Biochemical characterisation of pivotal enzymes involved in *Mycobacterium tuberculosis* cell wall biosynthesis

By

James Harrison

A thesis submitted to the

University of Birmingham

for the degree of

DOCTOR OF PHILOSOPHY

Institute of Microbiology and Infection

School of Biosciences

College of Life and Environmental Sciences

University of Birmingham

May 2016

ABSTRACT

Mycobacterium tuberculosis, the etiological agent of tuberculosis, has a unique cell envelope which accounts for its unusual low permeability and contributes to resistance against common antibiotics. The mycobacterial cell wall consists of a cross-linked network of peptidoglycan (PG) in which some of the muramic acid residues are adorned with a complex polysaccharide, arabinogalactan (AG), *via* a unique α-L-rhamnopyranose–(1→3)-α-D-GlcNAc-(1→P) linker unit. Whilst the cytoplasmic steps of mycobacterial cell wall biosynthesis have been largely delineated, the molecular processes that govern the flux of PG intermediates and the mechanism by which PG and AG pathways converge has remained elusive. We identified key conserved serine/threonine residues of MurC, as potential candidates for phospho-regulation by the cognate protein kinase, PknA. Pseudo-phosphorylated MurC mutants exhibited differential enzyme activity, suggesting that *M. tuberculosis* is capable of tight control of PG biosynthesis through phosphorylation of MurC. In addition, we have identified Lcp1, a mycobacterial orthologue of the LytR-CpsA-Psr (LCP) family of proteins found in Gram-positive bacteria, responsible for ligating cell wall teichoic acids to PG. We demonstrate that *lcp1* is an essential gene required for cell viability and show that recombinant Lcp1 is a phosphotransferase capable of ligating AG to PG in a cell free radiolabelled assay.

DECLARATION

The work presented in this thesis was carried out in the School of Biosciences at the University of Birmingham, U.K., B15 2TT during the period June 2012 to October 2015.

The work in this thesis is original, except where acknowledged by references.

No part of the work is being, or has been, submitted for a degree, diploma or any other qualification at any other University.

ADDITIONAL DECLARATION

Some parts of chapter 1 of this thesis, pertaining to the structure and biosynthesis of peptidoglycan (Section 1.2.2.1. and Section 1.2.2.2.), bear some similarity to portions of the publication "Assembly of the Mycobacterial Cell Wall, Jankute *et al*., 2015, Annual Review of Microbiology, Vol: 69, 405-423", in particular to the sections relating to peptidoglycan structure and biosynthesis. However, these similar sections included both within this thesis and within the publication, are my original work, with no significant contributions to the text by the co-authors. The authors listed below are in acknowledgement of the above.

Dr. Monika Jankute – Lead author

Prof. Gurdyal S. Besra – Corresponding author

James Harrison – Student

ACKNOWLEDGEMENTS

I would like to take this opportunity to thank the numerous people who have made the work involved and production of this thesis possible. Firstly, I would like to thank both of my supervisors, Prof. Gurdyal S. Besra and Dr. Luke J. Alderwick, for providing me the opportunity to undertake this research in such a brilliant laboratory. In particular, I would like to thank Luke, whom not only served as my undergraduate tutor and inspired me to pursue postgraduate study, but also continually provided advice, assistance and enlightenment with eternal patience.

I would also like to thank all members of the Besra lab, both past and present, for all the support: Amrita, Shipra, Vijay, Veemal, Pete, Usha, Cristian, Monika, Sarah, Natacha, Sid, Albel, Ramón, Szilvi, Bogdan, Giacomo, Padraic, Asma, Panchali, Pat, Jon and Kat.

A special thanks goes to Dr Georgina (George) Lloyd, whom not only taught me how to perform a lot of laboratory techniques, but continually provided assistance and looked out for me. She has also provided friendship, despite always having to be my chaperone when travelling, especially upon my visits to the University of Warwick.

I would also like to thank Prof. Todd. L. Lowary and Dr. Maju Joe of the University of Alberta for performing the chemical synthesis of mycobacterial linker unit analogues, compounds 1-4. Also, I would like to thank the Birmingham Biophysical Characterisation Facility for performing biophysical characterisation experiments. Further thanks go to the Dowson/Roper lab at the University of Warwick, for providing both substrates and access to equipment. Particular thanks to Dr. Adrian Lloyd for all the help and advice whilst there.

Finally, I would like to thank my friends and family for the extensive support in making this thesis possible.

This thesis is dedicated to my family

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

Å	Ångström
ABC	ATP binding cassette
acet	acetamide
Acp	Acyl carrier protein
AcpM	Mycobacterial acyl carrier protein
aDNA	ancient DNA
ADP	adenosine diphosphate
AFM	atomic-force microscopy
AG	Arabinogalactan
AGP	Arabinogalactan-Peptidoglycan
Ala/A	Alanine
AMP	adenosine monophosphate
AMPPNP	adenylyl-imidodiphosphate
Ara*f*	arabinofuranose
Arg/R	Arginine
Asp/D	Aspartate
ATP	adenosine triphosphate
AUC	analytical ultracentrifugation

B_{max}	maximum binding capacity
bp	base pairs
BSL	Biosafety Level
BTZ	Benzothiazinones
C	Carbon
c (s)	distribution of sedimentation coefficient
CD	Circular Dichroism
CESTET	Conditional Expression-Specialised Transduction Essentially Test
CFU	Colony Forming Units
Cg	*Corynebacterium glutamicum*
CPM	Counts per minute
CWP	Cell Wall Polymers
CWTA	Cell Wall Teichoic Acids
Cys/C	Cysteine
D-ala	D-alanine
DAT	diacyl trehalose
dcw	division/cell wall
dH_2O	deionised water
D-*iso*glu	D-*iso*glutamate

DNA	deoxyribonucleic acid
dNTP	deoxynucleotide triphosphate
DOTS	Directly Observed Therapy Shortcourse
DPA	β-D-arbinofuranosyl-1-monophosphoryldecaprenol
DPPR	decaprenylphosphoryl-5-β-D-phosphoribose
DPR	decaprenylphosphoryl-5-β-D-ribose
DPX	decaprenylphosphoryl-2-keto-β-D-*erythro*-pentofuranose
dTDP	thymidine diphosphate
DTT	dithiothreitol
dTTP	thymidine triphosphate
ECT	electron cryo-tomography
ES-MS	Electrospray Mass Spectrometry
FAS	Fatty acid synthase
Fbp	fibronectin-binding proteins
FDA	Food and Drug Administration (US)
g	gram
Gal	galactose
Gal*f*	galactofuranose
GalN	galactosamine

GalNAc	*N*-acetylgalactosamine
Gal*p*	galactopyranose
GFP	Green fluorescent protein
GG-P	Geranylgeranyl-monophosphate
GG-P-P	Geranylgeranyl-pyrophosphate
GL-1	Glycolipid 1 (GlcNAc-P-P-C_{50})
GL-2	Glycolipid 2 (Rha-GlcNAc-P-P-C_{50})
GL-3	Glycolipid 3 (Gal*f*-Rha-GlcNAc-P-P-C_{50})
GL-4	Glycolipid 4 (Gal*f*-Gal*f*-Rha-GlcNAc-P-P-C_{50})
GlcNAc	*N*-acetylglucosamine
GlfT	galactofuranosyl transferase
Gly/G	Glycine
GOI	gene of interest
G-P	Geranyl-monophophate
G-P-P	Geranyl-pyrophosphate
GTP	guanosine triphosphate
h	Hill coefficient/slope
H	Hydrogen
HBC	High Burden Countries

His/H	Histidine
HIV	Human Immunodeficiency Virus
HPO	hypertrophic pulmonary osteopathy
Hyg	Hygromycin
IC_{50}	half maximal inhibitory concentration
IMAC	immobilised metal ion affinity chromatography
ITF	Intrinsic Tryptophan Fluorescence
Kan	Kanamycin
kb	kilobase pairs
K_{cat}	turnover number
K_{cat}/K_m	specificity constant
K_d	dissociation constant
kDa	kilodaltons
K_i	inhibitory constant
K_m	Michaelis constant
L	litre
L-ala	L-alanine
LAM	Lipoarabinomannan
LB	Luria-Bertani

LCP	LytR-CpsA-Psr
Lcp1	Peptidoglycan-arabinogalactan Ligase
Leu/L	Leucine
LM	Lipomannan
L-ser	L-serine
LU	Linker Unit
M	Molar
m/z	mass/charge
M^{-1}	per molar
mA	milliamp
MA	Mycolic Acids
mAGP	mycolyl-arabinogalactan-peptidoglycan
ManNAc	*N*-acetylmannosamine
mCi/mmol	millicurie per millimole
m-DAP	*meso*-diaminopimelate
MDR	multi-drug resistant
Met/M	Methionine
Mg	Magnesium
mg	milligram

MIC	Minimum inhibitory concentration
MIM	mycobacterial inner membrane
min	minute
min^{-1}	per minute
mL	millilitre
mM	millimolar
MOM	mycobacterial outer membrane
Ms	*Mycobacterium smegmatis*
Mtb	*Mycobacterium tuberculosis*
Mur	Muramyl
MurNAc	*N*-acetylmuramic acid
MurNGlyc	*N*-glycolylmuramic acid
N	Nitrogen
NAD^+	oxidised nicotinamide adenine dinucleotide
NADH	reduced nicotinamide adenine dinucleotide
NADPH	reduced nicotinamide adenine dinucleotide phosphate
ng	nanogram
nm	nanometre
nM	nanomolar

NMR	Nuclear Magnetic Resonance
NOD	nuclotide-binding oligomerisation domain
O	Oxygen
P	Phosphate
P60	60 % Percoll cell wall fraction
PAGE	polyacrylamide gel electrophoresis
PASTA	PBP and Serine/Threonine associated
PAT	pentaacyl trehalose
PBP	penicillin-binding protein
PCR	polymerase chain reaction
PDIM	phthiocerol dimycocerosate
PEP	phosphoenolpyruvate
PG	Peptidoglycan
pH	"Power of Hydrogen"
phΔ^{Ms}*lcp1*	^{Ms}*lcp1* knock-out phage
Pi	inorganic phosphate
PKLDH	Pyruvate kinase/Lactate dehydrogenase
PPG	purified peptidoglycan
PPi	inorganic pyrophosphate

pRpp	5-phosphoribosyl-1-pyrophosphate
R_f	retention factor
Rha	rhamnose
Rha*p*	rhamnopyranose
RNA	ribonucleic acid
RNAP	ribonucleic acid polymerase
ROS	reactive oxygen species
S	Sulfur
S	Sedimentation Coefficient
SAM-MT	*S*-adenosyl-methionine-dependant methyltransferases
SDM	site-directed mutant/mutagenesis
SDS	Sodium dodecyl sulfate
Ser/S	Serine
SGL	sulfoglycolipid
STPK	Serine/Threonine protein kinase
TB	Tuberculosis
TBS	Tris-buffered saline
TDM	Trehalose dimycolate
TDR	Totally drug resistant

Thr/T	Threonine
TLC	Thin Layer Chromatography
TMHMM	Transmembrane hidden Markov model
TMM	Trehalose monomycolate
TSA	tryptic soy agar
TSB	tryptic soy broth
U	Unit
UDP	uridine diphosphate
V	rate
V	Volt
v/v	volume/volume
Val/V	Valine
V_{max}	maximum rate
w/v	weight/volume
WHO	World Health Organisation
WT	wild-type
XDR	extensively drug resistant
Δ	Delta = change
ΔF	change in fluorescence

Δ^{Ms}*lcp1*	^{Ms}*lcp1* knock-out mutant
%	Percentage
[^{14}C]	Radiolabelled Carbon-14
× g	times gravity
°	degrees
°C	degrees centigrade
µg/mL	microgram per millilitre
µL	microliter
µm	micrometre
µM	micromolar

PUBLISHED WORK ASSOCIATED WITH THIS THESIS

Alderwick, L.J., **Harrison, J**., Lloyd, G.S., Birch, H.L. (2015). The Mycobacterial Cell Wall – Peptidoglycan and Arabinogalactan. Cold Spring Harb Perspect Med doi: 10.1101/cshperpect.a021113

Jankute, M., Cox, J.A.G., **Harrison, J**., Besra, G.S. (2015). Assembly of the mycobacterial cell wall. Annual Review of Microbiology 69:405-423

Chapter 1

Introduction

1. INTRODUCTION

1.1. Tuberculosis

1.1.1. Etiology of Tuberculosis

Tuberculosis (TB) is a severe pulmonary disease that is prevalent globally, caused by the human pathogen *Mycobacterium tuberculosis* (Cambau and Drancourt, 2014). Primarily affecting the respiratory system, symptoms of active TB disease consist of persistent coughing, wasting, exhaustion and finger clubbing (through poor oxygenation). Later signs of the disease include coughing with sputum and/or blood and associated chest pain, shortness of breath, possible anaemia and potential leucocytosis (due to the immune response causing large numbers of leukocytes to migrate to the site of infection) (Knechel, 2009). Extrapulmonary TB has also become a problem in recent times, as the increase in incidence of Human Immunodeficiency Virus (HIV) allows the disease to take hold in organs, other than the lungs, much easier by spreading throughout the lymphatic and circulatory systems after an initial pulmonary infection. This kind of TB infection can be difficult to diagnose as symptoms are often non-specific, yet will most likely become progressively worse without treatment (Elder, 1992). Of all new cases reported in 2014, almost 900,000 were extrapulmonary disease cases (WHO, 2015).

Transmissible through airborne droplet nuclei, *M. tuberculosis* can infect the lungs once inhaled, and initiate innate immune response (Knechel, 2009). This causes macrophages to engulf the mycobacterial cells within phagosomes and try to destroy them through fusion with lysosomes, to form phagolysosomes (Smith, 2003). Inside the phagolysosomes, *M. tuberculosis* cells are bombarded with reactive oxygen species (ROS), lysosomal enzymes and even acidic pH, however, protection provided by the unique ‘waxy’ cell wall of *M. tuberculosis* allows it to survive, and even continue replicating slowly inside the macrophage

(Smith, 2003, Knechel, 2009). Released cytokines and antigen presentation by the macrophages begin to recruit T lymphocytes to the site of infection, which then begin to attempt to contain the spread of *M. tuberculosis* cells. The recruitment of monocytes, neutrophils as well as the lymphocytes causes the formation of a granuloma, with the infected macrophages in the centre (Smith, 2003, Sakamoto, 2012). Eventually, due to the hypoxic conditions of the granuloma, the central macrophages die and form necrotic lesions, in which the *M. tuberculosis* cells are able to adapt and survive (e.g. by regulating protein expression) (Knechel, 2009). Over time, the necrotic regions become caseous, characterised by low oxygen, low pH, limited nutrients as well as toxic fatty acids, where *M. tuberculosis* cells are unable to replicate (Smith, 2003, Knechel, 2009). These non-replicating *M. tuberculosis* cells can then remain in a dormant state for years, being kept in check by the host immune system by fibrosis and calcification of the granuloma (Knechel, 2009). However, if the infected host becomes immunocompromised or possess a less than adequate immune system to begin with, then the stability of the granuloma is compromised. The fibrous walls of the lesion begin to degrade, and the calcified tissue becomes liquefied. This allows the semi-liquid necrotic tissue to drain into blood vessels, and lesions fill with air. Because of this, the *M. tuberculosis* cells become free, and allow infected hosts to begin to spread the disease to others through droplet nuclei, as well as allowing the cells to spread to other areas of the lung and body, where the organisms can start new granulomas (Knechel, 2009).

TB is an extremely old disease, being prevalent for thousands of years. One of the oldest known confirmed cases of TB infection in humans is over 7000 years ago, in ancient Hungary. The skeletal remains of a young male adult were examined and found to display signs of hypertrophic pulmonary osteopathy (HPO) (Masson *et al.*, 2013). A form of hypertrophic osteoarthropathy, HPO is characterised by deposit of newer bone upon the diaphyses of long bones, which leaves identifiable differences between the two layers. A

common cause of HPO in modern times is intrathoracic cancer, but it is also an indicator of *M. tuberculosis* infection (Armstrong *et al*., 2007, Masson *et al*., 2013). The remains showed bone changes within the ribs and vertebrae also. Because of these symptoms, ancient DNA (aDNA) was extracted from various bones *via* polymerase chain reaction (PCR) amplification of specific *M. tuberculosis* complex regions, and then sequenced. Lipid biomarkers were also extracted and analysed from bone samples. Both aDNA and lipid biomarkers confirmed the presence of *M. tuberculosis*, and thus supports the hypothesis that the young male adult did have a TB infection (Masson *et al*., 2013). Similar techniques were used to prove the existence of *M. tuberculosis* within the skeletal remains of a woman and infant who were buried over 9000 years ago in an area in the Neolithic eastern Mediterranean (now submerged) (Hershkovitz *et al*., 2008). This suggests that the bacterium has evolved alongside humans for a very long time, and helps to explain why it is such an efficient human pathogen.

1.1.2. Taxonomy of *Mycobacterium tuberculosis*

The causative agent of TB in humans, *M. tuberculosis*, is an acid-fast staining bacterium that is part of the Actinomycetales order of the Actinobacteria phylum and class. It is a member of the Corynebacterineae suborder and the *Mycobacteriaceae* family. *M. tuberculosis* is therefore related to other members of the Corynebacterineae suborder (e.g. *Corynebacterium glutamicum*) and other mycobacterial species (e.g. *Mycobacterium smegmatis* and *Mycobacterium leprae*), and shares some genetic identity with these species. However, *M. tuberculosis* is a member of a group of closely related species, called the *Mycobacterium tuberculosis* complex. Members of this group are mycobacteria that cause TB or TB like symptoms in a wide range of hosts and therefore members have similar genetic identity to each other (as high as 99 % identity). Some examples of members of this complex are *Mycobacterium africanum*, *Mycobacterium canetti* (which both cause human TB)

Mycobacterium bovis (bovine TB) and *Mycobacterium microti* (vole TB) (Bouakaze *et al*., 2010). Furthermore, mutations can lead to changes in phenotype, giving rise to different strains within each species. *M. tuberculosis* H37Rv and CDC1551 differ in their pathogenic capability, since H37Rv can create larger, and a greater number of tubercles than CDC1551, which has an impact on the virulence of the latter strain (Manabe *et al*., 2003).

Because of the close genetic similarities between strains and species, it is possible to use different organisms as model systems to study *M. tuberculosis*. This is important, as it allows for mycobacterial study without the requirement of Biosafety Level (BSL) 3 laboratories and training, since non-pathogenic substitutes (such as *M. bovis* BCG, an attenuated strain) can be used in lower BSL laboratories (e.g. BSL 2) (Mahairas *et al*., 1996). It also means that research can proceed much faster with the use of quicker growing model organisms (such as *M. smegmatis*), since *M. tuberculosis* is extremely slow growing in comparison (*M. smegmatis* doubling time 2-5 h, *M. tuberculosis* doubling time 18-54 h) (Gill *et al*., 2009).

1.1.3. Epidemiology of TB disease

TB is prevalent worldwide, with both Africa and Asia having the highest incidence (approximately 86 % of all cases combined). Globally, there were more than 9.6 million estimated new cases of TB during 2014 (Figure 1.1) and 1.5 million deaths (1.1 million from HIV negative cases (Figure 1.2) and 390,000 deaths from HIV positive associated cases). Of these, 480,000 cases were estimated to be multi-drug resistant (MDR) TB, which resulted in 190,000 deaths (WHO, 2015). South East Asia and Western Pacific countries reported 58 % of all TB cases in 2014. Of these, India, Indonesia and China account for the highest incidence, 43 % of TB cases in 2014 (23 % (around 2.2 million), 10 % (around 960,000) and 10 % (around 960,000) respectively) were reported in these countries alone (WHO, 2015). African countries accounted for 28 %, with the remaining cases prevalent in the Eastern

Mediterranean (8 %), the Americas (3 %) and European regions (3 %) (Figure 1.1) (WHO, 2015). However, TB is a severe problem in the so called 22 'High Burden Countries' (HBC) (e.g. Russia, African countries such as Ethiopia and Kenya, and Asian countries such as Bangladesh), as these account for 83 % of all cases globally. These countries mostly have incidence rates of 150-300 per 100,000 of the population; however Mozambique and Zimbabwe have incidences of over 500 per 100,000. South Africa and Swaziland have a very severe incidence as it is estimated that at least 1 in 100 people will contract TB each year (1000+ per 100,000 of the population) (Figure 1.1) (WHO, 2015). However, after 2015 the list of HBCs will be increased, and split into 3 new lists for TB, MDR-TB and TB/HIV. These will encompass 30 countries in each list, which are defined as incidence of more than 10,000 cases annually for TB, or 1,000 cases for MDR-TB and TB/HIV (WHO, 2015).

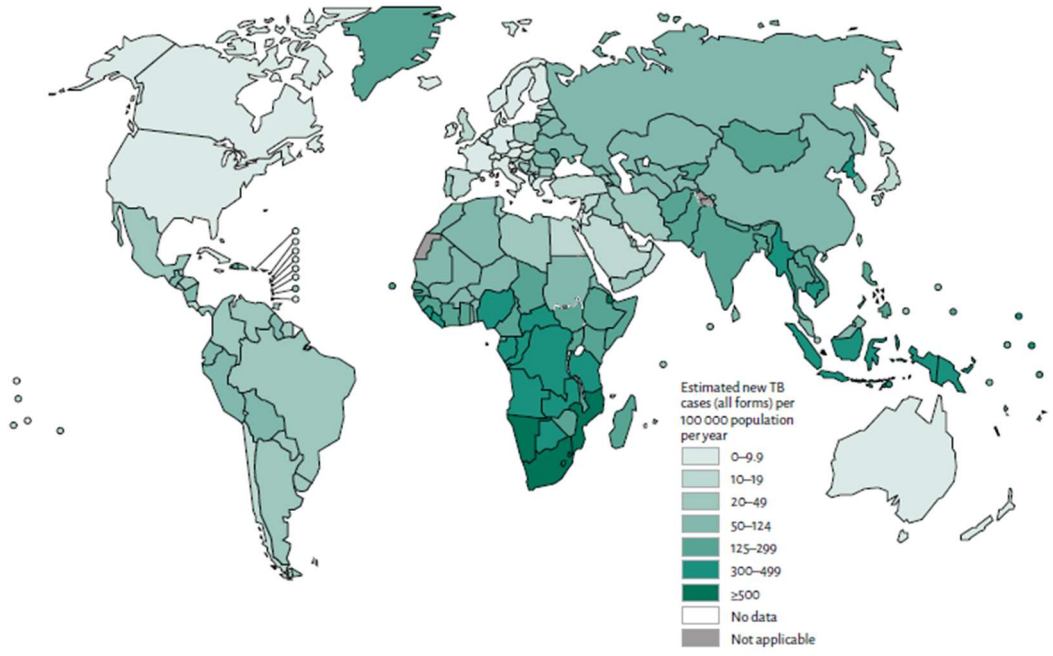

Figure 1.1: Map of global incidence of new TB cases in 2014. The highest incidence is found in sub-Saharan Africa (in particular South Africa), as well as countries in South East Asia. The lowest incidence of new cases are in Western Europe, North America and Australia (WHO, 2015).

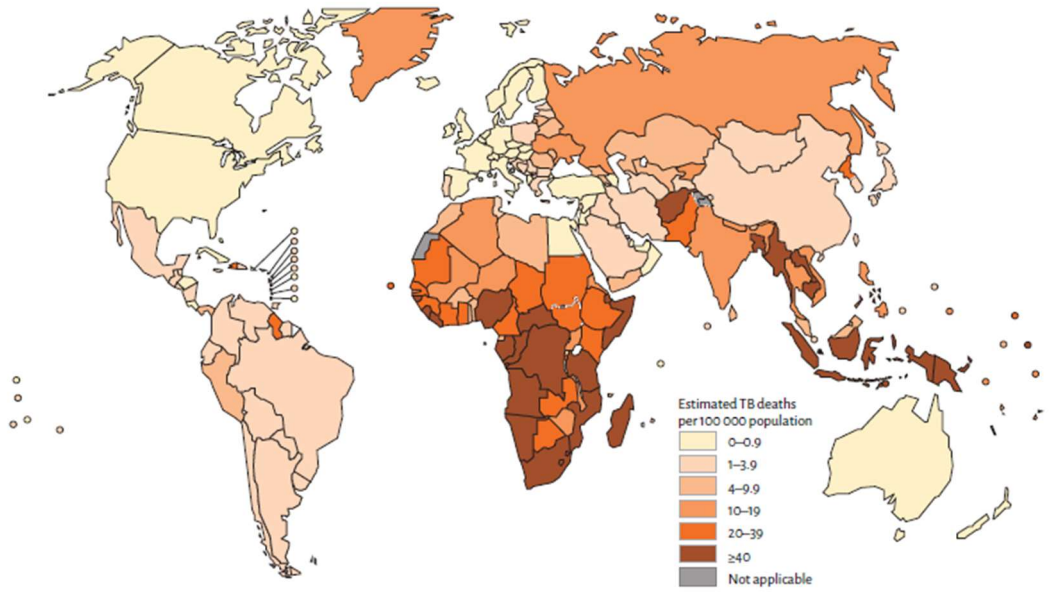

Figure 1.2: Map of deaths from HIV negative cases of TB worldwide in 2014. The highest incidence of death correlates with incidence of cases (figure 1.1), whereby the most deaths are in sub-Saharan Africa, as well as South East Asia, in particularly Indonesia. The lowest death rates are seen in Western Europe, the Americas and Australia (WHO, 2015).

The highest incidence of TB including both cases and deaths occur in men, but the disease is also very prevalent in women, as the number of new cases in women during 2014 was 3.2 million (33.3% of the total), whereas 1 million cases were in children aged 15 or below (10.4 % of the total) (WHO, 2015). TB is one of the top 3 causes of death in women in the world (along with other lower respiratory infections, the other 2 top causes being ischemic heart disease and stroke); with the disease killing around 480,000 (140,000 HIV positive deaths). Therefore, nearly half of all deaths by TB in 2014 were women (almost half of all HIV positive deaths) (WHO, 2015). However, 140,000 deaths in 2014 were children (55,000 HIV positive deaths), which emphasises that TB is a disease that can affect all (WHO, 2015).

1.1.4. Treatment of TB

The disease is treatable with a selection of anti-TB agents, but if ignored it can eventually cause death (Knechel, 2009). Without an effective treatment program, in cases of sputum smear-positive (cases where TB has been confirmed by microscopy of sputum samples) HIV

negative TB, it is likely that 70% of patients will die within 10 years of contracting the disease (Tiemersma *et al*., 2011). However, the treatment of TB is becoming increasingly difficult with the rise of MDR, extensively drug resistant (XDR) and now reported totally drug resistant (TDR) strains of the disease becoming globally more common (Udwadia *et al*., 2012, Velayati *et al.*, 2009). Drug resistance is believed to be largely due to mismanagement of TB cases and is defined as *in vivo* resistance to isoniazid and rifampicin (MDR) as well as any fluoroquinolone and at least one injectable drug (e.g. capreomycin, kanamycin, etc) (XDR) (Sotgiu *et al*., 2011).

Current front line drug treatments against TB are around 50 years old, but do not have a 100% success rate of curing HIV negative TB (WHO, 2012). The current treatment for TB is termed Directly Observed Therapy Shortcourse (DOTS), which was introduced by the World Health Organisation (WHO) in 1995 as a way to monitor treatment of patients and maintain adherence to the drug plan (Davies, 2003). DOTS is defined by WHO as a five element plan, consisting of: Political commitment and funding, Case confirmation through microbiological services, Standardised observed treatment, Consistent drug supply and Evaluation of the treatment (WHO, 2015). However, poor understanding of DOTS, and the underfunding by some countries leads to poor adherence to the program (Davies, 2003). The regime of drugs is also rigorous, requiring a 6 month course, of rifampicin, isoniazid, ethambutol and pyranzidamine for the first 2 months, followed by 4 months of just rifampicin and isoniazid. However, in cases of extrapulmonary TB treatment will often be extended to 12 months, and any case of drug resistance TB will require 20 months of drug treatment (WHO, 2012). Drug resistant TB requires a great deal of time to treat, down to drug susceptibility screening that must be performed before treatment can commence. However, delays in this testing can exacerbate the disease, not only for the individual, but it can allow for further spreading throughout the population, especially in countries designated as high burden (Rifat *et al*.,

2015). However, recent research suggests the use of β-lactam antibiotics can be used in conjunction with β-lactamase inhibitors to restore antibiotic activity in both MDR and XDR TB strains to the same level as susceptible strains of the disease (Hugonnet *et al.*, 2009). This is significant as many of these drugs already have FDA approval, so could be put into use rapidly. Nevertheless, full compliance with DOTS and the drug treatment regime still only provides a survival rate of 85 % (some cases less than 70 %), with MDR-TB survival rates being even lower, indicating the urgent requirement for new front line anti-TB agents (Davies, 2003, WHO, 2015).

1.1.4.1. Rifampicin

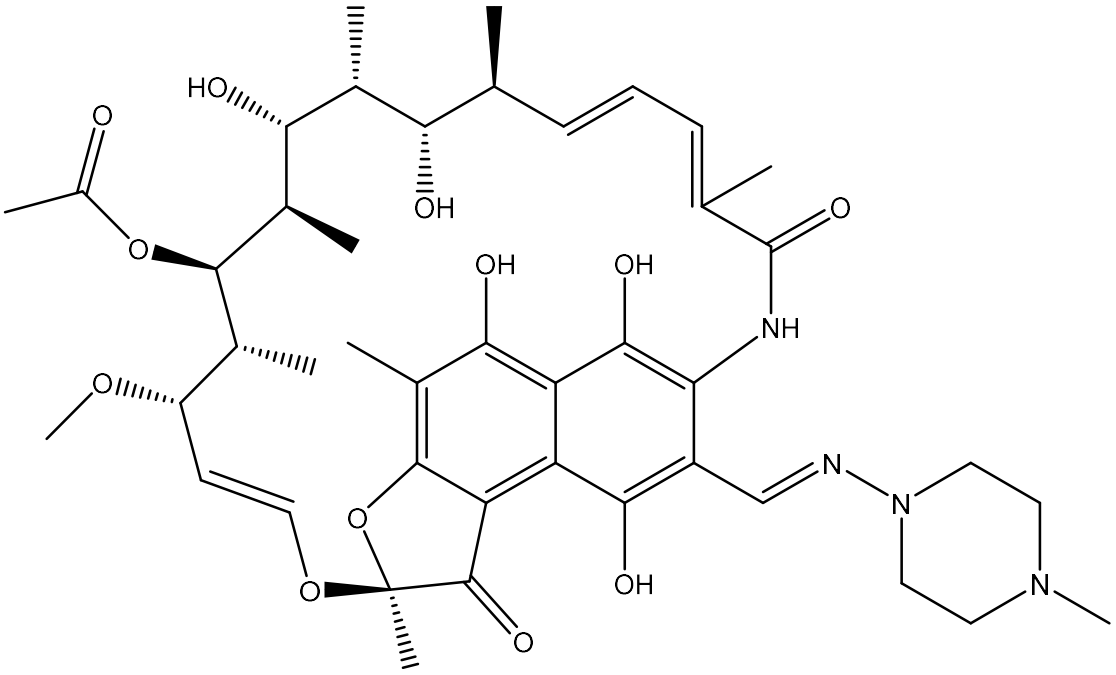

Figure 1.3: Structure of Rifampicin.

Each of the four front line drugs inhibit the growth of mycobacteria with distinct modes of action. The mechanism of action of rifampicin is the inhibition of RNA polymerase (RNAP) to block RNA synthesis (Figure 1.3) (McClure and Cech, 1978). This is achieved by the

binding of rifampicin to the β-subunit of RNAP, which causes the synthesis of RNA to arrest after the first phosphodiester bond is formed, due to steric hindrance in a process termed abortive initiation (Wehrli, 1983). Rifampicin additionally causes a 2-3 fold reduction in binding affinity of RNAP for nucleotide triphosphates, such as ATP, further reducing the ability of the cell to produce RNA (McClure and Cech, 1978). Resistance to rifampicin arises through mutations in the β-subunit gene *rpoB*, specifically within a 81 base pair region between codons 504-533 (Somoskovi *et al*., 2001). The most commonly found mutations occur at positions 516, 526 and 533 and are responsible for around 90 % of rifampicin resistant strains (Pang *et al*., 2013). The most frequently seen mutations are Ser531Leu and His526Asp, both of which provide high level resistance to rifampicin compared to other mutations, increasing the average MIC from 0.2 μg/mL, to >32 μg/mL (Somoskovi *et al*., 2001, van Ingen *et al*., 2011). However, mutations present in other parts of the 81 base pair region do not provide such high resistance as these 2 common mutations (however they can still provide a 5 fold increase in MIC) (Pang *et al*., 2013). An additional source of rifampicin resistance comes from the use of active efflux pumps by the cell, which can be attributed to around 5 % of resistance seen. *M. tuberculosis* that is rifampicin-resistant is believed to upregulate efflux pumps in order to provide additional low level resistance to the drug (Pang *et al*., 2013).

1.1.4.2. Isoniazid

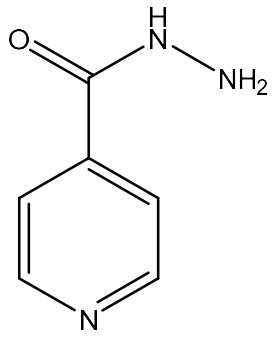

Figure 1.4: Structure of Isoniazid.

Isoniazid is a pro-drug that causes inhibition of mycolic acid biosynthesis, an essential component of the mycobacterial cell wall (Figure 1.4) (Section 1.2.4.). Isoniazid becomes activated by the catalase-peroxidase KatG, into numerous radicals, one of which (an isonicotinyl-acyl radical) binds to NAD^+/NADH to form an isoniazid-NADH adduct (Timmins and Deretic, 2006). This adduct then binds to the NADH-dependant enoyl-acyl carrier protein reductase, InhA (a part of the fatty acid synthase (FAS) II system), which causes a disruption to mycolic acid synthesis as fatty acids cannot become elongated (Marrakchi et al., 2000, Kremer *et al*., 2003). Since isoniazid is a pro-drug, there are numerous ways that the bacterium can acquire resistance. The most commonly found feature within resistant strains is a mutation in *katG*, Ser315Thr, found within 40 % of resistant isolates (Somoskovi *et al*., 2001). This mutation decreases the catalase-peroxidase ability of the enzyme to around 30-40 % activity, which is enough to continue function, but does not activate isoniazid (Marttila *et al*., 1998). An alternative method of resistance to isoniazid employed by the cell is mutations within the *inhA* gene. A biochemical study provides evidence that the Ser91Ala mutation in InhA causes a decrease in the binding affinity between the enzyme and the isoniazid-NADH adduct, due to a disruption of the hydrogen bonding network, showing a 30-fold increase in K_i value (172 nM compared to 5 nM for wild-type).

This decrease in binding affinity results in an increase in resistance of *Mycobacterium* to isoniazid (Vilcheze *et al*., 2006).

1.1.4.3. Ethambutol

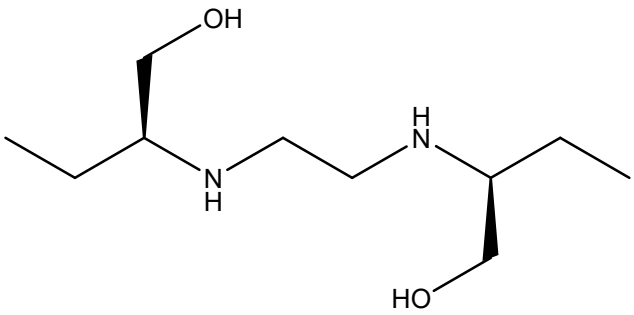

Figure 1.5: Structure of Ethambutol.

Ethambutol has a unique mode of action, whereby it inhibits proper biosynthesis of arabinogalactan, which is a crucial connective cell wall macromolecule (Figure 1.5) (Section 1.2.3.). The use of ethambutol on replicating mycobacteria showed that it caused an immediate accumulation of a major intermediate in arabinan biosynthesis, β-D-arbinofuranosyl-1-monophosphoryldecaprenol (DPA), as well as a build-up of trehalose monomycolates (TMM) and trehalose dimycolates (TDM), with no mycolylation of the cell wall occurring (Mikusova *et al*., 1995). This suggested that the action of ethambutol is to prevent the polymerisation of mycobacterial arabinan (which provides the site for mycolate attachment), and thus led to the identification of a set of arabinosyltransferases, known as EmbA and EmbB, as the target of ethambutol in *M. tuberculosis* (Belanger *et al*., 1996). Although very little is known about the site of action of ethambutol, EmbA and EmbB are known to work as a heterodimer to elongate the arabinan chain and it has been shown that overexpression of the *embCAB* locus of genes conveys resistance to the drug and restores arabinan biosynthesis (Belanger et al., 1996, Amin *et al*., 2008). A mutation in EmbB is commonly found at Met306, in around 66.7 % of resistant isolates (Ramaswamy *et al*., 2004).

Changes in this residue can give an MIC value generally twice as high as wild-type, since Met306Leu/Val gives an MIC of 40 µg/mL, whereas Met306Ile gives an MIC of 20 µg/mL. Cells with wild-type EmbB have an MIC of 5-15 µg/mL, which means these mutations lead to an increase in resistance (Sreevatsan *et al*., 1997). However, it has been shown that mutations at EmbB306 have been found in ethambutol sensitive strains, which has led to the suggestion that mutations at EmbB306 lead to a predisposition of the cell to acquire further mutations in different genes that lead to high-level resistance to ethambutol (Srivastava *et al*., 2009). In particular, mutations in *ubiA* have been shown to increase production of the intermediate DPA (Section 1.2.3.2), whereas mutations in *aftA* can lead to overexpression of *embC* (since *aftA* is immediately upstream of *embC*), both of which can interact with mutations in *embB* and lead to an increase in MIC value above that of wild-type (Safi *et al*., 2013).

1.1.4.4. Pyrazinamide

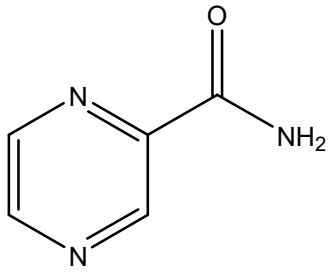

Figure 1.6: Structure of Pyrazinamide.

The mechanism of action of pyrazinamide is yet to be elucidated, however there are currently a number of theories behind how it takes action. Pyrazinamide (Figure 1.6) is known to be a pro-drug that is activated by pyrazinamidase into pyrazinoic acid (Konno *et al*., 1967). This accumulation of pyrazinoic acid is thought to cause disruption to the plasma membrane due to the acidic pH and disrupt membrane transport, although the significance of this remains unsolved (Zhang *et al*., 2003). More recent research has suggested a downstream target of pyrazinoic acid as ribosomal protein S1, RpsA, which has a role in *trans*-translation (the

rescue of ribosomes that have stalled mid-translation) (Shi *et al*., 2011). Resistance to pyrazinamide is usually associated with mutations within *pncA*, the gene encoding the pyrazinamidase (72-97 % of isolates) (Somoskovi *et al*., 2001). This is further evidenced due to the natural resistance of *M. bovis* and *M. bovis* BCG to pyrazinamide, since these bacteria carry a mutation within *pncA* (Cys169Gly) that prevents the drug from becoming activated (Scorpio and Zhang, 1996). However, *pncA* in *M. tuberculosis* has been found with mutations throughout the sequence that prevent activation of pyrazinamide (Scorpio and Zhang, 1996). Despite these mutations being highly prevalent, mutations have been identified within *rpsA* in isolates that do not contain any *pncA* mutations, yet still display pyrazinamide resistance (2-3 fold increase in MIC value) (Shi *et al*., 2011). This provides further evidence of RpsA as a target of the drug, but also shows a risk for an increase in resistance routes to pyrazinamide.

Each front line drug plays a crucial part in the combinational therapy against TB. Rifampicin is the most important drug within the regime, as it is bactericidal against both metabolically active and stationary *M. tuberculosis* cells, which therefore targets the majority of the infection (Arbex *et al*., 2010). As a cell wall targeting prodrug, isoniazid is effective in killing rapidly growing bacteria, yet slower growing cells are relatively unaffected (Zhang, 2005). Likewise, ethambutol has a bacteriostatic effect on rapidly growing *M. tuberculosis* (both intracellular and extracellular cells), but does not have any real effect on slower growing cells (Zhang, 2005). Finally, pyrazinamide is bactericidal against stationary *M. tuberculosis* and is used to sterilise the persistent bacterial populations within macrophages and prevents relapse of the disease from these cells (Arbex *et al*., 2010). The inclusion of pyrazinamide within the drug regime reduces the length of time of rifampicin treatment to 6 months overall (Arbex *et al*., 2010). The inclusion of isoniazid with rifampicin during the final 4 months of the regime, despite isoniazid only being active against rapidly growing bacteria, is because of the different stages of granulomas found within the lungs (Sakamoto, 2012). Some granulomas

will degrade sooner than others, meaning *M. tuberculosis* cells will become metabolically active at different times during the drug regime, therefore isoniazid must be present in order to combat these active persistent bacteria throughout treatment (Zhang, 2005, Arbex *et al*., 2010).

Such a complex combinational therapy is required against TB in order to avoid the rise of drug resistance and still be able to combat the disease even if it does arise (Cottarel and Wierzbowski, 2007). Each front line drug has a distinct target and mode of action, which prevents cross-resistance from becoming an issue. If a drug is used on its own, then *M. tuberculosis* would likely become resistant to it very quickly, whereas the combination of drugs reduces the likelihood of this occurring (Bakker-Woudenberg *et al*., 2005). However, new anti-tuberculosis drugs are urgently required, not just to try and tackle persistent cells in order to reduce the length of therapy, but also drugs that target some aspect of mycobacterial cell wall biosynthesis, in order to replace isoniazid or ethambutol in cases where resistance is found. In order to develop such antimicrobials, it is therefore incredibly important for us to understand the molecular architecture of such a complex structure as the mycobacterial cell wall, as well its biosynthesis, so that it may be exploited in the future.

1.2. Mycobacterial cell wall

1.2.1. Molecular architecture of the cell wall

The cell wall is an essential component to almost all prokaryotes, as it serves to protect the cell against a hostile external environment, as well as excessive osmotic turgor (Silhavy *et al*., 2010). Bacterial cell walls have historically been classified into one of two groups, depending upon if they retain Crystal violet during Gram staining, those that do are referred to as Gram-positive, and those that do not are known as Gram-negative (Bartholomew and Mittwer, 1952). However, in recent times the use of molecular genetics and cryo-electron microscopy are becoming increasingly more popular as methods of identification and classification of bacteria (Doolittle, 2005, Sani *et al*., 2010). The bacteria that are classified as Gram-negative (such as *Escherichia coli*) possess a cell wall that is comprised of a thin peptidoglycan (PG) layer, found within the periplasmic space, with a phospholipid/glycolipid (primarily lipopolysaccharide) outer membrane surrounding the cell (Silhavy *et al*., 2010). Both the periplasmic space and outer membrane are highly protein rich, many of these proteins are involved in nutrient transport both into and out of the cell (Rassam *et al*., 2015). The major difference between Gram-negative and Gram-positive cell walls is the lack of outer membrane in Gram-positive bacteria (such as *Bacillus subtilis*). Because of this, the cell is less protected from the environment and compensates by having a much thicker layer of PG (Silhavy *et al*., 2010). Anionic cell wall polymers (CWP) are found throughout the PG layer, making up 60 % of the total cell wall mass. These CWPs consist mostly of glycerol phosphate or ribitol phosphate bound to glycosyl and D-alaninyl ester residues, known as teichoic acids. Cell wall teichoic acids are covalently attached to PG, whereas lipoteichoic acids are intercalated into the inner membrane and extend into the PG (Neuhaus and Baddiley, 2003). However, the lack of outer membrane means that extracellular proteins need to be bound or

anchored to the inner plasma membrane to keep them in place. Alternatively proteins become associated with the PG layer, or teichoic acids to prevent them from diffusing away from the cell (Silhavy *et al*., 2010).

The cell wall of mycobacteria shares some structural features with Gram-positive species, but since they do not retain Gram stain, mycobacteria are classified as Gram-indeterminate bacteria. As such, mycobacteria resist de-colourisation of the Ziehl-Neelsen stain using acidic solvents after being heated and are therefore termed "Acid-fast". This is due to the relative impermeability of the mycolate matrix preventing the stain being removed (Chen *et al*., 2012). The *M. tuberculosis* cell wall possesses a thick PG layer surrounding the inner membrane, similar to Gram-positive species. However, in *M. tuberculosis* it is covalently attached to a branched macromolecule called arabinogalactan (AG), which is in turn bound to a range of mycolic acids (MA) on the non-reducing end of arabinan chains which cover the surface of the cell. Overall, this structure is known as the mycolyl-arabinogalactan-peptidoglycan (mAGP) complex (Figure 1.7) (Jankute *et al*., 2015). The mycolic acids on the outer most layer of the wall form a bilayer (with free unbound lipids) known as the mycobacterial outer membrane (MOM). Although not a true outer membrane like those possessed by Gram-negative bacteria, the MOM provides a waxy, largely impermeable barrier creating a periplasmic space between the inner membrane and the MOM (Hett and Rubin, 2008, Cook *et al*., 2009). The cell wall is further elaborated with the addition of lipomannan (LM) and lipoarabinomannan (LAM), which are intercalated glycolipids that are important to mycobacteria. These large structures begin with glycosylphosphatidylinositol (GPI), an acylated lipid anchor into either the inner membrane or outer membrane, which is bound to a branched mannan, further elongated by arabinan chains in the case of LAM. The arabinan branches are capped with units of mannose to complete the molecule, which are believed to be important in the virulence of mycobacteria, and *M. tuberculosis* in particular, such as

having a role in preventing apoptosis of the infected macrophage, as well as inhibiting phagosome-lysosome fusion (Briken *et al*., 2004, Jankute *et al*., 2012, Mazurek *et al*., 2012).

Recent work using cryo-electron microscopy had shed light on the dimensions of the mycobacterial cell wall, as well as its arrangement. It has been shown that overall size of the cell wall is around 30 nm in length, broken down to the MOM (7-8 nm), arabinogalactan-peptidoglycan (AGP) (6-7 nm) and then a region of periplasmic space, rich in LM and LAM, that has not been considered in previous models (14-17 nm) (Figure 1.7) (Meroueh *et al*., 2006, Hoffmann *et al*., 2008, Zuber *et al*., 2008, Sani *et al*., 2010). NMR data of synthetic PG fragments have suggested that rather than the traditional parallel to the plasma membrane structure of PG, it actually forms right handed helices consisting of 3 *N*-acetylglucosamine-*N*-acetylmuramic acid repeats per turn (Meroueh *et al*., 2006). This arrangement of PG suggests that the AGP of the mycobacterial cell wall adopts a "scaffold" structure, whereby the backbone of the PG and the galactan moiety is perpendicular to the plasma membrane and provides rational sites for arabinan attachment (and therefore mycolate attachment) within the dimensions of the cell wall (Figure 1.7) (Kaur *et al*., 2009, Minnikin *et al*., 2015).

However, despite a helical structure of PG being proposed, it remains controversial as numerous reports dispute this model in Gram-positive bacteria. Experiments performed using atomic-force microscopy (AFM) have shown evidence of a planar topography of PG glycan strands in both *Staphylococcus aureus* and *Lactococcus lactis,* similar to Gram-negative PG (Turner *et al*., 2010, Andre *et al*., 2010), whereas the same technique showed the PG glycan strands in the rod-shaped *B. subtilis* as 50 nm wide coiled-coils running around the short axis of the cell (Hayhurst *et al*., 2008). Yet, more recently AFM was combined with electron cryo-tomography (ECT) and *in silico* modelling, which showed that the cell wall of *B. subtilis* is uniformly dense, not organised into hollow coils, as well as a distinct lack of ~50 nm gaps for new coils to be inserted into the cell wall, casting doubt upon the coiled-coil model (Beeby *et*

al., 2013). Through further examination of the orientation of peptide cross-links within the *B. subtilis* PG, which were seen to run in parallel to the long axis of the cell, as well as molecular dynamic simulations of the way the PG folds once broken, it was concluded that the glycan strands more likely run in a circumferential planar topography (Beeby *et al*., 2013). However, this model does not seem capable of accommodating the long glycan strand lengths seen in previous AFM data, suggesting that further research is required before a definitive model is produced (Hayhurst *et al*., 2008).

Nonetheless, the overall architecture of rod-shaped *M. tuberculosis* PG remains elusive. Since there may not be one over-arching model that pertains to every instance of bacterial PG (Lovering *et al*., 2012, Turner *et al*., 2014), it is possible that with the spatial constraints of the cell wall, as well as the relatively heavy cross-linking in mycobacterial PG, the previously mentioned proposed helical model could be likely (Minnikin *et al*., 2015). Not only does this newly proposed model of the *M. tuberculosis* cell wall take into account the dimensions of each "layer" of the cell wall, as well as a periplasmic space between the mycobacterial inner membrane (MIM) and AGP, it also proposes an initial anchorage of LM and LAM within the MIM based upon a common lipid anchorage to phosphatidylinositol mannosides (PIMs) found within the MIM (Figure 1.7) (Minnikin *et al*., 2015). Previous studies suggest that a potential transfer to the MOM at a later date, by an as of yet undefined mechanism, could occur since lipoglycans have been found at the cell surface of mycobacteria, but they are also present in the membranes of other Actinobacteria which do not possess mycolic acids or an outer membrane (Pitarque *et al*., 2008, Minnikin *et al*., 2015). However, it should be noted that the proposed model structure is made up of data from various members of the Corynebacterineae, which despite their many structural similarities, may not pertain to *M. tuberculosis* cell wall structure in the long run (Figure 1.7) (Minnikin *et al*., 2015).

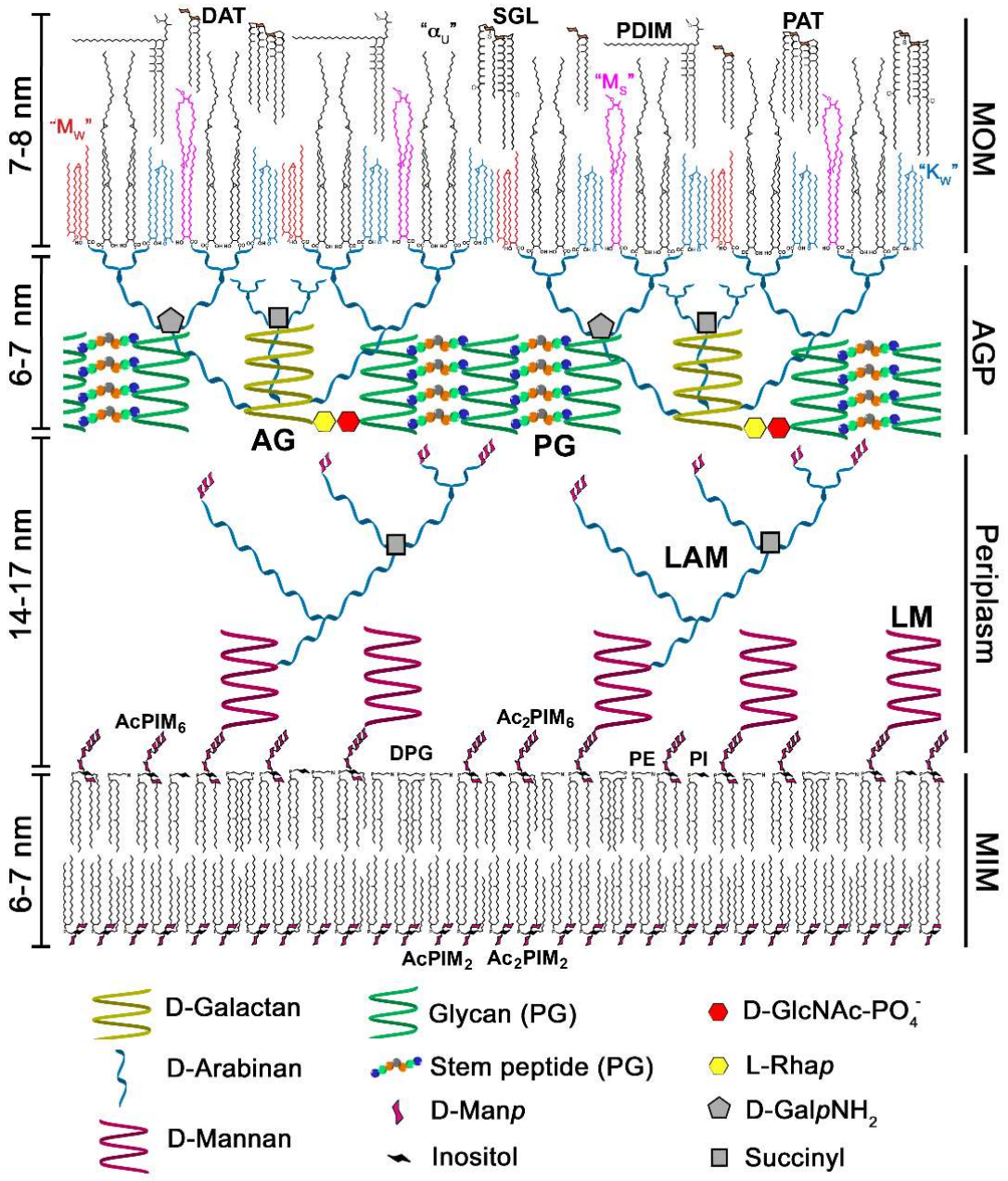

Figure 1.7: Structural representation of the mycobacterial cell wall, showing dimensions and interactions between each section. The mycobacterial inner membrane (MIM) is 6-7 nm in size and is interspersed with varying phosphatidylinositol mannosides (PIM). Within the periplasmic space (14-17 nm) LM and LAM are anchored to the MIM and extend throughout the space. The "scaffold" structure of AGP spans 6-7 nm and sits perpendicular to the MIM. Helical PG is bound to helical galactan through a unique linker region (α-L-Rha*p*(1→3)-α-D-GlcNAc-(1→P)), further bound to arabinan chains, which extend outwards and specific non-reducing ends become esterified with mycolic acids. These mycolates (α, keto (K) and methoxy (M)) of varying shapes (U,S and W) form a bilayer with free lipids (diacyl trehalose (DAT), pentaacyl trehalose (PAT), sulfoglycolipid (SGL) and phthiocerol dimycocerosate (PDIM)) on the outside of the cell which is referred to as the mycobacterial outer membrane (MOM) (7-8 nm). GalN and succinyl modifications are made to specific 3,5-Araf residues of arabinan chains of AG and LAM. The whole cell envelope spans roughly 30 nm in total from the cell surface. Figure adapted from Minnikin *et al*., 2015.

1.2.2. Peptidoglycan

1.2.2.1. Structure of Peptidoglycan

Present in almost all bacteria, PG is an important component of the cell wall, providing shape, rigidity and protection from internal osmotic pressure (Schleifer and Kandler, 1972). In addition, mycobacterial PG provides a foundation for additional covalently attached molecules that eventually form the mAGP complex (Figure 1.7). The glycan backbone of PG typically consists of *N*-acetylglucosamine (GlcNAc) bound to *N*-acetylmuramic acid (MurNAc) through β(1→4) glycosidic linkages, in a repeating fashion (Figure 1.8) (Brennan and Nikaido, 1995). However, it has been identified that mycobacterial PG can incorporate a hydroxylated version of MurNAc, *N*-glycolylmuramic acid (MurNGlyc), providing a mixture of two versions within the backbone structure (Figure 1.8) (Mahapatra *et al*., 2005, Raymond *et al*., 2005). The reason behind the modification of MurNAc to MurNGlyc is yet to be fully elucidated, but it is thought to be involved in increasing the overall strength of the PG by providing increased sites for hydrogen bonding, as well as reducing the susceptibility of the PG to lysozyme, providing enhanced protection to the cell (Brennan and Nikaido, 1995, Raymond *et al*., 2005). Nevertheless, the presence of MurNGlyc within the PG structure stimulates mammalian NOD-2 receptors more so than MurNAc, and therefore has a role in increasing NOD-2 related immunogenicity within the host (Coulombe *et al*., 2009).

