The role of LytR-CpsA-Psr proteins in cell envelope biogenesis of *Mycobacterium smegmatis*

Thesis submitted in accordance with the requirements of the University of Liverpool for the degree of Doctor in Philosophy

by

Abhipsa Sahu

October 2019

Abstract

Tuberculosis infection is one of the leading causes of mortality worldwide and is caused by *Mycobacterium tuberculosis* (*Mtb*). With an upsurge of multidrug-resistant tuberculosis, it is a global threat. Therefore, development of new drugs need immediate attention, and this needs identification of potential drug targets. The cell envelope of mycobacteria is one such attractive drug target owing to its role in maintaining the structural integrity and pathogenicity of the bacterium. The LytR-CpsA-Psr (LCP) family of proteins in *Mycobacterium* spp. have been shown to catalyze the coupling of arabinogalactan and peptidoglycan and possess pyrophosphatase activity.

The four LCP protein homologues present in *Mycobacterium smegmatis* (*Msmeg*), MSMEG_0107, MSMEG_1824, MSMEG_5775 and MSMEG_6421, have not been extensively investigated with the focus on the existence and interplay of multiple LCP proteins. In this study with this non-pathogenic model organism, all four LCP homologues were shown to possess pyrophosphatase activity, with a significant higher activity displayed by MSMEG_0107 and MSMEG_5775. In order to further study the role of the LCP proteins on the physiology of the bacterium, single and double deletion strains lacking of the three non-essential *lcp* genes were created along with the respective complemented strains. All the generated mutants showed different phenotypes in the different assays, but usually not very severe. However, the double-deletion *lcp* mutant, ΔΔ(0107+5775) was the most affected mutant strain and displayed a disrupted cell envelope as evident from deprived growth rate, slower cellular aggregation, diminished biofilm formation on air-liquid interface, altered morphology, as well as an increased susceptibility to surface detergent, lysozyme and a wide range of antibiotics. Thus, the loss of both MSMEG_0107 and MSMEG_5775 exhibited profound effects on the mycobacterial cell envelope, and therefore could be further investigated as a possible combined drug target by extending these studies in *Mtb*. A novel approach in this study is the detection of exposed mycobacterial Gal*f* moieties of arabinogalactan by EB-A2 monoclonal antibody, in the double *lcp* deletion mutant ΔΔ(0107+5775). Transcription profiling of all the *lcp* genes in the wild type strain and the mutants exhibited differential expression of these genes under both standard and stress conditions. A loss of MSMEG_5775 resulted in an upregulation of the other three *lcp* genes in comparison to the wild type strain under standard conditions. Under both acid and lysozyme stress, the loss of MSMEG_5775 downregulated all other *lcp* genes while loss of MSMEG_6421 upregulated these genes. Lastly, an *in silico* approach led to the identification of putative transcriptional factors in mycobacteria and related species which could be further investigated and experimentally confirmed. This study helped to understand the role of the *lcp* homologues in *Msmeg* better. From the differential expression studies, role of regulator(s) might be a significant approach to understand this family of proteins much better.

ACKNOWLEDGEMENTS

It is my pleasure to acknowledge numerous people who were instrumental during my four-year PhD journey.

First of all, I would like to express my gratitude to **Dr. Boris Tefsen**, Senior Associate Professor, Department of Biological Sciences, Xi'an Jiaotong-Liverpool University (XJTLU) and **Dr. Mal Horsburgh**, Senior Lecturer, Institute of Integrative Biology, University of Liverpool (UoL), for providing this opportunity and being very helpful as a supervisor and co-supervisor respectively. Their skilful guidance and stoic patience are greatly appreciated.

I am equally grateful to my assessors, **Dr. Sekar Raju**, Associate Professor, Department of Biological Sciences, XJTLU and, **Dr. Alan McCarthy**, Emeritus Professor of Microbiology, Institute of Integrative Biology, UoL for their research insights and advices throughout my work.

I would like to acknowledge the valuable inputs from **Dr. David Ruiz Carrillo**, **Dr. Ferdinand Kappes**, **Dr. David Chiu** and **Dr. Tatsuhiko Kadowaki**. I gratefully thank all the faculty and technicians from the Department of Biological Sciences, XJTLU, as well as, from the Institute of Integrative Biology, UoL, who have helped or supported me in every small way. I would also like to thank **Siauw Liang Liem** for introducing the mycobacterial gene deletion protocol in our laboratory.

A special mention goes to all my friends **Bhooma**, **Sam**, **Kiran**, **Sancy**, **Eliana**, **Nikki**, **Mervin**, **Eldho**, **Shiva** and all other PhD colleagues of XJTLU for being the best people around me. A special thanks goes to my Saudi Arabian friends **Hannah**, **Hind** and **Hashim** for being the kindest and the most wonderful people during my stay in UoL. It is also worthwhile to mention **Anna** who became a dearest friend during my visit to the FEMS Microbiology conference 2019. Apart, I thank **Nimesh** bhaiya, his wife **Ishori** Didi and their little kid for making me a part of their family in Suzhou. I will always cherish the fun times I had with all these people

Table of contents

List of tables

List of figures

List of abbreviations

ADS	Albumin Dextrose Saline
AG	Arabinogalactan
AMK	Amikacin
Amp	Ampicillin
APS	Ammonium persulfate
Ara*f*	Arabinofuranose
AraT	Arabinofuranosyltransferase
AZT	Azithromycin
BAC	Bacitracin
BSA	Bovine serum albumin
BSL	Biosafety level
CAP	Capreomycin
CESTET	Conditional expression-specialized transduction essentiality test
CIP	Ciprofloxacin
CL	Cardiolipin
CR	Complement receptor
CTL	Cytotoxic T lymphocyte
DAP	Meso-diaminopimelic acid
DAT	Diacyltrehaloses
Dec-P	Decaprenyl phosphate
DPA	Decaprenylphosphoryl-D-arabinose
DPR	Decaprenol-1-phosphoribose
dTDP-Rha	Deoxythymidine diphosphate-rhamnose
EDTA	Ethylene diamine tetraacetic acid
ELISA	Enzyme linked immunosorbent assay
EMB	Ethambutol
ERY	Erythromycin
ETH	Ethionamide
FASL	FAS ligand
FPLC	Fast Protein Liquid Chromatography
Gal*f*	Galactofuranose
Gal*p*	Galactopyranose
GL	Glycolipid
GlcNAc	*N*-acetylglucosamine
GlcNAc-1-P	*N*-acetylglucosamine-1phosphate
GPL	Glycopeptidolipids
GPP	Geranyl pyrophosphate
GM-CSF	Granulocyte–macrophage colony stimulating factor

GMQE	Global Model Quality Estimation
GNP	Gold nano particle
GspA	Glycosylated surface-linked protein A
GST	Glutathione S-transferase
H_2O_2	Hydrogen peroxide
His	Histidine
HIV	Human Immunodeficiency Virus
HRP	Horse radish peroxidase
HRV3C	Human Rhinovirus 3C
Hyg	Hygromycin
IPTG	Isopropyl β-D-1-thiogalactopyranoside
IFNγ	Interferon gamma
IL	Interleukin
INH	Isoniazid
iNOS	Inducible nitric oxide synthase
KAN	Kanamycin
LEV	Levofloxacin
LAM	Lipoarabinomannan
LCP	LytR-CpsA-Psr
LM	Lipomannan
LU	Linker unit
mAB	Monoclonal antibody
mAGP	Mycolyl-arabinogalactan-peptidoglycan
man-3-PAA	Mannotriose-polyacrylamide
Man-LAM	Mannose-capped lipoarabinomannan
MAIT	Mucosal-associated invariant T cell
MDR	Multidrug-resistant
MIC	Minimum inhibitory concentration
MOI	Multiplicity of infection
MOX	Moxifloxacin
MR	Mannose receptor
MurNAc	*N*-acetylmuramicacid
MurNGlyc	*N*-glycolylmuramic acid
NADPH	Nicotinamide adenine dinucleotide
NO	Nitric oxide
NOX2	Nicotinamide adenine dinucleotide oxidase
NOV	Novobiocin
LPS	Lipopolysachharide
OFX	Ofloxacin
P-GlcNAc-Rha	Phosphoryl-*N*-acetylglucosaminosyl-rhamnosyl

PAT	Polyacyltrehaloses
PBP	Penicillin-binding proteins
PBS	Phosphate buffer saline
PCR	Polymerase chain reaction
PDIM	Phthiocerol dimycocerosates
PE	Phosphatidylethanolamine
PEN	Penicillin
PEP	Phosphoenolpyruvate
PG	Peptidoglycan
PGL	Phenolic glycolipids
PI	Phosphatidylinositol
Pi	Inorganic phosphate
PIM	Phosphatidylinositol mannosides
PMA	Phorbol-12-myristate-13-acetate
POP	Prolyl oligopeptidase
pRpp	Phospho-α-D-ribosyl-1-pyrophosphate
PVC	Polyvinyl chloride
PZA	Pyrazinamide
qPCR	Quantitative polymerase chain reaction
REMA	Resazurin microtiter plate assay
RIF	Rifampicin
ROS	Reactive oxygen species
RNS	Reactive nitrogen species
RR-TB	Rifampicin resistant tuberculosis
SDS-PAGE	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
SEM	Scanning electron microscopy
SL	Sulfolipids
STR	Streptomycin
TB	Tuberculosis
TDR	Totally drug-resistant
TDM	Trehalose dimycolates
TEMED	N,N,N',N'-Tetramethylethylenediamine
TF	Transcription factors
T_H1	T helper cell
TLC	Thin layer chromatography
TM	Transmembrane
TMB	3,3'-5,5" - tetramethylbenzidine
TMM	Trehalose monomycolates
TNF	Tumour necrosis factor
UDP	Uridine diphosphate

VAN	Vancomycin
WHO	World Health Organization
WT	Wild type
WTA	Wall teichoic acid
XDR	Extensively drug-resistant

Chapter-1

General Introduction

1.1. Tuberculosis is a burden on global health

Tuberculosis (TB) infection is one of the leading causes of mortality and morbidity worldwide resulting from a single infectious agent (WHO, 2018). This persisting infection is even more facilitated by the co-occurrence of Human Immunodeficiency Virus (HIV)/ Acute Immunodeficiency Syndrome. According to the World Health Organization (WHO) Global TB Report 2018, an estimated 1.3 million deaths were reported among HIV-negative people in 2017, due to TB infection, whereas, about 0.3 million deaths were accounted in the HIV-positive population. It has been estimated that, the global occurrence of TB infection in 2017 was about 10 million. A total of 30 countries appeared on WHO's list of high TB burden countries, accounting for 87% of the total global cases (Fig 1.1).

Apart from increasing susceptibility of the HIV-infected population, the multidrug-resistant (MDR), extensively drug-resistant (XDR) and more recently, totally drug-resistant (TDR) TB strains have led to a public health crisis (WHO, 2018). MDR-TB accounts for TB infection that is resistant to first-line drugs such as rifampicin (RIF) and isoniazid (INH). XDR-TB accounts for the TB infection, which is resistant to any fluoroquinolones and second-line antibiotics such as streptomycin and ciprofloxacin, in addition to INH and RIF resistance. TDR-TB accounts for the infection which is completely resistant to all kinds of drugs. In 2017, about 0.56 million incidents were reported to be resistant to rifampicin (RR-TB), the most potent first line drug for TB and, 82% of these cases had MDR-TB (WHO, 2018). According to this report, India (24%), China (13%) and the Russian Federation (10%) have been accounted for about half of the global cases of MDR/RR-TB (Fig 1.2).
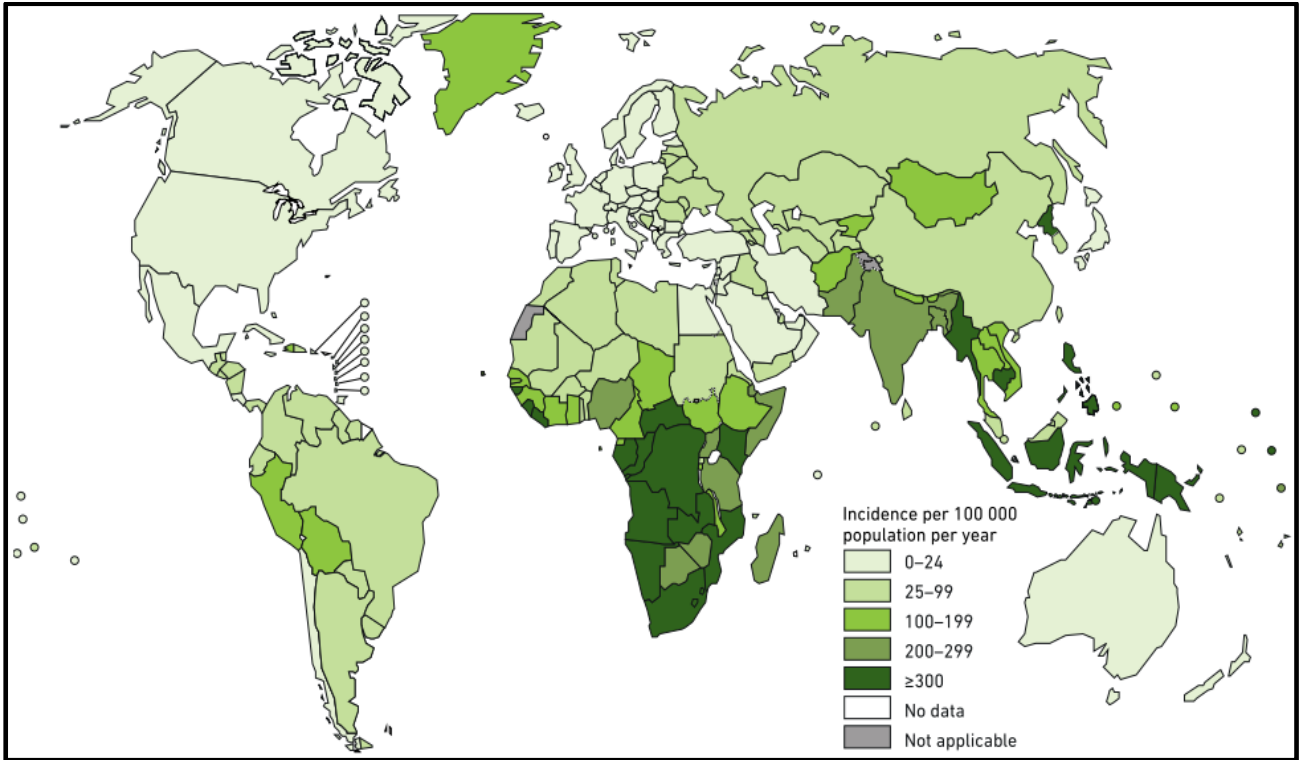
Fig 1.1. Global map representing the TB incidence rates in 2017 (WHO, 2018). The map has been adapted from the Global Tuberculosis Report, 2018. The 30 high TB burden countries have been represented with different color codes according to the incidence rates per 100,000 population per year. The number of incidence rates have also been indicated beside the color codes.
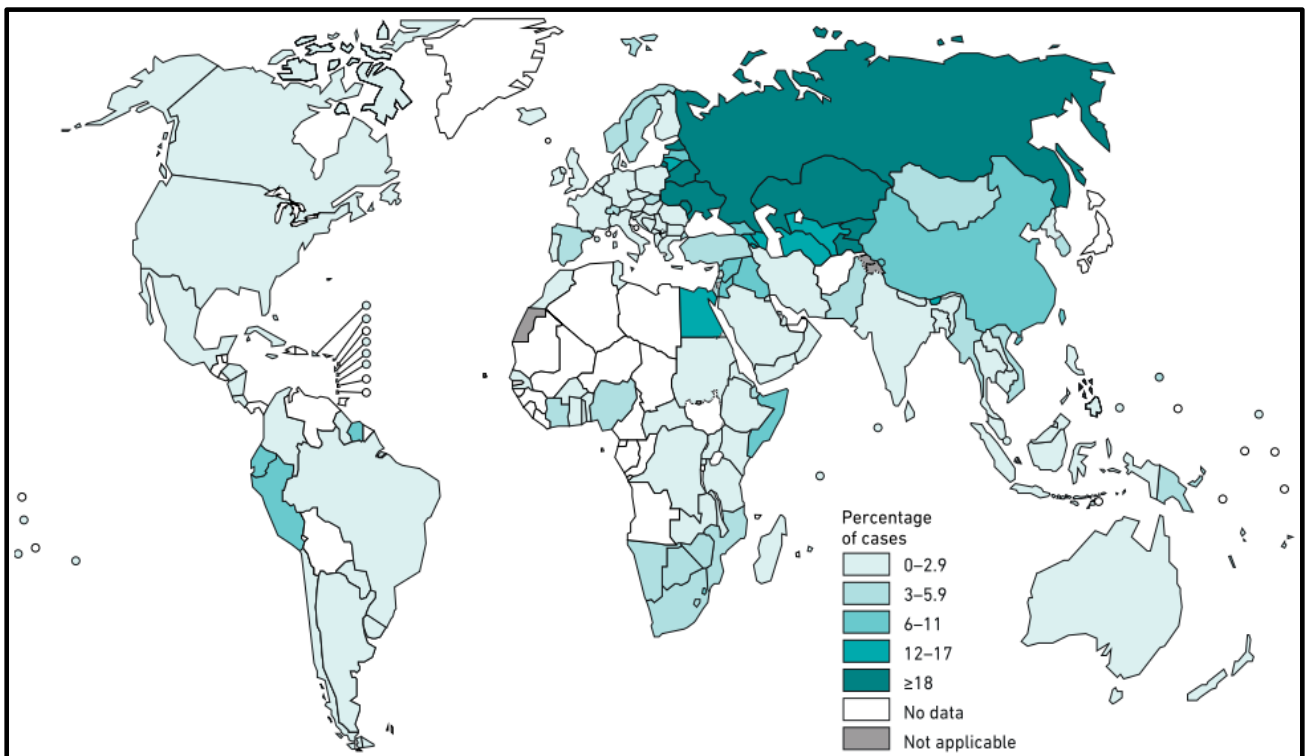
Fig 1.2. Global map representing percentage of new TB cases with MDR/RR-TB (WHO, 2018). The map has been adapted from the Global Tuberculosis Report, 2018. The figures are based on the most recent year for which data have been reported that varies among countries. Data covers the period 2002-2018.

1.2. Taxonomy of mycobacteria

The genus *Mycobacterium* consists of various members of *Mycobacterium tuberculosis* complex (MTBC) such as *Mycobacterium tuberculosis* (*Mtb*), *Mycobacterium bovis*, *Mycobacterium bovis* BCG, *Mycobacterium leprae*, *Mycobacterium canettii*, *Mycobacterium africanum* and *Mycobacterium pinnipedii* (Brosch *et al.*, 2002). Around 100 species have been found in the genus *Mycobacterium* amongst which a majority are saprophytic soil species (Brown-Elliott *et al.*, 2002) and the rest few are human pathogens such as *Mtb* that causes TB and *M. leprae* that causes leprosy. Additionally, other species such as *Mycobacterium avium complex* (MAC), *Mycobacterium fortuitum* and *Mycobacterium kansasii* cause infection and mortalities in immunocompromised individuals (Kiehn *et al.*, 1985). Recent studies have also established a novel genus *Mycolicibacterium*, into which the basonym *Mycobacterium smegmatis* (*Msmeg*) has been classified (Gupta *et al.*, 2018, Yamada *et al.*, 2018). Species within the *Mycobacterium* genus exhibit different growth rates and are categorized into slow or rapid growers (Lewin & Sharbati-Tehrani, 2005). *Mtb* and *M. leprae* exhibit a very slow growth rate and fastidious culturing processes, a distinct feature of this highly pathogenic species. The *Mtb* bacilli divides every 15 to 20 hours thus requires between 4-6 weeks to obtain visual colonies. This makes it an extremely slow grower compared to other bacteria, for instance *Escherichia coli* (*E. coli*) which has a doubling time of 20 minutes (Lewin & Sharbati-Tehrani, 2005) and *Msmeg* with a doubling time of approximately 3 hours, producing visible colonies in 3-5 days. The *Mycobacteriaceae* family that includes the genus *Mycobacterium*, is placed in the suborder *Corynebacterineae*, which is in turn included in Actinomycete taxon. *Corynebacterineae* is comprised of *Corynebacterium*, *Rhodococcus* and *Nocardia* all of which are Gram-positive, non-motile, aerobic and rod-shaped bacteria, and has characteristically high proportions of guanine and cytosine in their genomes. The most significant feature of *Corynebacterineae* is the existence a unique β-hydroxy-α-alkyl branched long chain fatty acids called mycolic acids (Minnikin & Goodfellow, 1980), which are the specific constituents of the cell envelope (Daffe & Draper, 1998, Dover *et al.*, 2004).

1.3. Virulence of *Mtb*

Mtb, the causative agent of TB, is considered as the world's most successful pathogen (Hingley-Wilson *et al.*, 2003), because it has infected a huge population globally, and has evolved with an ability to evade the host immune response and proliferate within the host. TB infection is primarily a pulmonary disease where the lungs are affected. It is instigated by the accumulation of *Mtb* present in aerosol droplets onto alveolar surfaces of the lung. This manifestation also affects the central nervous system and bones causing skeletal deformities.

The severity of the disease depends mainly on the efficacy of the host immune response. The virulence of *Mtb* infection is described in terms of "mortality" and "morbidity" (Smith, 2003). Mortality refers to the number of deaths in an infected population per year, whereas morbidity refers to the incidence rates of the disease in a population per year. Apart from these, the bacterial load in an infected host after initial infection is considered to be an important parameter to determine virulence, as it determines the fitness of the bacteria to survive host immune responses during infection. In order to measure virulence, it is thus important to understand the pathogenesis of TB infection.

1.4. Factors contributing to pathogenesis of *Mtb*

The immune system is a defense mechanism within the host, comprising of numerous biological processes that provides protection against invading infectious agents such as bacteria, viruses and parasites. The tubercle bacilli enters the body via the respiratory tract through inhalation of droplet nuclei, which are quite small (1-2 µm or less), thus allowing passage to the lower respiratory tract (Riley *et al.*, 1959). Large size droplets are barred from entering the lower respiratory tract quite efficiently by the physical barriers of nasopharynx and upper respiratory tract.

The *Mtb* cells infect the lungs during the initial stages of infection, which involves the innate immune responses of the infected host (Cooper *et al.*, 2011) (Fig 1.3). *Mtb* then propagates in lymph nodes which involves the adaptive immune response (Chackerian *et al.*, 2002, Reiley *et al.*, 2008, Wolf *et al.*, 2008). This is followed by the presentation of the bacterial antigens by dendritic cells in the lymph nodes, leading to propagation of the antigen-specific T cells. These T cells subsequently differentiate into effector T cells that migrate to the infected lungs, and stimulate the formation of granulomas along with other leukocytes. Granulomas represent organized structures consisting of macrophages, lymphocytes and fibroblasts (Flynn *et al.*, 2011), and the activation of macrophages occur inside the granulomas.

The receptor-mediated phagocytosis during the initial infection involves several host cell surface receptors such as, complement receptors, surfactant protein receptors, mannose receptors and scavenger receptors (Ernst *et al.*, 1998). The bacilli remains active inside the phagosome following phagocytosis, after which it matures to form phagolysosome (Fig 1.4). The primary host defense system plays an important role during this process by attacking the bacilli with hydrolytic enzymes of the lysosome. In most cases, this leads to the destruction of the organism, however some bacilli may evade this defense mechanism by inhibiting the maturation process of the phagosome, thus altering the internal environment. This allows a

logarithmic bacterial growth, leading to immune cell lysis. The T-lymphocytes then inhibits the growth of bacteria under the influence of interleukin-8 and cause necrosis. The necrotic material which is highly acidic inhibits most of the bacterial growth, leaving a few that can still survive in the macrophage debris, which stimulates the cell-mediated immune response. The cytokines then activate the macrophages, which destroy the internal bacilli. In most cases, a granulomatous lesion called 'tubercule' is formed where the disease enters a latent period.
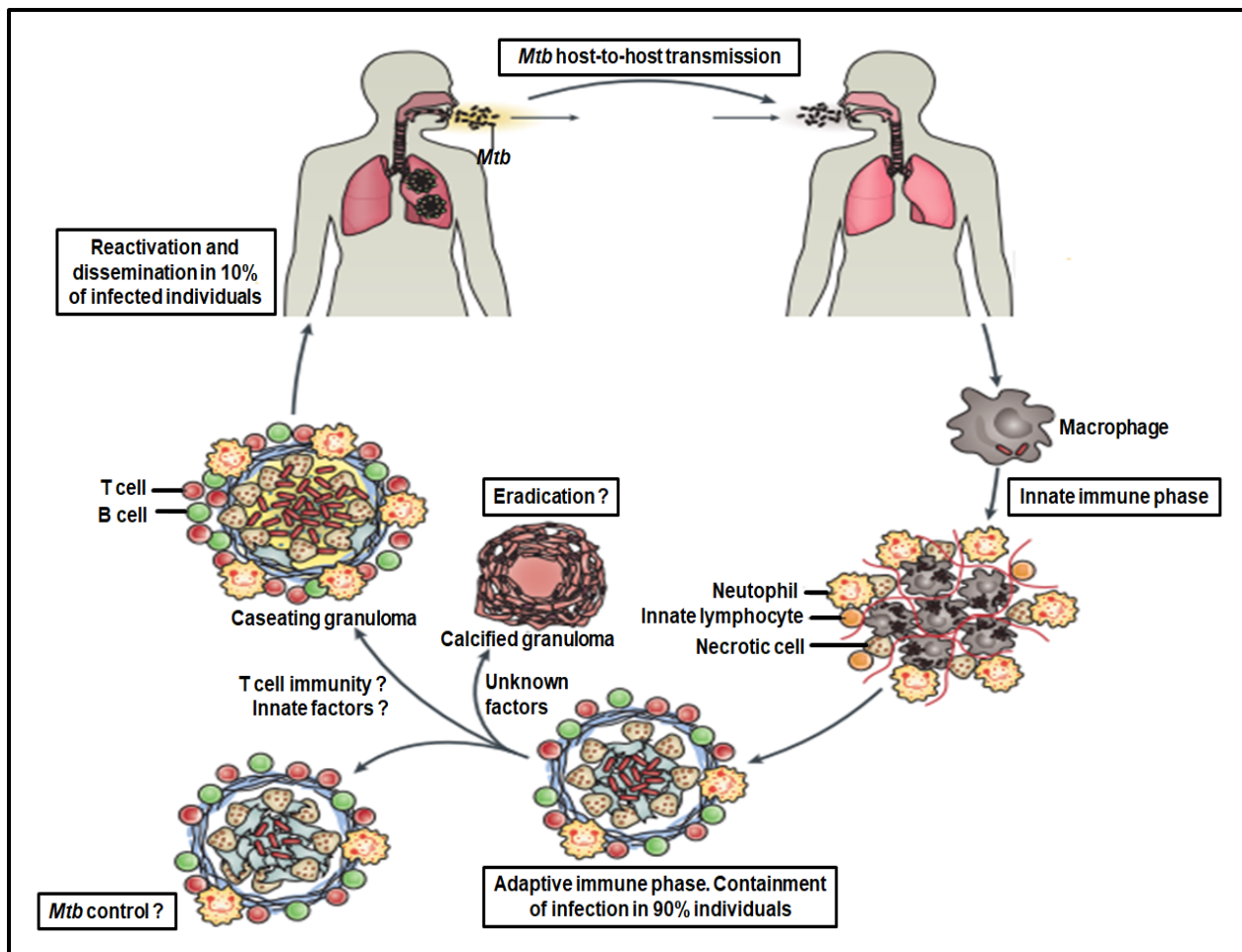
Fig 1.3. TB pathogenesis (Nunes-Alves *et al.*, 2014). The beginning of TB infection involves inhalation of aerosol droplets containing the *Mtb* bacilli. The primary infection stages comprises of the innate immune responses of the host where the inflammatory cells move to the lungs. Following *Mtb* bacilli propagation in the lymph node, the bacterial antigens are presented by the dendritic cells thus triggering the formation of T cells in the lung that are specific for the antigen. The T cells, B cells, activated macrophages and other leukocytes leads to granuloma formation that contains the *Mtb* bacilli. The infected individuals mostly acquire an asymptomatic latent infection. However, a minor fraction of these people progress towards an active disease, as the *Mtb* bacilli is released from the granulomas. Thus, when an active carrier of the disease coughs, infectious aerosol droplets are generated that transmits the infection.
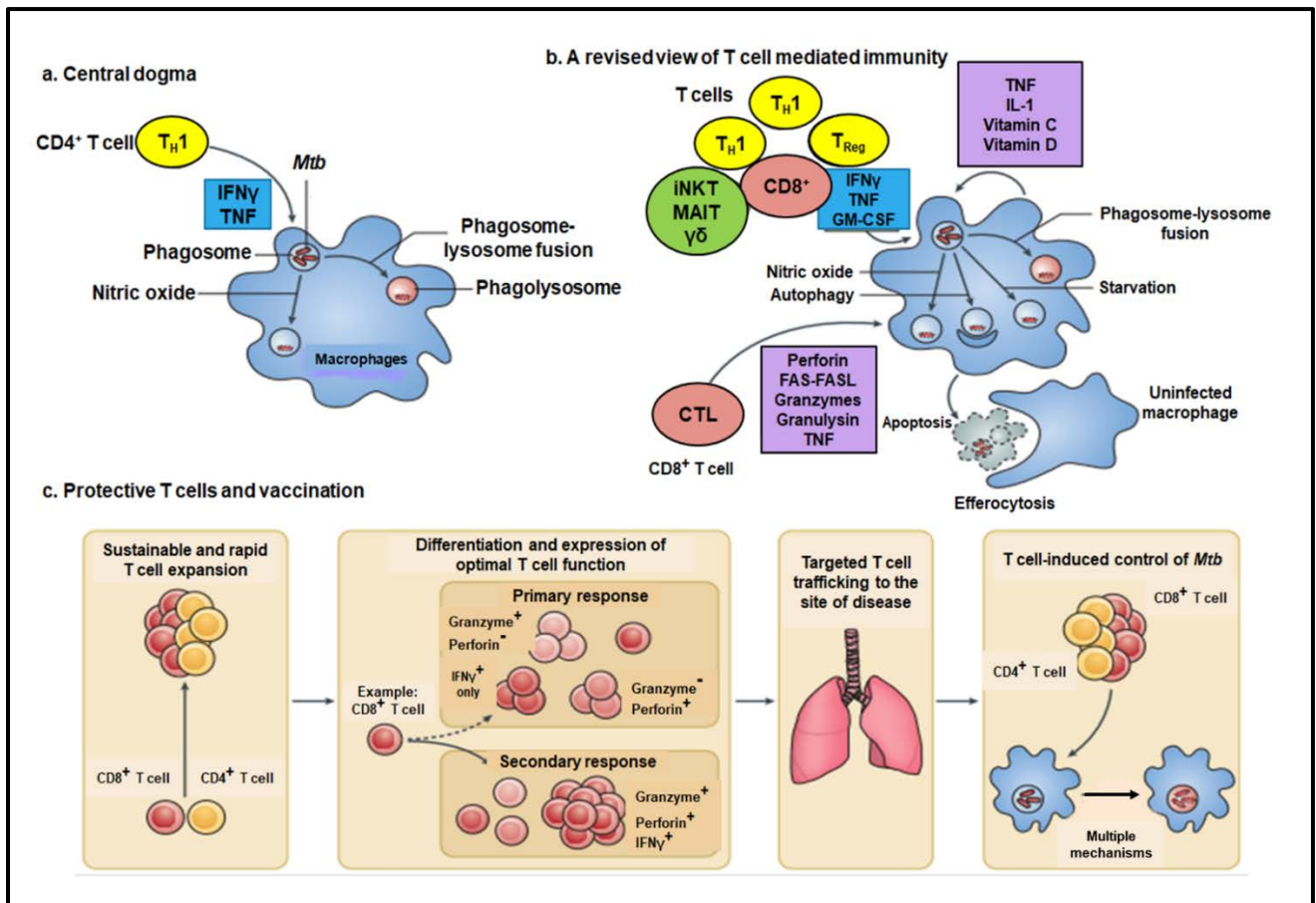
Fig 1.4. Paradigms of protective immunity to TB (Nunes-Alves et al., 2014). a) The TB defense mechanism includes the CD4[+]T cells to generate interferon-γ (IFNγ) (T helper 1 (T$_H$1) cells), which works in coordination with tumour necrosis factor (TNF; produced by the T cell or the macrophage). Both IFNγ and TNF function together to activate the macrophage antimicrobial activity to restrict *Mtb* growth. IFNγ activates two lytic pathways: the nitric oxide production and phagosome–lysosome fusion. These pathways acidifies the bacterial phagosome and destroys the *Mtb* bacilli. **b)** T cell-mediated immunity involves the defense mechanism that recruits additional T cells such as CD4[+]T cells, CD8[+]T cells, γδ T cells, mucosal-associated invariant T (MAIT) cells and CD1-restricted T cells which aids the T cells to kill the bacilli. Additionally, apoptotic cell death is contributed by cytokines such as granulocyte–macrophage colony-stimulating factor (GM-CSF), cytotoxic granules that deliver granzymes and granulysins, FAS ligand (FASL)–FAS mediated cytotoxic T lymphocyte (CTL) activity and TNF. Finally, the components of innate response such as interleukin-1 (IL-1) and vitamins, coordinate with the T cell generated cytokines. **c)** 'Protective T cells and vaccination' characterizes the defensive T cell responses. Vaccine-elicited memory T cells propagate faster to generate secondary effector T cells in order to undergo continuous proliferation following activation. Primary effector T cells are mainly expressed heterogeneously, however vaccination may result in more homogenous expression of effector functions during the memory response, making them more defensive. The primary effector and memory T cells efficiently move to the infection sites balanced with respect to T cell subsets thus limiting the potential for T cell exhaustion.

The effectiveness of this immune response is determined by several intrinsic factors such as the genetics of the immune system as well as extrinsic factors, e.g., damage to the immune system, improper nutritional balance and physiological state of the host. It is believed that immunosuppressed individuals like HIV-positive people infected with *Mtb*, have a 50% chance of developing reactivated TB infection at some time in their lives (Smith, 2003). In addition, higher incidence rates are found in places with lower nutrition status and healthcare. Thus, in order to combat this widespread infection, it is important for biologists to discover new drug targets and formulate new preventives or therapeutics, and to achieve this, an understanding of the mycobacterial structure and underlying mechanisms involved in host pathogenesis is inevitable.

1.5.   Anti-TB drugs and treatment

The most commonly used anti-TB drugs are the front-line drugs such as isoniazid (INH), ethambutol (EMB), rifampicin (RIF), streptomycin (STR) and pyrazinamide (PZA) (Table 1.1). However, the frontline anti-TB drugs may fail to cure TB for several reasons. For instance, relapse and spread of the disease contributes to the emergence of drug resistant TB bacilli. The emergence of MDR-TB is of great concern, because it is resistant to at least INH and RIF, and requires the use of second-line drugs that are difficult to obtain and are much more toxic as well as expensive than the frontline drugs (Espinal *et al.*, 2001). Therefore, an important strategy to prevent the emergence of MDR-TB is to detect and treat the drug sensitive or single drug resistant TB (Masjedi *et al.*, 2006).

The second-line anti-TB drugs are sub-divided into two categories (Table 1.2): i) Fluoroquinolones such as Ofloxacin (OFX), levofloxacin (LEV), moxifloxacin (MOX) and ciprofloxacin (CIP). ii) Aminoglycosides (or, injectable anti-TB drugs) such as kanamycin (KAN), amikacin (AMK) and capreomycin (CAP).

