A New Evolutionary Scenario for The *Mycobacterium tuberculosis* Complex. *Proceedings of the National Academy of Sciences*. 99(6):3684–3689. DOI: 10.1073/pnas.052548299.

Brynildsrud, O.B., Pepperell, C.S., Suffys, P., Grandjean, L., Monteserin, J., Debech, N., Bohlin, J., Alfsnes, K., *et al*. 2018. Global Expansion of *Mycobacterium tuberculosis* Lineage 4 Shaped by Colonial Migration and Local Adaptation. *Science Advances*. 4(10):eaat5869. DOI: 10.1126/sciadv.aat5869.

Cadieux, N., Bradbeer, C., Reeger-Schneider, E., Köster, W., Mohanty, A.K., Wiener, M.C. & Kadner, R.J. 2002. Identification of The Periplasmic Cobalamin-Binding Protein BtuF of *Escherichia coli*. *Journal of Bacteriology*. 184(3):706–717. DOI: 10.1128/JB.184.3.706-717.2002.

Cambau, E. & Drancourt, M. 2014. Steps Towards the Discovery of *Mycobacterium tuberculosis* by Robert Koch, 1882. *Clinical Microbiology and Infection*. 20(3):196–201. DOI: 10.1111/1469-0691.12555.

Cingolani, P., Platts, A., Wang, L.L., Coon, M., Nguyen, T., Wang, L., Land, S.J., Lu, X., *et al*. 2012. A Program for Annotating and Predicting the Effects of Single Nucleotide Polymorphisms, SnpEff: SNPs in the Genome of *Drosophila melanogaster* strain w^{1118}; *iso-2*; *iso-3*. *Fly*. 6(2):80–92. DOI: http://dx.doi.org/10.4161/fly.19695.

Cole, S.T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S. V, Eiglmeier, K., *et al*. 1998. Deciphering the Biology of *Mycobacterium tuberculosis* From the Complete Genome Sequence. *Nature*. 393(6685):537–544. DOI: 10.1038/31159.

Cole, S.T., Eiglmeier, K., Parkhill, J., James, K.D., Thomson, N.R., Wheeler, P.R., Honoré, N., Garnier, T., *et al*. 2001. Massive Gene Decay in The Leprosy Bacillus. *Nature*. 409:1007–1011. DOI: 10.1038/35059006.

Comas, I., Coscolla, M., Luo, T., Borrell, S., Holt, K.E., Parkhill, J., Malla, B., Berg, S., *et al*. 2013. Out-Of-Africa Migration and Neolithic Co-Expansion of *Mycobacterium tuberculosis* With Modern Humans. *Nature Genetics*. 45(10):1176–1182. DOI: 10.1038/ng.2744.

Comas, I., Hailu, E., Kiros, T., Bekele, S., Mekonnen, W., Gumi, B., Tschopp, R., Ameni, G., *et al*. 2015. Population Genomics of *Mycobacterium tuberculosis* in Ethiopia Contradicts the Virgin Soil Hypothesis for Human Tuberculosis in Sub-Saharan Africa. *Current Biology*. 25(24):3260–3266. DOI: 10.1016/j.cub.2015.10.061.

Coscolla, M. & Gagneux, S. 2014. Consequences of genomic diversity in *Mycobacterium tuberculosis*. *Seminars in Immunology*. 26(6):431–444. DOI: 10.1016/j.smim.2014.09.012.

Coscolla, M., Lewin, A., Metzger, S., Maetz-Rennsing, K., Calvignac-Spencer, S., Nitsche, A., Dabrowski, P.W., Radonic, A., *et al*. 2013. Novel *Mycobacterium tuberculosis* Complex Isolate from A Wild Chimpanzee. *Emerging Infectious Diseases*. 19(6):969–976. DOI: 10.3201/eid1906.121012.

Crespo, A., Blanco-Cabra, N. & Torrents, E. 2018. Aerobic Vitamin B_{12} Biosynthesis Is Essential for Pseudomonas aeruginosa Class II Ribonucleotide Reductase Activity During Planktonic and Biofilm Growth. *Frontiers in Microbiology*. 9(May):986. DOI: 10.3389/fmicb.2018.00986.

Croft, M.T., Lawrence, A.D., Raux-Deery, E., Warren, M.J. & Smith, A.G. 2005. Algae Acquire Vitamin B_{12} Through A Symbiotic Relationship with Bacteria. *Nature*. 438(7064):90–93. DOI: 10.1038/nature04056.

Daniel, T.M. 2015. Jean-Antoine Villemin And the Infectious Nature of Tuberculosis. *International Journal of Tuberculosis and Lung Disease*. 19(3):267–268.

Daniel, J., Maamar, H., Deb, C., Sirakova, T.D. & Kolattukudy, P.E. 2011. *Mycobacterium tuberculosis* Uses Host Triacylglycerol to Accumulate Lipid Droplets and Acquires A Dormancy-Like Phenotype in Lipid-Loaded Macrophages. *PLoS Pathogens*. 7(6). DOI: 10.1371/journal.ppat.1002093.

Danilchanka, O., Sun, J., Pavlenok, M., Maueroder, C., Speer, A., Siroy, A., Marrero, J., Trujillo, C., *et al*. 2014. An Outer Membrane Channel Protein of *Mycobacterium tuberculosis* With Exotoxin Activity. *Proceedings of the National Academy of Sciences*. 111(18):6750–6755. DOI: 10.1073/pnas.1400136111.

Danilchanka, O., Pires, D., Anes, E. & Niederweis, M. 2015. The *Mycobacterium tuberculosis* Outer Membrane Channel Protein CpnT Confers Susceptibility to Toxic Molecules. *Antimicrobial Agents and Chemotherapy*. 59(4):2328–2336. DOI: 10.1128/AAC.04222-14.

Dassanayake, R.S., Cabelli, D.E. & Brasch, N.E. 2013. Pulse Radiolysis Studies on The Reaction of The Reduced Vitamin B_{12} Complex Cob(II)Alamin With Superoxide. *ChemBioChem*. 14(9):1081–1083. DOI: 10.1002/cbic.201300229.

DeJesus, M.A., Gerrick, E.R., Xu, W., Park, S.W., Long, J.E., Boutte, C.C., Rubin, E.J., Schnappinger, D., *et al*. 2017. Comprehensive Essentiality Analysis of The *Mycobacterium tuberculosis* Genome Via Saturating Transposon Mutagenesis. *mBio*. 8(1):1–2. DOI: 10.1128/mBio.02133-16.

DeJesus, M.A., Ambadipudi, C., Baker, R., Sassetti, C. & Ioerger, T.R. 2015. TRANSIT – A Software Tool for Himar1 Tn-seq Analysis. *PLoS Computational Biology*. 11(10):1–17. DOI: 10.1371/journal.pcbi.1004401.

Delogu G, Cole ST, B.R. 2008. The PE And PPE Protein Families of *Mycobacterium tuberculosis*. *Handbook of tuberculosis*. 131–150.

Dixon, M.M., Huang, S., Matthews, R.G. & Ludwig, M. 1996. The Structure of the C-Terminal Domain of Methionine Synthase: Presenting S-Adenosylmethionine for Reductive Methylation of B_{12}. *Structure*. 4(11):1263–1275. DOI: 10.1016/S0969-2126(96)00135-9.

Domenech, P., Kobayashi, H., Levier, K., Walker, G.C. & Barry, C.E. 2009. Baca, An ABC Transporter Involved in Maintenance of Chronic Murine Infections with *Mycobacterium tuberculosis*. *Journal of Bacteriology*. 191(2):477–485. DOI: 10.1128/JB.01132-08.

Dowling, D.P., Miles, Z.D., Köhrer, C., Maiocco, S.J., Elliott, S.J., Bandarian, V. & Drennan, C.L. 2016. Molecular Basis of Cobalamin-Dependent RNA Modification. *Nucleic Acids Research*. 44(20):9965–9976. DOI: 10.1093/nar/gkw806.

Doxey, A.C., Kurtz, D.A., Lynch, M.D., Sauder, L.A. & Neufeld, J.D. 2015. Aquatic Metagenomes Implicate Thaumarchaeota in Global Cobalamin Production. *The ISME Journal*. 9:461–471. DOI: 10.1038/ismej.2014.142.