Attached to the muramyl components of the backbone, tetrapeptide side chains provide the mesh-like structure of PG through cross-linkages. The side chains in *M. tuberculosis* PG are L-alaninyl-D-*iso*glutaminyl-*meso*-diaminopimelyl-D-alanine, typically cross-linked (3→4) *meso*-diaminopimelate (*m*-DAP) to D-alanine (D-ala) (Figure 1.8) (Wietzerbin *et al*., 1974). *M. tuberculosis* PG shares many features with that of *E. coli*, being within the same PG group, A1γ, based upon the classification system of Schleifer and Kandler (1972). However,

mycobacterial PG also contains a high number of (3→3) cross-linkages between *m*-DAP and *m*-DAP throughout its structure, which increase in number (up to 80 % of all cross-linkages) upon entering stationary phase (Figure 1.8) (Lavollay *et al*., 2008). *M. tuberculosis* has the ability to adjust these cross-linkages from (3→4) to (3→3) without requiring it to perform *de novo* synthesis of PG, which could help to protect the cell from endopeptidase action (Lavollay *et al*., 2008). A further feature that makes mycobacterial PG unique is that it provides the site for AG attachment. Hydroxyls at the 6 position of muramyl residues within the PG backbone offer a site for a phosphodiester bond to the AG 'linker unit', α-L-Rhamnopyranose-(1→3)-α-D-GlcNAc-(1→P) (McNeil *et al*., 1990).

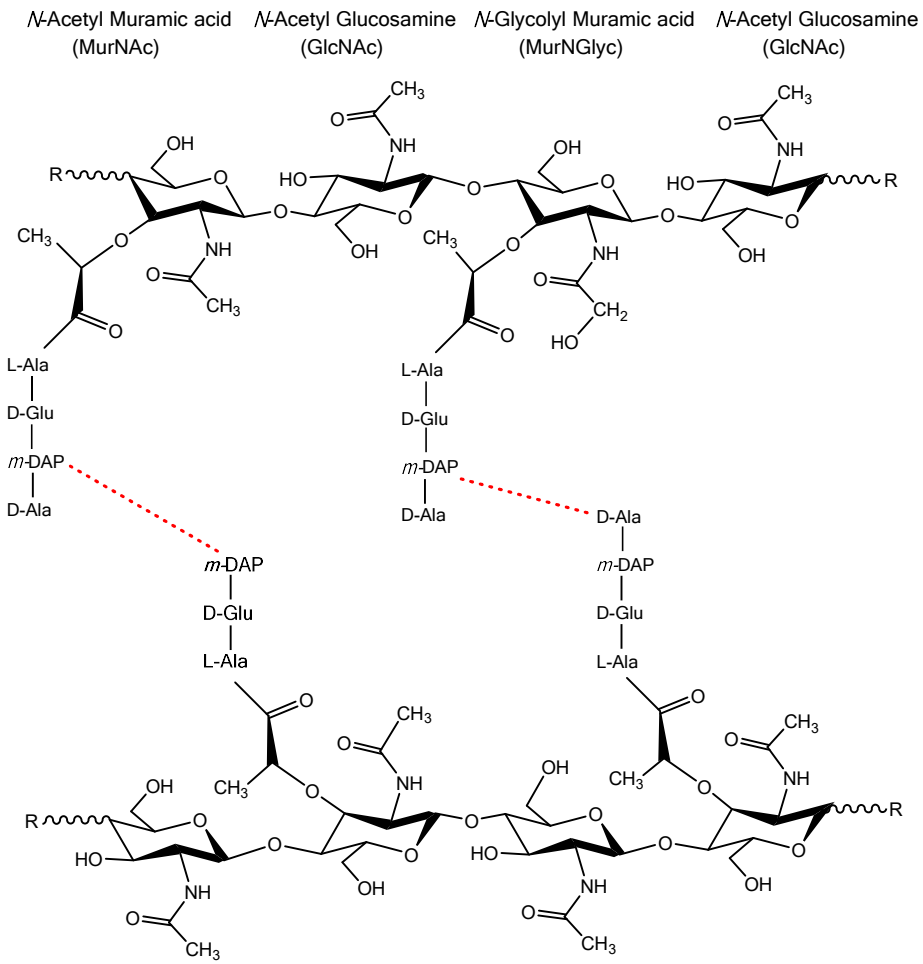

Figure 1.8: The structure of *M. tuberculosis* PG. Glycan backbone consists of repeating units of GlcNAc and MurNAc, with incorporated MurNGlyc (R side chains represent continuing glycan backbone). Tetrapeptide side chains consisting of L-ala, D-Glu, *m*-DAP and D-ala are cross-linked both *m*-DAP to D-ala (3→4) and *m*-DAP to *m*-DAP (3→3).

1.2.2.2. Synthesis of Peptidoglycan

An extensive number of proteins are utilised to synthesise PG (Table 1.1), being similar in most bacteria. Biosynthesis largely takes part within the cytoplasm, often proposed to start with UDP-GlcNAc. However, before this, UDP-GlcNAc is synthesised by 3 enzymes, initiated by GlmS which converts D-fructose-6-phosphate to D-glucosamine-6-phosphate (Badet *et al*., 1987). This is then acted on by the mutase GlmM/MrsA, which transfers the phosphate group from D-glucosamine-6-phosphate to D-glucosamine-1-phosphate (Li *et al*., 2011). The bifunctional action of GlmU is responsible for converting D-glucosamine-1-phosphate to *N*-acetylglucosamine-1-phosphate, *via* acyltransferase activity, and subsequently producing UDP-GlcNAc *via* uridylyltransferase activity, in order to produce the traditional 'starting point' for PG synthesis (Zhang *et al*., 2008).

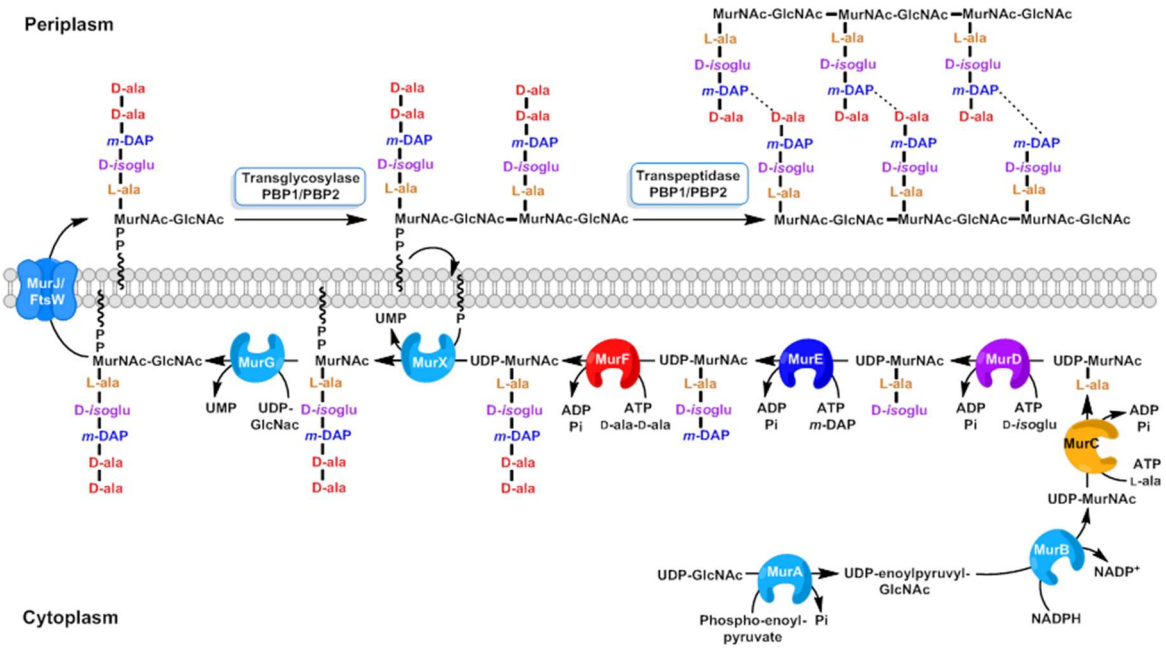

Figure 1.9: Mur ligase pathway showing the synthesis of PG on both sides of the plasma membrane. UDP-GlcNAc becomes converted into UDP-MurNAc through the action of MurA and MurB. At this point MurC begins the sequential addition of amino acids, followed by MurD, MurE and MurF to form UDP-MurNAc-pentapeptide. MurX then causes UDP-MurNAc-pentapeptide to become decaprenyl-phosphate linked, known as "Lipid I". The action of MurG adds GlcNAc to "Lipid I" to form "Lipid II", which then becomes translocated across the plasma membrane by the action of MurJ/FtsW. Once on the outside of the cell, "Lipid II" becomes transglycoslated and then transpeptidated by the PBP1/PBP2 (PonA1/PonA2) proteins, forming nascent PG.

From this point, substrates enter the sequential Mur ligase pathway (Table 1.1, Figure 1.9). UDP-GlcNAc is converted to enoylpyruvyl-UDP-N-acetylglucosamine (UDP-enoylpyruvylGlcNAc) by UDP-N-acetylglucosamine enolpyruvyl transferase activity of MurA, using phospho-enoyl-pyruvate as a substrate. This product is then reduced by UDP-*N*-acetylpyruvyl-glucosamine reductase (MurB), with the use of NADPH, to form UDP-MurNAc by conversion of the enolpyruvyl moiety to a lactoyl ether moiety (Benson *et al*., 1996). This completes the first 'stage' of the pathway. At this point, NamH can hydroxylate the methyl group of the MurNAc moiety of UDP-MurNAc into UDP-MurNGlyc to feed into the rest of the pathway (Raymond *et al*., 2005). From here amino acids are systematically adjoined to UDP-Muramyl (either MurNAc or MurNGlyc) to form a pentapeptide, by the ATP-dependent Mur ligases. The first of these being the addition of L-alanine (L-ala) by UDP-*N*-acetylmuramyl:L-alanine ligase (MurC) to form UDP-Muramyl-L-ala. The next step involves linking up D-*iso*glutamate (D-*iso*glu) to form UDP-Muramyl-L-ala-D-*iso*glu by the UDP-*N*-acetylmuramyl-L-alanine:D-*iso*glutamate ligase activity of MurD. The D-*iso*glu required by MurD is provided to the cell by the stereochemical conversion of L-glutamate to D-*iso*glu by the racemase MurI (Morayya *et al*., 2015). Next, *m*-DAP is attached by UDP-*N*-acetylmuramyl-L-alaninyl-D-*iso*glutamate: *meso*-diaminopimelate ligase (MurE) to form UDP-Muramyl-L-ala-D-*iso*glu-*m*-DAP. The final amino acid attached to the chain is the dipeptide D-alaninyl-D-alanine (D-ala-D-ala) by the UDP-*N*-acetylmuramyl-L-alaninyl-D-*iso*glutaminyl-*meso*-diaminopimelate: D-alaninyl-D-alanine ligase activity of MurF (Munshi *et al*., 2013). However, in order to provide the substrate for MurF, L-alanine must first be stereochemically inverted into D-alanine, by the racemase Alr (Strych *et al*., 2001). The D-alanine is then attached to another D-alanine, by the ligase Ddl, to produce D-ala-D-ala for attachment to the growing peptide chain (Figure 1.9) (Feng and Barletta, 2003). This final

addition gives rise to 'Park's nucleotide', UDP-Muramyl-L-ala-D-*iso*glu-*m*-DAP-D-ala-D-ala (Kurosu *et al*., 2007).

The third and final 'stage' of the Mur pathway involves linking the UDP-Muramyl-pentapeptide to the cell membrane and flipping it to the outside of the cell (Figure 1.9). A 50-carbon membrane bound lipid carrier containing 10 isoprenoid units (decaprenyl-phosphate) becomes attached to the UDP-Muramyl-pentapeptide *via* a transferase reaction performed by the integral membrane bound enzyme phospho-*N*-acetylmuramyl-pentapeptide-transferase (MraY/MurX), to create "Lipid I", a membrane bound pentapeptide precursor. GlcNAc is then transferred onto "Lipid I" by decaprenyldiphospho-muramylpentapeptide β-*N*-acetylglucosaminyltransferase (MurG) to create "Lipid II", a β 1-4 linked GlcNAc-Muramyl-pentapeptide. "Lipid II" will become the final monomeric unit of PG structure, and its synthesis is the last known step in the Mur ligase pathway (Smith, 2006, El Zoeiby et al., 2003, Bouhss et al., 1999).

The process which flips the monomer across the plasma membrane to the outside of the cell is currently a debated topic. It was initially thought that the protein involved in this mechanism was MurJ in *E. coli* (Ruiz, 2008), but this has since been debated to be the enzyme FtsW, as this has been shown to perform "Lipid II" flippase activity *in vitro* (Mohammadi *et al*., 2011). However, the Ruiz group have since shown data that shows MurJ having *in vivo* "Lipid II" flippase activity in *E. coli* (Sham *et al*., 2014). Nonetheless, further research is required for both proteins in order to truly identify and characterise the "Lipid II" flippase.

Once the monomeric units are outside the cell, they become polymerised by transglycosylation reactions, performed by the bifunctional proteins PonA1/PBP1 and PonA2/PBP2 (Table 1.1). This process involves the GlcNAc moiety of the "Lipid II" monomer becoming attached to the muramyl moiety of the growing PG backbone (Hett *et al*.,

2010). The second function of both PonA1 and PonA2 is D,D-transpeptidase activity, which form classical (3→4) cross-linkages between *m*-DAP and D-ala residues of neighbouring peptide chains, at the expense of cleavage of the terminal D-ala (Chang *et al*., 1990). Both PonA1 and PonA2 have additional domains, including a transmembrane helix, as well as a PASTA (PBP and Serine/Threonine associated) domain. However, these do not behave as traditional PASTA domains, as they do not bind to typical ligands such as β-lactams, muropeptides or nascent PG (Calvanese *et al*., 2014). Therefore further characterisation of these domains is required.

When *M. tuberculosis* enters a non-replicating phase, the cell rearranges the PG in order to increase the number of (3→3) cross-linkages between *m*-DAP and *m*-DAP, through the action of nonclassical L,D-transpeptidases (Table 1.1) (Cordillot *et al*., 2013). Within the *M. tuberculosis* genome, five paralogues have currently been identified, known as Ldt_{Mt1-5}, however, only Ldt_{Mt1} and Ldt_{Mt2} depleted strains showed any phenotypic difference within their PG structure, since Ldt_{Mt3}, Ldt_{Mt4} and Ldt_{Mt5} deficient strains showed no difference in growth rate or colony phenotype when compared to wild-type (Schoonmaker *et al*., 2014). These L,D-transpeptidases convert (3→4) cross-linkages to (3→3) through the cleavage of the remaining D-alanine residue on the donor peptide chain. This leaves behind a tripeptide side chain in the stationary phase PG (Sacco *et al*., 2010).

Table 1.1: List of genes, and their function, involved in the synthesis of Peptidoglycan (PG). *Genes are involved the synthesis of precursors to the pathway. ♦Genes are debated to have same role.

Gene name	Rv number	Role	Reference
*glmS**	Rv3436c	Converts D-fructose-6-P to D-GlcN-6-P	(Badet *et al*., 1987)
*glmM /mrsA**	Rv3441c	Converts D-GlcN-6-P to D-GlcN-1-P	(Li *et al*., 2011)
*glmU**	Rv1018c	Converts to D-GlcN-1-P to GlcNAc-1-P, then to UDP-GlcNAc	(Zhang *et al*., 2008)
murA	Rv1315	Converts UDP-GlcNAc to UDP-enoylpyruvylGlcNAc	(Kim *et al*., 1996)
murB	Rv0482	Converts UDP-enoylpyruvylGlcNAc to UDP-MurNAc	(Benson *et al*., 1996)
murC	Rv2152c	Attaches L-ala to UDP-MurNAc	(Mahapatra *et al*., 2000)
murD	Rv2155c	Adds D-*iso*glu to L-ala in peptide chain	(Smith, 2006)
murE	Rv2158c	Adds *m*-DAP to D-*iso*glu in peptide chain	(Basavannacharya *et al*., 2010)
murF	Rv2157c	Adds D-ala-D-ala to *m*-DAP in peptide chain	(Munshi *et al*., 2013)
murX /mraY	Rv2156c	Attaches lipid to MurNAc-pentapeptide chain ('Lipid I')	(Bouhss *et al*., 1999)
murG	Rv2153c	Adds UDP-GlcNAc to MurNAc-pentapeptide-lipid ('Lipid II')	(Trunkfield *et al*., 2010)
murJ♦	Rv3910 (probable)	Translocates 'Lipid II' across plasma membrane	(Ruiz, 2008)
ftsW♦	Rv2154c	Translocates 'Lipid II' across plasma membrane	(Mohammadi *et al.*, 2011)
ponA1 /pbp1	Rv0050	Bifunctional: transglycosylation of GlcNAc to MurNAc. Also transpeptidation to cross-link *m*-DAP to D-ala (3→4)	(Hett *et al*., 2010)
ponA2 /pbp2	Rv3682		(Chang *et al*., 1990)
ldt_{mt1}	Rv0116c	Nonclassical L,D-transpeptidase action, converts (3→4) *m*-DAP to D-ala cross-linkages to (3→3) *m*-DAP to *m*-DAP	(Schoonmaker *et al*., 2014)
ldt_{mt2}	Rv2518c		(Schoonmaker *et al*., 2014)
ldt_{mt3}	Rv1433		(Schoonmaker *et al*., 2014)
ldt_{mt4}	Rv0192		(Schoonmaker *et al*., 2014)
ldt_{mt4}	Rv0483		(Schoonmaker *et al*., 2014)
namH	Rv3808	Converts UDP-MurNAc to UDP-MurNGlyc	(Raymond *et al*., 2005)
*murI**	Rv1338	Converts L-glutamate to D-*iso*glu	(Morayya *et al*., 2015)
*alr**	Rv3423c	Converts L-ala to D-ala	(Strych *et al*., 2001)
*ddl**	Rv2981c	Ligates D-ala to D-ala	(Feng and Barletta, 2003)

1.2.3. Arabinogalactan

1.2.3.1. Structure of Arabinogalactan

Mycobacterial PG is further elaborated with a unique heteropolysaccharide termed arabinogalactan (AG) (Figure 1.7). AG is a branched polysaccharide, covalently attached to around 10-12 % of the muramyl units of PG *via* a phosphodiester bond between the C-6 hydroxyl of the muramyl residue and the GlcNAc of the AG 'linker unit' (LU) (McNeil *et al*., 1990). AG is comprised mostly of arabinose and galactose monomer units found only in their furanose configuration (Ara*f* and Gal*f*), and can be broken down into 3 seperate components, arabinan, galactan and the LU (Figure 1.10) (Crick *et al*., 2001, McNeil *et al*., 1987). The LU is comprised of P-GlcNAc (bound to decaprenyl-monophosphate before attachment to PG), in turn covalently bound to L-rhamnopyranose (Rha*p*) through an α(1→3) glycosidic bond (McNeil *et al*., 1990). The linear galactan constituent is made up of approximately 30 Gal*f* residues alternating β(1→5) and β(1→6) linkages, with the first Gal*f* unit bound to the Rha*p* residue of the LU in an α(1→4) fashion (McNeil *et al*., 1987, McNeil *et al*., 1990).

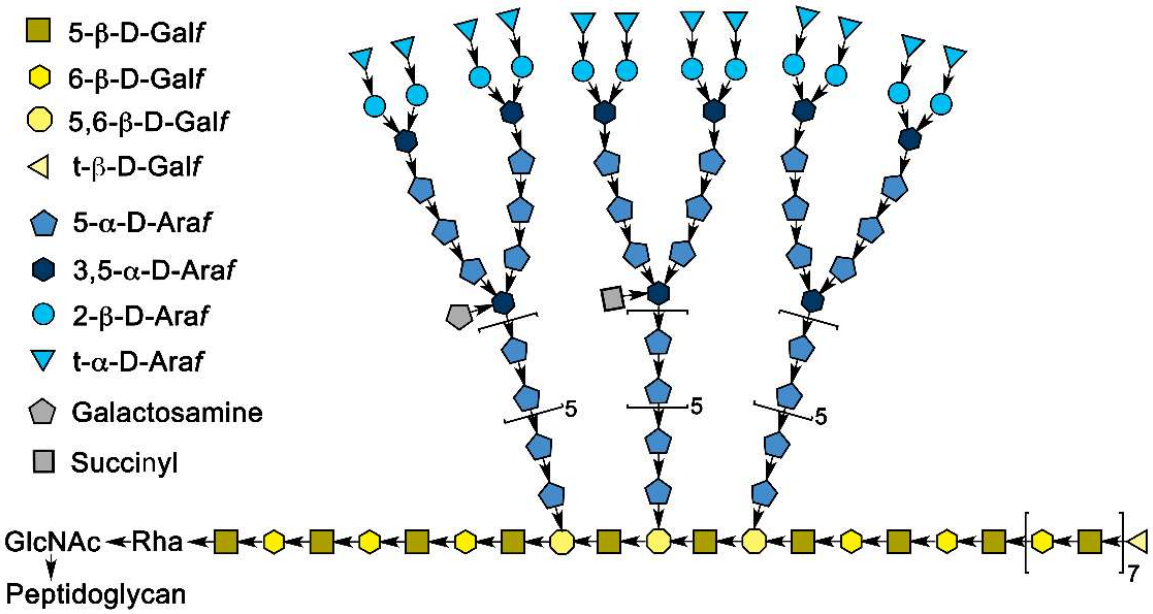

Figure 1.10: The structure of mycobacterial AG. PG bound linker unit (α-L-Rha*p*(1→3)-α-D-GlcNAc-(1→P)) is elaborated with approximately 30 Gal*f* units in alternating β(1→5), β(1→6) linkages to form mature galactan. At positions 8, 10 and 12 of the galactan chain, highly branched arabinan is attached, formed of α(1→5) linked Ara*f* residues, with 3,5 branch points introduced by α(1→3) linkages. The non-reducing termini are β(1→2) linked, forming Araf_6 terminal motif, which provides the site for mycolate attachment. AG is further modified by addition of GalN or succinyl residues at the 3,5-Ara*f* branch points. Figure adapted from Minnikin *et al*., 2015

Ara*f* residues were found to be attached to specific β(1→6) Gal*f* residues at the 8th, 10th and 12th positions of the galactan chain within the AG of *C. glutamicum*, which has largely been believed to be the case in *M. tuberculosis* AG, since mAGP architecture is largely conserved across members of the Corynebacterineae (Figure 1.10) (Besra *et al*., 1995, Alderwick *et al*., 2005, Dianiskova *et al*., 2011). However, more recent research has suggested that *M. tuberculosis* cell walls have only two arabinan chains that are attached to the galactan polymer, rather than three, a topic which remains debated (Bhamidi *et al*., 2011). Nonetheless, attached Ara*f* residues are elongated by α(1→5) linkages, with α(1→3) linkages introducing 3,5-Ara*f* branch points, which become further elongated themselves by additional α(1→5) linkages (Figure 1.10) (Amin *et al*., 2008, Birch *et al*., 2008). The highly branched arabinan chains are terminated at the nonreducing ends by a β(1→2) linkage, with the final structural

motif being [β-Ara*f*-(1→2)-α-Ara*f*]$_2$-3,5-α-Ara*f*-(1→5)-α-Ara*f*, a characteristic hexa-arabinofuranoside (Ara*f*$_6$) (McNeil *et al*., 1994). Of this motif, the available C-5 hydroxyls represent a site for mycolic acid attachment (Figure 1.7, Figure 1.10). However, only around two thirds of these motifs become mycolylated in the final cell wall (McNeil *et al*., 1991). AG is further adorned with a single galactosamine (GalN) residue at the C-2 position of some inner 3,5-Ara*f* residues (often the penultimate residue of the Ara*f*$_6$ motif), found as a single residue side chain on around one third of chains (Draper *et al*., 1997, Peng *et al*., 2012). Another substituent found on AG is the addition of succinyl residues, found attached at the same locations as GalN residues. These succinyl residues appear on a further one third of arabinan chains (Bhamidi *et al*., 2008). It is suggested that these substituents may have a role in strengthening anionic interactions between AG and rest of the cell wall, as well as potentially having a role in the orchestration of mycolylation of arabinan chains (Draper *et al*., 1997, Bhamidi *et al*., 2008).

1.2.3.2. Synthesis of Arabinogalactan

Biosynthesis of AG requires a plethora of biosynthetic machinery (Table 1.2) and is initiated within the cytoplasm, with the formation of the linker region by the enzyme WecA/Rfe. This enzyme catalyses the addition of GlcNAc-P (from UDP-GlcNAc) to decaprenyl-monophosphate (C_{50}-P), to form GlcNAc-P-P-C_{50}, known as glycolipid 1 (GL-1) (Mikusova *et al*., 1996). The linker is completed by the further addition of Rha*p* (from dTDP-β-L-Rha*p*) by the enzyme WbbL1, to form α-L-Rha*p*-(1→3)-α-D-GlcNAc-(1→P) (Rha-GlcNAc-P-P-C_{50}), known as glycolipid 2 (GL-2) (Figure 1.11) (Birch *et al*., 2008, McNeil *et al*., 1990). The dTDP-Rha that is required for completion of the LU comes from a series of reactions, which arise from the action of the products from the gene cluster *rmlABCD*. Firstly, dTTP is attached to α-D-glucose-1-P to form dTDP-α-D-glucose, and release inorganic pyrophosphate (PPi), due to the action of the enzyme RmlA (Ma *et al*., 1997). In the next step, dTDP-α-D-

glucose 4,6-dehydratase (RmlB) action forms dTDP-α-6-deoxy-D-*xylo*-4-hexulose which in turn becomes epimerised by RmlC to dTDP-β-6-deoxy-L-*lyxo*-4-hexulose (Stern *et al*., 1999, Ma *et al*., 2001). Finally, the action of RmlD forms dTDP-β-L-Rha*p*, which is ready for incorporation into the growing LU (Hoang *et al*., 1999).

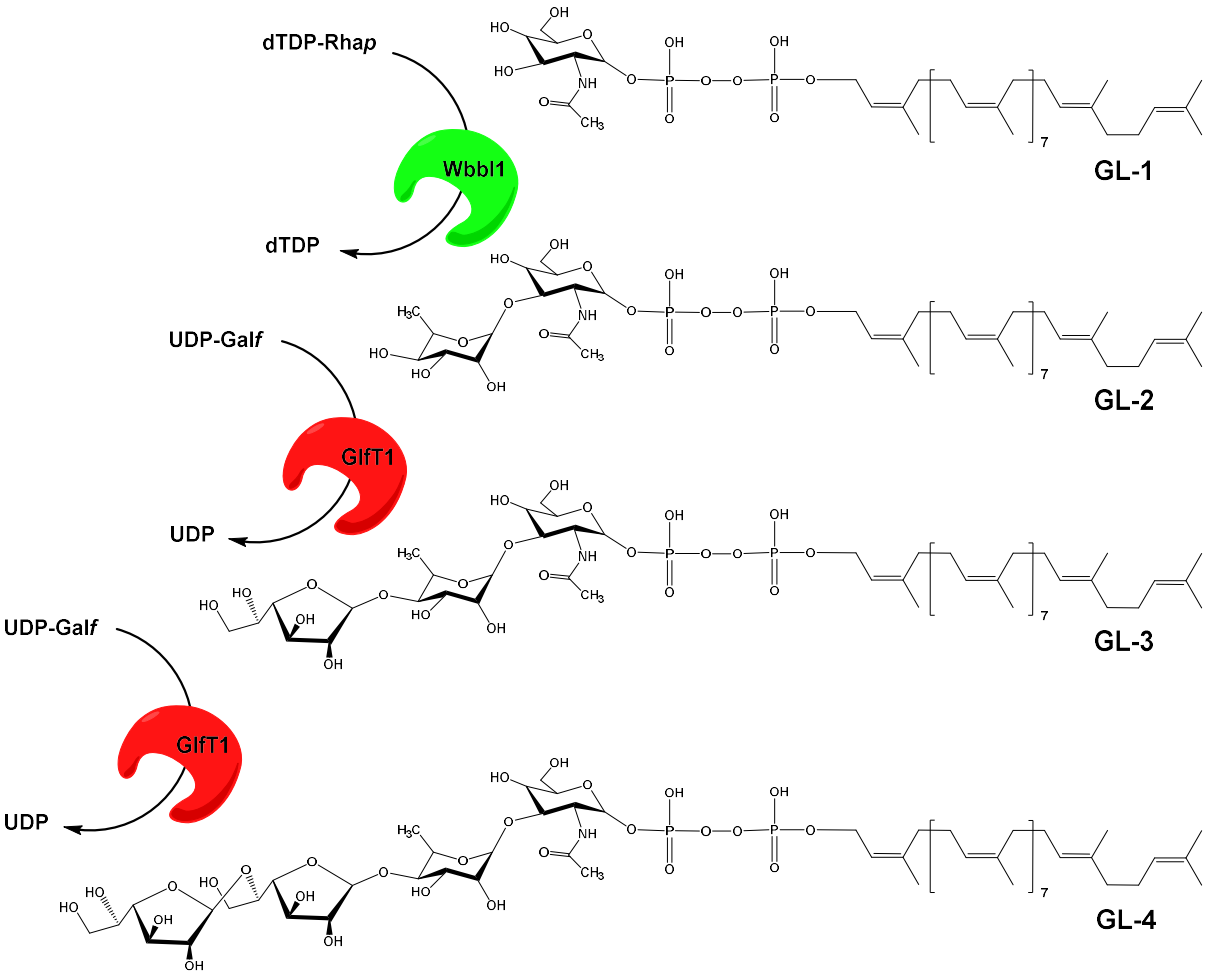

Figure 1.11: Structures of Glycolipid (GL) 1, 2, 3 and 4. GL-1 consists of decaprenyl-P-P-GlcNAc, which then becomes GL-2 with the addition of Rha*p* by Wbbl1. GlfT1 then adds a Gal*f* to form GL-3 and immediately adds an additional Gal*f* to form GL-4.

The precursor to galactan synthesis, the high energy donor UDP-Gal*f*, is made through a two-part reaction (Table 1.2). Firstly, UDP-glucopyranose is converted to UDP-galactopyranose (UDP-Gal*p*) through the epimerase action of GalE1. UDP-Gal*p* is then further converted to the furanose ring form UDP-Gal*f* by the enzyme Glf (Weston *et al*., 1997). UDP-

galactofuranosyl transferases (GlfT) then begin to build the galactan chain upon the LU, through the enzyme GlfT1, which begins with the catalysis of two units of UDP-Gal*f* onto the linker, in a bifunctional capacity. It attaches one Gal*f* to Rha-GlcNAc-P-P-C_{50} (known as glycolipid 3 (GL-3)), then immediately adds another to form Gal*f*-Gal*f*-Rha-GlcNAc-P-P-C_{50} (known as glycolipid 4 (GL-4)) (Figure 1.11) (Mikusova *et al*., 1996, Mikusova *et al*., 2006). GlfT2 then continues to add Gal*f* residues to the growing galactan chain, in a linear fashion of alternating 5-linked and 6-linked β-D-Gal*f*, until 30 Gal*f* residues are reached (Kremer *et al*., 2001a). The linear galactan is then translocated across the plasma membrane to the outside of the cell, by a conserved ABC transport system, the enzymes Wzt and Wzm (Dianiskova *et al*., 2011) (Table 1.2).

Once the completed galactan chain is outside the cell, the highly-branched arabinan can be added to the complex. All arabinose within the cell is attained from a single donor molecule, β-D-arabinofuranosyl-1-monophosphoryldecaprenol (DPA) (Alderwick *et al.*, 2005) (Table 1.2). The synthesis of DPA begins with 5-phosphoribosyl-1-pyrophosphate (pRpp), itself synthesised by the addition of pyrophosphate from ATP to ribose-5-phosphate by the enzyme PrsA (Alderwick *et al*., 2011b). Ribose-5-phosphate from the donor pRpp is attached to decaprenyl-phosphate by UbiA to form decaprenylphosphoryl-5-β-D-phosphoribose (DPPR) (Huang *et al*., 2005). Rv3807c is then believed to cause dephosphorylation of DPPR to give rise to decaprenylphosphoryl-5-β-D-ribose (DPR). An epimerisation reaction then occurs, catalysed by DprE1 (to form the keto-intermediate DPX) and DprE2, to create DPA (Batt *et al*., 2012, Grover *et al*., 2014). DprE1 oxidises the hydroxyl moiety at the C-2 of ribose of DPR, to form decaprenylphosphoryl-2-keto-β-D-*erythro*-pentofuranose (DPX). DprE2 then reduces the C-2 keto moiety of DPX to form the final donor, DPA (Mikusova *et al*., 2005).

The initial Ara*f* residues taken from DPA are attached to the C-5 hydroxyls of Gal*f* residues of galactan, at positions 8, 10 and 12, by the first arabinosyltransferase, AftA (Alderwick et al.,

2006). The arabinan chain is then prolonged by the heterodimer of EmbA and EmbB, which cause α(1→5) linkage between Ara*f* monomers (Amin *et al*., 2008). Internal α(1→3) branching is introduced by AftC, and also by the largest arabinosyltransferase AftD, which is thought to have other functions (e.g. acting as a scaffold for other multienzymatic complexes) due to its large size (Birch *et al*., 2008, Jankute *et al*., 2012). The arabinan chains are then capped by β(1→2) linkages that form the non-reducing Araf_6 motif, catalysed by AftB (Seidel *et al*., 2007). GalN residues are synthesised through a pathway beginning with PpgS, which utilises polyprenyl-phosphate and UDP-N-Acetylgalactosamine (UDP-GalNAc) to form polyprenyl-P-GalNAc. This then becomes deacetylated and translocated across the membrane by an as of yet unknown mechanism. Finally, Rv3779 transfers GalN onto the C-2 position of 3,5-Ara*f* residues of the arabinan chains (Skovierova *et al*., 2010). The enzymes and mechanisms involved in the attachment of succinyl residues to arabinan chains, currently remain unknown (Jankute *et al*., 2015).

Table 1.2: List of genes, and their function, involved in the synthesis of Arabinogalactan (AG). *Genes are involved the synthesis of precursors to the pathway.

Gene name	Rv number	Role	Reference
*rmlA**	Rv0334	Addition of dTTP + α-D-glucose-1-P to form dTDP-α-D-glucose + PPi	(Ma *et al*., 1997)
*rmlB**	Rv3464	Converts dTDP-α-D-glucose to dTDP-α-6-deoxy-D-*xylo*-4-hexulose	(Ma *et al*., 2001)
*rmlC**	Rv3465	Converts dTDP-α-6-deoxy-D-*xylo*-4-hexulose to dTDP-β-6-deoxy-L-*lyxo*-4-hexulose	(Stern *et al*., 1999)
*rmlD**	Rv3266	Uses NADH to convert dTDP-β-6-deoxy-L-*lyxo*-4-hexulose to dTDP-β-L-Rha*p*	(Hoang *et al*., 1999)
wecA /rfe	Rv1302	Attaches GlcNAc-P (from UDP-GlcNAc) to decaprenyl-monophosphate (C_{50}-P)	(McNeil *et al*., 1990)
wbbL1	Rv3265c	Addition of dTP-β-L-Rha*p* to GlcNAc-P-P-C_{50} to form 'linker unit'	(Birch *et al*., 2008)
*galE1**	Rv3634c	Conversion of UDP-glucose*p* to UDP-Gal*p*	(Weston *et al*., 1997)
*glf**	Rv3809c	Conversion of UDP-Gal*p* to UDP-Gal*f*	(Weston *et al*., 1997)
glfT1	Rv3782	Addition of initial Gal*f* residues onto linker unit	(Mikusova *et al*., 2006)
glfT2	Rv3808c	Continuation of Gal*f* addition up to Galf_{30} – linker unit	(Kremer *et al*., 2001a)
wzt	Rv3781	Flips Galf_{30} - linker unit across membrane alongside Wzm.	(Dianiskova *et al.*, 2011)
wzm	Rv3783	Flips Galf_{30} - linker unit across membrane alongside Wzt.	(Dianiskova *et al*., 2011)
aftA	Rv3792	Transfers the first Ara*f* unit onto the Galactan domain	(Alderwick et al., 2006)
aftB	Rv3805c	β(1→2) branching and capping of arabinan	(Seidel *et al.*, 2007)
aftC	Rv2673	Introduces α(1→3) branching in arabinan	(Birch *et al*., 2008)
aftD	Rv0236c	α(1→3) branching and potential scaffolding properties	(Skovierova *et al.*, 2009)
embA	Rv3794	Heterodimer with EmbB, α(1→5) addition of ara*f* to arabinan	(Amin *et al*., 2008)
embB	Rv3795	Heterodimer with EmbA, α(1→5) addition of ara*f* to arabinan	(Amin *et al*., 2008)
prsA	Rv1017c	Transfers PPi from ATP onto ribose-5-P to form pRpp	(Alderwick *et al*., 2011b)
*ubiA**	Rv3806c	Transfer of ribose-phosphate to decaprenyl-phosphate	(Huang *et al*., 2005)
? (unknown)*	Rv3807c	(Believed) Dephosphorylation of DPPR to form DPR	(Grover *et al*., 2014)
*dprE1**	Rv3790	First step of DP epimerisation DPR→DPX intermediate	(Batt *et al*., 2012)
*dprE2**	Rv3791	Second step of DP epimerisation DPX intermediate→DPA	(Batt *et al*., 2012)
*ppgS**	Rv3631	Addition of polyprenyl-P and UDP-GalNAc to form polyprenyl-P-GalNAc	(Skovierova *et al*., 2010)
? (unknown)	Rv3779	Attachment of GalN residues to arabinan chain at C-2 of 3,5-Ara*f*	(Skovierova *et al*., 2010)

1.2.4. Mycolic Acids

1.2.4.1. Structure of Mycolic Acids

The AG of mycobacteria is decorated with a variety of MAs. In *M. tuberculosis*, the MAs are involved in the permeability of the cell wall, as well as the ability of the cell to form biofilms (Ojha *et al*., 2005, Sambandan *et al*., 2013). They also have a role in the pathogenicity of *M. tuberculosis* (Glickman *et al*., 2000, Vander Beken *et al*., 2011). There are 3 types of MAs produced by *M. tuberculosis*, those being α-, methoxy- and keto- structures. The α-mycolates are by far the most abundant within the cell wall, at around 70 %, whereas the methoxy- and keto- forms account for 10-15 % each (Takayama *et al*., 2005). However, all MAs share the same basic structure of an alpha-alkyl chain (C_{26} in length) branching from a meromycolic backbone (up to around C_{60} in length) (Figure 1.12) (Marrakchi *et al*., 2014). The "mycolic motif" at the juncture of the two chains consists of a β-hydroxy acid group (consisting of both a carboxylic acid group as well as a hydroxyl group, separated by two carbon atoms) and is a conserved characteristic of MAs across the Corynebacterineae (Figure 1.12) (Lea-Smith *et al*., 2007). The α-mycolates are found only in a *cis*-cyclopropane configuration, whereas the methoxy- and keto-mycolates can be found as either *cis*- or *trans*-cyclopropane configurations, with a distal methyl group attached (Watanabe *et al*., 2002, Jankute *et al*., 2015). MAs become modified post-synthesis, with the blend of subtypes altering the fluidity and as a result the permeability of the MOM and therefore the cell wall if needed, as well as increasing the resistance of the cell wall to oxidative stresses (Yuan *et al*., 1995).

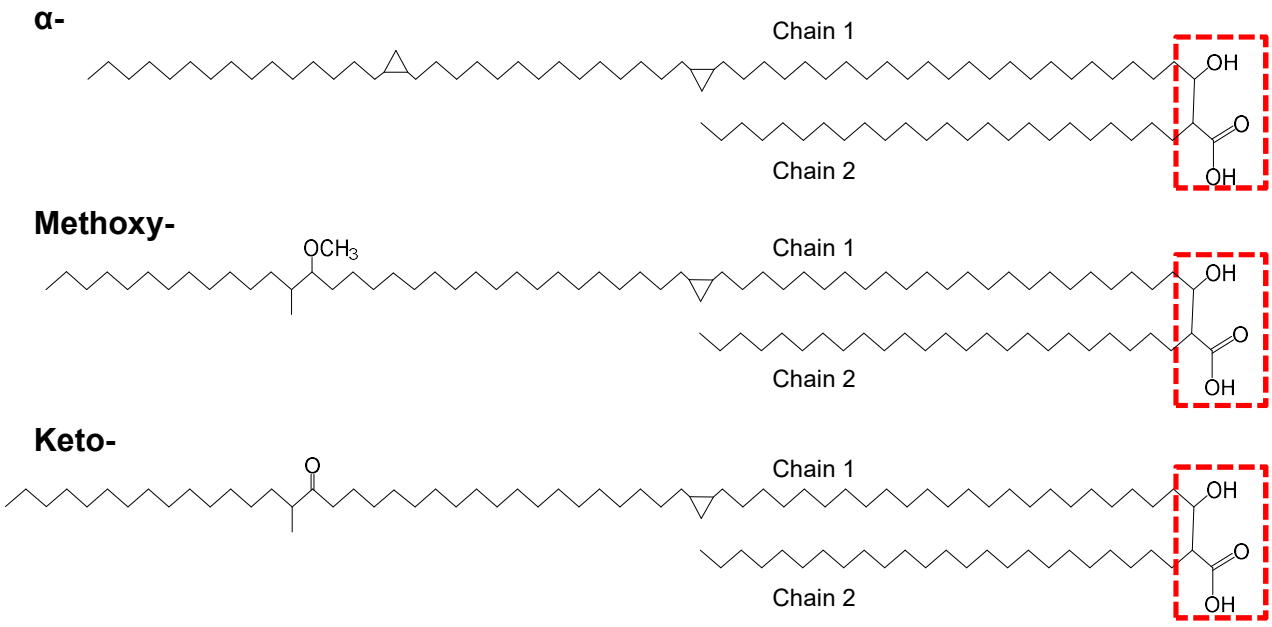

Figure 1.12: Basic structure of MAs from *M. tuberculosis*. Chain 1 represents the meromycolic backbone chain (up to C_{58}) containing *cis*-cyclopropane modifications and additional methoxy- and keto-groups within the corresponding MA subtype. Chain 2, the alpha-alkyl chain (C_{26}), is connected *via* the β-hydroxy acid "mycolic motif" (in red box).

1.2.4.2. Synthesis of Mycolic Acids

Both Acyl-CoA and Malonyl-CoA act as initial substrates for MA synthesis. Acyl/Malonyl-CoA is used by the Fatty Acid Synthase (FAS)-I system, a multifunctional enzyme complex, for *de novo* fatty acid synthesis (Figure 1.13) (Bhatt *et al.*, 2007). The FAS-I system can be found in eukaryotes and higher prokaryotes, and is comprised of several domains, each with a separate enzymatic activity. The order of discreet protein domains begins with an acyltransferase, and progresses with an enoyl reductase, dehydratase, malonyl/palmitoyl transferase, acyl carrier protein, β-ketoacyl reductase and finally a β-ketoacyl synthase. Elongation of the fatty acid chains proceeds by addition of 2 carbons at a time, and all intermediate products remain bound to the complex until being released (either at C_{12-18} for further elongation, or C_{24-26} for use in the complete mycolate) (Bhatt et al., 2007, Smith et al., 2003).

In order to shuttle FAS-I system products to further elongation, FabH binds to shorter (C_{12-18}) acyl-CoA chains and catalyses a condensation reaction with malonyl-AcpM (Mycobacterial Acyl carrier protein) to produce acyl-AcpM chains, elongated by C_2 (i.e. C_{14-20}). FabH then transports its product into the FAS-II system (Figure 1.13) (Bhatt *et al*., 2007). Malonyl-AcpM is generated by the transacylation of malonate from malonyl-CoA to holo-AcpM by FabD (Kremer *et al*., 2001b). The FAS-II pathway is, unlike FAS-I, a set of soluble enzymes found in prokaryotes and plants that perform fatty acid elongation (Table 1.3). However, unlike other bacteria, the *M. tuberculosis* FAS-II is not capable of *de novo* synthesis of fatty acids, and thus relies on the FAS-I system (Marrakchi *et al*., 2014, Bhatt *et al*., 2007). The FAS-II enzymes work in a similar fashion to the regions of FAS-I, with both systems utilising a C_2 group from malonyl-CoA/AcpM for additional elongation, however FAS-II can elongate the carbon chain to a much greater extent (Marrakchi *et al*., 2014). The first enzymes in the system are the condensing β-ketoacyl-Acp synthases, KasA and KasB (Figure 1.13). Although similar in sequence and function, it is believed that KasA catalyses the inital addition of carbons to the chain, whereas KasB is involved in the latter stages of elongation of longer meromycolate chains, due to them having different chain length specificities (Bhatt *et al*., 2007). After this, the β-ketoacyl-Acp reductase, MabA, reduces the β-ketoacyl-AcpM to β-hydroxyacyl-AcpM, which can then be used as a substrate for the heterodimeric β-hydroxyacyl-Acp dehydratases, HadAB and HadBC (Figure 1.13) (Marrakchi *et al*., 2014). The action of these enzymes produces an enoyl-AcpM, which is then reduced by NADH-dependent *trans*-2-enoyl-Acp reductase, InhA, to once again produce an acyl-AcpM which can undergo further elongation by KasB, etc (Figure 1.13) (Marrakchi *et al*., 2014).

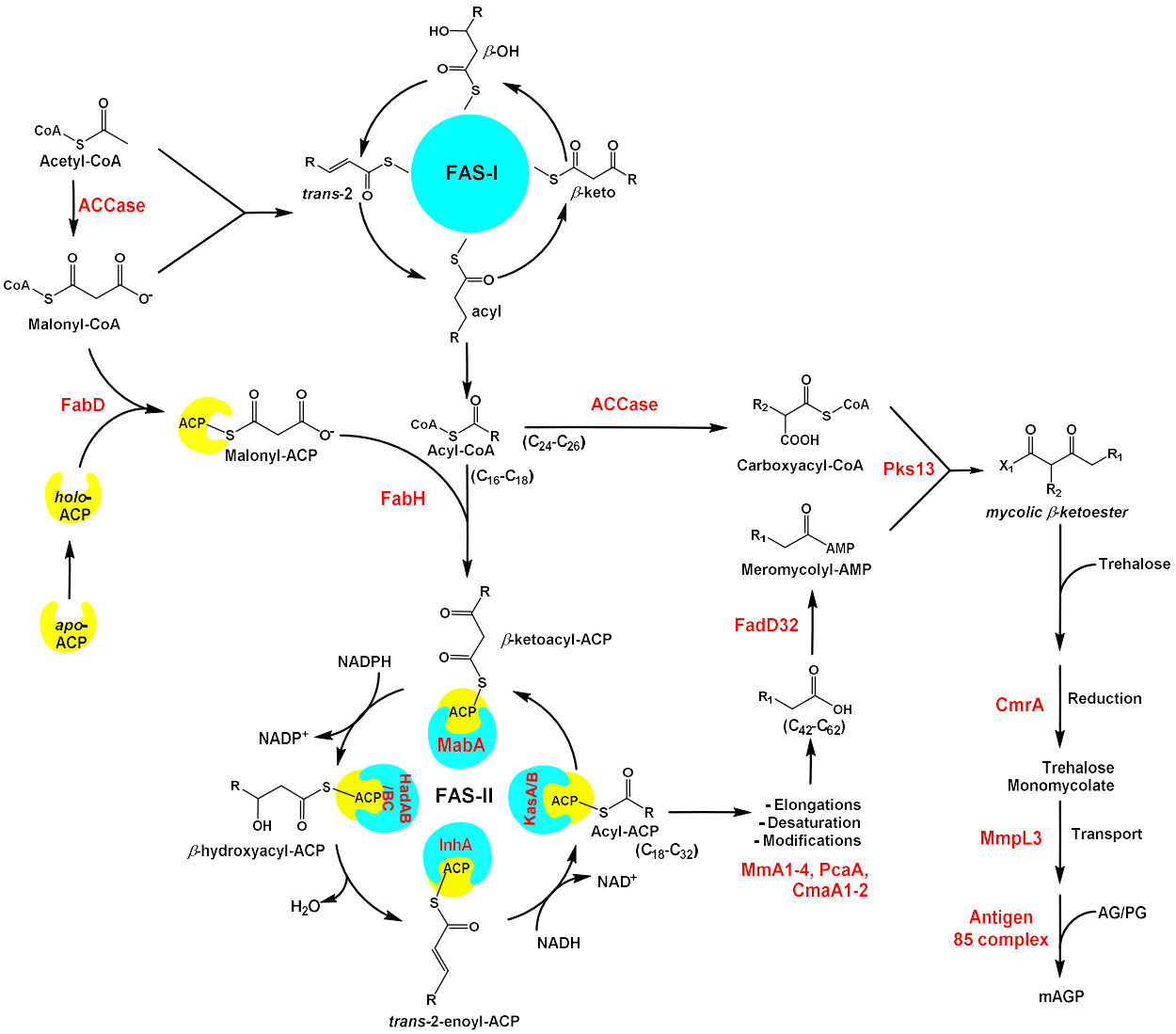

Figure 1.13: Synthesis pathway of MAs in *M. tuberculosis*. Acetyl-CoA and Malonyl-CoA are fed into the FAS-I system for *de novo* fatty acid synthesis. The products of FAS-I are then shuttled into the FAS-II system by FabH for further elongation. The long chain meromycolate product from FAS-II is then modified, before a condensation reaction performed by Pks13 combines the shorter α-alkyl chain from FAS-I to the meromycolate chain to form an intermediatary mycolic β-ketoester (3-oxo-C_{78}-mycolate). CmrA finally reduces the β-ketoester to form a secondary alcohol and mature mycolate. The mycolate is then conjugated to trehalose to form trehalose monomycolate (TMM), before being transported out of the cell by MmpL3 and finally attached to the cell wall by the antigen 85 complex to complete the mAGP complex.

Meromycolate chains become modified through the addition of double bonds, a feature of all MAs across each MA containing species (Marrakchi *et al.*, 2014). However, the mechanism in which *M. tuberculosis* introduces double bonds into its MAs remains elusive. In *Streptococcus pneumoniae* the enzymes FabM (a *trans*-2-*cis*-3-enoyl-ACP isomerase) along with FabZ (a dehydratase) were identified as being responsible for shunting fatty acid

synthesis towards an unsaturated state (Marrakchi *et al*., 2014). Two potential homologs of FabM were identified in *M. tuberculosis* as EchA10 and EchA11, which are proposed to perform the isomerisation of the *trans*-2 intermediate of the FAS-II system to the hypothetical *cis*-3-enoyl fatty acid, ultimately leading to a *cis*-double bond at the distal end of the meromycolate chain after further elongation (Takayama *et al*., 2005, Marrakchi *et al*., 2014). However, the fact that both EchA10 and EchA11 are both predicted to be non-essential enzymes casts doubt as their role as isomerases, since MAs are essential, the proteins that provide isomerase activity should also be essential, leaving the identity of a true mycobacterial FabM homolog currently unknown (Griffin *et al*., 2011, Marrakchi *et al*., 2014). An alternative method of introducing double bonds would be through aerobic desaturation of the meromycolate chain during synthesis, by an oxidative desaturase utilising both molecular oxygen and NADH (Marrakchi *et al*., 2014). Three potential desaturases have been identified within the *M. tuberculosis* genome, encoded by *desA1*, *desA2* and *desA3* respectively (Cole *et al*., 1998, Takayama *et al*., 2005). Through overexpression and biochemical analysis, DesA3 has been shown to be responsible for the synthesis of the unsaturated fatty acid oleic acid in *M. bovis* BCG, proving its capability of insertion of double bonds to long carbon chains (Phetsuksiri *et al*., 2003). However, no function has yet been elucidated for either DesA1 or DesA2 (Marrakchi *et al*., 2014).

The meromycolate chains become further modified by the action of a set of *S*-adenosyl-methionine-dependant methyltransferases (SAM-MTs). CmaA2, MmaA2 and PcaA all have roles in introducing cyclopropanation (*cis*- and *trans*-) at both the distal and proximal ends of the chain (George *et al*., 1995, Glickman *et al*., 2000, Glickman, 2003) (Table 1.3). MmaA3 and MmaA4 are SAM-MTs that are responsible for both forming methyl and hydroxyl branches and methoxylation of these to form methoxy- and keto-mycolate derivatives (Boissier *et al*., 2006). Once the meromycolate has matured, a condensation reaction

involving the type I polyketide synthase, Pks13, and a specific fatty-acyl-AMP ligase, FadD32, completes the synthesis of MAs. FadD32 converts the meromycoyl-S-AcpM into meromycoyl-AMP, whilst the C_{26}-acyl-CoA from FAS-I is carboxylated by Acyl-CoA carboxylase (ACCase) into 2-carboxyl-C_{26}-CoA (Figure 1.13) (Gago *et al*., 2006). Pks13 then binds both components of MA at its N terminus, and then transfers and condenses functional groups of the mycolate components within different domains of the enzyme *via* nucleophilic attack. The product of these reactions yields the intermediate 3-oxo-C_{78}-mycolate and CO_2. The intermediate is finally reduced by CmrA, into a secondary alcohol and thus the completed C_{78}-mycolate (Figure 1.13) (Bhatt et al., 2007, Lea-Smith *et al*., 2007). Completed MAs are conjugated to trehalose to form trehalose monomycolate (TMM) in a yet to be elucidated pathway. TMM is then transported across the plasma membrane by MmpL3 (Grzegorzewicz *et al*., 2012). Finally, MAs from TMM are attached to the AG at the Ara*f*$_6$ non-reducing termini, by the fibronectin-binding proteins (Fbp) (known as the antigen 85 complex), antigens 85A, 85B and 85C, thus completing the mAGP complex (Jackson *et al*., 1999, Puech *et al*., 2000).

Table 1.3: List of genes, and their function, involved in the synthesis of Mycolic Acids (MA).

Gene name	Rv number	Role	Reference
fas	Rv2524c	Complex of all components of FAS-I pathway of mycolic acid synthesis	(Smith *et al*., 2003)
fabD	Rv2243	Transacylation of malonate from malonyl-CoA to holo-AcpM to form malonyl-AcpM	(Kremer *et al*., 2001b)
fabH	Rv0533c	Condensation of $C_{12/14}$-Acyl-CoA with Malonyl-ACP to initiate FAS-II pathway	(Bhatt *et al*., 2007)
kasA	Rv2245	Initial addition of carbon to acyl-CoA primers in FAS-II pathway	(Bhatt *et al*., 2007)
kasB	Rv2246	Addition of carbon to longer acyl-CoA chains in FAS-II pathway	(Bhatt *et al*., 2007)
mabA	Rv1483	Reduction of β-ketoacyl-AcpM to β-hydroxyacyl-AcpM in FAS-II	(Takayama *et al*., 2005)
hadAB	Rv0635 Rv0636	Dehydration of β-hydroxyacyl-AcpM to enoyl-AcpM in FAS-II	(Marrakchi *et al*., 2014)
hadBC	Rv0636 Rv0637	Dehydration of β-hydroxyacyl-AcpM to enoyl-AcpM in FAS-II	(Marrakchi *et al*., 2014)
inhA	Rv1484	Reduction of enoyl-AcpM to acyl-AcpM in FAS-II	(Marrakchi *et al*., 2014)
fadD32	Rv3801c	Converts meromycolyl-S-AcpM to meromycolyl-AMP	(Bhatt *et al*., 2007)
pks13	Rv3800c	Condensation of 2-carboxyl-C_{26}-CoA and meromycolyl-AMP	(Bhatt *et al*., 2007)
cmrA	Rv2509	Reduction of 3-oxo-C_{78}-mycolate into C_{78}-mycolate	(Lea-Smith *et al*., 2007)
cmaA2	Rv0503c	Introduction of *trans*-cyclopropanation to MAs	(George *et al*., 1995)
mmaA2	Rv0644c	Introduction of distal cyclopropanation of α-MAs	(Glickman, 2003)
pcaA	Rv0470c	Introduces proximal cyclopropanation of α-MAs	(Glickman *et al*., 2000)
mmaA3	Rv0643c	Methoxylation of methyl branches of mycolate chains	(Boissier *et al*., 2006)
mmaA4	Rv0642c	Introduction of methyl and adjacent hydroxyl branches to mycolate chains	(Boissier *et al*., 2006)
mmpL3	Rv0206c	Transports MAs (as TMM) across the plasma membrane	(Grzegorzewicz *et al*., 2012)
fbpA/ a 85A	Rv3804c	Attachment of MAs to non-reducing termini of AG	(Jackson *et al*., 1999)
fbpB/ a 85B	Rv1886c		
fbpC/ a 85C	Rv0129c		

1.3. Aims and Objectives

Our understanding of crucial aspects of mycobacterial cell wall biosynthesis remains fragmented; particularly those associated with the regulation of the cytoplasmic assembly of peptidoglycan precursors and how AG is covalently attached to PG.