Table 1.1: Commonly used anti-TB frontline drugs and their targets. Adapted from (Kremer & Besra, 2002, Nachega & Chaisson, 2003, Abrahams & Besra, 2016)
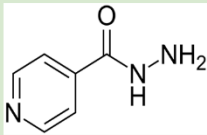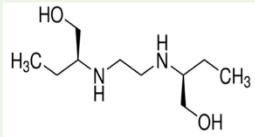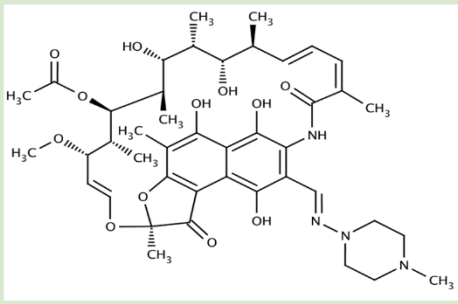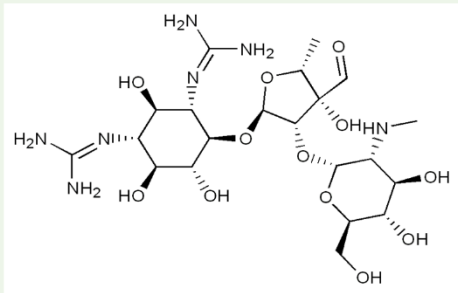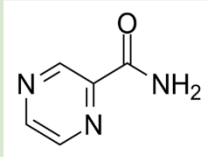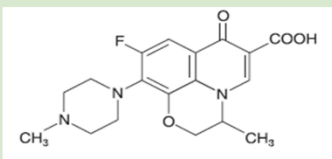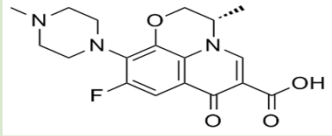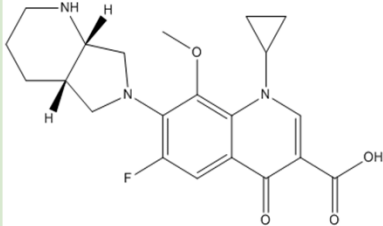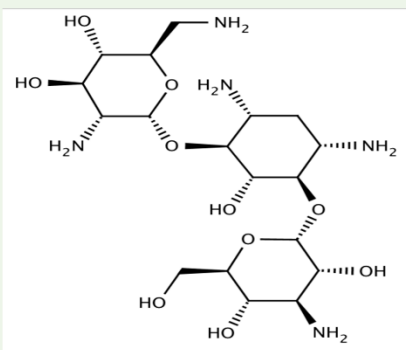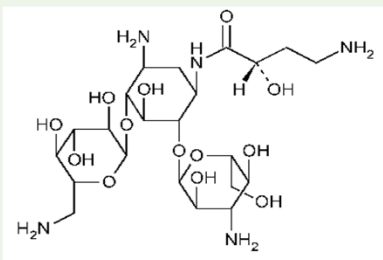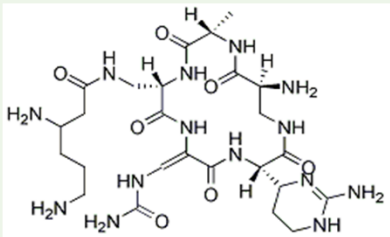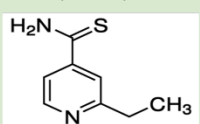
Drug	Mechanism of action	Genes associated with resistance	Gene function
Isoniazid (INH) (1952)	Primarily inhibits mycolic acid biosynthesis Also exhibits multiple effects on DNA, lipids, carbohydrates and NAD metabolism	*kat*G *inh*A *ndh* *oxy*R-*aph*C	Catalase-peroxidase Enoyl-Acp reductase NADH dehydrogenase II Alkyl hydroperoxidase
Ethambutol (EMB) (1961)	Inhibits arabinogalactan (AG) biosynthesis	*emb*CAB	Arabinosyltransferases
Rifampicin (RIF) (1966)	Inhibits transcription	*rpo*B	RNA polymerase
Streptomycin (STR) (1944)	Inhibits protein synthesis	*rps*L *rrs*	S12 ribosomal protein 16S rRNA
Pyrazinamide (PZA) (1952)	Acidifies the cytoplasm and de-energizes the membrane	*pnc*A	Nicotinamidase/pyrazinamide

Table 1.2: Second-line anti-TB drugs and their targets. Adapted from (Nath & Ryoo, 2013)

Drug	Mechanism of action	Genes associated with resistance	Gene function
Fluoroquinolones	Inhibits DNA gyrase, a type II topoisomerase	*gyr*A/*gyr*B	DNA gyrase
Ofloxacin (OFX)			
Levofloxacin (LEV)			
Moxifloxacin (MOX)			
Aminoglycosides	Inhibits protein synthesis		
Kanamycin (KAN)		*rrs*	16S rRNA
		cis	aminoglycoside acetyltransferase
Amikacin (AMK)		*rrs*	16S rRNA
Capreomycin (CAP)		*rrs*	16S rRNA
		tylA	rRNA methyltransferase
Ethionamide (ETH) (1956)	Inhibits mycolic acid biosynthesis	*eth*A *inh*A	Flavin mono-oxygenase Enoyl-Acp reductase

1.6. Mycobacterial cell envelope structure

The cell envelope is important for maintaining the structural integrity and the distinctive shape of a bacterial cell. It is found in both Gram-positive and Gram-negative bacteria, but differs largely owing to diverse biochemical and structural components. The mycobacterial cell envelope is a complex multi-layered structure and consists of three main components. The innermost layer is the plasma membrane composed of major phospholipids such as phosphatidylethanolamine (PE), cardiolipin (CL) and phosphatidylinositol (PI) (Puffal *et al.*, 2018). Except phosphatidylinositol mannosides (PIMs), which are specific to Actinomycetales, the lipid composition of the plasma membrane is similar to that of other prokaryotes, indicating that the general properties of lipid bilayers are applicable to mycobacteria. The outermost component is a loosely attached capsular-like structure outside the outer membrane of *Mycobacterium* that mainly consists of polysaccharides, proteins and solvent-extractable lipids (Lemassu & Daffe, 1994, Ortalo-Magne *et al.*, 1995, Sani *et al.*, 2010). The main capsular polysaccharide found in *Mtb* is α-glucan, a polymer of linear α-D-1–4-linked glucose units with core substitutions at position 6 on every 5 or 6 residues by α-D-1–4-linked oligoglucosides (Lemassu & Daffe, 1994, Dinadayala *et al.*, 2008, Sambou *et al.*, 2008). The surface-exposed lipids in the outer membrane also known as mycomembrane, is composed of trehalose monomycolates (TMM) and trehalose dimycolates (TDM), phthiocerol dimycocerosates (PDIM), sulfolipids (SL), diacyltrehaloses (DAT), polyacyltrehaloses (PAT) and phenolic glycolipids (PGL) (Jackson, 2014). In between the inner and outer membrane layers, a cell wall core is present that is comprised of peptidoglycan (PG) covalently attached to the heteropolysaccharide arabinogalactan (AG) via phosphoryl-*N*-acetylglucosaminosyl-rhamnosyl linkage units (P-GlcNAc-Rha) (Fig 1.5). AG in turn is esterified at its non-reducing ends to α-alkyl, β-hydroxy long-chain (C_{70}-C_{90}) mycolic acids to form the bulk of the inner leaflet of the outer membrane. Intercalated within this mycomembrane, are the non-covalently linked lipids and lipoglycans such as PDIMs and PGLs which have varied roles in signaling events, pathogenesis, and immune response (Jackson, 2014). Several enzymes in biosynthetic pathways of the key cell-envelope glycolipids, such as PIMs, lipomannan (LM), and lipoarabinomannan (LAM) have been shown or predicted to be essential in *Mtb* (Goude *et al.*, 2008, Griffin *et al.*, 2011). Apart from this, alterations in the LM/LAM structures make both *Mtb* and nonpathogenic *Msmeg* susceptible to β-lactam antibiotics, which otherwise cannot cross the mycobacterial cell wall efficiently (Jarlier & Nikaido, 1990). These observations suggest that PIMs/LM/ LAM play important structural roles within the cell envelope to maintain the permeability barrier. Furthermore, PIMs/LM/ LAM are critical for virulence, with recognized host factor interactions (Vergne *et al.*, 2014, Hmama *et al.*, 2015, Ishikawa *et al.*, 2017). Biosynthetic pathways of many cell envelope lipids and glycans start in the cytoplasm or the cytoplasmic side of the plasma membrane, followed by reactions taking place in the periplasmic space (Puffal *et al.*, 2018). The complete structure is referred to as

the mycolyl-arabinogalactan-peptidoglycan complex (mAGP). The cell envelope of *Mtb* is associated with pathogenesis via its modulation of both bacillary and host cellular processes including permeability, phagosome maturation, neutralization of free radicals, and alteration of host immune responses (Smith, 2003). This is accompanied by its associated components like glycopeptidolipids (GPL) and free mycolic acids which are involved in the formation of mycobacterial biofilms, and have been implicated in tolerance to antibiotics *in vitro* (Recht & Kolter, 2001, Ojha *et al.*, 2008). Thus, in order to understand these underlying mechanisms, it is important to understand the structure of the two major heteropolysaccharides AG and PG present in the mycobacterial cell envelope, and the interconnected biosynthetic pathways of various components present herein.
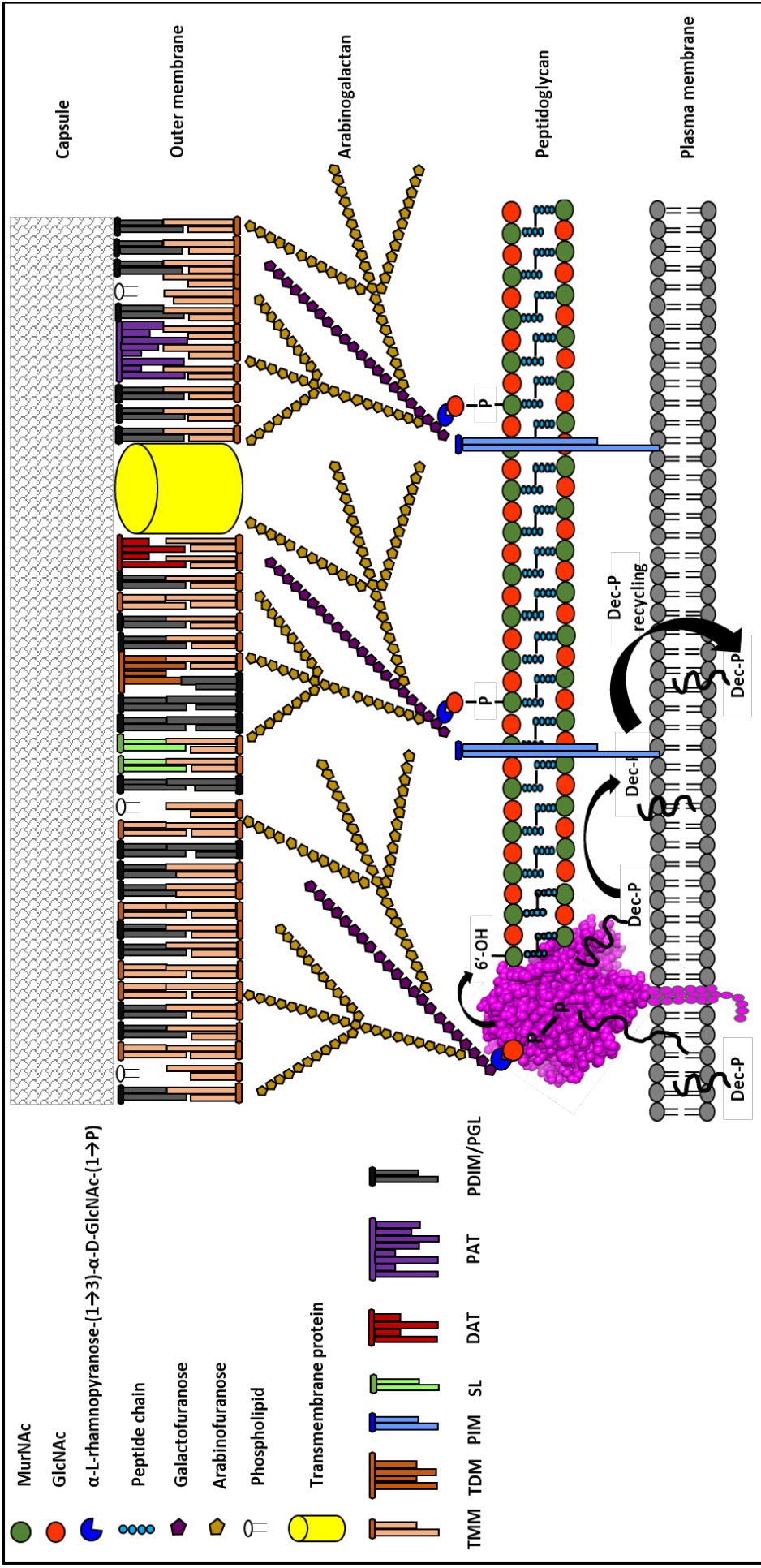
Fig 1.5. Schematic representation of the mycobacterial cell envelope with an emphasis on the role of LCP proteins in cell wall biogenesis. LCP proteins associates AG to the 6'-OH of MurNAc present in PG, thus resulting in the release of decaprenyl-1-monophosphate, which is recycled and used during the transfer of galactan polymers across plasma membrane. Upon linkage of the AG to PG, LCP proteins release the bound AG resulting in the formation of mature mAGP complex.

MurNAc- *N*-acetyl-muramic acid; GlcNAc- *N*-acetylglucosamine; TMM- Trehalose monomycolates; TDM- Trehalose dimycolates; PIM- Phosphatidylinositol mannosides; SL- Sulfolipids; DAT- Diacyltrehaloses; PAT- Polyacyltrehaloses; PDIM- Phthiocerol dimycocerosates; PGL- Phenolic glycolipids

1.7. Biogenesis of Peptidoglycan (PG) and Arabinogalactan (AG) in mycobacteria

The individual structures of PG and AG have been well studied before with a focus on the molecular genetics of their sub components and their individual role in cell wall assembly (Jankute *et al.*, 2014).

1.7.1. Peptidoglycan biosynthesis

The synthesis of PG in mycobacteria is similar to that in other bacteria (Pavelka *et al.*, 2014). The mycobacterial PG is comprised of alternating *N*-acetylglucosamine (GlcNAc) and modified muramic acid (Mur) residues, that are linked to each other in a β (1→4) configuration (Lederer *et al.*, 1975). Its mesh-like structural organization contributes to the rigidity of the cell, thus providing it endurance to withstand osmotic pressure maintaining cell integrity and cellular shape (Vollmer *et al.*, 2008). Unlike *E. coli* PG, in *Mtb* and *Msmeg*, the muramic acid residues contain a mixture of *N*-acetyl and N-glycolyl derivatives, whereby the *N*-acetyl function has been oxidised to a N-glycolyl function to form MurNGly (Mahapatra *et al.*, 2005a, Mahapatra *et al.*, 2005b, Raymond *et al.*, 2005). These additional glycolyl residues provide extra hydrogen bonding that strengthens the mesh-like structure (Brennan & Nikaido, 1995) and also protects the bacilli from lysozyme degradation. Other unique features of the mycobacterial PG are the amidation of the carboxylic acids in the peptide stems (Mahapatra *et al.*, 2005b) and additional glycine or serine residues (Vollmer *et al.*, 2008). The tetrapeptide side chains in PG is composed of **L**-alaninyl-**D**-isoglutaminyl-meso-diaminopimelyl-**D**-alanine (Petit *et al.*, 1969) that cross-links with identical short peptides of adjacent glycan chains (van Heijenoort, 2007). These cross-links consist of meso-diaminopimelic acid (DAP) and **D**-alanine bond, which are commonly found in most prokaryotes. Unlike *E. coli*, the PG in *Mtb* consists of muramic acid residues that serve as attachment sites for the galactan domain of the AG, while carbon-6 of some of the muramic acid residues form a phosphodiester bond and are linked to the α-**L**-rhamnopyranose–(1→3)-α-**D**-GlcNAc(1→P) linker unit (LU) of AG (McNeil *et al.*, 1990, Hancock *et al.*, 2002).

The mycobacterial PG biosynthesis has been summarized in Fig 1.6. The **first step** is the synthesis of uridine diphosphate-*N*-acetylglucosamine (UDP-GlcNAc). This is catalysed by GlmU that exhibits acetyltransferase and uridyltransferase activities (Zhang *et al.*, 2009). Here, the acetyl group from acetyl-CoA is transferred to glucosamine-1-phosphate (GlcN-1-P) to produce *N*-acetylglucosamine-1phosphate (GlcNAc-1-P). This is followed by the transfer of uridine-5′-monophosphate from UTP to GlcNAc-1-P to yield UDP-GlcNAc (Zhang *et al.*, 2009). The absence of GlcN-1-P from humans makes the acetyltransferase domain a potential drug

target (Mio *et al.*, 1998). The **second step** comprises of a sequential synthesis pathway that involves the formation of UDP-*N*-acetylmuramic acid (UDP-MurNAc)-pentapeptide, catalysed by the Mur ligases, MurA to MurF (Barreteau *et al.*, 2008). MurA, a UDP-*N*-acetylglucosamine 1-carboxyvinyltransferase, and MurB, a UDP-*N*-acetylenolpyruvoylglucosamine reductase, catalyze the synthesis of UDP-MurNAc from UDP-GlcNAc, by first transferring the enoylpyruvyl moiety of phosphoenolpyruvate (PEP) to the 3'-OH of UDP-GlcNAc, followed by reducing to a lactoyl ether moiety via NADPH. During this period, the hydroxylation of the UDP-MurNAc to UDP-*N*-glycolylmuramic acid (UDP-MurNGlyc) occurs which is catalysed by NamH, a UDP-*N*-acetylmuramic acid hydroxylase (Mahapatra *et al.*, 2005a). This unique structural variation in mycobacteria (and closely related species) is implicated in enhancing the intrinsic strength of PG, by potentially reducing the susceptibility to lysozyme and establishing sites for additional hydrogen bonding (Raymond *et al.*, 2005). In the **third step**, the pentapeptide chain is integrated onto the UDP-MurNAc/Glyc substrates by the sequential addition of **L**-alanine, **D**-isoglutamate, DAP and **D**-alanyl-**D**-alanine (generated by the ligase Ddl), by the ATP-dependent Mur ligases C-F respectively (Munshi *et al.*, 2013). This results in the formation of a muramyl-pentapeptide product, UDP-MurNAc/Glyc-**L**-Ala-**D**-isoGlu-*m*-DAP-**D**-Ala-**D**-Ala, also known as the Park's nucleotide (Kurosu *et al.*, 2007). In the **fourth step**, MurX (also known as MraY) catalyzes the translocation of Park's nucleotide to decaprenyl phosphate (C_{50}-P), resulting in the formation of Lipid I (Kurosu *et al.*, 2007). The **fifth and the final step** of PG synthesis is catalysed by the glycosyltransferase, MurG which leads to the generation of Lipid II, the monomeric building block of PG, via a β(1→4) linkage between GlcNAc (from UDP-GlcNAc) and MurNAc/Glyc of Lipid I (Mengin-Lecreulx *et al.*, 1991). The enzyme that catalyzes the translocation of Lipid II across the plasma membrane has still remained elusive. Currently, two different enzymes with flippase activity, MurJ and FtsW are considered for this translocation step (Mohammadi *et al.*, 2011, Sham *et al.*, 2014, Ruiz, 2015). After the translocation of Lipid II across the plasma membrane, the monofunctional and bifunctional penicillin-binding proteins (PBPs) polymerize the Lipid II (Sauvage *et al.*, 2008). The bifunctional PBPs, PonA1/PBP1 and PonA2/PBP2 possess transglycosylase and transpeptidase activities. PonA1 forms a linkage between the disaccharide building blocks of Lipid II and the pre-existing glycan chains resulting in the release of decaprenyl pyrophosphate. With the cleavage of the terminal **D**-Ala, PonA2 catalyzes the formation of (3→4) crosslinks, between *m*-DAP and **D**-Ala of the neighbouring pentapeptide chains. The monofunctional PBPs are involved in D,D-transpeptidation and D,D-carboxypeptidation, both leading to cleavage of the terminal **D**-Ala located in the peptide stem (Goffin & Ghuysen, 2002). *Mtb* PG primarily consists of (3→3) cross-links between two tetrapeptide stems (Lavollay *et al.*, 2008), whereas only 20% are (3→4) cross-links (Kumar *et al.*, 2012).
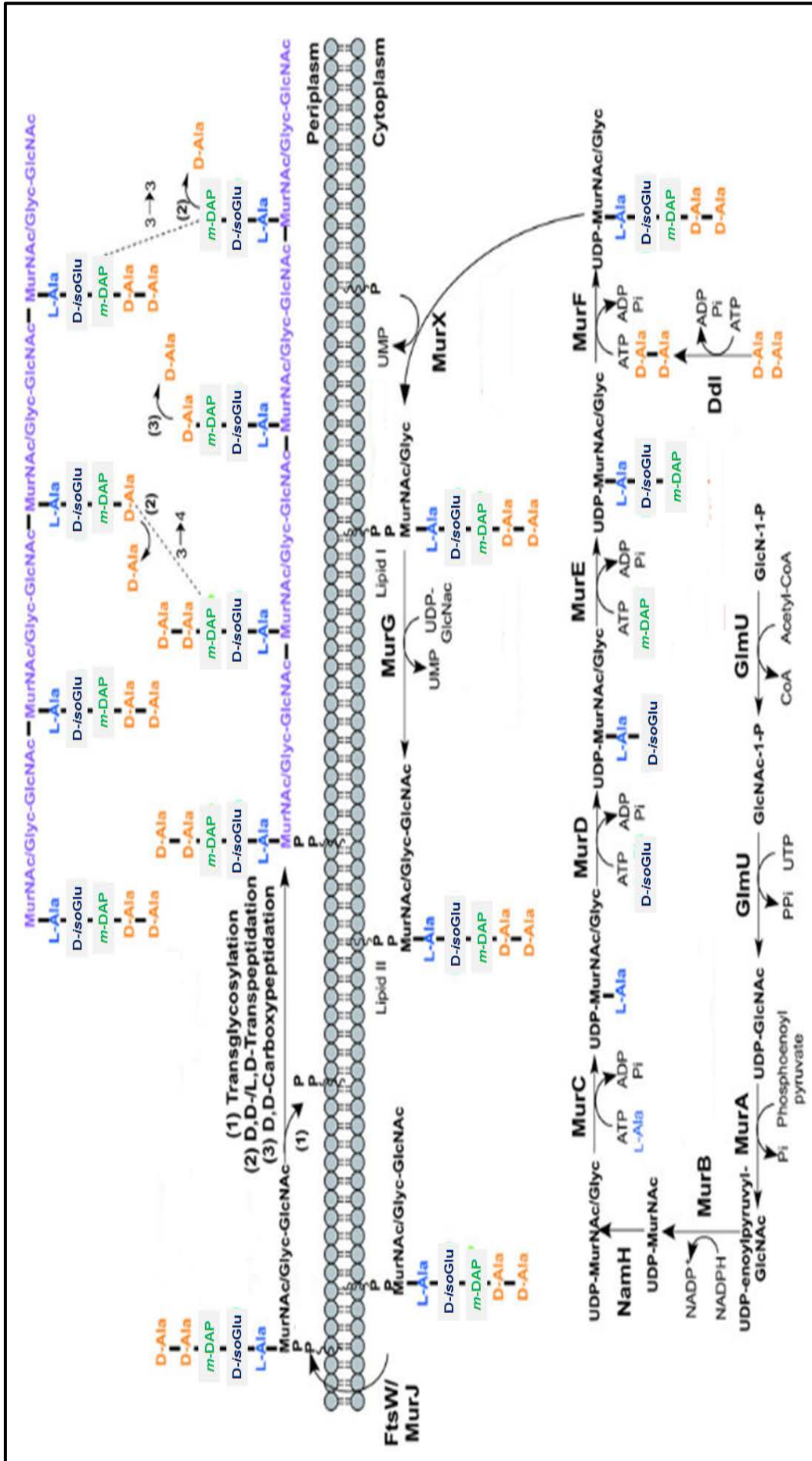
Fig 1.6. Mycobacterial peptidoglycan biosynthesis. The initiation of PG synthesis occurs in the cytoplasm with the synthesis of UDP-GlcNAc which is catalysed by the acetyltransferase and uridyltransferase activities of GlmU (Zhang *et al.,* 2009). This is followed by the synthesis of UDP-MurNAc-pentapeptide which is catalysed by a set of Mur ligases, MurA to MurF (Barreteau *et al.,* 2008). MurA aids in the transfer of enolpyruvate from phosphoenolpyruvate to the 3-OH of UDP linked GlcNAc, which is subsequently reduced to form UDP-MurNAc by MurB. At this point NamH, a UDP-N-acetylmuramic acid hydroxylase, hydroxylates UDP-MurNAc to UDP-N-glycolylmuramic acid (UDP-MurNGlyc). This is followed by the serial addition of L-alanine, D-glutamic acid, DAP and D-alanyl-D-alanine dipeptide to MurNAc, by MurC, MurD, MurE and MurF respectively, resulting in the formation of UDP-MurNAc-pentapeptide (Munshi *et al.,* 2013). Lipid I synthesis follows subsequently, with transfer of the modified MurNAc residue to a polyprenyl phosphate carrier lipid catalysed by MraY (MurX) (Ikeda *et al.,* 1991). This is followed with the addition of GlcNAc to MurNAc/Glyc of Lipid I by MurG via a β(1→4) linkage (Mengin-Lecreulx *et al.,* 1991). The final product, GlcNAc-MurNAc(pentapeptide)-diphosphorylundecaprenol (Lipid II) is then translocated across the membrane by two different enzymes with 'flippase' activity: MurJ and FtsW (Mohammadi *et al.,* 2011, Sham *et al.,* 2014, Ruiz, 2015). Further, Lipid II is incorporated into the growing PG by transglycosylases, and further modified by transpeptidases (Bhakta & Basu, 2002, Goffin & Ghuysen, 2002). **Adapted from (Abrahams & Besra, 2016).**

1.7.2. Linker Unit biosynthesis

The LU is covalently attached to the 6'-OH groups of about 10-12% muramic acid residues of PG through a phosphodiester bond (Amar & Vilkas, 1973), followed by the addition of galactofuranose (Gal*f*) and arabinofuranose (Ara*f*) residues (Mikusova *et al.*, 1996, Mikusova *et al.*, 2000, Yagi *et al.*, 2003) (Fig 1.7). The structural role of the LU in coupling AG to PG, as well as the presence of **L**-Rhamnose, a sugar absent in humans, makes the mycobacterial LU an attractive drug target (Ma *et al.*, 2002). However, in Gram-positive bacteria, the LU is **D**-ManNAc-(β1→4)-**D**-GlcNAc which is responsible for transferring AG to PG, and is slightly different with the presence of N- acetylmannosaminyl unit (Yokoyama *et al.*, 1986).
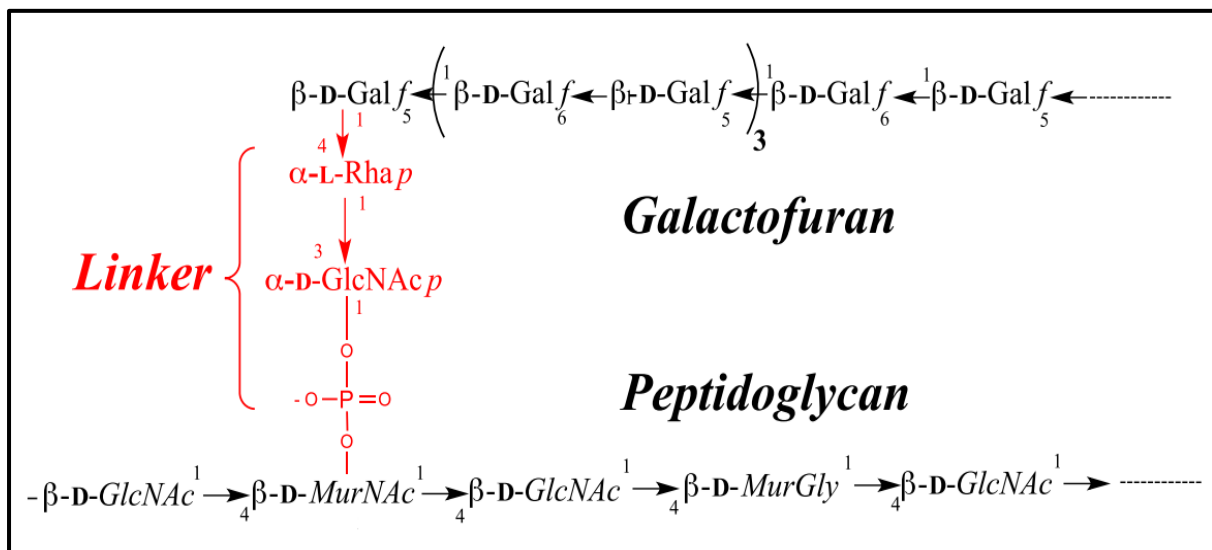
Fig 1.7. The Rha-GlcNAc linker unit in Mycobacterium (Grzegorzewicz *et al.*, 2016). The LU is the conduit between AG and PG. The reducing end of AG consists of the terminal sequence →5)-**D**-Galf-(1→4)-**L**-Rhap-(1→3)-**D**-GlcNAc, attached to muramyl-6-P

Several enzymes involved in the formation of the LU during AG synthesis have been previously reported in *Mtb* (Angala *et al.*, 2014). For instance, the gene *Rv1302* encoding a WecA-like transferase aids the transfer of GlcNAc 1-phosphate to decaprenyl phosphate resulting in the formation of Dec-P-P-GlcNAc (Jin *et al.*, 2010). Subsequently, a rhamnosyl residue from deoxythymidine diphosphate-rhamnose (dTDP-Rha) gets attached to the 3-position of GlcNAc of GL-1, forming glycolipid 2 (GL-2). This reaction is catalysed by WbbL1 (*Rv3265c*) resulting in the formation of the LU, i.e., Dec-P-P-GlcNAc-Rha (Mills *et al.*, 2004). Moreover, in *Msmeg*, WbbL1 (*MSMEG_1826*) has been highlighted essential for bacterial viability (Mills *et al.*, 2004). The product of WbbL1 utilises the nucleotide donor dTDP-rhamnose for the formation of GL-2, therefore the rhamnosyl biosynthetic pathway and several inhibitors in this pathway have been investigated by many researchers (Babaoglu *et al.*, 2003,

Kantardjieff *et al.*, 2004, Ma *et al.*, 2001). The synthesis of dTDP-Rha is a 4-stage linear pathway that utilizes the gene products of *rmlABCD*.

At first, RmlA (*Rv0334*) converts dTTP + α-**D**-glucose 1-phosphate to dTDP-glucose + PPi (Ma *et al.*, 1997). The product of RmlA activity is then transferred through three successive reactions which is catalysed by dTDP-**D**-glucose 4,6-dehydratase (Rv3464, RmlB), dTDP-4-keto-6-deoxy-**D**-glucose 3,5 epimerase (Rv3465, RmlC) and dTDP-Rha synthase (Rv3266, RmlD) (Hoang *et al.*, 1999, Ma *et al.*, 2001, Stern *et al.*, 1999). The essentiality of RmlB and RmlC genes for mycobacterial growth was shown by Li and colleagues (Li *et al.*, 2006).

1.7.3. Arabinogalactan biosynthesis

The mycobacterial AG biosynthesis has been summarized in Fig 1.8. The formation of LU is followed by AG synthesis where the synthesis of galactan is instigated in the cytoplasm on a decaprenyl phosphate (Dec-P) carrier lipid, where GlcNAc-phosphate is transferred from UDP-GlcNAc, forming C_{50}-P-P-GlcNAc, referred to as glycolipid 1 (GL-1) (Mikusova *et al.*, 1996). Two bi-functional galactosyltransferases (GlfT1 and GlfT2) encoded by *Rv3782* and *Rv3808c* respectively, are involved in adding Gal*f* residues to the LU (Alderwick *et al.*, 2008). GlfT1 first transfers Gal*f* from UDP-Gal*f* to the C-4 position of **L**-Rha, and then adds a second Gal*f* residue to the C-5 position of the primary Gal*f*, thus forming C_{50}-P-P-GlcNAc-**L**-Rha-Galf_2 (Mikusova *et al.*, 2006, Alderwick *et al.*, 2008, Belanova *et al.*, 2008). This is followed by the sequential transfer of Gal*f* residues to the growing galactan chain with alternating β(1→5) and β(1→6) glycosidic linkages, catalysed by GlfT2 (Kremer *et al.*, 2001, Rose *et al.*, 2006). The galactan chains consists of ~30 Gal*f* residues, forming C_{50}-P-P-GlcNAc-**L**-Rha-Galf_{30} (Daffe *et al.*, 1990).

The arabinosylation of AG occurs in the plasma membrane on its periplasmic side, which is catalysed by membrane-associated Dec-P arabinose-dependent glycosyltransferases (Angala *et al.*, 2014). The arabinan domain is a highly branched network built on a backbone of α(1→5) linked sugars with a number of α(1→3) linked residues forming 3,5-Ara*f* branch points (Daffe *et al.*, 1990). Further α(1→5) linked Ara*f* sugars are attached subsequently to this branch point with the non-reducing ends terminated by β(1→2) Ara*f* residues. The concluding structural motif is the characteristic hexa-arabinoside (Ara$_6$) present as (Ara*f*-β(1→2)-Ara*f*-α(1-)$_2$ → 3,5-Ara*f*-α(1→5)-Ara*f*-α(1→5), of which two-thirds are mycolated (McNeil *et al.*, 1994). A family of glycosylphosphoprenols which are associated with the plasma membrane, have been identified that operate as sugar donor substrates for several TB glycosyltransferases. For example, activated forms of **D**-ribofuranose (Wolucka *et al.*,

1994, Wolucka & de Hoffmann, 1995), **D**-mannopyranose (Takayama & Goldman, 1970) and **D**-arabinofuranose (Wolucka *et al.*, 1994, Wolucka & de Hoffmann, 1995) have been acknowledged.

Ara*f* residues are directly transferred onto C_{50}-P-P-GlcNAc-**L**-Rha-Gal*f*$_{30}$ from the lipid donor decaprenylphosphoryl-**D**-arabinose (DPA) (Wolucka *et al.*, 1994). The synthesis of DPA occurs sequentially in the cytoplasm and originates solely from phospho-α-**D**-ribosyl-1-pyrophosphate (pRpp). PrsA, a pRpp synthetase catalyses the transfer of pyrophosphate from ATP to C-1 of ribose-5-phosphate, resulting in pRpp (Alderwick *et al.*, 2011). This is followed by the addition of a decaprenyl moiety by UbiA (decaprenol-1-phosphate 5-phosphoribosyltransferase) leading to the formation of decaprenol-1-monophosphate 5-phosphoribose (Alderwick *et al.*, 2005, Huang *et al.*, 2005, Huang *et al.*, 2008). A putative phospholipid phosphatase encoded by Rv3807c, catalyses C-5 dephosphorylation, thus forming decaprenol-1-phosphoribose (DPR) (Jiang *et al.*, 2011). In the final step, a two-step oxidation/reduction activity of the decaprenylphosphoribose-2′-epimerase consisting of subunits DprE1 and DprE2, catalyzes the epimerization of ribose C-2 hydroxyl, resulting in the formation of DPA (Mikusova *et al.*, 2005).