Drain, P.K., Bajema, K.L., Dowdy, D., Dheda, K., Naidoo, K., Schumacher, S.G., Ma, S., Meermeier, E., *et al*. 2018. Incipient and Subclinical Tuberculosis: A Clinical Review of Early Stages and Progression of Infection. *Clinical Microbiology Reviews*. 31(4):1–24. DOI: 10.1128/CMR.00021-18.

Drennan, C.L., Huang, S., Drummond, J.T., Matthews, R.G. & Ludwig, M.L. 1994. How A Protein Binds B_{12}: A 3.0A X-Ray Structure of B_{12}-Binding Domains of Methionine Synthase. *Science*. 266(10):1660–1674.

Ehrt, S. & Schnappinger, D. 2009. Mycobacterial Survival Strategies in the Phagosome: Defense Against Host Stresses. *Cell Microbiology*. 11(8):1170–1178. DOI: 10.1111/j.1462-5822.2009.01335.x.

El-Gebali, S., Mistry, J., Bateman, A., Eddy, S.R., Luciani, A., Potter, S.C., Qureshi, M., Richardson, L.J., *et al*. 2018. The Pfam Protein Families Database In 2019. *Nucleic Acids Research*. DOI: 10.1093/nar/gky995.

Eoh, H. & Rhee, K.Y. 2014. Methylcitrate Cycle Defines the Bactericidal Essentiality of Isocitrate Lyase for Survival of *Mycobacterium tuberculosis* on Fatty Acids.

Proceedings of the National Academy of Sciences. 111(13):4976–4981. DOI: 10.1073/pnas.1400390111.

Escalante-Semerena, J. C., Suh, S. J., Roth, J. R. 1990. *cobA* Function is required for both *de novo* cobalamin biosynthesis and assimilation of exogenous corrinoids in *Salmonella typhimurium*. *Journal of Bacteriology*. 172: 273-280. DOI: 10.1128/jb.172.1.273-280.1990

Esmail, A., Sabur, N.F., Okpechi, I. & Dheda, K. 2018. Management of Drug-Resistant Tuberculosis in Special Subpopulations Including Those with HIV Co-Infection, Pregnancy, Diabetes, Organ-Specific Dysfunction, and in the Critically Ill. *Journal of Thoracic Disease*. 10(5):3102–3118. DOI: 10.21037/jtd.2018.05.11.

Fang, H., Kang, J. & Zhang, D. 2017. Microbial Production of Vitamin B_{12}: A Review and Future Perspectives. *Microbial Cell Factories*. 16(1):15. DOI: 10.1186/s12934-017-0631-y.

Ferguson, S.J. 2016. SAM – A Helping Hand in Many Places. *FEBS Letters*. 590(June):2536–2537. DOI: 10.1002/1873-3468.12249/.

Ferla, M.P. & Patrick, W.M. 2014. Bacterial Methionine Biosynthesis. *Microbiology (United Kingdom)*. 160(PART 8):1571–1584. DOI: 10.1099/mic.0.077826-0.

Ferrer, J.L., Ravanel, S., Robert, M. & Dumas, R. 2004. Crystal Structures of Cobalamin-Independent Methionine Synthase Complexed with Zinc, Homocysteine, and Methyltetrahydrofolate. *Journal of Biological Chemistry*. 279(43):44235–44238. DOI: 10.1074/jbc.C400325200.

Fleischmann, R.D., Dodson, R.J., Haft, D.H., Merkel, J.S., Nelson, W.C. & Fraser, C.M. 2006. *Direct submission*. Bethesda, MD: National Center for Biotechnology Information, NIH.

Forgacs, P., Wengenack, N.L., Hall, L., Zimmerman, S.K., Silverman, M.L. & Roberts, G.D. 2009. Tuberculosis and Trimethoprim-Sulfamethoxazole. *Antimicrobial Agents and Chemotherapy*. 53(11):4789–4793. DOI: 10.1128/AAC.01658-08.

Fu, T.M., Almqvist, J., Liang, Y.H., Li, L., Huang, Y. & Su, X.D. 2011. Crystal Structures of Cobalamin-Independent Methionine Synthase (MetE) from *Streptococcus mutans*: A Dynamic Zinc-Inversion Model. *Journal of Molecular Biology*. 412(4):688–697. DOI: 10.1016/j.jmb.2011.08.005.

Furukawa, K., Ramesh, A., Zhou, Z., Weinberg, Z., Vallery, T., Winkler, W.C. & Breaker, R.R. 2015. Bacterial Riboswitches Cooperatively Bind Ni^{2+} or Co^{2+} Ions and Control Expression of Heavy Metal Transporters. *Molecular Cell*. 57(6):1088–1098. DOI: 10.1016/j.molcel.2015.02.009.

Gagneux, S. 2018. Ecology and Evolution of *Mycobacterium tuberculosis*. *Nature Reviews Microbiology*. 16(4):202–213. DOI: 10.1038/nrmicro.2018.8.

Galagan, J.E. 2014. Genomic Insights into Tuberculosis. *Nature Reviews Genetics*. 15(5):307–320. DOI: 10.1038/nrg3664.

Van der Geize, R., Yam, K., Heuser, T., Wilbrink, M.H., Hara, H., Anderton, M.C., Sim, E., Dijkhuizen, L., *et al*. 2007. A Gene Cluster Encoding Cholesterol Catabolism in A Soil Actinomycete Provides Insight into *Mycobacterium tuberculosis* Survival in Macrophages. *Proceedings of the National Academy of Sciences*. 104(6):1947–1952. DOI: 10.1073/pnas.0605728104.

Gey van Pittius, N.C., Sampson, S.L., Lee, H., Kim, Y., Van Helden, P.D. & Warren, R.M. 2006. Evolution and Expansion of the *Mycobacterium tuberculosis* PE And PPE Multigene Families and Their Association with the Duplication of the ESAT-6 (ESX) Gene Cluster Regions. *BMC Evolutionary Biology*. 6(95). DOI: 10.1186/1471-2148-6-95.

Giedyk, M., Goliszewska, K. & Gryko, D. 2015. Vitamin B_{12} Catalysed Reactions. *Chem. Soc. Rev.* 44(11):3391–3404. DOI: 10.1039/C5CS00165J.

Gonzalez, J.C., Banerjee, R. V., Huang, S., Sumner, J.S. & Matthews, R.G. 1992. Comparison of Cobalamin-independent and Cobalamin-dependent Methionine Synthases from *Escherichia coli*: Two Solutions to the Same Chemical Problem. Biochemistry. 31(26):6045–6056. DOI: 10.1021/bi00141a013.

Gonzalo-Asensio, J., Malaga, W., Pawlik, A., Astarie-Dequeker, C., Passemar, C., Moreau, F., Laval, F., Daffe, M., *et al*. 2014. Evolutionary History of Tuberculosis Shaped by Conserved Mutations in the PhoPR Virulence Regulator. *Proceedings of the National Academy of Sciences*. 111(31):11491–11496. DOI: 10.1073/pnas.1406693111.

Gopinath, K., Venclovas, C., Ioerger, T.R., Sacchettini, J.C., McKinney, J.D., Mizrahi, V. & Warner, D.F. 2013. A vitamin B_{12} Transporter in *Mycobacterium tuberculosis*. *Open Biology*. 3(2):120175. DOI: 10.1098/rsob.120175.

Gopinath, K., Moosa, A., Mizrahi, V. & Warner, D.F. 2013. Vitamin B_{12} metabolism in *Mycobacterium tuberculosis*. *Future Microbiology*. 8(11):1405–1418.

Gopinath, K., Warner, D.F. & Mizrahi, V. 2015. Targeted Gene Knockout and Essentiality Testing by Homologous Recombination. *In*: *Mycobacteria Protocols*: Third Edition. V. 1285. T. Parish & D.M. Roberts, Eds. 131–149. DOI: 10.1007/978-1-4939-2450-9.

Goulding, C.W., Postigo, D. & Matthews, R.G. 1997. Cobalamin-dependent Methionine Synthase Is a Modular Protein with Distinct Regions for Binding Homocysteine, Methyltetrahydrofolate, Cobalamin, and Adenosylmethionine. *Biochemistry*. 36:8082–8091. DOI: 10.1021/BI9705164.

Gouzy, A., Larrouy-Maumus, G., Bottai, D., Levillain, F., Dumas, A., Wallach, J.B., Caire-Brandli, I., de Chastellier, C., *et al*. 2014. *Mycobacterium tuberculosis* Exploits Asparagine to Assimilate Nitrogen and Resist Acid Stress during Infection. *PLoS Pathogens*. 10(2). DOI: 10.1371/journal.ppat.1003928.