This thesis can be split into two major objectives:

1. Bioinformatics and biochemical analysis of pseudo-phosphorylated mutants of *M. tuberculosis* MurC, the first of the ATP-dependent Mur ligases involved in PG biosynthesis. This will assess whether phosphorylation of this enzyme has any impact upon the activity of the protein and therefore PG biosynthesis as a whole in *M. tuberculosis*.

2. Bioinformatic analysis to identify a putative phosphotransferase in *M. tuberculosis* responsible for attachment of AG to PG. We will adopt a molecular genetic approach to determine the essentiality of this phosphotransferase by generating a conditional deletion of its orthologue in *M. smegmatis*. We will also perform a comprehensive biochemical analysis on recombinant protein.

Chapter 2

Investigation into phosphoregulation of MurC ligase

2. INVESTIGATION INTO PHOSPHOREGULATION OF MURC LIGASE

2.1. Introduction

The Mur ligase pathway is a vital component in the biosynthesis of PG, responsible for the generation of the final PG 'monomer unit', "Lipid II", within the cytoplasm before its translocation across the plasma membrane and addition into the growing PG polymer (Figure 1.9) (Section 1.2.2.2.). ATP-dependent Mur ligases are a division of these proteins which are responsible for the addition of amino acid residues to UDP-MurNAc or UDP-MurNGlyc (in *M. tuberculosis*) to form a pentapeptide side chain, which later becomes cross-linked on the outside of the cell (Chang *et al*., 1990, Smith, 2006). These enzymes consist of MurC which catalyses the addition of the first residue L-alanine (L-ala), MurD which then adds D-glutamine (D-glu), MurE then attaches *meso*-diaminopimelate (*m*-DAP) and MurF which completes the side chain with D-alanine-D-alanine (D-ala-D-ala) attachment (Section 1.2.2.2.).

The ATP-dependent Mur ligases from *E. coli* exhibit similar structural features and have a high degree of sequence similarity. Each of these four enzymes consist of three different domains, each responsible for binding one of the three substrates involved in catalysis. The N-terminal domains (domain 1) show similarity to known dinucleotide binding Rossmann-fold like domains, which differ slightly in the orientation of β-sheets between MurC/D (mixed) and MurE/F (parallel) (Figure 2.1). Structural studies of these proteins have shown that domain 1 binds to the UDP-MurNAc moiety for each of the substrates utilised in this pathway (Smith, 2006). The central domain (domain 2) is consistent with ATP and GTP binding domains and the C-terminal domain (domain 3) has been shown to bind to the incoming amino acid residue to be attached to the growing peptide (Figure 2.1) (Smith, 2006, El Zoeiby *et al*., 2003). The active site for these enzymes is found at the junction between all 3 domains,

which is assisted by the ability of the protein to adjust the angle of its domains *via* hinge residues (such as phenylalanine-321 in MurC) (Smith, 2006). Because of the increasing chain length of the growing pentapeptide, MurE/F contains an extended loop within domain 1 that protrudes towards domain 3; which is absent in MurC/D (Figure 2.1). This loop interacts with UDP-MurNAc and is thought to provide additional space to accommodate the binding of a substrate with two or more additional amino acids (Smith, 2006).

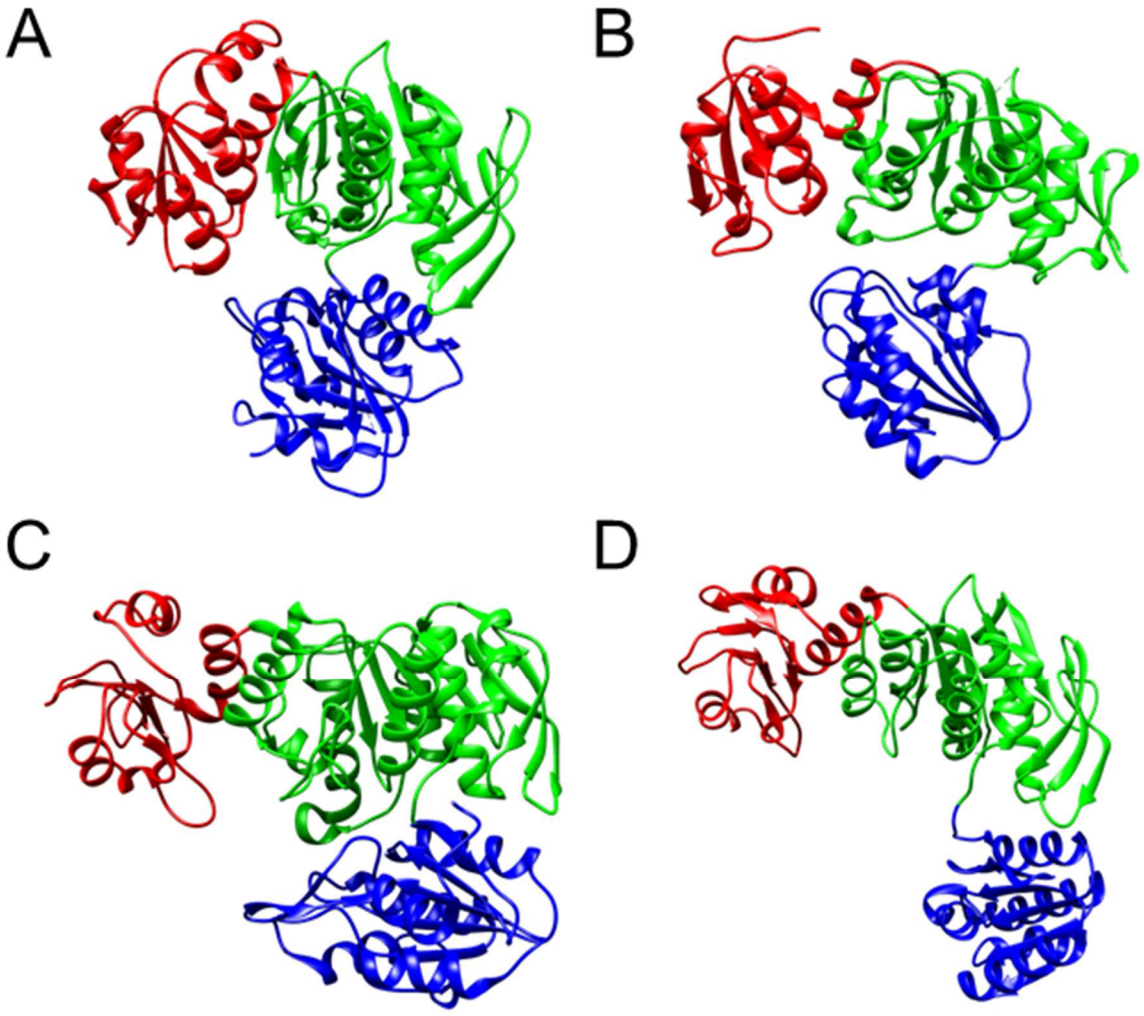

Figure 2.1: Comparison of *E. coli* ATP-dependent Mur ligase binding moieties. Domain 1 (red) is the UDP-MurNAc binding domain, domain 2 (green) is the ATP binding domain and domain 3 (blue) is the amino acid residue binding domain. A) MurC (PDB: 2F00) B) MurD (PDB: 2UAG) C) MurE (PDB: 1E8C) and D) MurF (PDB: 1GG4). All proteins share a similar overall structure, increasing in width as the size of the substrate increases throughout the pathway.

All of the ATP-dependent Mur ligases are believed to utilise the same mechanism to catalyse the addition of amino acid residues onto the growing peptide chain. Each of these enzymes sequentially bind substrates, beginning with the binding of ATP, followed by UDP-MurNAc/NGlyc (with any attached peptides) and finally the amino acid residue to be added, each binding to their specific sites within the protein (Figure 2.2) (Eveland *et al*., 1997). From here, the γ-phosphate group from ATP becomes covalently attached to the terminal carboxylate group of the UDP-Muramate (or equivalent carboxylate group of the attached amino acid residue), liberating ADP and creating an activated acyl-phosphate intermediate. Interestingly, both enzymatic and structural characterisation of these enzymes has implicated Mg^{2+} as playing an important role in stabilising the negative charges within the active site (Smith, 2006). Nucleophilic attack by the incoming amino acid then replaces the phosphate group, producing a peptide bond and therefore releasing both inorganic phosphate and the enzymatic product of the Mur ligase (Figure 2.2) (Falk *et al*., 1996, Bertrand *et al*., 1999). In the presence of high concentrations of ADP, it is possible for these ATP-dependent Mur ligases (such as MurC) to perform the reverse reaction and remove the amino acid residue from the UDP-MurNAc/NGlyc-peptide molecule (Liger *et al*., 1995).

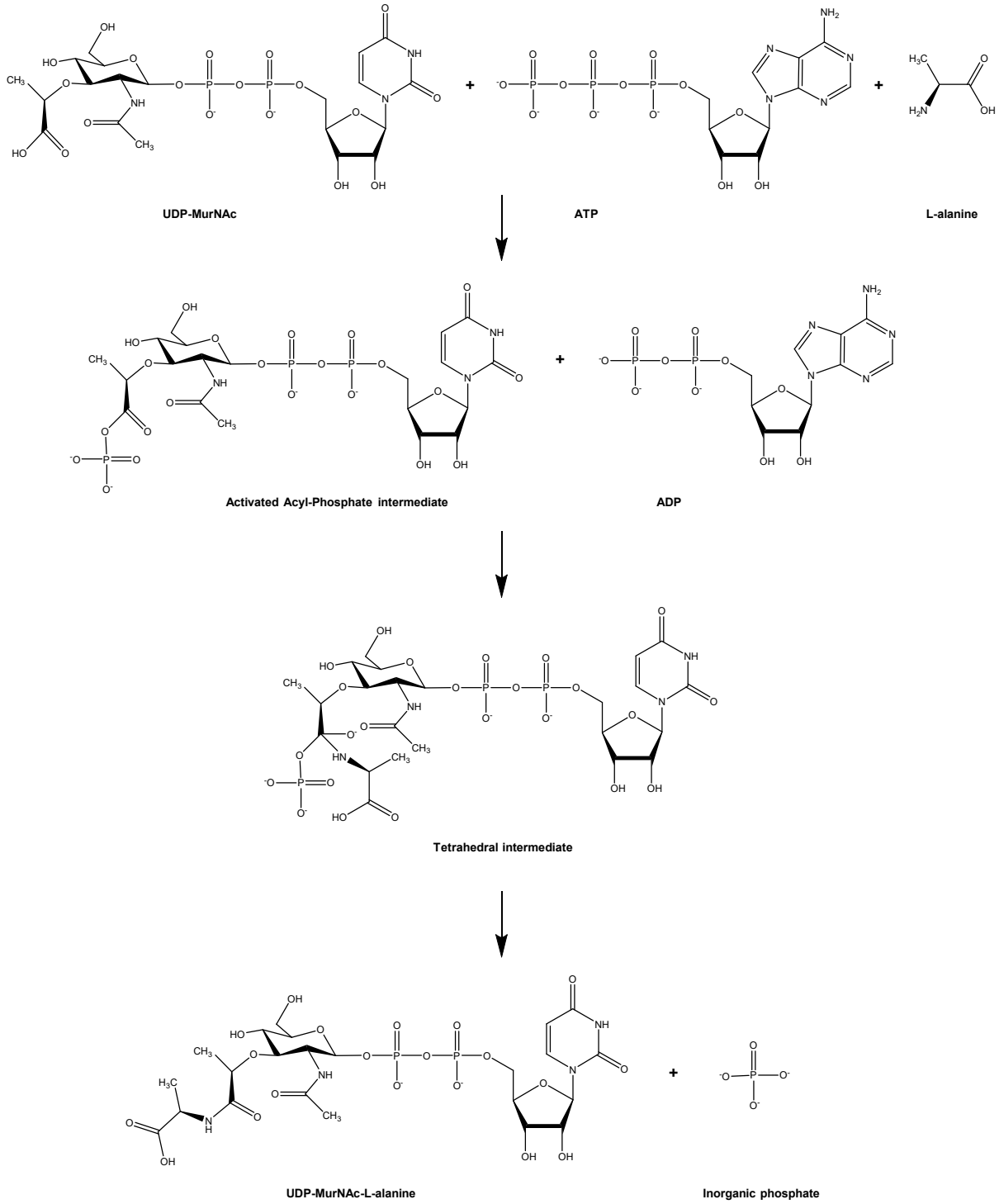

Figure 2.2: Mechanism of action of ATP-dependent MurC ligase. Substrates are bound in sequential order of ATP, then UDP-MurNAc (or UDP-MurNGlyc) and finally L-alanine. The γ-phosphate group from ATP then binds to terminal carboxylate group of UDP-MurNAc, creating an activated acyl-phosphate intermediate. Nucleophilic attack by the incoming L-alanine forms a tetrahedral intermediate before it replaces the phosphate group, producing a peptide bond and therefore creating UDP-MurNAc-L-alanine. Inorganic phosphate and ADP are released as by-products of the reaction.

Mycobacterial ATP-dependent Mur ligases have the ability to utilise UDP-MurNGlyc, as well as UDP-MurNAc, therefore integrating the hydroxylated backbone molecule into the structure of PG (Raymond *et al*., 2005). The first of these enzymes, MurC, has a preference for L-ala as its natural substrate, but is capable of utilising glycine (Gly) as well as L-serine (L-ser) *in vitro*, in the same way as MurC from other bacteria such as *E. coli*. However, the affinity of *M. tuberculosis* MurC for both Gly and L-ser is much lower than that for L-ala, suggesting the reason that L-ala is the preferred substrate (Munshi *et al*., 2013). The X-ray crystal structure of *M. tuberculosis* MurC is yet to be elucidated, however, homologous structures have been published from other bacteria such as *Haemophilus influenzae* and *E. coli*, with both structures consisting of the usual three domains as seen in other ATP-dependent Mur ligases (Mol *et al*., 2003, Deva *et al*., 2006). In *E. coli*, it was seen that MurC crystallises as a dimer, yet exists in solution in a dynamic equilibrium of both monomeric and dimeric forms, both of which show activity (Deva *et al*., 2006, Das *et al*., 2011).

The fact that MurC is at the start of the sequential ATP-dependent Mur ligase biosynthetic pathway, as well as that there is no eukaryotic equivalent, makes it a very attractive drug target. A series of compounds known as phosphinate transition-state analogues were tested as inhibitors against MurC (from *E. coli*), with the most potent compound providing an IC_{50} value of 49 μM (Reck *et al*., 2001). Unfortunately *in vivo* testing of these phosphinate compounds indicate that they have little antibacterial activity, presumably due to the inability of these compounds to cross the cytoplasmic membrane in order to reach their target destination (El Zoeiby *et al*., 2003).

In *C. glutamicum*, it has been observed that MurC is phosphorylated by a specific serine/threonine kinase called PknA. Radiochemical analysis indicated that Cg-MurC is intermittently phosphorylated at several locations which include Thr-51, Thr-120, Thr-133, Thr-167, Thr-362 and Thr-365 (Fiuza *et al*., 2008). This study also demonstrated that MurC

phosphorylation correlates with regulation of enzymatic activity using a radiochemical endpoint assay (Fiuza *et al.*, 2008). The addition of phosphate onto a protein can have a wide variety of effects (including up and down-regulation of activity). It can also have an effect in signalling pathways and protein degradation (Stock *et al*., 1989, Burnett and Kennedy, 1954). Phosphorylation is a useful mechanism of control by cells, as the addition of phosphate onto a protein residue, through phosphokinase action, can easily be reversed by removal of the phosphate by phosphatases (Burnett and Kennedy, 1954). This allows cells to maintain control over pathways that need to be tightly regulated, in order to prevent damage to the cell. This is important as it means that the cell has a means to regulate essential pathways such as PG biosynthesis. However, research into PknA activity on *M. tuberculosis* enzymes showed only MurD as a phosphorylation target through *in vitro* binding assays (Thakur and Chakraborti, 2008). Conversely, it has recently been shown by protein-protein interaction studies that both PknA and a different kinase, PknB, have the ability to interact with all of the ATP-dependent Mur ligases in *M. tuberculosis*, to varying degrees of intensity (Munshi *et al*., 2013). In contrast, *Streptococcus pneumoniae* uses StkP to phosphorylate MurC, which is a multiple PASTA-domain containing phosphokinase. It has been suggested that there are multiple Ser/Thr sites within the *C. glutamicum* and *M. tuberculosis* MurC/D and so these species use PknA in order to phosphorylate the sites on all 3 domains (Falk and Weisblum, 2013).

PknA (as well as PknB) belongs to a family of 11 proteins in *M. tuberculosis*, known as Serine/Threonine Protein Kinases (STPKs), which phosphorylate specific serine and threonine residues on various proteins within the cell (Prisic and Husson, 2014). The *M. tuberculosis* phosphoproteome is vast and STPKs have been found to be responsible for at least 500 phosphorylation events across 301 proteins, involved in a variety of different activities within the cell (Prisic *et al*., 2010). Of the 11 STPKs in *M. tuberculosis*, PknG and

PknK are two proteins that are distinct from the other nine as they do not contain a transmembrane domain, suggesting that these genes were acquired elsewhere than those transmembrane spanning STPKs, which likely all derived from a single common ancestral gene (Prisic and Husson, 2014). However, the extracellular sensor domains of the nine receptor-type STPKs share no sequence homology, indicating that they bind to distinctive signal molecules to each other (Prisic and Husson, 2014). Yet six of these STPKs (PknA,B,D,E,F and H), do share a similar core phosphorylation motif to each other found within an intracellular N-terminal kinase domain, with differences arising in proteins that require higher substrate specificity (Prisic and Husson, 2014). PknA and PknB are both encoded by adjacent genes within an operon that contains numerous cell wall synthetic gene products, such as PBPA (also a probable substrate of PknB) and RodA, which are involved in cell shape and division. Both PknA and PknB have been shown as essential in mycobacteria through transposon mutagenesis, with this essentially and their location within the genome suggesting that PknA and PknB have an important role in the regulation of cell growth, shape and cell wall biosynthesis (Sassetti *et al*., 2003, Prisic and Husson, 2014). Although PknB has been suggested to activate multiple STPKs, it has recently been shown that PknA is activated independently of PknB *in vivo* and its activation loop is likely to be autophosphorylated by PknA itself, indicating that PknA is not reliant on PknB activity (Baer *et al*., 2014, Nagarajan *et al*., 2015).

In this study, we have identified several homologous candidate residues for phosphorylation in *M. tuberculosis* MurC, based upon phosphorylation sites from the *C. glutamicum* protein, through extensive bioinformatic analysis. We therefore created pseudo-phosphorylated MurC mutants and biochemically analysed the effect of creating a negatively-charged region at these residues when compared to wild-type.

2.2. Results

2.2.1. *In silico* analysis of MurC and identification of candidate phosphorylation sites

A multiple protein sequence alignment (Section 5.17.) was used to compare MurC protein sequences between various bacteria including mycobacterial, Gram positive and Gram negative species (Figure 2.3). The varying mycobacterial MurC proteins show at least around 80 % sequence identity to *M. tuberculosis* MurC, whereas *C. glutamicum* MurC shows around 52 % identity (Figure 2.3, Table 2.1). The MurC sequences from *Pseudomonas aeruginosa*, *S. aureus* and *E. coli* all share around 30 % sequence identity to *M. tuberculosis* MurC. A second protein sequence alignment was performed between MurC sequences from *M. tuberculosis* and *C. glutamicum* in order to identify homologous residues between the two (Figure 2.4). Based upon reported phosphorylated residues in *C. glutamicum* MurC by Fiuza *et al*., 2008, five homologous residues were identified in *M. tuberculosis* MurC. These residues selected as possible phosphorylation sites were Thr-116, Thr-129, Ser-163, Thr-357 and Ser-360 (Figure 2.4). There was no homologous residue to Thr-51 (from *C. glutamicum* MurC (Fiuza *et al*., 2008)) in *M. tuberculosis* MurC, since the corresponding residue is Gly-47, which is unable to be phosphorylated by STPKs.

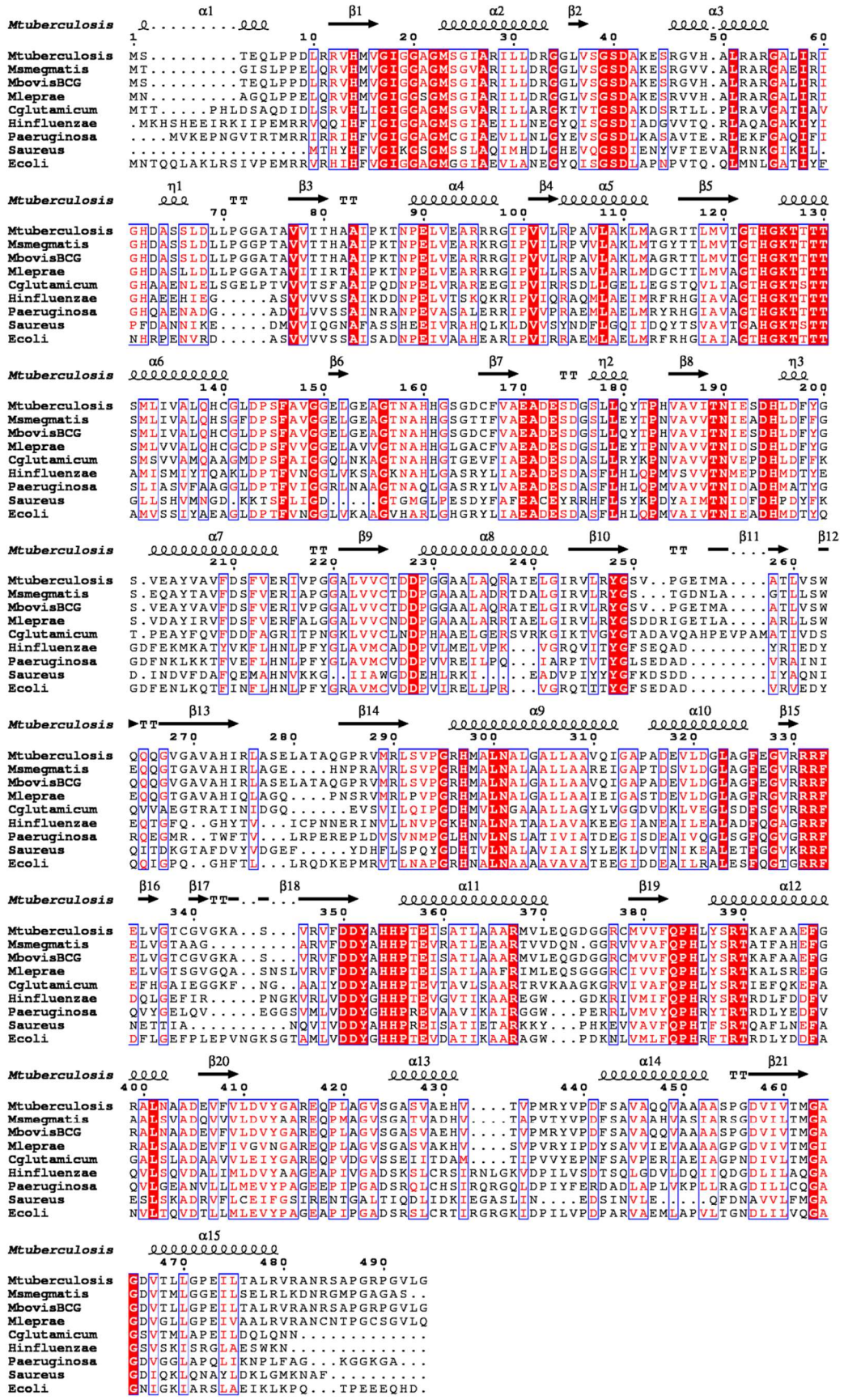

Figure 2.3: Protein sequence alignment of MurC from different bacterial species. *M. tuberculosis, M. smegmatis, M.bovis* BCG, *Mycobacterium. leprae, C. glutamicum, H. influenzae, Psuedomonas aeruginosa, Staphylococcus aureus* and *E. coli* display a high level of homology between sequences, indicated by residues in red. Residues highlighted in red indicate identity across all species aligned, showing conserved residues. Secondary structure information is seen above all sequences.

Table 2.1: Protein sequence identity of *M. tuberculosis* MurC with MurC from multiple different species. Highest identity is seen between *M. tuberculosis* and *M. bovis* BCG, as they have the exact same sequence (100 % identity), whereas the lowest identity is seen when compared to *S. aureus* (30.07 %). However, reasonably high identity is seen between all the mycobacterial species and *C. glutamicum*. There is still reasonable identity between the Gram positive and negative species, due to conserved residues between all species.

Species	Identity (%)
M. smegmatis	79.25
M. bovis BCG	100
M. leprae	80.61
C. glutamicum	52
H. influenzae	37.5
P. aeruginosa	36.93
S. aureus	30.07
E. coli	35.74

Figure 2.4: Protein sequence alignment of MurC from both *M. tuberculosis* and *C. glutamicum*. Residues highlighted in red are conserved residues, whereas residues highlighted blue represent phosphorylation sites identified in *C. glutamicum* by Fiuza *et al*., 2008 and homologous residues identified in the *M. tuberculosis* protein. The identified residues in *M. tuberculosis* MurC were Thr116, Thr129, Ser163, Thr357 and Ser360.

The sequences for both *C. glutamicum* (Cg) and *M. tuberculosis* (Mtb) MurC proteins were submitted for 3D homology modelling using the I-TASSER server (with a C-score of 1.11 and 0.785 respectively) (Section 5.17.). These homology models were then analysed in comparison to each other and also to the *H. influenzae* MurC crystal structure model bound to a non-hydrolysable ATP analogue, AMPPNP (PDB ID: 13P1), in order to identify where ATP would bind to the homology models (Figure 2.5) (Section 5.17.). We made use of the reported phosphorylation sites of Cg-MurC to identify the equivalent (putative) phosphorylation sites in Mtb-MurC (Figure 2.4) (Fiuza *et al*., 2008). In addition, we mapped

these phosphorylation sites into the I-TASSER homology models to investigate the surface exposure of the amino acids that ultimately become phosphorylated (Figure 2.5A,B Bottom). Structural alignment of both the Cg-MurC and Mtb-MurC homology models demonstrate that four of the five putative Mtb-MurC phosphorylation sites appear on the surface of the protein, with one (T129) being buried within the protein (Figure 2.5A,B Bottom).

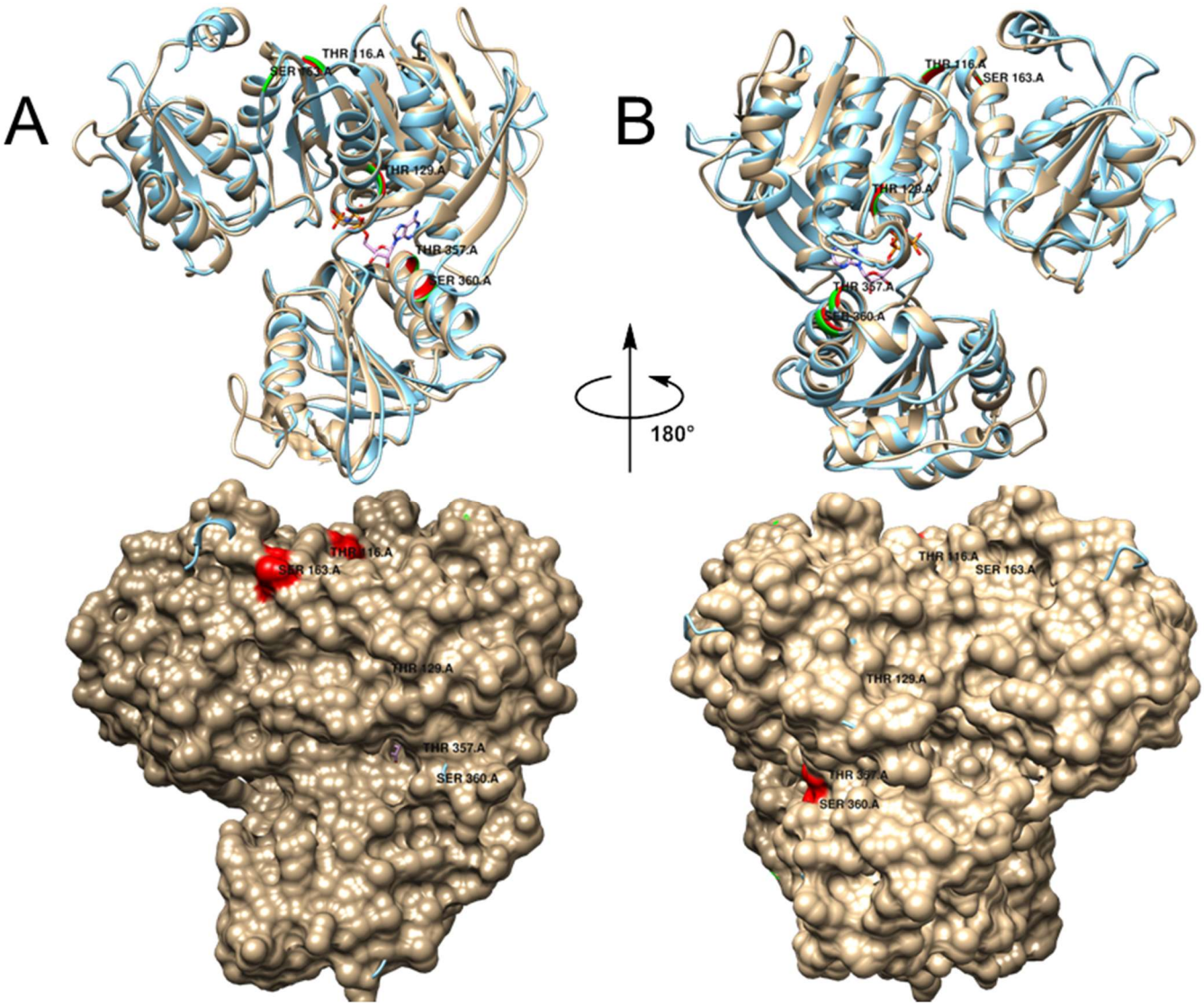

Figure 2.5: An overlay of I-TASSER generated homology models of Cg-MurC (blue) and Mtb-MurC (brown) showing location of ATP (AMPPNP) bound. Phosphorylated residues are highlighted green in the Cg-MurC model and selected residues are highlighted red in the Mtb-MurC model. A) Top – shows 'front' image of ribbon overlay of Cg and Mtb models, the highlighted residues overlay nearly identically and all show a close spatial proximity to the bound ATP analogue. Bottom – a spacefilling model of the 'front' of Mtb MurC showing surface selected residues in red. B) Top – shows 'back' image (180°) of ribbon overlay of Cg and Mtb models. Bottom – 'back' image (180°) of spacefilling model of Mtb MurC, showing selected surface residues in red. The only residue that is not found on the surface of the protein is T129.

The close spatial proximity between the selected phosphorylation residues in both the Cg-MurC and Mtb-MurC homology models and the bound ATP analogue suggest that alteration of these residues will have an effect on ATP binding by the enzyme (Figure 2.5A,B, Top). Both T116 and S163 are situated at the top of neighbouring loops that extend down toward the ATP binding site (Figure 2.5A). T116 is located at the start of β-strand 5, whereas S163 is found just before the start of β-strand 7, both of which form a part of a β-sheet (Figure 2.4). T129 is positioned within the middle portion of α-helix 6, which is embedded within the protein at the base of the ATP binding moiety (Figure 2.4, Figure 2.5). Both T357 and S360 residues are found within α-helix 11, which itself is actually situated within the L-ala binding moiety of Mtb-MurC (Figure 2.4). However, both residues are orientated on the outside edge of the α-helix which faces the ATP binding site, suggesting interaction with this substrate (Figure 2.5).

2.2.2. Purification of recombinant MurC protein

An N-terminally His-tagged *M. tuberculosis* MurC expression vector (pET28b-MurC) was constructed (Section 5.18.1.) and the selected homologous phosphorylation site residues (Section 2.2.1.) were mutated into aspartate through site-directed mutagenesis (SDM) (Section 5.6.). This gave rise to the constructs pET28b-MurC T116D, T129D, S163D, T357D and S360D alongside the MurC WT construct. All of these vectors were overexpressed in *E. coli* BL21 (DE3) (Section 5.4.1.) and purified through immobilised metal ion affinity chromatography (IMAC) (Figure 2.6) and subsequent anion exchange chromatography, if required (Section 5.11.).

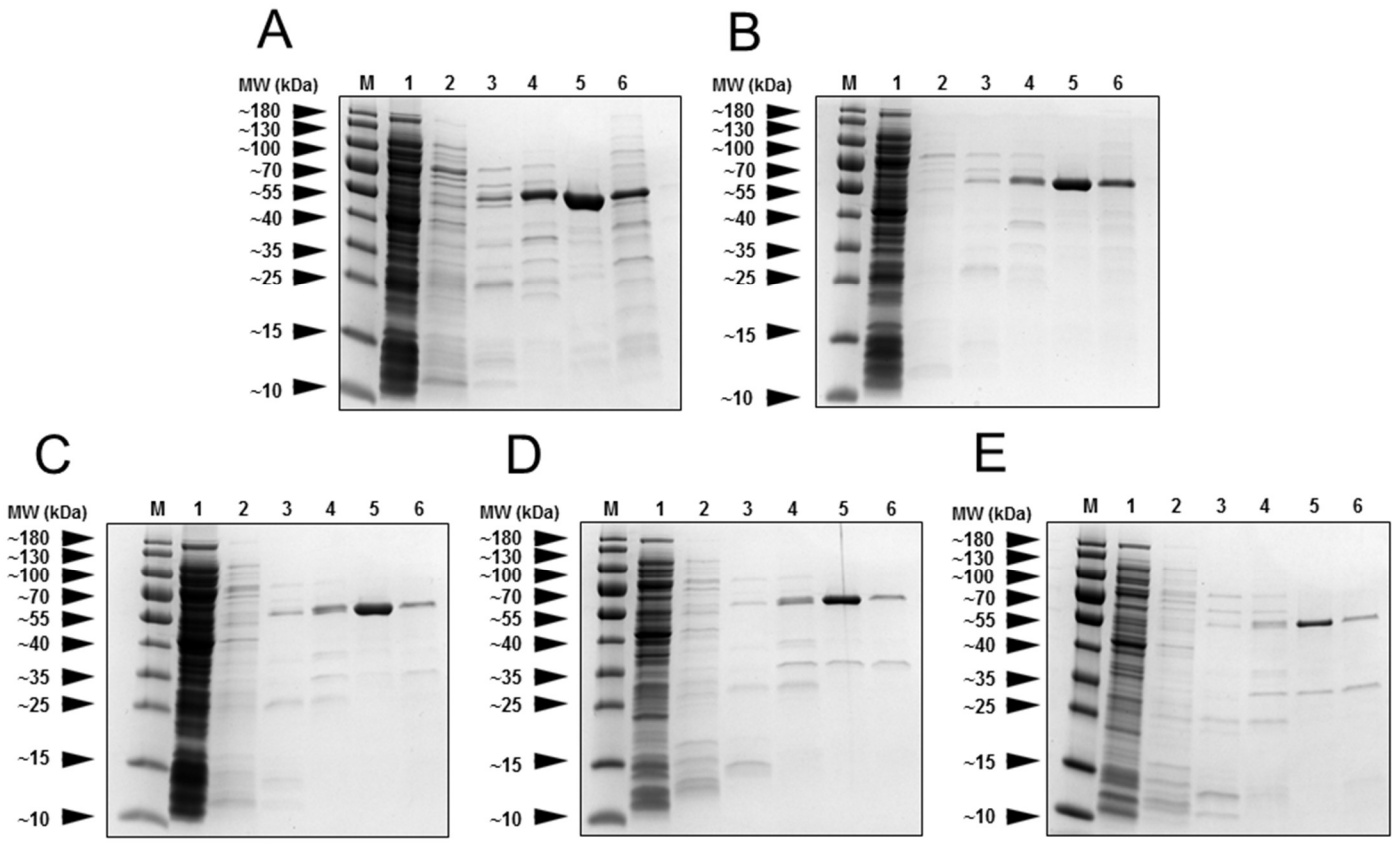

Figure 2.6: SDS-PAGE of MurC proteins. A) MurC WT, B) MurC T116D, C) MurC S163D, D) MurC T357D E) MurC S360D. MurC T129D was unable to be expressed or purified despite numerous attempts of optimisation. Lanes shown are protein marker (M), clarified lysate flow through (1), 50 mM imidazole elution (2), 75 mM imidazole elution (3), 125 mM imidazole elution (4), 150 mM imidazole elution (5) and 500 mM imidazole elution (6). Bands visible around 50-60 kDa in lane 5 represent recombinant MurC.

MurC WT as well as T116D, S163D, T357D and S360D pseudo-phosphorylated site-directed mutants all purified well and were stable in MurC 'storage' buffer (Section 5.1.11.) at concentrations up to 2.3 mg/mL. However, despite numerous attempts at optimising expression and purification, it was not possible to obtain recombinant MurC T129D, despite cells appearing phenotypically healthy.

2.2.3. Kinetic analysis of MurC pseudo-phosphorylation

In order to biochemically characterise each MurC pseudo-phosphorylated mutant, a kinetic assay was developed to assess the rate of MurC action (Section 5.16.). An enzyme-coupled assay was used, whereby the rate of MurC activity was directly linked to the action of the enzymes pyruvate kinase and lactate dehydrogenase. This was achieved through the utilisation of the ADP by-product of the reaction catalysed by MurC (Figure 2.2), being fed into a secondary reaction scheme which converts phosphoenolpyruvate (PEP) into pyruvate (catalysed by pyruvate kinase and the turnover of ADP to ATP), which in turn becomes converted into lactate (by lactate dehydrogenase) and the oxidation of NADH to NAD^+ (Figure 2.7). This turnover of ADP to ATP and NADH to NAD^+ occurs at a 1:1 ratio, meaning the rate of NADH conversion to NAD^+ is directly correlated to the rate of ADP release from the MurC reaction. Furthermore, the oxidation of NADH to NAD^+ can be directly monitored by a continuous spectrophotometric analysis at 340 nm, as a decrease in absorbance and the rate of the reaction observed as Δ340 nm vs time (Jenkins, 1991) (Section 5.16.).

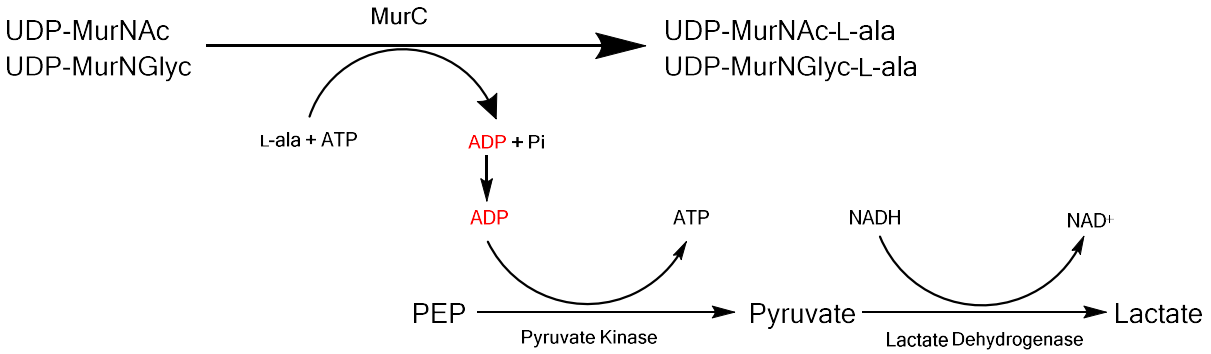

Figure 2.7: A reaction scheme for biochemical analysis of MurC pseudo-phosphorylated mutants. MurC utilises ATP to attach L-alanine to UDP-MurNAc/NGlyc and releases ADP and inorganic phosphate as by-products. The released ADP is then used by pyruvate kinase to convert PEP to pyruvate and releasing ATP. Pyruvate is then converted into lactate by lactate dehydrogenase through the oxidation of NADH to NAD^+. This reaction can be monitored by Δ340 nm.

We hypothesised that because *M. tuberculosis* MurC can utilise both UDP-MurNAc and UDP-MurNGlyc, as well as the fact that the pseudo-phosphorylation sites surround the ATP binding cleft (Figure 2.5), that there could be differences in MurC kinetics between substrates. Therefore we decided to assess the difference between the Michaelis-Menten kinetics of WT MurC and pseudo-phosphorylated MurC using UDP-Muramyl dependent as well as ATP dependent triplicate assays (Section 5.16.).

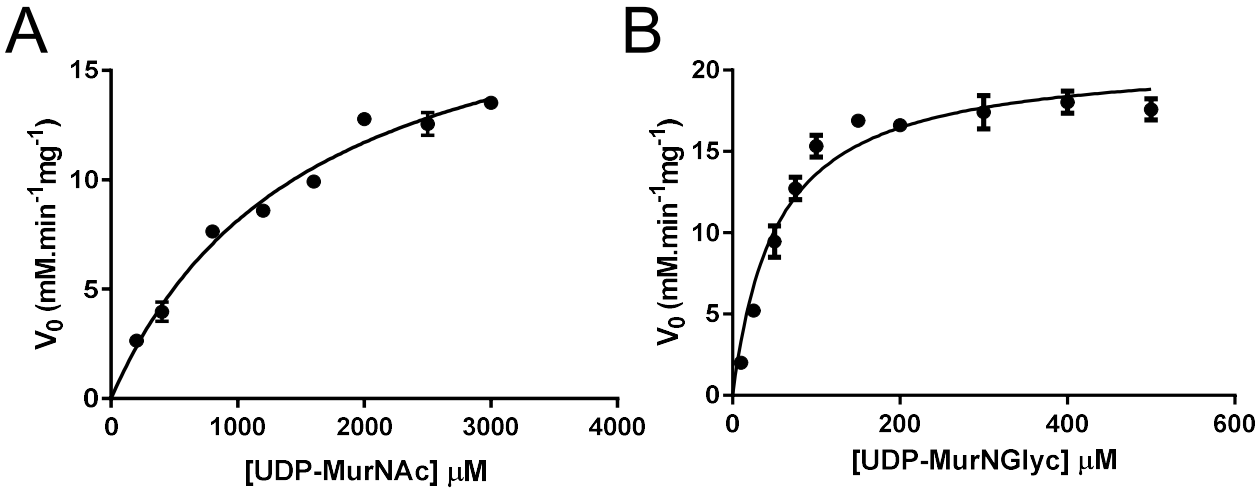

Figure 2.8: Michaelis-Menten graph of MurC WT with UDP-MurNAc (A) and UDP-MurNGlyc (B). The affinity of MurC WT for UDP-MurNGlyc was shown to be 30-fold more than for UDP-MurNAc, with K_m values of 52.98 ± 5.64 μM and 1,555 ± 272 μM respectively. R^2 values of A) 0.9661 and B) 0.9473.

The reaction velocity (V) of MurC was found to be hyperbolically dependent on both UDP-MurNAc and UDP-MurNGlyc at saturating levels of ATP and L-ala (Section 5.16.). This allowed for the Michaelis constant (K_m) to be deduced for each UDP-Muramyl substrate (using equation $V = \frac{V_{max} \times [Substrate]}{K_m + [Substrate]}$) with WT MurC, revealing a K_m for UDP-MurNAc of 1,555 ± 272 µM and a K_m of 52.98 ± 5.64 µM for UDP-MurNGlyc (Figure 2.8). Since there was a 30-fold difference in K_m between the two substrates, we decided to assess the UDP-Muramyl dependence of all pseudo-phosphorylated mutants using UDP-MurNGlyc.

The assay was repeated using each pseudo-phosphorylated mutant of MurC with UDP-MurNGlyc and saturating ATP and L-ala, with each MurC mutant displaying similar hyperbolic dependence to UDP-Muramyl as WT MurC. The reaction velocities (V, $mM.min^{-1}$) of each mutant were converted into specific activity of the enzymes ($mM.min^{-1}mg^{-1}$) to normalise the assays and allow for direct comparison between each protein (using equation $SA = \frac{V_{max}\ (mM.min^{-1})}{mg\ protein}$) (Figure 2.9). The turnover number (K_{cat}) and the specificity constant (K_{cat}/K_m) were calculated for each MurC mutant (using equation $K_{cat} = \frac{V_{max}\ (M)}{[Protein](M)}$) (Table 2.2).

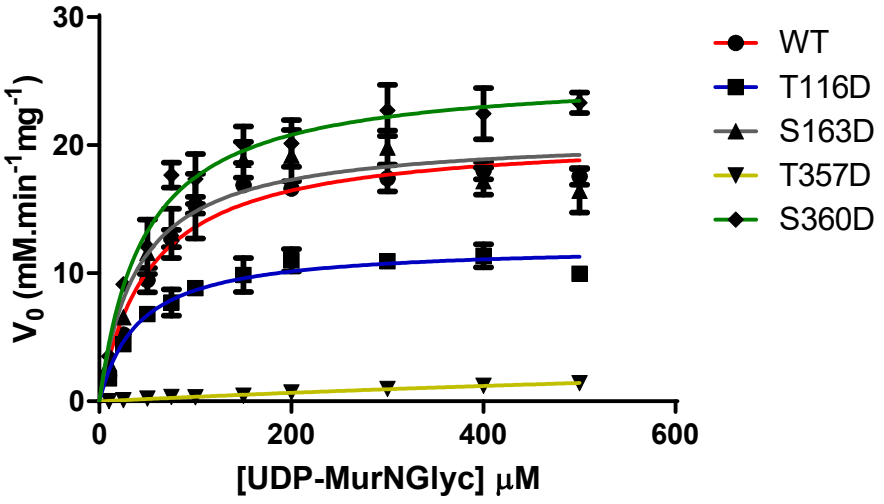

Figure 2.9: Michaelis-Menten graph displaying specific activities ($mM.min^{-1}mg^{-1}$) of MurC WT, T116D, S163D, T357D and S360D for UDP-MurNGlyc dependence. The K_m for each MurC variant is shown in Table 2.2. R^2 values for each MurC variant are 0.9473, 0.9094, 0.8427, 0.9608 and 0.9251 respectively, indicating high goodness of fit.

Table 2.2: Kinetic values for MurC WT and pseudo-phosphorylated mutants with UDP-MurNGlyc dependence. All mutants (except for T357D) display a higher affinity for the substrate than WT, however the specific activity for T116D is around half that of the other MurC variants. T357D has much lower affinity and activity than all of the other variants.

MurC Variant	K_m (μM)	V_{max} ($mM.min^{-1}$)	Specific Activity ($mM.min^{-1}mg^{-1}$)	K_{cat} (min^{-1})	K_{cat}/K_m ($min^{-1}M^{-1}$)
WT	52.98 (±7.631)	0.0348 (±0.00135)	20.81	212.84	4.02×10^6
T116D	40.99 (±7.282)	0.0127 (±0.00055)	12.22	124.51	3.04×10^6
S163D	40.47 (±10.05)	0.0119 (±0.00072)	20.81	212.88	5.26×10^6
T357D	1728 (±989.4)	0.0376 (±0.01776)	6.38	67.47	3.91×10^4
S360D	46.78 (±7.598)	0.0148 (±0.00062)	25.67	263.11	5.62×10^6

All MurC pseudo-phosphorylation mutants (apart from T357D) display a higher affinity for UDP-MurNGlyc than WT MurC. T116D, S163D and S360D all show a K_m within the range of 40-47 μM compared to 52.98 ±7.631 μM for WT (Table 2.2). The K_m of T357D of 1728 ±989.4 μM is more comparable to the WT K_m for UDP-MurNAc of 1555 μM. However, the specific activities of the MurC mutants correlate with this trend of affinities, apart from MurC T116D, which shows a specific activity of 12.22 $mM.min^{-1}mg^{-1}$, around half that of MurC WT (20.81 $mM.min^{-1}mg^{-1}$) indicating that each mg of MurC T116D displays around half the rate of that of MurC WT (Figure 2.9, Table 2.2). MurC T357D has a much lower specific activity than MurC WT, at 6.38 $mM.min^{-1}mg^{-1}$, around a third lower (Table 2.2). However, MurC S360D shows a higher specific activity than MurC WT, of 25.67 $mM.min^{-1}mg^{-1}$, indicating an increase of around a fifth in activity (Table 2.2). This disparity between MurC WT, S163D and S360D to both T116D and T357D is seen even clearer when both the K_{cat} and K_{cat}/K_m is considered. The K_{cat} indicates the maximum amount of reactions that can be performed by the enzyme active site per minute. MurC WT has a K_{cat} of 212.84 min^{-1}, whereas T116D is around half, at 124.51 min^{-1}, indicating around half the number of reactions

being performed by T116D. Again, T357D (67.47 min^{-1}) is around a third of that of MurC WT, showing a much lower number of reactions performed per minute (Table 2.2). Again, MurC S360D displays a higher K_{cat} (263.11 min^{-1}) than MurC WT (Table 2.2). Finally, the K_{cat}/K_m shows the efficiency of the enzyme, with a higher specificity constant meaning an enzyme shows a higher preference for the particular substrate. Interestingly, both MurC S163D and S360D show a higher K_{cat}/K_m than MurC WT, indicating a higher preference for UDP-MurNGlyc with pseudo-phosphorylation of these residues (Table 2.2). However, as expected, MurC T116D shows a lower K_{cat}/K_m (3.04×10^6 $min^{-1}M^{-1}$) than WT (4.02×10^6 $min^{-1}M^{-1}$), suggesting MurC T116D is less efficient than MurC WT (Table 2.2). MurC T357D shows a much lower K_{cat}/K_m than MurC WT of 3.91×10^4 $min^{-1}M^{-1}$, which is around 100 times lower, showing that MurC T357D is highly inefficient with UDP-MurNGlyc when compared to MurC WT (Table 2.2).

The assay was then repeated in order to assess the kinetic relationship of MurC WT and pseudo-phosphorylated mutants to ATP. This was initially performed using saturating concentrations of UDP-MurNGlyc and L-ala (Section 5.16.). Each MurC variant displayed hyperbolic dependence to ATP, similar to that of UDP-MurNGlyc (Figure 2.10). The kinetic parameters were calculated as before (Table 2.3).

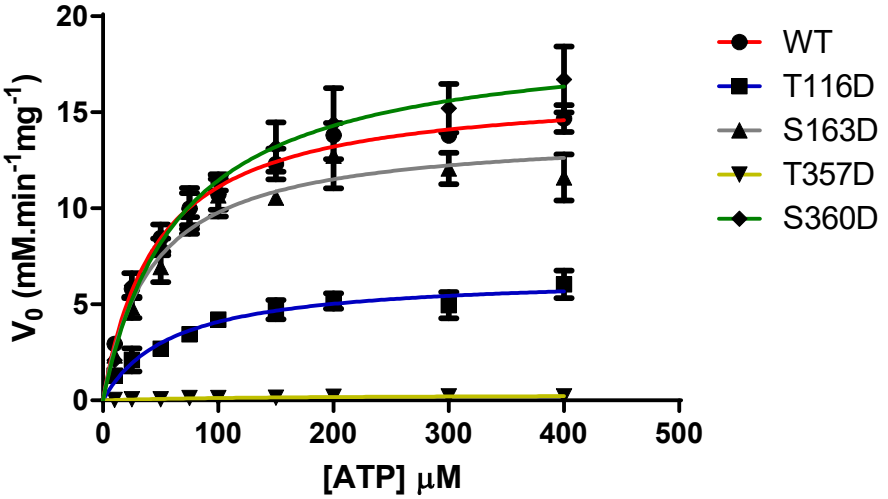

Figure 2.10: Michaelis-Menten graph displaying specific activities (mM.min^{-1}mg^{-1}) of MurC WT, T116D, S163D, T357D and S360D for ATP dependence in the presence of UDP-MurNGlyc. The K_m for each MurC variant is shown in Table 2.3. R^2 values for each MurC variant are 0.9391, 0.8814, 0.8939, 0.8924 and 0.8847 respectively, indicating high goodness of fit.

Table 2.3: Kinetic values for MurC WT and pseudo-phosphorylated mutants for ATP dependence in the presence of UDP-MurNGlyc. All MurC mutants show lower or similar affinity to ATP as WT, however they all show similar specific activity apart from both T116D and T357D, which are much lower than WT MurC. This is correlated by both K_{cat} and K_{cat}/K_m values.

MurC Variant	K_m (µM)	V_{max} (mM.min^{-1})	Specific Activity (mM.min^{-1}mg^{-1})	K_{cat} (min^{-1})	K_{cat}/K_m (min^{-1}M^{-1})
WT	46.80 (±5.70)	0.0272 (±0.00093)	16.30	166.36	3.56×10^6
T116D	56.76 (±13.29)	0.0068 (±0.00047)	6.53	66.67	1.18×10^6
S163D	43.09 (±9.03)	0.0080 (±0.00045)	13.99	143.11	3.32×10^6
T357D	142.6 (±35.29)	0.0018 (±0.00017)	0.30	3.23	2.27×10^4
S360D	66.42 (±12.51)	0.0110 (±0.00066)	19.05	195.56	2.94×10^6

All MurC mutants (except S163D, which is comparable to WT) show lower affinity for ATP in the presence of UDP-MurNGlyc than WT MurC (Table 2.3). T357D once again shows a much lower affinity (142.6 ±35.29 µM) than all other MurC variants, with a K_m around three times larger than that of WT. However, the specific activities of MurC WT, S163D and S360D are similar, ranging between 13.99-19.05 mM.min^{-1}mg^{-1}, whereas MurC T116D has a specific activity just under half of those including WT (6.53 mM.min^{-1}mg^{-1}). The specific activity of

MurC T357D is almost non-existent (0.30 $mM.min^{-1}mg^{-1}$), indicating very little activity by this protein (Figure 2.10, Table 2.3). Once again, MurC S360D displays the highest specific activity (19.05 $mM.min^{-1}mg^{-1}$), higher than MurC WT by around 15 % (Figure 2.10, Table 2.3). These figures are echoed by the K_{cat} values, which indicate that MurC T116D turns over just under half that of WT MurC (66.67 min^{-1} for T116D compared to 166.36 min^{-1} for WT), whereas T357D (3.23 min^{-1}) performs around 50 times less reactions per minute than WT MurC (Table 2.3). MurC S360D displays a higher K_{cat} than MurC WT (195.56 min^{-1}), performing more reactions per minute (Table 2.3). The efficiency of these enzymes shown by K_{cat}/K_m also corresponds with the rest of the kinetic figures for this assay. Once again, MurC WT and S163D show similar specificity constants to one another, however T116D (1.18×10^6 $min^{-1}M^{-1}$) is around three times lower than MurC WT (3.56×10^6 $min^{-1}M^{-1}$) (Table 2.3), indicating that MurC T116D shows lower preference to ATP in the presence of UDP-MurNGlyc than WT. MurC T357D shows even lower preference and very high inefficiency with ATP, since the K_{cat}/K_m for this protein is more than 150 times lower (at 2.27×10^4 $min^{-1}M^{-1}$) than MurC WT (Table 2.3). Interestingly, despite showing a higher activity and K_{cat} than MurC WT, MurC S360D actually displays a lower K_{cat}/K_m (2.94×10^6 $min^{-1}M^{-1}$), indicating that the protein is less efficient overall with ATP in the presence of UDP-MurNGlyc than MurC WT (Table 2.3).