The sole donor of Ara*f* residues involved in the biosynthesis of arabinans in the Actinomycetales involves β-**D**-arabinofuranosyl-1-monophosphodecaprenol (DPA) (Wolucka *et al.*, 1994, Xin *et al.*, 1997, Alderwick *et al.*, 2005), and the reaction is catalysed by a set of arabinofuranosyltransferases known as *embCAB* and AftA proteins (Wolucka *et al.*, 1994, Lee *et al.*, 1997), consisting of 13 transmembrane domains (Telenti *et al.*, 1997). This is a characteristic feature found only in *Corynebacterianeae* (Huang *et al.*, 2005, Mikusova *et al.*, 2005, Alderwick *et al.*, 2006a). AftA (Rv3792) is the first arabinofuranosyltransferase (AraT), to prime the galactan backbone by addition of arabinose from DPA onto the galactan chain (Alderwick *et al.*, 2006b), signifying that DPA is the only donor in the related organism *C. glutamicum* (Alderwick *et al.*, 2005). Apart from this, all the glycosyltransferases that are involved in the biosynthesis of both AG and LAM arabinans are integral membrane proteins, and hence classified as members of the glycosyltransferase family C (GT-C) (Berg *et al.*, 2007). The priming of the galactan backbone is followed by the attachment of α(1→5) linked Ara*f* residues. These reactions are catalysed by arabinofuranosyltransferase activities of EmbA and EmbB respectively (Telenti *et al.*, 1997, Radmacher *et al.*, 2005) which finally results in the formation of mature AG with the Ara$_6$ motif. EmbC, on the other hand is required for the elongation of the arabinan domain of LAM (Zhang *et al.*, 2003). The anti-TB drug ethambutol (EMB) specifically known to inhibit the *emb*CAB locus in *Mtb* (Telenti *et al.*, 1997). AftC (Rv2673/MSMEG_2785) transfers the Ara*f* residues from DPA to the arabinan domain resulting in the internal branching of AG by α(1→3)-linked Ara*f* residues. This forms a

branched arabinan domain which is distal to the non-reducing terminal Ara_6 motif specifically found in mycobacterial AG (Birch *et al.*, 2008). Finally, AftB (Rv3805c) catalyzes the transfer of Ara*f* residues from DPA to the arabinan domain to form terminal β(1→2) linked Ara*f* residues that represents the "end-point" for AG arabinan biosynthesis before attachment with mycolic acids (Seidel *et al.*, 2007). Recently, a fourth member of this GT-C family of AraTs known as AftD (Rv0236c), which is essential for the viability and growth in *Msmeg* (Skovierova *et al.*, 2009), was reported to function as an α(1→5) arabinofuranosyltransferase responsible for elongation of the α(→3) primed bifurcation strands of the arabinan found in AG (Alderwick *et al.*, 2018).

The transfer of succinyl and **D**-GalN residues to the inner arabinan units marks the accomplishment of the primary structure of AG. This is followed by the formation of polyprenol-P-**D**-GalNAc from polyprenyl-P and UDP-GalNAc, which is catalysed by PpgS, a polyprenyl-phospho-*N*-acetylgalactosaminyl synthase. The polyprenol-P-**D**-GalNAc is then translocated across the membrane (Skovierova *et al.*, 2010, Rana *et al.*, 2012). The transfer of **D**-GalN to AG at the C-2 position of 3,5-branched Ara*f* residue is then catalysed by the glycosyltransferase, encoded by Rv3779. The arabinosylation of AG is followed by the formation of 1-*O*-phosphoryl linkage by the transfer of GlcNAc of the LU of mature AG to the 6-position of a MurNAc residue of PG (Yagi *et al.*, 2003). This transfer reaction is believed to occur in the presence of newly synthesized PG undergoing concomitant cross-linking (Hancock *et al.*, 2002). However, in *Staphylococcus aureus* (*S. aureus*), the wall teichoic acid (WTA, a major cell wall polyanionic polymer replacing AG) precursors has been found to attach only to uncrosslinked PG which implies that transfer of these precursors by WTA ligases occur only at an early stage of cell wall biosynthesis (Schaefer *et al.*, 2018). On the other hand, in *C. glutamicum*, neither AG arabinosylation nor its mycolylation are required for its attachment to PG (Alderwick *et al.*, 2005, Alderwick *et al.*, 2006a). The final stage of AG biosynthesis involves the attachment of AG to PG. Until recently, the enzymes responsible for coupling AG to PG were unknown. However, the seminal paper by Kawai and co-workers strongly suggested a role for members of the so-called LCP protein family (Kawai *et al.*, 2011). Further investigations in *Mtb* and *C. glutamicum* led to the discovery of the enzyme Lcp1 (Rv3267) (Grzegorzewicz *et al.*, 2016, Harrison *et al.*, 2016) and Cg0847 (Baumgart *et al.*, 2016) respectively, that catalyzes this essential ligation.
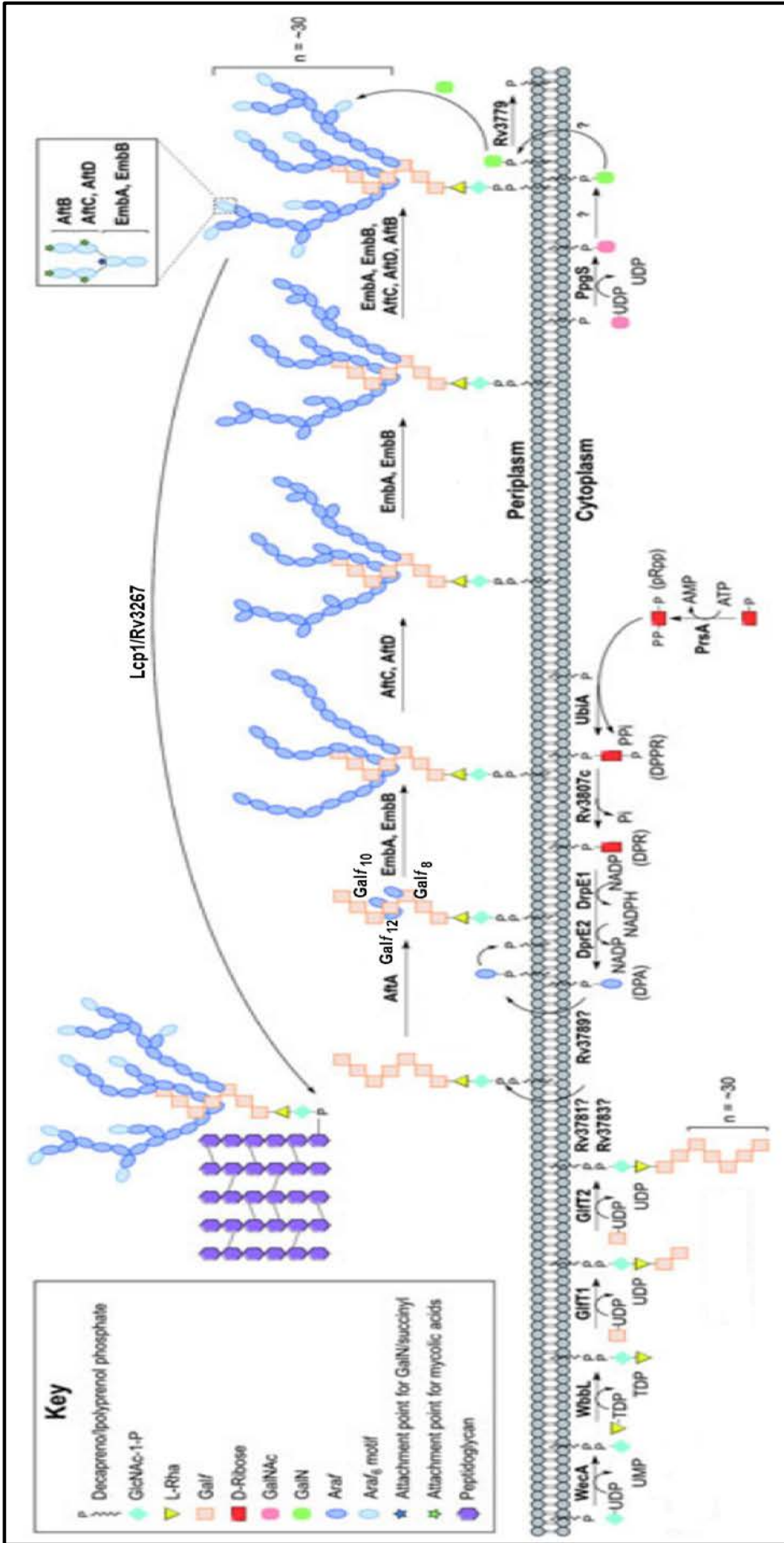
Fig 1.8. Schematic representation of Arabinogalactan biosynthesis in *Mtb*. The current knowledge of the roles of enzymes involved in AG biosynthesis has been represented here and described in the text. **Adapted from (Abrahams & Besra, 2016)**

1.8.   LytR-CpsA-Psr (LCP) proteins

Novel therapeutic agents and vaccines to complement and/or substitute the existing first-line treatment procedures are thus an immediate prerequisite to combat the threat of this disease. Keeping this in mind, substantial effort has been made to investigate the structure of mycobacterial cell envelope and its biosynthesis, to identify attractive drug targets. One such interesting target that has been studied by many researchers recently is the LCP family of proteins. The LCP proteins are found across all Gram-positive bacteria and in a few Gram-negative bacteria (Hubscher *et al.*, 2008). They are responsible for maintaining the structural integrity of the bacterial cell wall (Kawai *et al.*, 2011), thus promoting bacterial survival and propagation in the host organism. The Gram-positive bacterial genome often encodes several LCP proteins (in some cases, up to 11) (Baumgart *et al.*, 2016). Homologues of actinobacterial LCP proteins have been identified in the genomes of *Mtb*, *Mycobacterium marinum* (*M. marinum*), *Mycobacterium leprae* (*M. leprae*), *Mycobacterium bovis* (*M. bovis*), *Msmeg* and *C. glutamicum* (Table 1.3 and Fig 1.9).

Table 1.3: Orthologues of *lcp* genes across selected actinobacterial species

M. smegmatis	M. tuberculosis	M. marinum	M. leprae	M. bovis	C. glutamicum
MSMEG_0107	Rv3484	MMAR_4966	ML2247	Mb3514	-
MSMEG_1824	Rv3267	MMAR_1274	ML0750	Mb3295	Cg0847* (Cgl0740)**
MSMEG_5775	Rv0822c	MMAR_4858	ML2187	Mb0845c	Cg3210* (Cgl2902)**
MSMEG_6421	Rv3840	MMAR_5392	-	Mb3903	-

*The annotation of the genes as in Baumgart et al., 2016; **The annotation of the genes as in UNIPROT database (https://www.uniprot.org/)
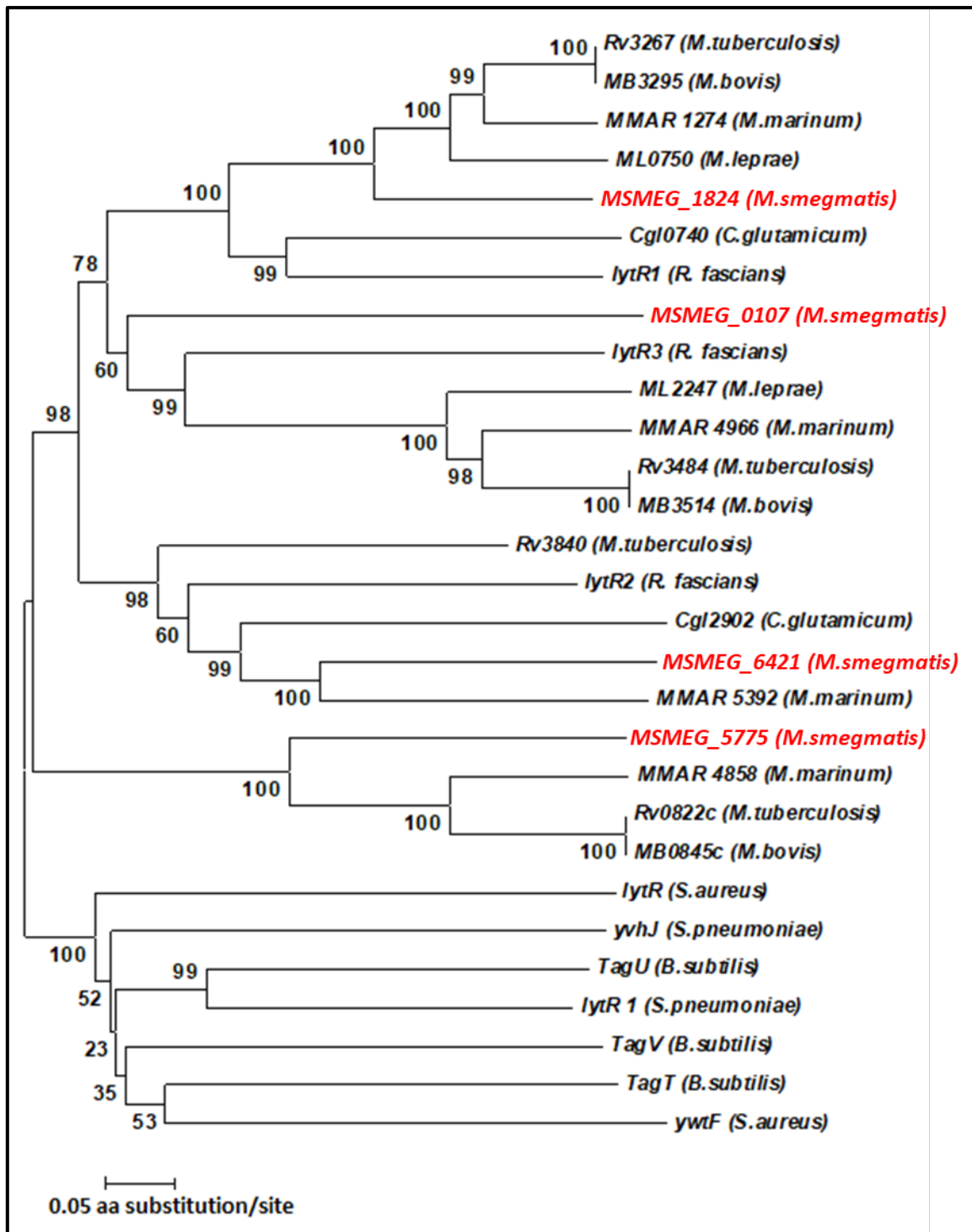
Fig 1.9. Evolutionary relationship of LCP proteins in *Msmeg* and other Gram-positive and actinobacterial species. Neighbor-Joining tree of the LCP proteins in various Gram positive organisms and actinobacterial species. The tree is constructed using ClustalW algorithm with a PAM protein weight matrix and a divergent cutoff of 70% in MEGA7 software (Kumar *et al.*, 2016). The four LCP homologs in *Msmeg* are in red.

aa- amino acid

The LCP proteins are named after the representatives of the family that were reported separately - LytR (Lazarevic *et al.*, 1992), CpsA (Cieslewicz *et al.*, 2001) and Psr (Ligozzi *et al.*, 1993). The role of LytR in *Bacillus subtilis* (*B. subtilis*) was initially described as an attenuator of both itself and the adjacent genes (Lazarevic *et al.*, 1992). However, this was later described to be misannotated, and instead suggested these proteins to be enzymes required for transferring wall teichoic acids and capsular polysaccharides onto PG (Kawai *et al.*, 2011, Chan *et al.*, 2014). CpsA, on the other hand was first studied in *Streptococcus agalactiae* (*S. agalactiae*), and proposed to have a role in transcription activation of the synthesis of capsular polysaccharides (Cieslewicz *et al.*, 2001). Initially Psr was suggested as a repressor of penicillin-binding protein 5 (PBP5) synthesis in *Enterococcus hirae* (Ligozzi *et al.*, 1993), however the effect of this protein on PBP5 synthesis, autolysis or β-lactam resistance could not be confirmed (Sapunaric *et al.*, 2003).

LCP proteins are involved in maintaining the structural integrity of the cell envelope in many Gram-positive bacteria, and strains devoid of these proteins have been associated with cell wall defects (Chan *et al.*, 2013, Chan *et al.*, 2014, Wang *et al.*, 2015). Apart from the effects of *lcp* genes on the cell envelope of Gram-positive organisms, the impacts have also been recently studied in actinobacteria including some species of mycobacteria. For instance, a conditional mutant of an essential *lcp* gene Cg0847, in *C. glutamicum* was shown to have severe morphological alterations, growth defects as well as reduced mycolic acids and AG, all of which indicates their role in structural integrity (Baumgart *et al.*, 2016). In *M. marinum*, a CpsA transposon mutant was shown to have altered bacterial colony morphology, sliding motility, cell surface hydrophobicity, cell wall permeability (Wang *et al.*, 2015). Here, the mutant also exhibited a decreased AG content, suggesting the role of CpsA in cell wall assembly. In *Mtb*, a mutant of Rv3484 (orthologue of *MSMEG_0107*) was shown to be essential for growth of *Mtb* in mice (Malm *et al.*, 2018). This mutant also showed higher resistance to lysozyme and other antibacterial compounds like meropenem/clavulanate that targets the PG synthesis. A different study on Rv3484 mutant demonstrated this LCP protein to aid *Mtb* from lysosomal clearance thus evading the host innate immunity (Koster *et al.*, 2017). Apart from the impact of these proteins on the cell envelope, their function as a ligase that couples AG to PG has also been recently described in many organisms such as *C. glutamicum*, *Mtb* and *S. aureus* (Baumgart *et al.*, 2016, Grzegorzewicz *et al.*, 2016, Harrison *et al.*, 2016, Schaefer *et al.*, 2017). Another study in *B.subtilis*, showed LCP proteins to anchor WTAs to PG *in vitro* (Gale *et al.*, 2017). All these findings imply the involvement of LCP proteins in transferring the saccharide entities from a lipid-polyprenol carrier to PG. However, in spite of extensive studies on the structural components involved in tethering AG and PG in the mycobacterial cell envelope, the exact sequence of events to couple these two major heteropolysaccharides still remains elusive.

The LCP proteins generally share a common structural organization i.e., an N-terminal cytoplasmic domain, a transmembrane region and an extracellular region that includes the LCP domain (Hubscher *et al.*, 2008). However, some do not have the transmembrane domain, such as Rv3840 in *Mtb*. The LCP domain is found predominantly in Firmicutes and Actinobacteria (Baumgart *et al.*, 2016). Apart from the LCP domain, a complementary LytR_C domain of unknown function is found at the C-terminal of most actinobacterial LCP proteins and, both these domains either occur separately or in association (Baumgart *et al.*, 2016). For instance, in *Msmeg*, all the three *lcp* genes except *MSMEG_6421* have the LytR_C domain (Fig 1.10).

1.9. Aims of the thesis

Although two different recent studies in *Mtb* have already shown *Rv3267* and *Rv3484* to couple AG to PG (Grzegorzewicz *et al.*, 2016, Harrison *et al.*, 2016), several aspects of the cell wall assembly involving this attachment during the final phase of AG synthesis have still remained elusive. Moreover, not all LCP homologues have been studied in depth in the non-pathogenic model organism *Msmeg*, and the role of orthologues of *MSMEG_6421* have not been reported in any species. In this PhD project, the following aims were considered:

i) To better understand the participation of single or multiple LCP proteins in exhibiting the same enzymatic function of coupling AG to PG.

ii) To investigate potential differences between essential and non-essential *lcp* genes in *Msmeg*.

Hence, in this whole study, the function of all the four *lcp* genes in *Msmeg* were comprehended by measuring the pyrophosphatase activity exhibited by each of these four homologues. Apart, this is the first comprehensive study to report the impact on the mycobacterial cell envelope by creating both single and double *lcp* deletion mutants. Mycobacteria are known for its survival fitness due to its tough cell envelope that helps it to evade several environmental stress. Therefore, in combination to understand the role of all four *lcp* genes in *Msmeg*, the differential expression between these four homologues as well as the effects of different stressors on the fitness of these mutants were extensively investigated. Thus, this study would help us to understand the role of all the four LCP proteins in this non-tubercle bacilli, implicating their possible and related mechanisms in cell envelope biogenesis in *Msmeg*.
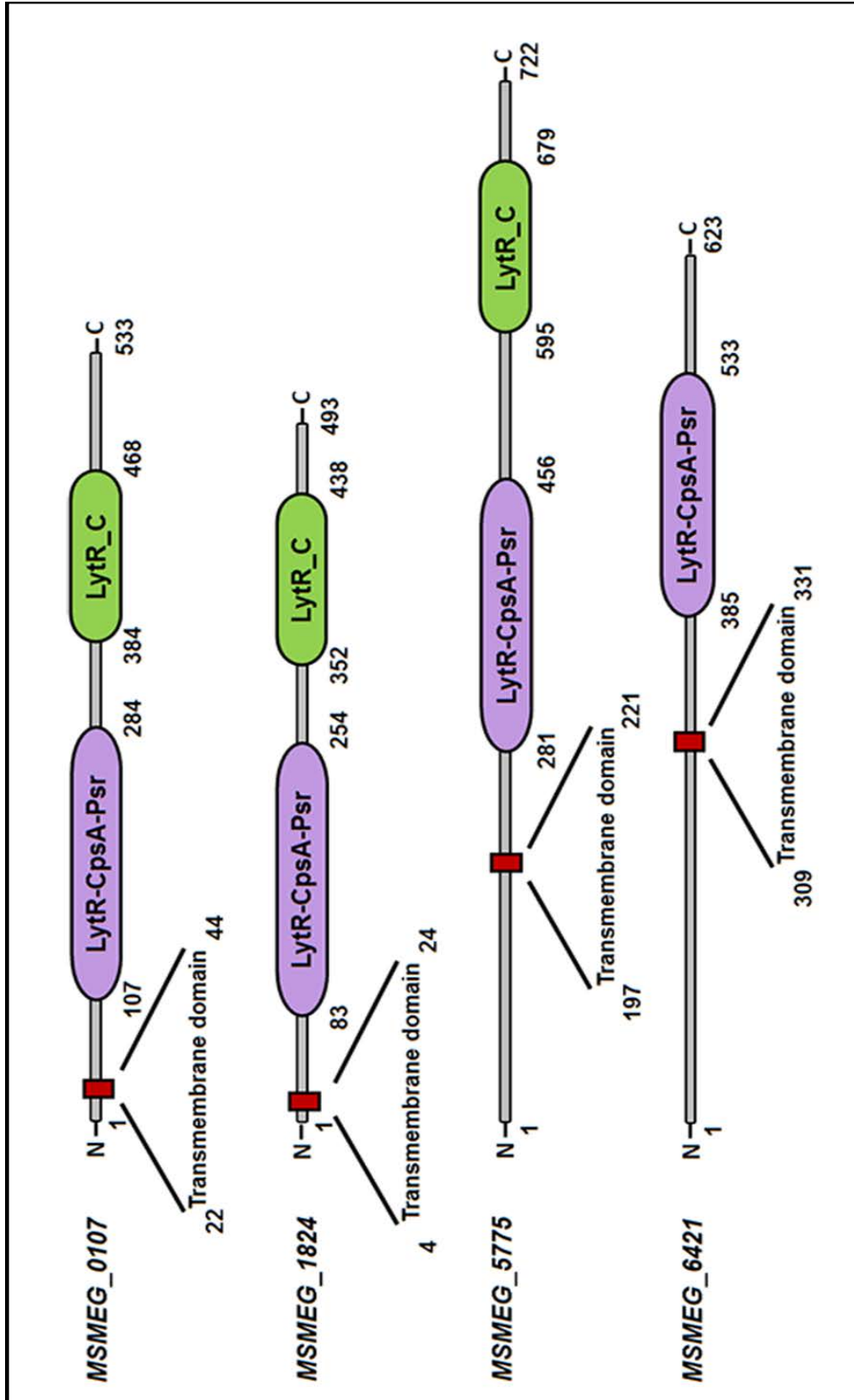
Fig 1.10. Schematic representation of the *lcp* genes in *Msmeg*. The domains of the full length protein sequence of each of the four LCP homologs in *Msmeg* were approximately mapped to scale after determining the transmembrane region from TMHMM server v. 2.0, and LytR-CpsA-Psr and LytR_C domains from UNIPROT database. The amino acid sequence spanning each of the domains has been numbered with respect to the full length LCP protein.

25

1.10. Outline of the thesis

The main goal of this thesis was to gain insights into the functional aspects of the *Msmeg lcp* genes, and the necessity of the four homologues in this non-pathogenic model organism.

In **chapter 2**, pyrophosphatase activity of the four LCP homologues in *Msmeg* is described, confirming what has been recently reported for LCP homologues in *Mtb* (Grzegorzewicz *et al.*, 2016, Harrison *et al.*, 2016) and *C. glutamicum* (Baumgart *et al.*, 2016). This activity is based on the background that, in the presence of a generic substrate geranyl pyrophosphate (GPP) and Mg^{+2}, the LCP proteins release inorganic phosphate, which is measured spectrophotometrically, to determine the extent of coupling between AG and PG by these proteins. MSMEG_0107 and MSMEG_5775 was found to release significantly higher amount of inorganic phosphate than the essential MSMEG_1824, indicating that both MSMEG_0107 and MSMEG_5775 have higher pyrophosphatase activity during the ligation of AG and PG in *Msmeg*.

In **chapter 3**, the impact of the *lcp* genes on physiology of *Msmeg* was studied by creating single and double deletion mutants of the non-essential *lcp* genes. Though a single deletion of these genes did not give profound effects in morphological features, however, the double deletion of *MSMEG_0107* and *MSMEG_5775* demonstrated a compromised cell envelope by exhibiting structural defects, slower growth rate, diminished biofilm on air-liquid interface, increased sensitivity to detergent and antibacterial agents like lysozyme and antibiotics, and lower aggregative properties. Finally, the exposed Gal*f* residues of AG were identified by EB-A2 Mab in the compromised cell envelope of the double *lcp* deletion mutant lacking both *MSMEG_0107* and *MSMEG_5775*.

In **chapter 4**, the effect of various environmental stresses on the mRNA expression levels of the four *lcp* genes in wild type and single and double-deletion mutants were investigated. For this study, quantitative RT-PCR method was used extensively to determine the differential expression of the *lcp* homologues in the *Msmeg* mutants at a late-log phase. Under standard conditions, ΔMSMEG_5775 exhibited an extraordinary upregulation of all the *lcp* genes than that in the wild type, suggesting MSMEG_5775 to play an important role in ligating AG and PG in *Msmeg*. Under stress conditions, such as in the presence of lysozyme and low pH, loss of *MSMEG_5775* downregulated other available *lcp* genes in the *Msmeg* genome, whereas, loss of *MSMEG_6421* upregulated them suggesting the role of a single bi-functional regulator or two different regulators. Finally, *in silico* investigation on the regulatory elements of *lcp* genes, revealed putative regulators which could be further investigated.

In the **general discussion**, the results are summarized and further discussed.

Chapter-2

All four LCP homologues in *Mycobacterium smegmatis* possess pyrophosphatase activity

Abhipsa Sahu, Ziwen Xie, David Ruiz-Carrillo and Boris Tefsen

Department of Biological Sciences, Xi'an Jiaotong-Liverpool University, Suzhou, Jiangsu, 215123, PR China

2.1.   Introduction

The composition and the complex architecture of the mycobacterial cell envelope distinguishes it from other prokaryotes. The cell envelope comprises of many physiological features that makes it susceptible and/or resistant to many anti-TB drugs (Barry *et al.*, 2007, Jackson *et al.*, 2013). Therefore, several studies and overviews on the cell envelope structure have recently aimed to gain insights into the biosynthetic machinery of this organism, and identify attractive drug targets (Kawai *et al.*, 2011, West *et al.*, 2011, Jankute *et al.*, 2012, North *et al.*, 2014, Abrahams & Besra, 2016, Bhat *et al.*, 2017). In the cell walls of mycobacteria, the PG is a crosslinked network of amino acids and glycans in which about 10-12% of MurNAc residues are covalently substituted at their C-6 position with the polysaccharide AG (Yokoyama *et al.*, 1986, McNeil *et al.*, 1990). As AG is acylated with mycolic acids (Mikusova *et al.*, 1996), these three different compounds thus form a complex network referred to as mAGP-complex (Jankute *et al.*, 2015).

The LCP family of proteins have paved way for many researchers to investigate their importance in cell envelope biogenesis in various Gram-positive pathogens. Recent advances have also shown the LCP proteins Rv3267 and Rv3484 in *Mtb*, and Cg0847 in *C. glutamicum* to be responsible for ligating AG and PG, in the later stages of cell wall biosynthesis (Baumgart *et al.*, 2016, Grzegorzewicz *et al.*, 2016, Harrison *et al.*, 2016), however, their significance or functional role in the non-pathogenic model organism, *Msmeg*, have not been studied yet. *Msmeg* is a fast-growing species that is amenable to genetic manipulation, easy to cultivate and handle in the laboratories devoid of BSL-3 facilities. There are many similarities between *Msmeg* and the much more virulent pathogens, with more than 2000 homologous genes shared with *Mtb*. These properties make it a very useful model organism for *Mtb* and other mycobacterial pathogens. Interestingly, *Msmeg* cell envelope has four LCP homologues, while other organisms can survive with only two or even one AG-PG ligase.

The aims of the study described in this chapter were to understand whether all four orthologues in *Msmeg* viz., MSMEG_0107, MSMEG_1824, MSMEG_5775 and MSMEG_6421 possess the same pyrosphosphatase activity as described for Rv 3267, Rv3484 and Rv0822c in *Mtb* (Grzegorzewicz *et al.*, 2016) and Cg0847 in *C. glutamicum* (Baumgart *et al.*, 2016). First, the LCP and LytR_C domains of mycobacterial and non-mycobacterial species were mapped via an *in silico* approach, with the goal of providing an understanding of the conservative nature of the LCP protein family in mycobacteria and other related organisms. Moreover, structural predictions of the four homologues were made using SWISS MODEL webserver. Finally, transmembrane-less variants of all four LCP homologues were purified and the enzymatic activity determined. Not only was it confirmed that MSMEG_0107, MSMEG_1824 and MSMEG_5775 possess the expected pyrophosphatase activity, but also MSMEG_6421, the least studied LCP homologue, was able to hydrolyze pyrophosphate at a comparable rate under the conditions tested. The dependency on Mg^{2+} was shown for all four enzymes and a mutation in an arginine residue in the suspected enzyme active site of MSMEG_1824 was shown to diminish the pyrophosphatase activity by 1.6-fold.

2.2. Materials and methods

2.2.1. Bacterial strains, plasmids and growth conditions

All bacterial strains and plasmids used in this chapter are listed in Table 2.1. All the reagents were purchased from Sigma-Aldrich, China unless otherwise stated. *Msmeg* wild-type strain mc^2155 was grown in 7H9 medium (Difco) supplemented with 10% (v/v) Albumin Dextrose Saline (ADS, 50 g of albumin, 20 g of dextrose and 8.5 g of sodium chloride in 1 liter of water), 0.5% glycerol (v/v) and 0.05% Tween-80 (v/v) (Solarbio, China). Plasmid propagation was performed using *E.coli* DH5-α (Tiangen Biotech Beijing Co. Ltd., China), genotype: F⁻, φ80, *lac*ZΔM15, Δ (*lac*ZYA-argF) U169 *end*A1, *rec*A1, *hsd*R17 (r_k^-, m_k^+) *sup*E44, λ⁻, thi-1, *gyr*A96, *rel*A1, *pho*A. For expression of fusion proteins with pGEX4T1 and pET21a vector constructs, *E.coli* BL21 (DE3) (Tiangen Biotech Beijing Co. Ltd., China) strain was used, genotype: F⁻ *omp*T *hsd*S_B (r_B^- m_B^-) *gal dcm* (DE3) pLysS Camʳ. In liquid culture, *E. coli* strains were grown in LB broth for 16-20 hours at 37°C and shaking at 200 rpm. In solid culture, *E. coli* strains were grown on LB agar plates for 16-20 hours at 37°C, and the plates stored at 4°C for up to one month. For long term storage of *E. coli* strains, fresh cultures in mid-log phase (optical density measured at 600 nm (OD_{600}) of 0.6 - 0.8) were frozen in 50% glycerol and stored at -80°C. Media was autoclaved for 15 min at 121°C and cooled to below 50°C prior to addition

of the appropriate antibiotic (in this chapter, all cultures were supplemented with a final concentration of 0.1 mg/mL of ampicillin), unless otherwise stated.

2.2.2. Phylogenetic analysis of LCP proteins

The nucleotide and amino acid sequences of the *lcp* genes of *Msmeg* mc^{2}155 strain and other organisms were taken from the NCBI database (accession numbers in Table 2.2) and UNIPROT Knowledgebase. The sequences were aligned with the Clustal W algorithm (Thompson *et al.*, 1994, Thompson *et al.*, 1997) in the MEGA7 software (Kumar *et al.*, 2016) using a PAM protein weight matrix and with a divergent cutoff of 70%. The amino acid sequence homology and divergence was determined using DNASTAR (version 15.3.0.66) program package (DNASTAR, Inc., USA) and the phylogenetic analysis was performed using MEGA7 software (Kumar *et al.*, 2016). The evolutionary history was inferred using the Neighbor-Joining method (Saitou & Nei, 1987) with confidence limits of the branching pattern evaluated by bootstrapping of 10,000 replicates of the original data sets. The evolutionary distances were computed using the *p*-distance method (Nei & Kumar, 2000) and are in the units of the number of amino acid differences per site. The rate variation among sites was modeled with a gamma distribution (shape parameter = 10). The analysis involved 29 amino acid sequences. All ambiguous positions were removed for each sequence pair.

2.2.3. Structural prediction of LCPΔTM proteins

For structural prediction, the primary full length sequences of the four LCP proteins in *Msmeg* were submitted for homology modelling in SWISS-MODEL available at https://swissmodel.expasy.org/. The models were built on the basis of their higher Global Model Quality Estimation (GMQE) scores instead of QMEAN scores. QMEAN (Benkert *et al.*, 2011) is a composite estimator based on different geometrical properties and provides both global (i.e. for the entire structure) and local (i.e. per residue) absolute quality estimates on the basis of one single model whereas, GMQE is a quality estimation which combines properties from the target–template alignment and the template search method. The resulting GMQE score is expressed as a number between 0 and 1, reflecting the expected accuracy of a model built with that alignment and template and the coverage of the target. Higher numbers indicate higher reliability. Once a model is built, the GMQE gets updated for this specific case by also taking into account the QMEAN score of the obtained model in order to increase reliability of the quality estimation. After protein homology modelling, PYMOL v. 2.3 was used to map the LCP domain and the conserved residues of the LCP proteins.

Table 2.1: Strains and plasmids used to study pyrophosphatase activity in *Msmeg* LCP proteins

Strain or plasmid	Description	Source or reference
E. coli		
DH5-α	strain used for general cloning procedures	
BL21 (DE3)	host for protein expression	
Msmeg		
mc2155	wild-type laboratory strain; DNA used as PCR template	(Snapper *et al.*, 1990)
Plasmids		
*MSMEG_0107*ΔTM	Amp^R; transmembrane-less *MSMEG_0107* cloned in pGEX-4T1	This study
*MSMEG_1824*ΔTM	Amp^R; transmembrane-less *MSMEG_1824* cloned in pGEX-4T1	This study
*mut-MSMEG_1824-R225A*ΔTM	Amp^R; transmembrane-less *MSMEG_1824* with amino acid substitution R→A at position 225 of the full length protein cloned in pGEX-4T1	This study
*MSMEG_6421*ΔTM	Amp^R; transmembrane-less *MSMEG_6421* cloned in pGEX-4T1	This study
*MSMEG_5775*ΔTM	Amp^R; transmembrane-less *MSMEG_5775* cloned in pET21a	This study

Amp^R- ampicillin resistance

Table 2.2: Strains and their Genbank IDs used for phylogenetic analysis

Organism	Strain	Genbank ID
Mycobacterium smegmatis	mc^{2}155	CP000480.1
Mycobacterium tuberculosis	H37Rv	AL123456.3
Mycobacterium marinum	M	CP000854.1
Mycobacterium leprae	TN	AL450380.1
Mycobacterium bovis	AF2122/97	LT708304.1
Bacillus subtilis	168	AL009126.3
Staphylococcus aureus	NCTC7887	UHAH01000002.1
Streptococcus pneumoniae	N	CRBU01000005.1
Rhodococcus fascians	D188	NZ_CP015235.1
Corynebacterium glutamicum	ATCC 13032	BX927147.1

2.2.4. Extraction of genomic DNA from *Msmeg*

Genomic DNA was extracted from *Msmeg* wild-type strain mc^{2}155 for use as template to amplify the *lcp* genes for cloning purposes. The bacteria were grown until OD_{600} reached 1, followed by centrifuging at 16,000 *g* for 5 min, and the medium discarded. The pellet was resuspended in 500 µL of lysis buffer (40 mM Tris-HAc pH 7.8, 20 mM NaAc, 1 mM EDTA, 10% Sodium dodecyl sulfate) and the cells were broken in a mini bead-beater (Biospec products) for a min using about 1 g of 5 mm sized glass beads. The broken cells were then treated with 300 µL of 5 M NaCl and centrifuged for another min at 16,000 *g*. The supernatant containing the nucleic acids was moved to another tube and treated with 800 µL of trichloromethane (Sinopharm Chemical Reagent Co. Ltd., China) and centrifuged for a min at 16,000 *g*. The supernatant was then treated with two volumes of 96% ethanol (Sinopharm Chemical Reagent Co. Ltd., China) and kept standing at room temperature for 30 min followed by centrifuging at 16,000 *g* for a min. The supernatant was then discarded and the pellet washed with 500 µL of 70% ethanol and centrifuged at 16,000 *g* for 2 min. After removal of the supernatant, traces of remaining ethanol were drained by inverting the tube and tapping on a paper towel followed by drying on air for about 10-15 min. The dried pellet was dissolved in 50 µL double-distilled water and used as a template. The genomic DNA concentration was measured using NanoDrop 2000 spectrophotometer (Thermo Scientific, China).