Greenhalgh, I. & Butler, A.R. 2017. Sanatoria Revisited: Sunlight and Health. *Journal of the Royal College of Physicians of Edinburgh*. 47(3):276–280. DOI: 10.4997/JRCPE.2017.314.

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(9):1–9. DOI: 10.1371/journal.ppat.1002251.

Griffin, J.E., Pandey, A.K., Gilmore, S.A., Mizrahi, V., McKinney, J.D., Bertozzi, C.R. & Sassetti, C.M. 2012. Cholesterol Catabolism by *Mycobacterium tuberculosis* Requires Transcriptional and Metabolic Adaptations. *Chemistry and Biology*. 19(2):218–227. DOI: 10.1016/j.chembiol.2011.12.016.

Gröschel, M.I., Sayes, F., Simeone, R., Majlessi, L. & Brosch, R. 2016. ESX Secretion Systems: Mycobacterial Evolution to Counter Host Immunity. *Nature Reviews Microbiology*. 14(11):677–691. DOI: 10.1038/nrmicro.2016.131.

Gryko, D., Giedyk, M., Jackowska, A., Równicki, M., Kolanowska, M. & Trylska, J. 2018. Vitamin B_{12} Transports Modified RNA into *E. coli* and *S. Typhimurium* Cells. *Chemical Communications*. 12:10–13. DOI: 10.1039/C8CC05064C.

Guirado, E. & Schlesinger, L.S. 2013. Modeling the *Mycobacterium tuberculosis* Granuloma – the Critical Battlefield in Host Immunity and Disease. *Frontiers in Immunology*. 4(APR):1–7. DOI: 10.3389/fimmu.2013.00098.

Gutierrez, M.C., Brisse, S., Brosch, R., Fabre, M., Omaïs, B., Marmiesse, M., Supply, P. & Vincent, V. 2005. Ancient Origin and Gene Mosaicism of the Progenitor of *Mycobacterium tuberculosis*. *PLoS Pathogens*. 1(1):0055–0061. DOI: 10.1371/journal.ppat.0010005.

Guzzo, M.B., Nguyen, H.T., Pham, T.H., Wyszczelska-Rokiel, M., Jakubowski, H., Wolff, K.A., Ogwang, S., Timpona, J.L., *et al*. 2016. Methylfolate Trap Promotes Bacterial Thymineless Death by Sulfa Drugs. *PLoS Pathogens*. 12(10):1–28. DOI: 10.1371/journal.ppat.1005949.

Haller, A., Soulière, M.F. & Micura, R. 2011. The Dynamic Nature of RNA as Key to Understanding Riboswitch Mechanisms. *Accounts of Chemical Research*. 44(12):1339–1348. DOI: 10.1021/ar200035g.

Halpern, J. 1985. Mechanisms of coenzyme B_{12}-dependent rearrangements. *Science*. 227(4689):869–875. doi:10.1126/science.2857503

Hamada, M., Shibata, C., Sakurai, K., Hosoyama, A., Oji, S., Teramoto, K. & Tamura, T. 2016. Reclassification of *Amycolicicoccus subflavus* as *Hoyosella subflava* comb. nov. and Emended Descriptions of the Genus Hoyosella and *Hoyosella altamirensis*. *International Journal of Systematic and Evolutionary Microbiology*. 66(11):4711–4715. DOI: 10.1099/ijsem.0.001415.

Hazra, A.B., Han, A.W., Mehta, A.P., Mok, K.C., Osadchiy, V., Begley, T.P. & Taga, M.E. 2015. Anaerobic Biosynthesis of the Lower Ligand of Vitamin B_{12}. *Proceedings of the National Academy of Sciences*. 112(34):10792–10797. DOI: 10.1073/pnas.1509132112.

van Helden, P.D., Victor, T.C., Warren, R.M. & Helden, E.G. Van. 2001. Isolation of DNA from *Mycobacterium tuberculosis*. *Methods in Molecular Medicine*. 54(15):19–29. DOI: 10.1385/1-59259-147-7:019.

Hermann, C., Giddey, A.D., Nel, A.J.M., Soares, N.C. & Blackburn, J.M. 2019. Cell Wall Enrichment Unveils Proteomic Changes in the Cell Wall During Treatment of *Mycobacterium smegmatis* with Sub-lethal Concentrations of Rifampicin. *Journal of Proteomics*. 191:166–179. DOI: 10.1016/j.jprot.2018.02.019.

Hershberg, R., Lipatov, M., Small, P.M., Sheffer, H., Niemann, S., Homolka, S., Roach, J.C., Kremer, K., *et al*. 2008. High Functional Diversity in *Mycobacterium tuberculosis* Driven by Genetic Drift and Human Demography. *PLoS Biology*. 6(12):2658–2671. DOI: 10.1371/journal.pbio.0060311.

Horswill, A. R., Dudding, Andrea R., Escalante-Semerena, J. C. 2001. Studies of Propionate Toxicity in *Salmonella enterica* Identify 2-Methylcitrate as a Potent Inhibitor of Cell Growth. *Journal of Biological Chemistry*. 276(22):19094-19101. DOI: 10.1074/jbc.M100244200

Houben, R.M.G.J. & Dodd, P.J. 2016. The Global Burden of Latent Tuberculosis Infection: A Re-estimation Using Mathematical Modelling. *PLOS Medicine*. 13(10):e1002152. DOI: 10.1371/journal.pmed.1002152.

Huang, L., Nazarova, E. V, Tan, S., Liu, Y. & Russell, D.G. 2018. Growth of *Mycobacterium tuberculosis in vivo* Segregates with Host Macrophage Metabolism and Ontogeny. *Journal of Experimental Medicine*. 215(4):1135–1152. DOI: 10.1084/jem.20172020.

Huang, T.S., Kunin, C.M., Yan, B.S., Chen, Y.S., Lee, S.S.J. & Syu, W. 2012. Susceptibility of *Mycobacterium tuberculosis* to Sulfamethoxazole, Trimethoprim and Their Combination Over A 12 Year Period in Taiwan. *Journal of Antimicrobial Chemotherapy*. 67(3):633–637. DOI: 10.1093/jac/dkr501.

Hunter, R.L. 2018. The Pathogenesis of Tuberculosis: The Early Infiltrate of Post-primary (Adult Pulmonary) Tuberculosis: A Distinct Disease Entity. *Frontiers in Immunology*. 9(2108):1–9. DOI: 10.3389/fimmu.2018.02108.

Ignatov, D. & Johansson, J. 2017. RNA-mediated Signal Perception in Pathogenic Bacteria. *Wiley Interdisciplinary Reviews: RNA*. 8(6):1–25. DOI: 10.1002/wrna.1429.

Imperiale, B.R., Moyano, R.D., Di Giulio, A.B., Romero, M.A., Alvarado Pinedo, M.F., Santangelo, M.P., Travería, G.E., Morcillo, N.S., *et al*. 2017. Genetic Diversity of *Mycobacterium avium* Complex Strains Isolated in Argentina by MIRU-VNTR. *Epidemiology and Infection*. 145(7):1382–1391. DOI: 10.1017/S0950268817000139.

James, B.W., Williams, A. & Marsh, P.D. 2000. The Physiology and Pathogenicity of *Mycobacterium tuberculosis* Grown Under Controlled Conditions in A Defined Medium. *Journal of Applied Microbiology*. 88(4):669–677. DOI: 10.1046/j.1365-2672.2000.01020.x.

Johnson, C. L. V., Pechonick, E., Park, S. D., Havemann, G. D., Leal, N. A., Bobik, T. A. 2001. Functional Genomic, Biochemical, and Genetic Characterization of the Salmonella *pduO* Gene, an ATP:Cob(I)alamin Adenosyltransferase Gene. *Journal of Bacteriology*. 183(5): 1577-1584. DOI: 10.1128/JB.183.5.1577

Johnson Jr, J.E., Reyes, F.E., Polaski, J.T. & Batey, R.T. 2012. B_{12} Cofactors Directly Stabilize An mRNA Regulatory Switch. *Nature*. 492(7427):133–137. DOI: 10.1038/nature11607.