The assay was repeated a final time to assess the kinetic relationship of MurC WT and pseudo-phosphorylated mutants to ATP dependence, this time in the presence of UDP-MurNAc and L-ala (Section 5.16.). However, due to a limited amount of UDP-MurNAc available for assay, as well as the high amount required per experiment, we were unable to assay MurC S163D with this substrate. Nonetheless, each assayed MurC variant once again displayed hyperbolic dependence to ATP (Figure 2.11) and the kinetic parameters were once again calculated (Table 2.4).

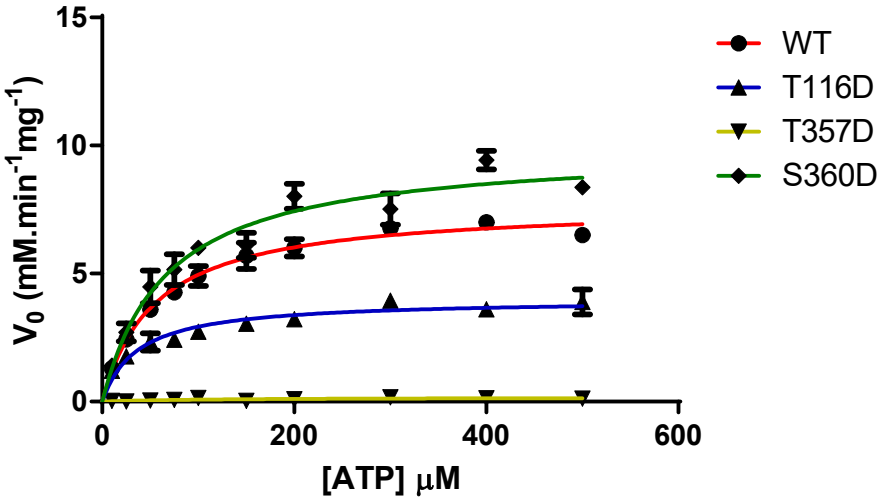

Figure 2.11: Michaelis-Menten graph displaying specific activities (mM.min^{-1}mg^{-1}) of MurC WT, T116D, T357D and S360D for ATP dependence in the presence of UDP-MurNAc. The K_m for MurC WT, T116D, T357D and S360D are shown in Table 2.4. R^2 values for each MurC variant are 0.9692, 0.8596, 0.3644 and 0.9284 respectively, indicating high goodness of fit for each variant (apart from T357D as values are too low to provide a better fit).

Table 2.4: Kinetic values for MurC WT and pseudo-phosphorylated mutants for ATP dependence in the presence of UDP-MurNAc. MurC T116D shows a higher affinity for ATP than WT, yet both S360D and T357D show a lower affinity than WT. MurC S360D shows the highest specific activity, being slightly higher than WT. However, T116D has specific activity just under half of that of WT, whereas T357D once again displays very low activity. These figures are correlated by both K_{cat} and K_{cat}/K_m.

MurC Variant	K_m (μM)	V_{max} (mM.min^{-1})	Specific Activity (mM.min^{-1}mg^{-1})	K_{cat} (min^{-1})	K_{cat}/K_m (min^{-1}M^{-1})
WT	55.46 (±5.635)	0.0129 (±0.00036)	7.70	78.90	1.42×10^6
T116D	36.70 (±7.192)	0.0042 (±0.00019)	3.99	41.18	1.12×10^6
T357D	113.6 (±98.52)	0.0010 (±0.00031)	0.17	1.79	1.58×10^4
S360D	66.79 (±10.97)	0.0057 (±0.00028)	9.92	101.33	1.52×10^6

The affinity of MurC T116D (36.70 ±7.192 μM) is higher for ATP in the presence of UDP-MurNAc than MurC WT (55.46 ±5.635 μM), whereas both S360D and T357D display higher K_m values than WT (Table 2.4). The specific activities of all assayed MurC variants show a similar trend as the previous assays, wherein MurC T116D displays a specific activity around half that of MurC WT and T357D shows an extremely low specific activity, whereas MurC

S360D shows an activity almost a third higher than MurC WT (Figure 2.11, Table 2.4). This trend is followed by the K_{cat}, showing a turnover number for MurC T116D (41.18 min^{-1}) again around half that of WT (78.90 min^{-1}). MurC T357D is shown to turnover 44 times less reactions (1.79 min^{-1}) than MurC WT (Table 2.4). MurC S360D displays a K_{cat} of 101.33 min^{-1}, showing an ability to perform more reactions per minute than MurC WT (Table 2.4). This is finally corroborated by the specificity constant values for these proteins, whereby the K_{cat}/K_m for both MurC WT and S360D are comparable (1.42×10^6 $min^{-1}M^{-1}$ and 1.52×10^6 $min^{-1}M^{-1}$ respectively) showing a similar efficiency to one another, but the K_{cat}/K_m for MurC T116D is slightly lower (1.12×10^6 $min^{-1}M^{-1}$), indicating a lower preference for ATP than WT (Table 2.4). However, the K_{cat}/K_m for MurC T357D is 90-fold lower than WT (1.58×10^4 $min^{-1}M^{-1}$) indicating the inefficiency of this pseudo-phosphorylated mutant.

2.2.4. Ligand binding of ATP to MurC pseudo-phosphorylated mutants

We utilised intrinsic tryptophan fluorescence (ITF) in order to investigate how pseudo-phosphorylation of MurC affects the binding of ATP to the protein (Section 5.15.). MurC WT and each pseudo-phosphorylated mutant displayed a varying degree of sigmoidal binding, indicating co-operative binding of ATP to MurC. Therefore, the data was fitted to a single site saturation model with a Hill slope, yielding both an equilibrium binding constant (K_d) and Hill coefficient/slope (h) for each protein. The Hill equation is used when a multi-site protein displays co-operative binding, with the Hill coefficient being derived from the equation ($Y = \frac{[Ligand]^h}{K_d+[Ligand]^h}$). If h>1 the protein shows positive co-operative binding, whereas if h<1 the protein displays negative co-operative binding. However, if h=1 then the binding is non-co-operative and the protein obeys the usual single site saturation binding equation (Stefan and Le Novere, 2013).

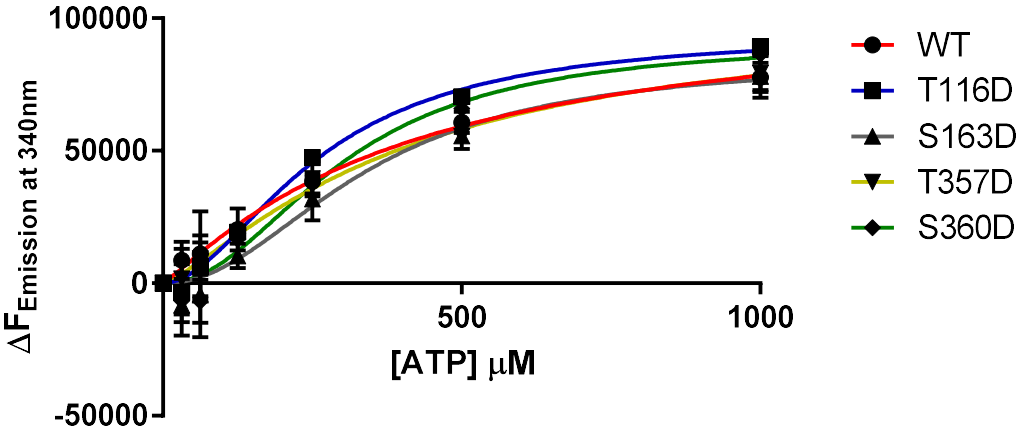

Figure 2.12: Saturation binding experiments to using ITF to assess ATP binding to MurC pseudo-phosphorylated mutants. Each protein displays a sigmoidal binding curve, indicative of co-operative binding. MurC WT, T116D, S163D, T357D and S360D gave K_d values of of 419 μM, 261.4 μM, 340.9 μM, 407.8 μM and 302.7 μM respectively. They each gave a Hill coefficient (h) of 1.121, 1.86, 2.109, 1.298 and 2.129 respectively, indicating that each protein exhibits positive co-operative binding. R^2 values for each MurC variant are 0.9451, 0.9610, 0.9360, 0.9731 and 0.9682 respectively, indicating high goodness of fit for each curve.

MurC WT shows a K_d of 419 μM for ATP and shows that h = 1.121, which indicates that MurC WT does show some positive co-operative binding of ATP (Figure 2.12). However, MurC T116D has almost twice the binding affinity for ATP than MurC WT, with a K_d of 261.4 μM. This correlates with an h value of 1.86, almost twice that of MurC WT (Figure 2.12). MurC S163D displays a K_d of 340.9 μM, which is still an increased binding affinity for ATP when compared to WT, yet this mutant has an h value of 2.109 indicating much higher positive co-operation of binding sites (Figure 2.12). MurC T357D behaves much more like MurC WT, with a comparable K_d of 407.8 μM and h = 1.298, showing less co-operative binding than the previous two mutants (Figure 2.12). However, MurC S360D displays a lower K_d of 302.7 μM, indicating a higher binding affinity for ATP, but a high h value of 2.129, showing the most positive co-operation between binding sites of all MurC variants (Figure 2.12). Therefore, MurC T116D, S163D and S360D show both higher binding affinity for ATP and also much higher positive co-operation between binding sites than MurC WT or T357D (Figure 2.12).

2.3. Discussion

The essential MurC enzyme is a key component in *M. tuberculosis* PG biosynthesis. As the first member of the sequential ATP-dependent Mur ligase pathway, it represents an interesting research topic, since disruption of any point in the pathway will cause PG biosynthesis to halt and will result in cell death (Munshi *et al*., 2013). The *murC* gene is among a cluster of genes known as the division/cell wall (*dcw*) operon, which contains each ATP-dependent Mur ligase (*murCDEF*), along with *murX* and *murG* in close proximity to each other, as well genes involved in cellular division processes, such as *ftsZ* and *ftsQ* (Munshi *et al*., 2013). MurC and the ATP-dependent Mur enzymes are thought to be a part of a wider superfamily of ATP-dependent amide ligases which include folyl-γ-polyglutamate ligases, which add glutamate to folate or folate derivatives in humans and the poly-γ-glutamate synthetase CapB (from *B. subilitis*) (Candela and Fouet, 2006, Lawrence *et al*., 2014). These enzymes share mechanistic similarities and some sequence homology to the Mur ligases across various bacterial and mammalian species (Eveland *et al.*, 1997). As the first of the ATP-dependent Mur ligases, it has been shown in *C. glutamicum* that the kinase PknA is utilised to phosphorylate specific sites of MurC in order to regulate its activity (Fiuza *et al*., 2008). We therefore hypothesised that *M. tuberculosis*, as a relative of *C. glutamicum*, would likely be capable of regulating MurC and therefore PG biosynthesis by phosphorylation of similar specific residues.

The site-directed mutagenesis of specific Ser/Thr sites to a negatively-charged amino acid can be performed to create phospho-mimics at these residues. This is particularly useful in situations such as this study, whereby the phosphorylation of specific individual residues is to be investigated, as well as the fact that definitive phosphorylation of MurC by PknA is yet to be shown. To assess the effect of phosphorylation of MurC, we decided to generate site-directed mutants of homologous specific Ser/Thr residues and substitute them with

negatively-charged aspartate residues. This approach to pseudo-phosphorylation has been performed successfully in the past, particularly in the analysis of the effects of phosphorylation at specific sites of the neuronal microtubule-associated protein tau. These previous studies showed that the substitution of Ser/Thr for Asp not only simulated the effects of phosphorylation upon tau, but also showed little difference between Asp or glutamate (a larger negatively-charged residue) substitutions (Leger *et al*., 1997, Gohar *et al*., 2009, Sun and Gamblin, 2009).

Through extensive bioinformatic analysis of MurC, homologous Ser/Thr residues were identified in the *M. tuberculosis* protein, comparable to those phosphorylated in the *C. glutamicum* enzyme (Figure 2.4). These residues were then mapped onto 3D homology models of both Mtb-MurC and Cg-MurC, then compared to one another (Figure 2.5). The two models overlay with a TM-score of 0.93544, indicating a good fit (since a TM-score of 1 means a perfect fit (Zhang and Skolnick, 2004)), with the selected phosphorylation sites occupying almost the exact same area in space between the two models (Figure 2.5A,B Top). Importantly, all of the selected residues in the Mtb MurC model appear to have a function in the binding of ATP, based on the location of bound AMPPNP, from the overlay of the *H. influenzae* MurC crystal structure. Both T116 and S163 appear to be the furthest from the bound ATP analogue, however, they are found at the opposite ends of two loops that appear to directly interact with ATP, which suggests they are likely candidates for phosphorylation *in vivo* (Figure 2.5). These residues also appear to be surface exposed by the use of a spacefilling model of Mtb-MurC (Figure 2.5A Bottom). Along with both T357 and S360, which also appear to be surface exposed, it increases the likelihood of these residues being phosphorylated *in vivo*, since any potential kinase that would act on these residues can reach them easier. T129 (and its Cg homolog T133) appears to be buried within the structure of

both Mtb and Cg spacefilling models, so it is unknown how any kinase would be capable of reaching this residue *in vivo* (Figure 2.5A,B Bottom).

The initial Michaelis-Menten kinetic data for WT MurC showed a high preference of the protein for UDP-MurNGlyc over UDP-MurNAc, with a 30-fold difference of affinity between the two (Figure 2.8). The K_m value of 1,555 ± 272 μM for UDP-MurNAc differs to the K_m of 23.5 ± 0.5 μM reported by Munshi *et al*., 2013, believed to be due to fact that they performed stopped assays for phosphate release after 30 min, whereas we performed assays that recorded the initial linear rates of reaction where the rate is proportional to the enzyme concentration. It is unlikely that they recorded initial rates, since it can be shown using their conditions and kinetic constants that MurC WT would have fully consumed all of the substrate before a measurement of activity could be made, which would lead to an inaccurate K_m being suggested. The acidic nature of the assay employed by Munshi *et al*., 2013 could also result in false positives in regard to activity, since UDP-Muramyl precursors tend to be susceptible to instability under these conditions. Both of these discrepancies are likely to account for the difference between the reported K_m and our data. Nonetheless, the significant difference in our affinities between UDP-Muramyl substrates led us to utilise UDP-MurNGlyc to perform UDP-Muramyl dependent assays using each pseudo-phosphorylated mutant of MurC (Figure 2.9). The same major trend was seen when the assays were repeated for ATP dependency using both UDP-MurNGlyc and UDP-MurNAc, whereby the activity of MurC T116D was reduced to around half that of WT, whereas the activity of MurC T357D was all but abrogated (Figure 2.10, Figure 2.11). Due to the vast quantities of UDP-MurNAc required for each assay, we were unable to perform any ATP dependency assays using MurC S163D and UDP-MurNAc, due to limiting substrate. However, since each of the other pseudo-phosphorylated mutants followed similar trends across each assay, we expect that MurC S163D would behave in a similar fashion to MurC WT.

The binding of ATP to MurC WT and pseudo-phosphorylated mutants was further probed using ITF. Each protein displayed a sigmoidal binding curve, indicative of co-operative binding between multiple binding sites, when at saturation with a defined B_{max} and binding affinity at the μM range (Figure 2.12). Because of the sigmoidal binding nature shown in these assays, we were also able to deduce a Hill coefficient from the binding of ATP to each MurC variant, which gives an insight into the co-operation between binding sites (Stefan and Le Novere, 2013). Each pseudo-phosphorylated MurC mutant displayed higher binding affinity for ATP, as well as increased positive co-operation between binding sites, than MurC WT, except for MurC T357D which behaved in a similar manner to MurC WT (Figure 2.12). Since MurC has been shown to exist in solution as both monomeric and dimeric forms, in *E. coli*, it is possible that the alteration of T116, S163 and S360 to aspartate increases the presence of dimerised MurC in solution, which would then afford these mutants a secondary ATP binding site and help to explain the increase in positive co-operative binding (Das *et al*., 2011). The increase in dimerisation of these MurC mutants is further supported by the fact that both T116 and S163 lie close to the dimer interface, between two MurC molecules (Figure 2.13). The negative charges introduced by these residues may cause an increase of interactions between opposing residues of the proteins or conformational change that favours dimerisation of the protein, for instance, T116 seemingly interacts with the N-terminal region of the opposing monomeric unit, so mutation of T to D may lead to an increase of non-covalent interactions (Figure 2.13). S360D may also cause a conformational change that resonates throughout the protein leading to an increased preference for dimers to form.

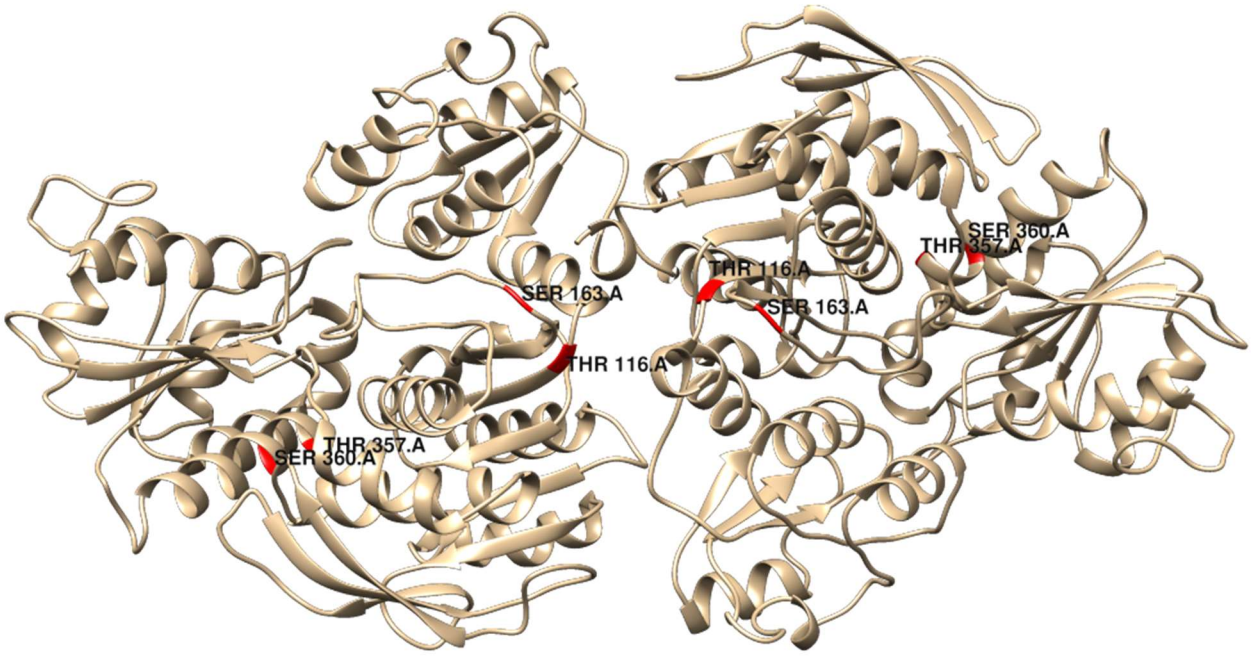

Figure 2.13: 3D-homology model representation of Mtb-MurC dimer. Both T116 and S163 are situated near to the dimer interface, suggesting that alteration of these residues may have an effect on dimerisation. The mutation of these residues to aspartate increases positive co-operation between binding sites suggesting that an increase of dimerisation has occurred when compared to MurC WT. The mutation of S360 to aspartate also increases co-operative binding, suggesting that a conformational change may result in an increase in dimer formation. Dimer model based upon *E. coli* MurC dimer crystal structure (PDB: 2f00).

This study provides a comprehensive biochemical insight into the outcomes of phosphorylation of *M. tuberculosis* MurC. Given the kinetic data results and the positions of the residues within the homology model, it is likely that *M. tuberculosis* could phosphorylate T116 and T357 in order to augment the activity of MurC. Phosphorylation of T116 would likely regulate the activity of the enzyme by around 50 %, whereas the modification of T357 would likely lead to an almost complete halt in enzymatic activity. It is likely that *in vivo* multiple residues are phosphorylated at once in order to maintain the level of activity of MurC that the cell requires at a particular time, since pseudo-phosphorylation of S360 displayed a very slight increase in activity of 15-20 %, compared to WT. Therefore by phosphorylating multiple sites at one time *M. tuberculosis* can possibly fine tune the activity of MurC beyond simply around 50 % activity, but to the level required.

It has been shown previously in *C. glutamicum* MurC that multiple threnonine phosphorylation sites are required for activity. Successive site-directed mutagenesis by Fiuza *et al*. (2008) revealed that the loss of more than three phosphorylation sites at once greatly decreased the activity of MurC, with a mutation of all six residues at once abolishing activity, in both *in vitro* assays as well as *in vivo* functional complementation experiments. It is therefore likely that each phosphorylation of MurC, at one or more sites, could differentially regulate the protein. This could potentially be investigated in Mtb-MurC by producing proteins containing multiple pseudo-phosphorylation mutations, including double, triple and quadruple mutations of T116, S163, T357 and S360, with multiple permutations of each mutant. These proteins would then require assaying in the same manner as the single site mutants (Section 2.2.3.) in order to compare the activity of MurC with increasing pseudo-phosphorylated residues and identify whether the regulation of MurC can be fine-tuned in *M. tuberculosis* by multiple phosphorylation sites.

The severe reduction in enzymatic activity caused by the pseudo-phosphorylation of T357 can be attributed to the location of the residue within the protein (Figure 2.14). Situated near to the enzymatic active site and ATP binding cleft, the change in charge at this position would more than likely cause a change in the electrostatic charges between neighbouring residues, in particular His124 and Glu192. The probable position of Asp357 (57.8% likelihood) is also a likely candidate for the position of a phosphate group attached to the native Thr357 (Figure 2.14). This means that the introduced negative charge would be closer to the negatively charged Glu192 than before, which would likely introduce repulsion between the residues and cause conformational and electrostatic disruption to the neighbouring active site, diminishing enzymatic activity (Figure 2.14).

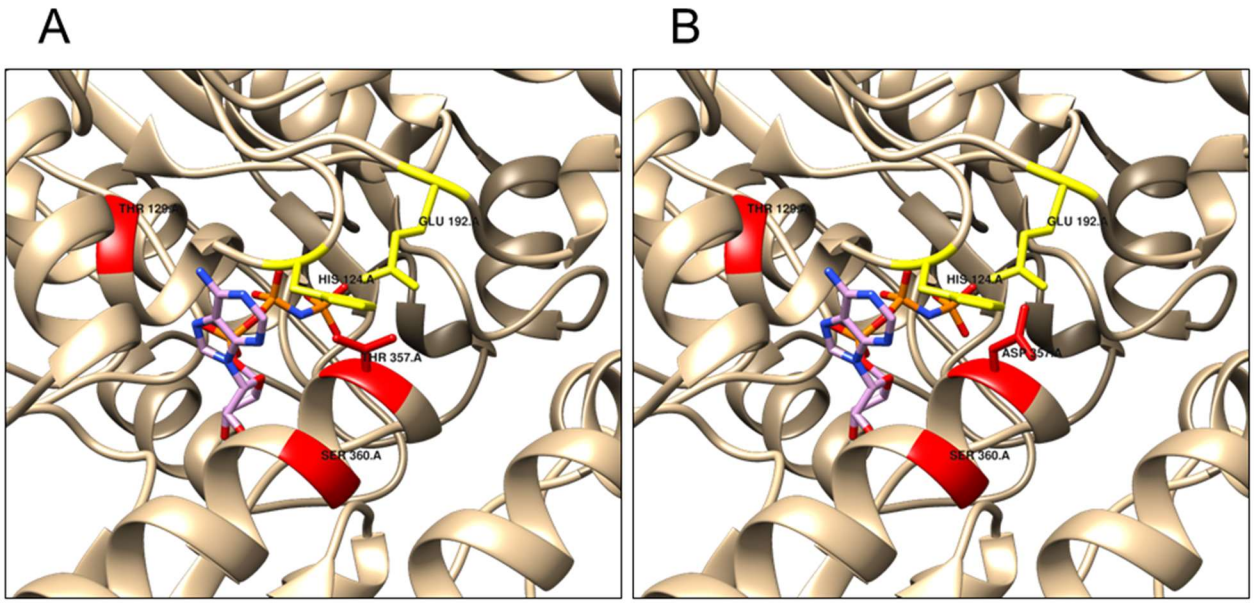

Figure 2.14: 3D-homology model of Mtb-MurC residue 357 with ATP bound. A) The predicted position of Thr357 in relation to nearby side chains. His124 and Glu192 are near to the enzyme active site. B) Mutation of Thr357 to Asp357 in the most probable position (57.8% likely) is also likely a very similar position of a phosphate group attached to Thr in the native protein. The introduction of a negative charge at this position would have an effect on the electrostatic charges within the protein, in particular the proximity to Glu192 and could easily cause disruption to the nearby active site.

The increased affinity for UDP-MurNGlyc is an important consideration concerning the phosphorylation of MurC. Non-replicative (dormant) *M. tuberculosis* cells inhabit granulomas within the host lungs, where the conditions are hypoxic and the bacilli halt growth. This is ended when granulomas break down through a compromise of the host immune system and the atmospheric oxygen rises (Section 1.1.1.). The increase in molecular oxygen would result in production of UDP-MurNGlyc through the hydroxylation of UDP-MurNAc by NamH (Raymond *et al*., 2005). We propose that during dormancy *M. tuberculosis* tightly regulates MurC activity by phosphorylation of the identified residues (e.g. T116 or T357) in order to maintain a low activity and along with phosphorylation of MurD, reduce PG biosynthesis. This low activity is also maintained by the lack of up-regulation of MurC during dormancy keeping protein abundance low (Murphy and Brown, 2007). Once the granuloma begins to break down and oxygen becomes introduced, MurC (and MurD) is de-phosphorylated in order to increase activity and the high affinity for the newly converted UDP-MurNGlyc drives

PG production forward in order to help push cell growth and to exit dormancy. We do not suggest that this is the only mechanism of *M. tuberculosis* to resume a replicative state, nor the only function of hydroxylation of UDP-MurNAc or phosphorylation of MurC. For instance, incorporation of MurNGlyc into the PG structure has been shown to increase the resistance of the cell wall to lysozyme assault (Section 1.2.2.1.). It is probable that *M. tuberculosis* could slow the rate of PG biosynthesis through phosphorylation of MurC, to allow time for NamH action to increase the pool of UDP-MurNGlyc for incorporation into the cell wall and increase the protection from lysozyme, therefore regulating the level of MurNGlyc found within the PG structure.

The discovery that pseudo-phosphorylation of specific sites of *M. tuberculosis* MurC has an effect on activity of the enzyme does raise a key issue, that of the identification of a potential *M. tuberculosis* MurC kinase. As previously mentioned (Section 2.1.), PknA has been shown to phosphorylate *M. tuberculosis* MurD, rather than MurC, but has recently been identified as capable of interacting with MurC (Thakur and Chakraborti, 2008, Munshi *et al*., 2013). However, it has not yet been proven that PknA is actually capable of phosphorylating MurC in *M. tuberculosis*, so the identity of a potential MurC kinase remains elusive.

Chapter 3

Identification of a putative mycobacterial AG/PG ligase

3. IDENTIFCATION OF A PUTATIVE MYCOBACTERIAL AG/PG LIGASE

3.1. Introduction

Much is known about the biosynthetic pathway that leads to the formation of mycobacterial AG and PG (Section 1.2.), however knowledge of the final steps of AGP construction, namely how AG and PG are covalently tethered, is missing. It is recognised that AG is covalently attached to 10-12 % of muramyl residues of PG, specifically at the C6-hydroxyl *via* a unique LU, also that AG only becomes bound to PG once the latter becomes fully matured through peptide cross-linkages by transpeptidase action (Amar and Vilkas, 1973, McNeil *et al*., 1990, Hancock *et al*., 2002). However, the enzymatic process of this final "ligation" of AG to PG *via* a phosphodiester bond remains unsolved.

Recently, a novel superfamily of proteins was identified called the LytR-CpsA-Psr (LCP) family, which are involved in the covalent attachment of anionic CWPs to the PG of Gram-positive bacteria (Kawai *et al*., 2011). Anionic CWPs consist of both cell wall teichoic acids (CWTAs), which are bound directly to PG *via* a phosphodiester bond and lipoteichoic acids (LTAs) which are membrane-anchored and span throughout the PG layer (Weidenmaier and Peschel, 2008). Both of these CWPs have important roles in orchestrating the placement of cell division machinery, as well as being directly involved in host-pathogen interactions (Weidenmaier and Peschel, 2008). Interestingly, Gram-positive CWPs share a strikingly similar chemical phenotype to the homologous AG and LM/LAM polysaccharides from the mycobacterial cell wall (Crick *et al*., 2001, Jankute *et al*., 2012). CWTAs also share a similar biosynthetic pathway to AG, being initially synthesised within the cytoplasm as a lipid-linked polymer, translocated across the plasma membrane, before being attached to the PG (Swoboda *et al*., 2010). The LU of Gram positive CWTAs are also very similar to the unique LU of mycobacterial AG (Figure 3.1). The disaccharide LU of Gram-positive bacteria, such

as *B. subtilis*, consists of β-D-*N*-acetylmannosamine (ManNAc)-(1→4)-α-D-GlcNAc-(1→P), which is similar to α-L-Rha*p*-(1→3)-α-D-GlcNAc-(1→P) of AG, notably they both share D-GlcNAc-(1→P) linking to PG (Kaya *et al*., 1984, McNeil *et al*., 1990, Weidenmaier and Peschel, 2008).

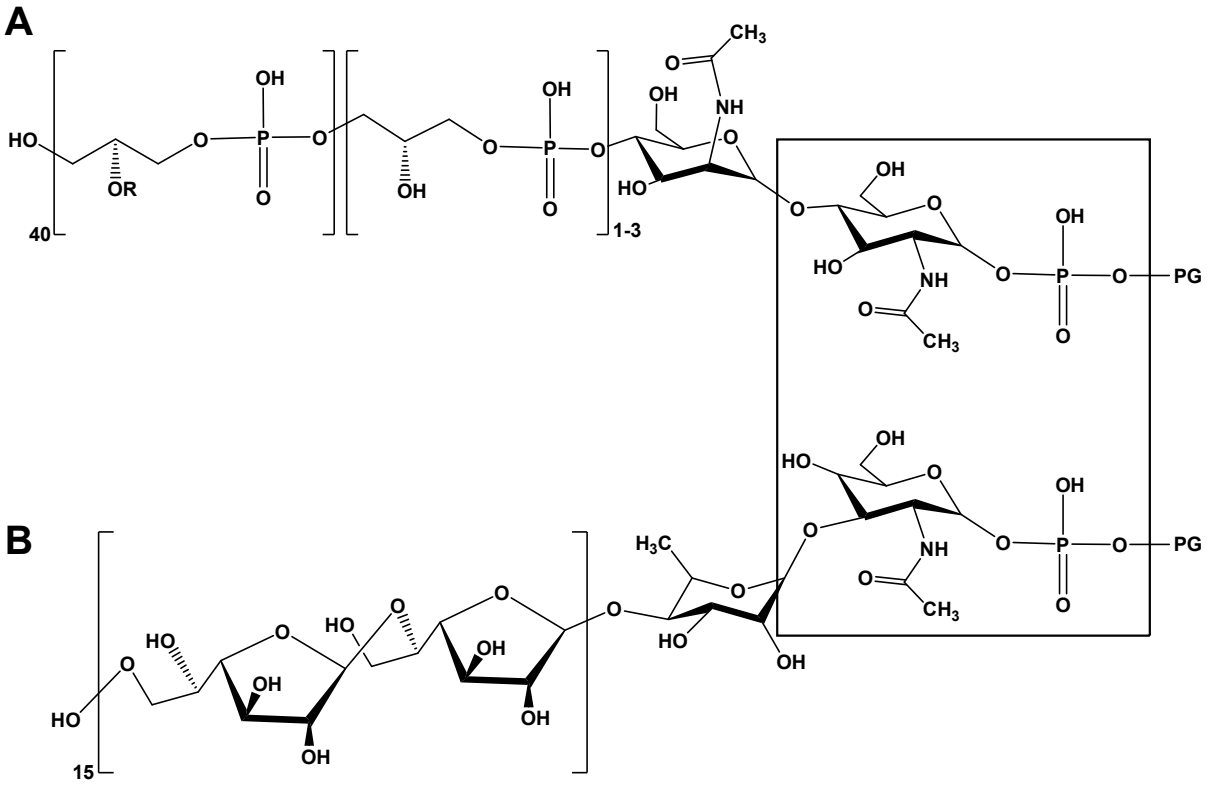

Figure 3.1: Comparison of LU structures. A) *B. subtilis* 168 CWTA showing (Glycerol-P(GroP))$_{40}$-(GroP)$_{1\text{-}3}$-β-D-ManNAc-(1→4)-α-D-GlcNAc-(1→P). R side chain represents H, D-Ala or Glucose B) *M. tuberculosis* AG showing ((1→6)-Gal*f*-(1→5)-Gal*f*)$_{15}$-α-L-Rha*p*-(1→3)-α-D-GlcNAc-(1→P). The two structures vary, but both LUs comprise of a disaccharide, comprising of either D-ManNAc (CWTA) or L-Rha*p* (AG) and a common D-GlcNAc-(1→P) forming a phosphodiester bond between the LU and C6-hydroxyls of muramyl residues of PG (boxed).

A recently characterised LCP protein identified in *Streptococcus pneumoniae*, Cps2A, acts as a phosphotransferase in the addition of CWTAs onto the PG (Kawai *et al*., 2011, Eberhardt *et al*., 2012). The crystal structure of Cps2A (solved to 1.69 Å) is described as having 2 domains, an LCP domain that is conserved across species (30% sequence identity with 3 LCP

family proteins from *B. subtilis*) and an accessory domain, with a sequence and structure that is not conserved at all, to the point where it is not present in the sequence of some other LCP proteins (Kawai *et al*., 2011). The core of the LCP domain of Cps2A is an α-β-α structure, a 5-stranded β-sheet surrounded by 2 α-helices, with 2 β-strands that extend from the protein core to form the interface between the domains. A highly conserved polyisoprenoid binding pocket containing hydrophobic residues was identified and it was noted that Cps2A co-eluted and crystallised with a bound decaprenyl-phosphate (from *E. coli* expression host), with the phosphate head group held in place by a variety of conserved charged residues (Kawai *et al*., 2011). Beyond this hydrophobic pocket there is a space available that would accommodate an oligosaccharide LU, in a cleft on the surface of the protein, which is then likely to interact with the conserved arginine pair Arg362 and Arg244 (Kawai *et al*., 2011). A groove was also identified between β-strands that was suggested to accommodate PG, at an almost 90° angle to CWTA (Kawai *et al*., 2011). Observed structural data also showed that Cps2A is magnesium dependent, with Mg^{2+} binding between 2 aspartate residues (Asp234 and Asp246). Mg^{2+} is suggested to stabilise the transition state intermediate, by neutralising the emerging negative charge (Kawai *et al*., 2011). The absence of Mg^{2+} (or the removal of Asp234 by site-directed mutagenesis to alanine) in colourimetric assays showed far less liberation of inorganic phosphate, indicating a reduction of enzyme activity and the importance of Mg^{2+} as well as key catalytic residues (such as Asp234) to enzymatic function (Kawai *et al*., 2011).

It has been identified that there are multiple LCP proteins present in various Gram-positive bacteria, which appear to have semi-redundant roles (Eberhardt *et al*., 2012). It has been shown that some of these homologs have a non-specific function (in *S. pneumoniae*), such as LytR and Psr, which can assist in the addition of polymers to the cell wall in the absence of Cps2A. However, it was noted that the homologous LCP genes cannot fully compliment the

function of genetically-altered strains depleted of a single LCP gene, as reduced cell wall material was recorded (Eberhardt *et al*., 2012). It was also shown that all 3 of these homologs are involved in some aspect of cell wall biosynthesis due to GFP-fused proteins aggregating at the mid-cell, where pneumococcal cell wall synthesis takes place, which in turn supports the hypothesis that these LCP proteins complete cell envelope maturation by attachment of CWTAs to PG (Eberhardt *et al*., 2012).

The *cpsA* gene was investigated in *Mycobacterium marinum* (one of four LCP-like proteins), by the use of a transposon insertion mutant, which had a kanamycin resistance cassette inserted into the predicted LCP domain of CpsA, which resulted in loss of protein action (Wang *et al*., 2015). The resulting mutant displayed phenotypes unlike the wild-type, such as altered colony morphology, sliding motility, reduced cell wall hydrophobicity and permeability and reduced virulence (Wang *et al*., 2015). Of note is that this *cpsA* mutant contained reduced AG within its cell walls, indicating a potential reduction, but not total abolition, of attachment to the cell wall (Wang *et al*., 2015). It is therefore highly probable that *M. tuberculosis* has a similar set of proteins that belong to the LCP family, in particular one which can identify the common D-GlcNAc-(1→P) of the LU and act as a peptidoglycan-arabinogalactan ligase (Figure 3.1).

In this study, we have identified Rv3267 as being a direct homolog of Cps2A. Our results demonstrate that the *M. smegmatis* orthologue (MSMEG1824) is an essential gene as determined using Conditional Expression-Specialised Transduction Essentiality Test (CESTET). Furthermore, a comprehensive biochemical characterisation of recombinant Rv3267 (henceforth termed Lcp1) demonstrates its ability to “ligate” AG to PG in a radiolabelled cell-free assay.

3.2. Results

3.2.1. Identification and *in silico* analysis of Lcp1

A bioinformatic analysis (Section 5.17.) of mycobacterial and corynebacterial genomes using *S. pneumoniae* Cps2A as a query revealed Rv3267 (^{Mtb}Lcp1) as a primary LCP homolog in *M. tuberculosis* showing 20 % sequence identity to Cps2A. A homologous protein was also identified in *M. smegmatis*, as MSMEG1824 (^{Ms}Lcp1) with 71 % sequence homology to ^{Mtb}Lcp1 (Figure 3.2). An alignment of *lcp1* and its surrounding genes (from *M. tuberculosis*, *M. smegmatis*, *Mycobacterium leprae* and *C. glutamicum*) showed close genetic synteny, indicating conserved functionality between these species (Figure 3.3A). Interestingly, the fact that the gene cluster is conserved in the degenerate genome of *M. leprae* indicates an important conserved function for these genes to the survival of mycobacteria (Sassetti *et al*., 2003). Furthermore, *lcp1* is found immediately upstream of several genes involved in cell wall biosynthesis, importantly *rmlD* and *wbbL1*, both of which encode for proteins involved in the biosynthesis of the AG LU (Section 1.2.3.).

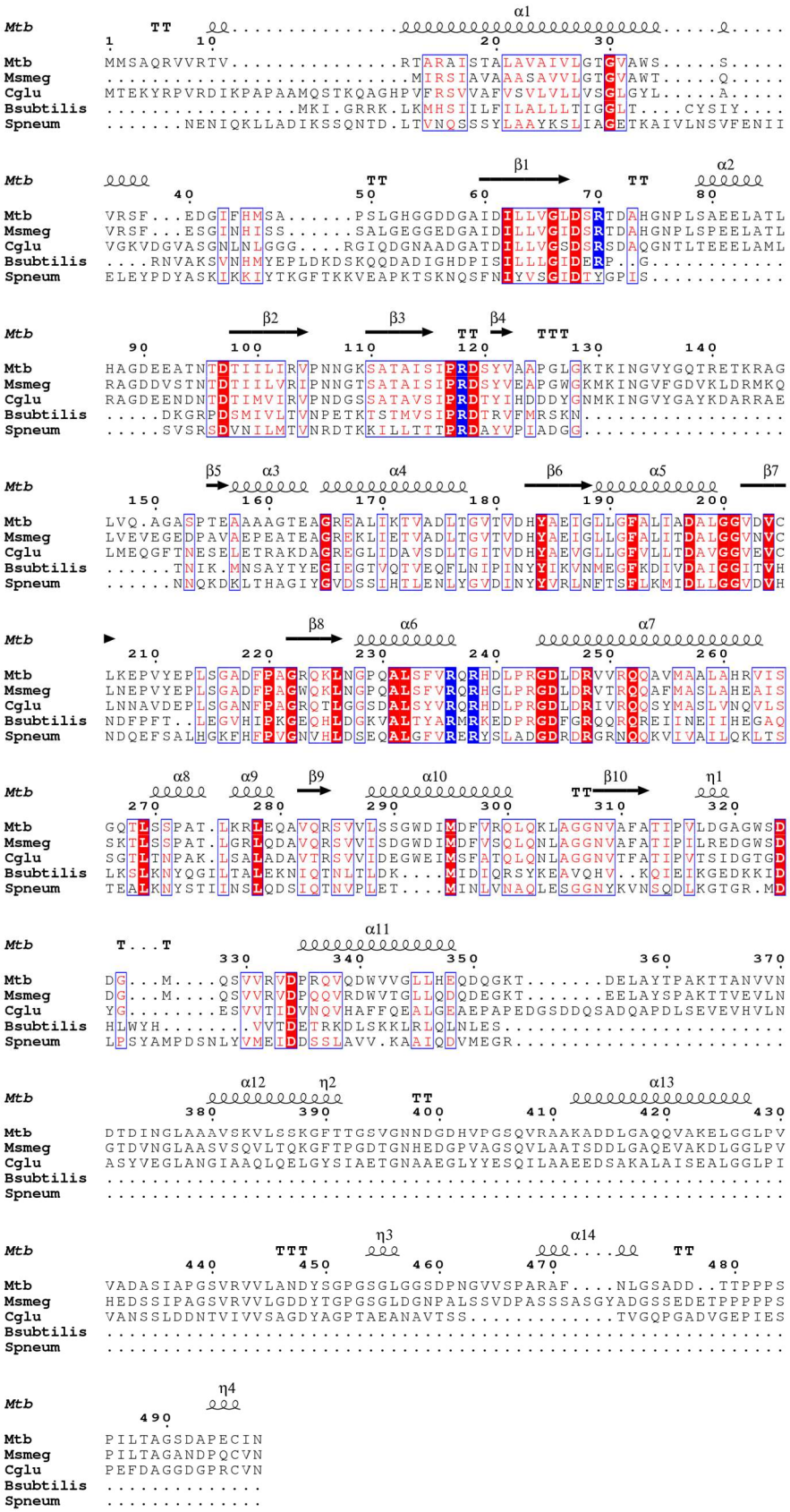

Figure 3.2: Multiple protein sequence alignment of Lcp1 from different bacterial species. *M. tuberculosis* (Rv3267), *M. smegmatis* (MSMEG1824), *C. glutamicum* (NCgl0708) and homologous proteins from *B. subtilis* (LytR) and *S. pneumoniae* (Cps2A) (minus 140 N-terminal residues representing non-conserved accessory domain) share some sequence homolgy. The predicted secondary structure of ^{Mtb}Lcp1 is shown along the top of the alignment. Completely conserved residues are highlighted red, whereas homologous residues mostly conserved are in red font. Residues in blue indicate conserved residues believed to be involved in polyisoprenoid binding.

A 3D homology model was produced from the ^{Mtb}Lcp1 protein sequence (UniProtKB - P96872_MYCTU) using the I-TASSER server (C-score of -2.36) with a predicted single transmembrane helix (between residues 1 and 36) omitted (Figure 3.3B,C) (Section 5.17.). The homology model was then analysed in comparison to the *S. pneumoniae* Cps2A crystal structure model (PDB ID: 2xxp) (Section 5.17.). The predicted N-terminal α-helix corresponds with what has been observed in other LCP proteins, such as *S. pneumoniae* Cps2A (Kawai *et al*., 2011), however the sequence and model of ^{Mtb}Lcp1 does not contain the non-conserved accessory domain as Cps2A. The LCP domain of ^{Mtb}Lcp1 is indicated to contain a 5-stranded β-sheet, surrounded by α-helices forming the core of the protein, similar to the structure of the LCP domain of Cps2A. The homology model also retains the highly conserved polyisoprenoid binding hydrophobic pocket (Figure 3.3C). Numerous amino acid residues were identified as catalytically important in Cps2A by Kawai *et al*. (2011), in particular the arginine residues Arg267, Arg362, Arg364 and Arg374, as well as Asp371 and Gln378. Using both the multiple sequence alignment and positions of residues within the homology model (Figure 3.2, 3.3C) we identified four residues in ^{Mtb}Lcp1 (Arg70, Arg118, Arg236 and Arg238) which were believed to be involved in polyisoprenoid-pyrophosphate binding and therefore catalytically important.

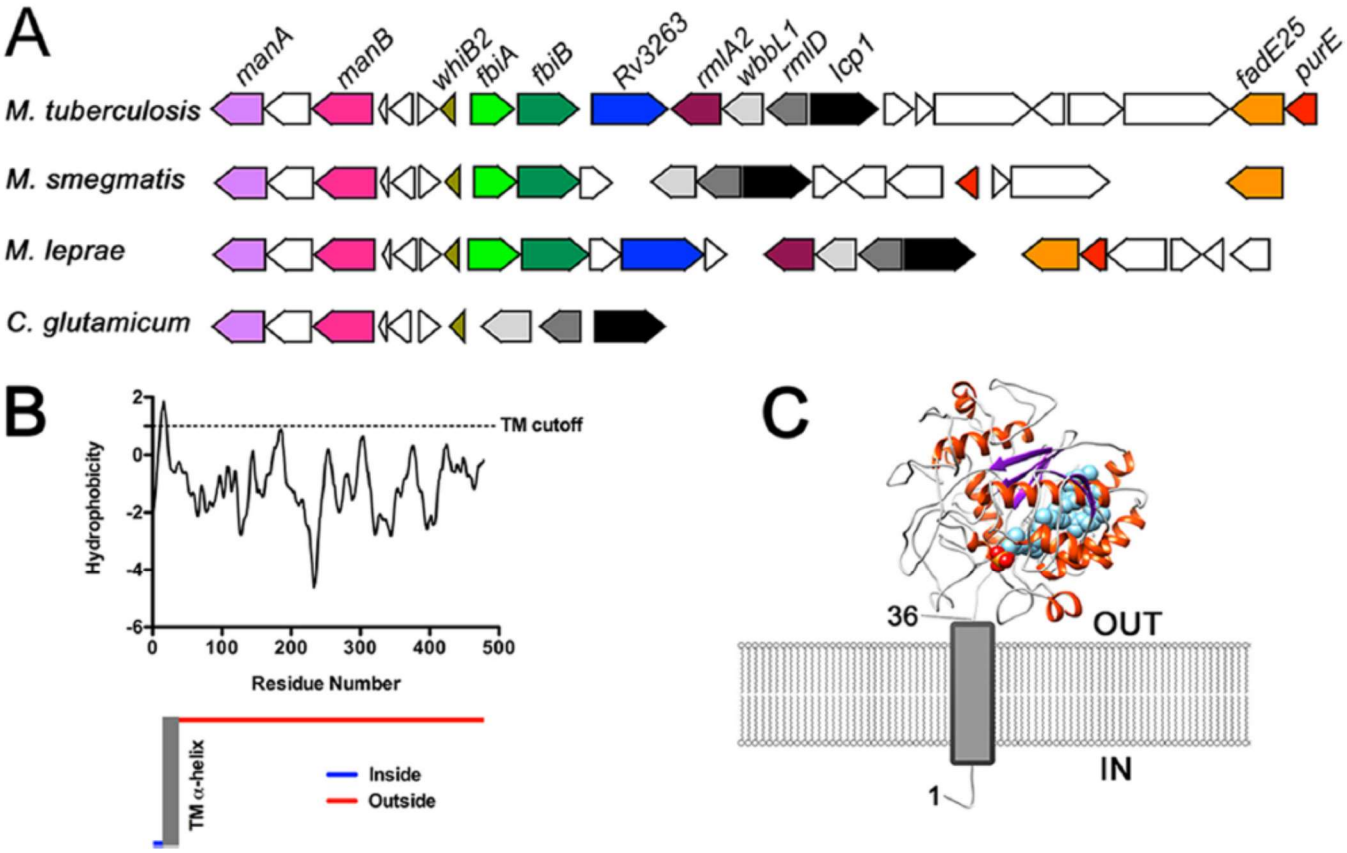

Figure 3.3: Comparison of the *lcp1* locus and *in silico* analysis of Lcp1 by TMHMM and I-TASSER. A) Mycobacterial *lcp1* locus alignment, where *lcp1* is black. The locus tags are *Rv3267* in *M. tuberculosis*, *MSMEG1824* in *M. smegmatis*, *ML0756* in *M. leprae* and *NCgl0708* in *C. glutamicum*. Each species shows genetic synteny, since *lcp1* is consistently immediately upstream of both *rmlD* (dark grey) and *wbbL1* (light grey). B) TMHMM server prediction of the hydrophobicity of ^{Mtb}Lcp1 (Section 5.17.). Predicted membrane toplogy showed an N-terminal transmembrane α-helix between residues 1 and 36, with the majority of the protein being extracellular. C) I-TASSER 3D model of ^{Mtb}Lcp1, positioned in relation to the plasma membrane, showing transmembrane domain between residues 1 and 36. Model also shows bound decaprenyl-phosphate.

3.2.2. Essentially of ^{Ms}Lcp1 in *M. smegmatis*

Since Lcp1 is suspected to have an important role in the attachment of AG to PG, we hypothesised that *lcp1* is likely to be an essential gene in mycobacteria. To test this hypothesis, we created a knockout phage (phΔ^{Ms}*lcp1*) (Section 5.19.) in order to replace *MSMEG1824* (^{Ms}*lcp1*) with a hygromycin resistance cassette in *M. smegmatis* (Bhatt and Jacobs, 2009). As expected, there were no direct transductants, indicating an inability to directly generate a Δ^{Ms}*lcp1* null mutant. Therefore, a merodiploid strain of *M. smegmatis* harbouring a secondary acetamide inducible copy of ^{Ms}*lcp1* (Section 5.18.2.) was used for transduction with phΔ^{Ms}*lcp1* by CESTET (Bhatt and Jacobs, 2009). Only when acetamide was present in the growth media were Δ^{Ms}*lcp1* knockout mutants generated.

To confirm that a conditional Δ^{Ms}*lcp1* mutant had been created, cells were grown in the presence of acetamide and genomic DNA was extracted (Section 5.9.). As previously explained, if the genomic copy of ^{Ms}*lcp1* had been successfully knocked out, then it would be replaced with a larger hygromycin resistance cassette (Figure 3.4A). Therefore, PCR was performed on the extracted genomic DNA using primers that flank either side of genomic ^{Ms}*lcp1*, which was then analysed by agarose gel electrophoresis (Section 5.5., Section 5.7.) (Figure 3.4B). Primers for *murC* amplification were also used as a control experiment to confirm presence of *M. smegmatis* template DNA (Section 5.5.1.) (Figure 3.4B). This PCR showed a band at around 3.5 kb using primers for the flanking regions of ^{Ms}*lcp1* from *M. smegmatis* WT genomic DNA (Figure 3.4B, Lane 2). However, the same experiment using genomic DNA from the Δ^{Ms}*lcp1* mutant revealed a band at around 6 kb, corresponding to the larger size of the hygromycin resistance cassette replacing ^{Ms}*lcp1* (Figure 3.4A, B, Lane 3). The Δ^{Ms}*lcp1* mutant DNA also showed a band for *murC*, at around 1.5 kb, confirming that the DNA is indeed from *M. smegmatis* cells (Figure 3.4B, Lane 4). The increase in size between

the bands from both *M. smegmatis* WT and Δ*Mslcp1* genomic DNA indicate that *Mslcp1* was successfully knocked out in this conditional mutant.

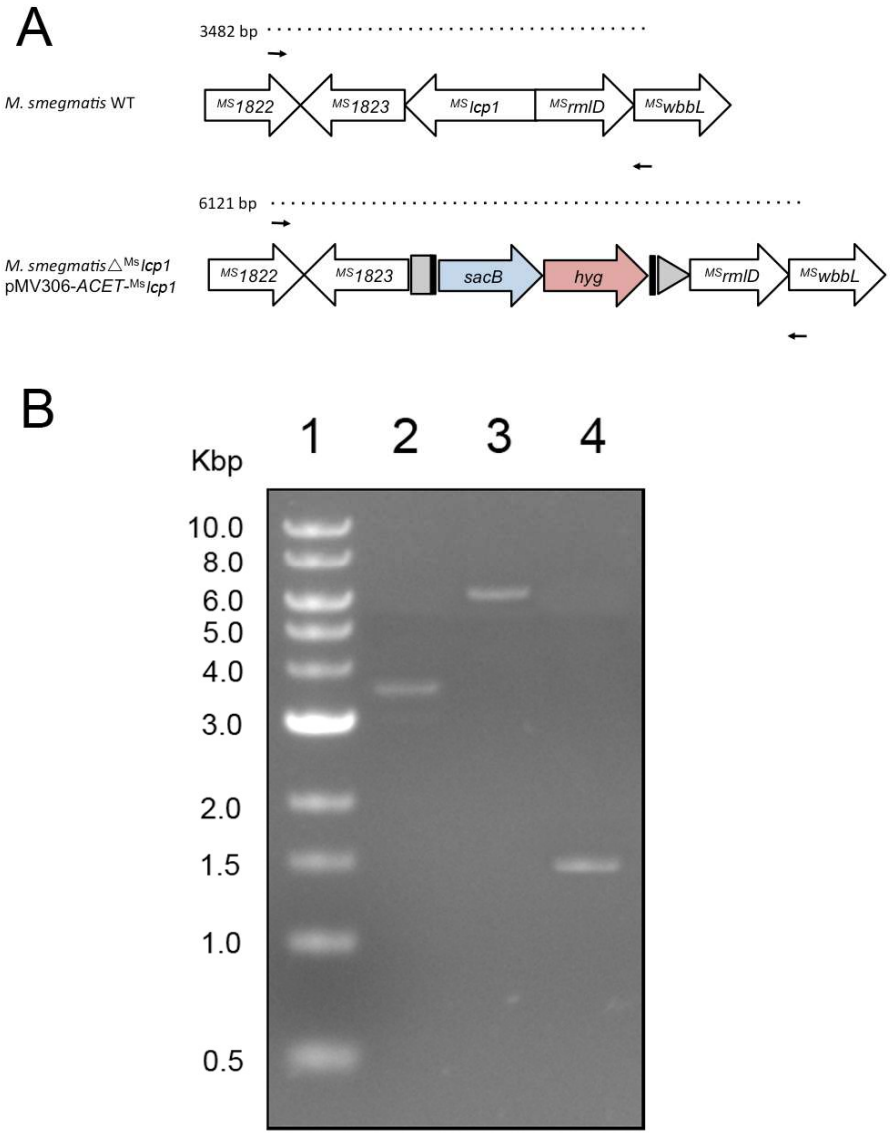

Figure 3.4: Confirmation of generation of a Δ*Mslcp1* conditional knock-out mutant. A) Using primers *Mslcp1-Van*91I L-LL and R-RR (Section 5.5.1.), the amplified region from *M. smegmatis* WT would be 3482 bp, whereas in the *M. smegmatis* Δ*Mslcp1* conditional mutant, the increased size of the hygromycin resisitance cassette would cause this region to be 6121 bp in size. B) Agarose gel electrophoresis of PCR from both *M. smegmatis* WT and Δ*Mslcp1*. Lane 1 shows a 1 kb DNA ladder, Lane 2 shows *M. smegmatis* WT with flanking primers, Lane 3 shows *M. smegmatis* Δ*Mslcp1* with flanking primers and Lane 4 shows *M. smegmatis* Δ*Mslcp1* with *murC* amplifying primers. Lane 2 shows the correct sized band of around 3.5 kb, whereas Lane 3 corresponds to the size of the region containing the hygromycin resistance cassette. Lane 4 shows the correct sized band for *murC*, indicating that the conditional mutant DNA is indeed *M. smegmatis* and that the conditional mutant is indeed a true Δ*Mslcp1* knock out.

The *M. smegmatis* Δ^{Ms}*lcp1* mutant cells grown on agar plates in the presence of acetamide showed a difference in colony morphology when compared to *M. smegmatis* WT cells on the same plates (Figure 3.5A). The essentiality of ^{Ms}*lcp1* was further investigated through subsequent passages in liquid media, in the absence of the acetamide inducer (Figure 3.5B). In the absence of acetamide, there was a visibly rapid loss in cell viability, indicating the requirement of the second pMV306-*ACET*- ^{Ms}*lcp1* copy of the gene is required for cell growth, which supports the hypothesis of the essentiality of ^{Ms}*lcp1* (Figure 3.5B).