2.2.5. Cloning of *lcp* genes devoid of transmembrane domain

2.2.5.1. *In silico* cloning of *lcp*ΔTM in expression vector

The transmembrane domains are highly hydrophobic in nature that tend to interfere during protein expression, hence *Msmeg lcp* genes were cloned devoid of their N-terminal transmembrane domain. The transmembrane spanning region in the amino acid sequence was predicted using the TMHMM server available at the Expasy Bioinformatics Resource Portal website. The recombinant *lcp*ΔTM plasmids were constructed using SnapGene software (from GSL Biotech; available at snapgene.com, San Diego, California, USA) and all the molecular features namely, promoter sequence, antibiotic resistance marker, primer recognition sites and restriction sites were marked in the construct (Fig 2.1). The recombinant *lcp*ΔTM genes, *MSMEG_0107*ΔTM (corresponding to amino acid residues 51-533 of the total protein), *MSMEG_1824*ΔTM (corresponding to amino residues 25-493 of the total protein), *MSMEG_5775*ΔTM (corresponding to amino acid residues 222-722 of the total protein) and *MSMEG_6421*ΔTM (corresponding to amino acid residues 332-623 of the total protein) were constructed using the primers mentioned in Table 2.3. These primers harbor an N-terminal glutathione S-transferase (GST) tag, hexahistidine (6XHis) tag and a human Rhinovirus 3C (HRV3C) protease site. Restriction sites were added to the 5' end of the primer sequence.

2.2.5.2. Polymerase Chain Reaction and agarose gel electrophoresis

For cloning, a two-step PCR was performed. In the first PCR, primer sets MSMEG_0107-F-pGEX4T1 and MSMEG_0107-R-*Sma*I were used to amplify *MSMEG_0107*ΔTM. In the second PCR, the gel purified PCR product from the first PCR was used as a template, using primer sets MSMEG_0107-F-*EcoR*I and MSMEG_0107-R-*Sma*I. Similarly, other *lcp*ΔTM genes were amplified. A typical 50 μL reaction mixture contained 1X Q5 Reaction buffer, 300 μM of dNTPs, 0.5 μM of each primer, 1X Q5 GC enhancer, 0.008 U/ μL Q5$^{®}$ High-Fidelity DNA Polymerase (NEB) and 0.5-1 μg of template DNA. PCRs were performed using a Veriti 96 well Thermal cycler (Applied Biosystems, Thermo Fisher Scientific, China) with cycling conditions: 35 cycles of denaturation, annealing and extension at 95°C, 70°C and 72°C for 10 sec, 30 sec and 1 kb/min respectively.

PCR products were confirmed by agarose gel electrophoresis using 0.8-1% agarose in 1X TAE (40 mM Tris-acetic acid, 10 mM EDTA pH 8.0) buffer containing 0.01% Gel-Red (Beyotime, China) and a 1 kb DNA Ladder (Takara Bio Inc., China) was used as DNA marker.

The gel loaded with samples and marker was run for about 1 hour at 100 volts, until an appropriate resolution was achieved. The gel was then imaged using a UV transilluminator system (Bio-Rad Gel Doc™ XR⁺ Imaging System, China).
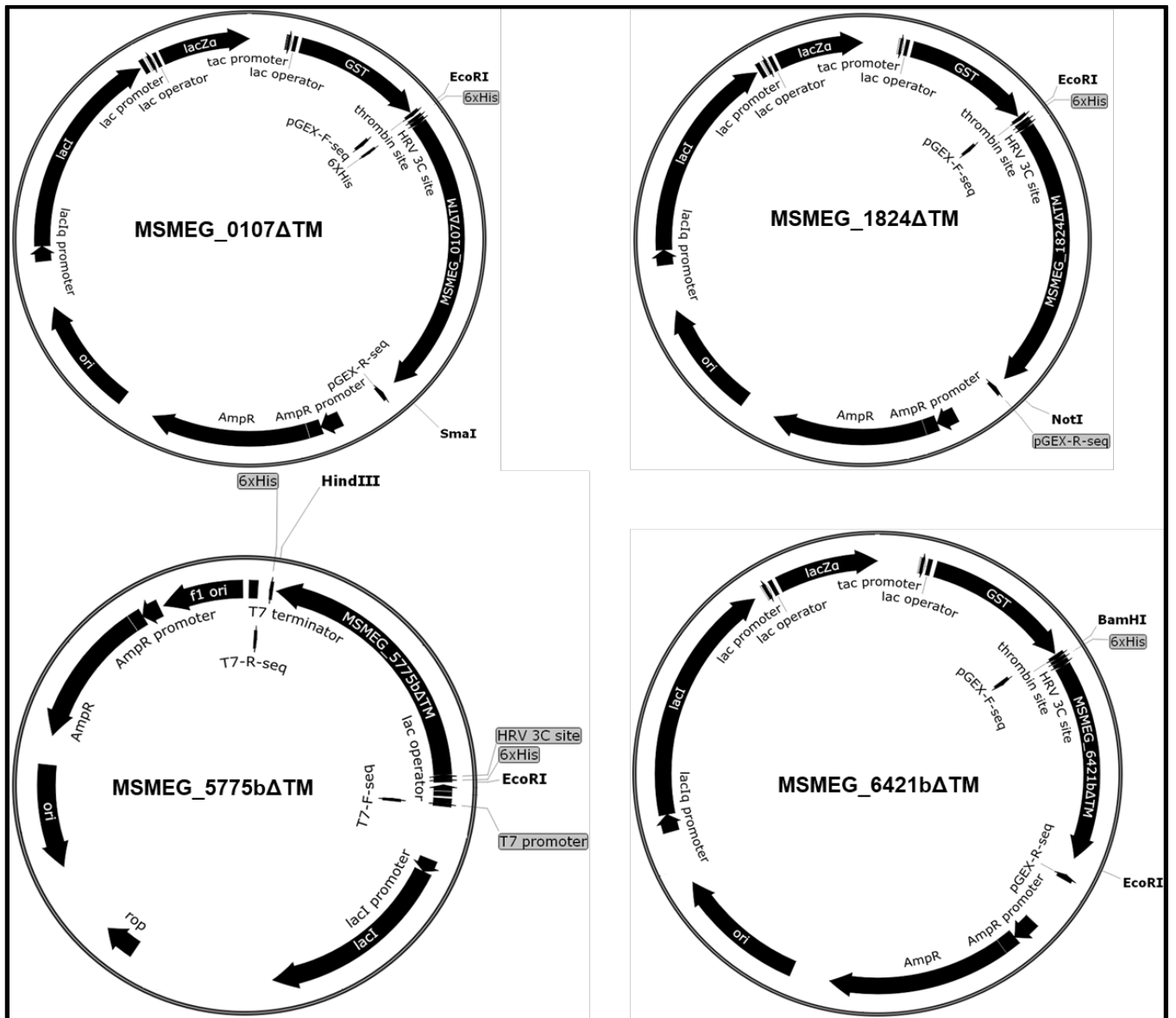
Fig 2.1. Construction of all the four recombinant LCPΔTM proteins in *Msmeg*. Transmembrane (TM) of *lcp* genes were predicted using TMHMM webserver. The subsequent transmembrane-less (ΔTM) gene sequences were used to construct *in silico* cloned maps of the *lcp*ΔTM proteins. An HRV3C cleavage site and a 6X*his*-tag was cloned at the N-terminal of *MSMEG_0107*ΔTM, *MSMEG_1824*ΔTM and *MSMEG_6421*ΔTM in pGEX 4T-1 plasmid in such a way that the Tac promoter and GST-tag already present in the plasmid is located in the N-terminal of the his-HRV3C-*lcp*ΔTM protein. In case of *MSMEG_5775*ΔTM, pET21a plasmid was used for cloning and, the 6X*his*-HRV3C-*MSMEG_5775*ΔTM was cloned with the T7 promoter in the N-terminal. An additional 6XHis-tag from the plasmid is located at the C-terminal of the fused protein. The primers used for sequencing the recombinant plasmid have been marked at specific location of the construct.

2.2.5.3. Gel extraction of DNA

DNA fragments of appropriate length were cut with a scalpel under a UV trans-illuminator ULTRA-LUM (Carson, USA) and purified using QIAquick Gel Extraction Kit (Qiagen, China) following the manufacturer's instructions. Samples were eluted in 30-50 µL of double-distilled water.

2.2.5.4. Restriction digestion, ligation and transformation

All restriction enzymes used were purchased from NEB. Typically, 1 µg of the appropriate purified expression vector and the entire purified PCR product were digested with the appropriate 10X NEB buffer and 10 U each of the corresponding restriction enzymes for double digestion, in a final reaction volume of 50 µL, according to the manufacturer's instructions. The digested vector and the insert DNA were then resolved on agarose gel, and the desired DNA fragments were excised and purified by gel extraction (section 2.2.5.3).

2.2.5.5. Ligation of insert DNA into expression vector

Ligation reactions of the digested vector and the insert DNA were performed using 0.5 µL T4 DNA ligase (NEB) and 1 µL of 10X ligase buffer, in a 10 µL reaction. A typical ligation reaction consisted of 3:1 molar ratio of insert to vector. Reactions were mixed well before incubating at room temperature (approximately 25°C) for 1 hour. Negative control reactions were devoid of the target gene insert.

Table 2.3: Primers used for cloning and sequencing the recombinant plasmids for subsequent pyrophosphatase assay

Primer name	Sequence (5' - 3')
Cloning primers	
MSMEG_0107-F-pGEX4T1	GGGCCCGACGGCATCACGACGTCCAA
MSMEG_0107-R-*Sma*I	CGGCCCGGGTCATTTCACGCACGGGATGCC
MSMEG_0107-F-*EcoR*I	CTCGAATTCCACCATCATCATCATCATCTTGAGGTTCTTTTCCAGGGGCCCGACGGCATCACGACGTCCAA
MSMEG_1824-F-pGEX4T1	GGGCCCCGGTCGTTCGAGTCCGGTATC
MSMEG_1824-R-*Not*I	CGGGCGGCCGCTCAGTTCACGCACTGCGG
MSMEG_1824-F-*EcoR*I	CCGGAATTCCACCATCATCATCATCATCTTGAGGTTCTTTTCCAGGGGCCCCGGTCGTTCGAGTCCGGTATC
MSMEG_5775-F-pET21a	GGGCCCCAGTCGTCGAAGAACGACTC
MSMEG_5775-R-*Hind*III	CGGAAGCTTTCATTCGCAGGTGGCGTC
MSMEG_5775-F-*EcoR*I	CTCGAATTCCACCATCATCATCATCATCTTGAGGTTCTTTTCCAGGGGCCCCAGTCGTCGAAG
MSMEG_6421-F-pGEX4T1	GGGCCCGACACGTCACTGCAACG
MSMEG_6421-R-*EcoR*I	CGGGAATTCTCAATTGACGGCATTTTCCAGG
MSMEG_6421-F-*BamH*I	CTCGGATCCCACCATCATCATCATCATCTTGAGGTTCTTTTCCAGGGGCCCGACACGTCACTG
MSMEG_1824-R225A-F	GCGTTGAGCTTCGTGGCCCAGCGCCACGGCCTG
MSMEG_1824-R225A-R	CAGGCCGTGGCGCTGGGCCACGAAGCTCAACGC
Sequencing primers	
pGEX-F-seq	GGGCTGGCAAGCCACGTTTGGTG
pGEX-R-seq	CCGGGAGCTGCATGTGTCAGAGG
T7-F-seq	TAATACGACTCACTATAGGG
T7-R-seq	GCTAGTTATTGCTCAGCGG

2.2.5.6. Transformation of vector constructs into *E. coli*

10 µL of the ligation mixture was mixed with 20 µL of competent DH5-α and incubated on ice for 30 min, 90 seconds at 42°C on a heat block, and on ice for 3 min before adding 500 µL of LB broth. The cells were incubated at 37°C for 1 hour, centrifuged at 16,000 *g* for 1 min and 50 µL of the media resuspended with the pellets were plated onto LB agar plates supplemented with ampicillin and were incubated at 37°C overnight.

2.2.6. Colony PCR, restriction analysis and sequencing

Several single transformed colonies were resuspended with a 10 µL PCR reaction mix, using the reaction components and cycling conditions as mentioned in section 2.2.5.2. The presence of an inserted gene was confirmed by agarose gel electrophoresis. Restriction analysis of a positive transformant was performed as in section 2.2.5.4 using 10-20 ng of the extracted recombinant plasmid DNA, and the presence of insert was confirmed by agarose gel electrophoresis. Sequencing reactions for the recombinants were done by Sangon Biotech Co. Ltd, Shanghai, China. The sequences were retrieved and the insert confirmed using Chromas Lite, version 2.6.6. (Technelysium Pty Ltd, South Brisbane, Australia).

2.2.7. Site-directed mutagenesis for construction of mut-*MSMEG_1824-R225A*ΔTM

For cloning *mut-MSMEG_1824-R225A*ΔTM, primers MSMEG_1824-R225A-F and MSMEG_1824-R225A-R were used for PCR, using *MSMEG_1824*ΔTM plasmid DNA as the template. The primers were designed in such a manner that both of them contain the desired mutation in the middle with 15 nucleotide bases of correct sequence on both sides (Fig 2.2) and, terminating one or more C or G bases (Table 2.3), desired mutations are in red). PCR reaction was performed using Q5® High-Fidelity DNA Polymerase (NEB) as described in section 2.2.5.2. Typical PCR cycling conditions for site-directed mutagenesis had 18 cycles of denaturation, annealing and extension at 95°C, 60°C and 68°C for 20 sec, 10 sec and 1 min/kb of plasmid length respectively.

Following PCR, 20 U of *Dpn*I restriction enzyme (NEB) was added directly to the PCR amplified product, resuspended well and incubated immediately for 1 hour at 37°C, to digest the parental (i.e. the non-mutated) supercoiled double stranded DNA. 2 µL of this digested product was then transformed in DH5-α (section 2.2.5.6) and plated on LB agar supplemented with ampicillin. Colony PCR, restriction analysis and sequencing were done as in section 2.2.6.
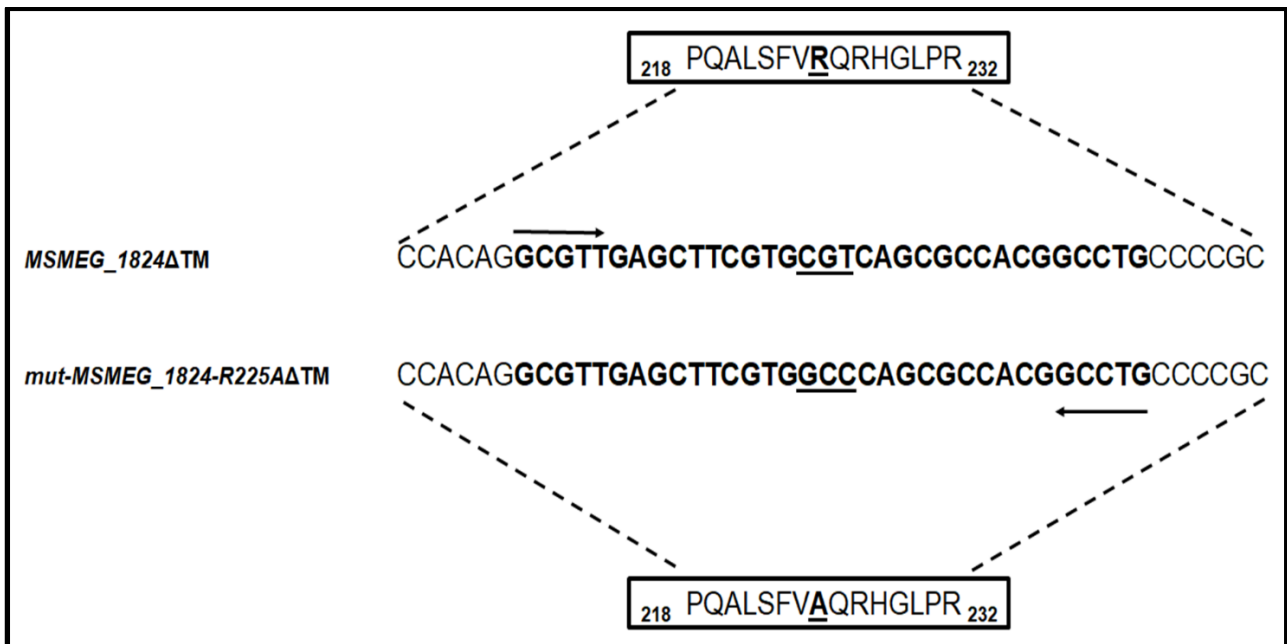
Fig 2.2. Site-directed mutagenesis in *mut-MSMEG_1824*-R225AΔTM. The amino acid stretch from residues 218-232 of the full length protein is represented with the mutated residue (bold and underlined). The corresponding nucleotides containing the mutation are depicted in bold and underlined. The primer sequences are in bold. The arrows represent the orientation of the primers.

2.2.8. Small-scale expression of truncated LCP proteins

Expression vectors containing the *lcp*ΔTM genes were transformed in BL21 (DE3) following the same protocol as for DH5-α (section 2.2.5.6), followed by small-scale expression of the LCPΔTM proteins. For this, a single colony was picked from the BL21 (DE3) transformants and cultured overnight in 5 mL of LB broth supplemented with 0.1 mg/mL ampicillin, at 37°C and shaking at 200 rpm. For small-scale expression, 200 μL of the overnight culture was inoculated into 5 mL of fresh LB broth supplemented with ampicillin and, cultured at 37°C and shaking at 200 rpm, for about 1.5 hours or, until the OD_{600} reached 0.6. Then, the cells were induced with Isopropyl β-D-1-thiogalactopyranoside (IPTG) at a final concentration of 0.2 mM and grown further for 3 hours. Then, 1.5 mL of the induced culture was taken, centrifuged at 16,000 *g* for a min and then the supernatant discarded. Protein samples were then prepared by dissolving the bacterial pellet in 100 μL of 1X Laemmli sample buffer (12 mM Tris-HCl pH 6.8, 0.4% SDS, 2% glycerol, 1% β-mercaptoethanol, 0.002% OrangeG dye) and incubated in a heat block at 95°C for 10 min. The samples were then cooled on ice prior to separation by SDS-PAGE (section 2.2.9).

2.2.9. Sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE)

The protein samples were separated on 10% (v/v) acrylamide gels by SDS-PAGE under denaturing and reducing conditions using the Mini-Protean$^®$ Tetra Gel System (Bio-Rad, China). The resolving gel contained 375 mM Tris-HCl pH 8.8, 10% (v/v) acrylamide (30% acrylamide: Bis-acrylamide (29:1) (Beyotime, China), 0.1% (w/v) SDS, 0.1% (w/v) ammonium persulphate (APS) and 0.1% N,N,N´,N´-Tetramethylethylenediamine (TEMED). The stacking gel consisted of 187.5 mM Tris-HCl pH 6.8, 4.95% (v/v) acrylamide, 0.1% (w/v) SDS, 0.1% (w/v) Ammonium persulfate (APS) and 0.1% TEMED. The SDS-PAGE running buffer consisted of 25 mM Tris-base, 192 mM glycine and 1% (w/v) SDS. PageRuler™ Prestained Protein Ladder (10-180 kDa) (ThermoFisher Scientific, China) was used as protein marker. Gel was run at 100 V for approximately 90 min or, until the visible pre-stained protein marker was properly resolved. For direct visualization, SDS-PAGE gels were stained with Coomassie blue staining solution R250 (Beyotime, China) for 20 min at room temperature, on a mildly shaking platform. This was followed by destaining the gel overnight at 4°C with 40% (v/v) methanol and 10% (v/v) acetic acid, to get rid of excess dye from the non-binding regions.

2.2.10. Large scale expression and His-LCPΔTM fusion protein purification

A single colony from a freshly transformed *lcp*ΔTM in BL21 (DE3) was cultured overnight in 500 mL of LB broth supplemented with ampicillin. 100 mL of the overnight culture was inoculated in each of four flasks with 1 L of fresh LB broth and ampicillin, and cultured until OD_{600} reached 0.6, followed by inducing with 0.2 mM IPTG. The induced culture was further grown overnight, at 16°C and shaking at 200 rpm and the overexpressed culture was centrifuged at 6000 g for 20 min at 4°C, followed by resuspension of the pellets in 120 mL of ice cold lysis buffer (50mM Tris-HCl pH 8.0, 500 mM NaCl and 10% glycerol) supplemented with 0.1 mM DNaseI (40 mM Tris-HCl pH 8.0, 10 mM $MgSO_4$, 1 mM $CaCl_2$) (Promega) and 4 mL of 1X EDTA-free protease inhibitor cocktail (BBI Life Sciences). The cells were disrupted by sonicating the lysates on ice for 3 min at amplitude: 30, pulse-on: 5 sec and pulse-off: 10 sec (QSonica, China). The lysed cells were then centrifuged at 40,000 g for 45 min at 4°C. The supernatant containing the protein of interest was filtered through a 0.22 μm filter (Millipore) and subjected to affinity chromatography using (Nickel-nitrilotriacetic acid) Ni-NTA agarose beads (Bio-Rad, China). Subsequently, 2 mL of the 50% slurry of beads were first washed with molecular biology grade water and equilibrated with 20 mL elution buffer (50 mM Tris-HCl pH 8.0, 250 mM NaCl, 150 mM immidazole and 5% glycerol). The column was equilibrated with 20 mL lysis buffer. The filtered protein lysates were run through the column

and the flow through collected in 50 mL falcon tubes. The column was washed thrice with 10 mL of lysis buffer each time, with the final wash in lysis buffer containing 0.15 mM n-dodecyl β-D-maltoside (BBI Life Sciences, China) (Grzegorzewicz *et al.*, 2016). The protein was then eluted with 10 mL of elution buffer and concentrated to half the volume using a 10 kDa cut-off Ultra-15 Centrifugal Filter Unit (Amicon®, China). The 5 mL concentrated protein was further purified using a HiLoad™ 16/600 Superdex™ 200 pg Fast Protein Liquid Chromatography (FPLC) column coupled to the ÄKTAprime plus design system (GE Healthcare Life Sciences, China). The eluted fractions containing the protein of interest were determined by the chromatogram at 280 nm and confirmed by SDS-PAGE (Section 2.2.9). Then, the eluted fractions containing the desired protein were pooled and concentrated prior to storing at -80°C.

2.2.11. Enzymatic activity of LCPΔTM proteins

The pyrophosphatase activity of the truncated LCP proteins was determined using Geranyl pyrophosphate (GPP) as a substrate. The inorganic phosphate released was determined by P$_i$per™ phosphate assay kit (Molecular Probes) (Grzegorzewicz *et al.*, 2016) (Fig 2.3). The LCPΔTM reaction mixture contained 50 mM Tris-HCl pH 8.0, 20 mM MgCl$_2$, 270 µM GPP and 1 µM of purified LCPΔTM proteins and incubated at 37°C for 16 hours, in a 96-well plate. Sterilized Tris buffer, MgCl$_2$ and double-distilled water was used for this purpose. Inorganic pyrophosphatase from yeast (NEB) at a final concentration of 0.002 U was used as a positive control, and incubated at 25°C for equal incubation time periods as that of LCPΔTM proteins. For negative control, gel filtrated 6XHis-tagged prolyl oligopeptidase (POP) was used (a kind gift from Dr. David Ruiz Carrillo), which is a cytosolic serine peptidase involved in the maturation and degradation of peptide hormones and neuropeptides isolated from *Msmeg*. After the 16 hour incubation period, phosphate standards ranging from 0-250 µM were prepared in duplicate for calculating the calibration curve. The Amplex detection reagent was then prepared as per the manufacturer's instructions (Fig 2.3), protected from light, and added to all wells containing the samples and standards. Since this detection assay is continuous, the absorbance was measured at 560 nm at several time points (0 min, 30 min, 1 hour and 2 hour) after addition of the Amplex detection reagent (see Fig 2.9). Such continuous assays are particularly helpful to study enzyme kinetics. However in this study, only a single concentration of the GPP substrate was used [as in (Grzegorzewicz *et al.*, 2016)] for the sole purpose of identifying the pyrophosphatase activity exhibited by all the four LCP proteins in *Msmeg*. Future studies could determine the enzyme kinetics of these proteins by using different concentrations of the substrate, as well as different incubation times. As indicated in the manufacturer's manual of the detection assay, the ideal time of incubation was required to

be established, and T=2h was shown as the optimal time point to determine the Pi values in our hands, since longer incubation times gave a drop in values (data not shown). Hence in our study, we used T=2h to estimate the pyrophosphatase activity of all the *Msmeg* LCP proteins.
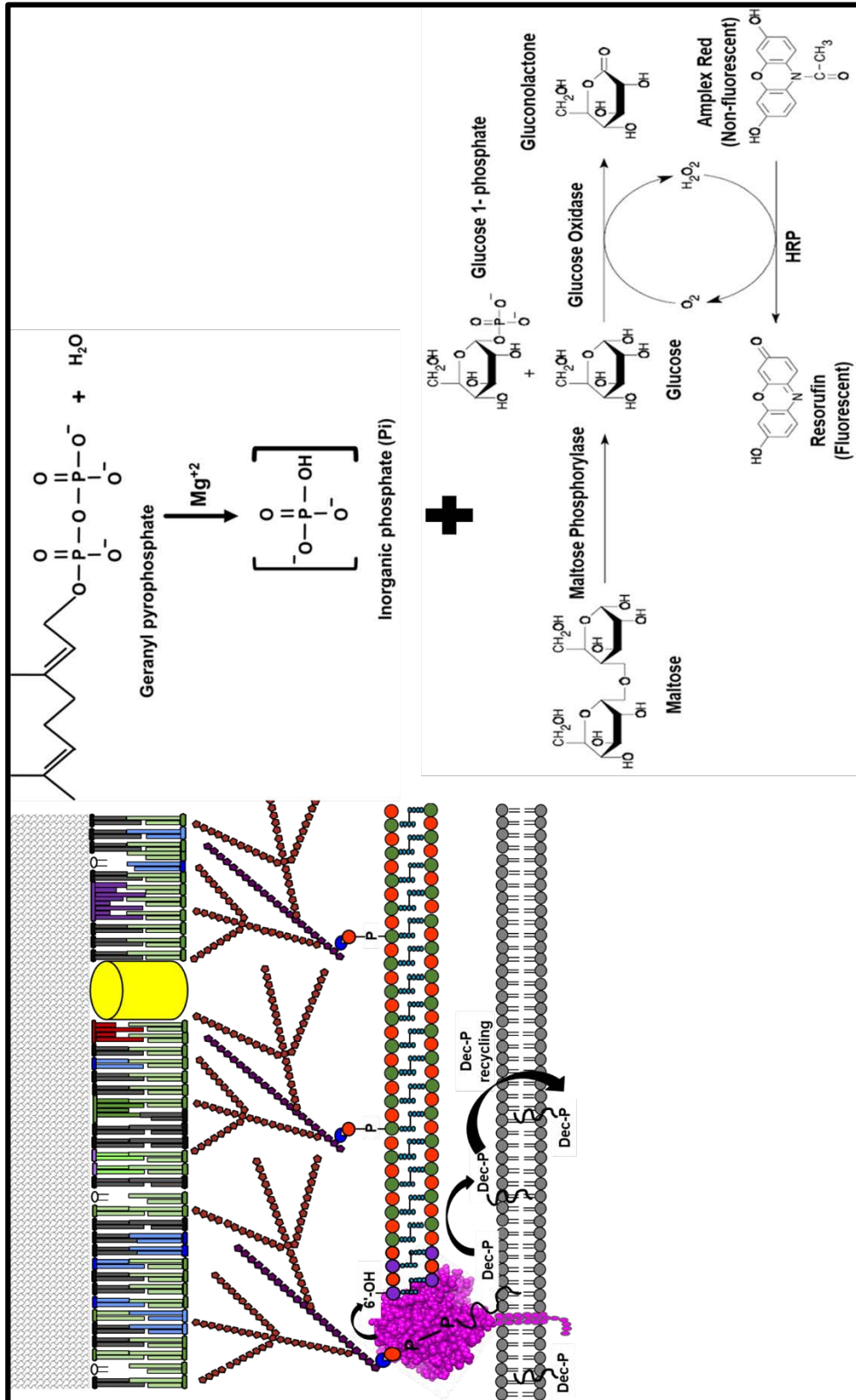
Fig 2.3. Schematic representation to determine pyrophosphatase activity in recombinant LCPΔTM proteins. The decaprenyl-1-monophosphate, an inorganic phosphate (P_i) is released after LCP proteins associate AG to the 6'OH of MurNAc present in PG, *in vivo*. This P_i can be measured by the $P_i$per™ Phosphate Assay Kit (Molecular Probes). In this study, the kit was used to determine the inorganic phosphate released with the cleavage of GPP by the *Msmeg* LCPΔTM proteins. The reaction is Mg^{+2} dependent. In the P_iPer Phosphate Assay, maltose phosphorylase converts maltose (in the presence of P_i) to glucose-1-phosphate and glucose. Then, glucose oxidase converts the glucose to gluconolactone and H_2O_2. Finally, with horseradish peroxidase (HRP) as a catalyst, the H_2O_2 reacts with the Amplex® Red reagent (10-acetyl-3,7-dihydroxyphenoxazine) to generate resorufin, which is measured spectrophotometrically at 560 nm. The resulting increase in absorption is proportional to the amount of P_i in the sample.

40

2.3. Results and discussion

2.3.1. Phylogenetic analysis of LCP protein family

The LCP family of proteins has been previously described to have four homologues in *Msmeg* (Harrison *et al.*, 2016). The four *Msmeg* homologues viz., MSMEG_0107, MSMEG_1824, MSMEG_5775 and MSMEG_6421 correspond to the four homologues of *Mtb* (Table 1.3 and Fig 1.9). In this study, the focus was on the domains of the *Msmeg* LCP protein viz., LCP and the Lytr_C domain, as these domains are highly conserved and have been concurred to possess catalytic activity of the LCP protein (Kawai *et al.*, 2011, Baumgart *et al.*, 2016, Grzegorzewicz *et al.*, 2016, Harrison *et al.*, 2016).

In this study, the amino acid sequences of the LCP and LytR_C domains across various mycobacterial and non-mycobacterial species were retrieved from UNIPROT and aligned. The non-mycobacterial organisms were selected from a previous phylogenetic study on the LCP family of proteins (Hubscher *et al.*, 2008). Almost 75% of LCP domain was found to be conserved in *Msmeg*, and six residues (S103, P105, R106, D107, R225 and R227 of MSMEG_1824 full length protein sequence) were completely conserved in all the species (Fig 2.4), which is in congruence with the findings in *S. pneumoniae* (Kawai *et al.*, 2011) and *Mtb* (Grzegorzewicz *et al.*, 2016). The three positively-charged arginine residues amongst these conserved residues have been found to interact with the pyrophosphate head group of the octaprenyl pyrophosphate in *S. pneumoniae* (Kawai *et al.*, 2011). Apart from this, absence of three charged residues that interact with the pyrophosphate head group of the polyprenyl-pyrophosphate is also observed in Rv3840, viz., D218, R221 and Q225 of the sequence in the alignment (Fig 2.4).

The LCP and the LytR_C domains have been found to occur separately as well as in combination, suggesting the presence of LytR_C is not an absolute requirement within LCP protein family members (Baumgart *et al.*, 2016). For instance, in MSMEG_6421, only the LCP domain is found, exactly like in its *Mtb* orthologue, Rv3480. The other *Msmeg* and *Mtb* LCP proteins have both the domains. While the LCP domain has been found to be abundantly distributed amongst Firmicutes, the LytR_C domain has been found widely in Actinobacteria (Baumgart *et al.*, 2016) suggesting the LytR_C domain to play an important role in *Msmeg* and other mycobacterial species. In a recent study in *Mtb*, Rv2700 that contains only LytR_C domain but not LCP domain, was found to exhibit decreased growth rate, increased sensitivity to PG-targeting antibiotics and increased envelope permeability thus proposed to be involved in cell integrity and virulence (Ballister *et al.*, 2019). Given that both the domains together

might be important for LCP function in *C. glutamicum* (Baumgart *et al.*, 2016), five highly conserved residues were found in the LytR_C domain across all species (Fig 2.5). The C-terminus CXN motif of the LCP protein (Baumgart *et al.*, 2016) was found to be conserved in MSMEG_1824, however in MSMEG_0107, the asparagine residue was found to be replaced by a lysine residue ($N_{533} \rightarrow K_{533}$ of the full-length amino acid sequence of MSMEG_0107). The other two *Msmeg lcp* genes were found to be devoid of this motif. Though the loss of the C-terminal are shown to have no effect on the enzymatic activity of the protein in *C. glutamicum*, however phenotypic changes have been associated with it (Baumgart *et al.*, 2016). On this note, further investigations are a definite necessity to determine the function of the LytR_C domain and the C-terminus of LCP proteins in *Msmeg*.

2.3.2. Structural prediction of *Msmeg* LCP proteins

Homology based modelling for all the four LCP proteins of *Msmeg* is shown in Fig 2.6. Only a part covering the LCP domain was retrieved from the reference model, in all cases to build the models in *Msmeg*. The six residues that are conserved in the LCP domain across all species has been mapped on to the three-dimensional structure. Though higher QMEAN scores are generally considered for reliability during homology modelling, however, not all LCP proteins of *Msmeg* gave a QMEAN score. The focus was to map the LCP domain and the six residues that are highly conserved in the LCP domain, in these 3D predictions. These two intended features were conserved for all the LCP proteins in *Msmeg* hence, the GMQE scores were considered for LCP protein structure modelling. The MSMEG_0107 model was built based on a reference model of *S. pneumoniae* serotype 2 strain D39 LCP family protein Cps2A containing a bound octaprenyl monophosphate lipid, with PDB ID 4DE8 and GMQE score 0.3. MSMEG_1824 was homologous to Cps2A (PDB ID 2xxq) from *S. pneumoniae* serotype 2 strain D39 with a GMQE score 0.34. The LCP domain in MSMEG_5775 was mapped using the LCP family protein Cps2A of *S. pneumoniae* with PDB ID 3TEP and GMQE score 0.18 whereas, the structural modelling MSMEG_6421 was done using PDB ID 3PE5 from *Clostridium leptum* strain DSM 753 with a GMQE score 0.23. It was interesting to notice that in MSMEG_6421, the residues R503 and R505 are distinct and further away from rest of the residues, in comparison to other LCP proteins, where all the six residues are adjacent to each other. We assume that this difference could be a reason for MSMEG_6421 to display lower pyrophosphatase activity (Section 2.3.4), as there might be less stable interaction with the pyrophosphate head group.

Fig 2.4. Global alignment of the LCP domain of LCP proteins. The six conserved residues in the LCP domain of the LCP proteins are highlighted in yellow. Amino acid numbering in this figure starts at the first residue of the LCP domain, which is identical to residue 83 in MSMEG_1824 shown in Fig 1.10. The mutated conserved arginine residue in mut-MSMEG_1824-R225AΔTM is indicated with a red arrow and represents position 225 of the full length amino acid sequence of MSMEG_1824.