Jones, A.R. 2017. The Photochemistry and Photobiology of Vitamin B_{12}. *Photochem. Photobiol. Sci.* 16(6):820–834. DOI: 10.1039/C7PP00054E.

Jones, P., Binns, D., Chang, H.Y., Fraser, M., Li, W., McAnulla, C., McWilliam, H., Maslen, J., *et al*. 2014. InterProScan 5: Genome-scale Protein Function Classification. *Bioinformatics.* 30(9):1236–1240. DOI: 10.1093/bioinformatics/btu031.

Kang, Z., Zhang, J., Zhou, J., Qi, Q., Du, G. & Chen, J. 2012. Recent Advances in Microbial Production of δ-aminolevulinic Acid and Vitamin B_{12}. *Biotechnology Advances*. 30(6):1533–1542. DOI: 10.1016/j.biotechadv.2012.04.003.

Kapopoulou, A., Lew, J.M. & Cole, S.T. 2011. The MycoBrowser Portal: A Comprehensive and Manually Annotated Resource for Mycobacterial Genomes. *Tuberculosis*. 91(1):8–13. DOI: 10.1016/j.tube.2010.09.006.

Karasseva, V., Weiszfeiler, J.G. & Lengyel, Z. 1977. Synthesis of Vitamin B_{12} by Various Species of Mycobacteria. *Zentralblatt für Bakteriologie, Parasitenkunde, Infektionskrankheiten und Hygiene.* Erste Abte (Reihe A: Medizinische Mikrobiologie und Parasitologie 239):514–520.

Kaufmann, S.H. 2001. How Can Immunology Contribute to the Control of Tuberculosis? *Nature Reviews Immunology*. 1(1):20–30. DOI: 10.1038/35095558.

Keck, B., Munder, M. & Renz, P. 1998. Biosynthesis of Cobalamin in *Salmonella typhimurium*: Transformation of Riboflavin into the 5,6-dimethylbenzimidazole Moiety. *Archives of Microbiology*. 171(1):66–68. DOI: 10.1007/s002030050679.

Koch, A.S., Brites, D., Stucki, D., Evans, J.C., Seldon, R., Heekes, A., Mulder, N., Nicol, M., *et al*. 2017. The Influence of HIV on the Evolution of *Mycobacterium tuberculosis*. *Molecular Biology and Evolution*. 34(7):1654–1668. DOI: 10.1093/molbev/msx107.

Koeck, J.L., Fabre, M., Simon, F., Daffé, M., Garnotel, É., Matan, A.B., Gérôme, P., Bernatas, J.J., *et al*. 2011. Clinical Characteristics of the Smooth Tubercle Bacilli "*Mycobacterium canettii*" Infection Suggest the Existence of An Environmental

Reservoir. *Clinical Microbiology and Infection*. 17(7):1013–1019. DOI: 10.1111/j.1469-0691.2010.03347.x.

Korkhov, V.M., Mireku, S.A., Veprintsev, D.B. & Locher, K.P. 2014. Structure of AMP-PNP-bound BtuCD and Mechanism of ATP-powered Vitamin B_{12} Transport by BtuCD-F. *Nature Structural and Molecular Biology*. 21(12):1097–1099. DOI: 10.1038/nsmb.2918.

Koutmos, M., Pejchal, R., Bomer, T.M., Matthews, R.G., Smith, J.L. & Ludwig, M.L. 2008. Metal Active Site Elasticity Linked to Activation of Homocysteine in Methionine Synthases. *Proceedings of the National Academy of Sciences of the United States of America*. 105(9):3286–91. DOI: 10.1073/pnas.0709960105.

Koutmos, M., Datta, S., Pattridge, K. A., Smith, J. L., Matthews, R. G. 2009. Insights into the reactivation of cobalamin-dependent methionine synthase. *Proceedings of the National Academy of Sciences*. 106(14): 18527-18532. DOI: 10.1073/pnas.0906132106

Krebs, H., Hems, R. & Tyler, B. 1976. The Regulation of Folate and Methionine Metabolism. *The Biochemical Journal*. 158(2):341–353.

Kucukyildirim, S., Long, H., Sung, W., Miller, S.F., Doak, T.G. & Lynch, M. 2016. The Rate and Spectrum of Spontaneous Mutations in *Mycobacterium smegmatis*, a Bacterium Naturally Devoid of the Postreplicative Mismatch Repair Pathway. *G3: Genes, Genomes, Genetics*. 6(7):2157–2163. DOI: 10.1534/g3.116.030130.

Lawrence, A. D., Deery, E., McLean, K. J., Munro, A. W., Pickersgill, R. W., Rigby, S. E. J., Warren, M. J. 2008. Identification, characterization, and structure/function analysis of a corrin reductase involved in adenosylcobalamin biosynthesis. *Journal of Biological Chemistry*. 283(16):10813-10821. DOI: 10.1074/jbc.M710431200

Lawrence, A.D., Nemoto-smith, E., Deery, E., Baker, J.A., Schroeder, S., Brown, D.G., Tullet, J.M.A., Howard, M.J., *et al.* 2018. Construction of Fluorescent Analogs to Follow the Uptake and Distribution of Cobalamin (Vitamin B_{12}) in Bacteria, Worms, and Plants. *Cell Chemical Biology*. 25(8):941–951. DOI: 10.1016/j.chembiol.2018.04.012.

Lawrence, J.G., Roth, J.R., Lawrence, J.G. & Roth, J.R. 1995. The Cobalamin (Coenzyme B_{12}) Biosynthetic Genes of *Escherichia coli*. *Journal of Bacteriology*. 177(22):6371–6380.

Lee, W., VanderVen, B.C., Fahey, R.J. & Russell, D.G. 2013. Intracellular *Mycobacterium tuberculosis* Exploits Host-derived Fatty Acids to Limit Metabolic Stress. *Journal of Biological Chemistry*. 288(10):6788–6800. DOI: 10.1074/jbc.M112.445056.

Lewinson, O., Lee, A.T., Locher, K.P. & Rees, D.C. 2010. A Distinct Mechanism for the ABC Transporter BtuCD-BtuF Revealed by the Dynamics of Complex Formation. *Nature Structural and Molecular Biology*. 17(3):332–338. DOI: 10.1038/nsmb.1770.

Li, H. & Durbin, R. 2009. Fast and Accurate Short Read Alignment with Burrows-Wheeler Transform. *Bioinformatics*. 25(14):1754–1760. DOI: 10.1093/bioinformatics/btp324.

Li, H. & Durbin, R. 2010. Fast and Accurate Long-read Alignment with Burrows-Wheeler Transform. *Bioinformatics*. 26(5):589–595. DOI: 10.1093/bioinformatics/btp698.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., *et al*. 2009. The Sequence Alignment/Map Format and SAMtools. *Bioinformatics*. 25(16):2078–2079. DOI: 10.1093/bioinformatics/btp352.

Li, K., Li, G., Bradbury, L.M.T., Hanson, A.D. & Bruner, S.D. 2016. Crystal Structure of the Homocysteine Methyltransferase MmuM from *Escherichia coli*. *Biochemical Journal*. 473(3):277–284. DOI: 10.1042/BJ20150980.

Libecco, J.A. & Powell, K.R. 2004. Trimethoprim/Sulfamethoxazole: Clinical Update. *Pediatrics in Review*. 25(11):375–380.

Lin, P.L. & Flynn, J.L. 2018. The End of the Binary Era: Revisiting the Spectrum of Tuberculosis. *The Journal of Immunology*. 201(9):2541–2548. DOI: 10.4049/jimmunol.1800993.

Lin, P.L., Ford, C.B., Coleman, M.T., Myers, A.J., Gawande, R., Ioerger, T., Sacchettini, J., Fortune, S.M., *et al*. 2014. Sterilization of Granulomas isCcommon in Active and Latent Tuberculosis Despite Within-host Variability in Bacterial Killing. *Nature Medicine*. 20(1):75–9. DOI: 10.1038/nm.3412.

Locher, K.P., Lee, A.T. & Rees, D.C. 2002. The *E. coli* BtuCD Structure: A Framework for ABC Transporter Architecture and Mechanism. *Science*. 296(5570):1091–1098. DOI: 10.1126/science.1071142.

Lopez, B., Aguilar, D., Orozco, H., Burger, M., Espitia, C., Ritacco, V., Barrera, L., Kremer, K., *et al.* 2003. A Marked Difference in Pathogenesis and Immune Response Induced by Different *Mycobacterium tuberculosis* Genotypes. *Clinical and Experimental Immunology*. 133:30–37.