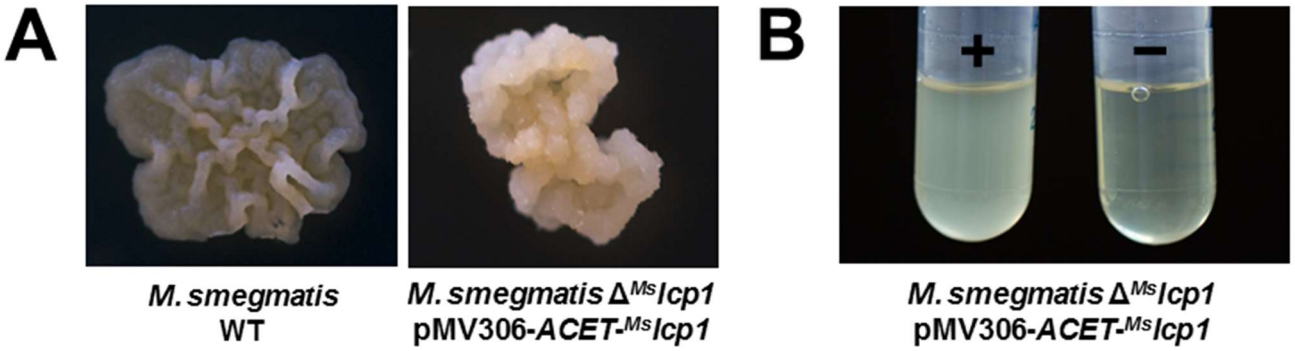

Figure 3.5: Essentiality of ^{Ms}*lcp1* in *M. smegmatis* on solid and liquid media. A) The growth and colony formation of *M. smegmatis* WT and Δ^{Ms}*lcp1* mutant on tryptic soy agar (TSA). Pictures show a clear difference in shape and fitness of colonies, even though the mutant is grown in the presence of acetamide, inducing the second copy of ^{Ms}*lcp1*. B) Growth of *M. smegmatis* Δ^{Ms}*lcp1* mutant in tryptic soy broth (TSB) in the presence (+) and the absence (−) of acetamide. Without acetamide, the Δ^{Ms}*lcp1* mutant rapidly declines in cell viability, shown by lack of growth in TSB.

The *M. smegmatis* Δ^{Ms}*lcp1* mutant was grown both in the presence and absence of acetamide over 55 h, with samples of both periodically being plated on TSA containing acetamide in order to check for colony forming units (CFU) (Figure 3.6). The Δ^{Ms}*lcp1* cells in the absence of acetamide rapidly begin to decline in CFU count after 20-25 h, whereas those grown in the presence of acetamide continue to grow unhindered (Figure 3.6). After around 24 h, there is a whole log difference in CFU between the Δ^{Ms}*lcp1* and WT cells. After the 55 h growth period the there is a 4 log difference between CFU counts for both samples, with WT cells reaching 10^8 CFU, whereas Δ^{Ms}*lcp1* cells drop to a CFU count of below 10^4 (Figure 3.6). Since the

number of CFU is decreasing over time in the absence of acetamide, it suggests that the cells are dying rather than becoming static, indicating that the loss of ^{Ms}Lcp1 to the cell is lethal (Figure 3.6). Because of the rapid loss in cell viability of the the Δ*^{Ms}lcp1* mutant in the absence of acetamide, it was not possible to phenotypically characterise the expected resultant lesions in the cell wall, due to a lack of available cellular material. We therefore decided to biochemically characterise recombinant ^{Mtb}Lcp1.

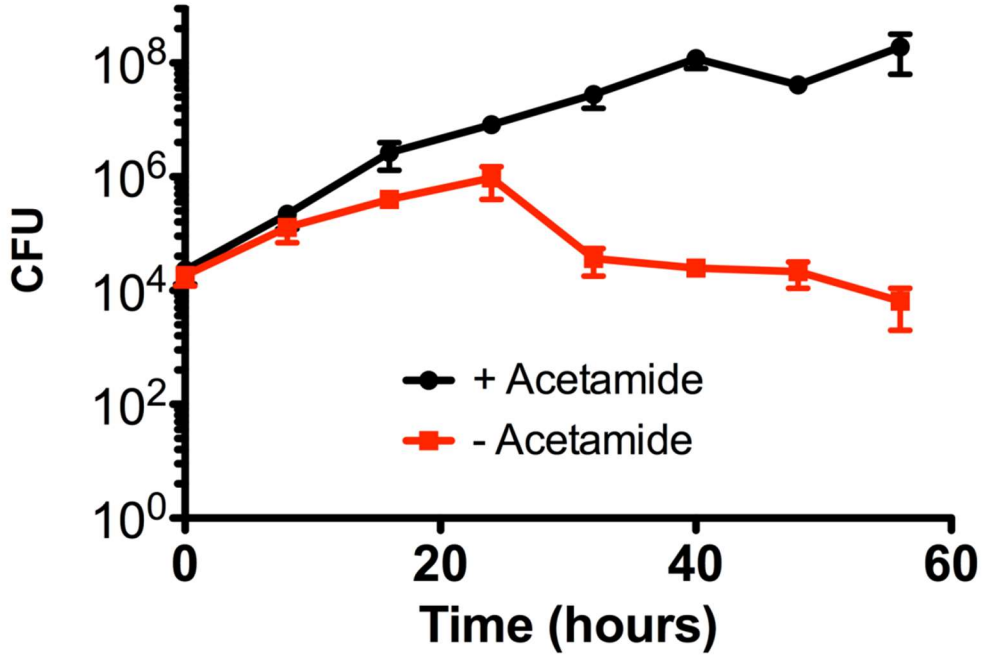

Figure 3.6: ***M. smegmatis* Δ^{Ms}*lcp1* CFU growth curve, in both the presence and absence of acetamide.** When acetamide is removed from the growth medium, the cells immediately begin to die, as evidenced by the rapid reduction of CFU over the first 20 h. After 55 h, the cells in the absence of acetamide are well in decline, whereas the cells grown with acetamide continue to grow unhindered across the entire timespan, indicating that the cells are healthy and in optimum growth conditions, showing that the lack of acetamide and therefore ^{Ms}Lcp1, results in increased cell death.

3.2.3. Purification of recombinant ^{Mtb}Lcp1 protein

An N-terminally His-tagged ^{Mtb}Lcp1 (minus transmembrane domain) expression vector (pET28b-^{Mtb}Lcp1) was constructed (Section 5.18.1.). Four identified amino acid residues with a predicted role in polyisoprenoid-phosphate binding (Section 3.2.1.) were mutated into alanine through site-directed mutagenesis (SDM) of the pET28b-^{Mtb}Lcp1 vector (Section 5.6.). This gave rise to the constructs pET28b-^{Mtb}Lcp1 WT, R70A, R118A, R236A and R238A. All of these vectors were overexpressed in *E. coli* BL21 (DE3) (Section 5.4.1.) and purified through immobilised metal ion affinity chromatography (IMAC) (Figure 3.7) and subsequent anion exchange chromatography, if required (Section 5.11.).

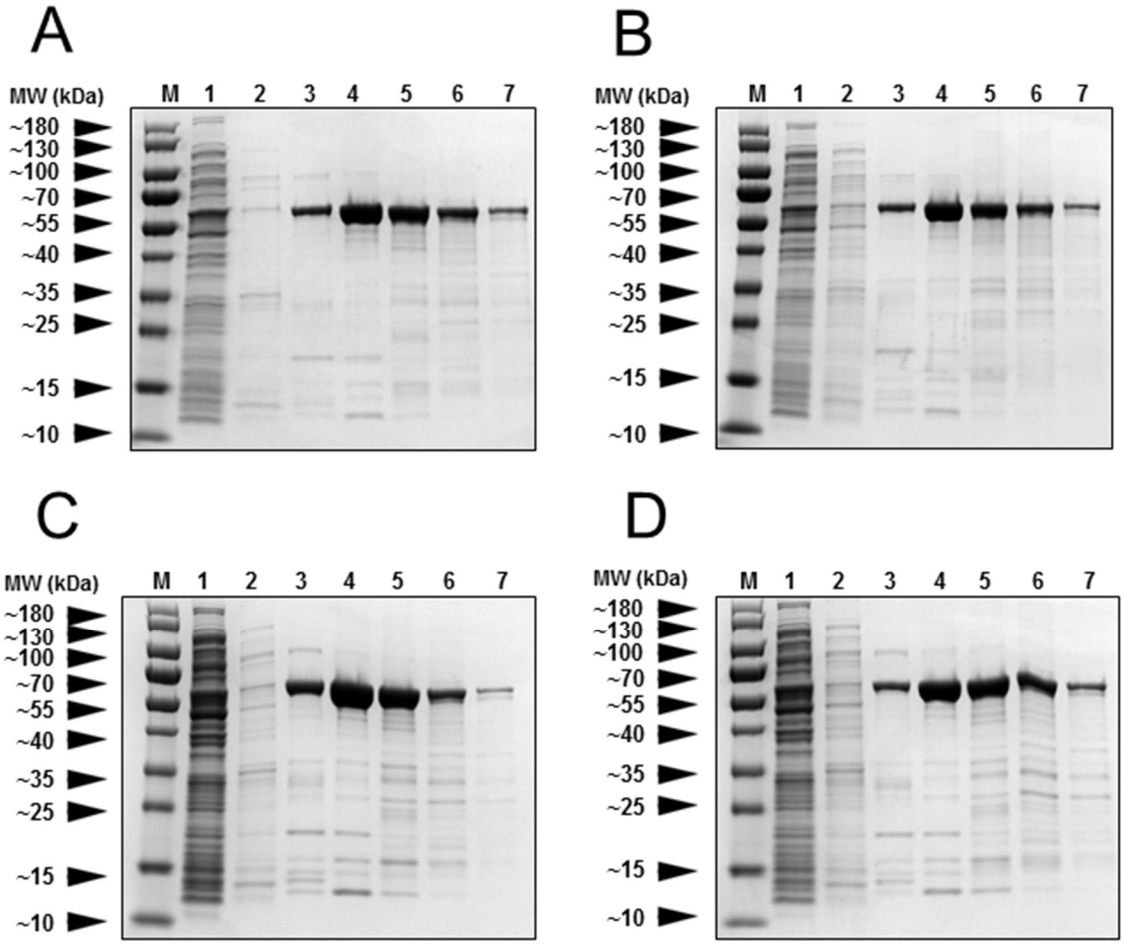

Figure 3.7: SDS-PAGE of ^{Mtb}Lcp1. A) ^{Mtb}Lcp1 WT, B) ^{Mtb}Lcp1R70A, C) ^{Mtb}Lcp1 R118A, D) ^{Mtb}Lcp1 R236A. ^{Mtb}Lcp1 R238A was unable to be expressed or purified despite numerous attempts of optimisation. Lanes shown are protein marker (M), clarified lysate flow through (1), 50 mM imidazole elution (2), 75 mM imidazole elution (3), 100 mM imidazole elution (4), 200 mM imidazole elution (5), 250 mM imidazole elution (6) and 500 mM imidazole elution (7). Bands visible around 50-60 kDa represent recombinant ^{Mtb}Lcp1.

^{Mtb}Lcp1 WT, R70A, R118A and R236A all purified abundantly (Figure 3.7) and were stable at concentrations of 50 mg/mL and higher, however, it was not possible to obtain recombinant ^{Mtb}Lcp1 R238A, despite numerous attempts at optimising expression and purification. Recombinant purified ^{Mtb}Lcp1 WT and the site-directed mutants (SDMs) (R70A, R118A and R236A) were all stable when dialysed into 100 % dH_2O for a period of 1 to 2 weeks at 4 °C.

Since it was observed that *S. pneumoniae* Cps2A co-eluted with a polyisoprenoid-phosphate (Section 3.1.) and the ^{Mtb}Lcp1 homology model predicted a hydrophobic binding cavity (Section 3.2.1.), we investigated whether ^{Mtb}Lcp1 also co-eluted with a polyisoprenoid. We performed an organic solvent extraction of ^{Mtb}Lcp1 and analysed the extracted sample by TLC (Figure 3.8), which revealed a spot with the same R_f value as the known standard decaprenyl-monophosphate (Section 5.21.). These extractions were also submitted for mass spectrometric analysis (Section 5.21.), which clearly indicated an m/z of 777, which directly correlates with the expected mass of decaprenyl-1-monophosphate (Figure 3.9). This indicates that ^{Mtb}Lcp1 also co-purifies with a bound polyisoprenoid-phosphate, in a similar fashion to other LCP proteins (Section 3.1.). We repeated an identical extraction and analysis approach for each of the three recombinant mutant ^{Mtb}Lcp1 proteins using the same amount of protein as the original experiment. Surprisingly, all three mutants also co-eluted with decaprenyl-monophosphate as indicated by TLC and mass spectrometric analysis (Figure 3.8, Figure 3.9).

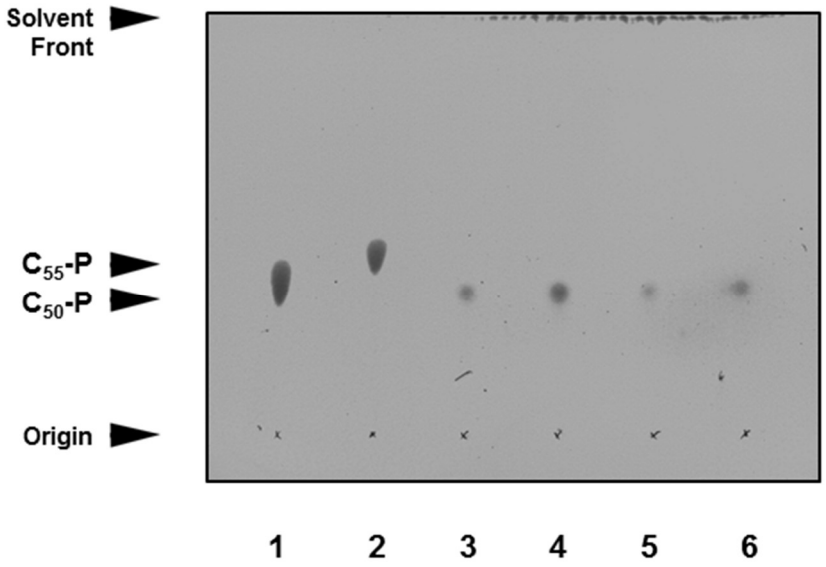

Figure 3.8: TLC plate showing ^{Mtb}Lcp1 WT and SDM bound lipid. Decaprenyl-monophosphate (Lane 1, C_{50}-P) and undecaprenyl-monophosphate (Lane 2, C_{55}-P) shown running alongside solvent extracted samples from ^{Mtb}Lcp1 WT (Lane 3), R70A (Lane 4), R118A (Lane 5) and R236A (Lane 6) using the solvent system $CHCl_3$:CH_3OH:NH_4OH:H_2O (65:25:0.5:3.6, v/v/v/v) (Section 5.14.). All ^{Mtb}Lcp1 extractions ran at the same R_f value as decaprenyl-monophosphate, indicating that ^{Mtb}Lcp1 WT and SDMs co-purify with a bound polyisoprenoid, likely to be decaprenyl-monophosphate.

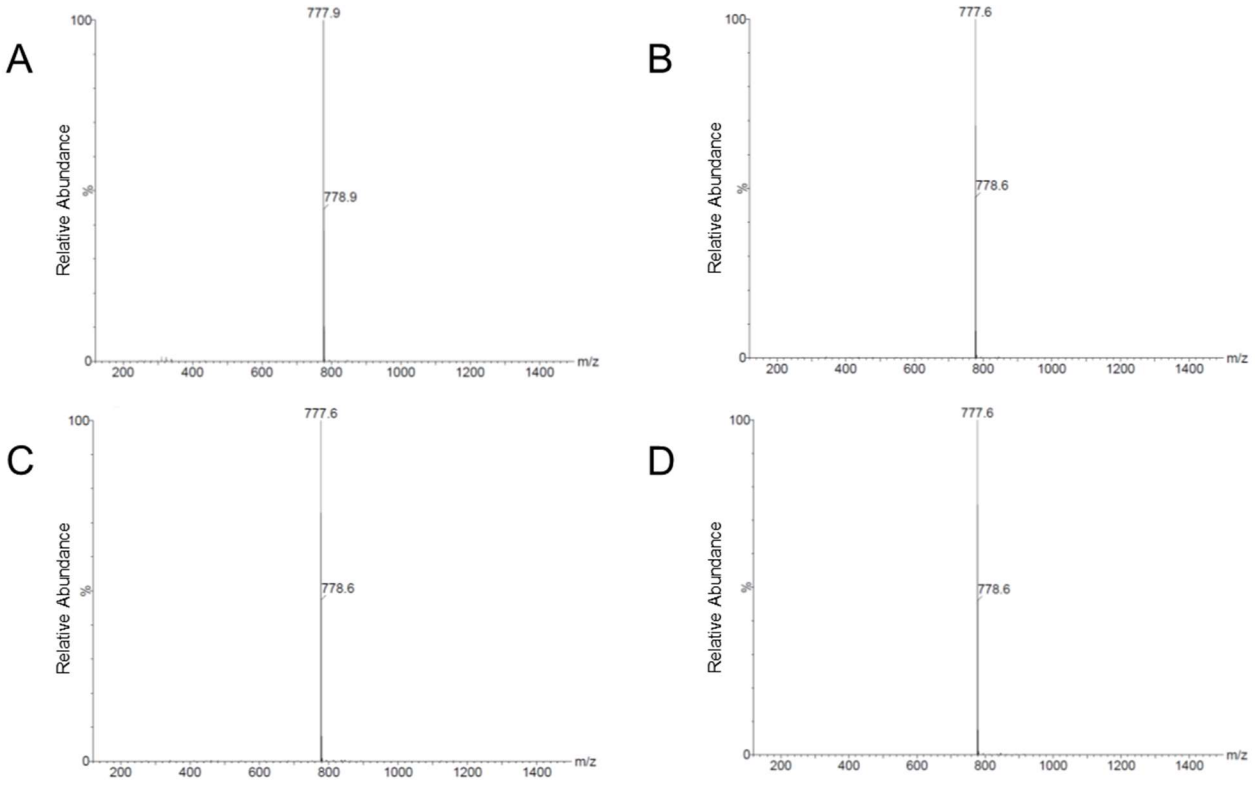

Figure 3.9: Mass spectra of ^{Mtb}Lcp1 WT and SDM. A single peak at 777 m/z is seen, which corresponds to the expected mass of decaprenyl-1-monophosphate in the negative ion mode. A) ^{Mtb}Lcp1 WT. B) ^{Mtb}Lcp1 R70A. C) ^{Mtb}Lcp1 R118A. D) ^{Mtb}Lcp1 R236A.

Since the ^{Mtb}Lcp1 SDMs harbour mutated residues that are predicted to be important for the binding of decaprenyl-monophosphate, it was important for us to eliminate the possibility that these mutated recombinant proteins were folded incorrectly. Therefore we conducted a far-UV circular dichroism (CD) (Section 5.22.) analysis of both ^{Mtb}Lcp1 WT and each of the three SDMs.

An overlay of the spectra from each sample of ^{Mtb}Lcp1 (Figure 3.10) indicated that each protein was folded with the same degree of secondary structure. The only major differences in spectra were at the low wavelength end (190 nm) and were attributed to system noise.

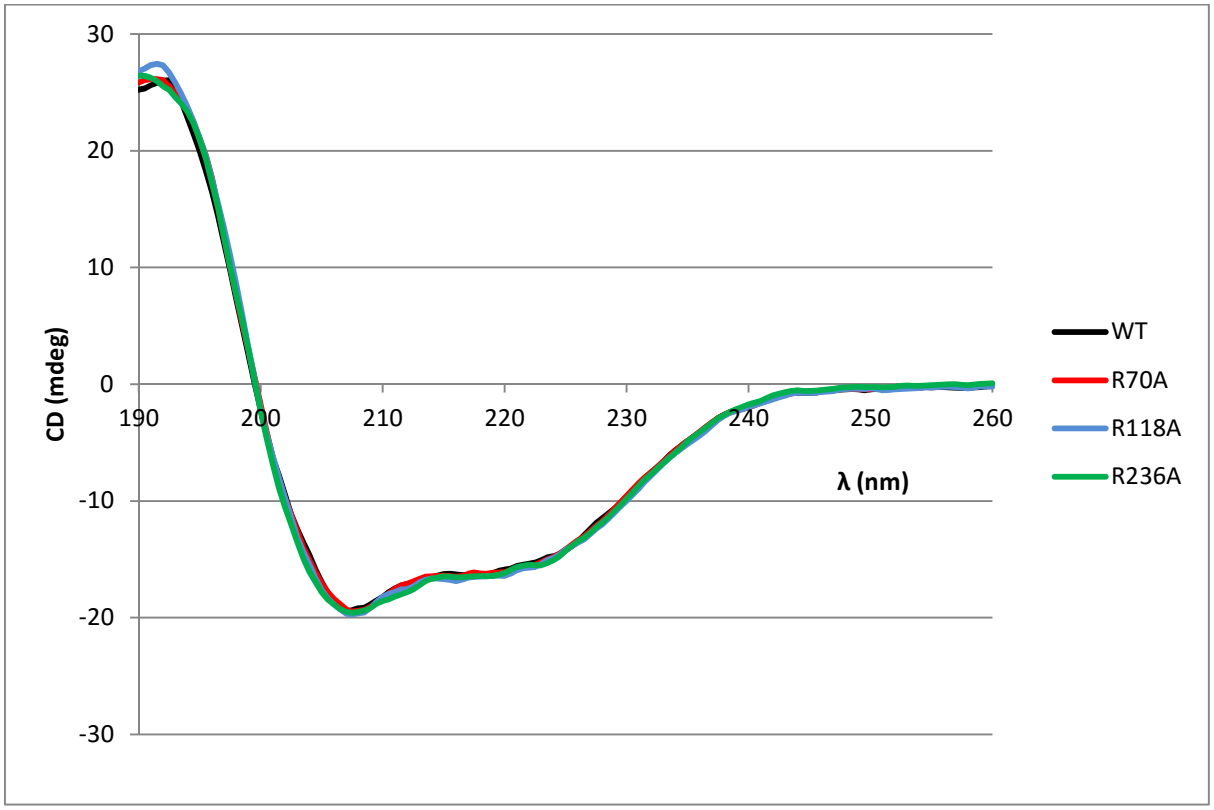

Figure 3.10: Far-UV CD spectra of ^{Mtb}Lcp1 WT (black) and SDMs (R70A red, R118A blue, R236A green), between 260 and 190 nm. Each spectrum overlays extremely well, indicating that each protein is folded into the same secondary structure as ^{Mtb}Lcp1 WT, and thus each SDM protein is folded correctly.

Further biophysical characterisation of ^{Mtb}Lcp1 could be obtained through analytical ultracentrifugation (AUC), in order to probe how ^{Mtb}Lcp1 self-assembles in solution (Section 5.22.). We performed a set of sedimentation velocity experiments using ^{Mtb}Lcp1 WT and SDMs alone, and also ^{Mtb}Lcp1 WT in the presence of varying concentrations of a synthetic carbohydrate ligand, compound 1, one of a panel of analogues representing various stages of the mycobacterial LU (Figure 3.13), synthesised to investigate ^{Mtb}Lcp1 and its interaction with its chemically complex and labile natural substrate (Section 3.2.4.). When the ^{Mtb}Lcp1 proteins were in solution on their own, the distribution (c (s)) of the sedimentation coefficient (S) indicated a single major molecular species for all samples at 3.2-3.3 S (Figure 3.11), which correspond to apparent molecular weights (Table 3.1) fitting the calculated molecular weight of ^{Mtb}Lcp1, of around 48.5 kDa, indicating ^{Mtb}Lcp1 WT and SDMs are likely monomeric in solution. When increasing ligand concentration was added to ^{Mtb}Lcp1 WT, there was a slight increase in S with increased ligand concentration (from 3.2 S to 3.5 S) (Figure 3.12), however, this did not cause a significant change in apparent molecular weights (Table 3.2) and it is probable that ^{Mtb}Lcp1 remains a monomer in solution in the presence of carbohydrate ligands, such as the AG LU.

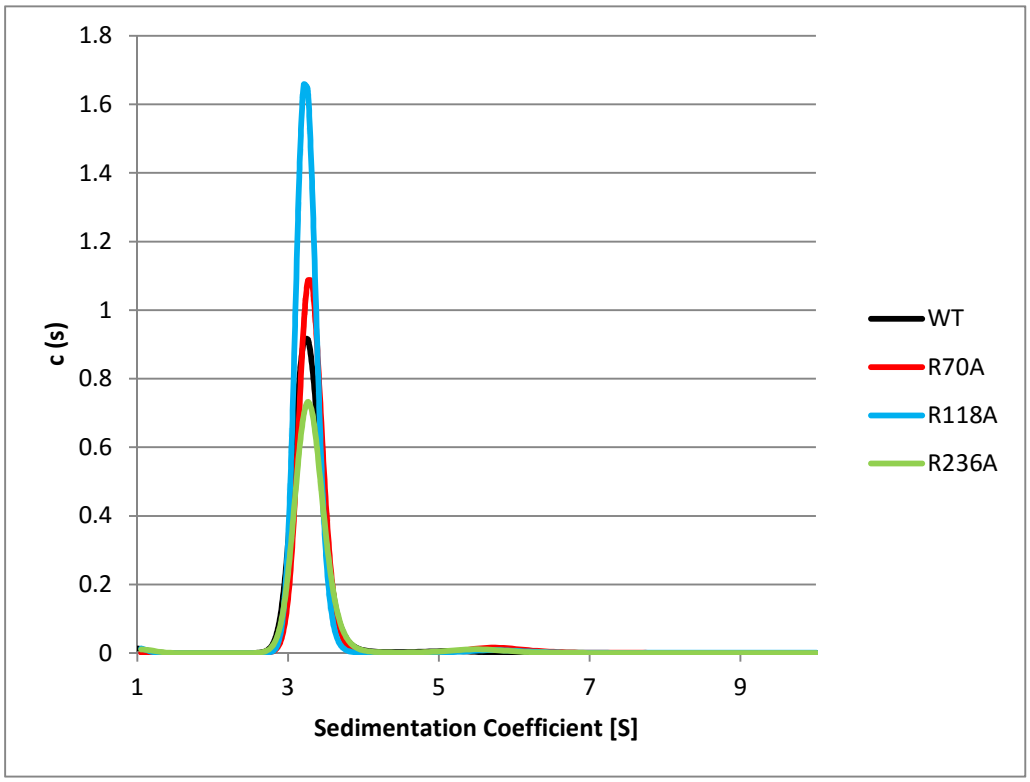

Figure 3.11: Self-assembly analysis of ^{Mtb}Lcp1 WT (black) and SDMs (R70A red, R118A blue, R236A green) by AUC in sedimentation velocity mode. The peaks for each protein fell within 3.2 S and 3.3 S and gave apparent molecular weights (Table 3.1) that fit to the calculated molecular weight of ^{Mtb}Lcp1 in its monomeric state.

Table 3.1: Apparent molecular weights of ^{Mtb}Lcp1 WT and SDMs calculated through sedimentation velocity experiments by AUC. Each molecular weight can be fitted to the calculated molecular weight of ^{Mtb}Lcp1, 48.5 kDa.

Sample	Molecular Weight (kDa)
^{Mtb}Lcp1 WT	50.8
^{Mtb}Lcp1 R70A	47.5
^{Mtb}Lcp1 R118A	48.4
^{Mtb}Lcp1 R236A	50.0

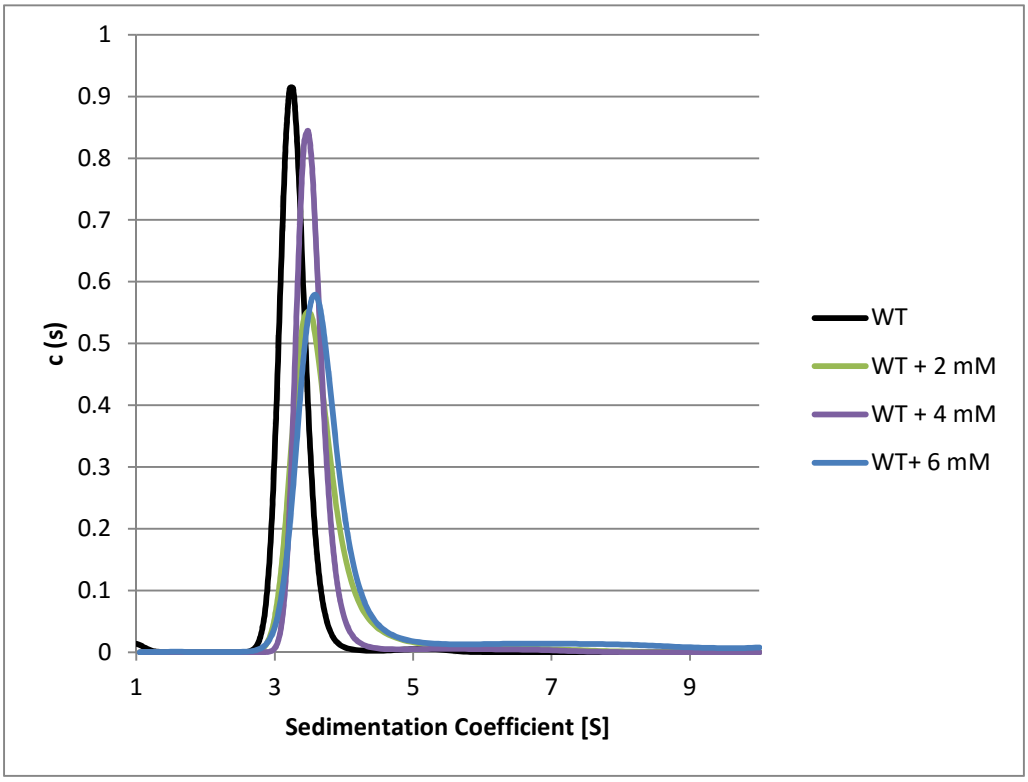

Figure 3.12: Self-assembly analysis of ^{Mtb}Lcp1 WT (black) and increasing carbohydrate ligand (compound 1) concentration (2 mM green, 4 mM purple, 6 mM blue) by AUC in sedimentation velocity mode. The peak for WT protein on its own was at 3.2 S, with increasing ligand concentration raising the sedimentation coefficient to 3.5 S. However, the apparent molecular weight of protein with each ligand concentration (Table 3.2) did not significantly alter from the calculated molecular weight of monomeric ^{Mtb}Lcp1.

Table 3.2: Apparent molecular weights of ^{Mtb}Lcp1 WT with increasing concentration of carbohydrate ligand (compound 1) calculated through sedimentation velocity experiments by AUC. Each molecular weight can be fitted to the calculated molecular weight of ^{Mtb}Lcp1, 48.5 kDa.

Sample	Molecular Weight (kDa)
^{Mtb}Lcp1 WT	50.8
^{Mtb}Lcp1 WT + 2 mM compound 1	54.5
^{Mtb}Lcp1 WT + 4 mM compound 1	49.8
^{Mtb}Lcp1 WT + 4 mM compound 1	51.2

3.2.4. Carbohydrate binding of linker unit mimetics to ^{Mtb}Lcp1

Our laboratory has previously synthesised neo-glycolipid acceptors, modelled upon motifs found within mycobacterial AG, in order to probe the ligand binding abilities of proteins involved in cell wall structure, as well as any enzymatic function associated with them (Alderwick *et al*., 2008, Alderwick *et al*., 2011a). Therefore, it was possible to synthesise an array of analogues of the mycobacterial LU (Compounds 1-4) (Figure 3.13), which could be used to investigate ^{Mtb}Lcp1 and its interaction with its putative natural substrate decaprenyl-pyrophosphoryl-GlcNAc-Rha*p*-Gal*f*$_{30}$. Compounds 1-4 are LU mimetics synthesised to represent a variety of stages of LU biosynthesis (Section 1.2.3.2.), since the physiological substrate is chemically complex and labile, they offered an alternative means to assess this important biochemical interaction. Unstable decaprenyl-pyrophosphate groups were substituted with C_8 alkyl chains, since compounds containing these have previously been shown to be successful acceptors (Lee *et al*., 1997). Compound 1 (Rha-GlcNAc-*O*-C_8) is a representation of glycolipid (GL)-2 and resembles the disaccharide L-Rha*p*-(1→3)-α-D-GlcNAc of the mycobacterial LU, with a shortened representation of the decaprenyl-pyrophosphoryl moiety. Compound 2 (Gal*f*$_2$-Rha-GlcNAc-*O*-C_8) represents GL-4 and compound 3 (Gal*f*$_3$-Rha-GlcNAc-*O*-C_8) contains a single Gal*f* residue addition to compound 2, representing the increasing galactan chain. However, compound 4 (Gal*f*-Rha-*O*-C_8) represents the terminal L-Rha*p* of the LU, attached to a single Gal*f* residue (Figure 3.13).

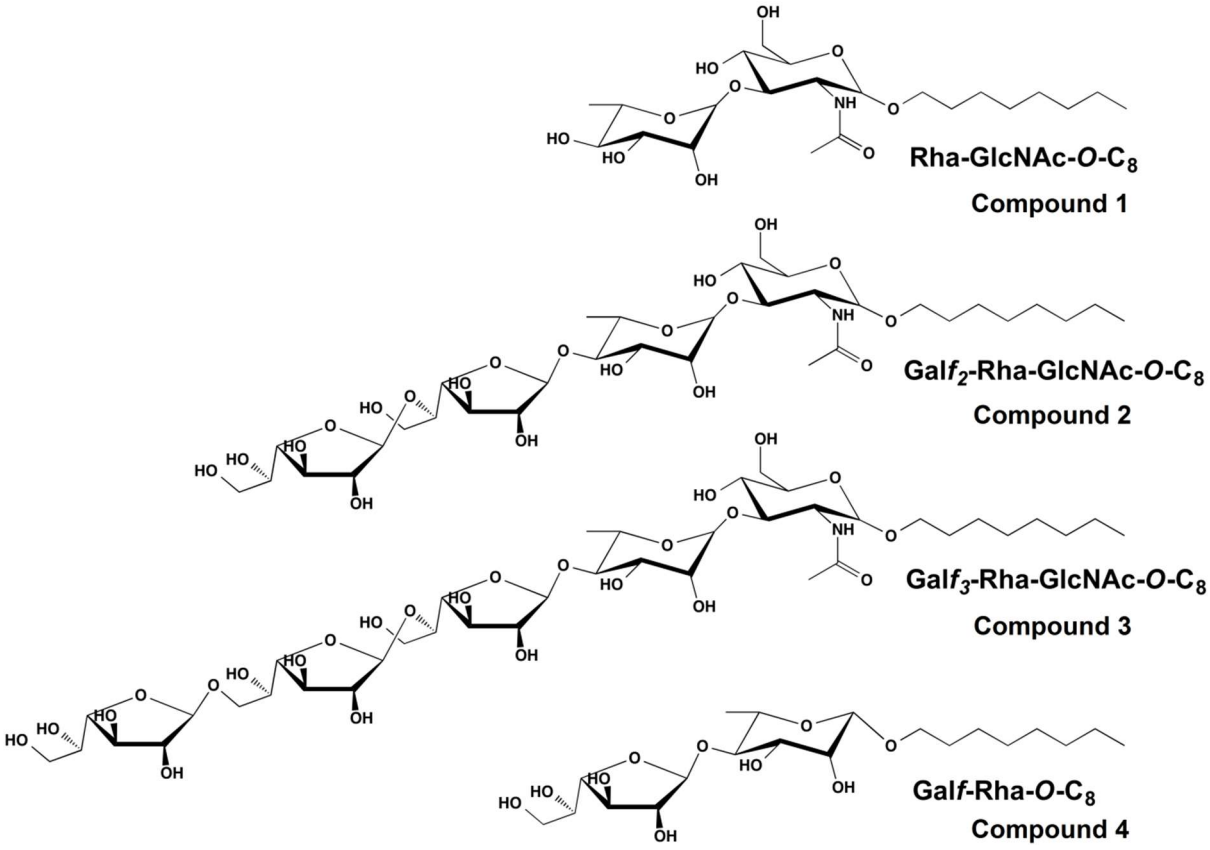

Figure 3.13: Structures of novel chemical scaffolds used as ligands to probe ^{Mtb}Lcp1 interaction with mycobacterial LU. Compounds 1 and 2 represent GL-2 and GL-4 of LU/galactan assembly respectively (Section 1.2.3.2.). Compound 3 represents an extension of GL-4 by a single Gal*f* residue. Compound 4 represents the terminal Rha*p* of the mycobacterial LU bound to a single Gal*f* residue.

We used intrinsic tryptophan fluorescence (ITF) to probe the binding of ^{Mtb}Lcp1 WT with compounds 1-4 (Section 5.15.). Binding data was fitted to a single site saturation model to yield equilibrium binding constants (K_d) for each compound (Figure 3.14). The K_d for compound 1 was 57.68 µM, which increased ~10 fold in affinity with the addition of 2 Gal*f* residues, providing a K_d of 5.13 µM for compound 2. Further elongation of the galactan chain with a third Gal*f* residue in compound 3 resulted in a K_d of 20.39 µM. Compound 4 displayed a much lower binding affinity, providing a K_d of 97.61 µM, as well as ~10 fold decrease in binding capacity (B_{max}).

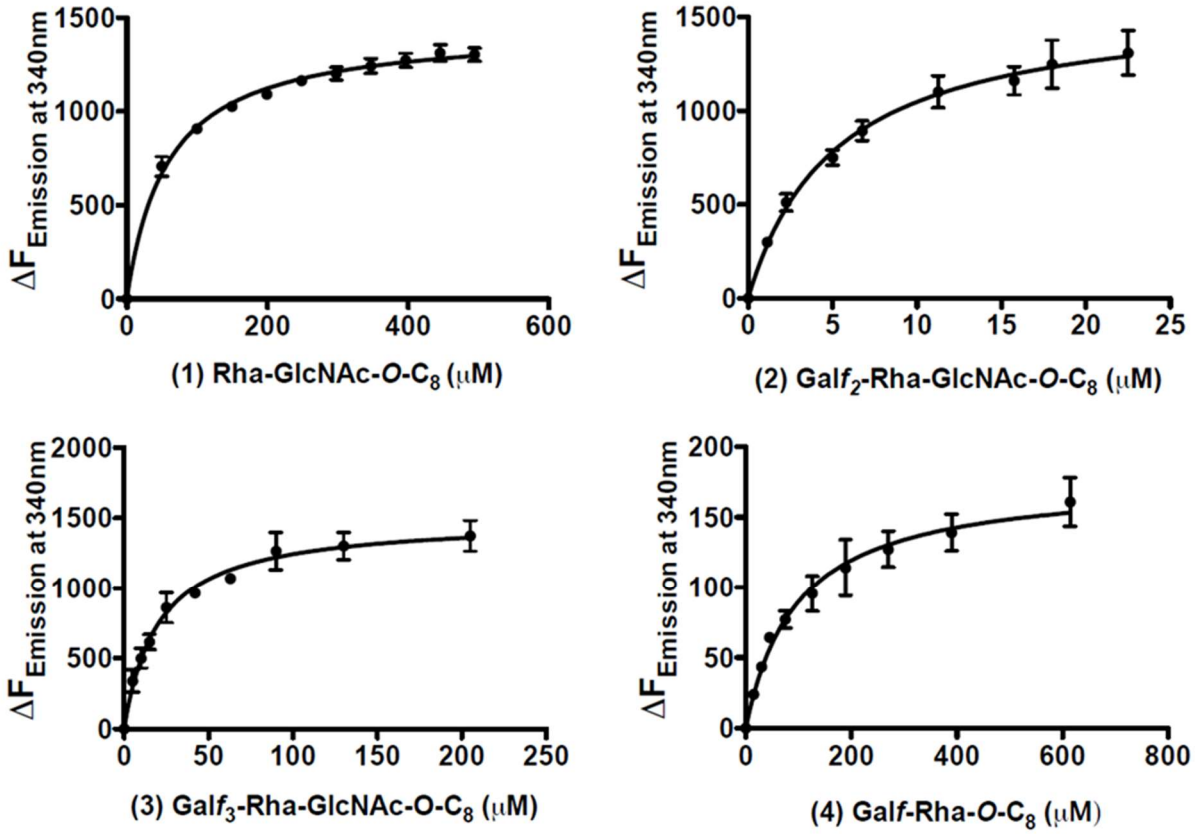

Figure 3.14: Saturation binding experiments using ITF to assess ^{Mtb}Lcp1 WT ligand binding. Each experiment used 26 nM ^{Mtb}Lcp1, all experiments performed in triplicate. (1) Binding of compound 1 (Rha-GlcNAc-*O*-C_8) gives a K_d of 57.68 μM. (2) Compound 2 (Gal*f*$_2$-Rha-GlcNAc-*O*-C_8) binding gives a K_d of 5.13 μM. (3) Binding of compound 3 (Gal*f*$_3$-Rha-GlcNAc-*O*-C_8) gives a K_d of 20.39 μM. (4) Compound 4 (Gal*f*-Rha-*O*-C_8) gives no significant binding affinity, providing a K_d of 97.61 μM and a much reduced binding capacity (B_{max}).

The data shows that ^{Mtb}Lcp1 is capable of binding the disaccharide compound 1, representing the LU. It also shows that ^{Mtb}Lcp1 has a high affinity for LU with additional Gal*f* residues attached (compound 2 and 3), suggesting that the presence of Gal*f* residues is important in substrate binding. The fact that compound 4 displays a significantly lower binding affinity to ^{Mtb}Lcp1 suggests that D-GlcNAc is an important component of ^{Mtb}Lcp1 substrate recognition (Figure 3.14). However, when the selected arginine residues were mutated to alanine, a change in binding was seen for both compounds 1 and 4 (compounds 2 and 3 were omitted due to depletion of available ligands) (Figure 3.15).

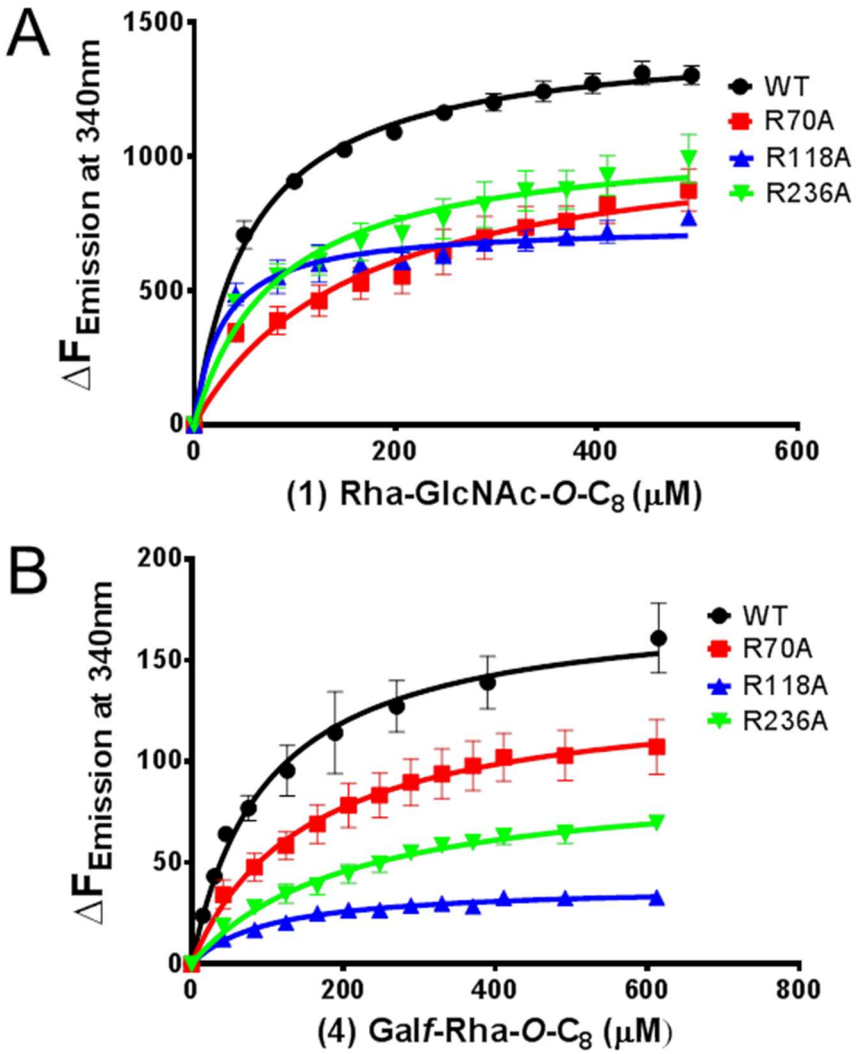

Figure 3.15: Saturation binding experiments using ITF to assess ^{Mtb}Lcp1 SDM binding. For compound 1 (A), the K_d of each of the ^{Mtb}Lcp1 variants were 57.68 μM for WT, 161.5 μM for R70A, 27.76 μM for R118A and 82.43 μM for R236A. However, the B_{max} for each mutant was lower than the 1,448 that WT displayed, with R70A being 1,102, R118A being 744.5 and R236A being 1,077. For compound 4 (B), the K_d of each of the ^{Mtb}Lcp1 variants were 97.61 μM for WT, 150.1 μM for R70A, 99.07 μM for R118A and 206.4 μM for R236A. Again, each mutant ^{Mtb}Lcp1 displayed a lower B_{max} than the 177.6 for WT, with R70A being 135.4, R118A being 38.61 and R236A being 92.79. The data for both compounds with each ^{Mtb}Lcp1 variant shows that mutation of these residues results in a decrease in overall binding of the LU mimetics.

The mutation at R70 resulted in a binding affinity around three fold lower for compound 1 (161.5 μM) than for ^{Mtb}Lcp1 WT, as well as a lower B_{max} of 1,102, compared to that of 1,448 for WT (Figure 3.15A). ^{Mtb}Lcp1 R70A shows a similar trend in binding compound 4, whereby it displays a higher K_d than WT of 150.1 μM and a slight reduction in the already significantly reduced binding capacity (Figure 3.15B). Interestingly, ^{Mtb}Lcp1 R118A displays

a K_d of almost half that of WT when binding compound 1 (27.76 μM), yet also shows a B_{max} of around half that of WT (744.5), suggesting that this mutation increases the binding affinity, whilst lowering the overall binding capacity (Figure 3.15A). When ^{Mtb}Lcp1 R118A binds compound 4, the K_d remains similar to that of WT (99.07 μM), however the binding capacity is again dramatically reduced to 38.61 (Figure 3.15B). Finally, ^{Mtb}Lcp1 R236A displays a lower binding affinity than WT (82.43 μM) and a lower B_{max} of 1,077 when binding compound 1, behaving in a similar fashion to ^{Mtb}Lcp1 R70A (Figure 3.15A). Yet when binding compound 4, ^{Mtb}Lcp1 R236A shows a K_d around twice that of WT of 206.4 μM and a binding capacity almost half of WT (92.79) (Figure 3.15B). All of these results suggest that mutation of these specific residues in ^{Mtb}Lcp1 has a detrimental effect on LU binding.

3.2.5. ^{Mtb}Lcp1 cell-free activity assay

We developed a cell-free biochemical assay in order to assess whether ^{Mtb}Lcp1 was capable of ligating AG to PG. Since ^{Mtb}Lcp1 was predicted to contain an N-terminal transmembrane helix (Section 3.2.1.), attachment is likely to occur at or around the cytoplasmic membrane, we therefore proposed that membranes prepared from mycobacteria would contain some endogenous AG to PG attachment activity, which could be investigated by using a [^{14}C]-carbohydrate radiolabelled assay. A previous method used to monitor the synthesis of radiolabelled decaprenyl-pyrophosphoryl intermediates of mycobacterial cell wall biosynthesis in our laboratory was modified to directly investigate the attachment of [^{14}C]-AG intermediates to PG, *in vitro* (Birch *et al*., 2009). Mycobacterial membranes (Section 5.23.1.) and cell wall fraction (rich in endogenous cell wall biosynthetic enzymes) (Section 5.23.2.) were mixed with [^{14}C]-galactopyranose, with the endogenous Glf mutase activity within the assay mix enough to allow for direct incorporation of [^{14}C]-Gal*f* into cell wall intermediates, as was previously demonstrated (Birch *et al*., 2009). Purified PG (PPG) from *M. smegmatis* (Section 5.23.3.), free of any covalently bound AG, was also added before initiation of the assay by addition of recombinant ^{Mtb}Lcp1 WT and SDMs. The products of the assay were then solvent extracted through a complex fractionation process (Section 5.23.4.) before analysis. The first extraction from the assay mix was the aqueous fraction, which contained any remaining unused radiolabel.

The organic fraction contained samples extracted using $CHCl_3$:CH_3OH (2:1, v/v), which were analysed through TLC (Figure 3.16A) (Section 5.23.5.). This analysis revealed three major bands migrating towards the bottom of the plate, with bands corresponding to GL-3 and GL-4 (Section 1.2.3.2.) and decaprenyl-pyrophosphoryl-GlcNAc-Rha*p*-galactan, which remained at the origin. Interestingly, the intensity of these bands decreased with an increase in the amount of ^{Mtb}Lcp1 present in the sample (Figure 3.16B). When the assay was repeated in the absence of PPG, the intensity of all bands appeared to remain at similar levels as those in the absence of ^{Mtb}Lcp1 (Figure 3.16A, Lanes 2 and 5). This was corroborated by densitometry of the TLC plate, and analysis of the radioactivity of each band (Figure 3.16B). Therefore, the data shows that with the addition of ^{Mtb}Lcp1, the relative amount of [^{14}C]-Gal*f* labelled material is reduced in the organic fraction. This trend is seen not just for ^{Mtb}Lcp1 WT (Figure 3.16) but also for all SDMs in repeated assays.

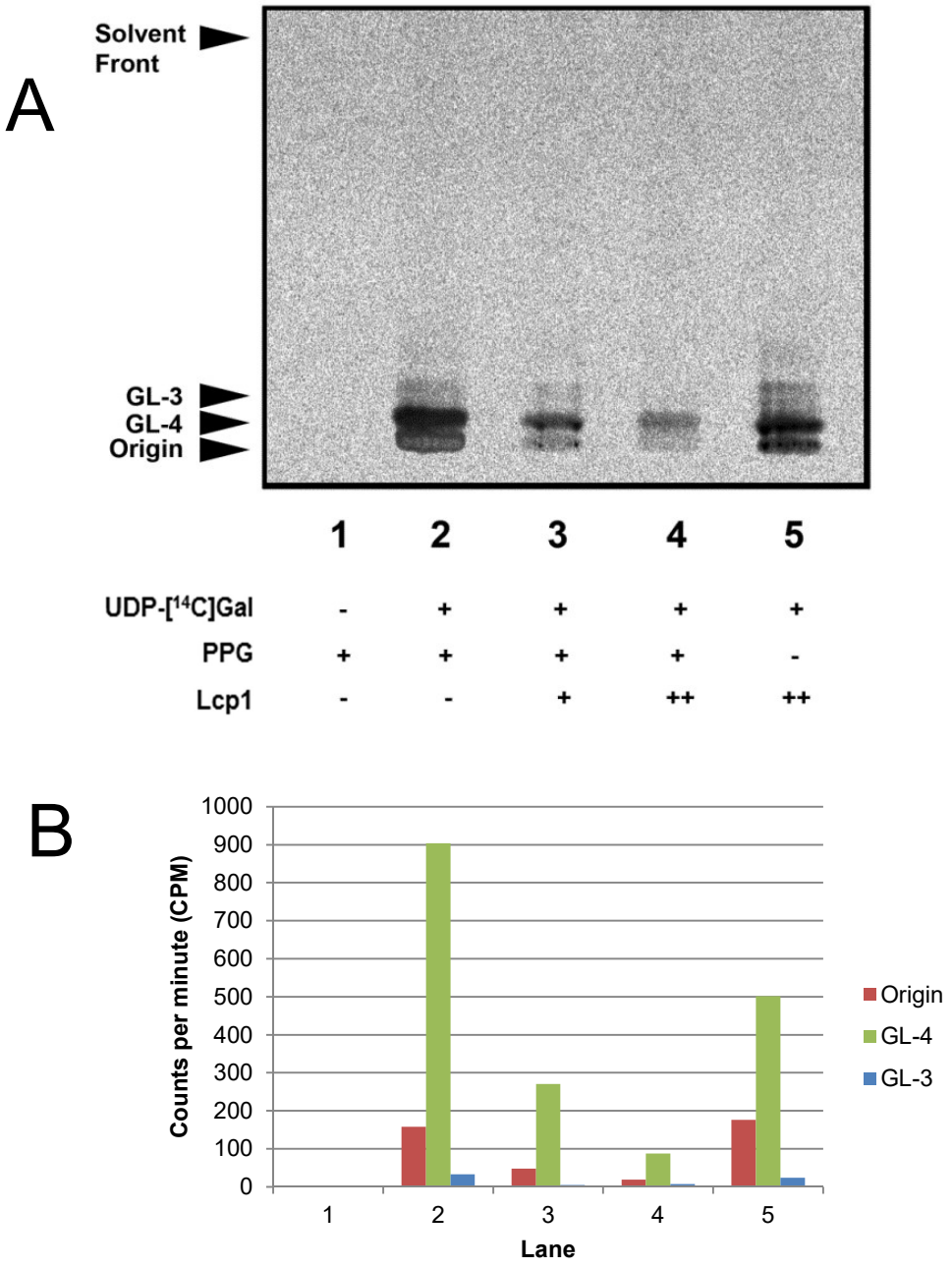

Figure 3.16: Organic solvent extraction analysis. A) Organic solvent extracted fraction from [^{14}C]-labelled assay analysed by TLC (Section 5.14.), developed in $CHCl_3$:CH_3OH:NH_4OH:H_2O (65:25:0.5:3.6, v/v/v/v). As Lcp1 concentration increases, the bands corresponding to GL-4 and GL-3 and mature galactan at the origin, appear less intense, suggesting radioactivity and therefore substrates are being utilised. When there is an absence of PPG, the bands are much more intense, regardless of Lcp1 concentration, suggesting reduction in radioactivity moving through the system. Bands present are a result of ^{Mtb}Lcp1 WT, but are representative of all SDMs. B) Radioactivity present in all bands, showing the true counts per minute (CPM) within each band from the total radioactivity in the sample. The radioactivity within each band decreases when the amount of Lcp1 is increased. However, in the absence of PPG in lane 5, the level of radioactivity within each band remains high, at comparable levels to the no Lcp1 control in lane 2. This corresponds with what can be seen on the TLC plate.

The H_2O:C_2H_5OH:$(C_2H_5)_2O$:C_5H_5N:NH_4OH (15:15:5:1:0.017, v/v/v/v/v) extraction, termed 'E-soak' fraction, was believed to contain free unbound AG generated by the biosynthetic enzymes within the assay mixture. This was analysed by western blot (Figure 3.17A), using a primary antibody (CS-35) that recognised and bound to the hexa-arabinofuranoside terminal motif ([β-Ara*f*-(1→2)-α-Ara*f*]$_2$-3,5-α-Ara*f*-(1→5)-α-Ara*f*) of AG and LAM (Section 1.2.3.1., Section 5.23.5.). The secondary antibody had alkaline phosphatase activity and was visualised by development with BCIP solution (Sigma Aldrich), which identified bands of varying size (Figure 3.17A). This not only showed that arabinan was present in the 'E-soak' fraction, but also that these arabinan chains were "capped" with terminal Ara*f*$_6$ domains at varying lengths. The blot was then subjected to autoradiography, which showed the location of [^{14}C]-Gal*f* in relation to the developed arabinan bands (Figure 3.17B). The radiolabelled bands of the autoradiograph match the locations of the developed bands on the blot, indicating bound [^{14}C]-galactan to arabinan, showing the presence of AG within the 'E-soak' fraction. Importantly, the intensity of the radioactive bands decreased with an increase in the amount of ^{Mtb}Lcp1, as proven using densitometry (Figure 3.17C). However, when the assay was repeated in the absence of PPG, the intensity of the [^{14}C]-galactan bands increased (Figure 3.17B,C, Lane 5). Therefore, as ^{Mtb}Lcp1 is increased, the relative amount of [^{14}C]-Gal*f* labelled material within the 'E-soak' fraction becomes decreased, which is a trend that is followed when the assay is repeated using ^{Mtb}Lcp1 SDMs.