LytR_C domain

```
MSMEG 0107 (Msmeg)     G A T L D V I N G S T Y T G L A A K L Q T T F A A D N F T E G Q I G D A E A L E P - A T T I D Y G P G A D A A A H S L A [60]
MSMEG 1824 (Msmeg)     K T T V E V L N G T D V N G L A A S V S Q V L T Q K G F T P G D T G N H E D G P V A G S Q V L A A T S D D L G A Q E V A [60]
MSMEG 5775 (Msmeg)     Q V T V K V S N S T G Q S G L A A T A A G E L E Q H G F N V T T P - D D Y P G P L D A T T V F F S P - - - - G N E E A A [60]
Rv3484 (Mtub)          S T V V D V V N A G S I S G L A S Q V S G A L L K R G Y T A G Q V R D R E S G D P F T T A I E Y G A - - - - G A E T D A [60]
Rv3267 (Mtub)          K T T A N V V N D T D I N G L A A A V S K V L S S K G F T T G S V G N N D G D H V P G S Q V R A A K A D D L G A Q Q V A [60]
Rv0822c (Mtub)         E V T V Q V S N S T G Q A G L A T T A T D Q L K R N G F N V M A P - D D Y P S S L L A T T V F F S P - - - - G N E Q A A [60]
MMAR 4966 (M.marinum)  S T V V D V V N A G S I S G L A S Q V S N A L R K R G Y T T D Q V R D R K S G D P A A T T I K Y G P - - - - G A Q A D A [60]
MMAR 1274 (M.marinum)  K T T A D V V N D T D I N G L A A A V S D V L S A K G F T T G S V G N N E A G H V K L S Q V R A A T T D D L G A Q Q V A [60]
MMAR 4858 (M.marinum)  E V T V Q V S N G T G T T G L A A T A A S Q L E R N G F N V M A P - D D Y P N S L Q T T T V L F A P - - - - G N E Q A A [60]
ML2247 (M.leprae)      S T V V D V I N A G R L T G L A S E V F Q T L K K H G Y T A G Q V R D R S S G E P T T T S I Q Y G A - - - - G A D T D A [60]
ML0750 (M.leprae)      K I T A N V A N D T D I N G L A A A V S D I L T T K G F I A G S V A N N Q G S H V K S S Q V R A A K S D D L G A Q K V S [60]
MB3514 (M.bovis)       S T V V D V V N A G S I S G L A S Q V S G A L L K R G Y T A G Q V R D R E S G D P F T T A I E Y G A - - - - G A E T D A [60]
MB3295 (M.bovis)       K T T A N V V N D T D I N G L A A A V S K V L S S K G F T T G S V G N N D G D H V P G S Q V R A A K A D D L G A Q Q V A [60]
MB0845c (M.bovis)      E V T V Q V S N S T G Q A G L A T T A T D Q L K R N G F N V M A P - D D Y P S S L L A T T V F F S P - - - - G N E Q A A [60]
Cgl0740 (C.glutamicum) E V E V H V L N A S Y V E G L A N G I A A Q L Q E L G Y S I A E T G N A A E G L Y Y E S Q I L A A E E D S A K A L A I S [60]
lytR1 (R. fascians)    T V T V D V A N D S G I E G L A N G V A G S L T A L G Y V G G T I G N Y E G S S I S T S R I Q T A D V D S D A A R A V S [60]

MSMEG 0107 (Msmeg)     Y L G A D M F T S T T S G A S E S E S Y G P S E T A E L T P I T T V A A T A T G T Y T P A P T D L S Q M T A S G I P C V K [181]
MSMEG 1824 (Msmeg)     - - - - - - - - S G L D G N P A L S S V D P A S S S A S G Y A D - - G S S E D E T P P P P P S P I L T A G A N D P Q C V N [181]
MSMEG 5775 (Msmeg)     - - - - - - - - D L S V T N A A D A T C E - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - [181]
Rv3484 (Mtub)          - - - - - - - - P D E A T A A A T S T - - - - - - - E T S T Y P L Y G G G - - T T T D P T P D Q G A P I D G G G V P C V N [181]
Rv3267 (Mtub)          - - - - - - - - S G L G G S D P N G V V S P A R A F N L G S - - - - - - - - A D D T T P P P S P I L T A G S D A P E C I N [181]
Rv0822c (Mtub)         - - - - - - - - D L T V T N A A D T T C E - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - [181]
MMAR 4966 (M.marinum)  - - - - - - - - P D D - S D D D T T T Y P K P N A Y - - - - - - - Y N G G D D S S S Y P A P D H G Q P I D G G G V P C V N [181]
MMAR 1274 (M.marinum)  - - - - - - - - S G L G G - - - - - A M A P A H A S N P G - - - - - - - - - Y S D T A P P P S P I L T A G S D Q P E C I N [181]
MMAR 4858 (M.marinum)  - - - - - - - - D L T V T N A A D T T C E - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - [181]
ML2247 (M.leprae)      - - - - - - - - D E D S T M D E T I T S Q - - - - K V G T Y P P Y S Y N - - T S P Y P T P D Q G K P I D G G G I P C V N [181]
ML0750 (M.leprae)      - - - - - - - - S G L F G - - - N G S I A S D Y A S N P G A S - - - - - - - A D P N I P V P S P I L T A G S D Q P E C I N [181]
MB3514 (M.bovis)       - - - - - - - - P D E A T A A A T S T - - - - - - - E T S T Y P L Y G G G - - T T T D P T P D Q G A P I D G G G V P C V N [181]
MB3295 (M.bovis)       - - - - - - - - S G L G G S D P N G V V S P A R A F N L G S - - - - - - - - A D D T T P P P S P I L T A G S D A P E C I N [181]
MB0845c (M.bovis)      - - - - - - - - D L T V T N A A D T T C E - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - [181]
Cgl0740 (C.glutamicum) - - - - - - - - A E A N A V T S S T V G Q P G A D V G E P I E - - - - - - - - - - - - - - - - - - S P E F D A G G D G P R C V N [181]
lytR1 (R. fascians)    - - - - - - - - S G Q A A T A E T V A P E A G - - - - - - - - - - - - - - - - - - - T A V P E E P A A P I D A G S T G P M C V N [181]
```

Fig 2.5. Global alignment showing the fully conserved residues in the LytR_C domain of LCP proteins. The fully conserved residues in the LytR_C domain are highlighted in yellow. The partially conserved C-terminus motif is highlighted in blue.

Fig 2.6. A-D. Three-dimensional structural modelling (cartoon style) of the LCP homologues in *Msmeg*. The structures were generated in PYMOL v. 2.3 after homology modelling with LCP proteins across different organisms, using SWISS-MODEL. The reference PDB IDs used for homology modelling of MSMEG_0107, MSMEG_1824, MSMEG_5775 and MSMEG_6421 were 4DE8, 2xxq, 3TEP and 3PE5 respectively. The LCP domain is highlighted in red. The six conserved residues across all LCP proteins are mapped in each of the four LCP homologues in *Msmeg*, indicated with different colors. The arginine residue (blue) is the site of mutation in mut-MSMEG_1824-R225AΔTM.

45

2.3.3. Overexpression and purification of *lcp*ΔTM genes in *Msmeg*

The *Msmeg lcp*ΔTM genes were cloned successfully into appropriate vectors and the presence of insert in the recombinants were confirmed by agarose gel electrophoresis (Fig 2.7). The full-length nucleotide sequences of the *lcp* genes, viz., *MSMEG_0107*, *MSMEG_1824*, *MSMEG_5775* and *MSMEG_6421* are 1602 bp, 1482 bp, 2169 bp and 1872 bp, respectively. A single transmembrane (Fig 1.10) helical structure in each of the four LCP proteins was predicted in all the *Msmeg lcp* homologues. While *MSMEG_0107* and *MSMEG_1824* have their transmembrane domains at their N-terminal, *MSMEG_5775* and *MSMEG_6421* have their transmembrane domains approximately in the middle of the protein. Interestingly, in contrast to *Msmeg*, the orthologue of *MSMEG_6421* in *Mtb* i.e., *Rv3840* doesn't possess a transmembrane domain (Baumgart *et al.*, 2016). Since the presence of a transmembrane helix decreases the solubility of proteins due to its hydrophobic attributes, it becomes quite difficult to extract such proteins after bacterial lysis. On the other hand, the N-terminal domain of LcpA protein in *C. glutamicum* (orthologue of MSMEG_1824) does not seem to carry a pyrophosphatase function (Baumgart *et al.*, 2016), which implies that the sequence upstream of the transmembrane domain are not required to demonstrate pyrophosphatase activity in LCP proteins. Hence, the transmembrane domain and N-terminus were excised for subsequent overexpression of all *Msmeg* LCP proteins, using a similar approach as described previously (Grzegorzewicz *et al.*, 2016). The resulting transmembrane-less *lcp* genes, were thus resized to 1449 bp, 1407 bp, 1503 bp and 876 bp, respectively. After introducing the N-terminal 6XHis tag and HRV3C cleavage site to the transmembrane-less protein, the full length transmembrane-less proteins were termed LCPΔTM proteins i.e., MSMEG_0107ΔTM, MSMEG_1824ΔTM, MSMEG_5775ΔTM and MSMEG_6421ΔTM and were 1503 bp, 1458 bp, 1521 bp and 894 bp respectively (Fig 2.7). The recombinant proteins were further fused to an N-terminal GST-tag, which has a size of 26 kDa, thus the recombinant proteins produced were predicted to be 81 kDa, 79 kDa, 82 kDa and 59 kDa, respectively, which was confirmed by SDS-PAGE after gel filtration (Fig 2.8). Initially the HRV-3C cleavage site was cleaved to obtain LCPΔTM, and subsequent pyrophosphatase activity was determined however, these cleaved proteins were found to degrade rapidly even after storing at -80°C. The fused proteins were then opted and it was found that, not only they produced enzymatic activity but, could also be stored longer. Hence, the fused GST-tagged LCPΔTM were used for determining the enzymatic activity in all experiments.

It is also interesting to note the oligomerization of MSMEG_0107ΔTM, MSMEG_1824ΔTM and MSMEG_5775ΔTM, a feature shown for LcpA in *C. glutamicum* as well (Baumgart *et al.*,

2016). This phenomenon is identified by the presence of multiple peaks during gel filtration chromatography (Fig 2.8). However, MSMEG_6421ΔTM has a single well-defined peak and apparently lacks this oligomerization ability. In *C. glutamicum*, ΔTMCps2A for instance, has been reported to be monomeric (Baumgart *et al.*, 2016). Considering the fact that Cps2A lacks a LytR_C domain, it is possible that this domain is involved in dimerization or oligomerization.
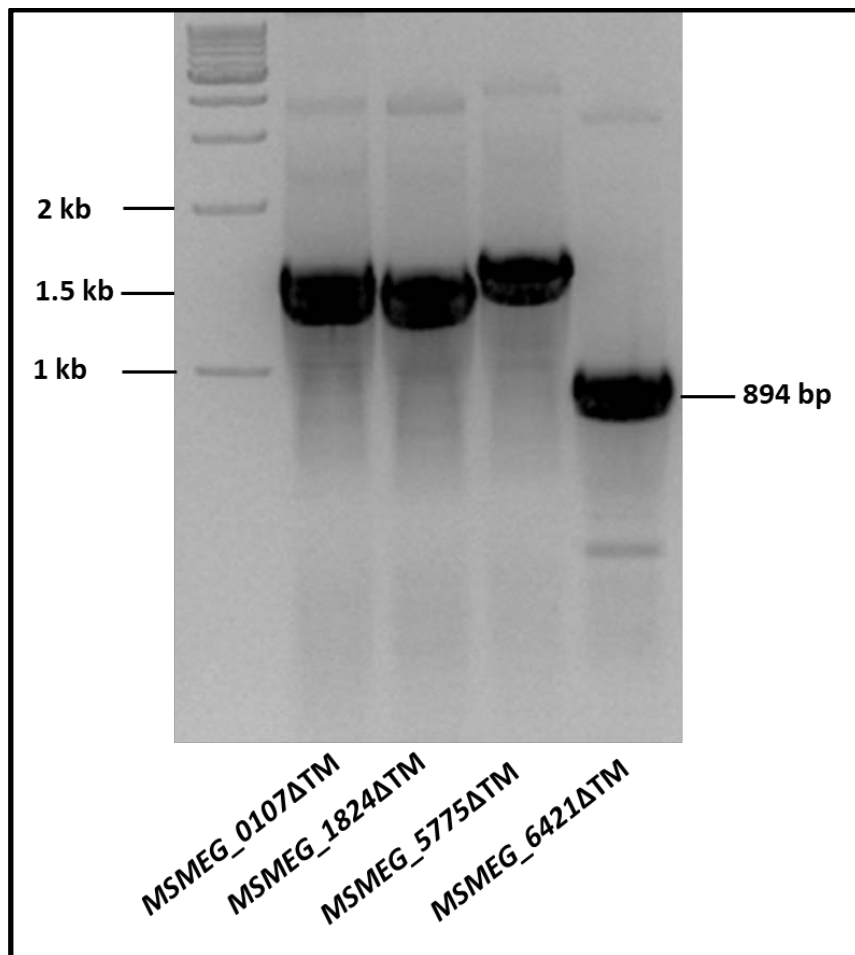
Fig 2.7. Agarose gel electrophoresis of PCR-amplified recombinant *Msmeg lcp*ΔTM plasmids. The cloned genes were confirmed using PCR prior to sequencing. Lane 1: 1 kb DNA ladder. Lane 2: *MSMEG_0107*ΔTM - 1503 bp, Lane 3: *MSMEG_1824*ΔTM -1458 bp, Lane 4: *MSMEG_5775*ΔTM -1521 bp and Lane 5: *MSMEG 6421*ΔTM - 894 bp

A *MSMEG_0107ΔTM*

B *MSMEG_1824ΔTM*

C mut-*MSMEG_1824*ΔTM

Pool fractions → 36 37 38 39

79kDa —

56.20 min
16.14 mAu

D *MSMEG_5775*ΔTM

Pool fractions → 48 49 50 51 52 53

82kDa —

74.82 min
6.41 mAu
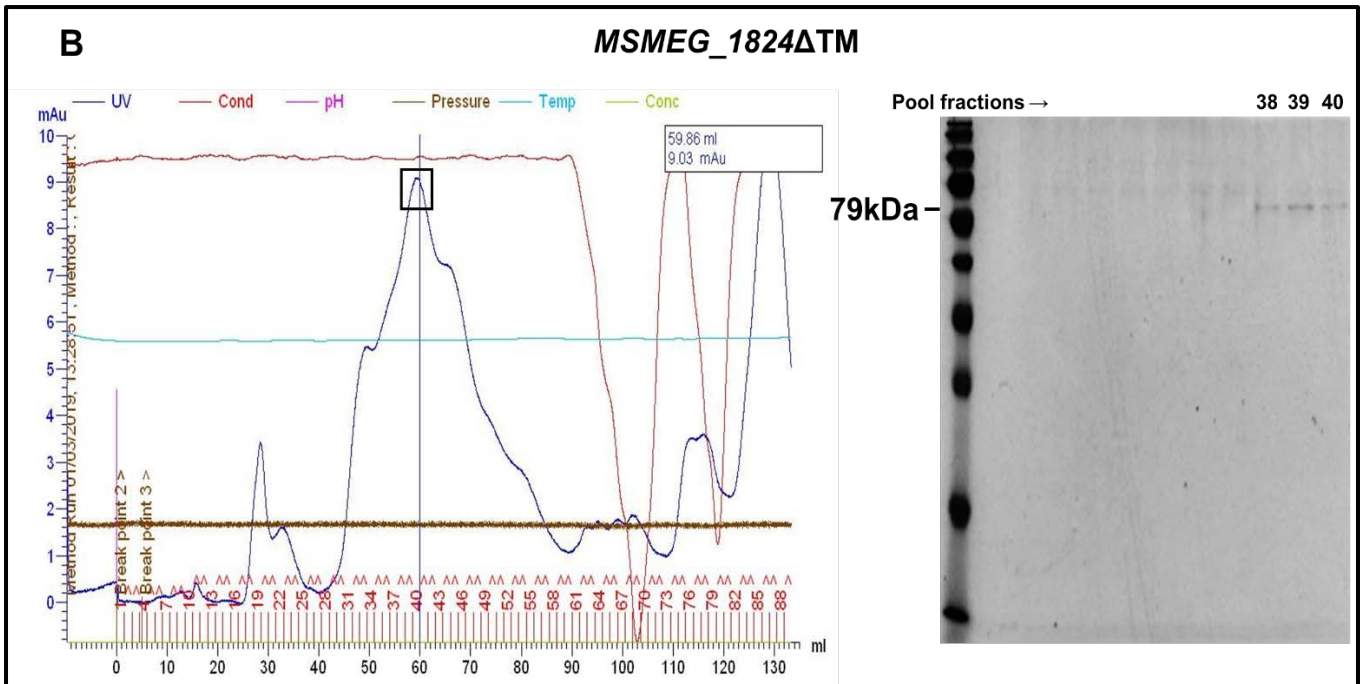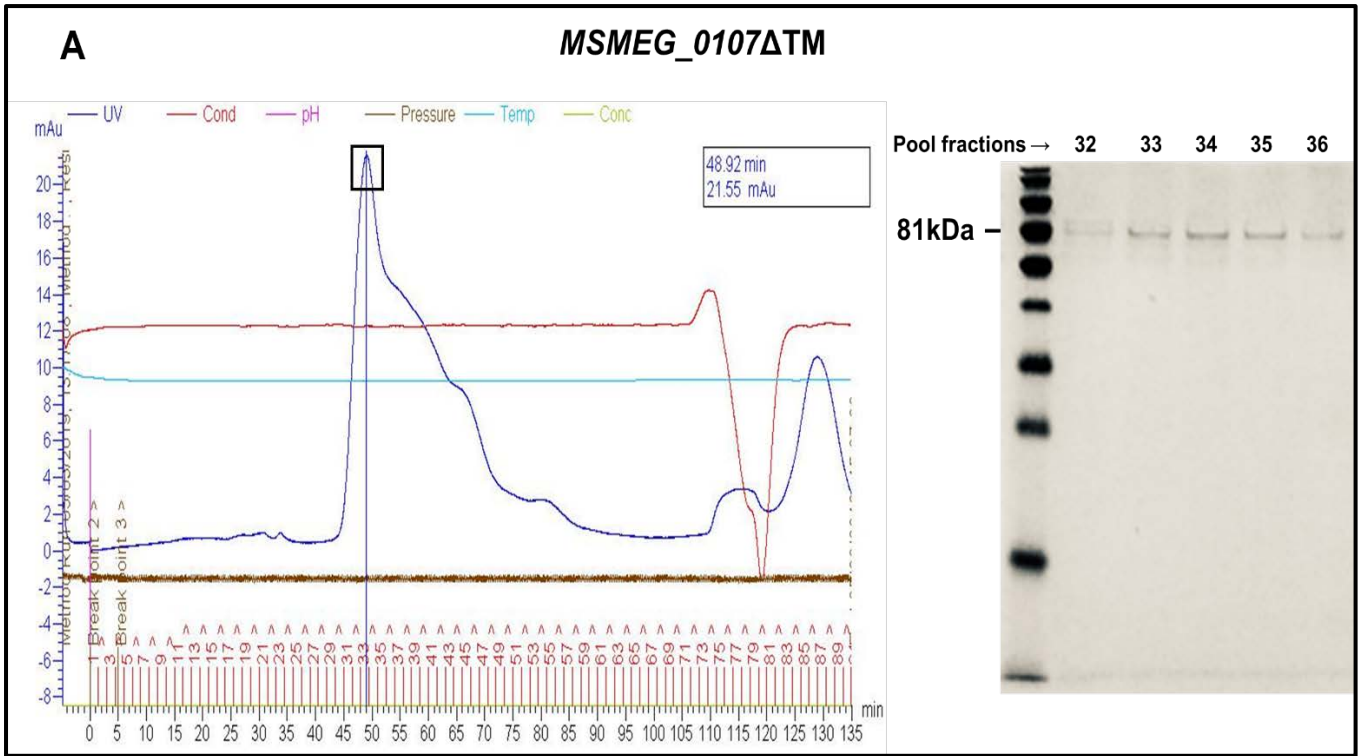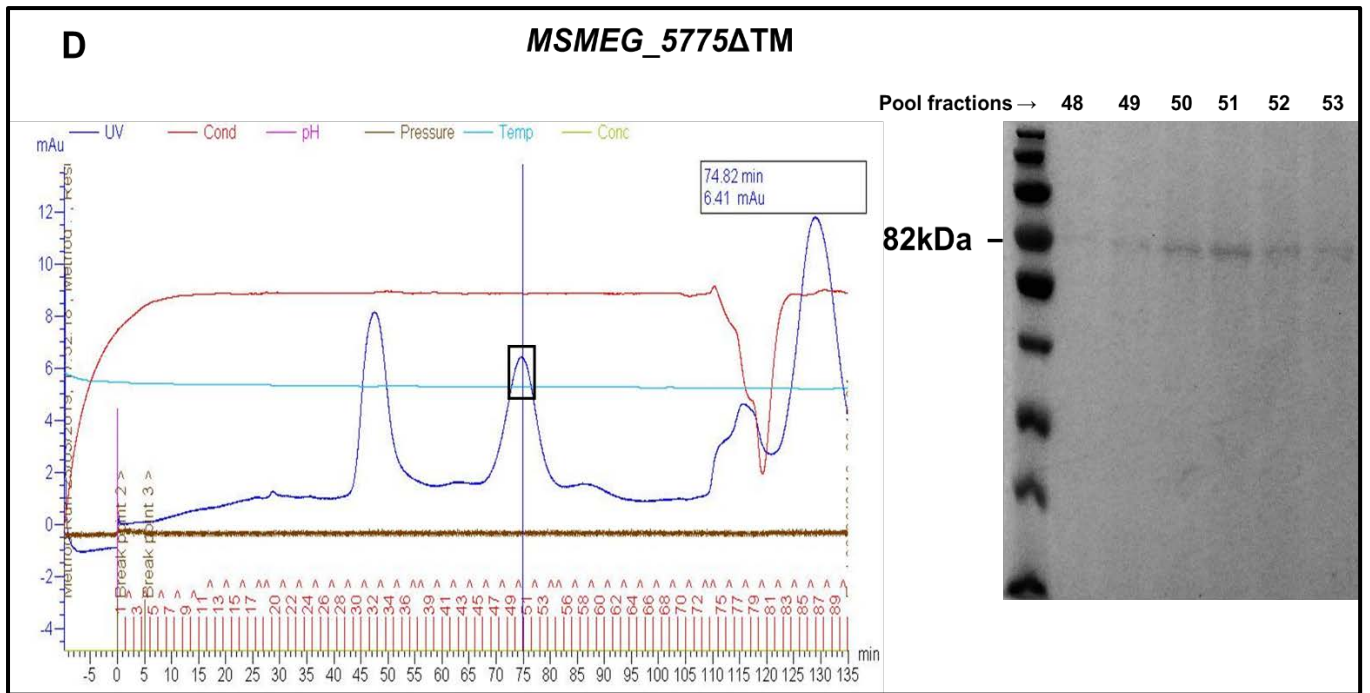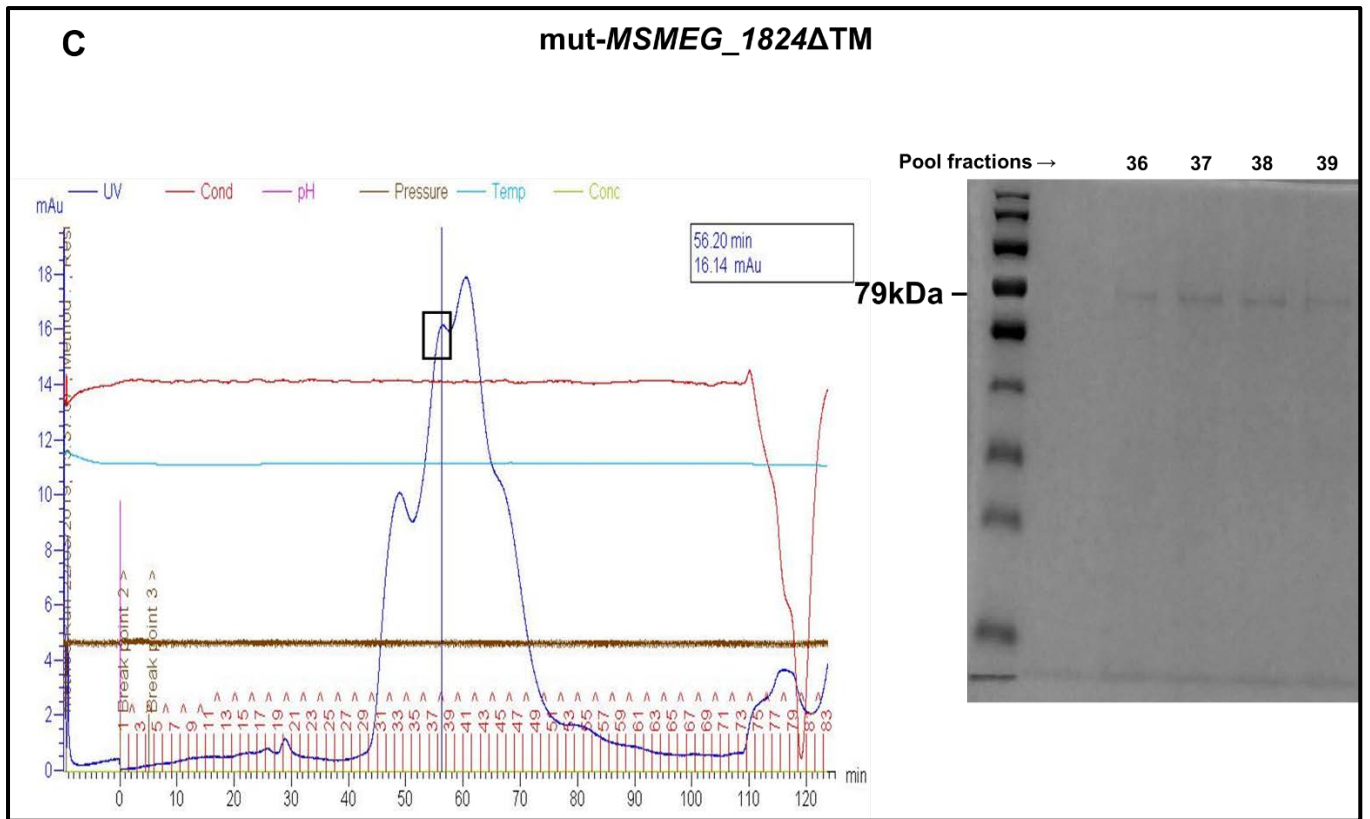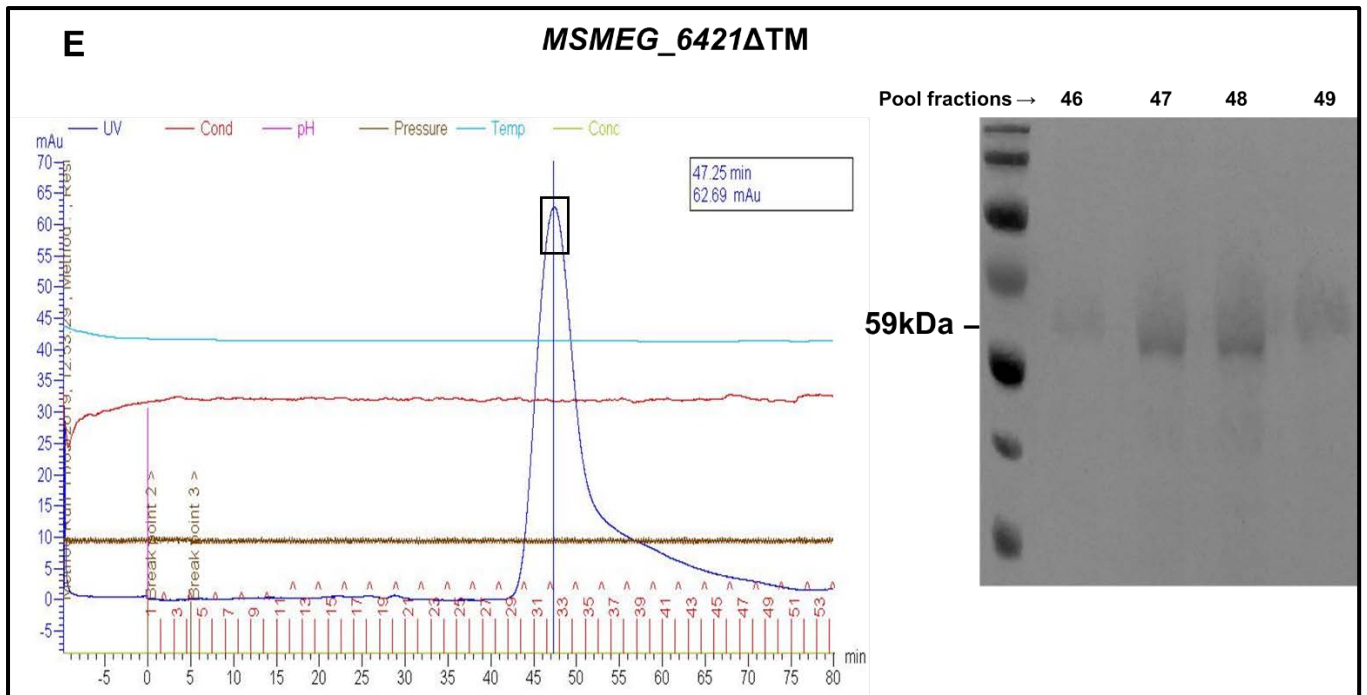
Fig 2.8. A-E. Fast Protein Liquid Chromatography (FPLC) of the GST-tagged LCPΔTM proteins in *Msmeg.* Ni-NTA columns were first used to purify the crude protein lysates by affinity chromatography. 10 mL of the protein eluent of each sample was further concentrated to half the volume, using a 10 kDa cut-off Ultra-15 Centrifugal Filter Unit. The 5 mL concentrated protein was further purified using a HiLoad™ 16/600 Superdex™ 200 pg Fast Protein Liquid Chromatography (FPLC) column coupled to the ÄKTAprime plus design system. The FPLC-eluted fractions containing the protein of interest were determined by the chromatogram at 280 nm and confirmed by SDS-PAGE. The square block on the peak represents the fractions used for SDS-PAGE that were subsequently pooled for use in the pyrophosphatase assay. Fractions 32-36, 38-40, 36-39, 48-53 and 46-49 are seen in the SDS-PAGE of MSMEG_0107ΔTM, MSMEG_1824ΔTM, mut-MSMEG_1824ΔTM, MSMEG_5775bΔTM and MSMEG_6421ΔTM. The corresponding GST-tagged LCPΔTM proteins are shown along with their molecular weight.

2.3.4. LCP proteins in *Msmeg* show pyrophosphatase activity *in vitro*

Previous studies in *Mtb* (Grzegorzewicz *et al.*, 2016, Harrison *et al.*, 2016) have demonstrated that the LCP proteins display pyrophosphatase activity on the artificial substrate geranyl pyrophosphate (GPP). LCP proteins are believed to transfer decaprenyl phosphate-linked intermediates of glycopolymers to C-6 hydroxyl of MurNAc residues of PG. This process is known as phosphotransferase reaction which releases the decaprenyl phosphates by cleaving the pyrophosphate group. This enzyme activity of the LCP proteins is Mg^{2+} dependent and has been demonstrated previously in selected orthologues across different organisms (Kawai *et al.*, 2011, Baumgart *et al.*, 2016). The amino acids responsible for binding to the pyrophosphate functional group, octaprenyl phosphate chains and the magnesium ions (Kawai *et al.*, 2011, Grzegorzewicz *et al.*, 2016) were found to be conserved in the primary sequences of all the LCP proteins in *Msmeg* (Fig 2.4). In this enzymatic reaction, absorbance of the inorganic phosphate release was measured at 560 nm using the $P_i$per^{TM} phosphate assay kit as has been previously described (Grzegorzewicz *et al.*, 2016). The mechanism of the enzymatic reaction using this kit depends on the amount of free inorganic phosphates released after incubating the LCPΔTM proteins at 37°C for 16 hours with GPP, $MgCl_2$ and Tris-HCl buffer pH 8. The inorganic phosphate released during the reaction reacts with maltose to form glucose and glucose 1-phosphate in the presence of maltose phosphorylase. The glucose is converted to gluconolactone and hydrogen peroxide (HRP) in the presence of glucose oxidase. In the final step, the HRP as a catalyst, reacts with the non-fluorescent Amplex Red reagent to generate a highly fluorescent product, resorufin (Fig 2.3). The amount of fluorescence or absorption obtained is directly proportional to the amount of inorganic phosphate released. However, the time point to determine the enzymatic activity, after addition of the kit reagents, was not clear from the study by Grzegorzewicz and colleagues. According to our experiments, the absorbance values at OD_{560} to determine enzymatic activity was found to be negative at T=0h for all the sample proteins and controls. The Pi values gradually increased after T= 0h until T= 2h of adding the Amplex reagent mixture (Fig 2.9 a,b), after which the values started decreasing. This pattern was seen in the previous experiments, the data of which is not included in this study. Therefore, the time point T= 2h was considered the peak time for determining the pyrophosphatase activity at OD_{560}.

LCP proteins are known to have pyrophosphatase activity, which means that decaprenyl phosphate-linked glycopolymer molecules are transferred to C-6 hydroxyl of MurNAc residues within PG. This leads to the release of decaprenylphosphate. This release of inorganic phosphate by hydrolyzing the geranyl pyrophosphate is measured spectrophotometrically, to determine the enzymatic activity of LCP proteins. However, biochemical analysis of

recombinant Lcp1 in *Mtb* has demonstrated the co-purification of decaprenyl-1-monophosphate (Harrison *et al.*, 2016). Hence, the estimation of inorganic phosphate released in this assay might be the net amount accumulated after GPP cleavage and co-purified decaprenyl-1-monophosphate. Overall, all the four LCP homologues in *Msmeg* was found to display pyrophosphatase activity in the range from 1.6-6.8 nmoles of released Pi/μM protein/minute (Fig 2.10), suggesting that all four proteins participate in releasing the newly bound AG from the Dec-P-P carrier lipid to form mature AG (Harrison *et al.*, 2016). The LCPΔTM proteins, MSMEG_0107ΔTM and MSMEG_5775ΔTM displayed about 1.4- and 2-fold higher pyrophosphatase activity on GPP than the MSMEG_1824ΔTM however, MSMEG_6421ΔTM showed about 2.2-fold lesser enzymatic activity than this essential LCP protein (Fig 2.10). In *Mtb*, Rv3484 (MSMEG_0107) showed the highest pyrophosphatase activity of the three LCP proteins tested (Grzegorzewicz *et al.*, 2016) whereas, in *C. glutamicum*, Cg0847/LcpA (MSMEG_1824) was shown to have pyrophosphatase activity (Baumgart *et al.*, 2016). These findings indicate that, perhaps different species use a specific set of these LCP proteins to aid the formation of mature AG after the release of premature AG from Dec-P-P carrier lipid. Therefore in this study, this might suggest that, MSMEG_0107 and MSMEG_5775 are involved in generating abundant mature AG compared to MSMEG_1824 or MSMEG_6421.

It is very interesting that the number of LCP homologues differ from organism to organism. For example, *Msmeg*, *Mtb* and *M. marinum* each have four LCP homologues in their genome whereas, *M. leprae* and *M. bovis* have three each. *C. glutamicum* has the least with only two LCP proteins in its genome. There could be two possibilities behind this condition. i) The organism might require several LCP homologues at the same moment to perform its function most optimally, or ii) one LCP protein might compensate the function of the others, during various environmental stresses.

A mutant of *MSMEG_1824*ΔTM was created by site-directed mutagenesis to detect a reduction in the pyrophosphatase activity. Here, the arginine amino acid that interacts with the pyrophosphate head group (represented by the red arrow in Fig 2.4) was substituted with an alanine residue i.e., R225A (Fig 2.2), corresponding to the R138A substitution in *C. glutamicum* (Baumgart *et al.*, 2016) and R362 in *B. subtilis*, where it was found to interact with the available oxygen from the phosphate to form a positively charged groove on the protein surface (Kawai *et al.*, 2011). This arginine amino acid is conserved across all LCP orthologues in Gram-positive organisms (Fig 2.4). The mutated *MSMEG_1824*ΔTM had a 1.6-fold lower enzymatic activity compared to the wild type 1824, however this difference was not significant (P=0.1). But our results are comparable to the lowered enzymatic activity levels seen in the R138A mutant in *C. glutamicum* in the presence of $MgCl_2$ and GPP (Baumgart *et al.*, 2016).

As described in *C. glutamicum*, this ariginine residue has a critical role in LCP function, as this mutant had a compromised growth (Baumgart *et al.*, 2016).