Lovewell, R.R., Sassetti, C.M. & VanderVen, B.C. 2016. Chewing the Fat: Lipid Metabolism and Homeostasis During *M. tuberculosis* Infection. Current Opinion in Microbiology. 29:30–36. DOI: 10.1016/j.mib.2015.10.002.

Lussier, A., Bastet, L., Chauvier, A. & Lafontaine, D.A. 2015. A Kissing Loop Is Important for BtuB Riboswitch Ligand Sensing and Regulatory Control. *Journal of Biological Chemistry*. 290(44):26739–26751. DOI: 10.1074/jbc.M115.684134.

Markowitz, V.M., Chen, I.M.A., Palaniappan, K., Chu, K., Szeto, E., Grechkin, Y., Ratner, A., Jacob, B., *et al*. 2012. IMG: The Integrated Microbial Genomes Database and Comparative Analysis System. *Nucleic Acids Research*. 40(D1):115–122. DOI: 10.1093/nar/gkr1044.

Marrero, J., Rhee, K.Y., Schnappinger, D., Pethe, K. & Ehrt, S. 2010. Gluconeogenic Carbon Flow of Tricarboxylic Acid Cycle Intermediates Is Critical for *Mycobacterium tuberculosis* to Establish and Maintain Infection. *Proceedings of the National Academy of Sciences*. 107(21):9819–9824. DOI: 10.1073/pnas.1000715107.

Mattes, T. A., Deery, E., Warren, M. J., Escalante-Semerena, J. C. 2017. Cobalamin Biosynthesis and Insertion. In *Encyclopedia of Inorganic and Bioinorganic Chemistry*. Scott, R.A. (ed.) Chichester: John Wiley & Sons Ltd, pp. 1–24. DOI: 10.1002/9781119951438.eibc2489

Matthews, R.G. & Goulding, C.W. 1997. Enzyme-catalyzed Methyl Transfers to Thiols: The Role of Zinc. *Current Opinion in Chemical Biology*. 1(3):332–339. DOI: 10.1016/S1367-5931(97)80070-1.

McKinney, J.D., Höner zu Bentrup, K., Muñoz-Elías, E.J., Miczak, a, Chen, B., Chan, W.T., Swenson, D., Sacchettini, J.C., *et al*. 2000. Persistence of *Mycobacterium tuberculosis* in Macrophages and Mice Requires the Glyoxylate Shunt Enzyme Isocitrate Lyase. *Nature*. 406(6797):735–738. DOI: 10.1038/35021074.

Merker, M., Blin, C., Mona, S., Duforet-Frebourg, N., Lecher, S., Willery, E., Blum, M.G.B., Rüsch-Gerdes, S., *et al*. 2015. Evolutionary History and Global Spread of the *Mycobacterium tuberculosis* Beijing Lineage. *Nature Genetics*. 47(3):242–249. DOI: 10.1038/ng.3195.

Minias, A., Minias, P. & Dziadek, J. 2018. Purifying Selective Pressure Suggests the Functionality of a Vitamin B_{12} Biosynthesis Pathway in a Global Population of *Mycobacterium tuberculosis*. *Genome Biology and Evolution*. 10(9):2326–2337. DOI: 10.1093/gbe/evy153/5061319.

Mokrousov, I. 2013. Insights into the Origin, Emergence, and Current Spread of a Successful Russian Clone of *Mycobacterium tuberculosis*. *Clinical Microbiology Reviews*. 26(2):342–360. DOI: 10.1128/CMR.00087-12.

Moore, S.J. & Warren, M.J. 2012. The Anaerobic Biosynthesis of Vitamin B_{12}. *Biochemical Society Transactions*. 40(3):581–586. DOI: 10.1042/BST20120066.

Mostowy, S., Cousins, D., Brinkman, J., Aranaz, A. & Behr, M.A. 2002. Genomic Deletions Suggest a Phylogeny for the *Mycobacterium tuberculosis* Complex. *The Journal of Infectious Diseases*. 186(1):74–80. DOI: 10.1086/341068.

Muñoz-Elías, E.J., Upton, A.M., Cherian, J. & McKinney, J.D. 2006. Role of the Methylcitrate Cycle in *Mycobacterium tuberculosis* Metabolism, Intracellular Growth, and Virulence. *Molecular Microbiology*. 60(5):1109–1122. DOI: 10.1111/j.1365-2958.2006.05155.x.

Nahvi, A., Barrick, J.E. & Breaker, R.R. 2004. Coenzyme B_{12} Riboswitches are Widespread Genetic Control Elements in Prokaryotes. *Nucleic Acids Research*. 32(1):143–150. DOI: 10.1093/nar/gkh167.

Nebenzahl-Guimaraes, H., Yimer, S.A., van Soolingen, D., Brosch, R., Holm-Hansen, C. & de Beer, J. 2016. Genomic Characterization of *Mycobacterium tuberculosis* Lineage 7 and a Proposed Name: '*Aethiops vetus*'. *Microbial Genomics*. 2(6):1–8. DOI: 10.1099/mgen.0.000063.

Nechooshtan, G., Elgrably-Weiss, M., Sheaffer, A., Westhof, E. & Altuvia, S. 2009. A pH-Responsive Riboregulator. *Genes and Development*. 23(22):2650–2662. DOI: 10.1101/gad.552209.

Neyrolles, O., Wolschendorf, F., Mitra, A. & Niederweis, M. 2015. Mycobacteria, Metals, and the Macrophage. *Immunological Reviews*. 364:249–263.

Niederweis, M., Danilchanka, O., Huff, J., Hoffmann, C. & Engelhardt, H. 2010. Mycobacterial Outer Membranes: In Search of Proteins. *Trends in Microbiology*. 18(3):109–116. DOI: 10.1016/j.tim.2009.12.005.

Nielsen, M.J., Rasmussen, M.R., Andersen, C.B.F., Nexø, E. & Moestrup, S.K. 2012. Vitamin B_{12} Transport from Food to the Body's Cells – a Sophisticated, Multistep Pathway. *Nature Reviews Gastroenterology & Hepatology*. 9(6):345–54. DOI: 10.1038/nrgastro.2012.76.

Noinaj, N., Guillier, M., Barnard, T.J. & Buchanan, S.K. 2010. TonB-dependent Transporters: Regulation, Structure, and Function. *Annual Review of Microbiology*. 64(1):43–60. DOI: 10.1146/annurev.micro.112408.134247.

Nou, X. & Kadner, R. J. 2000. Adenosylcobalamin inhibits ribosome binding to *btuB* RNA. *Proceedings of the National Academy of Sciences*. 97(13): 7190-7195. DOI: 10.1073/pnas.130013897

Novichkov, P.S., Kazakov, A.E., Ravcheev, D.A., Leyn, S.A., Kovaleva, G.Y., Sutormin, R.A., Kazanov, M.D., Riehl, W., *et al*. 2013. RegPrecise 3.0--a Resource for Genome-scale Exploration of Transcriptional Regulation in Bacteria. *BMC Genomics*. 14(1):745. DOI: 10.1186/1471-2164-14-745.

O'Neill, M.B., Kitchen, A., Zarley, A., Aylward, W., Eldholm, V. & Pepperell, C.S. 2017. Lineage Specific Histories of *Mycobacterium tuberculosis* Dispersal in Africa and Eurasia. *bioRxiv 210161*. DOI: 10.1128/JB.00935-15.Editor.

Ogata, H., Goto, S., Sato, K., Fujibuchi, W., Bono, H. & Kanehisa, M. 1999. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Research*. 27(1):29–34. DOI: 10.1093/nar/27.1.29.

Oh-hama, T., Stolowich, N.J. & Scott, A.I. 1988. 5-Aminolevulinic Acid Formation from Glutamate via the C5 Pathway in *Clostridium thermoaceticum*. *FEBS letters*. 228(1):89–93. DOI: 10.1016/0014-5793(88)80591-X.

Orgeur, M. & Brosch, R. 2018. Evolution of Virulence in the *Mycobacterium tuberculosis* Complex. *Current Opinion in Microbiology*. 41:68–75. DOI: 10.1016/j.mib.2017.11.021.