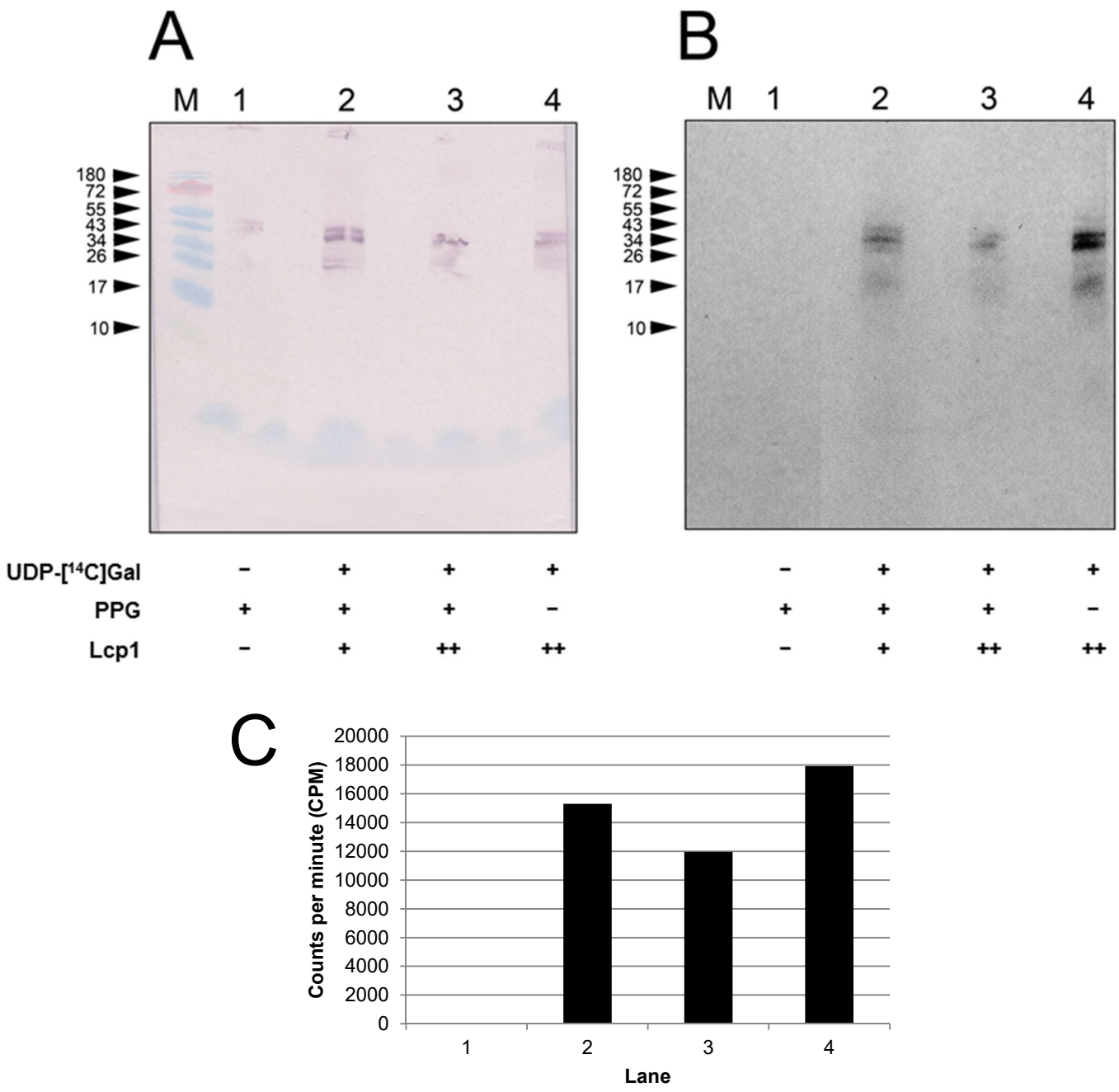

Figure 3.17: 'E-soak' extraction analysis. A) Western blot of 'E-soak' fraction after Tris-Tricine SDS-PAGE (Section 5.23.5.). Lane M is protein marker. The primary antibody used was CS-35, which binds to the terminal Ara*f*6 motif of the arabinan of AG and LAM, the secondary antibody was SC-2070 which can be visualised using BCIP solution (Sigma Aldrich). Blot shows terminal domains of arabinan at varying lengths. B) Autoradiography of 'E-soak' western blot A. Bands show [^{14}C]-galactan at the same locations as arabinan on the blot, indicating bound AG. C) Radioactivity present in each lane in counts per minute (CPM), based on the total radiation in the sample. Radioactivity decreases with increased Lcp1, whereas the absence of PPG causes an increase in remaining radioactivity, which corresponds with what is seen in B.

The final insoluble fraction contains PG and any PG linked [^{14}C]-Gal*f* labelled material. Analysis of per-*O*-acetylated alditol acetates of this material by TLC (Section 5.23.5.), revealed a major band which migrated to the position of per-*O*-acetylated galactitol acetate (Figure 3.18A). Bands were also identified to correspond with per-*O*-acetylated glucitol acetate, among others, which indicated that some UDP-[^{14}C]-Gal had been metabolised into other sugars under these conditions, which is a phenomenon that has been previously reported (Yagi *et al*., 2003). The addition of ^{Mtb}Lcp1 into the assay mix resulted in an increase in intensity of the per-*O*-acetylated galactitol acetate band, which seemingly increased further with increased ^{Mtb}Lcp1 (Figure 3.18A, Lane 3 and 4). This was confirmed through densitometry, since there is higher radioactivity within these bands, than in the absence of ^{Mtb}Lcp1 (Figure 3.18B Lane 2, 3 and 4). However, when the assay was repeated in the absence of PPG, the incorporation of [^{14}C]-Gal material had diminished, since the band corresponding to per-*O*-acetylated galactitol acetate was of very low intensity, once again confirmed by densitometry, showing much lower radioactivity within this band (Figure 3.18A, B, Lane 5). The same trend continued for the SDMs of ^{Mtb}Lcp1 where the addition of ^{Mtb}Lcp1 led to an increase in intensity of per-*O*-acetylated galactitol acetate bands, except in absence of PPG (Figure 3.18, Lanes 6-8 (R70A), Lanes 9-11 (R118A), Lanes 12-14 (R236A)).

The assay data as a whole shows two major findings. Firstly, that the attachment of [^{14}C]-Gal*f* cell wall material to PG can be directly monitored. Secondly, that recombinant ^{Mtb}Lcp1 is able to utilise decaprenyl-pyrophosphoryl-GlcNAc-Rha*p*-[^{14}C]-Galf_n as a substrate for attachment to PG, which is unaffected by the presence of bound arabinan, or by mutations to predicted residues believed to be involved in polyisoprenoid-phosphate binding.

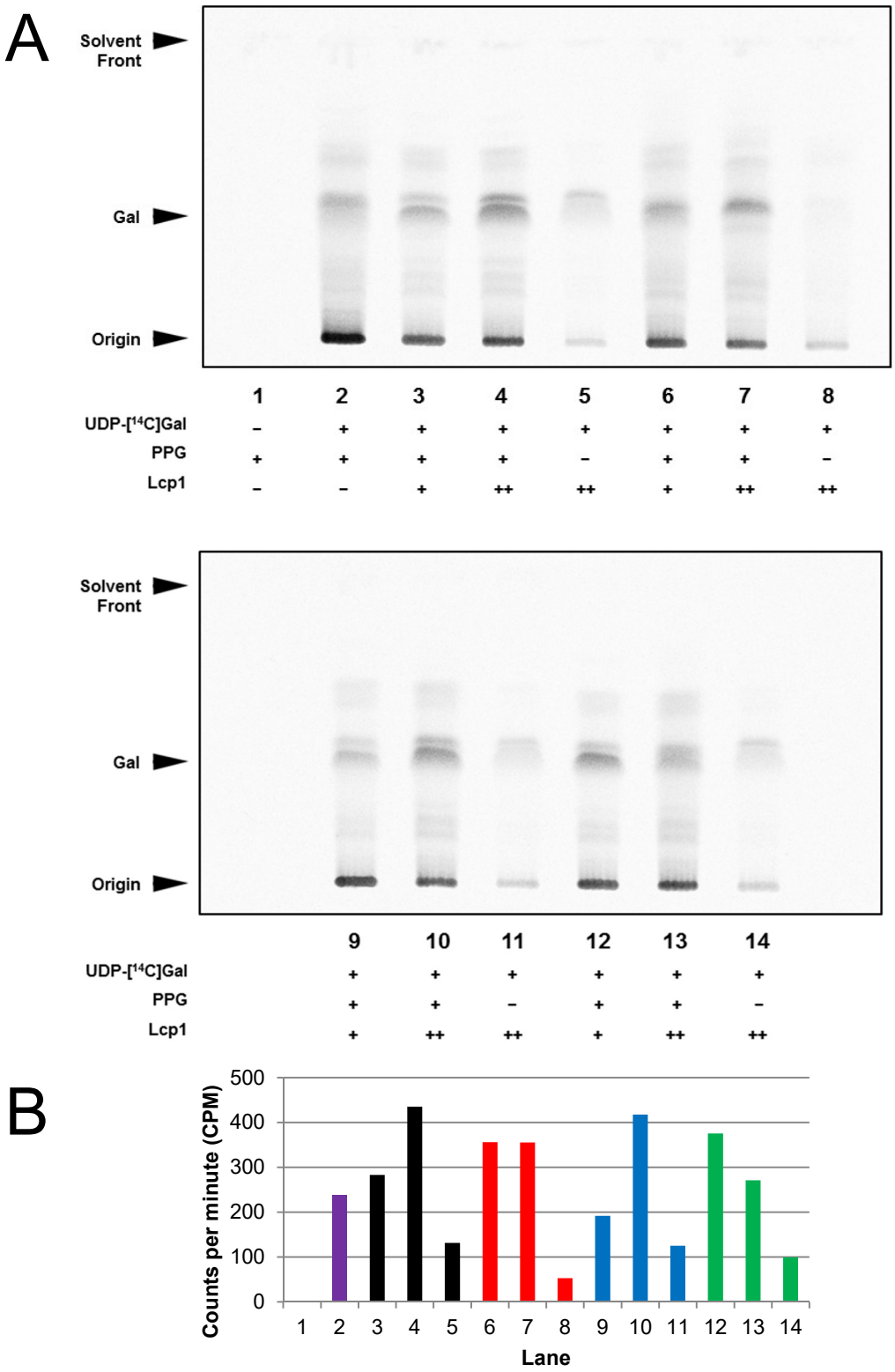

Figure 3.18: Insoluble material analysis. A) Insoluble pellet material was analysed by TLC, after hydrolysis, reduction and per-*O*-acetylation (Section 5.23.5.). The resulting per-*O*-acetylated alditol acetates were developed by TLC using ethyl acetate:hexane (4:6, v/v). As the amount of Lcp1 increases, so does the intensity of per-*O*-acetylated galactitol acetate bands. In the absence of PPG, the intensity of these bands diminish, suggesting both Lcp1 and PPG are required for addition of [^{14}C]-Gal onto PG. B) Radioactivity present in all bands, showing the true counts per minute (CPM) within each band from the total radioactivity in the sample (No Lcp1 = purple, ^{Mtb}Lcp1 WT = black, R70A = red, R118A = blue, R236A = green). The same trend is seen across all varations of ^{Mtb}Lcp1, whereby an increase of Lcp1 corresponds to an increase of radioactivity (over that of the no Lcp1 control (Lane 2)), except when there is an absence of PPG. Lane 13 shows lower CPM than expected but still a higher amount of radioactivity than seen in Lane 2.

3.2.6. Phosphatase activity of ^{Mtb}Lcp1

It has previously been shown that LCP proteins have a phosphatase activity, that is capable of producing inorganic phosphate (Pi) as a by-product of cleavage of the pyrophosphate head group of the bound polyisoprenoid, and the formation of the new phosphodiester bond (Kawai *et al*., 2011, Eberhardt *et al*., 2012). With this in mind, an assay was devised to assess whether recombinant ^{Mtb}Lcp1 also possessed phosphatase activity.

Geranyl-pyrophosphate (G-P-P) and geranylgeranyl-pyrophosphate (GG-P-P) were utilised as decaprenyl-pyrophosphate mimics in a phosphate cleavage assay (Section 5.24.), the products of these assays were then analysed by TLC (Figure 3.19). In the absence of ^{Mtb}Lcp1, there is little to no cleavage of phosphate from the pyrophosphate head group of either G-P-P or GG-P-P, despite degradation seen at the top of each TLC (Figure 3.19A, B). However, when the assay is repeated in the presence of ^{Mtb}Lcp1 (WT or SDM) an additional band appeared on each TLC plate, corresponding to geranyl-monophosphate (G-P) (Figure 3.19A) and geranylgeranyl-monophosphate (GG-P) (Figure 3.19B). Therefore, the data suggests that recombinant ^{Mtb}Lcp1 does possess phosphatase activity comparable to previously characterised LCP proteins

Additional bands migrating toward the top of the TLC remain unidentified (Figure 3.19). Mass spectrometry of the products produced in each experiment provided inconclusive data as to the characteristics of these bands. However, the positions of G-PP, GG-PP, G-P and GG-P were identified by R_f values (Kawai *et al*., 2011).

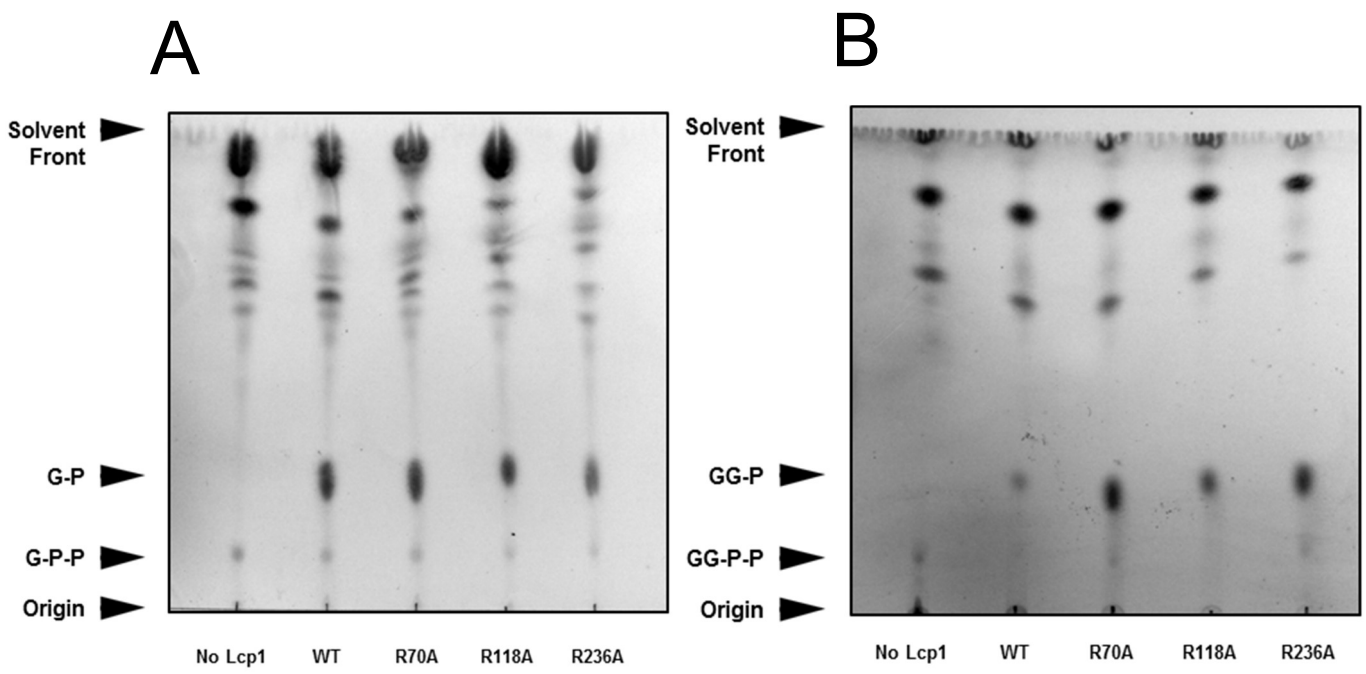

Figure 3.19: Analysis of ^{Mtb}Lcp1 phosphatase assay products by TLC. TLC plate (Section 5.14.) developed in $CHCl_3$:CH_3OH:NH_4OH:H_2O (65:25:0.5:3.6, v/v/v/v). A) Analysis of G-P-P products shows a band for G-P present when ^{Mtb}Lcp1 is added to the assay mix, which is missing in the absence of the protein. B) Analysis of GG-P-P products shows a band for GG-P present when ^{Mtb}Lcp1 is added to the assay mix, which is missing in the absence of the protein. Both of these analyses suggest a phosphatase activity for ^{Mtb}Lcp1.

3.3. Discussion

The LU of mycobacterial AG is remarkably similar to that of CWTA in Gram-positive bacteria, both sharing a common D-GlcNAc (1→P) at the PG bound end (Figure 3.1). These common LU are then both covalently bound to the C6-hydroxyl of specific muramyl residues of PG through a phosphodiester bond. Other commonalities exist between mycobacterial AG and Gram-positive CWTA, in that both biosynthetic pathways begin with assembly of polymers onto polyprenyl-pyrophosphate lipid carriers within the cytoplasm, before transport to the extracellular side of the membrane (Yokoyama *et al*., 1986, Jankute *et al*., 2015). A recently discovered superfamily of proteins was found to be responsible for the attachment of CWTA to PG through phosphotransferase activity, known as LCP proteins (Kawai *et al*., 2011). In this regard, we hypothesised that *M. tuberculosis* may harbour a similar protein of the LCP family in order to attach AG to PG within its cell wall architecture.

Through bioinformatic analysis of members of the Corynebacterineae suborder, we identified ^{Mtb}Lcp1 (Rv3267) as the primary putative LCP protein in *M. tuberculosis*, sharing 20 % sequence identity with Cps2A from *S. pneumoniae*. There are two additional putative LCP orthologs in *M. tuberculosis*, Rv3434 and Rv0822c, which share 15 % and 12 % identity with Cps2A from *S. pneumoniae* respectively and also share 36 % and 26 % identity with ^{Mtb}Lcp1, respectively. Furthermore, both orthologues of ^{Mtb}Lcp1 are predicted to contain a single N-terminal α-helix, as well as LCP and LytR_C superfamily domains similar to ^{Mtb}Lcp1. Because *lcp1* is found immediately downstream of two genes involved in LU biosynthesis, *wbbL1* and *rmlD*, we decided to utilise *M. smegmatis* as a model system to investigate the molecular genetics of the gene, in order to assess its essentiality and role as a primary LCP orthologue involved in mycobacterial cell wall assembly.

^{Mtb}*lcp1* was purported to be an essential gene through a high throughput screen of the *M. tuberculosis* H37Rv genome, however this needed to be examined experimentally (Xu *et al*., 2013). *M. smegmatis* has been used previously as a model organism to successfully study the molecular genetics of mycobacterial cell wall biosynthesis and therefore a clean deletion of ^{Ms}*lcp1* in *M. smegmatis* was attempted (Birch *et al*., 2008, Bhatt and Jacobs, 2009). However, as anticipated, it was not possible to generate a direct null mutant, as this would result in a non-viable, lethal phenotype because of the essential nature of ^{Ms}*lcp1*, as observed (Section 3.2.2.). The essentiality of ^{Ms}*lcp1* was confirmed by performing a CESTET experiment (Figure 3.4, Figure 3.5). The generation of null mutants of genes involved in the biosynthesis of the cell wall is technically challenging, since a non-viable phenotype is often returned due to lesions within the essential cell envelope structure. In *S. aureus*, the first two genes involved in CWTA biosynthesis (*tarO* and *tarA*) are non-essential and can be deleted without altering growth (D'Elia *et al*., 2006a, D'Elia *et al*., 2009). Despite this, deletion of any subsequent CWTA synthesising enzyme (*tarBDFIJH*) resulted in a non-viable, lethal phenotype, unless the mutant also contained an inactivation of either *tarO* or *tarA* (D'Elia *et al*., 2006b). A proposed hypothesis for this lethality, which is overcome when CWTA synthesis is prevented (by inactivation of *tarO*), is the sequestering of a scarce shared building block, such as the polyisoprenoid lipid carrier which is required not only for CWTA biosynthesis, but also the biosynthesis of "Lipid II" in PG assembly, (D'Elia *et al*., 2006b).

The polyisoprenoid lipid carrier, decaprenyl-phosphate, in Corynebacterineae such as *M. tuberculosis* is required for not only "Lipid II" in PG biosynthesis, but also in the production of DPA required for arabinan biosynthesis. Benzothiazinones (BTZs) are a new class of compound that are potent anti-mycobacterial agents (Batt *et al*., 2012). They are particularly potent inhibitors of mycobacterial growth since they target DprE1, which is involved in the epimerisation of DPR to DPA, thus inhibiting arabinan biosynthesis (Section 1.2.3.2.).

However, it has recently been concluded that BTZs mode of action is through the blocking of decaprenyl-phosphate recycling, consequently preventing cell wall biosynthesis from occurring properly (Grover *et al*., 2014). With this in mind, the identification and characterisation of the enzyme responsible for the attachment of AG to PG in *M. tuberculosis* would likely be a potential drug target, since its inhibition would prevent cell wall maturation and reduce the availability of decaprenyl-phosphate, leading to cell lethality.

The presence of a hydrophobic polyisoprenoid binding pocket within the ^{Mtb}Lcp1 homology model (Figure 3.2C) is consistent with similar conserved cavities within other LCP proteins (Kawai *et al*., 2011, Eberhardt *et al*., 2012). This is evident through biochemical analysis of recombinant ^{Mtb}Lcp1 which was shown to co-elute with decaprenyl-1-monophosphate, likely bound within this binding pocket (Figure 3.9). The biochemistry of ^{Mtb}Lcp1 was probed further with the use of a chemically synthesised panel of LU analogues, each representing chemical structures found near the start of the mycobacterial AG biosynthetic pathway (Section 1.2.3.2.). ITF was used to probe the ligand binding properties of these compounds against ^{Mtb}Lcp1, with compounds 1-3 showing a ligand binding relationship that reached saturation with a defined B_{max} and binding affinity at the low µM range. Compound 1 was synthesised to represent the disaccharide 'core' of the LU, L-Rha*p*-(1→3)-D-GlcNAc, which recorded a K_d of 58.71 µM. However, with the addition of 2 Gal*f* residues in the tetrasaccharide Galf_2-Rha-GlcNAc-*O*-C_8 (compound 2) the binding affinity increased 10-fold, giving a K_d of 5.13 µM. When this molecule was increased by a further Gal*f* residue (compound 3) there was a 4-fold decrease in K_d to 20.39 µM (Figure 3.14). Due to technical reasons the chemical synthesis of Gal*f*-Rha-GlcNAc-*O*-C_8 was not possible, which meant the study of the ligand binding properties of this single Gal*f*-LU structure were unable to be studied. A control experiment was performed with the use of compound 4, representing the LU in the absence of D-GlcNAc, which displayed very low binding affinity (K_d of 97.61 µM)

in comparison to the other compounds, as well as a diminished binding capacity, suggesting that D-GlcNAc is a vital residue in the biophysical interaction between ^{Mtb}Lcp1 and ligand. Since ^{Mtb}Lcp1 is capable of binding not only the LU (compound 1) but also LU plus multiple Gal*f* residues (compound 2-3), the data suggests that it is highly likely that ^{Mtb}Lcp1 might catalyse the phosphotransfer of AG to PG upon encountering a true physiological substrate, consisting of decaprenyl-pyrophosphate-LU and at least one (if not two) Gal*f* residues attached (Figure 3.21).

Mutations to residues believed to be involved in polyisoprenoid-phosphate binding resulted in an overall reduction in binding capacity of both compounds 1 and 4, for each ^{Mtb}Lcp1 variant (Figure 3.15). In particular, the mutation of R118A resulted in ^{Mtb}Lcp1 displaying a much higher binding affinity for compound 1 than the WT protein, with a K_d 27.76 μM, yet also shows a much lower binding capacity, with a B_{max} of 744.5, compared to 1,448 for WT (Figure 3.15A). This trend continued with compound 4, whereby ^{Mtb}Lcp1 R118A showed a further diminished binding capacity of 38.61 (Figure 3.15B). The fact that these mutations affect the binding of LU mimetics, as well as the fact that each mutant protein has been shown to co-elute with a bound decaprenyl-1-monophosphate, suggests that these specific residues actually have a larger role in LU recognition, rather than simply polyisoprenoid-phosphate binding. The position of these residues within the homology model of ^{Mtb}Lcp1 helps to support this, since the side chains of R118 and R236 in particular extend toward the phosphate head group of decaprenyl-phosphate when it is bound in its suspected binding pocket (Figure 3.20). Therefore it is likely that these residues would be capable of interaction with D-GlcNAc-P of the LU region and play a role in substrate binding (Figure 3.20).

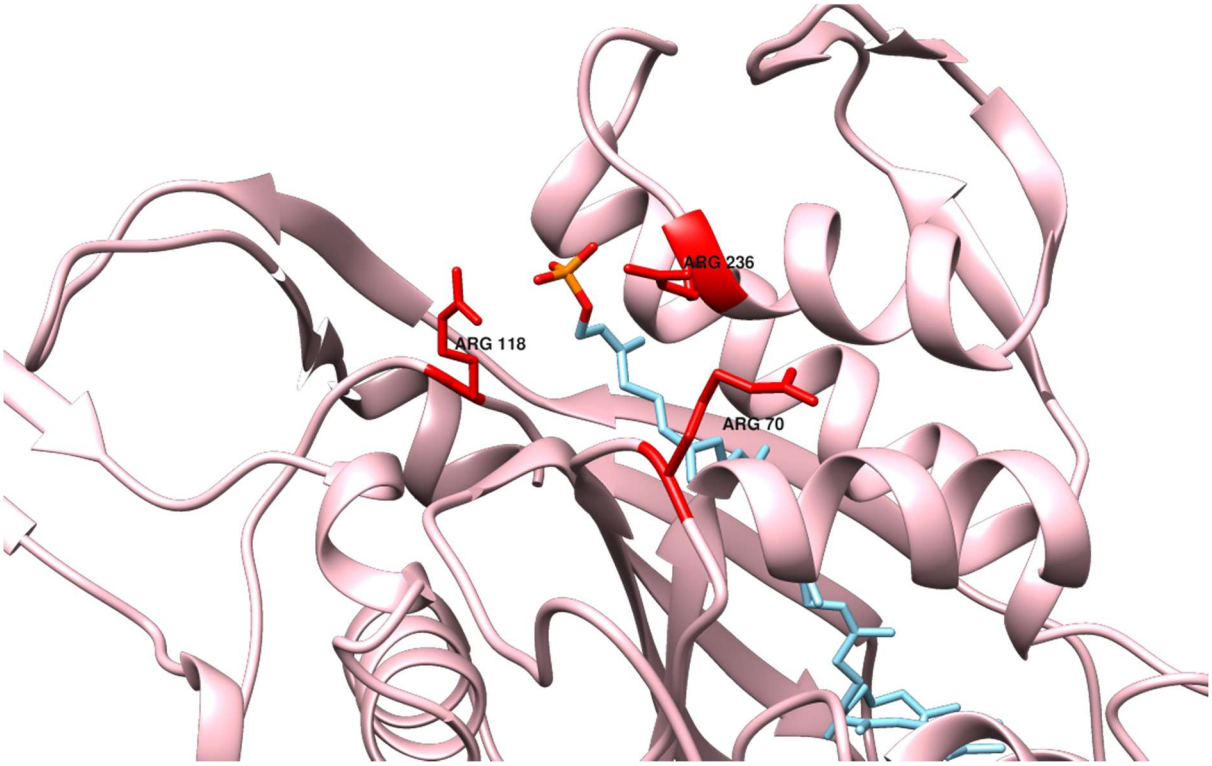

Figure 3.20: Positions of R70, R118 and R236 within homology model of ^{Mtb}Lcp1 with bound decaprenyl-phosphate. All three arginine residues are within the vicinity of the phosphate head group of decaprenyl-phosphate, in particular R118 and R236 whose side chains extend toward the phosphate group. It is likely that theses residues could be involved in LU binding since they would be in the correct location to interact with D-GlcNAc-P of the LU.

We devised a cell-free functional assay to assess the activity of ^{Mtb}Lcp1 and measure its ability to attach AG to PG. Chemical analysis of fractionated samples from an assay mix containing [^{14}C]-Gal radiolabelled products showed that lipid-linked AG precursors (Figure 3.16), as well as lipid-linked mature AG (Figure 3.17) decrease in an inverse relationship to the amount of ^{Mtb}Lcp1 present. A similar pattern was displayed in the insoluble material, showing a titratable dependant increase in per-*O*-acetalyated-[^{14}C]-galacitol acetate with an increase of ^{Mtb}Lcp1 in the assay mix. This was seen regardless of any mutations to residues potentially important in LU and polyisoprenoid-phosphate binding, and indicates an increase in AG bound to PG in these samples (Figure 3.18). Interestingly, it was seen that in the absence of PPG, there was little to no ^{Mtb}Lcp1 action seen throughout the assay. This is of

note, as it has previously been suggested that attachment between AG and PG requires concomitant biosynthesis of both cell wall polymers in order to take place (Hancock *et al*., 2002). It has also been shown that ^{Mtb}Lcp1 displays some phosphatase activity which is indicative of the action of LCP family proteins, since phosphate release would occur upon formation of a phosphodiester bond between CWTA/AG and PG (Figure 3.19). However, the results of our phosphatase assay are contrary to those shown by Kawai *et al*., 2011, since their results suggest that Cps2A is incapable of cleaving G-P-P, which is attributed to blockage of the active site by exogenous polyisoprenoid-pyrophosphate. Since we have shown that ^{Mtb}Lcp1 co-elutes with a bound decaprenyl-phosphate, our results seem to suggest that this bound lipid can be displaced easily in order to allow other substrates to bind (Figure 3.8, Figure 3.9). This is also evidenced by the fact that ^{Mtb}Lcp1 shows activity in the cell-free assay, proving its capability of binding to glycolipid 3/4 substrates and utilising them within the assay (Figure 3.16). Conversely, an alternative reason for the activity of ^{Mtb}Lcp1 may be that it could purify as a heterogeneous mix of both decaprenyl-phosphate bound and un-bound states, with the 'un-bound' protein providing the activity seen.

We suggest a putative mechanism for phosphotransfer by ^{Mtb}Lcp1 (Figure 3.21), initiated by the hydroxyl group of the phosphate bound to the decaprenyl moiety of the LU becoming deprotonated by an as of yet unconfirmed arginine residue (R70, R118, R236 or R238). We propose the hydrogen of the phosphate group binds to the double bonded NH/NH_2 of the arginine residue, with the resulting positive charge passing to the opposing NH_2 group. A neighbouring aspartate residue (D97, D119 or D245) then removes the extra proton to balance the charge, whilst the electrons between the hydroxyl group pass to the oxygen, breaking the bond between the atoms (Figure 3.21). The oxygen of the C6 hydroxyl from a MurNAc residue within PG then attacks the LU phosphate bound to the GlcNAc moiety by nucleophilic substitution, resulting in electrons passing from phosphate to its bound oxygen to

break the bond (Figure 3.21). Meanwhile, the newly negatively charged oxygen from the neighbouring phosphate group attacks the hydrogen of the MurNAc C6 hydroxyl, removing it and culminating in the bond between the AG LU and PG and the liberation of decaprenyl-monophosphate (Figure 3.21). However, this mechanism remains uncertain without the full structure of ^{Mtb}Lcp1 being solved.

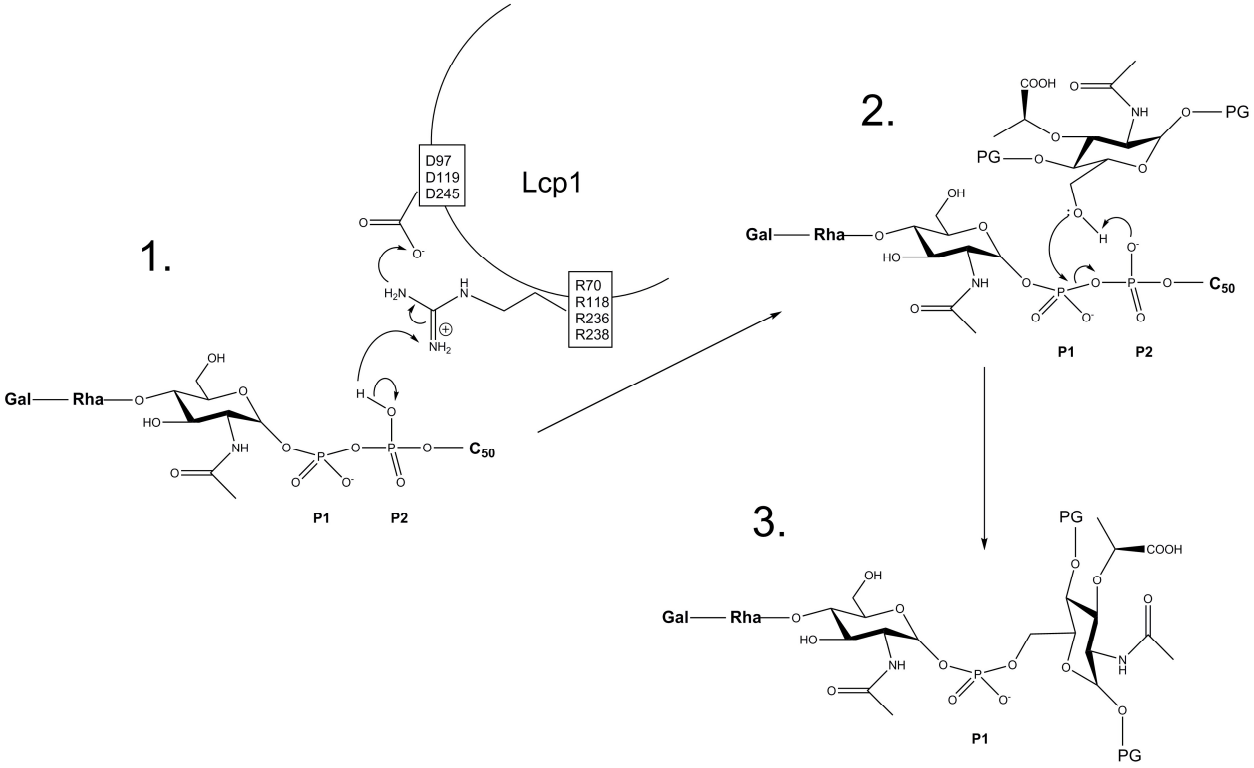

Figure 3.21: Potential arrow pushing mechanism of ligase activity by ^{Mtb}Lcp1. 1) Deprotonation of phosphate group attached to the decaprenyl moiety (C_{50}) (P2) by arginine of ^{Mtb}Lcp1 (either R70, R118, R236 or R238). Electrons are passed to oxygen of P2 to break bond as proton is attached to NH group of arginine. Proton is then removed from active arginine by neighbouring aspartate (either D97, D119 or D245). 2) The C6 hydroxyl of MurNAc from PG then acts as a nucleophile, attacking the phosphate attached to the LU moiety (P1). The negative charge attached to P2 then attacks the proton of the C6 hydroxyl , as the phosphodiester bond between the phosphate groups is broken. 3) P1 becomes covalently bound to C6 of MurNAc (PG) as decaprenyl-phosphate is liberated.

This study provides compelling genetic and biochemical evidence that ^{Mtb}Lcp1 is the enzyme responsible for the convergence of two major cell wall biosynthetic pathways, culminating in a phosphodiester bond between AG and PG in mycobacteria (Figure 3.22).The discovery of ^{Mtb}Lcp1 sheds new light on a key reaction during late stage cell wall assembly and is likely to

be a critical protein for all Corynebacterineae that possess cell walls containing AG attached to PG *via* a phosphodiester bound LU. However, this discovery does raise the issue of when arabinan becomes bound to galactan to form AG, before or after attachment to PG, since ^{Mtb}Lcp1 has been shown to be capable of utilising both arabinan bound and non-arabinan bound LU-galactan. As a consequence of this, the timing of mycolylation of the arabinan chains is also questioned. If arabinan is bound to galactan before "ligation" to PG, then it could also be mycolylated at this point with a potential interaction between ^{Mtb}Lcp1 and antigen 85 complex. Conversely, this could all occur once ^{Mtb}Lcp1 has completed the attachment of the LU-galactan to PG and has left the area.

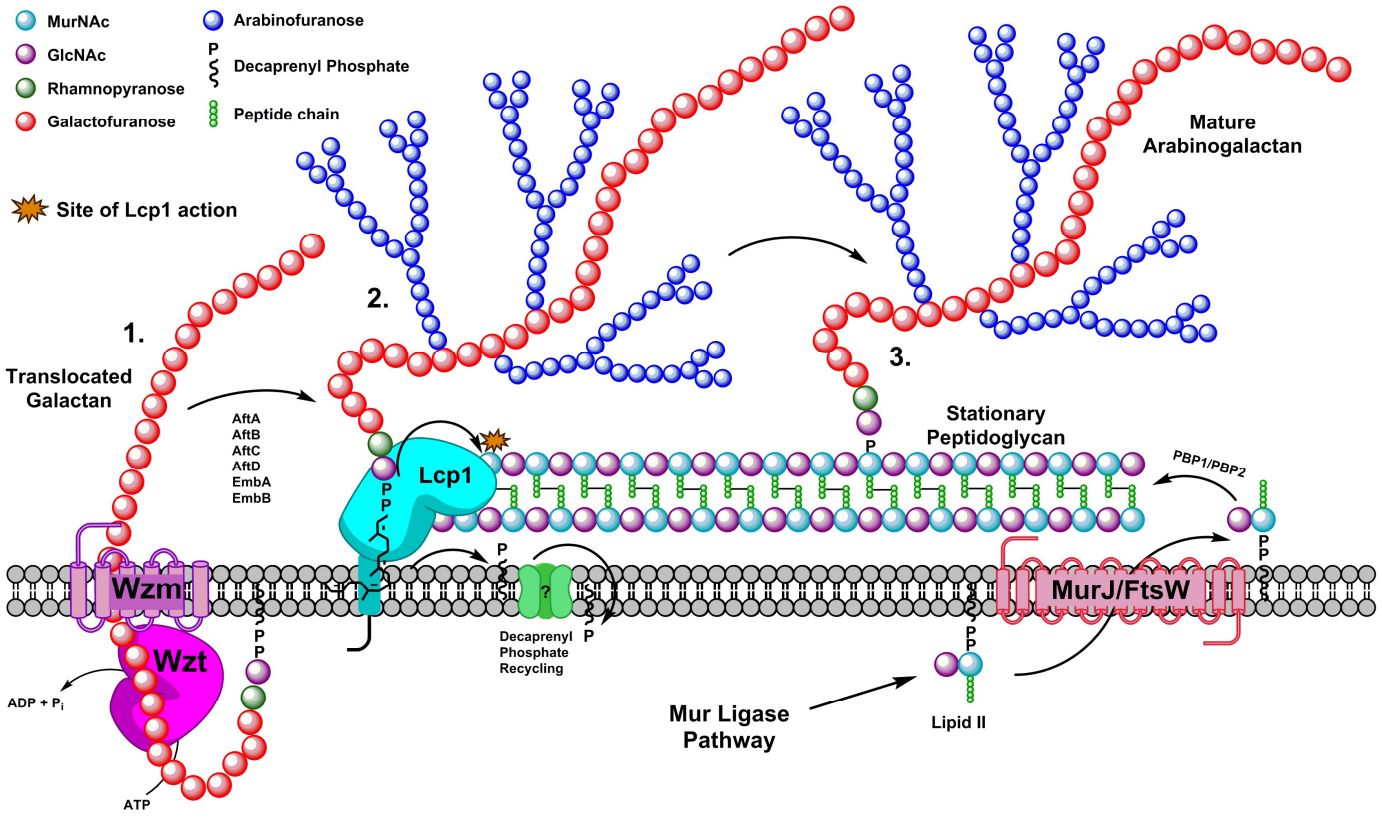

Figure 3.22: Proposed mechanism of action of Lcp1 for attachment of AG to PG in the latter stages of cell wall biosynthesis. 1) Translocation of galactan polymer across plasma membrane. Lcp1 then recognises and binds the decaprenyl-pyrophosphoryl-LU (i.e. represented by compounds 2 and 3). 2) Phosphotransferase activity links AG to the 6'-OH of muramyl residues, releasing decaprenyl-phosphate to be recycled. 3) Newly bound AG is released by Lcp1 to go on to form mature mAGP. Additional synthesis of new PG, requiring AG addition by Lcp1, takes place simultaneously.

Chapter 4

General Discussion

4. GENERAL DISCUSSION

4.1. Conclusions and Future perspectives

TB remains a global health crisis, which is only becoming exacerbated in recent times with the rise of MDR, XDR and even TDR strains of *M. tuberculosis* around the world. The effectiveness of this pathogen is due, in part, to its intricate carbohydrate and lipid-rich cell wall, which not only provides protection to the bacterium, but also plays an important role in its virulence. It is therefore vital to have a complete understanding of the mycobacterial cell wall, not just to advance scientific understanding, but also to help assist in the ongoing chemotherapeutic effort against TB by identifying potential new drug targets. Although many of the biosynthetic pathways of the mycobacterial cell wall have been fairly well defined in recent years, numerous voids in knowledge still exist. Therefore, this thesis aims to address the key themes of regulation and completion of crucial aspects of mycobacterial cell wall biosynthesis, which have previously been overlooked by scientific investigation.

In chapter 2, we investigated how *M. tuberculosis* can control PG biosynthesis and therefore cell wall assembly. We have proven that phosphorylation of *M. tuberculosis* MurC has an effect on protein kinetics and is capable of downregulating its activity. Pseudo-phosphorylated mutants simulated phosphorylation of specific sites within MurC, through substitution of specific serine/threonine residues (homologous to those known to be phosphorylated in *C. glutamicum* by PknA) to aspartate (Fiuza *et al*., 2008). We used the I-TASSER server to produce homology models of both *C. glutamicum* and *M. tuberculosis* MurC, which were subsequently analysed against the crystal structure of *H. influenzae* MurC (Mol *et al*., 2003). Both the Cg-MurC and Mtb-MurC models aligned against each other with a TM-score of 0.93544, indicating a good fit, with both the phosphorylated Cg-MurC residues occupying the same space as the homologous residues in the Mtb-MurC model. All of these

residues occupied space surrounding the ATP binding domain, suggesting that phosphorylation of these residues would likely interfere with ATP binding *in vivo*.

The introduction of a negative charge at key residues of MurC, namely T116 and T357, has a profound effect on the Michaelis-Menten kinetics of the enzyme, reducing activity by least around 50 %, when compared to MurC WT. Conversely, pseudo-phosphorylation of S360 resulted in a slight up-regulation of MurC activity of around 15-20 % when compared to MurC WT, as opposed to the down-regulation displayed by the mutation of some of the other selected residues. A similar trend was seen regardless of whether the UDP-Muramyl or ATP dependence of MurC was assayed. We also discovered that phosphorylation of the specific sites of *M. tuberculosis* MurC have the potential to increase co-operative binding of ATP, as well as binding affinity, when assayed by ITF. This increase in co-operation between binding sites could be explained by an increase in dimerisation of MurC, caused by the pseudo-phosphorylation of T116, S163 and S360 residues, since *E. coli* MurC has shown the ability to form dimers in solution, yet retain functional activity (Das *et al*., 2011). Furthermore, we noted that *M. tuberculosis* MurC displayed a much higher affinity for UDP-MurNGlyc over UDP-MurNAc, showing a 30-fold difference between K_m values. We suggested that this difference in affinity, along with de-phosphorylation of down-regulated MurC, could be utilised by *M. tuberculosis* as a means to exit dormancy and resume active replication upon the degradation of the granuloma.

In chapter 3, we probed the final union of both the AG and PG biosynthetic pathways through the identification of a putative AG/PG ligase, Lcp1. Firstly, bioinformatic analysis of Corynebacterineae revealed Rv3267 (^{Mtb}Lcp1) as the primary LCP homolog of *S. pneumoniae* Cps2A in *M. tuberculosis* and MSMEG1824 (^{Ms}Lcp1) as the primary orthologue in *M. smegmatis*. We created a homology model of ^{Mtb}Lcp1 (using the I-TASSER server) and analysed it against the crystal structure of Cps2A (Fiuza *et al*., 2008). We saw that the model

complied with many of the structural features seen in other LCP proteins, only lacking the non-conserved accessory domain. The essentiality of ^{Ms}*lcp1* was proved through a CESTET experiment, by the creation of a conditional Δ^{Ms}*lcp1* knock-out mutant. Without the acetamide induced second copy of ^{Ms}*lcp1*, the Δ^{Ms}*lcp1* cell viability rapidly diminished, proving how vital ^{Ms}*lcp1* is to the survival of *M. smegmatis* and in turn *M. tuberculosis*.

The presence of co-eluted polyisoprenoid-phosphate is indicative of LCP proteins (Kawai *et al*., 2011) and as such, we identified that recombinant ^{Mtb}Lcp1 WT and three site-directed mutant proteins (^{Mtb}Lcp1 R70A, R118A and R236A) all co-eluted with a bound decaprenyl-1-monophosphate, despite the fact that the mutated residues were originally believed to be involved in polyisoprenoid-phosphate binding. Each of these ^{Mtb}Lcp1 variants were also shown to be capable of *in vitro* phosphatase activity, able to cleave the pyrophosphate bond of G-P-P and GG-P-P, an ability which Cps2A was shown incapable of (Kawai *et al*., 2011). It is possible that the co-eluted decaprenyl-1-monophosphate is easily displaced from ^{Mtb}Lcp1, allowing for surrogate polyisoprenoid-phosphate binding and permitting phosphate cleavage. However, it is equally likely that ^{Mtb}Lcp1 is purified as a mixture of both decaprenyl-1-monophosphate bound and un-bound units and it is the 'un-bound' protein that displays this activity.

The ideal ^{Mtb}Lcp1 LU substrate structure was investigated using both ITF and a panel of synthetic LU analogues. We were able to show that the preferred substrate was Gal*f*$_2$-Rha-GlcNAc-*O*-C$_8$ (compound 2), which displayed a binding affinity 4 times higher than any other compound. However, in a control experiment using compound 4, representing the LU without D-GlcNAc, the binding affinity diminished, indicating the importance of this conserved monosaccharide to substrate recognition. The ligand binding properties of each ^{Mtb}Lcp1 SDM were also probed, revealing a reduced binding capacity of each mutant protein for compounds 1 and 4, suggesting a role in LU binding for the mutated residues. This data shows that it is

highly likely *in vivo*, that ^{Mtb}Lcp1 action would depend on encountering a true physiological substrate consisting of decaprenyl-pyrophosphate-LU and a galactan chain consisting of at least one, if not two, Gal*f* residues.

We were able to prove the ability of ^{Mtb}Lcp1 to 'ligate' AG to PG by developing a cell-free radiolabelled functional assay. In the absence of recombinant ^{Mtb}Lcp1, a build-up of [^{14}C]-Gal*f* labelled substrates was seen, including both GL-4 and AG, whereas an increase in ^{Mtb}Lcp1 correlated with a large decrease in these substrates. Likewise, the addition of recombinant ^{Mtb}Lcp1 correlated with an increase of [^{14}C]-Gal*f* labelled products within the final insoluble material (the PG), therefore indicating that the increased presence of ^{Mtb}Lcp1 results in an increase of radiolabelled AG addition to PG. We also discovered that in the absence of PPG within the starting assay mix, ^{Mtb}Lcp1 activity was abolished, demonstrating that ^{Mtb}Lcp1 activity is dependent upon pre-existing PG.

Both of the major proteins discussed in this thesis play an important role in *M. tuberculosis* cell wall biosynthesis, which is a major factor during cell division. Being a rod-shaped organism, *M. tuberculosis* traditionally divides symmetrically at the septum formed at the mid cell (Hett and Rubin, 2008). However, it has recently been shown that around 20 % of mid-log replicating *M. tuberculosis* divides asymmetrically, due to deviation in septum location, resulting in viable, but non-identical short/long daughter cells (Vijay *et al*., 2014). It has also been noted that in cases of drug-resistant *M. tuberculosis*, cells are capable of branching division, whereby daughter cells begin to grow on the sides of mother cells, rather than splitting at the septum. In XDR and TDR strains, this branched replication becomes further adapted as daughter cells become small and spherical in shape, believed to be due to harsh environmental conditions (Farnia *et al*., 2010). Septum formation is initiated by FtsZ, a tubulin-like homolog, which binds GTP to form protofilaments, which then become polymerised to form a complex known as the Z-ring (Margolin, 2005). In *M. tuberculosis*,

FtsZ has been shown to display slower GTP-dependent polymerisation, as well as weaker GTPase activity by polymers than other bacteria. These polymers of FtsZ were also seen to be much more stable than the homologous polymers in *E. coli* (White *et al*., 2000). Once the Z-ring has formed, it begins recruitment of membrane-associated cell division proteins, with one, FtsW, forming a unique interaction with FtsZ in *M. tuberculosis*, believed to be responsible for anchoring the Z-ring to the plasma membrane (Datta *et al*., 2002). PG formation then begins at the septum, as the Z-ring begins to contract, believed to be possibly down to loss of FtsZ subunits in the polymer (Margolin, 2005, Hett and Rubin, 2008). This completes the septum and results in the sealing of the daughter cells (Hett and Rubin, 2008). The daughter cells separate as a V-shape, due to the growth of the cell wall at the septum, with PG hydrolysis allowing the cells to 'snap' apart from one another. This results in the outer layer of the cell wall becoming ruptured which 'scars' the new cell (Dahl, 2004, Thanky *et al*., 2007).

Opposed to most rod shaped bacteria, such as *E. coli*, which have nascent PG continuing through the cell and inert PG at the poles, mycobacteria grow with nascent PG being at their poles and inert PG running throughout the length of the cell (Thanky *et al*., 2007). Therefore, PG biosynthetic machinery would also be located at the poles of the cell where PG synthesis is taking place. It is therefore highly likely that other enzymes involved in cell wall biosynthesis would also locate to the poles where cell growth is occurring. Since the FtsZ polymer or Z-ring causes constriction and pinches the septum, as well as recruitment of the cell wall assembly machinery, it is probable that FtsZ remains at the poles long after cell division (Hett and Rubin, 2008). Since there is no cytoskeletal MreB homolog in mycobacteria to compete with FtsZ for cell shape definition, as well as the fact that the Z-ring is highly stable, could mean that FtsZ is responsible for recruiting and maintaining the cell

wall biosynthetic enzymes at the poles of the cell, allowing for nascent PG to form at each end of the cell (Thanky et al., 2007, Hett and Rubin, 2008).

It would be of interest to the outcomes of this thesis to see where Lcp1 is localised throughout the mycobacterial cell. We would assume that it would be located at the poles of the cell, where cell wall biosynthesis occurs to drive cellular growth, perhaps recruited by FtsZ if it truly does remain at the poles past cell division. This could be investigated by the creation of an Lcp1-GFP conjugation, which could then be visualised by fluorescence microscopy in order to track Lcp1 within live cells. This method could perhaps be utilised to see whether Lcp1 interacts with other cell wall biosynthetic enzymes by creating various coloured fluorescent tag conjugations of proteins such as members of the antigen 85 complex or Wzt/Wzm, etc. It would then be interesting to see whether these additionally tagged proteins co-locate with Lcp1. Furthermore, if the FtsZ polymer does remain at the poles of the cell to orchestrate cell wall growth, then the constriction of the Z-ring during cell division would create a 'pinch point' within the cell wall, which would likely be carried forward past the point of cell separation. In terms of the dimensions of the mycobacterial cell wall (Section 1.2.1.), this 'pinch point' would bring the components of the cell wall much closer to the plasma membrane. This would be in important consideration for membrane anchored proteins, such as Lcp1, since there would be no way that these proteins could act upon mAGP components across a 14-17 nm periplasmic space. With a pinching of the cell wall, these proteins would then be able to reach their sites of action and perform their roles in cell wall assembly.

We do not know how mycobacterial cells perform the regulation of such an important cell wall assembly protein as Lcp1. However, since we have proven that *M. tuberculosis* MurC can be phosphorylated in order to regulate its activity, it is likely that, if phosphorylation of MurC does indeed occur *in vivo*, Lcp1 would indirectly be regulated by MurC action. This

would likely transpire since we have proven that Lcp1 requires pre-existing PG for protein action, so if MurC is down-regulated by phosphorylation of key residues (T116 or T357), then the down-regulation of PG biosynthesis would also lead to the consequence of reduced Lcp1 activity, as there would be no PG to "ligate" AG to. It is likely however, that there are many more measures *M. tuberculosis* takes to regulate the protein that combines two crucial, nutrient demanding cell wall biosynthesis pathways.

Chapter 5

Materials and Methods

5. MATERIALS AND METHODS

5.1. General chemical and media preparation

All Chemicals and Solvents were from Sigma-Aldrich, Bio-Rad or Fisher Chemicals, unless otherwise stated. Restriction enzymes were from Thermo Scientific.

5.1.1. Luria-Bertani (LB) Broth

37 g of LB Broth (Merck Millipore) in 1 L Water (H_2O), autoclaved at 121 °C for 15 min.

5.1.2. Luria-Bertani (LB) Agar

37 g of LB Broth (Merck Millipore), 15 g Bacto Agar (BD, Difco) in 1 L H_2O, autoclaved at 121 °C for 15 min. Incubated at 55 °C, then poured into petri dishes (25 mL/plate).

5.1.3. Terrific Broth

47.6 g Terrific Broth (Merck Millipore) in 1 L H_2O, with additional 4 mL glycerol (100 % v/v). Autoclaved at 121 °C for 15 min.

5.1.4. NZY+ Broth

10 g NZ amine (casein hydrolysate), 5 g yeast extract, 5 g Sodium Chloride (NaCl) in 1 L H_2O, autoclaved at 121 °C for 15 min. Once cooled, 12.5 mL Magnesium Chloride ($MgCl_2$) (1 M), 12.5 mL Magnesium Sulfate ($MgSO_4$) (1 M) and 10 mL glucose (2 M) were filter sterilised (through 0.2 µm filter), and supplemented into media.

5.1.5. Tryptic Soy Broth (TSB)

30 g of TSB (BD, Difco) in 1 L H_2O, autoclaved at 121 °C for 15 min.

5.1.6. Tryptic Soy Agar (TSA)

30 g of TSB (BD, Difco), 15 g Bacto Agar (BD, Difco) in 1 L H_2O, autoclaved at 121 °C for 15 min. Incubated at 55 °C, then poured into petri dishes (25 mL/plate).

5.1.7. Middlebrook 7H9 Basal Agar

4.7 g Middlebrook 7H9 Broth (BD, Difco), 15 g Bacto Agar (BD, Difco) in 1 L H_2O (without ADC supplement). Autoclaved at 121 °C for 15 min. Incubated at 55 °C, then poured into petri dishes (20 mL/plate).

5.1.8. Middlebrook 7H9 'Top' Agar

4.7 g Middlebrook 7H9 Broth (BD, Difco), 6 g Bacto Agar (BD, Difco) in 1 L H_2O (without ADC supplement). Autoclaved at 121 °C for 15 min.

5.1.9. Transformation Buffers

5.1.9.1. Transformation Buffer 1 (TFB1)

100 mM Rubidium Chloride ($RbCl_2$), 50 mM Manganese Chloride ($MnCl_2$), 30 mM Potassium Acetate (KAc), 10 mM Calcium Chloride ($CaCl_2$), 15 % glycerol (v/v), pH 5.8.