In order to confirm the magnesium-dependence of the LCP proteins, 0.1 M EDTA was added to chelate the Mg^{2+} ion. This inhibits the enzymatic activity of a cation-dependent protein, and this phenomenon is well observed in all the *Msmeg* LCP proteins in this study (Fig 2.10). The results on Mg^{2+}-binding affinity of LCP proteins in this study is also in congruence with previous studies (Kawai *et al.*, 2011, Grzegorzewicz *et al.*, 2016). This suggests that the LCP proteins in both *Mtb* and *Msmeg* are Mg^{+2} dependent. However, the effect of EDTA in *C. glutamicum* LcpA exhibited an opposite effect though, and the enzymatic activity in all the wild type and mutants increased, after the addition of EDTA (Baumgart *et al.*, 2016). Apart from this, several attempts were also made to crystallize the GST-tagged *Msmeg* LCPΔTM proteins, but they proved unsuccessful.
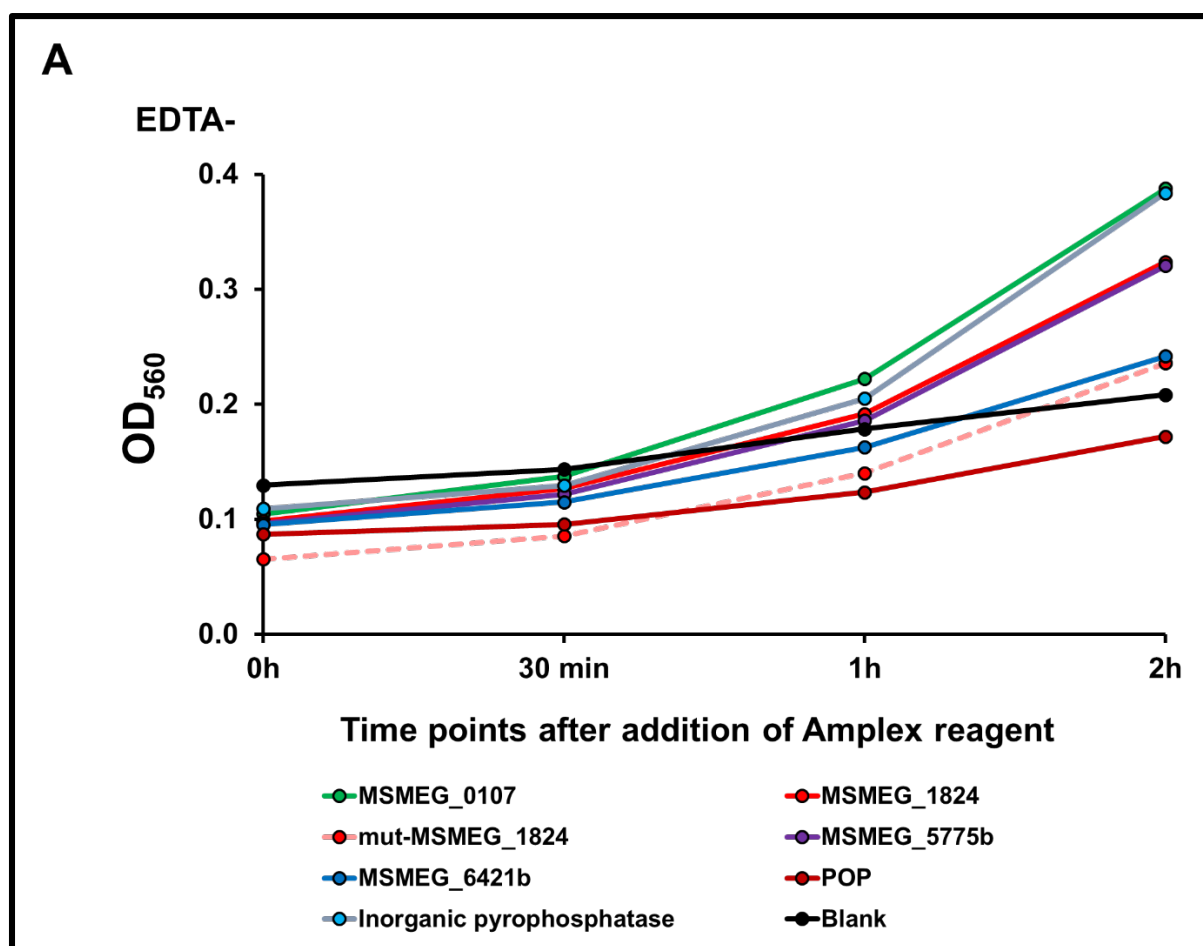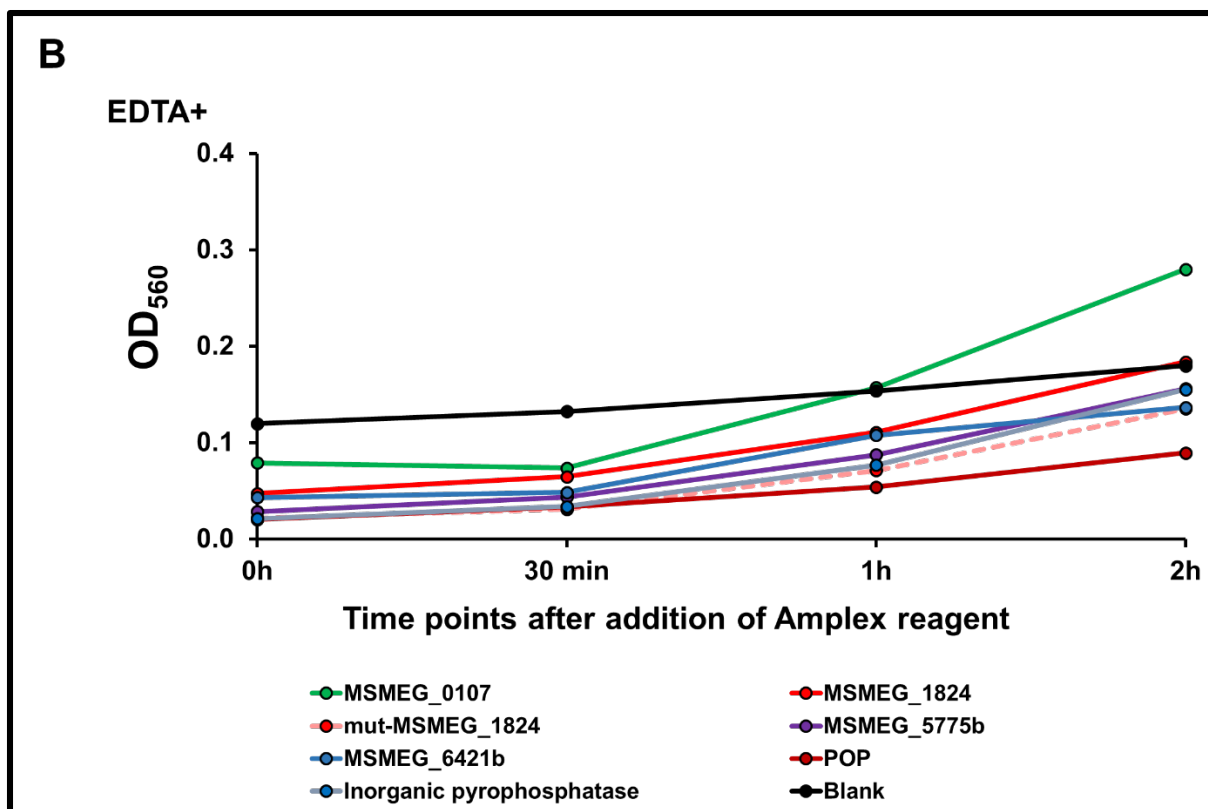
Fig 2.9. Detection of inorganic phosphate (Pi) at the corresponding OD$_{560}$ using continuous Amplex detection assay. OD$_{560}$ of the produced resorufin corresponding to the Pi released from the discontinuous pyrophosphatase assay is shown in the **A)** absence and **B)** presence of EDTA at various time points after addition of the Amplex detection reagent. The detection reagent was added to the 16h-incubated pyrophosphatase reaction mixture and the OD$_{560}$ was measured continuously at four time points (T=0h, 30 min, 1h and 2h) in a microplate reader. There is a gradual increase in the OD$_{560}$ from T=30 min until T=2h for the samples and controls, after which the OD$_{560}$ value drops (data not shown). Hence in our study, T=2h is used to detect the maximum Pi in every condition as shown in Fig 2.10. The data shown here is an average of three independent experiments.
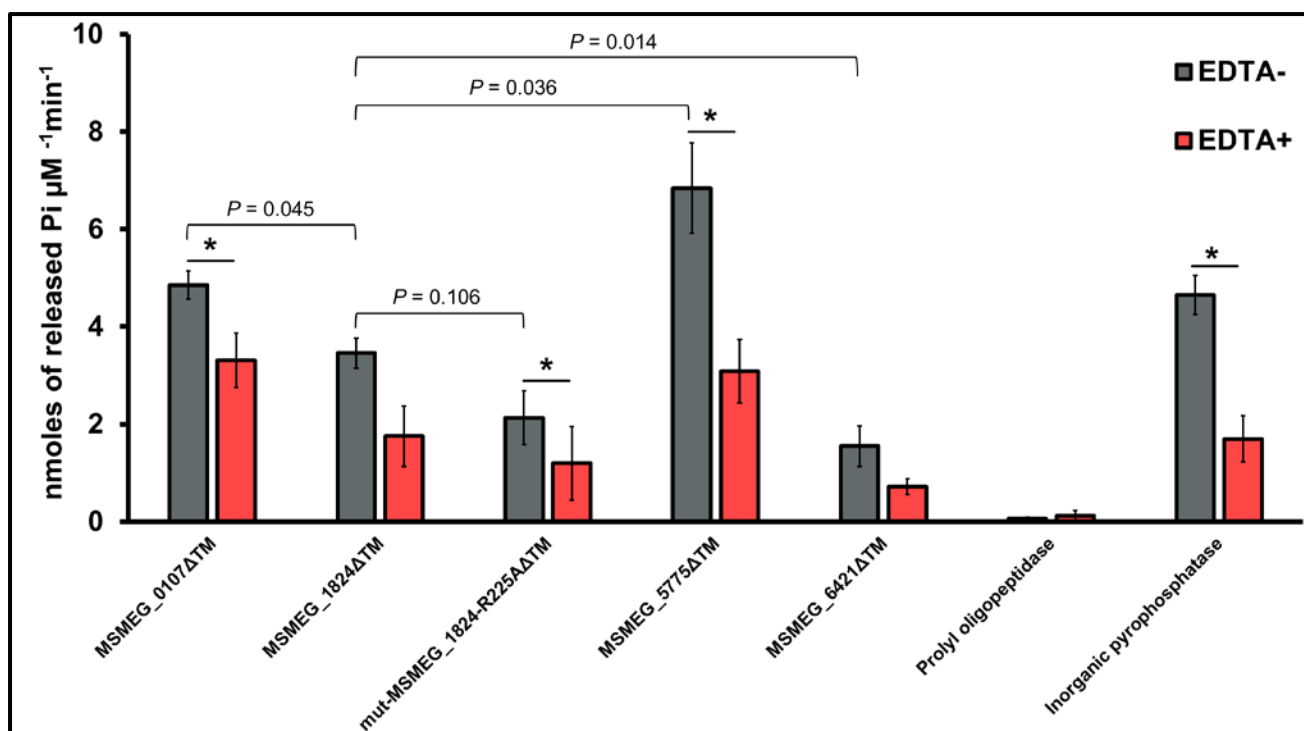
Fig 2.10. Pyrophosphatase activity of the GST-tagged LCPΔTM proteins in *Msmeg*. The release of inorganic phosphate (Pi) upon hydrolysis of the pyrophosphate phosphoanhydride bond of GPP is determined by measuring the absorbance at 560 nm. Reaction mixtures contained 1 μM of the fused recombinant LCPΔTM protein and were incubated for 16h in the absence (black bars) or presence (red bars) of 10 mM EDTA. Inorganic pyrophosphatase from yeast is used as the positive control. Prolyl oligopeptidase is the negative control. The data represents the detected Pi at T=2h (see Fig 2.9. A,B). The HRV-3C cleaved proteins were found to degrade rapidly even after storing at -80°C hence in our study, GST-tagged LCPΔTM proteins have been used which had a longer storage period. The data represented is an average of three independent experiments. Error bars represent ±SD and asterisk represents statistical significance ($P \leq$ 0.01).

2.4. Conclusions

The results presented in this chapter demonstrate that all the four LCP proteins in *Msmeg*, have pyrophosphatase activity in varied levels. This also indicates that all of the four *Msmeg* LCP homologues participate in releasing the newly bound AG from the Dec-P-P carrier lipid to form mature AG. However, there were noticeable variances in the enzymatic activity of the corresponding LCP homologues in *Msmeg* and *Mtb*, as compared with previous studies indicating either a lower activity in a non-pathogen or because of different purification methods. Although we have not used varied concentrations of the enzymes and the GPP substrate in the current study, such an approach could be considered in future to determine the enzyme kinetics. Since, the enzymatic role of the LCP proteins have not been previously studied in the non-pathogenic mycobacteria, our study contributes to understanding the role of these proteins in the non-tubercle bacterial cell wall biosynthesis. This study along with another previous study in LCP proteins have anticipated interesting role of LytR_C domain in protein oligomerization, and thus requires further investigation on this domain.

Chapter-3

Impact of deletion of *lcp* genes on the physiology of *Mycobacterium smegmatis*

Abhipsa Sahu[a], Ziwen Xie[a], Yuanchao Li[a], Xinyue Wang[a], Roy Ummels[b], Siauw Liang Liem[a,b], Wilbert Bitter[b], Mal Horsburgh[c] and Boris Tefsen[a]

[a]Department of Biological Sciences, Xi'an Jiaotong-Liverpool University, Suzhou, Jiangsu, 215123, PR China

[b]Department of Medical Microbiology and Infection Control, VU University Medical Center, Amsterdam, Netherlands

[c]Institute of Integrative Biology, University of Liverpool, Crown Street, Liverpool, L69 7ZB UK

3.1. Introduction

TB, which is caused by the highly infectious human pathogen *Mtb*, proliferates within the host environment and evades its immune defense mechanism (Stanley & Cox, 2013). This, along with its attribute to developing resistance to anti-TB drugs, makes *Mtb* a highly successful pathogen. Therefore, it is of utmost importance that we extend our understanding of the pathogenicity and physiology of the tubercle bacillus in the hope of priming novel therapeutic approaches to combat the growing threat of TB infection. The characteristic feature of mycobacteria is its complex cell wall which is a well-recognized drug target. The lipid-rich mycolic acid layer and the carbohydrate-rich capsular, PG and AG layers of the cell wall (Fig 1.5) not only serve as a permeability barrier, but also provides protection against hydrophilic compounds, and are essential in the pathogenesis and survival of *Mtb*. In this regard, the biosynthetic machinery and components of the cell wall assembly serve as attractive drug targets. One such family of proteins is the LCP family of proteins, which is widely distributed amongst many Gram-positive bacteria (Hubscher *et al.*, 2008).

LCP proteins are involved in cell envelope maintenance and bacterial virulence in many Gram-positive bacteria, and their loss has been implicated with various defects in the cell wall structure. For instance, BrpA, an LCP protein of *Streptococcus mutans*, plays important roles in biofilm formation, autolysis, and cell division, and the loss of this protein considerably affects the acid and oxidative stress tolerance (Bitoun *et al.*, 2012). CpsA in many pathogenic species

of *Streptococcus* is involved in the synthesis of capsular polysaccharides, protecting the organism from host defense mechanisms including phagocytosis, complement deposition, and clearance by mucus (Marques *et al.*, 1992, Nelson *et al.*, 2007, Beiter *et al.*, 2008). ConR, an LCP protein in *Anabaena* sp., is involved in vegetative cell septum formation under specific growth conditions (Mella-Herrera *et al.*, 2011). Apart from this, LCP proteins in *S. aureus* are essential for cell division, autolysis, and β-lactam resistance (Over *et al.*, 2011). For instance, in *S. aureus*, β-lactam resistance was shown to be contributed by MsrR (Rossi *et al.*, 2003), whereas both MsrR of *S. aureus* and Psr of *Enterococcus faecalis* were described to augment virulence in *Caenorhabditis elegans* (Bae *et al.*, 2004, Maadani *et al.*, 2007).

In *Actinomyces oris*, these proteins are involved in protein glycosylation, i.e., transferring the glycan strands from a polyprenoid lipid coupled with glycan, to a glycosylated surface-linked protein A (GspA), before attaching to the cell wall (Wu *et al.*, 2014). In *S. aureus*, the LCP proteins have shown to transfer the secondary cell wall polymers like teichoic acids onto the glycan strands of PG whereas in *Streptococcus*, capsular polysaccharides are transferred to C6-hydroxyl of MurNAc in PG (Chan *et al.*, 2014).

Mutants completely devoid of these proteins have been shown to release type 5 capsular polysaccharides and wall teichoic acids into the medium (Chan *et al.*, 2013, Chan *et al.*, 2014). In *Streptococcus agalactiae* and *M. marinum*, inactivation of CpsA has shown to be associated with substantial attenuation in zebrafish (Hanson *et al.*, 2012, Wang *et al.*, 2015). The inactivated CpsA in *M. marinum* resulting from a transposon insertion has shown to be involved with altered colony morphology, sliding motility, surface hydrophobicity, permeability and reduced AG content (Wang *et al.*, 2015). A conditional mutant of an essential *lcp* gene in *C. glutamicum* has also been implicated with severe morphological alterations, growth defects as well as reduced mycolic acids and AG, all of which indicates their role in structural integrity (Baumgart *et al.*, 2016).

Apart from the effects of the *lcp* genes on the cell envelope of Gram-positive organisms, as described above, their function as a ligase that couples AG to PG has also been recently described in *C. glutamicum* and *Mtb* (Baumgart *et al.*, 2016, Grzegorzewicz *et al.*, 2016, Harrison *et al.*, 2016). In chapter 2, this pyrophosphatase activity were also demonstrated for all the four LCP homologues found in *Msmeg*. However, the role and potential interplay of these LCP proteins was required to be addressed by examining the physiological changes in *Msmeg* mutants that are missing in one or more of the LCP proteins. Therefore in this study, a panel of single and double *lcp* deletion mutants was created, to understand the influence of these genes on the surface morphology and other physiological characteristics in *Msmeg*.

3.2. Materials and methods

3.2.1. Bacterial strains, plasmids and growth conditions

All bacterial strains and plasmids used in this chapter are listed in Table 3.1. All the reagents were purchased from Sigma-Aldrich, China unless otherwise stated. *Msmeg* wild type, mutant and complemented strains were grown in 7H9 medium (Difco) supplemented with 10% (v/v) ADS (50 g of albumin, 20 g of dextrose and 8.5 g of sodium chloride in 1 liter of water), 0.5% glycerol (v/v), 0.05% Tween-80 (v/v) (Solarbio, China) and appropriate antibiotics as mentioned below. For long term storage of *Msmeg* strains, fresh cultures in mid-log phase (optical density measured at 600 nm (OD_{600}) of 1 - 1.5) were frozen in 50% glycerol and stored at -80°C. They were then grown overnight at 37°C, shaking at 200 rpm. The seed strains were stored at 4°C for a month. For every experiment, seed strain was inoculated in fresh 7H9 medium at a dilution of 1:100, supplemented with Tween-80 and antibiotics. In solid culture, the bacterial strains were grown in 7H10 plates supplemented with 10% ADS, Tween-80 and antibiotics. Recombinant plasmid constructs were propagated using *E. coli* DH5-α. In this chapter, the marked double deletion mutants ΔΔ(0107+5775), ΔΔ(0107+6421) and ΔΔ(5775+6421), as well as complemented strains of single deletion mutants viz., c-0107, c-5775 and c-6421 were supplemented with 0.25 µg/mL of hygromycin (Hyg). The complemented strains of double deletion mutants, c-(0107+5775), c-(0107+6421) and c-(5775+6421) were supplemented with a final concentration of 0.25 µg/mL of kanamycin (Kan).

3.2.2. Preparation of *Msmeg* competent cells

The mycobacterial cells were grown overnight in 7H9 medium supplemented with ADS, Tween-80 and antibiotics (wherever required) after which they were placed in ice for about 2 hours. The cells were then centrifuged at 16,000 *g* for 10 min at 4°C and the pellet was washed twice with 10% glycerol. Finally, the cells were resuspended in 10% glycerol and stored in pre-cooled tubes at -80°C.

Table 3.1: Strains and plasmids used to study the impact of deletion mutants on cell envelope physiology

Strain or plasmid	Description	Source or reference
E. coli		
DH5-α	strain used for general cloning procedures	
M. smegmatis		
mc2155	wild-type laboratory strain; DNA used as PCR template	(Snapper *et al.*, 1990)
Δ0107	MSMEG_0107 single deletion mutant	This study
Δ5775	MSMEG_5775 single deletion mutant	This study
Δ6421	MSMEG_6421 single deletion mutant	This study
ΔΔ(0107+5775)	Hyg^R; MSMEG_0107-MSMEG_5775 double deletion mutant	This study
ΔΔ(0107+6421)	Hyg^R; MSMEG_0107-MSMEG_6421 double deletion mutant	This study
ΔΔ(5775+6421)	Hyg^R; MSMEG_5775-MSMEG_6421 double deletion mutant	This study
c-0107	Hyg^R; MSMEG_0107 complemented strain	This study
c-5775	Hyg^R; MSMEG_5775 complemented strain	This study
c-6421	Hyg^R; MSMEG_6421 complemented strain	This study
c-(0107+5775)	Kan^R; MSMEG_0107-MSMEG_5775 complemented strain	This study
c-(0107+6421)	Kan^R; MSMEG_0107-MSMEG_6421 complemented strain	This study
c-(5775+6421)	Kan^R; MSMEG_5775-MSMEG_6421 complemented strain	This study
Plasmids		
pML2424	Hyg^R; temperature sensitive plasmid for creating deletion construct	A gift from Michael Niederweis
pML2714	Kan^R; cre expression plasmid for removing the cassette	A gift from Michael Niederweis
pML2424-0107	Hyg^R; MSMEG_0107 deletion construct	This study
pML2424-5775	Hyg^R; MSMEG_5775 deletion construct	This study
pML2424-6421	Hyg^R; MSMEG_6421 deletion construct	This study
pSMT3	Hyg^R; complementation vector of single lcp deletion mutant	A gift from Michael Niederweis
pSMT3-mspA-LipYtb	Kan^R; for complement construct of double lcp deletion mutant	A gift from Michael Niederweis
pSMT3-0107	Hyg^R; MSMEG_0107 complementation construct	This study
pSMT3-5775	Hyg^R; MSMEG_5775 complementation construct	This study
pSMT3-6421	Hyg^R; MSMEG_6421 complementation construct	This study
pSMT3-0107-5775	Kan^R; MSMEG_0107-MSMEG_5775 complementation construct	This study
pSMT3-0107-6421	Kan^R; MSMEG_0107-MSMEG_6421 complementation construct	This study
pSMT3-5775-6421	Kan^R; MSMEG_5775-MSMEG_6421 complementation construct	This study

3.2.3. Construction of single and double *lcp* deletion mutants

3.2.3.1. Creation of gene-inactivation constructs

To construct single and double deletion mutants, LB broth and LB agar were used whenever required. The gene deletion strategy in *Msmeg* was adapted from the Niederweis group (Ofer *et al.*, 2012) and requires pML2424 plasmid (kindly provided by Michael Niederweis, University of Alabama, Birmingham). This plasmid contains a temperature-sensitive PAL5000 origin of replication and the *sacB* gene as a counter-selectable marker. In addition, this vector contains Hyg resistance as the positive selection marker and, two reporter genes (*gfp* and *Tdtomato*), which indicates the integration of the plasmid and loss of the plasmid backbone. The reporter genes enable distinction between spontaneous Hyg-resistant mutants and allelic exchange clones on a plate. For preparing a single *lcp* gene deletion construct, the gene upstream and downstream of the target gene to be deleted were amplified using knockout construct primers (Table 3.2) and wild type mc^{2}155 genomic DNA (refer to section 2.2.4) as template. PCR products were confirmed by gel electrophoresis followed by gel extraction (section 2.2.5.2 and 2.2.5.3). The pML2424 plasmid was digested with *Spe*I (NEB) and *Nsi*I (NEB) to yield the large fragment of vector and digested with *Pac*I (NEB) and *Swa*I (NEB) to yield the other small fragment of the vector. The four fragments thus produced were fused together using the In-Fusion® HD Cloning Plus CE kit (Clontech) as per the manufacturer's instructions, transformed (refer section 2.2.5.6) and selected on LB/Hyg plates. The transformants were confirmed by colony PCR and/or restriction analysis (refer section 2.2.6). Selected transformants were inoculated into LB medium, their plasmids isolated using a QIAprep Spin Miniprep Kit (Qiagen) according to manufacturer's protocol and finally sequenced (Sangon Biotech Co. Ltd, Shanghai, China).

Table 3.2: Primers used for cloning, generating and confirming knockout constructs

Primer name	Sequence (5' - 3')
Gene amplification primers	
MSMEG_5775-F-HindIII	GCGAAGCTTGTGAACGATCCGGTGGACGA
MSMEG_5775-R-ClaI	GCTAGATCGATTCATTCGCAGGTGGCGTCGGC
MSMEG_6421-F-PstI	GTATTCTGCAGGTGGCGCAGCGGAATCG
MSMEG_6421-R-HindIII	GGCGCAAGCTTTCAATTGACGGCATTTTCCAG
MSMEG_0107-F-HindIII	GCGGCAAGCTTTCATTTCACGCACGG
MSMEG_0107-R-ClaI	GCCGCATCGATATGAACTCACCTG
MSMEG_5775-F-NheI	GCGGCTAGCGTGAACGATCCGGTGGACGA
MSMEG_5775-R-PmeI	CGTGTTTAAACTCATTCGCAGGTGGCGTCGG
MSMEG_0107-F-PmeI	CGGGTTTAAACATGAACTCACCTGCCCATGC
MSMEG_0107-R-SpeI	CGACTAGTTCATTTCACGCACGGGATGC
MSMEG_6421-F-NheI	GCGGCTAGCGTGGCGCAGCGGAATC
MSMEG_6421-R-PmeI	CGTGTTTAAACTCAATTGACGGCATTTTCCAGG
Knockout construct primers	
MSMEG_0107-UF-SpeI-Fw	ATATTGATCCACTAGTCGAGCATGAACGCAACGAGA
MSMEG_0107-UF-SwaI-Rev	TATACGAAGTTATTTAAATTATGCTGCCGATGGCGTGAA
MSMEG_0107-DF-PacI-Fw	TATACGAAGTTATTAATTAACACAGACCTGAGCCAGATGA
MSMEG_0107-DF-NsiI-Rev	GACAATAACCATGCATCGATGGCCAAGGAGCTGTAT
MSMEG_5775-UF-SpeI-Fw	ATATTGATCCACTAGTTCCCGGTGACGGTCAAGTTC
MSMEG_5775-UF-SwaI-Rev	TATACGAAGTTATTTAAATCGTCGTCCACCGGATCGTTC
MSMEG_5775-DF-PacI-Fw	TATACGAAGTTATTAATTAACATCCCGACCACGGGTTACG
MSMEG_5775-DF-NsiI-Rev	GACAATAACCATGCATGCTGGCGCGACTTCCTGATG
MSMEG_6421-UF-SpeI-Fw	ATATTGATCCACTAGTGATTCCACCGGCGCGGAGTC
MSMEG_6421-UF-SwaI-Rev	TATACGAAGTTATTTAAATGTGGCGGCTACCGCAACCAG
MSMEG_6421-DF-PacI-Fw	TATACGAAGTTATTAATTAAGACGTGGTGATGTGGGACAG
MSMEG_6421-DF-NsiI-Rev	GACAATAACCATGCATCTGGAACCGTCCGGTCAAGC
Knockout confirmation primers	
KO-0107 gene F	CGTTCTCGGCACTGGACTCA
KO-0107 gene R	GAACTCCGCGAACGCAATCC
KO-5775 gene F	ACAGTCGGAGGCCGATAGGA
KO-5775 gene R	GCGGTGTTGGTGGTGGATTC
KO-6421 gene F	CTGCGCCAACTCGACATCAG
KO-6421 gen R	ATCGCGTGATTGCGCTCCTC
KO-6421 del F	GCACCGACACGATCCTCTTG
KO-6421 del R	CGTCCGTCGAGTTCCTGACA
pML2424-cas1	GCGTGCGAACGCACAGATCA
pML2424-cas2	CAAGTACCGCCACCTAACAA
qPCR primers	
5775-qPCR-F	CAGCAGTTGTTCCTGTCGTC
5775-qPCR-R	TGACCACGTTGTTGAGCTTG
6421-qPCR-F	CATCGCCCACACTCTGATC
6421-qPCR-R	ACCGTTCATCTCAGCAATGTAC
0107-qPCR-F	TCGTTGAGGCATACGGTGATC
0107-qPCR-R	GGAATCCCCATCGACCATTT
smeg-rpoD-qPCR-F	GTGTGGGACGAGGAAGAGTC
smeg-rpoD-qPCR-R	ACCTCTTCTTCGGCGTTGAG

Table 3.3: Primers used for sequencing the complement constructs

Primer name	Sequence (5' - 3')
Sequencing primers	
pSMT3-5775-Fseq-3381	CAGCTGCTGGGATTACACATGG
pSMT3-5775-Fseq-4119	CGAAGAACGACTCCCTCAACAAG
pSMT3-5775-Fseq-4870	CGCCGACAGCTACGTCGACAACATC
pSMT3-6421-Fseq-3350	AGAGACCACATGGTCCTTCTTG
pSMT3-6421-Fseq-3941	GCCACGTACATTGCTGAGATGA
pSMT3-6421-Fseq-4680	CCGGTTACGGCGAGGACAAGATCA
pSMT3-0107-Fseq-3381	CAGCTGCTGGGATTACACATGG
pSMT3-0107-Fseq-4128	TCGAGCGCCACGTTCTGCTT

Table 3.4: Primers used to study relative expression of cytokines produced in the wild type, deletion mutants and their complemented strains

Primer name	Sequence (5' - 3')
Cytokine qPCR primers	
TNF-α-qPCR-F	GATCAATCGGCCCGACTATC
TNF-α-qPCR-R	TCCTCACAGGGCAATGATCC
IL-1β-F	TGGCAATGAGGATGACTTGT
IL-1β-R	GTGGTGGTCGGAGATTCGTA
IL-6-F	GCCACTCACCTCTTCAGAACG
IL-6-R	CCGTCGAGGATGTACCGAATT
IL-10-F	GACTTTAAGGGTTACCTGGGTTG
IL-10-R	TCACATGCGCCTTGATGTCTG
GAPDH-F	CCATGTTCGTCATGGGTGTG
GADPH-R	GGTGCTAAGCAGTTGGTGGTG

3.2.3.2. Transformation and gene deletion

Confirmed single *lcp* gene deletion constructs were transformed into *Msmeg* strain mc²155, using standard electroporation protocol however, the recovery and plating temperature was 30°C. The reporter gene activity of the transformants was confirmed by fluorescence microscopy, followed by fresh plating of a positive transformant that displayed both green and red fluorescence (from *gfp* and *TdTomato*). Diluted cultures of the positive transformant were plated on LB/Hyg plates, supplemented with 10% (w/v) sucrose and incubated at 40°C. At this restrictive temperature, double-crossover clones resulting from homologous recombination, express GFP but not TdTomato. The *sacB* gene enables this selection of double-crossover mutants in the presence of sucrose, resulting in successful deletion of the target gene (Fig 3.1). This strain was termed marked deletion mutant, because of the presence of cassette that contains both *gfp* and *hyg* resistance marker. Another plasmid pML2714, was subsequently used to facilitate excision of this cassette from the chromosome via *loxP* site-specific recombination catalysed by the *cre* recombinase, resulting in an unmarked deletion mutant strain. All the deletion mutants were confirmed by RT-PCR (section 3.2.5).

3.2.3.3. Construction of double deletion mutant strains

Double deletion mutants were made by transforming the pML2424 based recombinant plasmid construct of a second target *lcp* gene into competent cells of an unmarked single deletion strain. Thus, the pML2424-0107 construct was transformed into competent unmarked Δ5775 strain resulting in double deletion mutant strain ΔΔ(0107+5775). Similarly, ΔΔ(0107+6421) and ΔΔ(5775+6421) were constructed by transforming pML2424-0107 into the unmarked Δ6421 strain and pML2424-5775 into the unmarked Δ6421 strain, respectively. Similar procedures as described in section 3.2.3.2 were carried out to create and confirm the double deletion strains.
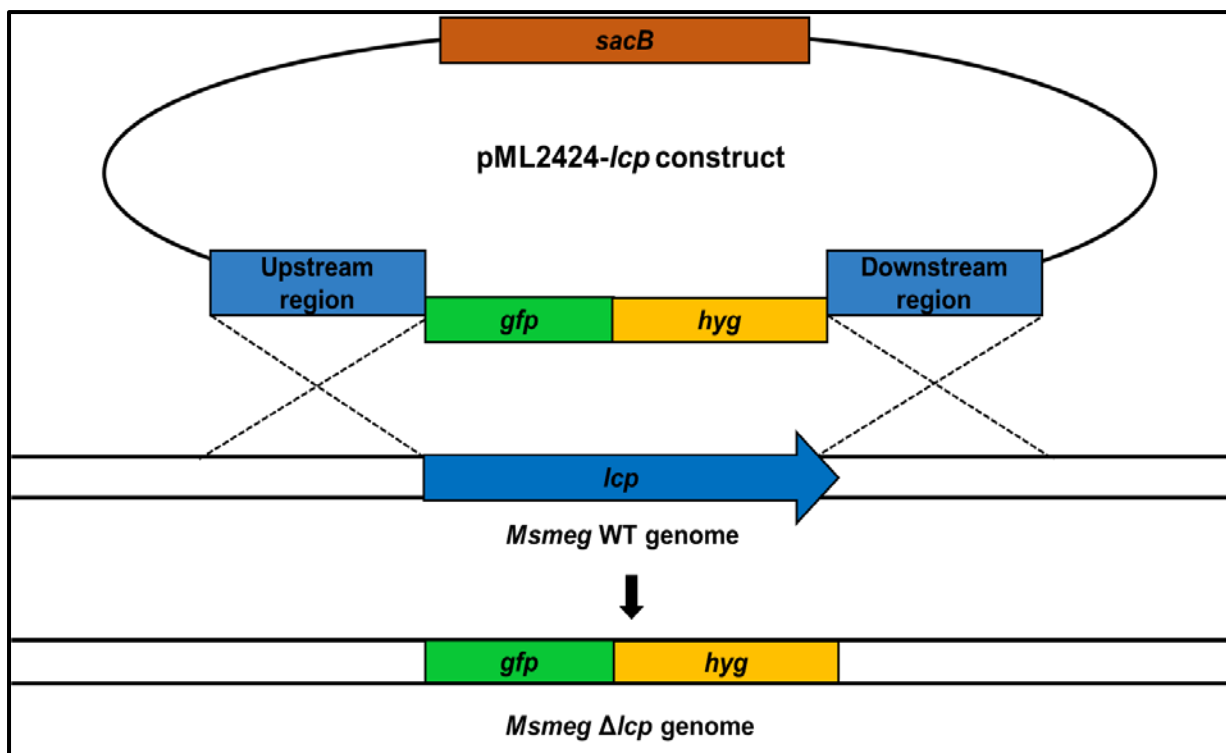
Fig 3.1. Schematic representation of the construction of a single *lcp* deletion mutant. Levansucrase, the product of *sacB* gene present in pML2424 temperature-sensitive plasmid catalyzes the conversion of sucrose to levan, which is toxic to the bacteria, resulting in defective growth in media containing sucrose. A hygromycin (*hyg*) resistance marker ensures positive selection on Hyg plates. The reporter genes *gfp* and *TdTomato* are fluorescent markers. Together, *gfp* and *hyg* is termed as cassette. Cloning of the upstream and downstream regions of target gene on either side of the cassette, and transforming into the wild type (WT) genome enables homologous recombination in the WT genome at both ends. This leads to allelic replacement of target gene with the cassette, and subsequent removal of the target gene.