Padmanabhan, S., Jost, M., Drennan, C.L. & Elías-Arnanz, M. 2017. A New Facet of Vitamin B_{12}: Gene Regulation by Cobalamin-based Photoreceptors. *Annual Review of Biochemistry*. 86:485–514. DOI: 10.1146/annurev-biochem-061516-044500.

Padovani, D., Labunska, T., Palfey, B. A., Ballou, D. P., Banerjee, R. 2008. Adenosyltransferase tailors and delivers coenzyme B_{12}. *Nature Chemical Biology*. 4(3):194–196. DOI:10.1038/nchembio.67

Palmer, A.M., Kamynina, E., Field, M.S. & Stover, P.J. 2017. Folate rescues vitamin B_{12} Depletion-induced Inhibition of Nuclear Thymidylate Biosynthesis and Genome Instability. *Proceedings of the National Academy of Sciences*. 114(20):e4095–E4102. DOI: 10.1073/pnas.1619582114.

Parish, T. & Stoker, N.G. 2000. Use of a Flexible Cassette Method to Generate a Double Unmarked *Mycobacterium tuberculosis tlyA plcABC* Mutant by Gene Replacement. *Microbiology*. 146(8):1969–1975. DOI: 10.1099/00221287-146-8-1969.

Parks, J.M., Johs, A., Podar, M., Bridou, R., Hurt, R. a., Smith, S.D., Tomanicek, S.J., Qian, Y., *et al*. 2013. The Genetic Basis for Bacterial Mercury Methylation. *Science*. 339(6125):1332–5. DOI: 10.1126/science.1230667.

Payne, K. A.P., Quezada, C. P., Fisher, K., Dunstan, M. S., Collins, F. A., Sjuts, H., Levy, C., Hay, S., Rigby, S.E.J.. Leys, D. 2015. Reductive dehalogenase structure suggests a mechanism for B_{12}-dependent dehalogenation. 517(7535): 513-516. DOI: 10.1038/nature13901

Pejchal, R. & Ludwig, M.L. 2005. Cobalamin-independent Methionine Synthase (MetE): A Face-to-face Double Barrel That Evolved by Gene Duplication. *PLoS Biology*. 3(2):0254–0265. DOI: 10.1371/journal.pbio.0030031.

Perdrizet, G.A., Artsimovitch, I., Furman, R., Sosnick, T.R. & Pan, T. 2012. Transcriptional Pausing Coordinates Folding of the Aptamer Domain and the Expression Platform of a Riboswitch. *Proceedings of the National Academy of Sciences*. 109(9):3323–3328. DOI: 10.1073/pnas.1113086109.

Peselis, A. & Serganov, A. 2012. Structural Insights into Ligand Binding and Gene Expression Control by an Adenosylcobalamin Riboswitch. *Nature Structural & Molecular Biology*. 19(11):1182–1184. DOI: 10.1038/nsmb.2405.

Peyron, P., Vaubourgeix, J., Poquet, Y., Levillain, F., Botanch, C., Bardou, F., Daffé, M., Emile, J.F., *et al*. 2008. Foamy Macrophages from Tuberculous Patients' Granulomas Constitute a Nutrient-rich Reservoir for *M. tuberculosis* Persistence. *PLoS Pathogens*. 4(11):1–14. DOI: 10.1371/journal.ppat.1000204.

Polaski, J.T., Webster, S.M., Johnson, J.E. & Batey, R.T. 2017. Cobalamin Riboswitches Exhibit a Broad Range of Ability to Discriminate Between Methylcobalamin and Adenosylcobalamin. *Journal of Biological Chemistry*. 292(28):11650–11658. DOI: 10.1074/jbc.M117.787176.

Pym, A.S., Brodin, P., Brosch, R., Huerre, M. & Cole, S.T. 2002. Loss of RD1 Contributed to the Attenuation of the Live Tuberculosis Vaccines *Mycobacterium bovis* BCG and *Mycobacterium microti*. *Molecular Microbiology*. 46(3):709–717. DOI: 10.1046/j.1365-2958.2002.03237.x.

Quadri, L.E.N. 2014. Biosynthesis of Mycobacterial Lipids by Polyketide Synthases and Beyond. *Critical Reviews in Biochemistry and Molecular Biology*. 49(3):179–211. DOI: 10.3109/10409238.2014.896859.

Quan, T.P., Bawa, Z., Foster, D., Walker, T., Elias, C. del O., Rathod, P., MMM Informatics Group, Iqbal, Z., *et al*. 2018. Evaluation of Whole-Genome Sequencing for Mycobacterial Species Identification and Drug Susceptibility Testing in a Clinical Setting: A Large-Scale Prospective Assessment of Performance against Line Probe Assays and Phenotyping. *Journal of Clinical Microbiology*. 56(e01480-17):1–14. DOI: 10.1128/JCM.01480-17.

Rad, M.E., Bifani, P., Martin, C., Kremer, K., Samper, S., Rauzier, J., Kreiswirth, B., Blazquez, J., *et al*. 2003. Mutations in Putative Mutator Genes of *Mycobacterium tuberculosis* Strains of the W-Beijing Family. *Emerging Infectious Diseases*. 9(7):838–845. DOI: 10.3201/eid0907.020589.

Ramagopalan, S. V., Goldacre, R., Skingsley, A., Conlon, C. & Goldacre, M.J. 2013. Associations Between Selected Immune-mediated Diseases and Tuberculosis: Record-linkage Studies. *BMC Medicine*. 11(1). DOI: 10.1186/1741-7015-11-97.

Randall, P.J., Hsu, N.-J., Quesniaux, V., Ryffel, B. & Jacobs, M. 2015. *Mycobacterium tuberculosis* Infection of the 'Non-classical Immune Cell'. *Immunology and Cell Biology*. 93(9):789–795. DOI: 10.1038/icb.2015.43.

Raux, E., Lanois, A., Warren, M. J., Rambach, A., Thermes, C. 1998. Cobalamin (vitamin B_{12}) biosynthesis: identification and characterization of a *Bacillus megaterium cobI* operon. *The Biochemical Journal*. 335:159–166.

Ravnum, S. & Andersson, D. I. 1997. Vitamin B_{12} repression of the *btuB* gene in *Salmonella typhimurium* is mediated via a translational control which requires leader and coding sequences. Molecular Microbiology. 23:35-42. DOI: 10.1046/j.1365-2958.1997.1761543.x

Rempel, S., Colucci, E., Gier, J.W. de, Guskov, A. & Slotboom, D.J. 2018. Cysteine-mediated Decyanation of Vitamin B_{12} by the Predicted Membrane Transporter BtuM. *Nature Communications*. 9(3038):1–8. DOI: 10.1038/s41467-018-05441-9.

Renz, P. 1970. Riboflavin as Precursor in the Biosynthesis of 5,6-dimethylbenzimidazole Moiety of Vitamin B_{12}. *FEBS Letters*. 6(3):187–189.

Reyrat, J.-M. & Kahn, D. 2001. *Mycobacterium smegmatis*: An Absurd Model for Tuberculosis? *Trends in Microbiology*. 9(10):472–473. DOI: 10.1016/S0966-842X(01)02169-2.

Rhee, K.Y., Carvalho, L.P.S. de, Bryk, R., Ehrt, S., Marrero, J., Park, S.W., Schnappinger, D., Venugopal, A., *et al*. 2011. Central Carbon Metabolism in *Mycobacterium tuberculosis*: An Unexpected Frontier. *Trends in Microbiology*. 19(7):307–314. DOI: 10.1016/j.tim.2011.03.008.

Rindi, L. & Garzelli, C. 2014. Genetic Diversity and Phylogeny of *Mycobacterium avium*. *Infection, Genetics and Evolution*. 21:375–383. DOI: 10.1016/j.meegid.2013.12.007.

Rocco, C. J., Escalante-Semerena, J. C. 2010. In *Salmonella enterica*, 2-methylcitrate blocks gluconeogenesis. *Journal of Bacteriology*. 192(3):771-778

Rock, J.M., Hopkins, F.F., Chavez, A., Diallo, M., Chase, M.R., Gerrick, E.R., Pritchard, J.R., Church, G.M., *et al*. 2017. Programmable Transcriptional Repression in Mycobacteria Using an Orthogonal CRISPR Interference Platform. *Nature Microbiology*. 2(16274). DOI: 10.1038/nmicrobiol.2016.274.