5.1.9.2. Transformation Buffer 2 (TFB2)

10 mM MOPS, 10 mM $RbCl_2$, 75 mM $CaCl_2$, 15 % glycerol (v/v), pH 6.8.

5.1.10. Protein Purification Buffer (2 ×)

100 mM Potassium Phosphate (KH_2PO_4), 600 mM NaCl, pH 7.9

5.1.10.1. Lysis Buffer

50 mM KH_2PO_4, 300 mM NaCl, 20 mM Imidazole, pH 7.9.

5.1.11. Dialysis Buffers

MurC 'Storage' Buffer - 30 mM HEPES pH 7.6, 50 mM Potassium Chloride (KCl), 1 mM $MgCl_2$, 1 mM EDTA, 50 % glycerol (v/v), made up to 2 L.

Lcp1 Buffer - 25 mM Tris HCl pH 7.9, 10 mM NaCl, 10 % glycerol (v/v), made up to 2 L.

5.1.12. MurC Reaction Buffer

50 mM HEPES pH 7.6, 10 mM $MgCl_2$.

5.1.13. Mycobacteriophage (MP) Buffer

50 mM Tris HCl pH 7.6, 150 mM NaCl, 10 mM $MgCl_2$, 2 mM $CaCl_2$, sterilised through a 0.2 μm filter.

5.1.14. Buffer 'A'

50 mM MOPS (NaOH pH 7.9), 10 mM $MgCl_2$, 5 mM β-Mercaptoethanol.

5.1.15. Western Transfer Buffer

25 mM Tris, 190 mM Glycine pH 7.5, 10 % Methanol in 1 L H_2O.

5.1.16. Tris Buffered Saline (TBS)

20 mM Tris pH 7.5, 150 mM NaCl in 1 L H_2O. Can be made with 0.05 % Tween 20 (TBS-T).

5.2. Bacterial Strains and Conditions

E. coli Top 10 cells were used for propagation of plasmid DNA. These cells were routinely grown in LB broth, or LB agar at 37 °C. *E. coli* BL21 (DE3) cells were used for the overproduction of recombinant protein, grown in Terrific Broth at 37 °C. *E. coli* XL10 Gold cells were used during site-directed mutagenesis protocols, grown in NYZ+ broth, and LB agar at 37 °C. *M. smegmatis* mc^2155 was grown in TSB (containing Tween 80 (0.05 % v/v)) and TSA at 37 °C. All cells were supplemented with the appropriate antibiotics (50 µg/mL Kanamycin (Kan) and/or 100 µg/mL Hygromycin (Hyg)) when required.

5.3. Preparation of competent cells

5.3.1. *E. coli* chemically competent cells

E. coli cells were grown at 37 °C to OD_{600} = 0.4-0.5, before incubation on ice for 10 min. Cells were pelleted by centrifugation at 3,300 × *g* for 15 min and the supernatant was discarded. From this point, pelleted cells were kept on ice and were then re-suspended in 0.4 volume of ice cold TFB1 (Section 5.1.9.1.). The suspension was then incubated on ice for 15 min, before centrifugation at 3,300 × *g* for 15 min at 4 °C. The supernatant was discarded, and the cells were re-suspended in 0.02 volume of ice cold TFB2 (Section 5.1.9.2.). The suspension was then incubated on ice for 30-60 min, before being aliquoted into microcentrifuge tubes. Each aliquot was then flash frozen in liquid nitrogen (N_2) and stored at -80 °C.

5.3.2. *M. smegmatis* electrocompetent cells

M. smegmatis cells were grown at 37 °C to OD_{600} = 0.4-0.5, then pelleted by centrifugation at 3,300 × *g* for 15 min at 4 °C. The supernatant was discarded, and the pellet was re-suspended in 0.5 volume of ice cold 10 % glycerol (v/v). The suspension was then incubated on ice for

10 min. Cells were pelleted by centrifugation at 3,300 × *g* for 15 min at 4 °C, before being re-suspended in 0.3 volume of ice cold 10 % glycerol (v/v). The suspension was incubated on ice for 10 min. This trend of centrifugation, resuspension and incubation continued, reducing the volume of 10 % glycerol in each round (0.15 volume, 0.1 volume and 0.04 volume successively). Once the cells had been incubated in 0.04 volume 10 % glycerol (v/v), cells were aliquoted into microcentrifuge tubes. Each aliquot was then flash frozen in liquid N_2 and stored at -80 °C.

5.4. Transformation of bacterial cells

5.4.1. Transformation of *E. coli* cells by heat shock method

Chemically competent *E. coli* cells (100 µL) (Section 5.3.1.) were added on top of 1 µL plasmid DNA and incubated on ice for 15 min. The cell suspension was then subjected to 'heat shock', incubation at 42 °C for 1 min. Cells were then returned to ice for a further 2 min, before 250 µL LB broth was added and the suspension was incubated at 37 °C for 1 h. The suspension was then plated onto LB agar plates (supplemented with correct antibiotic) and incubated overnight at 37 °C.

5.4.2. Transformation of *M. smegmatis* cells by electroporation

Electrocompetent *M. smegmatis* cells (Section 5.3.2.) and plasmid DNA were incubated on ice until thawed. After this, 2 µL of plasmid DNA was inserted into a 1 mm electroporation cuvette (Cell Projects) and incubated on ice for 20 min. Then 200 µL of *M. smegmatis* cells were added to the cuvette and incubated on ice for 30 min. The cuvette was then inserted into an electroporator 2510 (Eppendorf) and pulsed at 1,800 V before incubation on ice for 15 min. Cells were rescued by the addition of 250 µL TSB and incubation at 37 °C for 4 h. After

incubation, cells were plated onto TSA plates (supplemented with correct antibiotic) and incubated at 37 °C for 2-3 days.

5.5. Polymerase Chain Reaction (PCR)

PCR was performed in 20 µL reactions, consisting of 0.5 µM forward primer, 0.5 µM reverse primer (Section 5.5.1.), 250 ng genomic DNA (*M. smegmatis*/*M. tuberculosis*), 200 µM dNTP mix, 1.5 mM $MgCl_2$, 3 % dimethyl sulfoxide (DMSO), 1 U/(50 µL reaction) Phusion DNA polymerase (New England Biolabs), 1× Phusion 5× GC buffer (New England Biolabs) and made up to 20 µL with nuclease free H_2O. Each reaction was then subjected to cycling conditions (Table 5.1) for gene amplification, inside a Mastercycler Gradient thermocycler (Eppendorf).

Table 5.1: PCR cycling conditions for *MSMEG1824* amplification. * R = 3 deg/s + G = 8 deg.

Cycle	Temp (°C)	Time (sec)
1	98	30
2	98	10
3	72*	30
	Back to cycle 2 × 35 times	
4	72	600
Hold at 4 °C		

5.5.1. Primers used for PCR

Table 5.2: Primers designed for amplification of *MSMEG1824* from *M. smegmatis*

Forward	5'-GATCGATCGATATCTTGATCAGGTCCATTGCTGTGGCCGCA-3'
Reverse	5'-GATCGATCAAGCTTTCAGTTCACGCACTGCGGGTCGTTGGC-3'

Table 5.3: Primers designed for amplification of *rv3267* from *M. tuberculosis*

Forward	5'-GATCGATCCATATGGTGATGTCTGCGCAACGTGTGGTTCGT-3'
Reverse	5'-GATCGATCAAGCTTTCAGTTGATGCACTCCGGCGCGTCGGA-3'

Table 5.4: Primers designed for amplification of *murC* from *M. tuberculosis* (TB) and *M. smegmatis* (MS)

TB Forward	5'-GATCGATCCATATGGTGAGCACCGAGCAGTTGCCGCCCGAT-3'
TB Reverse	5'-GATCGATCAAGCTTTCATCCCAGCACCCCCGGACGGCCGGG-3'
MS Forward	5'-GATCGATCCATATGATGACGGGCATCTCGCTGCCACCGGAG-3'
MS Reverse	5'-GATCGATCAAGCTTTCACGACGCGCCCGCTCCCGGCATACC-3'

Table 5.5: Primers designed for ^{Ms}*lcp1* Van91I-containing flanking regions. For creation of Δ^{Ms}*lcp1* phasmid (Section 5.19.) and for confirmation of genetic knock-out, through PCR of Δ^{Ms}*lcp1* genomic DNA.

F-LL	5'-TTTTTTTTCCATAAATTGGCGACCACCAGGGGGCGGGCATCGTCCA-3'
R-LR	5'-TTTTTTTTCCATTTCTTGGACCGGGGCAGGCGGCATGGTCGGGCGC-3'
F-RL	5'-TTTTTTTTCCATAGATTGGACGACGCCACCGGTGAACGCATCGAGC-3'
R-RR	5'-TTTTTTTTCCATCTTTTGGGCCGACGACAGGCTGCTCGACGACTAC-3'

5.6. Site-Directed Mutagenesis (SDM)

Site-Directed Mutagenesis (SDM) was performed using Quikchange II XL kit (Agilent Technologies). Reaction mixtures were made, consisting of 0.5 µL (10 ng) double-stranded DNA template (plasmid-gene DNA), 1 µL (125 ng) forward primer, 1 µL (125 ng) reverse primer (Section 5.6.1.), 1 µL dNTP mix, 5 µL 10 × reaction buffer (Agilent Technologies), 3 µL QuikSolution (Agilent Technologies) and 38.5 µL nuclease-free H_2O to give a final volume of 50 µL. Once mixed, 1 µL (2.5 U/µL) *pfuUltra* high fidelity polymerase (Agilent Technologies) was added and the reaction was subjected to cycling conditions (Table 5.6) inside a Mastercycler Gradient thermocycler (Eppendorf). Reactions were immediately incubated on ice following cycling conditions to cool to ≤ 37 °C, before 1 µL *Dpn*I (Agilent Technologies) was added to reactions, and incubated at 37 °C for 1 h. *E. coli* XL10 Gold cells were then transformed (Section 5.4.1.) using 2 µL of *Dpn*I-digested DNA mixture. After plasmid extraction (Section 5.10.), samples were submitted for DNA sequencing to confirm SDM (MWG Eurofins).

Table 5.6: PCR cycling conditions for Quikchange II XL protocol. *pET28b = 5.3 kb, *murC/lcp1* = 1.5 kb. Therefore total size is 6.8 kb = 7 minutes elongation step.

Segment	Cycles	Temp (°C)	Time
1	1	95	1 min
2	18	95	50 sec
		60	50 sec
		68	7 min*
3	1	68	7 min

5.6.1. Primers used for SDM

5.6.1.1. MurC primers

Table 5.7: Primers designed to produce Serine/Threonine to Aspartate mutants of *M. tuberculosis murC*.

T116D Forward	5'-GTTGATGGCCGGGCGCGACACATTGATGGTCACCG-3'
T116D Reverse	5'-CGGTGACCATCAATGTGTCGCGCCCGGCCATCAAC-3'
T129D Forward	5'-GCACGGCAAGACAACGGACACGTCCATGCTGATCG-3'
T129D Reverse	5'-CGATCAGCATGGACGTGTCCGTTGTCTTGCCGTGC-3'
S163D Forward	5'-CAACGCCCATCACGGCGATGGCGACTGTTTCGTCG-3'
S163D Reverse	5'-CGACGAAACAGTCGCCATCGCCGTGATGGGCGTTG-3'
T357D Forward	5'-CTACGCCCACCACCCGGACGAGATCAGCGCGACAC-3'
T357D Reverse	5'-GTGTCGCGCTGATCTCGTCCGGGTGGTGGGCGTAG-3'
S360D Forward	5'-CCACCCGACGGAGATCGACGCGACACTGGCGGCGG-3'
S360D Reverse	5'-CCGCCGCCAGTGTCGCGTCGATCTCCGTCGGGTGG-3'

5.6.1.2. Lcp1 primers

Table 5.8: Primers designed to produce Arginine to Alanine mutants of *M. tuberculosis lcp1*.

R70A Forward	5'-GGTCGGCCTGGACAGCGCTACCGACGCGCACGGCA-3'
R70A Reverse	5'-TGCCGTGCGCGTCGGTAGCGCTGTCCAGGCCGACC-3'
R118A Forward	5'-CGCAATCTCTATACCGGCGGACTCCTACGTCGCGG-3'
R118A Reverse	5'-CCGCGACGTAGGAGTCCGCCGGTATAGAGATTGCG-3'
R236A Forward	5'-AGCGCTCAGCTTCGTTGCCCAGCGGCATGATCTGC-3'
R236A Reverse	5'-GCAGATCATGCCGCTGGGCAACGAAGCTGAGCGCT-3'
R238A Forward	5'-CAGCTTCGTTCGCCAGGCGCATGATCTGCCCCGCG-3'
R238A Reverse	5'-CGCGGGGCAGATCATGCGCCTGGCGAACGAAGCTG-3'

5.7. Agarose gel DNA electrophoresis

DNA samples were added to Loading Dye (5 % glycerol (v/v), 0.04 % Bromophenol Blue), and loaded onto 1 % TAE (40 mM Tris-acetate, 1 mM EDTA, pH 8.0) agarose gel, containing 5 % Midori Green dye (v/v) (Nippon Genetics) alongside 1 kb DNA ladder (New England Biolabs). Gel was run in 1 × TAE buffer at 140 V, 400 mA for 50 min. The gel was then visualised using a Gel Doc XR (Bio-Rad) with Image Lab software.

5.8. DNA extraction from agarose gel

DNA was extracted from agarose gel with the use of QIAquick gel extraction kit (Qiagen). The DNA band was excised from the agarose gel using a scalpel, but by minimising the amount of excess agarose in the slice. The agarose slice was then weighed and 3 volumes of Buffer QG was added to every 1 volume of agarose (3 µL: 1 mg). The agarose was then incubated at 50 °C for 10 min, until the agarose had dissolved. The pH of the solution was then checked, by assuring the mixture was yellow in colour (indicator shows pH ≤ 7.5). One gel volume of isopropanol was added to the solution and mixed, before being transferred to a QIAquick spin column. The column was centrifuged at 18,000 × *g* for 1 min, and the flow through was discarded. This was repeated until all the sample was passed through the spin column. The spin column was then washed with 500 µL Buffer QG and centrifuged at 18,000 × *g* for 1 min. The flow through was discarded before 750 µL Buffer PE (Qiagen) was added to the spin column and was centrifuged at 18,000 × *g* for 1 min. The flow through was discarded and the spin column was again centrifuged at 18,000 × *g* for 1 min, to extract any residual flow through, which was then also discarded. The spin column was then inserted into a microcentrifuge tube, with 50 µL H_2O being added to the column for 2 min. The tube was then centrifuged at 18,000 × *g* for 1 min and the DNA in the flow through was retained and stored at -20 °C.

5.8.1. 'DNA clean up'

DNA that contained enzymes and buffers was cleaned by making use of QIAquick gel extraction kit (Qiagen). The DNA sample was mixed with 500 µL Buffer QG (Qiagen) and transferred to a QIAquick spin column. The column was then centrifuged at 18,000 × *g* for 1 min and the flow through was discarded. The spin column was then washed with 750 µL Buffer PE (Qiagen) by centrifugation at 18,000 × *g* for 1 min. The flow through was discarded and the spin column was again centrifuged at 18,000 × *g* for 1 min, to extract any residual flow through, which was then also discarded. The spin column was then inserted into a microcentrifuge tube, with 50 µL H_2O being added to the column for 2 min. The tube was then centrifuged at 18,000 × *g* for 1 min and the cleaned DNA in the flow through was retained and stored at -20 °C.

5.9. Genomic DNA extraction

Genomic DNA was extracted from 10 mL of bacterial culture, centrifuged at 3,300 × *g* for 15 min. Cells were re-suspended in 450 µL Buffer P1 (Qiagen) (or 50 mM Tris HCl pH 8, 10 mM EDTA and 100 µg/mL RNase A) and 50 µL of 10 mg/mL lysozyme, incubated at 37 °C overnight. 100 µL 10 % SDS was added to the suspension, alongside 50 µL 10 mg/mL proteinase K. The suspension was gently mixed and incubated at 55 °C for 4 h. 200 µL 5 M NaCl was added to the suspension, and was incubated at 65 °C for 15 min. Once the suspension was cooled, 500 µL chloroform ($CHCl_3$) was added, gently mixed for 5 min, and then centrifuged at 18,000 × *g* for 5 min. The aqueous layer was taken, 350 µL ice cold isopropanol added and gently mixed by inversion until DNA precipitation. Mixture was centrifuged at 18,000 × *g* for 10 min at 4 °C. The supernatant was removed, 1 mL ice cold 70 % ethanol (C_2H_5OH) was added and the tube was inverted gently. Sample was centrifuged at

18,000 × *g* for 5 min at 4 °C. Supernatant was removed and the DNA pellet was dried at room temperature for 20 min. The DNA was re-suspended in 50 µL H_2O.

5.10. Plasmid DNA extraction

Plasmid DNA was extracted from 5 mL bacterial culture with use of QIAprep spin miniprep kit (Qiagen). Cells were pelleted by centrifugation at 3,300 × *g* for 15 min and supernatant discarded. Pellets were re-suspended in 250 µL Buffer P1 containing RNAse A (Qiagen) and transferred into microcentrifuge tubes. After this, 250 µL Buffer P2 (Qiagen) was added and mixed by tube inversion, before the addition of 350 µL Buffer N3 (Qiagen) which was also mixed by tube inversion. Mixtures were then centrifuged at 18,000 × *g* for 10 min, and the supernatant was extracted and transferred into a miniprep spin column (Qiagen). The spin column was centrifuged at 18,000 × *g* for 1 min and the flow through was discarded. The spin column was then washed with 500 µL Buffer PB (Qiagen), by centrifugation at 18,000 × *g* for 1 min and the flow through was discarded. A final wash step of 750 µL Buffer PE (Qiagen) was then added to the spin column, and was centrifuged at 18,000 × *g* for 1 min, with the flow through being discarded. The spin column was again centrifuged at 18,000 × *g* for 1 min, to extract any residual flow through, which was then discarded. The spin column was then inserted into a microcentrifuge tube, with 50 µL H_2O being added to the column for 2 min. The tube was then centrifuged at 18,000 × *g* for 1 min and the plasmid DNA in the flow through was retained and stored at -20 °C.

5.11. Recombinant protein purification

Transformed *E. coli* BL21(DE3) were induced for protein production, once OD_{600} = 0.5 was reached, by the addition of 1 mL 1 M IPTG (0.1 M final concentration). Cultures were then incubated at 16 °C overnight with shaking. Cells were then harvested from *via* centrifugation at 6,000 × *g* for 15 min at 4 °C. The supernatant was discarded and the pellet was re-

suspended in Phosphate Buffered Saline (PBS). The re-suspended cells were then centrifuged at 3,300 × *g* for 18 min at 4 °C. The supernatant was once again discarded and the cell pellets were stored at -20 °C.

An imidazole gradient was produced from 'Purification Buffer' (2 ×) (Section 5.1.10.) and 2 M Imidazole pH 7.9. Protein containing pellet was defrosted then re-suspended in 20-30 mL 'lysis buffer' (Section 5.1.10.1.), with a proteinase inhibitor tablet (Roche) added. The mixture was subsequently sonicated using a soniprep 150 (MSE) (30 sec on, 30 sec off, 10 cycles). The solution was then centrifuged at 21,000 × *g* for 40 min at 4 °C, and the supernatant was retained (clarified lysate). Protein was purified using a HisTrap 1 mL/5 mL chelating HP column (GE healthcare) charged with 0.1 M $NiCl_2$ (binds to 6x Histidine tag). The clarified lysate/flow-through and imidazole fractions were collected.

Fractions that required further purification were loaded onto a 1 mL Q HP (anion exchange) column (GE healthcare) after dialysis. Protein was eluted using a NaCl gradient, made up from the same buffer system as the protein dialysis buffer (Section 5.1.11.). All fractions were collected.

5.12. SDS-PAGE

Protein fractions were mixed with SDS loading dye (2 % SDS, 125 mM Tris pH 6.8, 715 mM β-mercaptoethanol, 20 % glycerol made up to 1 mL with H_2O) and incubated at 100 °C for 10 min. Boiled samples, alongside PageRuler (Thermo Scientific) protein marker, were then ran on mini-PROTEAN TGX precast (Bio-Rad) gels at 200 V, 50 mA for 30 min. Gels were then incubated in Instant Blue stain (Expedeon) at room temperature with shaking for 1 h.

5.13. Western Blot

Once subjected to SDS-PAGE (Section 5.12.) samples are transferred onto nitrocellulose membrane in Western transfer buffer (Section 5.1.15.) at 20 V, 300 mA for 1 h. The membrane was then blocked with 5 % milk (Source of Bovine Serum Albumin (BSA)) in TBS-T (Section 5.1.16.) or Western blocking buffer (Roche), before incubation at room temperature with 0.1 % milk. The primary antibody (mouse anti-His, Qiagen) was added (1:1,000), the membrane was then washed in TBS-T (Section 5.1.16) and the secondary antibody (goat anti-mouse, Sigma Aldrich) was added (1:25,000). The membrane was further washed in TBS-T and then in TBS (Section 5.1.16.) only. Bands were then developed by incubation in BCIP solution.

5.14. Thin Layer Chromatography (TLC)

Samples were re-suspended in 10 µL solvent and loaded onto aluminium-backed silica gel plates (5735 silica gel 60 F254, Merck). The plates were then placed into a running tank containing 0.5 cm depth of appropriate solvent mixture for samples to migrate across the plate. The plates were removed from the running tank once the solvent front had reached around 1 cm from the top of the plate, before being left to dry. Once dry, plates were visualised using phosphomolybdic acid and heat. In the case of radiolabelled TLC, the plates were developed on a storage phosphor screen (Kodak). Bands were then visualised on a Personal Molecular Imager (Bio-Rad) and compared to known standards.

5.15. Ligand Binding Assays

Purified protein, total volume of 400 µL in a quartz cuvette, was allowed to equilibrate to 25 °C, inside a Fluorescence Spectrophotometer (Hitachi F-7000), with magnetic stirring (10 min). Intrinsic Tryptophan Fluorescence (ITF) was measured at excitation wavelength of 280

nm, and emission wavelength of 300-400 nm ($F_{emission}$), with an excitation and emission slit width of 5 nm. An emission maximum ($F_{emission}^{max}$) was recorded at 340 nm. Ligands were added to the cuvette in increasing concentration by 1 µL additions. The change in fluorescence emission ($\Delta F_{emission}$) was calculated by the subtraction of $F_{emission}$ (recorded 2 min after addition of ligand) from $F_{emission}^{max}$, this data was then plotted against ligand concentration [L]. Data analysis was performed using Microsoft Excel and then by fitting $\Delta F_{emission}$ at 340 nm vs [L] to a one-site specific binding equation using Graphpad Prism 5: $\Delta F_{emission} = F_{max} \times [L]/(K_d + [L])$.

5.16. Kinetic analysis of MurC site-directed mutants

The assay mixture for the kinetic analysis of MurC was made up to a total volume of 200 µL in a quartz cuvette, with 1 mm path length. The assay mixture consisted of MurC 'Reaction buffer' (Section 5.1.12.), 50 mM KCl, 1 mM dithiothreitol (DTT), 0.2 mM nicotinamide adenine dinucleotide (NADH), 2 mM phosphoenolpyruvate (PEP), 2 µL Pyruvate Kinase/Lactate Dehydrogenase (PKLDH, PK = 600-1000 U/mL, LDH = 900-1400 U/mL), 10 mM L-alanine, 0.1 µM MurC enzyme. In assays to assess the dependence of MurC on UDP-muramyl substrate, adenosine triphosphate (ATP), pH 7.5, was fixed at 1 mM. Assays to assess the dependence of MurC on ATP concentration fixed UDP-N-acetylmuramic acid at 1.5 mM and UDP-N-glycolylmuramic acid at 100 µM. The variable substrate concentration ranged between 10-400 µM. All assays were made up to final volume with H_2O, where L-alanine was added as the final component to initiate reactions. UDP-muramyl substrates were provided by BacWAN (University of Warwick). All assays were perfomed at the University of Warwick using a Cary 100 UV/Visible double beam spectrophotometer (Varian/Agilent) equipped with a thermostatically controlled cell changer, set to 37 °C, reading at 340 nm. Data collection was performed using Cary WinUV software. Data analysis was performed

using Microsoft Excel and by fitting to a Michaelis-Menten model using Graphpad Prism software ($Y = V_{max} \times X / (K_m + X)$).

5.17. Bioinformatic analytical tools

Homologous genes were identified using BLAST analysis against known search sequences (Altschul *et al*., 1990). Protein sequences were submitted to the I-TASSER server for 3D modelling prediction (Zhang, 2008, Roy *et al*., 2010, Yang *et al*., 2015, Yang and Zhang, 2015), as well as to the TMHMM server for transmembrane helix prediction (Sonnhammer *et al*., 1998). 3D models were then analysed against known homologous x-ray crystal structures using UCSF Chimera software (Pettersen *et al*., 2004). Multiple protein sequences were then aligned for comparison using ClustalW2 (Goujon *et al*., 2010), which could then be used to identify residues of interest within the protein sequences.

5.18. Production of plasmid vectors

5.18.1. Expression vectors

The gene to be expressed was amplified *via* PCR (Section 5.5.) with restriction sites corresponding to the selected expression plasmid. The gene product was then ran on 1 % TAE agarose gel electrophoresis (Section 5.7.). Bands at the correct molecular weight were excised and extracted (Section 5.8.). The extracted gene product as well as the expression plasmid was digested with the corresponding restriction enzymes for 1 h at 37 °C. A 'DNA clean-up' (Section 5.8.1.) was performed to remove the enzymes and buffer. A ligation was then performed between the gene product and expression vector using T4 DNA ligase (New England Biolabs) and the ligation mixtures were then transformed into *E. coli* Top 10 cells (Section 5.4.1.). Colonies that grew were selected and grown in 5 mL LB broth (50 µg/mL Kan) at 37 °C for 18 h, with shaking. The following day, plasmid DNA was extracted

(Section 5.10.) and the vector was digested using the corresponding restriction enzymes for 1 h at 37 °C. Samples were ran on 1 % TAE agarose gel electrophoresis (Section 5.7.). Lanes that showed bands at the correct molecular weight (for plasmid and gene product) were successful ligations and were therefore retained. Samples were submitted for DNA sequencing for confirmation of complete vector (MWG Eurofins).

5.18.2. Genetic knock-out mutant 'rescue' plasmid

The gene of interest (GOI) was amplified *via* PCR (Section 5.5.), with restriction sites for *Hind*III and *Eco*RV, with the product being ran on 1 % TAE agarose gel electrophoresis (Section 5.7.). Bands at the correct molecular weight were excised and extracted (Section 5.8.). The extracted gene was then digested with *Hind*III/*Eco*RV restriction enzymes for 1 h at 37 °C. A 'DNA clean-up' (Section 5.8.1.) was performed to remove the enzymes and buffer. The plasmids pSD26 and pMV306 were both double digested (pSD26 – *Eco*RV/*Xba*I, pMV306 – *Xba*I/*Hind*III) for 1 h at 37 °C. The digestion was then ran on 1 % TAE agarose gel electrophoresis (Section 5.7.) and bands at the correct molecular weight for the acetamidase gene (from pSD26) and cut pMV306 were excised and extracted (Section 5.8.)

A triple ligation between cut pMV306, acetamidase gene and GOI was performed using T4 DNA ligase (New England Biolabs). The ligation mixtures were then transformed into *E. coli* Top 10 cells (Section 5.4.1.) and colonies that grew were selected and grown in 5 mL LB broth (50 µg/mL Kan) at 37 °C for 18 h, with shaking. The following day, plasmid DNA was extracted (Section 5.10.) and the vector was triple digested using *Xba*I/*Hind*III/*Eco*RV enzymes for 1 h at 37 °C. Samples were ran on 1 % TAE agarose gel electrophoresis (Section 5.7.). Lanes that showed bands at the correct molecular weight (for plasmid, acetamidase gene and GOI) were successful ligations and were therefore retained. Samples were submitted for DNA sequencing for confirmation of complete vector (MWG Eurofins).

5.19. Production of genetic knock-out conditional mutant

Electrocompetent *M. smegmatis* MC^2 155 (5.3.2.) were electroporated (Section 5.4.2.) with the 'rescue' plasmid construct (Section 5.18.), as well as the original uncut pMV306 plasmid. The electroporated cells were plated onto TSA plates (50 µg/mL Kan) and incubated at 37 °C for 4 days. Successfully electroporated cells containing the construct were stored as a glycerol stock (500 µL) at -80 °C. The remaining cell culture were once again made electrocompetent (Section 5.3.2.) and also stored at -80 °C.

For the generation of the Δ*GOI* phasmid, around 1 kb of flanking region sequences both up and downstream of the GOI were cloned (with primers containing *Van*91I restriction sites on 5' ends (Section 5.5.1.)). These fragments were then digested with *Van*91I, ligated into *Van*91I digested p0004S (A gift from T. Hsu and W.R. Jacobs Jr., Albert Einstein College of Medicine, NY) and subsequently sequenced, by Dr G.S. Lloyd. The correctly sequenced DNA was then linearised by *Pac*I digestion and packaged into the temperature sensitive mycobacteriophage phAE159 by Mr Albel Singh, to yield the knockout phage phΔ*GOI* phasmid DNA.

In order to produce high-titre phage for transduction, phΔ*GOI* phasmid was electoporated into electrocompetent *M. smegmatis* WT cells (Section 5.4.2.), which were then incubated in 1 mL TSB at 30 °C for 4 h. After this, 100 µL of electroporated cells were added to 0.5 mL pre-grown *M. smegmatis* culture (OD_{600}=0.8-1), whilst the remaining cells were added to 100 µL of the pre-grown *M. smegmatis* culture. Both mixtures were added to separate tubes each containing 4 mL 'top agar' (Section 5.1.8.) molten at around 50 °C. This 'top agar' mixture is then added atop of 7H9 basal agar plates (Section 5.1.7.) and incubated at 30 °C for 3 days. After this, 1-2 plaques were extracted and added to 200 µL MP buffer (Section 5.1.13.) and phage was extracted by incubation at room temperature for at least 3 h. Once extracted, 5 µL

of lysate (as well as 5 µL of a 1:5 dilution in a separate tube) was added to 4 mL 'top agar' (containing *M. smegmatis* WT), which was added on top of 7H9 basal agar plates, and then incubated at 30 °C for 3 days. The plate containing around 1000 plaques (known as 'lacy' pattern) was selected and 5 mL MP buffer was added atop of the plate for at least 4 h at room temperature, to extract phage. Lysate was then filtered through a 0.2 µm syringe filter, and stored at 4 °C. The titre of the phage was checked by producing a 10-fold dilution series, whereby 100 µL of each dilution was spotted onto *M. smegmatis* WT containing 'top agar', on top of 7H9 basal agar. These plates were then incubated at 30 °C for 2-3 days, and the numbers of plaques were counted, to calculate the number of plaque forming units (PFU) titre of the lysate.

Electroporated *M. smegmatis* (containing 'rescue' plasmid (Section 5.18.)) was streaked from glycerol stock onto TSA plates (containing 50 µg/mL Kan) alongside *M. smegmatis* WT which were streaked onto antibiotic free TSA plates. All plates were incubated at 37 °C for 2 days. Colonies were picked using a sterile loop, and used to inoculate 5 mL TSB containing Tween 80 (0.05 % v/v) (+ Kan for vector containing cells). These cultures were then incubated at 37 °C with shaking for 2 days. Once grown, 500 µL of each 5 mL culture was used to seed into 50 mL TSB (containing the same additives as 5 mL cultures). These were incubated at 37 °C with shaking overnight. The following day, 50 mL cultures were centrifuged at 3,300 × *g* for 15 min, supernatant discarded and the pellet was washed twice with MP buffer to remove Tween 80 (3,300 × *g* for 15 min performed between washes). The pellets were then re-suspended in 2 mL MP buffer, with the addition of 1 mL MP buffer containing high-titre ph∆*GOI*. The suspensions were then incubated at 37 °C overnight, to induce lysogeny. The following day suspensions were centrifuged at 3,300 × *g* for 15 min, and the pellets were re-suspended in 2 mL TSB containing Tween 80 (0.05 % v/v) and acetamide (acet) (0.2 % w/v) each and incubated at 37 °C for 5 h. Cells were then plated

across TSB agar plates containing either Hyg only for WT, or Hyg, Kan and acet for 'rescue' plasmid containing cells. All plates were incubated at 37 °C for 3 days.

5.20. Growth of mutant strains for validation of genetic knock out

Hyg and Kan resistant colonies from knock out plates (Section 5.19.) were streaked onto TSB agar containing Kan, Hyg and acet and incubated at 37 °C for 3 days. Single colonies from these plates were looped and used to inoculate 5 mL TSB containing Kan, Hyg, acet and +/- Tween 80 (0.05 %). Cultures were incubated at 37 °C for 2 days. 2 mL of culture was centrifuged at 3,300 × *g* for 15 min and the pellet was twice washed in 10 mL TSB containing Tween 80 (0.05 %) (to remove acet). The pellet was re-suspended in 2 mL TSB with Tween 80 (0.05 %), and mixed with 8 mL TSB containing Hyg, Kan and Tween. Cultures were then incubated at 37 °C for 1-2 days to deplete intracellular GOI. 30 µL of the 10 mL culture was used to inoculate separate 3 mL TSB (containing Hyg, Kan and Tween 80) with +/- acet. Cultures were incubated at 37 °C overnight, and the following day were checked for growth.

5.21. Extraction of bound decaprenyl-1-monophosphate from ^{Mtb}Lcp1

2 mL of recombinant ^{Mtb}Lcp1 (41.2 µM) (WT and site-directed mutants) was mixed with 2 mL $CHCl_3$ and 1 mL methanol (CH_3OH) and incubated at 60 °C for 2 h. The samples were then centrifuged at 5,000 × *g* for 10 min. The lower organic phase was then collected and dried under nitrogen, before being re-suspended in 100 µL $CHCl_3$:CH_3OH (2:1, v/v). Samples were then analysed by TLC (Section 5.14.) and electrospray mass spectrometry (ES-MS) in the negative mode using a Micromass LCT mass spectrometer.

5.22. Biophysical analysis of ^{Mtb}Lcp1

Circular Dichroism (CD) was performed on WT and mutant ^{Mtb}Lcp1 at a concentration of around 2 µM. CD was recorded using a Jasco J-715 spectropolarimeter, using 300 µL protein sample in a 1 mm quartz cuvette. Spectra were recorded at 1 nm bandwidth, at 50 nm/min between a range of 260-190 nm. Once the spectra were recorded, 9 pt Savitzky-Golay smoothing was applied. CD was performed by the Birmingham Biophysical Characterisation Facility (BBCF). Data was analysed using Microsoft Excel software.

Analytical ultracentrifugation (AUC) was performed using WT and mutant ^{Mtb}Lcp1, as well as ^{Mtb}Lcp1 WT with increasing concentration of 'Compound 1' ligand (2, 4 and 6 mM respectively). All samples were between 13-19 µM of protein (between 0.35 and 0.5 absorbance at A280). 500 µL of each sample was centrifuged at 110,000 × *g* inside a Beckman XLI Ultracentrifuge. AUC was performed by the BBCF. All data was analysed using SEDFIT and Microsoft Excel software.

5.23. Cell-free ^{Mtb}Lcp1 protein activity assay

5.23.1. Extraction of *M. smegmatis* membranes

M. smegmatis WT was grown in 1 L TSB at 37 °C for 2 days. Cells were then harvested *via* centrifugation at 6,000 × *g* for 20 min at 4 °C. Supernatant was discarded, and cells were re-suspended in 30 mL PBS. Suspended cells were centrifuged at 3,300 × *g* for 15 min and the supernatant was discarded. *M. smegmatis* pellet was re-suspended in 24 mL 'Buffer A' (Section 5.1.14.), and then sonicated (60 on, 90 off, 15 cycles). Lysate was then centrifuged at 21,000 × *g* for 30 min at 4 °C. Supernatant and pellet were then split, but retained (pellet used for P60 cell wall fraction (Section 5.23.2.)). The supernatant was then centrifuged at 100,000 × *g* for 1 h 30 min. The supernatant was then discarded and 200 µL 'Buffer A' was added to

the top of the pellet and left on ice overnight. Pellets were then homogenised and transferred into a microcentrifuge tube. The suspension was further homogenised into a uniform preparation (not containing bubbles). Membrane preparation was aliquoted into 4 x 250 µL and stored at -20 °C.

5.23.2. Extraction of P60 cell wall fraction

The separated pellet (Section 5.23.1.) was re-suspended in 24 mL 'Buffer A' (Section 5.1.14.) and split into 2 centrifuge tubes (12 mL in each) and 18 mL Percoll (60 %) was added to each. Samples were centrifuged at 15,000 × *g* for 50 min. The middle layer was then extracted from each tube and combined using a plastic pipette. Combined samples were then equally separated into 2 centrifuge tubes and made up to 30 mL with 'Buffer A', and centrifuged at 39,000 × *g* for 20 min. Supernatant was discarded and pellet washed with 30 mL 'Buffer A'. Centrifugation at 39,000 × *g* for 20 min was repeated twice more with 30 mL 'Buffer A' washes in-between. Pellets were finally combined in 5 mL 'Buffer A' and stored at -20 °C.

5.23.3. Preparation of nascent mycobacterial peptidoglycan

M. smegmatis WT was grown to 1 L in TSB (Section 5.23.1.), and harvested *via* centrifugation at 6,000 × *g* for 15 min. Pelleted *M. smegmatis* was re-suspended in H_2O before addition of $CHCl_3$:CH_3OH (1:1 v/v) to create a $CHCl_3$:CH_3OH:H_2O (10:10:3 v/v/v) single phase. The suspension was centrifuged at 3,300 × *g* for 15 min, the resulting supernatant was discarded and the remaining pellet was re-suspended in 30-40 mL H_2O. The cell suspension was sonicated (Soniprep 150, MSE) (60 on, 90 off, 15 cycles), before the addition of 1 % (v/v) Triton X100 and incubation at room temperature for 16 h. The suspension was then centrifuged at 3,300 × *g* for 15 min. The remaining insoluble material was extracted three times with 2 % (w/v) SDS in PBS at 95 °C for 1 h, before being

consecutively washed with H_2O, 80 % (v/v) $(CH_3)_2CO$ in H_2O, then 100 % $(CH_3)_2CO$ and finally lyophilised to yield purified cell wall mAGP preparation.

Mycolic acids were removed from the mAGP by addition of 0.5 % (w/v) KOH in CH_3OH, incubated at 37 °C for 4 days. The suspension was then centrifuged at 27,000 × *g* for 20 min in order to collect the treated mAGP, which was then washed repeatedly with 100 % CH_3OH before the pellet was finally recovered by centrifugation at 27,000 × *g* for 20 min. The mycolic acids were then extracted from the pellet by treatment with diethyl ether ($(C_2H_5)_2O$) and then centrifugation at 27,000 × *g*, which was repeated three times to result in supernatant containing mycolic acid methyl esters, and pellet containing the remaining AGP complex. The pellet material was then hydrolysed with 0.2 M H_2SO_4 at 85 °C for 30 min, then neutralised with $BaCO_3$. After centrifugation at 27,000 × *g* for 30 min, the supernatant (containing solubilised arabinogalactan (AG)) was removed and the remaining insoluble pellet material (containing highly purified peptidoglycan (PPG)) was repeatedly washed with H_2O, before being lyophilised and stored at -20 °C.

5.23.4. Assay procedure

The assay mixture consisted of 0.5 mg purified *M. smegmatis* membranes (Section 5.23.1.), 1 mg P60 cell wall fraction (Section 5.23.2.), 0.2 mM ATP, 0.2 mM NADH, 15 mg PPG (Section 5.23.3.), ^{Mtb}Lcp1 (0.3 µM (+) or 0.9 µM (++)) and reactions were initiated by the addition of 0.1 µCi UDP-[^{14}C]-D-Galactose (Specific activity 300 mCi/mmol, ARC radiochemicals) (final volume 200 µL, made up with 'Buffer A' (Section 5.1.14.)). All reactions were incubated at 37 °C for 17 h. Reactions were quenched by the addition of 1333 µL of $CHCl_3$:CH_3OH (1:1 v/v) to form a $CHCl_3$:CH_3OH:H_2O (10:10:3 v/v/v) single phase. Samples were centrifuged at 18,000 × *g* for 15 min, supernatant was extracted, and the pellet retained for further processing. The supernatant was dried under nitrogen, before re-

suspension in 2 mL $CHCl_3$:H_2O (1:1 v/v). The lower organic phase was separated from the upper aqueous phase, and both extractions were retained for later analysis (Figure 5.1). The earlier insoluble pellet was re-suspended in 1 mL H_2O:C_2H_5OH:$(C_2H_5)_2O$:pyridine(C_5H_5N):NH_4OH (15:15:5:1:0.017, v/v/v/v/v) (named 'E-soak'), and incubated at room temperature for 2 h. Samples were then centrifuged at 18,000 × *g* for 15 min and the supernatant was extracted and kept for later analysis (Figure 5.1). The remaining insoluble pellet was re-suspended in 200 µL H_2O for further analysis. Each extraction stage (Upper aqueous, Lower organic, 'E-soak' and insoluble pellet) were quantified for radioactivity through liquid scintillation counting, by using 10 % of material with 5 mL EcoScintA (National Diagnostics).

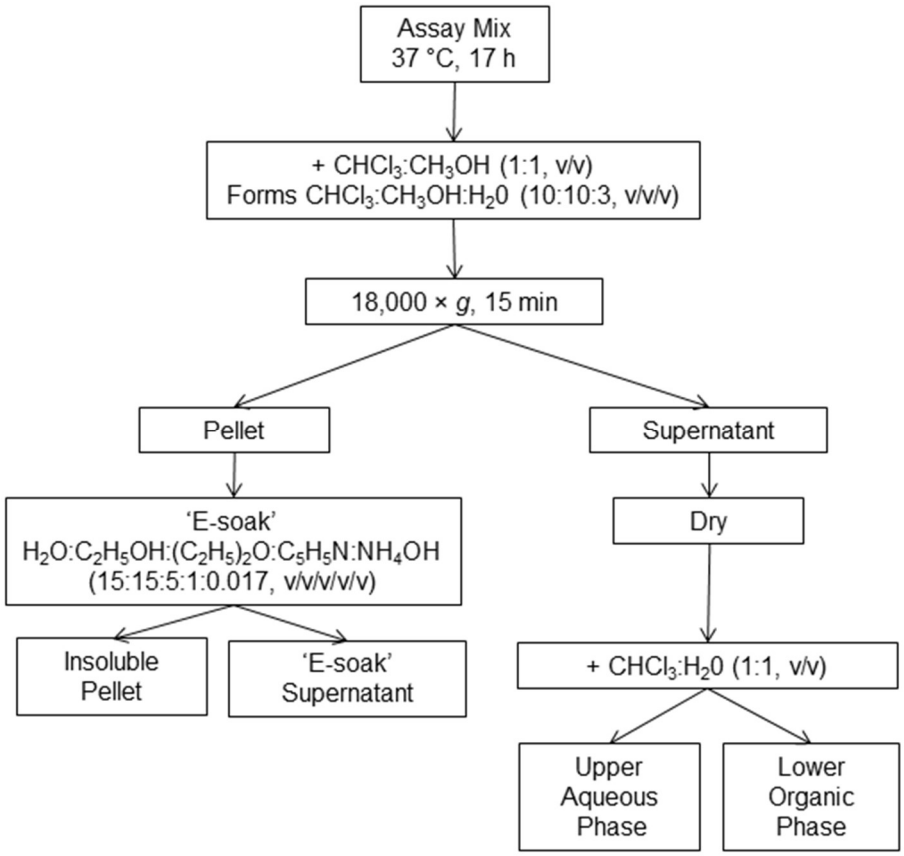

Figure 5.1: A flow-chart showing the fractionation steps of Lcp1 activity assay. The assay mixture was incubated at 37 °C for 18 h before being quenched with $CHCl_3$:CH_3OH (1:1 v/v) to form a $CHCl_3$:CH_3OH:H_2O (10:10:3 v/v/v) single phase. This was then separated into soluble and insoluble fractions. The soluble fraction was again separated into aqueous and organic fractions using $CHCl_3$:H_2O (1:1 v/v). The insoluble fraction was solubilised in 'E-soak', before once again being separated into soluble and insoluble fractions.

5.23.5. Analysis of assay extractions

The Lower organic and 'E-soak' samples were dried under nitrogen. The Lower organic material samples were then re-suspended in 10 µL $CHCl_3$. This resuspension as well as 10 µL of the Upper aqueous material were ran on TLC plates (Section 5.14.).

'E-soak' material was re-suspended in 30 µL LDS loading dye with DTT (Expedeon), before being loaded onto mini-PROTEAN Tris-Tricine (Bio-Rad) SDS-PAGE (200 V, 80 mA, 1.5 h). The gel was then transferred onto a 0.2 µm nitrocellulose membrane. Transfer was run at 20 V, 300 mA for 1.5 h. Western blot was performed (Section 5.13.), but the primary antibody used was CS-35 (mouse anti-hexa-arabinofuranoside) (1:1000 dilution)) and the secondary antibody was goat anti-mouse SC-2070 (1:25,000). After visualisation, the blot was then additionally visualised by autoradiography, through exposure to X-ray film (Kodak X-Omat). Bands on the membrane and film were then analysed against each other.

The insoluble material was hydrolysed in 2 M trifluoroacetic acid at 120 °C for 2 h, before being dried under nitrogen. Samples were then reduced by addition of 100 µL sodium borohydride ($NaBH_4$) (10 mg/mL in 1 M NH_4OH:C_2H_5OH (1:1 v/v)) incubated at room temperature for 2 h. Glacial acetic acid was then added (3 drops) before being dried under nitrogen. A further 3 drops of 10 % glacial acetic acid in CH_3OH was added to samples before being dried under nitrogen. 100 % CH_3OH was then added by 3 drops before being dried under nitrogen, which was repeated, in order to thoroughly dry samples. Once dry, samples were per-*O*-acetylated by the addition of 100 µL acetic anhydride and incubation at 120 °C for 1 h. Once cooled, 100 µL toluene was added and samples were dried under nitrogen. Samples were then re-suspended in 4 mL $CHCl_3$:H_2O (1:1 v/v) and centrifuged at 3,300 × *g* for 15 min. The lower phase was extracted and dried under nitrogen. The resulting per-*O*-

acetylated alditol acetates were re-suspended in 10 µL $CHCl_3$ and ran on a TLC plate (Section 5.14.).

5.24. Phosphatase activity assay

Geranyl pyrophosphate (GPP) and Geranylgeranyl pyrophosphate (GGPP) ammonium salts were added to separate tubes (25 µL of 1 mg/mL stock, Sigma Aldrich) and dried under nitrogen. Samples were then re-suspended in 'Buffer A' (Section 5.1.14.) before the addition of ^{Mtb}Lcp1 WT and site-directed mutants to a final volume of 200 µL. Samples were then incubated at 37 °C for 3 h, before the addition of 1 mL $CHCl_3$. The lower phase was then extracted from each sample, and dried under nitrogen. Each sample was then re-suspended in 10 µL $CHCl_3$ and ran on a TLC plate (Section 5.14.).

Chapter 6

References

6. REFERENCES

ALDERWICK, L. J., DOVER, L. G., VEERAPEN, N., GURCHA, S. S., KREMER, L., ROPER, D. L., PATHAK, A. K., REYNOLDS, R. C. & BESRA, G. S. 2008. Expression, purification and characterisation of soluble GlfT and the identification of a novel galactofuranosyltransferase Rv3782 involved in priming GlfT-mediated galactan polymerisation in Mycobacterium tuberculosis. *Protein Expr Purif,* 58**,** 332-41.

ALDERWICK, L. J., LLOYD, G. S., GHADBANE, H., MAY, J. W., BHATT, A., EGGELING, L., FUTTERER, K. & BESRA, G. S. 2011a. The C-terminal domain of the Arabinosyltransferase Mycobacterium tuberculosis EmbC is a lectin-like carbohydrate binding module. *PLoS Pathog,* 7**,** e1001299.

ALDERWICK, L. J., LLOYD, G. S., LLOYD, A. J., LOVERING, A. L., EGGELING, L. & BESRA, G. S. 2011b. Biochemical characterization of the Mycobacterium tuberculosis phosphoribosyl-1-pyrophosphate synthetase. *Glycobiology,* 21**,** 410-25.

ALDERWICK, L. J., RADMACHER, E., SEIDEL, M., GANDE, R., HITCHEN, P. G., MORRIS, H. R., DELL, A., SAHM, H., EGGELING, L. & BESRA, G. S. 2005. Deletion of Cg-emb in corynebacterianeae leads to a novel truncated cell wall arabinogalactan, whereas inactivation of Cg-ubiA results in an arabinan-deficient mutant with a cell wall galactan core. *J Biol Chem,* 280**,** 32362-71.

ALDERWICK, L. J., SEIDEL, M., SAHM, H., BESRA, G. S. & EGGELING, L. 2006. Identification of a novel arabinofuranosyltransferase (AftA) involved in cell wall arabinan biosynthesis in Mycobacterium tuberculosis. *J Biol Chem,* 281**,** 15653-61.

ALTSCHUL, S. F., GISH, W., MILLER, W., MYERS, E. W. & LIPMAN, D. J. 1990. Basic local alignment search tool. *J Mol Biol,* 215**,** 403-10.

AMAR, C. & VILKAS, E. 1973. [Isolation of arabinose phosphate from the walls of Mycobacterium tuberculosis H 37 Ra]. *C R Acad Sci Hebd Seances Acad Sci D,* 277**,** 1949-51.

AMIN, A. G., GOUDE, R., SHI, L., ZHANG, J., CHATTERJEE, D. & PARISH, T. 2008. EmbA is an essential arabinosyltransferase in Mycobacterium tuberculosis. *Microbiology,* 154**,** 240-8.

ANDRE, G., KULAKAUSKAS, S., CHAPOT-CHARTIER, M. P., NAVET, B., DEGHORAIN, M., BERNARD, E., HOLS, P. & DUFRENE, Y. F. 2010. Imaging the nanoscale organization of peptidoglycan in living Lactococcus lactis cells. *Nat Commun,* 1**,** 27.

ARBEX, M. A., VARELLA MDE, C., SIQUEIRA, H. R. & MELLO, F. A. 2010. Antituberculosis drugs: drug interactions, adverse effects, and use in special situations. Part 1: first-line drugs. *J Bras Pneumol,* 36**,** 626-40.

ARMSTRONG, D. J., MCCAUSLAND, E. M. & WRIGHT, G. D. 2007. Hypertrophic pulmonary osteoarthropathy (HPOA) (Pierre Marie-Bamberger syndrome): two cases presenting as acute inflammatory arthritis. Description and review of the literature. *Rheumatol Int,* 27**,** 399-402.

BADET, B., VERMOOTE, P., HAUMONT, P. Y., LEDERER, F. & LEGOFFIC, F. 1987. Glucosamine synthetase from Escherichia coli: purification, properties, and glutamine-utilizing site location. *Biochemistry,* 26**,** 1940-8.

BAER, C. E., IAVARONE, A. T., ALBER, T. & SASSETTI, C. M. 2014. Biochemical and spatial coincidence in the provisional Ser/Thr protein kinase interaction network of Mycobacterium tuberculosis. *J Biol Chem,* 289**,** 20422-33.

BAKKER-WOUDENBERG, I. A., VAN VIANEN, W., VAN SOOLINGEN, D., VERBRUGH, H. A. & VAN AGTMAEL, M. A. 2005. Antimycobacterial agents differ with respect to their bacteriostatic versus bactericidal activities in relation to time of exposure, mycobacterial growth phase, and their use in combination. *Antimicrob Agents Chemother,* 49**,** 2387-98.

BARTHOLOMEW, J. W. & MITTWER, T. 1952. The Gram stain. *Bacteriol Rev,* 16**,** 1-29.

BASAVANNACHARYA, C., ROBERTSON, G., MUNSHI, T., KEEP, N. H. & BHAKTA, S. 2010. ATP-dependent MurE ligase in Mycobacterium tuberculosis: biochemical and structural characterisation. *Tuberculosis (Edinb),* 90**,** 16-24.

BATT, S. M., JABEEN, T., BHOWRUTH, V., QUILL, L., LUND, P. A., EGGELING, L., ALDERWICK, L. J., FUTTERER, K. & BESRA, G. S. 2012. Structural basis of inhibition of Mycobacterium tuberculosis DprE1 by benzothiazinone inhibitors. *Proc Natl Acad Sci U S A,* 109**,** 11354-9.

BEEBY, M., GUMBART, J. C., ROUX, B. & JENSEN, G. J. 2013. Architecture and assembly of the Gram-positive cell wall. *Mol Microbiol,* 88**,** 664-72.

BELANGER, A. E., BESRA, G. S., FORD, M. E., MIKUSOVA, K., BELISLE, J. T., BRENNAN, P. J. & INAMINE, J. M. 1996. The embAB genes of Mycobacterium avium encode an arabinosyl transferase involved in cell wall arabinan biosynthesis that is the target for the antimycobacterial drug ethambutol. *Proc Natl Acad Sci U S A,* 93**,** 11919-24.

BENSON, T. E., WALSH, C. T. & HOGLE, J. M. 1996. The structure of the substrate-free form of MurB, an essential enzyme for the synthesis of bacterial cell walls. *Structure,* 4**,** 47-54.

BERTRAND, J. A., AUGER, G., MARTIN, L., FANCHON, E., BLANOT, D., LE BELLER, D., VAN HEIJENOORT, J. & DIDEBERG, O. 1999. Determination of the MurD mechanism through crystallographic analysis of enzyme complexes. *J Mol Biol,* 289**,** 579-90.

BESRA, G. S., KHOO, K. H., MCNEIL, M. R., DELL, A., MORRIS, H. R. & BRENNAN, P. J. 1995. A new interpretation of the structure of the mycolyl-arabinogalactan complex of Mycobacterium tuberculosis as revealed through characterization of oligoglycosylalditol fragments by fast-atom bombardment mass spectrometry and 1H nuclear magnetic resonance spectroscopy. *Biochemistry,* 34**,** 4257-66.

BHAMIDI, S., SCHERMAN, M. S., JONES, V., CRICK, D. C., BELISLE, J. T., BRENNAN, P. J. & MCNEIL, M. R. 2011. Detailed structural and quantitative analysis reveals the spatial organization of the cell walls of in vivo grown Mycobacterium leprae and in vitro grown Mycobacterium tuberculosis. *J Biol Chem,* 286**,** 23168-77.

BHAMIDI, S., SCHERMAN, M. S., RITHNER, C. D., PRENNI, J. E., CHATTERJEE, D., KHOO, K. H. & MCNEIL, M. R. 2008. The identification and location of succinyl residues and the characterization of the interior arabinan region allow for a model of the complete primary structure of Mycobacterium tuberculosis mycolyl arabinogalactan. *J Biol Chem,* 283**,** 12992-3000.

BHATT, A. & JACOBS, W. R., JR. 2009. Gene essentiality testing in mycobacterium smegmatis using specialized transduction. *Methods Mol Biol,* 465**,** 325-36.

BHATT, A., MOLLE, V., BESRA, G. S., JACOBS, W. R., JR. & KREMER, L. 2007. The Mycobacterium tuberculosis FAS-II condensing enzymes: their role in mycolic acid biosynthesis, acid-fastness, pathogenesis and in future drug development. *Mol Microbiol,* 64**,** 1442-54.

BIRCH, H. L., ALDERWICK, L. J., BHATT, A., RITTMANN, D., KRUMBACH, K., SINGH, A., BAI, Y., LOWARY, T. L., EGGELING, L. & BESRA, G. S. 2008.

Biosynthesis of mycobacterial arabinogalactan: identification of a novel alpha(1-->3) arabinofuranosyltransferase. *Mol Microbiol,* 69**,** 1191-206.

BIRCH, H. L., ALDERWICK, L. J., RITTMANN, D., KRUMBACH, K., ETTERICH, H., GRZEGORZEWICZ, A., MCNEIL, M. R., EGGELING, L. & BESRA, G. S. 2009. Identification of a terminal rhamnopyranosyltransferase (RptA) involved in Corynebacterium glutamicum cell wall biosynthesis. *J Bacteriol,* 191**,** 4879-87.