3.2.4. Construction of complemented strains of single and double *lcp* deletion strains

The complemented strains of single *lcp* gene deletion mutants were constructed using the plasmid pSMT3 that contains a cassette with a *gfp* reporter and a *hyg* resistance gene whereas, to create complemented strains of double deletion mutants, pSMT3-*mspA-lipYtb* that contains a *gfp* and kanamycin resistance gene was used. For complementation of single deletion mutants, the target gene was fused downstream of *gfp* reporter gene and expressed under the control of the constitutive promoter *Phsp*60. For complementation of double deletion mutants, the two pre-cloned adjacent *Mtb* genes *mspA* and *lipYtb* were replaced with the target *lcp* genes. The constructs were confirmed by colony PCR, restriction analysis and subsequently sequenced using sequencing primers (Table 3.3) as described in section 2.2.6. The confirmed plasmid constructs were transformed by standard electroporation procedure to create complementation strains of the *lcp* deletion mutants.

3.2.5. mRNA extraction and reverse transcription- PCR (RT-PCR)

The wild type and mutant strains overnight grown cultures were centrifuged at 11,000 g for 2-3 min and the supernatant discarded. The pellet was resuspended in 100 µL TE (10 mM Tris-HCl pH 8.0 plus 0.1 mM EDTA) buffer and lysozyme at a final concentration of 3 mg/mL to disrupt the mycobacterial cells. The mRNA was then extracted using the manufacturer's instructions in Innuprep Micro RNA kit (Analytikjena, China). The extracted mRNA was reverse transcribed to cDNA using GoScript™ Reverse Transcription System (Promega, China) according to the manufacturer's instructions. GoTaq® qPCR Master Mix (Promega, China) and primer sequences listed in Table 3.2 were used for qPCR (cycling conditions: 95°C for 2 minutes followed by 40 cycles at 95°C for 15 seconds and at 60°C about 1 minute) in QuantStudio™ 5 System (ThermoFisher, China). In this study, SigA (*rpo*D) was used as the reference gene for confirming the mutants, and GAPDH was used as a reference gene for cytokine qPCR. The relative expression analyses are based on three replicates of at least three independent experiments.

3.2.6. Growth kinetics

Growth of the wild type strain, single and double *lcp* deletion mutants as well as the complementation strains were measured in 96-well plates. The overnight cultures were adjusted to OD_{600} of 0.5 and diluted further 1:20. Subsequently, a total volume of 200 µL of each strain was plated in triplicate and incubated at 37°C for 60 hours in a microplate reader (Varioskan LUX, Thermo Fisher), with readings taken every 2 hours at OD_{600}. The growth curve was plotted using Microsoft Excel.

3.2.7. Scanning electron microscopy (SEM)

For SEM, the wild type strain and the double *lcp* deletion mutant ΔΔ(0107+5775), were fixed with 2.5% (v/v) glutaraldehyde (Macklin) in PBS and sent to Yuan test (QingDao) for further processing and imaging. According to the company's protocol, the fixed cells were first washed in PBS for 15-30 min, followed by dehydrating them. The dehydration process is done by incubating the cells consecutively in an ascending acetone series (30%, 50%, 70%, 90%, and 100%) for 15-20 min each and the last step repeated thrice. The samples were then dried in liquid CO_2 to a critical point followed by sputter coating with a 10-nm gold/palladium layer. Samples were subsequently analyzed using an environmental scanning electron microscope

(ESEM XL 30 FEG, FEI, Philips, Eindhoven, Netherlands) in a high vacuum environment with a 20 kV acceleration voltage.

3.2.8. Colony morphology

The colony morphology of the wild type, *lcp* deletion mutants and their complemented strains were analyzed on 7H10 plates supplemented with 10% ADS, appropriate antibiotics (section 3.2.1) of 0.5 µg/mL concentration and 100 µg/mL Congo red (Cangelosi *et al.*, 1999, Klepp *et al.*, 2012). Briefly, 5 µL of overnight cultures were adjusted to OD_{600} of 0.5 and spotted on these plates in triplicate and incubated for 3 days at 37°C, in plastic bags. The colonies were subsequently inspected by naked eye and each colony diameter was determined by averaging two perpendicular measurements. The experiment was done at least thrice.

3.2.9. Biofilm formation and quantification

For biofilm formation on liquid-air interface, the method was adapted from previous studies in *Msmeg* with few modifications (Bharati *et al.*, 2012, Zanfardino *et al.*, 2016). Briefly, adjusted stationary phase culture (OD_{600} of 0.5) was inoculated into a petri dish containing 20 mL 7H9 medium at a 1:100 dilution, supplemented with 10% ADS and appropriate antibiotics (section 3.2.1), but devoid of Tween-80, and incubated in standing condition for 4-5 days at 37°C. For biofilm quantification, crystal violet assay was performed as described earlier (Bharati *et al.*, 2012) with few modifications. Briefly, overnight cultures were washed with Sauton's medium and diluted to a final OD_{600} of 0.05 in Sauton's medium supplemented with 2% glucose, followed by plating 200 µL of the cultures into each well and incubating at 37°C for 5 days. Here, 8 wells were assayed for each strain in one experiment and the assay was repeated at least thrice. The samples were discarded from the wells and washed 2-3 times with water. The biofilm adhering to the walls of the plate was stained with 1% crystal violet solution and incubated for 1 hour at room temperature. The excess dye was washed 3-4 times with water and the plate air dried. The bound dye was then solubilized in 300 µL 80% (v/v) ethanol and the A_{550} was measured using Biotek Synergy 2 microplate reader, and the results analyzed using the Gen5 software (Biotek, Winooski, USA).

3.2.10. Cellular aggregation

The method is a modified protocol from a previous study (Yang *et al.*, 2017). Saturated cultures of the strains were grown in 7H9 medium supplemented with 10% ADS but without Tween-80. Cells were washed thrice with PBS and finally resuspended in 5 mL PBS. The cellular aggregates were broken by vortexing the resuspended cells for 1 min and the initial OD_{600} (T=0 min) of the cell suspension was measured using a spectrophotometer. Subsequent readings were taken without resuspending the cells, for 15 minutes at an interval of 1 min. The results were analyzed by first normalizing each time point to T=0 min and subsequently calculating the percentage reduction of each time point from initial OD_{600}. For visual inspection, the washed and vortexed cells were kept stand still for 4 minutes and the image captured.

3.2.11. Extraction and purification of mycobacterial lipids for thin layer chromatography

This method in *Msmeg* is an adaptation from an unpublished protocol shared by Nicholas P. West (University of Queensland). Briefly, overnight cultures (OD_{600} of 1.5 - 2) of the strains in 7H9 medium were spun down at 11,000 *g* for 10 min and washed 2-3 times with PBS. The bacterial pellet was air dried and weighed. 1 g of the bacterial cells were resuspended in 1 mL mixture of chloroform:methanol (2:1) in PTFE-capped glass centrifuge tubes (Fisher Scientific), and incubated overnight at room temperature on a rotating wheel. The bacterial suspension was centrifuged for 30 min at 11,000 *g*. The resulting supernatant containing a mixture of polar and apolar lipids was carefully separated from the pellet and air dried. This dried supernatant was dissolved in 50 µL of chloroform:methanol at a ratio of 2:1 and were spotted on to the TLC silica gel 60 (Merck Millipore, China). The TLC plates were developed in chloroform:methanol:water at a ratio of 20:4:0.5, and subsequently stained with a solution containing 8% aqueous phosphoric acid (v/v) and 10% copper (II) sulfate (w/v) (Churchward *et al.*, 2008). The plates were then charred using a heat gun until the lipids were revealed.

3.2.12. Determination of minimum inhibitory concentration (MIC)

The MIC values of various antibiotics for each strain under study, were determined using the resazurin (7-hydroxy-10-oxidophenoxazin-10-ium-3-one) microtiter plate assay (REMA), which is a widely used method for determining MICs in mycobacteria (Palomino *et al.*, 2002). This method was adapted from a previous method done in *Msmeg* (Agrawal *et al.*, 2015) with

few modifications. Briefly, overnight cultures were adjusted (OD$_{600}$ of 0.5) and diluted 1000 times. Then, 50 µL of 7H9 medium supplemented with 10% ADS was added to all wells in a Polystyrene 96-well plate (Corning®) except the first row. Antibiotic concentration double the requirement was prepared, and 100 µL was added to the wells in the first row. This was serially diluted to half the concentration in the subsequent wells, by mixing 50 µL of the antibiotic to 50 µL of the medium until the penultimate row. The last row containing only medium was the negative control i.e. without antibiotics. Then, 50 µL of the diluted overnight cultures was added to each well thus attaining the desired antibiotic concentration in each well. The plate was sealed in a plastic bag and incubated at 37°C for 48h after which 30 µL of filter-sterilized 0.2 mg/mL resazurin dye was added to each well and incubated again overnight at 37°C. If live bacteria are present in any of the wells, the dye changes it color to pink; otherwise, it retains its blue color. A lack of metabolic activity indicates inhibitory activity of the compound tested but not necessarily its bactericidal activity. The MIC was defined as the minimum antibiotic concentration at which the color of resazurin did not change i.e., the first well showing bacterial growth inhibition. Three independent experiments were performed each in duplicate.

3.2.13. Sodium dodecyl sulfate (SDS) sensitivity assay

Sensitivity of the wild type, *lcp* deletion mutants and their complemented strains to the detergent sodium dodecyl sulfate (0.1% SDS) was determined as in a previous study (Meng *et al.*, 2017). Briefly, overnight cultures were adjusted to OD$_{600}$ of 0.5 and further diluted to 1:10. Then, 100 µL of this diluted bacterial suspension was spread over a 7H10 plate supplemented with 10% ADS, using a cotton swab. 5 µL of different concentrations of SDS was pipetted on to sterilized Whatman discs and placed on the plate followed by incubation at 37°C for 48 hours. The sensitivity of the strains to the detergent was estimated by measuring the diameter of the inhibition zone. The experiment was done in triplicates and was repeated at least thrice to determine the average sensitivity zone diameter.

3.2.14. Lysozyme sensitivity assay

The resazurin assay was used to determine the susceptibility of the *lcp* deletion mutants under lysozyme stress, as done in a previous study (Malm *et al.*, 2018). Briefly, overnight exponential cultures were adjusted to OD$_{600}$ of 0.5 and diluted further to 1:20. Subsequently, 100 µL of this diluted culture was added to wells in a 96-well plate (Corning, USA) and different

concentration of lysozyme was added to each of the bacterial strains, as indicated in the results. The plate was incubated at 37°C for 24 hours in a sealed plastic bag after which 30 µL of 0.2 mg/mL resazurin was added to the wells and the plate was incubated further at 37°C overnight. The fluorescence was detected by Varioskan LUX fluorescent microplate reader (Thermo Fisher) and emission detected at 590 nm using an excitation wavelength of 540 nm. Background fluorescence was corrected by subtracting values obtained from wells containing medium only. Experiments were performed in triplicate and repeated at least thrice.

3.2.15. Immunological response to *lcp* mutants in differentiated THP-1 cells

The human monocytic cell line THP-1 was obtained from the Cell Bank of the Academy of Sciences (Shanghai, China, http://www.sibcb.ac.cn/ep6-1.asp). Cells were maintained in RPMI-1640 (Gibco) supplemented with 10% heat-inactivated fetal bovine serum (FBS, Gibco), 2 mM L-glutamine (Gibco), 100 U/mL penicillin and 100 µg/mL streptomycin (HyClone) at 37°C with 5% CO_2 and passaged twice a week. The wild type and the mutant cells were seeded at a density of 5 x 10^4 cells/well in a 96-well flat-bottom plate with medium containing 100 µg/mL phorbol-12-myristate-13-acetate (PMA, Promega) for 48 h. Lipopolysachharide (LPS) from *E. coli* was used as a positive control in this experiment. Differentiated THP-1 cells were infected at a multiplicity of infection (MOI) of 10 for 2 hours followed by adding 100 µg/mL gentamycin per well to kill extracellular bacteria. At the time points, 2h, 4h and 8h (T_2, T_4 and T_8), mRNA was isolated from macrophages using Innuprep mRNA isolation kit (Analytikjena) for measuring relative expression by qPCR. The primers used for cytokine qPCR are listed in Table 3.4. This is a preliminary experiment and has been conducted once with three replicate for the qPCR relative expression. The reference gene GAPDH was used to normalize gene expression.

3.2.16. Detection of AG and mannose components in the cell envelope of *lcp* mutants

A direct ELISA was performed in a NUNC maxisorp 96-well ELISA plate (Thermo Fisher Scientific) to detect AG and mannosylated biomolecules of the wild type and the *lcp* deletion mutants by EB-A2 mAB and ConA lectin, respectively. The commercial Platelia™ Aspergillus kit used to detect Aspergillosis, contains the rat antigalactofuran (anti-Gal*f*) IgM EB-A2 mAB. Briefly, overnight cultures were washed 2-3 times with PBS, suspended to a final concentration of 1.5 X 10^8 cells/well in 50 mM carbonate coating buffer and 150 µL of this

bacterial suspension was added into each well of a 96-well plate. The reference Gal*f* embedded in gold nano particle (Gal*f*-GNP) or mannotriose-polyacrylamide (man-3-PAA) was diluted in carbonate buffer at indicated concentration and plated. After coating overnight at 4°C, the plates are dried and washed twice with PBS containing 0.05% Tween-20 (PBST). The plates were then blocked with 1% BSA in PBST for 1h at room temperature. 150 µL of 1:5 dilution of HRP conjugated EB-A2 (stock concentration 1 mg/mL) in PBST or, 2 µg/mL final concentration of ConA was added to each well and incubated at 4°C overnight. The plates were then washed 2-3 times with PBST and 200 µL TMB substrate solution was subsequently added to each well. The plates were incubated for 30 min at room temperature with gentle shaking and 50 µL of 0.8M H_2SO_4 (stop solution) was then added. The absorbance was measured at 450 nm using a microplate reader.

3.2.17. Statistical analysis

Microsoft excel was used for statistical analysis. Data were expressed as the mean ± SD and student's two-tailed t-test was used to determine significant differences between the groups. *P*-values in a range of 0.0005-0.05 were considered to be significant, as indicated in the experimental figures.

3.3.   Results and discussion

3.3.1. Construction of deletion mutants and their complemented strains

The use of pML2424, a temperature sensitive plasmid, in deletion of single and double *lcp* genes in our study, is facilitated by homologous recombination between flanking sequences upstream and downstream of the target gene to be deleted (Ofer *et al.*, 2012). The reporter gene *sacB*, plays a significant role in selecting the double crossover mutants, by acting as a counter-selectable marker. Levansucrase, a secretory enzyme encoded by *sacB* gene, catalyzes the hydrolysis of sucrose into levan, which is toxic to most Gram-negative bacteria (Sangiliyandi *et al.*, 1999), a few Gram-positive and many other bacteria (Jager *et al.*, 1992, Pelicic *et al.*, 1996). This toxic product is conferred by a cysteine residue in *sacB* gene (Senthilkumar *et al.*, 2003) resulting in defective growth of bacteria in the presence of sucrose. This strategy was successfully applied to create single *lcp* deletion mutant strains Δ0107, Δ5775 and Δ6421, as well as double *lcp* deletion strains ΔΔ(0107+5775), ΔΔ(0107+6421)

and ΔΔ(5775+6421) (Table 3.1). This study is the first of its kind to construct and study double *lcp* deletion mutants, as this would give more insights into understanding the impact of deleting two non-essential *lcp* genes from the *Msmeg* genome, on the mycobacterial cell physiology and possible altered chemical composition in the cell envelope. However, using this strategy, several attempts to delete *MSMEG_1824* were unsuccessful. During this study, the paper published by Harrison (Harrison *et al.*, 2016) confirmed our suspicion that *MSMEG_1824* is an essential gene, so these attempts were halted. Though conditional expression-specialized transduction essentiality test (CESTET) has been previously established to be a suitable tool to create efficient deletion of essential genes utilizing an inducible acetamidase promoter (Bhatt & Jacobs, 2009), and has been used to create a conditional deletion of the essential *lcp* gene *MSMEG_1824* in *Msmeg* (Harrison *et al.*, 2016), however, this conditional mutant could not be created in our lab due to unavailability of resources. Furthermore, several attempts were made to generate a triple deletion mutant ΔΔΔ(0107+5775+6421) but they did not yield any colonies after transformation of pML2424-0107 construct into the unmarked mutant ΔΔ(5775+6421), and is thus anticipated to be lethal. Consequently, our current study includes three out of four single *lcp* deletion mutants, three double *lcp* deletion mutants and complemented strains of all these mutants. All the mutants were confirmed by PCR with site-specific primers on genomic DNA (Fig 3.3) and for lack of mRNA expression by quantitative RealTime PCR (Fig 3.2). The complement constructs for creating complemented strains were confirmed by restriction analysis (Fig 3.4) and sequencing. Finally, all the mutants and the complemented strains of the single *lcp* deletion mutants were confirmed for their transcription levels by real-time PCR (Fig 3.5). As of now, we do not have enough quantitative real-time PCR data on confirmation of complementation strains of the double *lcp* deletion mutants, and so aim to pursue that as a future work.

The double-crossover deletion mutants generated using the plasmid pML2424 has a cassette containing both *gfp* and *hyg* resistance marker. Hence, these mutants were termed marked mutants. In order to construct a double deletion mutant, it is required to remove the cassette from the chromosome of the marked single deletion mutant. Therefore, another plasmid pML2714 that facilitates the excision of this cassette from the chromosome via *loxP* site-specific recombination, was used. The resultant unmarked single *lcp* deletion mutant was then used as a template to introduce another pML2424-*lcp* gene construct to create a double deletion mutant. However, for unknown reasons, the unmarked double *lcp* deletion mutants ΔΔ(0107+5775) and ΔΔ(0107+6421) could not be created, meaning that the cassette containing *gfp* and *hyg* was still present within the *Msmeg* genome of these marked mutants. Although the cassette could be excised in ΔΔ(5775+6421), rendering it unmarked, it was decided to use the marked ΔΔ(5775+6421) throughout our study in order to maintain similar

culturing conditions for the three double deletion mutants to be used in various assays. In order to rule out any discrepancy between the marked and unmarked strains in any assays, both the marked and unmarked strains of ΔΔ(5775+6421) were confirmed by using them in few assays viz., morphological characteristics determination, growth kinetics, sliding motility and biofilm formation. Similar phenotypes between the marked and unmarked strains of each of these double deletion mutants were observed. The complemented strains of single and double *lcp* deletion mutants were constructed with the single constitutive promoter, *Phsp*60 which is present in the complementation plasmid pSMT3. *Phsp60* controls expression of the reintroduced *lcp* genes in all complementation strains in this study.

In our results to confirm the expression levels of the mutants and their respective complemented strains, all the single and double *lcp* deletion mutants showed the expected null expression (Fig 3.5), however, the complemented strains displayed several folds higher expression levels than the wild type strain (12-fold increase in c-0107, P=.001; 19-fold increase in c-5775, P=.002 and about 2.6-fold increase in c-6421) (Fig 3.5, A-C). This overexpression in the complemented strains has been associated with the presence of *Phsp60* promoter in pSMT3 plasmid (Movahedzadeh & Bitter, 2009).

A

I

0   500   1000   1500   2000   2500   2630

13 geneF

MSMEG_0107

2093

2600 geneR

492

Deleted region

2044

2093

II

MNSPAHALTKPRRRTPRRAVR**LARRALVGLTAAAILAGTGMGWV**TYHGALDGITTSNALAGESGS
VGDTENILIMGLDSRLDQHGNPLPEDVYQALHAGDETVGGY**NANVLIVVHLPGDGGPATAFSIPRD**
DYVDLAGCPSGICKGKVKQAYGFAYQAEMEATESEEEDPATREQKAREAGRKAEIATVRNLLGIPI
DHFVEVTLGAFFEIAKAVAPITVCLNDDTSDPYSGADFRKGVQQIDAAQAMAFVRQRRDINDENFT
DLDRTRRQQAFLAALVSALRQSGALSSPSKLRGLLDVTKQNVALDAGFDLASFTKRASALTDAPPT
LYTLPIKEFGQNAYGEDINIIDVPTIRKIVHDLVSKDDPATTTATPDKPELNGH**GATLDVINGSTYTGLA**
AKLQTTFAADNFTEGQIGDAEALEPATTIDYGPGADAAAHSLATELGVTATPSDWVLADTVRLTIG
TDFPGTDYLGADMFTSTTSGASESESYGPSETAELTPITTVAATATGTYTPA<mark>PTDLSQMTASGIPCVK</mark>

III

1 kb DNA marker

2 kb — 2587 bp

1 kb — 1034 bp

WT   Δ0107

B

I

0   500   1000   1500   2000   2500

71 geneF

302

MSMEG_5775

1983 geneR

2470

323

Deleted region

1819

302

2470

II

<mark>MNDPVDD</mark>DAEPTGNHADGVTVADLIAKVAGSDGATPRRSRRRRAAEPDPEPEFEAQAEPAPKPTP
EPEILPVAPEPPASADPTEVLQAIRYDEYPGDHAGEYPGDHPDAAYAGAADPNLDAIFEPAFAPMPF
AGVTAAHPDLDSCDTEIIPAVPAADEHADLPDTHRDSWRRRTRPQHAGTDDFDEPHRRHPNRKAMI
VGRAAAAMIAALTLVLTGGAWQWQSSKNDSLNKVAALDLGSRDILDPNAQFGDENFLIVGMDSRF
GENSQMGAGSTDDADGA**RSDTIMLVNIPASRERVVAVSFPRDLAITPMKCEPWDPETGTYGPIWD**
EASGTYGPDEVYTETKLNSAFAFGGPKCLVKVIQKISGLSVNRFMAVDFAGFSKMVDALGGVEVC
STTPLEDYELGTVLANAGRQIVDGHTALNYVRARQVTTEYNGDYGRIKRQQLFLSSLLRSLISKDT
FFSLSKLNNVVNMFIADSYVDNIKTKDLVDLGQSVQGVNAGRITFVT<mark>IPTTGYADEWGNEQPRMDD</mark>
<mark>VRALFDAIINDDPLPGEKNPDNTPVPGTPESTTNTAQAPQSTSPAPQPAHAAPAPGSAEMVDTVTT</mark>
<mark>EPQ</mark>QVTVKVSNSTGQSGLAATAAGELEQHGFNVTTPDDYPGPLDATTVFFSPGNEEAAATVASSF
SNAAIERVAGLGDVVQVVLGTDFMAVNPPSPSGSAIQVKLVRGTSSTPTQLPEDLSVTNAADATCE

III

100 bp DNA marker

500 bp — 416 bp

Δ5775

Fig 3.2. Confirmation of the single *lcp* deletion mutants in *Msmeg*. Deletion mutants were created in *Msmeg* for **A.** Δ0107 **B.** Δ5775 and **C.** Δ6421. **A-I, B-I** and **C-I** represents the sequence spanning the deletion region (blue rectangular box), the remnant portion of the target *lcp* genes that could not be deleted (pink rectangular box) and, the primers used to confirm the deletion. The *lcp* gene is represented by a black arrow, and all the start and end sites are numbered in accordance to the full length sequence shown here. **A-II, B-II** and **C-II** are amino acid sequences of the respective *lcp* genes in *Msmeg*. The deleted region is highlighted in pink, transmembrane domain is highlighted in blue, LCP domain is bold and underlined and LytR_C domain is in red. **A-III, B-III** and **C-III** shows the confirmation of single lcp deletion mutants by agarose gel electrophoresis.

74
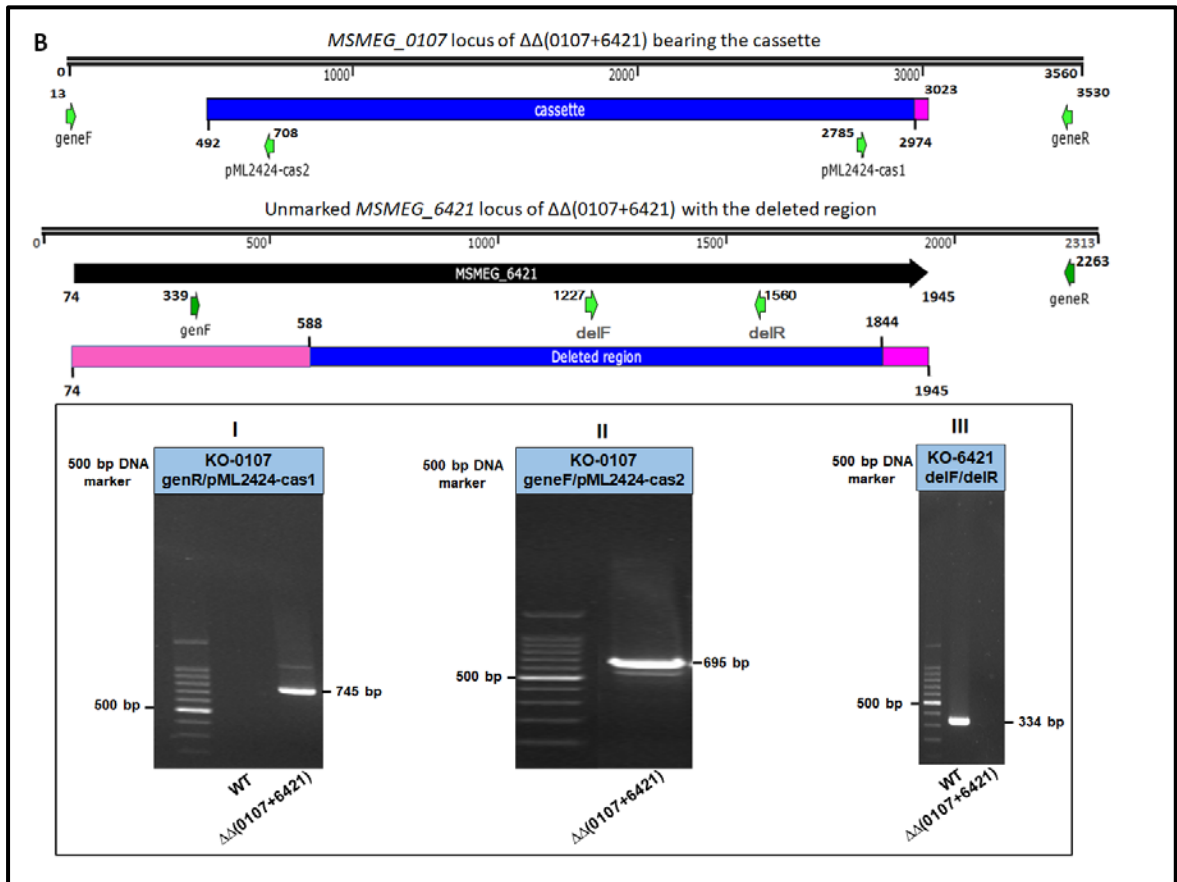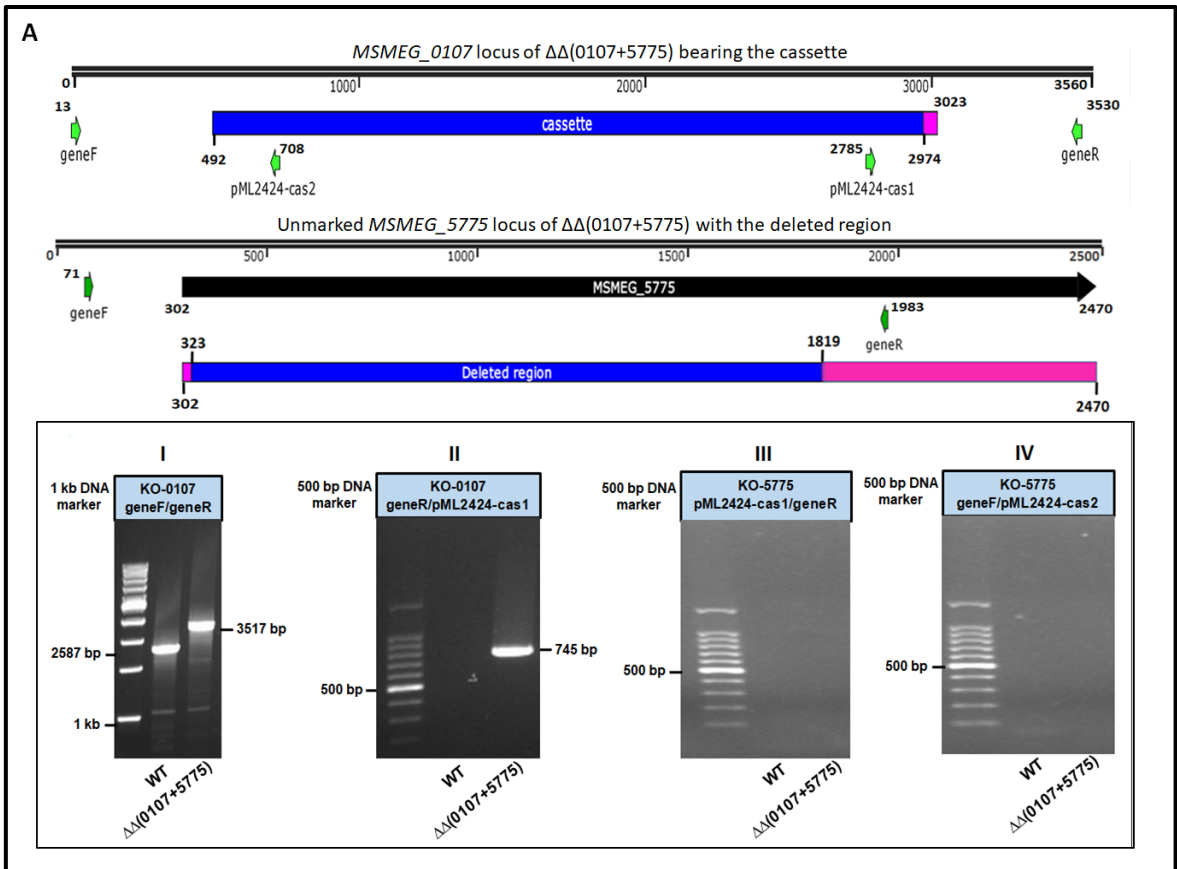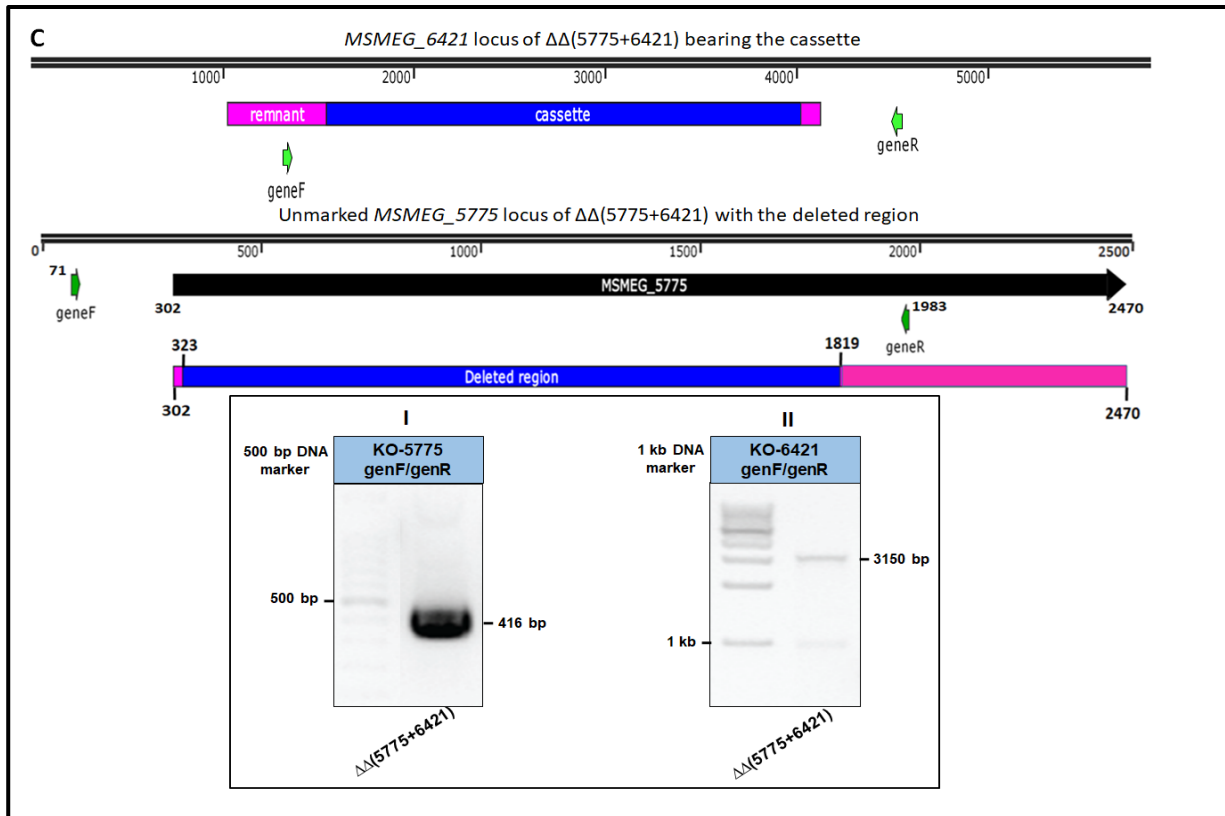
Fig 3.3. Confirmation of the double *lcp* deletion mutants in *Msmeg*. The marked double *lcp* deletion mutants **A.** ΔΔ(0107+5775) **B.** ΔΔ(0107+6421) **and C.** ΔΔ(5775+6421) were created using an unmarked single *lcp* deletion mutant. pML2424-0107 containing the cassette was electroporated into unmarked single deletion mutant Δ5775 and Δ6421 to create ΔΔ(0107+5775) and ΔΔ(0107+6421) respectively, whereas pML2424-6421 was used to construct ΔΔ(5775+6421). The primer pairs used to confirm the deletion by agarose gel electrophoresis, are represented for each mutant. The target genes are represented by black arrow, remnant portions of the undeleted target gene is in pink and the deletion part is represented by blue box. The numbers of start and end site of each gene or primer are in accordance with the full length sequence taken here.
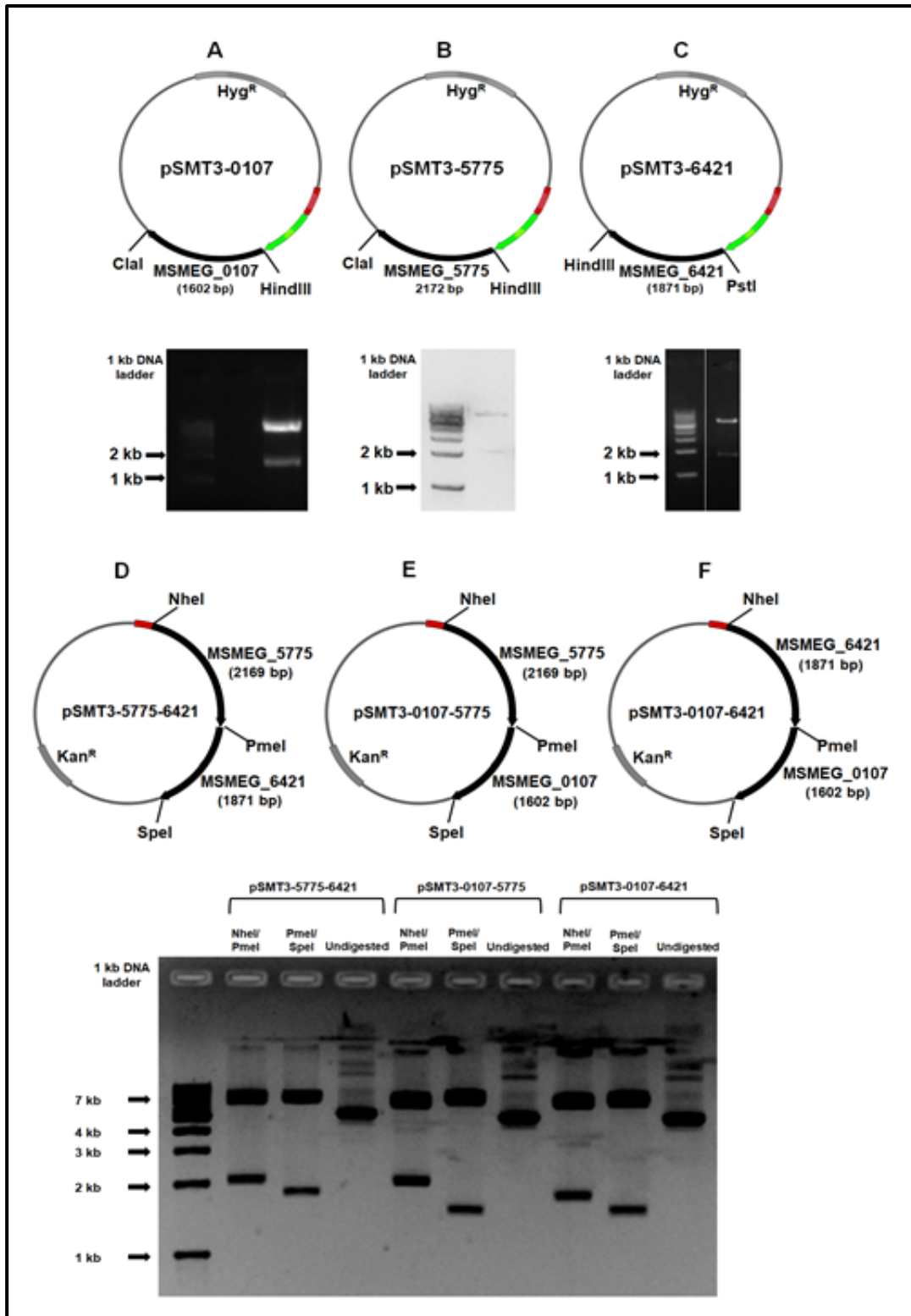
Fig 3.4. Confirmation of the complement constructs by restriction analysis. The complement constructs containing single (**A-C**) and double (**D-F**) *lcp* genes were created using pSMT3 and pSMT3-mspA-LipYtb vector respectively. The recombinant plasmid maps shown here are mapped to scale using Snapgene software. Black arrow(s) on the map represents the target gene cloned into the vector. The size of the respective gene(s) are shown in parentheses. Green arrow represents Green fluorescent protein (GFP) and red box represents Phsp60 promoter. The antibiotic resistance marker gene has been labeled with a grey box. The restriction sites used for cloning the gene(s) have been marked on their respective locations on the plasmid map. The complement constructs were confirmed by restriction analysis on DNA agarose gel, using these restriction enzymes.
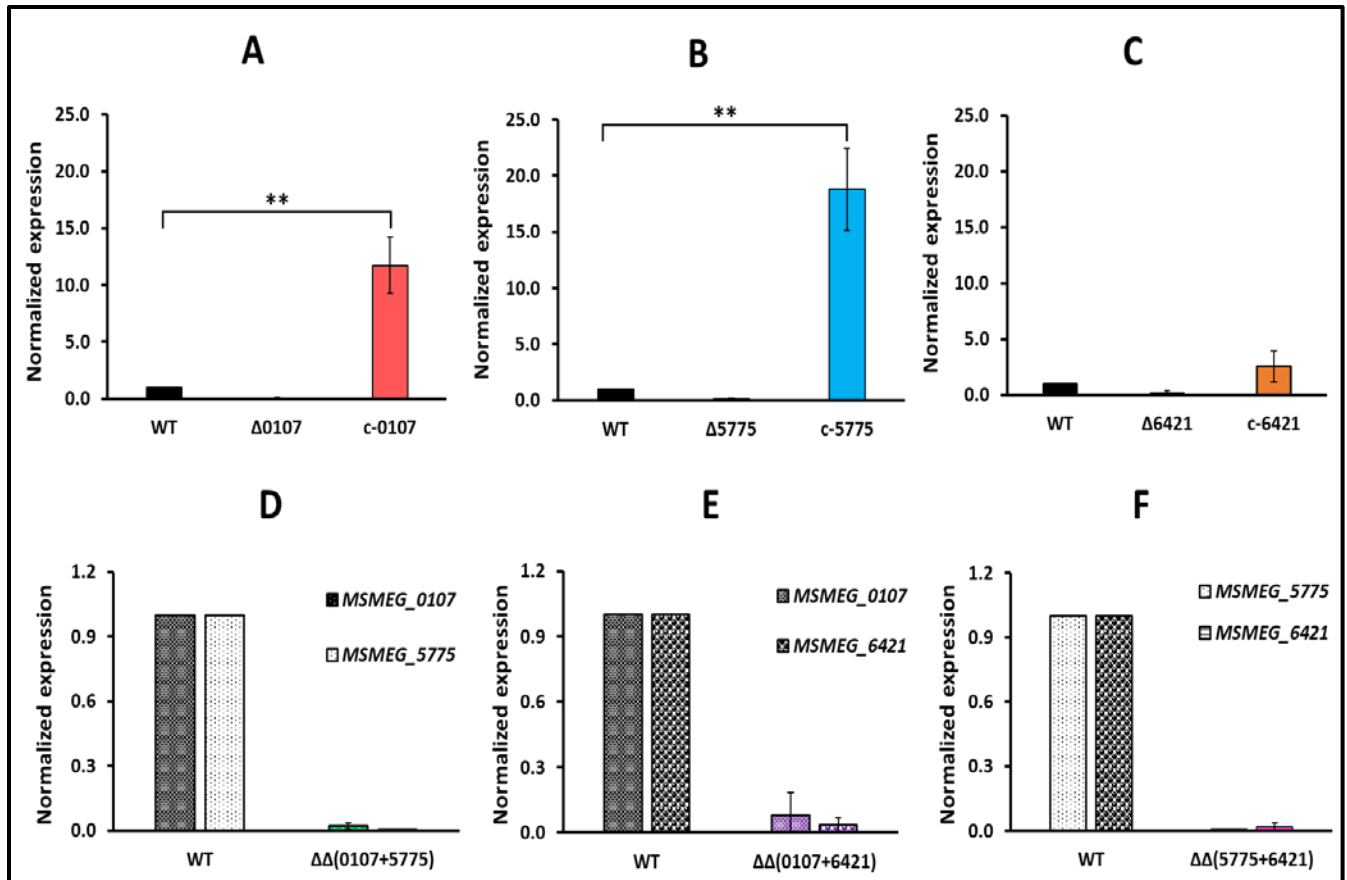
Fig 3.5. Confirmation of *Msmeg* single and double *lcp* deletion mutants by real-time PCR. Quantitative real-time PCR analysis was performed on the single *lcp* deletion mutants. **A.** Δ0107, **B.** Δ5775 and **C.** Δ6421 for their null expression, and on their respective complemented strains to check their phenotypic restoration to wild type levels. The data represented here is an average of five independent experiments each performed in triplicate for Δ0107, Δ5775 and their respective complemented strains whereas, for Δ6421 and its complemented strain, an average of six independent experiments each performed in triplicate is shown. The transcription levels in the double *lcp* deletion mutants ΔΔ(0107+5775), ΔΔ(0107+6421) and ΔΔ(5775+6421) represented here is an average of three independent experiments performed in triplicate **(D-F).** The relative expression values of the mutant and the complemented strains are normalized to wild type expression levels. Error bars denote mean ±SD values between the three independent experiments and asterisks represent statistical significance (** $P \leq .005$).