Rodionov, D.A., Vitreschak, A.G., Mironov, A.A. & Gelfand, M.S. 2003. Comparative Genomics of the Vitamin B_{12} Metabolism and Regulation in Prokaryotes. *Journal of Biological Chemistry*. 278(42):41148–41159. DOI: 10.1074/jbc.M305837200.

Romine, M.F., Rodionov, D.A., Maezato, Y., Anderson, L.N., Nandhikonda, P., Rodionova, I.A., Carre, A., Li, X., *et al.* 2017. Elucidation of Roles for Vitamin B_{12} in Regulation of Folate, Ubiquinone, and Methionine Metabolism. *Proceedings of the National Academy of Sciences*. 114(7):E1205–E1214. DOI: 10.1073/PNAS.1612360114.

Russell, D.G. 2011. *Mycobacterium tuberculosis* and the Intimate Discourse of a Chronic Infection. *Immunological Reviews*. 240(1):252–268. DOI: 10.1111/j.1600-065X.2010.00984.x.

Russell, D.G., VanderVen, B.C., Lee, W., Abramovitch, R.B., Kim, M.J., Homolka, S., Niemann, S. & Rohde, K.H. 2010. *Mycobacterium tuberculosis* Wears What It Eats. *Cell Host and Microbe*. 8(1):68–76. DOI: 10.1016/j.chom.2010.06.002.

Sampson, S.L. 2011. Mycobacterial PE/PPE proteins at the Host-pathogen Interface. *Clinical and Developmental Immunology*. 2011(497203). DOI: 10.1155/2011/497203.

Sasaki, S., Takeshita, F., Okuda, K. & Ishii, N. 2001. Mycobacterium leprae and Leprosy: A Compendium. *Microbiology and Immunology*. 45(11):729–736. DOI: 10.1111/j.1348-0421.2001.tb01308.x.

Savvi, S., Warner, D.F., Kana, B.D., McKinney, J.D., Mizrahi, V. & Dawes, S.S. 2008. Functional Characterization of a Vitamin B_{12}-Dependent Methylmalonyl Pathway in *Mycobacterium tuberculosis*: Implications for Propionate Metabolism During Growth on Fatty Acids. *Journal of Bacteriology*. 190(11):3886–3895. DOI: 10.1128/JB.01767-07.

Schroeder, J.W., Yeesin, P., Simmons, L.A. & Wang, J.D. 2018. Sources of Spontaneous Mutagenesis in Bacteria. *Critical Reviews in Biochemistry and Molecular Biology*. 53(1):29–48. DOI: 10.1080/10409238.2017.1394262.

Serganov, A., Yuan, Y.R., Pikovskaya, O., Polonskaia, A., Malinina, L., Phan, A.T., Hobartner, C., Micura, R., *et al*. 2004. Structural Basis for Discriminative Regulation of Gene Expression by Adenine- and Guanine-sensing mRNAs. *Chemistry and Biology*. 11(12):1729–1741. DOI: 10.1016/j.chembiol.2004.11.018.

Shelton, A.N., Seth, E.C., Mok, K.C., Han, A.W., Jackson, S.N., Haft, D.R. & Taga, M.E. 2018. Uneven Distribution of Cobamide Biosynthesis and Dependence in Bacteria Predicted by Comparative Genomics. *The ISME Journal*. DOI: doi: http://dx.doi.org/10.1101/342006.

Shepard, C.C. 1960. The Experimental Disease That Follows the Injection of Human Leprosy Bacilli into Foot-pads of Mice. *Journal of Experimental Medicine*. 112:445–454.

Sherwood, A. V & Henkin, T.M. 2016. Riboswitch-mediated Gene Regulation: Novel RNA Architectures Dictate Gene Expression Responses. *Annual Review of Microbiology*. 70:361–374. DOI: 10.1146/annurev-micro-091014-104306.

Shiloh, M.U. & DiGiuseppe Champion, P.A. 2010. To Catch a Killer. What Can Mycobacterial Models Teach Us About *Mycobacterium tuberculosis* Pathogenesis? *Current Opinion in Microbiology*. 13(1):86–92. DOI: 10.1016/j.mib.2009.11.006.

Singh, V., Brecik, M., Mukherjee, R., Evans, J.C., Svetlíková, Z., Blaško, J., Surade, S., Blackburn, J., *et al*. 2015. The Complex Mechanism of Antimycobacterial Action of 5-fluorouracil. *Chemistry and Biology*. 22(1):63–75. DOI: 10.1016/j.chembiol.2014.11.006.

Smith, A.D., Warren, M.J. & Refsum, H. 2018. Vitamin B_{12}. *Advances in Food and Nutrition Research*. 83:215–279. DOI: 10.1016/bs.afnr.2017.11.005.

Snapper, S.B., Melton, R.E., Mustafa, S., Kieser, T. & Jr, W.R.J. 1990. Isolation and Characterization of Efficient Plasmid Transformation Mutants of *Mycobacterium smegmatis*. *Molecular Microbiology*. 4(11):1911–1919. DOI: 10.1111/j.1365-2958.1990.tb02040.x.

van Soolingen, D., Hoogenboezem, T., De Haas, P.E.W., Hermans, P.W.M., Koedam, M.A., Teppema, K.S., Brennan, P.J., Besra, G.S., *et al*. 1997. A Novel Pathogenic Taxon of the *Mycobacterium tuberculosis* Complex, Canetti: Characterization of an Exceptional Isolate from Africa. *International Journal of Systematic Bacteriology*. 47(4):1236–1245. DOI: 10.1099/00207713-47-4-1236.

Soulière, M.F., Haller, A., Santner, T. & Micura, R. 2013. New Insights into Gene Regulation – High-Resolution Structures of Cobalamin Riboswitches. *Angewandte Chemie – International Edition*. 52(7):1874–1877. DOI: 10.1002/anie.201208167.

Springer, B., Bottger, E.C., Kirschner, P. & Richard J. Wallace, J. 1995. Phylogeny of the *Mycobacterium chelonae*-Like Organism Based on Partial Sequencing of the 16S rRNA Gene and Proposal of *Mycobacterium mucogenicum* sp. nov. *International Journal of Systematic Bacteriology*. 45(2):262–267.

Srivastava, S., Ernst, J.D. & Desvignes, L. 2014. Beyond Macrophages: The Diversity of Mononuclear Cells in Tuberculosis. *Immunological Reviews*. 262(1):179–192. DOI: 10.1111/imr.12217.

Sun, J., Siroy, A., Lokareddy, R.K., Speer, A., Doornbos, K.S., Cingolani, P. & Niederweis, M. 2015. The Tuberculosis Necrotizing Toxin Kills Macrophages by Hydrolyzing NAD. *Nature Structural & Molecular Biology*. 22(9):672–678. DOI: 10.1038/nsmb.3064.

Supply, P. & Brosch, R. 2017. The Biology and Epidemiology of *Mycobacterium canettii*. *In: Advances in Experimental Medicine and Biology*. V. 1019. S. Gagneux, Ed. Springer International Publishing. 27–41. DOI: 10.1007/978-3-319-64371-7.

Supply, P., Marceau, M., Mangenot, S., Roche, D., Rouanet, C., Khanna, V., Majlessi, L., Criscuolo, A., *et al*. 2013. Genomic Analysis of Smooth Tubercle Bacilli Provides Insights into Ancestry and Pathoadaptation of *Mycobacterium tuberculosis*. *Nature Genetics*. 45(2):172–179. DOI: 10.1038/ng.2517.

Tailleux, L., Waddell, S.J., Pelizzola, M., Mortellaro, A., Withers, M., Tanne, A., Castagnoli, P.R., Gicquel, B., *et al*. 2008. Probing Host Pathogen Cross-talk by Transcriptional Profiling of Both *Mycobacterium tuberculosis* and Infected Human Dendritic Cells and Macrophages. *PloS One*. 3(1):e1403. DOI: 10.1371/journal.pone.0001403.

Thanbichler, M., Neuhierl, B. & Böck, A. 1999. S-Methylmethionine Metabolism in *Escherichia coli*. *Journal of Bacteriology*. 181(2):662–665.

Tientcheu, L.D., Koch, A., Ndengane, M., Andoseh, G., Kampmann, B. & Wilkinson, R.J. 2017. Immunological Consequences of Strain Variation Within the *Mycobacterium tuberculosis* Complex. *European Journal of Immunology*. 47(3):432–445. DOI: 10.1002/eji.201646562.