BOISSIER, F., BARDOU, F., GUILLET, V., UTTENWEILER-JOSEPH, S., DAFFE, M., QUEMARD, A. & MOUREY, L. 2006. Further insight into S-adenosylmethionine-dependent methyltransferases: structural characterization of Hma, an enzyme essential for the biosynthesis of oxygenated mycolic acids in Mycobacterium tuberculosis. *J Biol Chem,* 281**,** 4434-45.

BOUAKAZE, C., KEYSER, C., DE MARTINO, S. J., SOUGAKOFF, W., VEZIRIS, N., DABERNAT, H. & LUDES, B. 2010. Identification and genotyping of Mycobacterium tuberculosis complex species by use of a SNaPshot Minisequencing-based assay. *J Clin Microbiol,* 48**,** 1758-66.

BOUHSS, A., MENGIN-LECREULX, D., LE BELLER, D. & VAN HEIJENOORT, J. 1999. Topological analysis of the MraY protein catalysing the first membrane step of peptidoglycan synthesis. *Mol Microbiol,* 34**,** 576-85.

BRENNAN, P. J. & NIKAIDO, H. 1995. The envelope of mycobacteria. *Annu Rev Biochem,* 64**,** 29-63.

BRIKEN, V., PORCELLI, S. A., BESRA, G. S. & KREMER, L. 2004. Mycobacterial lipoarabinomannan and related lipoglycans: from biogenesis to modulation of the immune response. *Mol Microbiol,* 53**,** 391-403.

BURNETT, G. & KENNEDY, E. P. 1954. The enzymatic phosphorylation of proteins. *J Biol Chem,* 211**,** 969-80.

CALVANESE, L., FALCIGNO, L., MAGLIONE, C., MARASCO, D., RUGGIERO, A., SQUEGLIA, F., BERISIO, R. & D'AURIA, G. 2014. Structural and binding properties of the PASTA domain of PonA2, a key penicillin binding protein from Mycobacterium tuberculosis. *Biopolymers,* 101**,** 712-9.

CAMBAU, E. & DRANCOURT, M. 2014. Steps towards the discovery of Mycobacterium tuberculosis by Robert Koch, 1882. *Clin Microbiol Infect,* 20**,** 196-201.

CANDELA, T. & FOUET, A. 2006. Poly-gamma-glutamate in bacteria. *Mol Microbiol,* 60**,** 1091-8.

CHANG, Y. H., LABGOLD, M. R. & RICHARDS, J. H. 1990. Altering enzymatic activity: recruitment of carboxypeptidase activity into an RTEM beta-lactamase/penicillin-binding protein 5 chimera. *Proc Natl Acad Sci U S A,* 87**,** 2823-7.

CHEN, P., SHI, M., FENG, G. D., LIU, J. Y., WANG, B. J., SHI, X. D., MA, L., LIU, X. D., YANG, Y. N., DAI, W., LIU, T. T., HE, Y., LI, J. G., HAO, X. K. & ZHAO, G. 2012. A highly efficient Ziehl-Neelsen stain: identifying de novo intracellular Mycobacterium tuberculosis and improving detection of extracellular M. tuberculosis in cerebrospinal fluid. *J Clin Microbiol,* 50**,** 1166-70.

COLE, S. T., BROSCH, R., PARKHILL, J., GARNIER, T., CHURCHER, C., HARRIS, D., GORDON, S. V., EIGLMEIER, K., GAS, S., BARRY, C. E., 3RD, TEKAIA, F., BADCOCK, K., BASHAM, D., BROWN, D., CHILLINGWORTH, T., CONNOR, R., DAVIES, R., DEVLIN, K., FELTWELL, T., GENTLES, S., HAMLIN, N., HOLROYD, S., HORNSBY, T., JAGELS, K., KROGH, A., MCLEAN, J., MOULE, S., MURPHY, L., OLIVER, K., OSBORNE, J., QUAIL, M. A., RAJANDREAM, M. A., ROGERS, J., RUTTER, S., SEEGER, K., SKELTON, J., SQUARES, R., SQUARES, S., SULSTON, J. E., TAYLOR, K., WHITEHEAD, S. & BARRELL, B.

G. 1998. Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence. *Nature,* 393**,** 537-44.

COOK, G. M., BERNEY, M., GEBHARD, S., HEINEMANN, M., COX, R. A., DANILCHANKA, O. & NIEDERWEIS, M. 2009. Physiology of mycobacteria. *Adv Microb Physiol,* 55**,** 81-182, 318-9.

CORDILLOT, M., DUBEE, V., TRIBOULET, S., DUBOST, L., MARIE, A., HUGONNET, J. E., ARTHUR, M. & MAINARDI, J. L. 2013. In vitro cross-linking of Mycobacterium tuberculosis peptidoglycan by L,D-transpeptidases and inactivation of these enzymes by carbapenems. *Antimicrob Agents Chemother,* 57**,** 5940-5.

COTTAREL, G. & WIERZBOWSKI, J. 2007. Combination drugs, an emerging option for antibacterial therapy. *Trends Biotechnol,* 25**,** 547-55.

COULOMBE, F., DIVANGAHI, M., VEYRIER, F., DE LESELEUC, L., GLEASON, J. L., YANG, Y., KELLIHER, M. A., PANDEY, A. K., SASSETTI, C. M., REED, M. B. & BEHR, M. A. 2009. Increased NOD2-mediated recognition of N-glycolyl muramyl dipeptide. *J Exp Med,* 206**,** 1709-16.

CRICK, D. C., MAHAPATRA, S. & BRENNAN, P. J. 2001. Biosynthesis of the arabinogalactan-peptidoglycan complex of Mycobacterium tuberculosis. *Glycobiology,* 11**,** 107R-118R.

D'ELIA, M. A., HENDERSON, J. A., BEVERIDGE, T. J., HEINRICHS, D. E. & BROWN, E. D. 2009. The N-acetylmannosamine transferase catalyzes the first committed step of teichoic acid assembly in Bacillus subtilis and Staphylococcus aureus. *J Bacteriol,* 191**,** 4030-4.

D'ELIA, M. A., MILLAR, K. E., BEVERIDGE, T. J. & BROWN, E. D. 2006a. Wall teichoic acid polymers are dispensable for cell viability in Bacillus subtilis. *J Bacteriol,* 188**,** 8313-6.

D'ELIA, M. A., PEREIRA, M. P., CHUNG, Y. S., ZHAO, W., CHAU, A., KENNEY, T. J., SULAVIK, M. C., BLACK, T. A. & BROWN, E. D. 2006b. Lesions in teichoic acid biosynthesis in Staphylococcus aureus lead to a lethal gain of function in the otherwise dispensable pathway. *J Bacteriol,* 188**,** 4183-9.

DAHL, J. L. 2004. Electron microscopy analysis of Mycobacterium tuberculosis cell division. *FEMS Microbiol Lett,* 240**,** 15-20.

DAS, D., HERVE, M., FEUERHELM, J., FARR, C. L., CHIU, H. J., ELSLIGER, M. A., KNUTH, M. W., KLOCK, H. E., MILLER, M. D., GODZIK, A., LESLEY, S. A., DEACON, A. M., MENGIN-LECREULX, D. & WILSON, I. A. 2011. Structure and function of the first full-length murein peptide ligase (Mpl) cell wall recycling protein. *PLoS One,* 6**,** e17624.

DATTA, P., DASGUPTA, A., BHAKTA, S. & BASU, J. 2002. Interaction between FtsZ and FtsW of Mycobacterium tuberculosis. *J Biol Chem,* 277**,** 24983-7.

DAVIES, P. D. 2003. The role of DOTS in tuberculosis treatment and control. *Am J Respir Med,* 2**,** 203-9.

DEVA, T., BAKER, E. N., SQUIRE, C. J. & SMITH, C. A. 2006. Structure of Escherichia coli UDP-N-acetylmuramoyl:L-alanine ligase (MurC). *Acta Crystallogr D Biol Crystallogr,* 62**,** 1466-74.

DIANISKOVA, P., KORDULAKOVA, J., SKOVIEROVA, H., KAUR, D., JACKSON, M., BRENNAN, P. J. & MIKUSOVA, K. 2011. Investigation of ABC transporter from mycobacterial arabinogalactan biosynthetic cluster. *Gen Physiol Biophys,* 30**,** 239-50.

DOOLITTLE, R. F. 2005. Evolutionary aspects of whole-genome biology. *Curr Opin Struct Biol,* 15**,** 248-53.

DRAPER, P., KHOO, K. H., CHATTERJEE, D., DELL, A. & MORRIS, H. R. 1997. Galactosamine in walls of slow-growing mycobacteria. *Biochem J,* 327 (Pt 2)**,** 519-25.

EBERHARDT, A., HOYLAND, C. N., VOLLMER, D., BISLE, S., CLEVERLEY, R. M., JOHNSBORG, O., HAVARSTEIN, L. S., LEWIS, R. J. & VOLLMER, W. 2012. Attachment of capsular polysaccharide to the cell wall in Streptococcus pneumoniae. *Microb Drug Resist,* 18**,** 240-55.

EL ZOEIBY, A., SANSCHAGRIN, F. & LEVESQUE, R. C. 2003. Structure and function of the Mur enzymes: development of novel inhibitors. *Mol Microbiol,* 47**,** 1-12.

ELDER, N. C. 1992. Extrapulmonary tuberculosis. A review. *Arch Fam Med,* 1**,** 91-8.

EVELAND, S. S., POMPLIANO, D. L. & ANDERSON, M. S. 1997. Conditionally lethal Escherichia coli murein mutants contain point defects that map to regions conserved among murein and folyl poly-gamma-glutamate ligases: identification of a ligase superfamily. *Biochemistry,* 36**,** 6223-9.

FALK, P. J., ERVIN, K. M., VOLK, K. S. & HO, H. T. 1996. Biochemical evidence for the formation of a covalent acyl-phosphate linkage between UDP-N-acetylmuramate and ATP in the Escherichia coli UDP-N-acetylmuramate:L-alanine ligase-catalyzed reaction. *Biochemistry,* 35**,** 1417-22.

FALK, S. P. & WEISBLUM, B. 2013. Phosphorylation of the Streptococcus pneumoniae cell wall biosynthesis enzyme MurC by a eukaryotic-like Ser/Thr kinase. *FEMS Microbiol Lett,* 340**,** 19-23.

FARNIA, P., MOHAMMAD, R. M., MERZA, M. A., TABARSI, P., ZHAVNERKO, G. K., IBRAHIM, T. A., KUAN, H. O., GHANAVEI, J., FARNIA, P., RANJBAR, R., POLESCHUYK, N. N., TITOV, L. P., OWLIA, P., KAZAMPOUR, M., SETAREH, M., SHEIKOLSLAMI, M., MIGLIORI, G. B. & VELAYATI, A. A. 2010. Growth and cell-division in extensive (XDR) and extremely drug resistant (XXDR) tuberculosis strains: transmission and atomic force observation. *Int J Clin Exp Med,* 3**,** 308-14.

FENG, Z. & BARLETTA, R. G. 2003. Roles of Mycobacterium smegmatis D-alanine:D-alanine ligase and D-alanine racemase in the mechanisms of action of and resistance to the peptidoglycan inhibitor D-cycloserine. *Antimicrob Agents Chemother,* 47**,** 283-91.

FIUZA, M., CANOVA, M. J., PATIN, D., LETEK, M., ZANELLA-CLEON, I., BECCHI, M., MATEOS, L. M., MENGIN-LECREULX, D., MOLLE, V. & GIL, J. A. 2008. The MurC ligase essential for peptidoglycan biosynthesis is regulated by the serine/threonine protein kinase PknA in Corynebacterium glutamicum. *J Biol Chem,* 283**,** 36553-63.

GAGO, G., KURTH, D., DIACOVICH, L., TSAI, S. C. & GRAMAJO, H. 2006. Biochemical and structural characterization of an essential acyl coenzyme A carboxylase from Mycobacterium tuberculosis. *J Bacteriol,* 188**,** 477-86.

GEORGE, K. M., YUAN, Y., SHERMAN, D. R. & BARRY, C. E., 3RD 1995. The biosynthesis of cyclopropanated mycolic acids in Mycobacterium tuberculosis. Identification and functional analysis of CMAS-2. *J Biol Chem,* 270**,** 27292-8.

GILL, W. P., HARIK, N. S., WHIDDON, M. R., LIAO, R. P., MITTLER, J. E. & SHERMAN, D. R. 2009. A replication clock for Mycobacterium tuberculosis. *Nat Med,* 15**,** 211-4.

GLICKMAN, M. S. 2003. The mmaA2 gene of Mycobacterium tuberculosis encodes the distal cyclopropane synthase of the alpha-mycolic acid. *J Biol Chem,* 278**,** 7844-9.

GLICKMAN, M. S., COX, J. S. & JACOBS, W. R., JR. 2000. A novel mycolic acid cyclopropane synthetase is required for cording, persistence, and virulence of Mycobacterium tuberculosis. *Mol Cell,* 5**,** 717-27.

GOHAR, M., YANG, W., STRONG, W., VOLKENING, K., LEYSTRA-LANTZ, C. & STRONG, M. J. 2009. Tau phosphorylation at threonine-175 leads to fibril formation and enhanced cell death: implications for amyotrophic lateral sclerosis with cognitive impairment. *J Neurochem,* 108**,** 634-43.

GOUJON, M., MCWILLIAM, H., LI, W., VALENTIN, F., SQUIZZATO, S., PAERN, J. & LOPEZ, R. 2010. A new bioinformatics analysis tools framework at EMBL-EBI. *Nucleic Acids Res,* 38**,** W695-9.

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 Pathog,* 7**,** e1002251.

GROVER, S., ALDERWICK, L. J., MISHRA, A. K., KRUMBACH, K., MARIENHAGEN, J., EGGELING, L., BHATT, A. & BESRA, G. S. 2014. Benzothiazinones mediate killing of Corynebacterineae by blocking decaprenyl phosphate recycling involved in cell wall biosynthesis. *J Biol Chem,* 289**,** 6177-87.

GRZEGORZEWICZ, A. E., PHAM, H., GUNDI, V. A., SCHERMAN, M. S., NORTH, E. J., HESS, T., JONES, V., GRUPPO, V., BORN, S. E., KORDULAKOVA, J., CHAVADI, S. S., MORISSEAU, C., LENAERTS, A. J., LEE, R. E., MCNEIL, M. R. & JACKSON, M. 2012. Inhibition of mycolic acid transport across the Mycobacterium tuberculosis plasma membrane. *Nat Chem Biol,* 8**,** 334-41.

HANCOCK, I. C., CARMAN, S., BESRA, G. S., BRENNAN, P. J. & WAITE, E. 2002. Ligation of arabinogalactan to peptidoglycan in the cell wall of Mycobacterium smegmatis requires concomitant synthesis of the two wall polymers. *Microbiology,* 148**,** 3059-67.

HAYHURST, E. J., KAILAS, L., HOBBS, J. K. & FOSTER, S. J. 2008. Cell wall peptidoglycan architecture in Bacillus subtilis. *Proc Natl Acad Sci U S A,* 105**,** 14603-8.

HERSHKOVITZ, I., DONOGHUE, H. D., MINNIKIN, D. E., BESRA, G. S., LEE, O. Y., GERNAEY, A. M., GALILI, E., ESHED, V., GREENBLATT, C. L., LEMMA, E., BAR-GAL, G. K. & SPIGELMAN, M. 2008. Detection and molecular characterization of 9,000-year-old Mycobacterium tuberculosis from a Neolithic settlement in the Eastern Mediterranean. *PLoS One,* 3**,** e3426.

HETT, E. C., CHAO, M. C. & RUBIN, E. J. 2010. Interaction and modulation of two antagonistic cell wall enzymes of mycobacteria. *PLoS Pathog,* 6**,** e1001020.

HETT, E. C. & RUBIN, E. J. 2008. Bacterial growth and cell division: a mycobacterial perspective. *Microbiol Mol Biol Rev,* 72**,** 126-56, table of contents.

HOANG, T. T., MA, Y., STERN, R. J., MCNEIL, M. R. & SCHWEIZER, H. P. 1999. Construction and use of low-copy number T7 expression vectors for purification of problem proteins: purification of mycobacterium tuberculosis RmlD and pseudomonas aeruginosa LasI and RhlI proteins, and functional analysis of purified RhlI. *Gene,* 237**,** 361-71.

HOFFMANN, C., LEIS, A., NIEDERWEIS, M., PLITZKO, J. M. & ENGELHARDT, H. 2008. Disclosure of the mycobacterial outer membrane: cryo-electron tomography and vitreous sections reveal the lipid bilayer structure. *Proc Natl Acad Sci U S A,* 105**,** 3963-7.

HUANG, H., SCHERMAN, M. S., D'HAEZE, W., VEREECKE, D., HOLSTERS, M., CRICK, D. C. & MCNEIL, M. R. 2005. Identification and active expression of the Mycobacterium tuberculosis gene encoding 5-phospho-{alpha}-d-ribose-1-diphosphate: decaprenyl-phosphate 5-phosphoribosyltransferase, the first enzyme

committed to decaprenylphosphoryl-d-arabinose synthesis. *J Biol Chem,* 280**,** 24539-43.

HUGONNET, J. E., TREMBLAY, L. W., BOSHOFF, H. I., BARRY, C. E., 3RD & BLANCHARD, J. S. 2009. Meropenem-clavulanate is effective against extensively drug-resistant Mycobacterium tuberculosis. *Science,* 323**,** 1215-8.

JACKSON, M., RAYNAUD, C., LANEELLE, M. A., GUILHOT, C., LAURENT-WINTER, C., ENSERGUEIX, D., GICQUEL, B. & DAFFE, M. 1999. Inactivation of the antigen 85C gene profoundly affects the mycolate content and alters the permeability of the Mycobacterium tuberculosis cell envelope. *Mol Microbiol,* 31**,** 1573-87.

JANKUTE, M., COX, J. A., HARRISON, J. & BESRA, G. S. 2015. Assembly of the Mycobacterial Cell Wall. *Annu Rev Microbiol,* 69**,** 405-23.

JANKUTE, M., GROVER, S., RANA, A. K. & BESRA, G. S. 2012. Arabinogalactan and lipoarabinomannan biosynthesis: structure, biogenesis and their potential as drug targets. *Future Microbiol,* 7**,** 129-47.

JENKINS, W. T. 1991. The pyruvate kinase-coupled assay for ATPases: a critical analysis. *Anal Biochem,* 194**,** 136-9.

KAUR, D., GUERIN, M. E., SKOVIEROVA, H., BRENNAN, P. J. & JACKSON, M. 2009. Chapter 2: Biogenesis of the cell wall and other glycoconjugates of Mycobacterium tuberculosis. *Adv Appl Microbiol,* 69**,** 23-78.

KAWAI, Y., MARLES-WRIGHT, J., CLEVERLEY, R. M., EMMINS, R., ISHIKAWA, S., KUWANO, M., HEINZ, N., BUI, N. K., HOYLAND, C. N., OGASAWARA, N., LEWIS, R. J., VOLLMER, W., DANIEL, R. A. & ERRINGTON, J. 2011. A widespread family of bacterial cell wall assembly proteins. *EMBO J,* 30**,** 4931-41.

KAYA, S., YOKOYAMA, K., ARAKI, Y. & ITO, E. 1984. N-acetylmannosaminyl(1----4)N-acetylglucosamine, a linkage unit between glycerol teichoic acid and peptidoglycan in cell walls of several Bacillus strains. *J Bacteriol,* 158**,** 990-6.

KIM, D. H., LEES, W. J., KEMPSELL, K. E., LANE, W. S., DUNCAN, K. & WALSH, C. T. 1996. Characterization of a Cys115 to Asp substitution in the Escherichia coli cell wall biosynthetic enzyme UDP-GlcNAc enolpyruvyl transferase (MurA) that confers resistance to inactivation by the antibiotic fosfomycin. *Biochemistry,* 35**,** 4923-8.

KNECHEL, N. A. 2009. Tuberculosis: pathophysiology, clinical features, and diagnosis. *Crit Care Nurse,* 29**,** 34-43; quiz 44.

KONNO, K., FELDMANN, F. M. & MCDERMOTT, W. 1967. Pyrazinamide susceptibility and amidase activity of tubercle bacilli. *Am Rev Respir Dis,* 95**,** 461-9.

KREMER, L., DOVER, L. G., MORBIDONI, H. R., VILCHEZE, C., MAUGHAN, W. N., BAULARD, A., TU, S. C., HONORE, N., DERETIC, V., SACCHETTINI, J. C., LOCHT, C., JACOBS, W. R., JR. & BESRA, G. S. 2003. Inhibition of InhA activity, but not KasA activity, induces formation of a KasA-containing complex in mycobacteria. *J Biol Chem,* 278**,** 20547-54.

KREMER, L., DOVER, L. G., MOREHOUSE, C., HITCHIN, P., EVERETT, M., MORRIS, H. R., DELL, A., BRENNAN, P. J., MCNEIL, M. R., FLAHERTY, C., DUNCAN, K. & BESRA, G. S. 2001a. Galactan biosynthesis in Mycobacterium tuberculosis. Identification of a bifunctional UDP-galactofuranosyltransferase. *J Biol Chem,* 276**,** 26430-40.

KREMER, L., NAMPOOTHIRI, K. M., LESJEAN, S., DOVER, L. G., GRAHAM, S., BETTS, J., BRENNAN, P. J., MINNIKIN, D. E., LOCHT, C. & BESRA, G. S. 2001b. Biochemical characterization of acyl carrier protein (AcpM) and malonyl-CoA:AcpM transacylase (mtFabD), two major components of Mycobacterium tuberculosis fatty acid synthase II. *J Biol Chem,* 276**,** 27967-74.

KUROSU, M., MAHAPATRA, S., NARAYANASAMY, P. & CRICK, D. C. 2007. Chemoenzymatic synthesis of Park's nucleotide: toward the development of high-throughput screening for MraY inhibitors. *Tetrahedron Letters,* 48**,** 799-803.

LAVOLLAY, M., ARTHUR, M., FOURGEAUD, M., DUBOST, L., MARIE, A., VEZIRIS, N., BLANOT, D., GUTMANN, L. & MAINARDI, J. L. 2008. The peptidoglycan of stationary-phase Mycobacterium tuberculosis predominantly contains cross-links generated by L,D-transpeptidation. *J Bacteriol,* 190**,** 4360-6.

LAWRENCE, S. A., TITUS, S. A., FERGUSON, J., HEINEMAN, A. L., TAYLOR, S. M. & MORAN, R. G. 2014. Mammalian mitochondrial and cytosolic folylpolyglutamate synthetase maintain the subcellular compartmentalization of folates. *J Biol Chem,* 289**,** 29386-96.

LEA-SMITH, D. J., PYKE, J. S., TULL, D., MCCONVILLE, M. J., COPPEL, R. L. & CRELLIN, P. K. 2007. The reductase that catalyzes mycolic motif synthesis is required for efficient attachment of mycolic acids to arabinogalactan. *J Biol Chem,* 282**,** 11000-8.

LEE, R. E., BRENNAN, P. J. & BESRA, G. S. 1997. Mycobacterial arabinan biosynthesis: the use of synthetic arabinoside acceptors in the development of an arabinosyl transfer assay. *Glycobiology,* 7**,** 1121-8.

LEGER, J., KEMPF, M., LEE, G. & BRANDT, R. 1997. Conversion of serine to aspartate imitates phosphorylation-induced changes in the structure and function of microtubule-associated protein tau. *J Biol Chem,* 272**,** 8441-6.

LI, Y., ZHOU, Y., MA, Y. & LI, X. 2011. Design and synthesis of novel cell wall inhibitors of Mycobacterium tuberculosis GlmM and GlmU. *Carbohydr Res,* 346**,** 1714-20.

LIGER, D., MASSON, A., BLANOT, D., VAN HEIJENOORT, J. & PARQUET, C. 1995. Over-production, purification and properties of the uridine-diphosphate-N-acetylmuramate:L-alanine ligase from Escherichia coli. *Eur J Biochem,* 230**,** 80-7.

LOVERING, A. L., SAFADI, S. S. & STRYNADKA, N. C. 2012. Structural perspective of peptidoglycan biosynthesis and assembly. *Annu Rev Biochem,* 81**,** 451-78.

MA, Y., MILLS, J. A., BELISLE, J. T., VISSA, V., HOWELL, M., BOWLIN, K., SCHERMAN, M. S. & MCNEIL, M. 1997. Determination of the pathway for rhamnose biosynthesis in mycobacteria: cloning, sequencing and expression of the Mycobacterium tuberculosis gene encoding alpha-D-glucose-1-phosphate thymidylyltransferase. *Microbiology,* 143 (Pt 3)**,** 937-45.

MA, Y., STERN, R. J., SCHERMAN, M. S., VISSA, V. D., YAN, W., JONES, V. C., ZHANG, F., FRANZBLAU, S. G., LEWIS, W. H. & MCNEIL, M. R. 2001. Drug targeting Mycobacterium tuberculosis cell wall synthesis: genetics of dTDP-rhamnose synthetic enzymes and development of a microtiter plate-based screen for inhibitors of conversion of dTDP-glucose to dTDP-rhamnose. *Antimicrob Agents Chemother,* 45**,** 1407-16.

MAHAIRAS, G. G., SABO, P. J., HICKEY, M. J., SINGH, D. C. & STOVER, C. K. 1996. Molecular analysis of genetic differences between Mycobacterium bovis BCG and virulent M. bovis. *J Bacteriol,* 178**,** 1274-82.

MAHAPATRA, S., CRICK, D. C. & BRENNAN, P. J. 2000. Comparison of the UDP-N-acetylmuramate:L-alanine ligase enzymes from Mycobacterium tuberculosis and Mycobacterium leprae. *J Bacteriol,* 182**,** 6827-30.

MAHAPATRA, S., SCHERMAN, H., BRENNAN, P. J. & CRICK, D. C. 2005. N Glycolylation of the nucleotide precursors of peptidoglycan biosynthesis of Mycobacterium spp. is altered by drug treatment. *J Bacteriol,* 187**,** 2341-7.

MANABE, Y. C., DANNENBERG, A. M., JR., TYAGI, S. K., HATEM, C. L., YODER, M., WOOLWINE, S. C., ZOOK, B. C., PITT, M. L. & BISHAI, W. R. 2003. Different

strains of Mycobacterium tuberculosis cause various spectrums of disease in the rabbit model of tuberculosis. *Infect Immun,* 71**,** 6004-11.

MARGOLIN, W. 2005. FtsZ and the division of prokaryotic cells and organelles. *Nat Rev Mol Cell Biol,* 6**,** 862-71.

MARRAKCHI, H., LANEELLE, G. & QUEMARD, A. 2000. InhA, a target of the antituberculous drug isoniazid, is involved in a mycobacterial fatty acid elongation system, FAS-II. *Microbiology,* 146 (Pt 2)**,** 289-96.

MARRAKCHI, H., LANEELLE, M. A. & DAFFE, M. 2014. Mycolic acids: structures, biosynthesis, and beyond. *Chem Biol,* 21**,** 67-85.

MARTTILA, H. J., SOINI, H., EEROLA, E., VYSHNEVSKAYA, E., VYSHNEVSKIY, B. I., OTTEN, T. F., VASILYEF, A. V. & VILJANEN, M. K. 1998. A Ser315Thr substitution in KatG is predominant in genetically heterogeneous multidrug-resistant Mycobacterium tuberculosis isolates originating from the St. Petersburg area in Russia. *Antimicrob Agents Chemother,* 42**,** 2443-5.

MASSON, M., MOLNAR, E., DONOGHUE, H. D., BESRA, G. S., MINNIKIN, D. E., WU, H. H., LEE, O. Y., BULL, I. D. & PALFI, G. 2013. Osteological and biomolecular evidence of a 7000-year-old case of hypertrophic pulmonary osteopathy secondary to tuberculosis from neolithic hungary. *PLoS One,* 8**,** e78252.

MAZUREK, J., IGNATOWICZ, L., KALLENIUS, G., SVENSON, S. B., PAWLOWSKI, A. & HAMASUR, B. 2012. Divergent effects of mycobacterial cell wall glycolipids on maturation and function of human monocyte-derived dendritic cells. *PLoS One,* 7**,** e42515.

MCCLURE, W. R. & CECH, C. L. 1978. On the mechanism of rifampicin inhibition of RNA synthesis. *J Biol Chem,* 253**,** 8949-56.

MCNEIL, M., DAFFE, M. & BRENNAN, P. J. 1990. Evidence for the nature of the link between the arabinogalactan and peptidoglycan of mycobacterial cell walls. *J Biol Chem,* 265**,** 18200-6.

MCNEIL, M., DAFFE, M. & BRENNAN, P. J. 1991. Location of the mycolyl ester substituents in the cell walls of mycobacteria. *J Biol Chem,* 266**,** 13217-23.

MCNEIL, M., WALLNER, S. J., HUNTER, S. W. & BRENNAN, P. J. 1987. Demonstration that the galactosyl and arabinosyl residues in the cell-wall arabinogalactan of Mycobacterium leprae and Mycobacterium tuberculosis are furanoid. *Carbohydr Res,* 166**,** 299-308.

MCNEIL, M. R., ROBUCK, K. G., HARTER, M. & BRENNAN, P. J. 1994. Enzymatic evidence for the presence of a critical terminal hexa-arabinoside in the cell walls of Mycobacterium tuberculosis. *Glycobiology,* 4**,** 165-73.

MEROUEH, S. O., BENCZE, K. Z., HESEK, D., LEE, M., FISHER, J. F., STEMMLER, T. L. & MOBASHERY, S. 2006. Three-dimensional structure of the bacterial cell wall peptidoglycan. *Proc Natl Acad Sci U S A,* 103**,** 4404-9.

MIKUSOVA, K., BELANOVA, M., KORDULAKOVA, J., HONDA, K., MCNEIL, M. R., MAHAPATRA, S., CRICK, D. C. & BRENNAN, P. J. 2006. Identification of a novel galactosyl transferase involved in biosynthesis of the mycobacterial cell wall. *J Bacteriol,* 188**,** 6592-8.

MIKUSOVA, K., HUANG, H., YAGI, T., HOLSTERS, M., VEREECKE, D., D'HAEZE, W., SCHERMAN, M. S., BRENNAN, P. J., MCNEIL, M. R. & CRICK, D. C. 2005. Decaprenylphosphoryl arabinofuranose, the donor of the D-arabinofuranosyl residues of mycobacterial arabinan, is formed via a two-step epimerization of decaprenylphosphoryl ribose. *J Bacteriol,* 187**,** 8020-5.

MIKUSOVA, K., MIKUS, M., BESRA, G. S., HANCOCK, I. & BRENNAN, P. J. 1996. Biosynthesis of the linkage region of the mycobacterial cell wall. *J Biol Chem,* 271**,** 7820-8.

MIKUSOVA, K., SLAYDEN, R. A., BESRA, G. S. & BRENNAN, P. J. 1995. Biogenesis of the mycobacterial cell wall and the site of action of ethambutol. *Antimicrob Agents Chemother,* 39**,** 2484-9.

MINNIKIN, D. E., LEE, O. Y., WU, H. H., NATARAJ, V., DONOGHUE, H. D., RIDELL, M., WATANABE, M., ALDERWICK, L., BHATT, A. & BESRA, G. S. 2015. Pathophysiological Implications of Cell Envelope Structure in Mycobacterium tuberculosis and Related Taxa. *Tuberculosis - Expanding Knowledge, Dr. Wellman Ribón (Ed.)***,** 145-175.

MOHAMMADI, T., VAN DAM, V., SIJBRANDI, R., VERNET, T., ZAPUN, A., BOUHSS, A., DIEPEVEEN-DE BRUIN, M., NGUYEN-DISTECHE, M., DE KRUIJFF, B. & BREUKINK, E. 2011. Identification of FtsW as a transporter of lipid-linked cell wall precursors across the membrane. *Embo Journal,* 30**,** 1425-1432.

MOL, C. D., BROOUN, A., DOUGAN, D. R., HILGERS, M. T., TARI, L. W., WIJNANDS, R. A., KNUTH, M. W., MCREE, D. E. & SWANSON, R. V. 2003. Crystal structures of active fully assembled substrate- and product-bound complexes of UDP-N-acetylmuramic acid:L-alanine ligase (MurC) from Haemophilus influenzae. *J Bacteriol,* 185**,** 4152-62.

MORAYYA, S., AWASTHY, D., YADAV, R., AMBADY, A. & SHARMA, U. 2015. Revisiting the essentiality of glutamate racemase in Mycobacterium tuberculosis. *Gene,* 555**,** 269-76.

MUNSHI, T., GUPTA, A., EVANGELOPOULOS, D., GUZMAN, J. D., GIBBONS, S., KEEP, N. H. & BHAKTA, S. 2013. Characterisation of ATP-dependent Mur ligases involved in the biogenesis of cell wall peptidoglycan in Mycobacterium tuberculosis. *PLoS One,* 8**,** e60143.

MURPHY, D. J. & BROWN, J. R. 2007. Identification of gene targets against dormant phase Mycobacterium tuberculosis infections. *BMC Infect Dis,* 7**,** 84.

NAGARAJAN, S. N., UPADHYAY, S., CHAWLA, Y., KHAN, S., NAZ, S., SUBRAMANIAN, J., GANDOTRA, S. & NANDICOORI, V. K. 2015. Protein kinase A (PknA) of Mycobacterium tuberculosis is independently activated and is critical for growth in vitro and survival of the pathogen in the host. *J Biol Chem,* 290**,** 9626-45.

NEUHAUS, F. C. & BADDILEY, J. 2003. A continuum of anionic charge: structures and functions of D-alanyl-teichoic acids in gram-positive bacteria. *Microbiol Mol Biol Rev,* 67**,** 686-723.

OJHA, A., ANAND, M., BHATT, A., KREMER, L., JACOBS, W. R., JR. & HATFULL, G. F. 2005. GroEL1: a dedicated chaperone involved in mycolic acid biosynthesis during biofilm formation in mycobacteria. *Cell,* 123**,** 861-73.

PANG, Y., LU, J., WANG, Y., SONG, Y., WANG, S. & ZHAO, Y. 2013. Study of the rifampin monoresistance mechanism in Mycobacterium tuberculosis. *Antimicrob Agents Chemother,* 57**,** 893-900.

PENG, W., ZOU, L., BHAMIDI, S., MCNEIL, M. R. & LOWARY, T. L. 2012. The galactosamine residue in mycobacterial arabinogalactan is alpha-linked. *J Org Chem,* 77**,** 9826-32.

PETTERSEN, E. F., GODDARD, T. D., HUANG, C. C., COUCH, G. S., GREENBLATT, D. M., MENG, E. C. & FERRIN, T. E. 2004. UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem,* 25**,** 1605-12.

PHETSUKSIRI, B., JACKSON, M., SCHERMAN, H., MCNEIL, M., BESRA, G. S., BAULARD, A. R., SLAYDEN, R. A., DEBARBER, A. E., BARRY, C. E., 3RD,

BAIRD, M. S., CRICK, D. C. & BRENNAN, P. J. 2003. Unique mechanism of action of the thiourea drug isoxyl on Mycobacterium tuberculosis. *J Biol Chem,* 278**,** 53123-30.

PITARQUE, S., LARROUY-MAUMUS, G., PAYRE, B., JACKSON, M., PUZO, G. & NIGOU, J. 2008. The immunomodulatory lipoglycans, lipoarabinomannan and lipomannan, are exposed at the mycobacterial cell surface. *Tuberculosis (Edinb),* 88**,** 560-5.

PRISIC, S., DANKWA, S., SCHWARTZ, D., CHOU, M. F., LOCASALE, J. W., KANG, C. M., BEMIS, G., CHURCH, G. M., STEEN, H. & HUSSON, R. N. 2010. Extensive phosphorylation with overlapping specificity by Mycobacterium tuberculosis serine/threonine protein kinases. *Proc Natl Acad Sci U S A,* 107**,** 7521-6.

PRISIC, S. & HUSSON, R. N. 2014. Mycobacterium tuberculosis Serine/Threonine Protein Kinases. *Microbiol Spectr,* 2.

PUECH, V., BAYAN, N., SALIM, K., LEBLON, G. & DAFFE, M. 2000. Characterization of the in vivo acceptors of the mycoloyl residues transferred by the corynebacterial PS1 and the related mycobacterial antigens 85. *Mol Microbiol,* 35**,** 1026-41.

RAMASWAMY, S. V., DOU, S. J., RENDON, A., YANG, Z., CAVE, M. D. & GRAVISS, E. A. 2004. Genotypic analysis of multidrug-resistant Mycobacterium tuberculosis isolates from Monterrey, Mexico. *J Med Microbiol,* 53**,** 107-13.

RASSAM, P., COPELAND, N. A., BIRKHOLZ, O., TOTH, C., CHAVENT, M., DUNCAN, A. L., CROSS, S. J., HOUSDEN, N. G., KAMINSKA, R., SEGER, U., QUINN, D. M., GARROD, T. J., SANSOM, M. S., PIEHLER, J., BAUMANN, C. G. & KLEANTHOUS, C. 2015. Supramolecular assemblies underpin turnover of outer membrane proteins in bacteria. *Nature,* 523**,** 333-6.

RAYMOND, J. B., MAHAPATRA, S., CRICK, D. C. & PAVELKA, M. S., JR. 2005. Identification of the namH gene, encoding the hydroxylase responsible for the N-glycolylation of the mycobacterial peptidoglycan. *J Biol Chem,* 280**,** 326-33.

RECK, F., MARMOR, S., FISHER, S. & WUONOLA, M. A. 2001. Inhibitors of the bacterial cell wall biosynthesis enzyme MurC. *Bioorg Med Chem Lett,* 11**,** 1451-4.

RIFAT, M., HALL, J., OLDMEADOW, C., HUSAIN, A. & MILTON, A. H. 2015. Health system delay in treatment of multidrug resistant tuberculosis patients in Bangladesh. *BMC Infect Dis,* 15**,** 526.

ROY, A., KUCUKURAL, A. & ZHANG, Y. 2010. I-TASSER: a unified platform for automated protein structure and function prediction. *Nat Protoc,* 5**,** 725-38.

RUIZ, N. 2008. Bioinformatics identification of MurJ (MviN) as the peptidoglycan lipid II flippase in Escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America,* 105**,** 15553-15557.

SACCO, E., HUGONNET, J. E., JOSSEAUME, N., CREMNITER, J., DUBOST, L., MARIE, A., PATIN, D., BLANOT, D., RICE, L. B., MAINARDI, J. L. & ARTHUR, M. 2010. Activation of the L,D-transpeptidation peptidoglycan cross-linking pathway by a metallo-D,D-carboxypeptidase in Enterococcus faecium. *Mol Microbiol,* 75**,** 874-85.

SAFI, H., LINGARAJU, S., AMIN, A., KIM, S., JONES, M., HOLMES, M., MCNEIL, M., PETERSON, S. N., CHATTERJEE, D., FLEISCHMANN, R. & ALLAND, D. 2013. Evolution of high-level ethambutol-resistant tuberculosis through interacting mutations in decaprenylphosphoryl-beta-D-arabinose biosynthetic and utilization pathway genes. *Nat Genet,* 45**,** 1190-7.

SAKAMOTO, K. 2012. The pathology of Mycobacterium tuberculosis infection. *Vet Pathol,* 49**,** 423-39.

SAMBANDAN, D., DAO, D. N., WEINRICK, B. C., VILCHEZE, C., GURCHA, S. S., OJHA, A., KREMER, L., BESRA, G. S., HATFULL, G. F. & JACOBS, W. R., JR. 2013. Keto-mycolic acid-dependent pellicle formation confers tolerance to drug-sensitive Mycobacterium tuberculosis. *MBio,* 4**,** e00222-13.

SANI, M., HOUBEN, E. N., GEURTSEN, J., PIERSON, J., DE PUNDER, K., VAN ZON, M., WEVER, B., PIERSMA, S. R., JIMENEZ, C. R., DAFFE, M., APPELMELK, B. J., BITTER, W., VAN DER WEL, N. & PETERS, P. J. 2010. Direct visualization by cryo-EM of the mycobacterial capsular layer: a labile structure containing ESX-1-secreted proteins. *PLoS Pathog,* 6**,** e1000794.

SASSETTI, C. M., BOYD, D. H. & RUBIN, E. J. 2003. Genes required for mycobacterial growth defined by high density mutagenesis. *Mol Microbiol,* 48**,** 77-84.

SCHLEIFER, K. H. & KANDLER, O. 1972. Peptidoglycan Types of Bacterial Cell-Walls and Their Taxonomic Implications. *Bacteriological Reviews,* 36**,** 407-477.

SCHOONMAKER, M. K., BISHAI, W. R. & LAMICHHANE, G. 2014. Nonclassical transpeptidases of Mycobacterium tuberculosis alter cell size, morphology, the cytosolic matrix, protein localization, virulence, and resistance to beta-lactams. *J Bacteriol,* 196**,** 1394-402.

SCORPIO, A. & ZHANG, Y. 1996. Mutations in pncA, a gene encoding pyrazinamidase/nicotinamidase, cause resistance to the antituberculous drug pyrazinamide in tubercle bacillus. *Nat Med,* 2**,** 662-7.

SEIDEL, M., ALDERWICK, L. J., BIRCH, H. L., SAHM, H., EGGELING, L. & BESRA, G. S. 2007. Identification of a novel arabinofuranosyltransferase AftB involved in a terminal step of cell wall arabinan biosynthesis in Corynebacterianeae, such as Corynebacterium glutamicum and Mycobacterium tuberculosis. *J Biol Chem,* 282**,** 14729-40.

SHAM, L. T., BUTLER, E. K., LEBAR, M. D., KAHNE, D., BERNHARDT, T. G. & RUIZ, N. 2014. Bacterial cell wall. MurJ is the flippase of lipid-linked precursors for peptidoglycan biogenesis. *Science,* 345**,** 220-2.

SHI, W., ZHANG, X., JIANG, X., YUAN, H., LEE, J. S., BARRY, C. E., 3RD, WANG, H., ZHANG, W. & ZHANG, Y. 2011. Pyrazinamide inhibits trans-translation in Mycobacterium tuberculosis. *Science,* 333**,** 1630-2.

SILHAVY, T. J., KAHNE, D. & WALKER, S. 2010. The bacterial cell envelope. *Cold Spring Harb Perspect Biol,* 2**,** a000414.

SKOVIEROVA, H., LARROUY-MAUMUS, G., PHAM, H., BELANOVA, M., BARILONE, N., DASGUPTA, A., MIKUSOVA, K., GICQUEL, B., GILLERON, M., BRENNAN, P. J., PUZO, G., NIGOU, J. & JACKSON, M. 2010. Biosynthetic origin of the galactosamine substituent of Arabinogalactan in Mycobacterium tuberculosis. *J Biol Chem,* 285**,** 41348-55.

SKOVIEROVA, H., LARROUY-MAUMUS, G., ZHANG, J., KAUR, D., BARILONE, N., KORDULAKOVA, J., GILLERON, M., GUADAGNINI, S., BELANOVA, M., PREVOST, M. C., GICQUEL, B., PUZO, G., CHATTERJEE, D., BRENNAN, P. J., NIGOU, J. & JACKSON, M. 2009. AftD, a novel essential arabinofuranosyltransferase from mycobacteria. *Glycobiology,* 19**,** 1235-47.

SMITH, C. A. 2006. Structure, function and dynamics in the mur family of bacterial cell wall ligases. *J Mol Biol,* 362**,** 640-55.

SMITH, I. 2003. Mycobacterium tuberculosis pathogenesis and molecular determinants of virulence. *Clin Microbiol Rev,* 16**,** 463-96.

SMITH, S., WITKOWSKI, A. & JOSHI, A. K. 2003. Structural and functional organization of the animal fatty acid synthase. *Prog Lipid Res,* 42**,** 289-317.

SOMOSKOVI, A., PARSONS, L. M. & SALFINGER, M. 2001. The molecular basis of resistance to isoniazid, rifampin, and pyrazinamide in Mycobacterium tuberculosis. *Respir Res,* 2**,** 164-8.

SONNHAMMER, E. L., VON HEIJNE, G. & KROGH, A. 1998. A hidden Markov model for predicting transmembrane helices in protein sequences. *Proc Int Conf Intell Syst Mol Biol,* 6**,** 175-82.

SOTGIU, G., D'AMBROSIO, L., CENTIS, R., BOTHAMLEY, G., CIRILLO, D. M., DE LORENZO, S., GUENTHER, G., KLIIMAN, K., MUETTERLEIN, R., SPINU, V., VILLAR, M., ZELLWEGER, J. P., SANDGREN, A., HUITRIC, E., LANGE, C., MANISSERO, D. & MIGLIORI, G. B. 2011. TB and M/XDR-TB infection control in European TB reference centres: the Achilles' heel? *Eur Respir J,* 38**,** 1221-3.

SREEVATSAN, S., STOCKBAUER, K. E., PAN, X., KREISWIRTH, B. N., MOGHAZEH, S. L., JACOBS, W. R., JR., TELENTI, A. & MUSSER, J. M. 1997. Ethambutol resistance in Mycobacterium tuberculosis: critical role of embB mutations. *Antimicrob Agents Chemother,* 41**,** 1677-81.

SRIVASTAVA, S., AYYAGARI, A., DHOLE, T. N., NYATI, K. K. & DWIVEDI, S. K. 2009. emb nucleotide polymorphisms and the role of embB306 mutations in Mycobacterium tuberculosis resistance to ethambutol. *Int J Med Microbiol,* 299**,** 269-80.

STEFAN, M. I. & LE NOVERE, N. 2013. Cooperative binding. *PLoS Comput Biol,* 9**,** e1003106.

STERN, R. J., LEE, T. Y., LEE, T. J., YAN, W., SCHERMAN, M. S., VISSA, V. D., KIM, S. K., WANNER, B. L. & MCNEIL, M. R. 1999. Conversion of dTDP-4-keto-6-deoxyglucose to free dTDP-4-keto-rhamnose by the rmlC gene products of Escherichia coli and Mycobacterium tuberculosis. *Microbiology,* 145 (Pt 3)**,** 663-71.

STOCK, J. B., NINFA, A. J. & STOCK, A. M. 1989. Protein phosphorylation and regulation of adaptive responses in bacteria. *Microbiol Rev,* 53**,** 450-90.

STRYCH, U., PENLAND, R. L., JIMENEZ, M., KRAUSE, K. L. & BENEDIK, M. J. 2001. Characterization of the alanine racemases from two mycobacteria. *FEMS Microbiol Lett,* 196**,** 93-8.

SUN, Q. & GAMBLIN, T. C. 2009. Pseudohyperphosphorylation causing AD-like changes in tau has significant effects on its polymerization. *Biochemistry,* 48**,** 6002-11.

SWOBODA, J. G., CAMPBELL, J., MEREDITH, T. C. & WALKER, S. 2010. Wall teichoic acid function, biosynthesis, and inhibition. *Chembiochem,* 11**,** 35-45.

TAKAYAMA, K., WANG, C. & BESRA, G. S. 2005. Pathway to synthesis and processing of mycolic acids in Mycobacterium tuberculosis. *Clin Microbiol Rev,* 18**,** 81-101.

THAKUR, M. & CHAKRABORTI, P. K. 2008. Ability of PknA, a mycobacterial eukaryotic-type serine/threonine kinase, to transphosphorylate MurD, a ligase involved in the process of peptidoglycan biosynthesis. *Biochem J,* 415**,** 27-33.

THANKY, N. R., YOUNG, D. B. & ROBERTSON, B. D. 2007. Unusual features of the cell cycle in mycobacteria: polar-restricted growth and the snapping-model of cell division. *Tuberculosis (Edinb),* 87**,** 231-6.

TIEMERSMA, E. W., VAN DER WERF, M. J., BORGDORFF, M. W., WILLIAMS, B. G. & NAGELKERKE, N. J. 2011. Natural history of tuberculosis: duration and fatality of untreated pulmonary tuberculosis in HIV negative patients: a systematic review. *PLoS One,* 6**,** e17601.

TIMMINS, G. S. & DERETIC, V. 2006. Mechanisms of action of isoniazid. *Mol Microbiol,* 62**,** 1220-7.

TRUNKFIELD, A. E., GURCHA, S. S., BESRA, G. S. & BUGG, T. D. 2010. Inhibition of Escherichia coli glycosyltransferase MurG and Mycobacterium tuberculosis Gal transferase by uridine-linked transition state mimics. *Bioorg Med Chem,* 18**,** 2651-63.

TURNER, R. D., RATCLIFFE, E. C., WHEELER, R., GOLESTANIAN, R., HOBBS, J. K. & FOSTER, S. J. 2010. Peptidoglycan architecture can specify division planes in Staphylococcus aureus. *Nat Commun,* 1**,** 26.

TURNER, R. D., VOLLMER, W. & FOSTER, S. J. 2014. Different walls for rods and balls: the diversity of peptidoglycan. *Mol Microbiol,* 91**,** 862-74.

UDWADIA, Z. F., AMALE, R. A., AJBANI, K. K. & RODRIGUES, C. 2012. Totally drug-resistant tuberculosis in India. *Clin Infect Dis,* 54**,** 579-81.

VAN INGEN, J., AARNOUTSE, R. E., DONALD, P. R., DIACON, A. H., DAWSON, R., PLEMPER VAN BALEN, G., GILLESPIE, S. H. & BOEREE, M. J. 2011. Why Do We Use 600 mg of Rifampicin in Tuberculosis Treatment? *Clin Infect Dis,* 52**,** e194-9.

VANDER BEKEN, S., AL DULAYYMI, J. R., NAESSENS, T., KOZA, G., MAZA-IGLESIAS, M., ROWLES, R., THEUNISSEN, C., DE MEDTS, J., LANCKACKER, E., BAIRD, M. S. & GROOTEN, J. 2011. Molecular structure of the Mycobacterium tuberculosis virulence factor, mycolic acid, determines the elicited inflammatory pattern. *Eur J Immunol,* 41**,** 450-60.

VELAYATI, A. A., MASJEDI, M. R., FARNIA, P., TABARSI, P., GHANAVI, J., ZIAZARIFI, A. H. & HOFFNER, S. E. 2009. Emergence of new forms of totally drug-resistant tuberculosis bacilli: super extensively drug-resistant tuberculosis or totally drug-resistant strains in iran. *Chest,* 136**,** 420-5.

VIJAY, S., NAGARAJA, M., SEBASTIAN, J. & AJITKUMAR, P. 2014. Asymmetric cell division in Mycobacterium tuberculosis and its unique features. *Arch Microbiol,* 196**,** 157-68.

VILCHEZE, C., WANG, F., ARAI, M., HAZBON, M. H., COLANGELI, R., KREMER, L., WEISBROD, T. R., ALLAND, D., SACCHETTINI, J. C. & JACOBS, W. R., JR. 2006. Transfer of a point mutation in Mycobacterium tuberculosis inhA resolves the target of isoniazid. *Nat Med,* 12**,** 1027-9.

WANG, Q., ZHU, L., JONES, V., WANG, C., HUA, Y., SHI, X., FENG, X., JACKSON, M., NIU, C. & GAO, Q. 2015. CpsA, a LytR-CpsA-Psr Family Protein in Mycobacterium marinum, Is Required for Cell Wall Integrity and Virulence. *Infect Immun,* 83**,** 2844-54.

WATANABE, M., AOYAGI, Y., MITOME, H., FUJITA, T., NAOKI, H., RIDELL, M. & MINNIKIN, D. E. 2002. Location of functional groups in mycobacterial meromycolate chains; the recognition of new structural principles in mycolic acids. *Microbiology,* 148**,** 1881-902.

WEHRLI, W. 1983. Rifampin: mechanisms of action and resistance. *Rev Infect Dis,* 5 Suppl 3**,** S407-11.

WEIDENMAIER, C. & PESCHEL, A. 2008. Teichoic acids and related cell-wall glycopolymers in Gram-positive physiology and host interactions. *Nat Rev Microbiol,* 6**,** 276-87.

WESTON, A., STERN, R. J., LEE, R. E., NASSAU, P. M., MONSEY, D., MARTIN, S. L., SCHERMAN, M. S., BESRA, G. S., DUNCAN, K. & MCNEIL, M. R. 1997. Biosynthetic origin of mycobacterial cell wall galactofuranosyl residues. *Tuber Lung Dis,* 78**,** 123-31.

WHITE, E. L., ROSS, L. J., REYNOLDS, R. C., SEITZ, L. E., MOORE, G. D. & BORHANI, D. W. 2000. Slow polymerization of Mycobacterium tuberculosis FtsZ. *J Bacteriol,* 182**,** 4028-34.

WHO 2012. Global Tuberculosis Report 2012.

WHO 2015. Global Tuberculosis Report 2015.

WIETZERBIN, J., DAS, B. C., PETIT, J. F., LEDERER, E., LEYH-BOUILLE, M. & GHUYSEN, J. M. 1974. Occurrence of D-alanyl-(D)-meso-diaminopimelic acid and meso-diaminopimelyl-meso-diaminopimelic acid interpeptide linkages in the peptidoglycan of Mycobacteria. *Biochemistry,* 13**,** 3471-6.

XU, G., LIU, B., WANG, F., WEI, C., ZHANG, Y., SHENG, J., WANG, G. & LI, F. 2013. High-throughput screen of essential gene modules in Mycobacterium tuberculosis: a bibliometric approach. *BMC Infect Dis,* 13**,** 227.

YAGI, T., MAHAPATRA, S., MIKUSOVA, K., CRICK, D. C. & BRENNAN, P. J. 2003. Polymerization of mycobacterial arabinogalactan and ligation to peptidoglycan. *J Biol Chem,* 278**,** 26497-504.

YANG, J., YAN, R., ROY, A., XU, D., POISSON, J. & ZHANG, Y. 2015. The I-TASSER Suite: protein structure and function prediction. *Nat Methods,* 12**,** 7-8.

YANG, J. & ZHANG, Y. 2015. I-TASSER server: new development for protein structure and function predictions. *Nucleic Acids Res,* 43**,** W174-81.

YOKOYAMA, K., MIYASHITA, T., ARAKI, Y. & ITO, E. 1986. Structure and functions of linkage unit intermediates in the biosynthesis of ribitol teichoic acids in Staphylococcus aureus H and Bacillus subtilis W23. *Eur J Biochem,* 161**,** 479-89.

YUAN, Y., LEE, R. E., BESRA, G. S., BELISLE, J. T. & BARRY, C. E., 3RD 1995. Identification of a gene involved in the biosynthesis of cyclopropanated mycolic acids in Mycobacterium tuberculosis. *Proc Natl Acad Sci U S A,* 92**,** 6630-4.

ZHANG, W., JONES, V. C., SCHERMAN, M. S., MAHAPATRA, S., CRICK, D., BHAMIDI, S., XIN, Y., MCNEIL, M. R. & MA, Y. 2008. Expression, essentiality, and a microtiter plate assay for mycobacterial GlmU, the bifunctional glucosamine-1-phosphate acetyltransferase and N-acetylglucosamine-1-phosphate uridyltransferase. *Int J Biochem Cell Biol,* 40**,** 2560-71.

ZHANG, Y. 2005. The magic bullets and tuberculosis drug targets. *Annu Rev Pharmacol Toxicol,* 45**,** 529-64.

ZHANG, Y. 2008. I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics,* 9**,** 40.

ZHANG, Y. & SKOLNICK, J. 2004. Scoring function for automated assessment of protein structure template quality. *Proteins,* 57**,** 702-10.

ZHANG, Y., WADE, M. M., SCORPIO, A., ZHANG, H. & SUN, Z. 2003. Mode of action of pyrazinamide: disruption of Mycobacterium tuberculosis membrane transport and energetics by pyrazinoic acid. *J Antimicrob Chemother,* 52**,** 790-5.

ZUBER, B., CHAMI, M., HOUSSIN, C., DUBOCHET, J., GRIFFITHS, G. & DAFFE, M. 2008. Direct visualization of the outer membrane of mycobacteria and corynebacteria in their native state. *J Bacteriol,* 190**,** 5672-80.