Turner, R.D., Chiu, C., Churchyard, G.J., Esmail, H., Lewinsohn, D.M., Gandhi, N.R. & Fennelly, K.P. 2017. Tuberculosis Infectiousness and Host Susceptibility. *Journal of Infectious Diseases*. 216(suppl_6):S636–S643. DOI: 10.1093/infdis/jix361.

Ubhi, D.K. & Robertus, J.D. 2015. The Cobalamin-independent Methionine Synthase Enzyme Captured in a Substrate-induced Closed Conformation. *Journal of Molecular Biology*. 427(4):901–909. DOI: 10.1016/j.jmb.2014.12.014.

Ubhi, D., Kago, G., Monzingo, A.F. & Robertus, J.D. 2014. Structural Analysis of a Fungal Methionine Synthase with Substrates and Inhibitors. *Journal of Molecular Biology*. 426(8):1839–1847. DOI: 10.1016/j.jmb.2014.02.006.

Upton, A.M. & McKinney, J.D. 2007. Role of the Methylcitrate Cycle in Propionate Metabolism and Detoxification in *Mycobacterium smegmatis*. *Microbiology*. 153(12):3973–3982. DOI: 10.1099/mic.0.2007/011726-0.

Uribe-Querol, E. & Rosales, C. 2017. Control of Phagocytosis by Microbial Pathogens. *Frontiers in Immunology*. 8(1368). DOI: 10.3389/fimmu.2017.01368.

Vanderven, B.C., Huang, L.U., Rohde, K.H. & Russell, D.G. 2016. The Minimal Unit of Infection: *Mycobacterium tuberculosis* in the Macrophage. *Microbiology Spectrum*. 4(6):1–18. DOI: 10.1128/microbiolspec.TBTB2-0025-2016.

Vasava, M.S., Bhoi, M.N., Rathwa, S.K., Borad, M.A., Nair, S.G. & Patel, H.D. 2017. Drug Development Against Tuberculosis: Past, Present and Future. *Indian Journal of Tuberculosis*. 64(4):252–275. DOI: 10.1016/j.ijtb.2017.03.002.

Vergne, I., Chua, J., Lee, H.-H., Lucas, M., Belisle, J. & Deretic, V. 2005. Mechanism of Phagolysosome Biogenesis Block by Viable *Mycobacterium tuberculosis*. *Proceedings of the National Academy of Sciences*. 102(11):4033–4038. DOI: 10.1073/pnas.0409716102.

Vitreschak, A.G., Rodionov, D.A., Mironov, A.A. & Gelfand, M.S. 2003. Regulation of the Vitamin B12 Metabolism and Transport in Bacteria by a Conserved RNA Structural Element. *RNA*. 9:1084–1097. DOI: 10.1261/rna.5710303.conserved.

Vordermeier, H.M., Hewinson, R.G., Wilkinson, R.J., Wilkinson, K.A., Gideon, H.P., Young, D.B. & Sampson, S.L. 2012. Conserved Immune Recognition Hierarchy of Mycobacterial PE/PPE Proteins During Infection in Natural Hosts. *PLoS One*. 7(8). DOI: 10.1371/journal.pone.0040890.

Voss, G., Casimiro, D., Neyrolles, O., Williams, A., Kaufmann, S.H.E., McShane, H., Hatherill, M. & Fletcher, H.A. 2018. Progress and Challenges in TB Vaccine Development. *F1000Research*. 7(Mvi):199. DOI: 10.12688/f1000research.13588.1.

Warner, D.F., Savvi, S., Mizrahi, V. & Dawes, S.S. 2007. A Riboswitch Regulates Expression of the Coenzyme B_{12}-independent Methionine Synthase in *Mycobacterium tuberculosis*: Implications for Differential Methionine Synthase Function in Strains H37Rv and CDC1551. *Journal of Bacteriology*. 189(9):3655–3659. DOI: 10.1128/JB.00040-07.

Warren, M.J., Raux, E., Schubert, H.L. & Escalante-Semerena, J.C. 2002. The Biosynthesis of Adenosylcobalamin (vitamin B_{12}). *Natural Product Reports*. 19:390–412. DOI: 10.1039/b108967f.

Weerdenburg, E.M., Abdallah, A.M., Rangkuti, F., El Ghany, M.A., Otto, T.D., Adroub, S.A., Molenaar, D., Ummels, R., *et al*. 2015. Genome-wide Transposon Mutagenesis Indicates That *Mycobacterium marinum* Customizes Its Virulence Mechanisms for Survival and Replication in Different Hosts. *Infection and Immunity*. 83(5):1778–1788. DOI: 10.1128/IAI.03050-14.

Weissbach, H., Ladd, J. N., Volcani, B. E., Smyth, R. D., Barker, H. A. 1960. Structure of the Adenylcobamide Coenzyme: Degradation by Cyanide, Acid, and Light. *The Journal of Biological Chemistry*. 235(5): 1462-1473

de Wet, T.J., Gobe, I., Mhlanga, M.M. & Warner, D.F. 2018. CRISPRi-Seq for Improved Identification, Targeting and Phenotyping of Essential Mycobacterial Genes. *bioRxiv 358275*. DOI: https://doi.org/10.1101/358275.

Wheatley, R.W., Ng, K.K.S. & Kapoor, M. 2016. Fungal Cobalamin-independent Methionine Synthase: Insights from the Model Organism, *Neurospora crassa*. *Archives of Biochemistry and Biophysics*. 590:125–137. DOI: 10.1016/j.abb.2015.11.037.

WHO Tuberculosis Programme. 1994. *WHO Tuberculosis Programme: Framework for Effective Tuberculosis Control*. Geneva. Available: http://www.who.int/iris/handle/10665/58717.

Wirth, T., Hildebrand, F., Allix-Béguec, C., Wölbeling, F., Kubica, T., Kremer, K., Van Soolingen, D., Rüsch-Gerdes, S., *et al*. 2008. Origin Spread and Demography of the *Mycobacterium tuberculosis* Complex. *PLoS Pathogens*. 4(9). DOI: 10.1371/journal.ppat.1000160.

Woodson, J. D., Escalante-Semerena, J. C. 2003. CbiZ, an amidohydrolase enzyme required for salvaging the coenzyme B_{12} precursor cobinamide in archaea. *Proceedings of the National Academy of Sciences*. 101(10): 3591-3596. DOI: 10.1073/pnas.0305939101

World Health Organization. 2018. *Global tuberculosis report 2018*. Geneva. Available: http://www.who.int/iris/handle/10665/274453.

Ying, J., Wang, H., Bao, B., Zhang, Y., Zhang, J., Zhang, C., Li, A., Lu, J., *et al*. 2015. Molecular Variation and Horizontal Gene Transfer of the Homocysteine Methyltransferase Gene *mmuM* and its Distribution in Clinical Pathogens. *International Journal of Biological Sciences*. 11(1):11–21. DOI: 10.7150/ijbs.10320.

Young, D.B., Comas, I. & de Carvalho, L.P.S. 2015. Phylogenetic Analysis of Vitamin B_{12}-related Metabolism in *Mycobacterium tuberculosis*. *Frontiers in Molecular Biosciences*. 2(6):1–14. DOI: 10.3389/fmolb.2015.00006.

Zhang, Y., Rodionov, D.A., Gelfand, M.S. & Gladyshev, V.N. 2009. Comparative Genomic Analyses of Nickel, Cobalt and Vitamin B_{12} Utilization. *BMC Genomics*. 10:78. DOI: 10.1186/1471-2164-10-78.

Zignol, M., Cabibbe, A.M., Dean, A.S., Glaziou, P., Alikhanova, N., Ama, C., Andres, S., Barbova, A., *et al*. 2018. Genetic Sequencing for Surveillance of Drug Resistance in Tuberculosis in Highly Endemic Countries: A Multi-Country Population-Based Surveillance Study. *The Lancet Infectious Diseases*. 18(6):675–683. DOI: 10.1016/S1473-3099(18)30073-2.

Zumbo, A., Palucci, I., Cascioferro, A., Sali, M., Ventura, M., D'Alfonso, P., Iantomasi, R., Di Sante, G., *et al*. 2013. Functional Dissection of Protein Domains Involved in the Immunomodulatory Properties of PE_PGRS33 of *Mycobacterium tuberculosis*. *Pathogens and Disease*. 69(3):232–239. DOI: 10.1111/2049-632X.12096.