3.3.2. *Msmeg* double deletion mutant ΔΔ(0107+5775) exhibits slower growth rate

To determine growth rate of the wild type, deletion mutants and their complemented strains, 7H9 medium was used at all times, supplemented with 10% ADS and 0.5% glycerol for enrichment, and 0.05% Tween-80 to prevent cell clumping and the strains incubated at 37°C for 60 hours. While all the single deletion mutants, the double deletion mutants ΔΔ(0107+6421) and ΔΔ(5775+6421), and all of their respective complemented strains had similar growth patterns compared to the wild type, the double *lcp* deletion mutant, ΔΔ(0107+5775) displayed a slower growth rate (Fig 3.6). This suggests that, with the loss of both *lcp* genes *MSMEG_0107* and

MSMEG_5775, the cells divide much slower rate. A 2-4 fold reduced growth rate of this mutant was observed compared to the wild type strain, during the log phase. However, the complementation strains of the double deletion mutants ΔΔ(0107+5775) and ΔΔ(5775+6421) did not revert back to the wild type growth characteristics and exhibited the slowest growth of all. The constitutive *Phsp60* promoter of the plasmid pSMT3 used to generate c-(0107+5775) and c-(5775+6421) that results in an overexpression of the incorporated *lcp* genes compared to the wild type strain (Fig 3.5), might be the cause for this unexpected detrimental effect on the growth of these complemented strains.
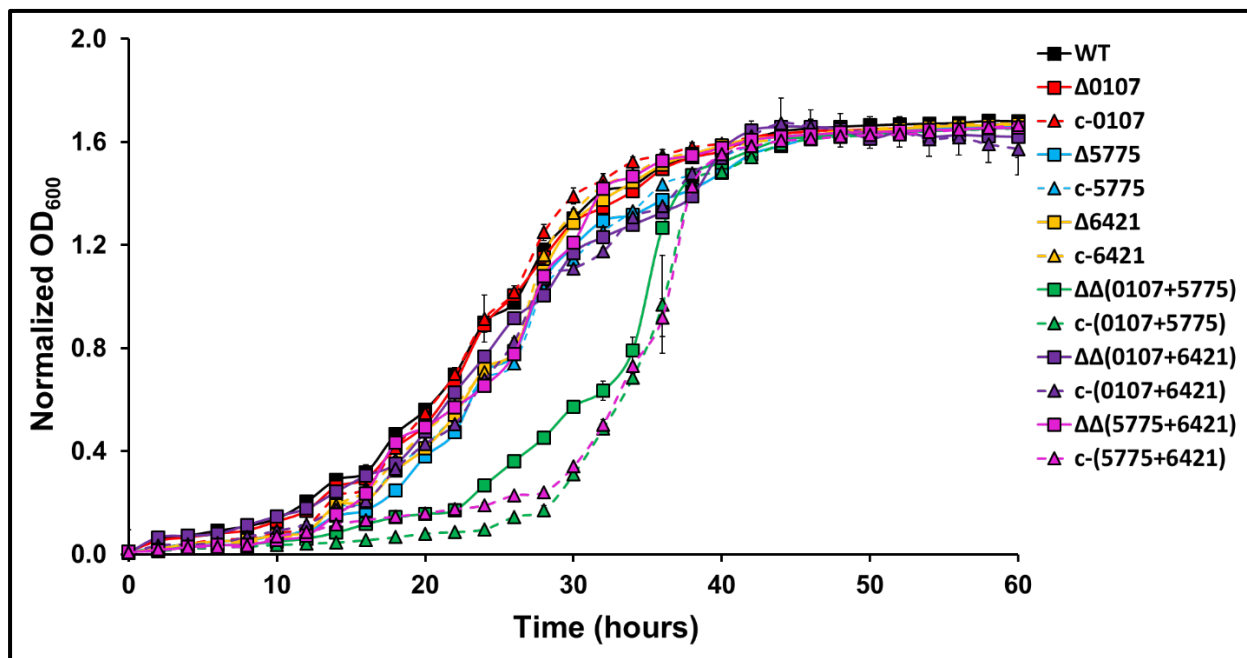
Fig 3.6. Growth curve of the wild type strain and *lcp* deletion mutants. The data represents continuous growth of the strains for 60 hours at continuous shaking conditions at 37°C, and the absorbance recorded every 2 hours. The double *lcp* deletion mutant ΔΔ(0107+5775) displays the slowest growth rate amongst all mutants whereas, its complemented strain c-(0107+5775) along with c-(5775+6421) displays a growth rate even slower than ΔΔ(0107+5775). The error bars represent ± SD of three independent experiments each performed in 3 replicates.

3.3.3. Scanning electron microscopy of the compromised double *lcp* deletion mutant

Based on the detrimental growth curve result of the double *lcp* deletion mutant ΔΔ(0107+5775), it was hypothesized that the cell envelope of this mutant might have been compromised. Hence, we were interested to further investigate the surface alterations of this mutant by SEM. Both the strains were grown to an OD_{600} of 2 before processing them for SEM. Few cells of the mutant had punctured ends while others were over-inflated in the middle or at the terminals (Fig 3.7). Some cells also showed septum-like features either at the centre or at the terminals. These features were absent in the *Msmeg* wild type. In *C. glutamicum*, the layers of the cell envelope of an *lcp* mutant (*Cg0847*, orthologue of *MSMEG_1824*) was reported to be scraped (Baumgart *et al.*, 2016). In our study, a couple of such scraped cells were found in a total of five pictures taken for ΔΔ(0107+5775). However, the *Msmeg* double *lcp* deletion mutant ΔΔ(0107+5775) seemed to be rounder and bulkier than the wild type strain. However, this data was generated from one independent experiment and hence needs to be reproduced. Since, a late-log phase culture was used for this SEM study, it would be worthwhile to find the surface features of this mutant at all growth phases and hence, further investigations are required in this regard.
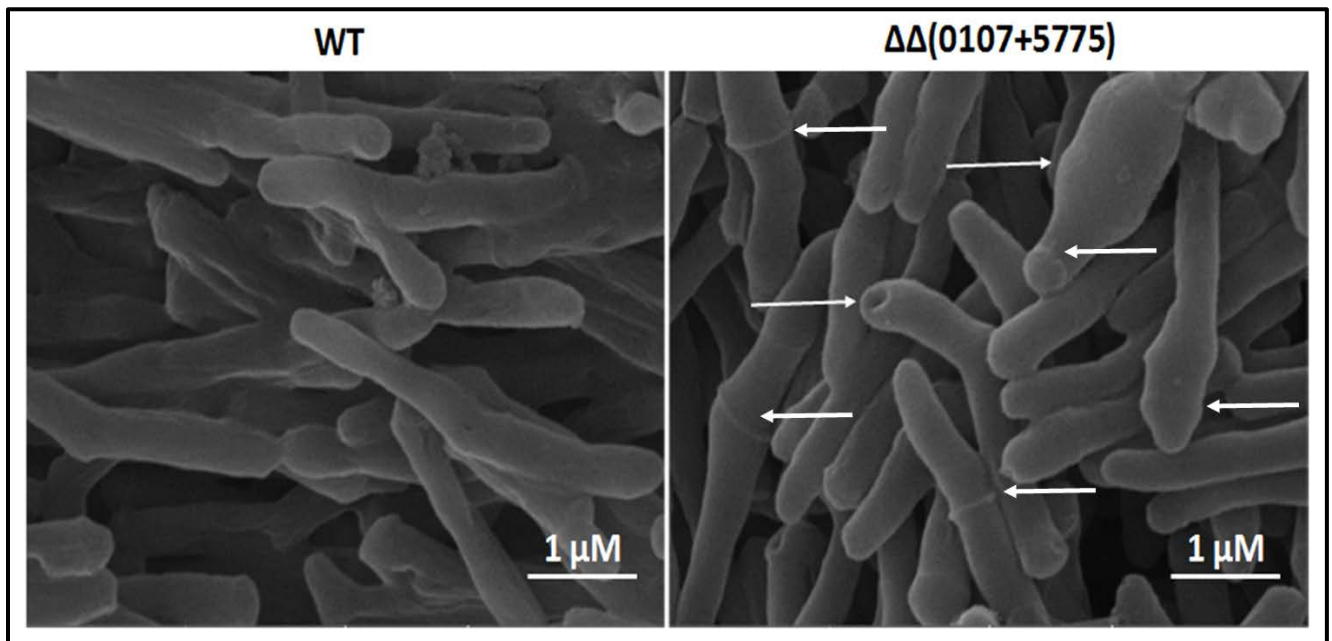
Fig 3.7. Scanning electron microscopy of *Msmeg* wild type and *lcp* double deletion mutant ΔΔ(0107+5775). Images were captured at a high voltage of 20 kV with 40,000 magnification. The arrows indicate the altered cell surface features in the mutant.

3.3.4. The double *lcp* deletion mutant ΔΔ(0107+5775) is altered in colony morphology

To characterize the *lcp* deletion mutants, their colony morphology and diameter of growth were analyzed on Middlebrook 7H10 along with Congo red (Klepp *et al.*, 2012). Congo red not only improves visualization of distinct alterations in cell morphology but also evaluates any changes in the lipid and lipoprotein contents of the cell envelope (Cangelosi *et al.*, 1999). This was advantageous in our study since changes in the cell surface hydrophobicity of the *lcp* deletion mutants were anticipated. On comparing the single *lcp* deletion mutants with the wild type strain mc^{2}155, it was observed that Δ0107 and Δ6421 had no significant visible alterations in their morphology (Fig 3.8, A). They all had common features of i) a dense, dried and complex network of three dimensional crenulated structures interconnected to each other ii) a distinct central zone with a network of smaller crenulated structures iii) an outer zone comprising of a network with larger thread-like structures and, iv) coarse and shrunk uneven edges. In contrast, the colony formed by Δ5775, showed a significant decrease in the crenulated network both in the central and outer zone and, the edges appeared smoother but still uneven. Since we did not find any literature describing these thread-like networks formed in *Msmeg*, these different phenotypes could not be explained easily. However, it is anticipated that these formations are likely to be associated with hydrophobicity properties of the strain, as Congo red dye is known to associate with the lipophilic regions of the mycobacterial cell envelope and delivers a measure of total hydrophobicity of the cell (Jankute *et al.*, 2017). When comparing the double *lcp* deletion mutants with the wild type, striking differences were observed between ΔΔ(0107+5775) and the wild type strain. The dense network of crenulated structures were almost completely absent and the whole colony appeared smooth and moist. This suggests the overall hydrophobicity of this mutant might have been diminished due to the deletion of the genes *MSMEG_0107* and *MSMEG_5775*. The colony of ΔΔ(0107+6421) seemed more dense. This indicates that, *MSMEG_1824* and *MSMEG_5775* might have compensated for the loss of the two other *Msmeg lcp* homologues in this strain and thus were able to maintain the hydrophobicity of the bacteria at levels comparable to that of the wild type. On the other hand, in ΔΔ(5775+6421), the crenulated network was still seen, though less dense and with smooth edges, and the colony had increased by 1.2-folds (Fig 3.8, B). The complemented strain of ΔΔ(5775+6421) could restore the size and dense crenulated features of the wild type phenotype however, it lacked a visible central zone. An explanation for this incomplete restoration to wild type phenotype might be in the constant expression levels of both genes due to the constitutively active Phsp60 promoter instead of differentially regulated expression patterns during different growth phases of the colony like in the wild type strain.

Strikingly, c-(0107+6421) appeared similar to its mutant ΔΔ(0107+6421) but, the measure of the diameter was about 1.5-fold reduced. Altogether, the observed changes in morphology suggest cell envelope alterations due to the loss of *lcp* genes that are also probably associated with the hydrophobicity of the mycobacterial cell. Therefore, the hydrophobic attributes of these mutants were further investigated by examining its biofilm forming ability, which is generally seen in mycobacteria.
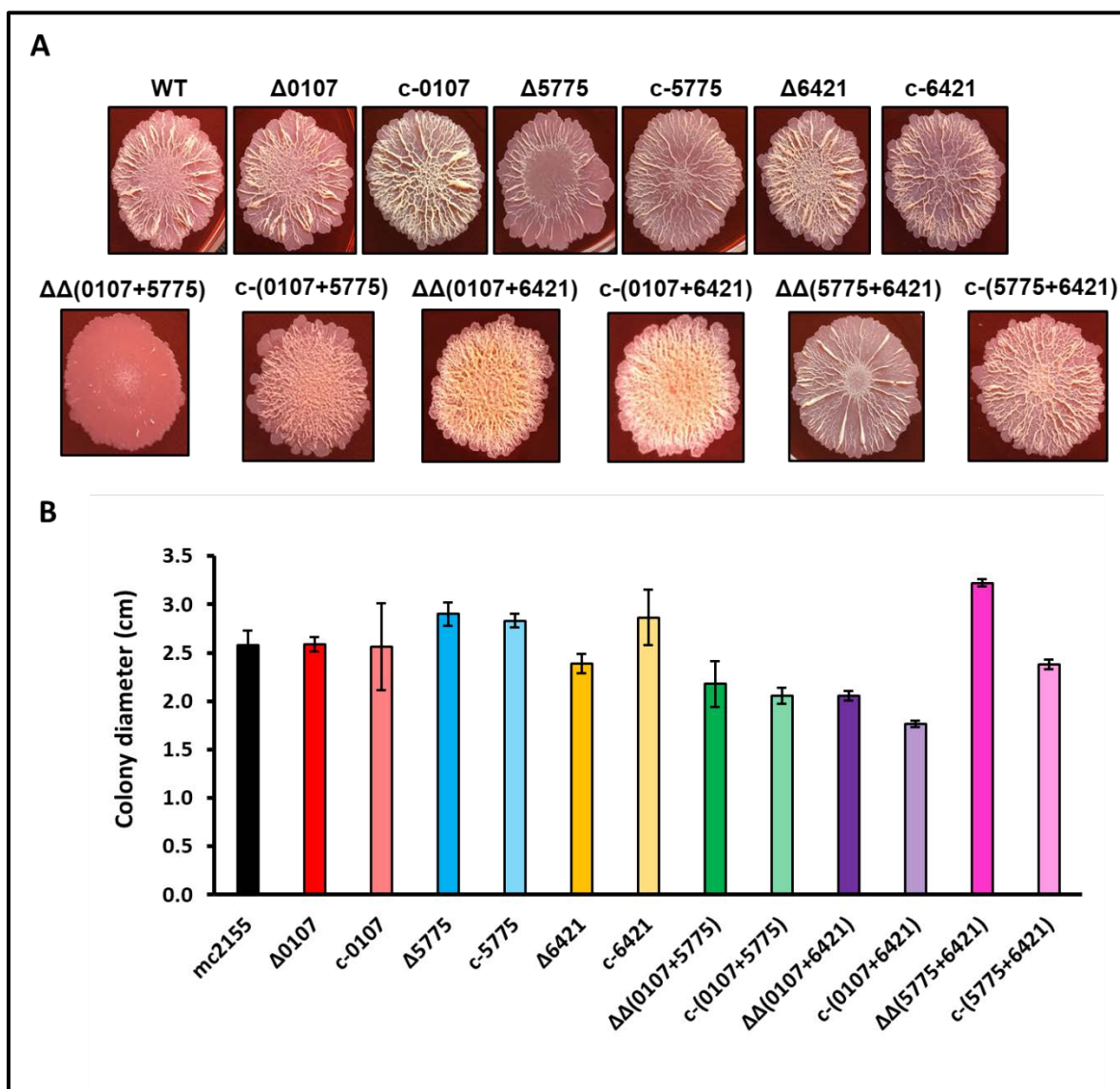
Fig 3.8. Morphology of *lcp* deletion mutants on congo red plates. A. Colony morphology on congo red plates. The plates for the marked double-deletion mutants and the complemented strains were supplemented with appropriate antibiotics (section 3.2.1). In both the experiments, 5 μL of OD_{600} of 0.5 adjusted overnight culture is pipetted onto the plate, and incubated in sealed plastic bags for 3-4 days. **B.** Quantification of colonies by measuring the diameter of each colony from two perpendicular angles using a ruler, and the average recorded. The data shown is the average of three independent experiments, each done in triplicates. The error bars indicate standard deviation.

3.3.5. ∆∆(0107+5775) exhibits enhanced biofilm adherence on plastic but diminished biofilm formation on air-liquid interface

A biofilm is described as a thin and slimy layer of bacterial cells that either adhere to each other or to a solid surface. The formation of bacterial biofilms is believed to contribute to virulence and acts as a survival mechanism (Lai *et al.*, 2018). Bacterial biofilms consist of extracellular polymeric elements such as proteins, nucleic acids, lipids and polysaccharides (Sousa *et al.*, 2015), that act together as a physical barrier to antimicrobial agents (Shi *et al.*, 2011). They can be formed on solid surfaces because of the bacteria's ability to adhere to abiotic surfaces such as polystyrene or polyvinyl chloride surfaces, or at the air-liquid interface called as pellicles (Donlan, 2002).

Since an altered colony morphology of the double *lcp* deletion mutant ∆∆(0107+5775) on Congo red was observed, it was hypothesized that the crenulated structures which were lost in this mutant could be features associating to its hydrophobicity. Hence, this mutant was investigated for its ability to form biofilm. It was observed that this mutant had lost its ability to form biofilm on the air-liquid interface, whereas ∆∆(5775+6421) showed an enhanced biofilm formation (Fig 3.9, A). A distinct feature of the mycobacterial cell wall is that around 60% of its weight comprises of lipids, which mainly includes long chain fatty acids containing 60 to 90 carbons, also called mycolic acids (Bendinger *et al.*, 1993, Jarlier & Nikaido, 1994). The extreme hydrophobicity of mycobacterial cells is rendered by this high lipid content, which is also responsible for impermeability of many chemical disinfectants and some antibiotics into the cell (Borrego *et al.*, 2000), and also confining uptake of nutrients into the organisms, thus contributing to slower growth rate of mycobacteria (Draper, 1984). This again indicates reduced hydrophobicity in the mutant that lacks both *MSMEG_0107* and *MSMEG_5775*. Therefore, it was speculated that a deletion of *MSMEG_0107* and *MSMEG_5775* might have had a negative effect on the amount of lipids on the mycobacterial cell envelope, especially in the mycolic acid layer which has been implicated in this pellicle (biofilm on air-liquid interface) formation (Ojha *et al.*, 2005). Similarly in ∆∆(5775+6421), *MSMEG_1824* and *MSMEG_5775* might have overexpressed and tried to compensate the loss of the other two genes. From this observation, it seems that the three *lcp* genes compensate the loss of one another. However, when both *MSMEG_0107* and *MSMEG_5775* are absent from the genome, the organism loses its ability to form biofilm, suggesting that these two genes together are important for maintaining the cell wall integrity. The biofilm formation was further quantified in a polystyrene 96-well plate using 1% crystal violet. There was about 8-fold increase (*P*=.001) in the amount of biofilm adhering to the walls of the plate by ∆∆(0107+5775), and about 4.5-fold increase by ∆∆(5775+6421) compared to the wild type (*P*=.002) (Fig 3.9, B). The crystal violet is known to bind to the hydrophobic features of the

mycobacterial cell envelope and therefore, it is suspected that an increased biofilm adherence to the polystyrene plate by the double *lcp* deletion mutant ΔΔ(0107+5775), is a result of increased GPL levels in the *Msmeg* cell envelope associated with the loss of both *MSMEG_0107* and *MSMEG_5775*. The complemented strains for these two double *lcp* deletion mutants showed a decreased adherence to the well, but did not restore to wild type levels completely. The single *lcp* deletion mutant Δ0107 showed enhanced and large aggregates of biofilm, whereas Δ5775, Δ6421, ΔΔ(0107+6421) and ΔΔ(5775+6421) displayed overall increase in biofilm formation on the air-liquid interface (inlet of Fig 3.9, A and Appendix). However, there was no significant change in the biofilm adherence of the two single deletion mutants, in the 96-well plate. This discrepancy in biofilm formation on the air-liquid interface and on the walls of the 96-well plate is due to the association of pellicle (biofilm formation on air-liquid interface) with lipids within the mycolic acid layer (Ojha *et al.*, 2005), whereas the biofilm adherence on polystyrene 96-well plate has been mostly correlated to the amount of GPLs present in the cell wall of non-tubercle mycobacterial species (Recht *et al.*, 2000, Recht & Kolter, 2001, Schorey & Sweet, 2008). Thus, it would be interesting to investigate the association between these surface hydrophobic features and the LCP family of proteins. One fastest and easiest way to establish this correlation is by investigating the relative expression of genes encoding GPL proteins in *lcp* deletion mutants. We propose to investigate this in future. Apart from this, reduced biofilm formation have been previously established to be associated with a defective extracellular matrix and a defective induction of short-chain mycolic acids (Mathew *et al.*, 2006), as well as with increased aggregation (Gupta *et al.*, 2015). Therefore, the aggregative properties of the *lcp* mutants were further investigated to confirm whether they are in line with the observed hydrophobicity.
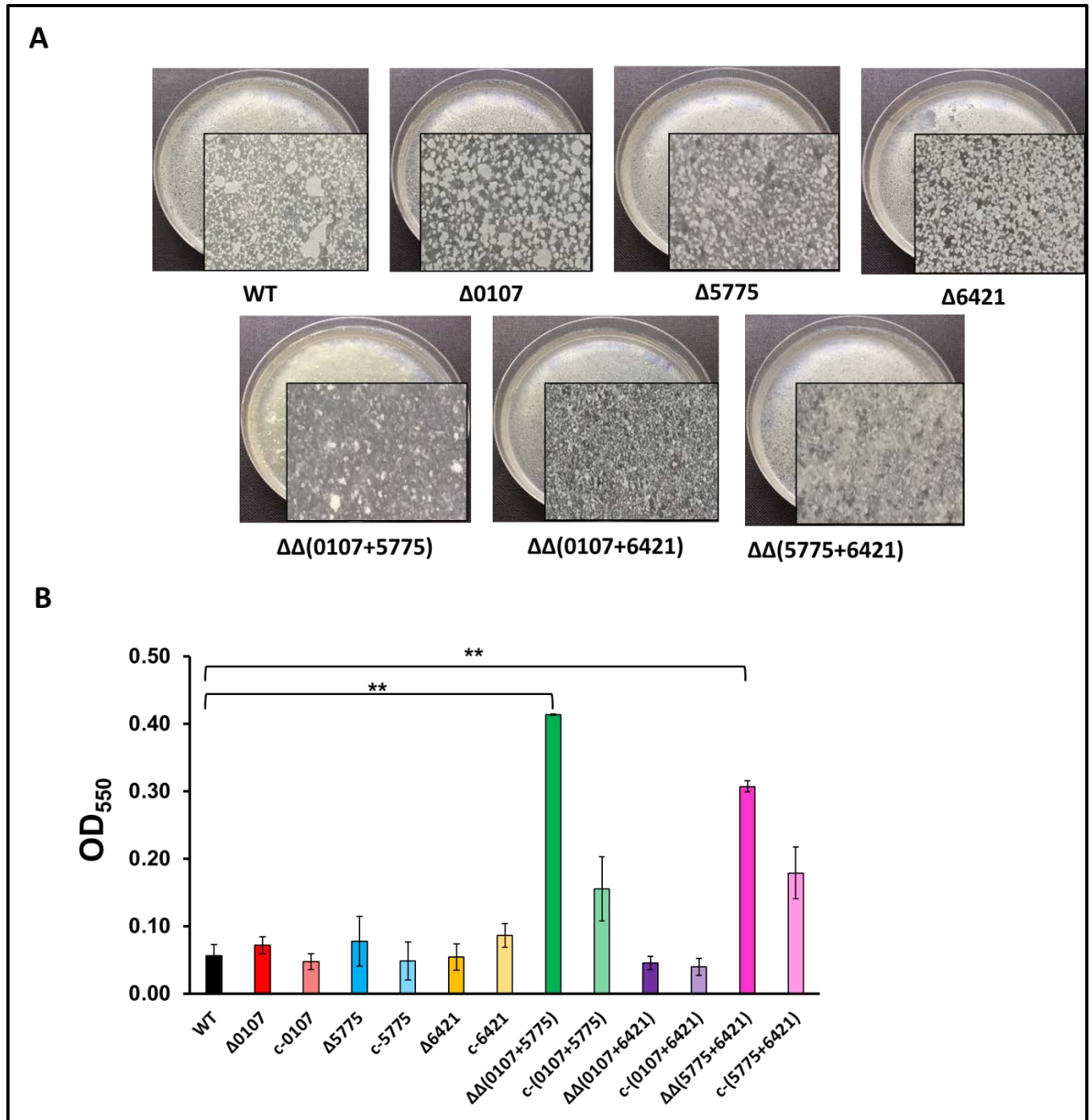
Fig 3.9. Biofilm determination. A. Air-liquid interface biofilm determination in wild type and *lcp* deletion mutants. The biofilm observed on the surface of 7H9 medium supplemented with ADS but without Tween-80, is after incubation of the plates at 37°C for 4 days. The data is a representative of one of the three independent experiments. **B.** Quantification of adhering biofilms with 1% crystal violet solution. The data represented here is a quantification of the adherent biofilms formed in a Polystyrene 96-well plate after 5 days of incubation at 37°C. The amount of biofilm accumulated by the bacteria is determined by measuring the absorbance at OD_{550}. The data is an average of three independent experiments performed in at least 3 replicates in each experiment. The error bars represent standard deviation and asterisks represent statistical significance (**, $P \leq .005$).

3.3.6. Deletion of both *MSMEG_0107* and *MSMEG_5775* diminishes aggregative properties

Cellular aggregation is demarcated by the accumulation of cells to form fairly stable, interconnected, multicellular associations under physiological conditions (Borrego *et al.*, 2000). The hydrophobic attribute of the cell envelope contributes to cell clumping and formation of large cell aggregates in mycobacteria, thus allowing to settle down when kept stationary. The turbidimetric method used to measure the cell aggregation displayed about 2-fold and 2.3-fold slower aggregation rate of the single deletion mutant Δ5775 and the double *lcp* deletion mutant ΔΔ(0107+5775) compared with the wild type strain, at T=5 min (Fig 3.10, A and B). In a complete time window of 15 min, ΔΔ(0107+5775) settled down slower at a uniform rate until T=10 min after which it remained stationary. On the other hand, Δ5775 settled down even slower throughout, and mostly appeared to remain uniformly dispersed in the suspension until the last minute. The single deletion mutants, Δ0107 and Δ6421, showed higher aggregation rate than the wild type after T=2 min indicating higher cell surface hydrophobicity in these strains. This is in congruence with the biofilm formation of these two single deletion mutants on the air-liquid interface. Though the aggregation rate of the double deletion mutant ΔΔ(0107+6421) was slower than the wild type initially (until T=5 min), the rate became faster than the wild type thereafter. However, ΔΔ(5775+6421) aggregated slower and parallel to the wild type throughout. It was also interesting to notice that all the complemented strains of the double deletion mutants aggregated faster than the wild type after about T=4 min. Since such turbidimetric assays are prone to higher standard deviations when reproduced in multiple experiments, one such representative experiment has been presented here from two independent experiments.

GPLs in the envelope has been previously reported to inhibit aggregation (Yang *et al.*, 2017). Though our data does not include major studies on GPLs, however it is evident that the double deletion mutant ΔΔ(0107+5775) with an enhanced biofilm formation on polystyrene plate has diminished aggregative properties. This is in congruence with a previous study where reduced biofilms have been associated with increased aggregation (Gupta *et al.*, 2015).
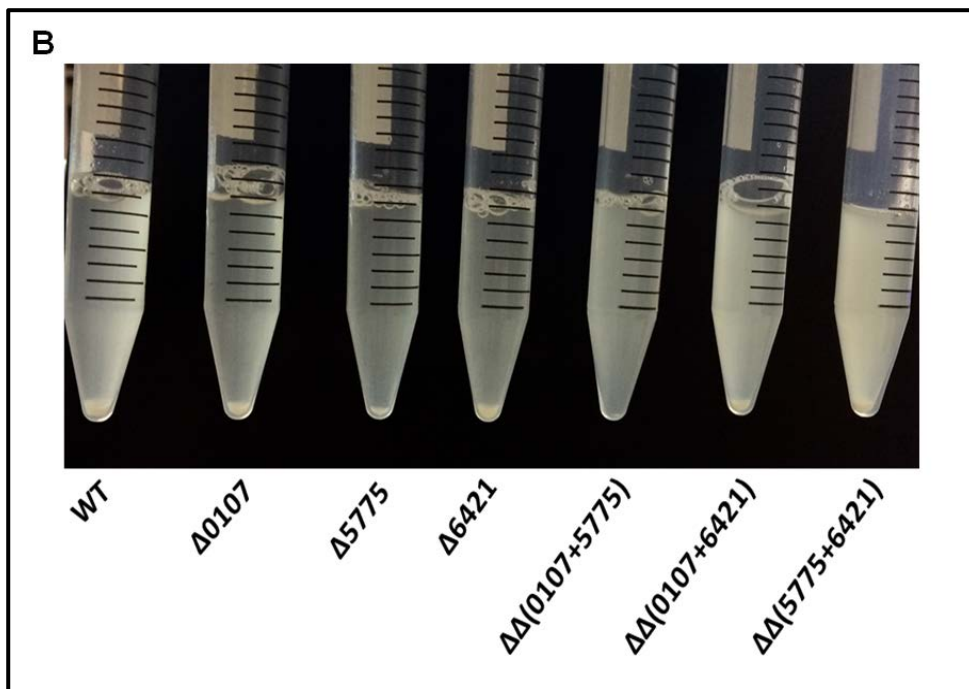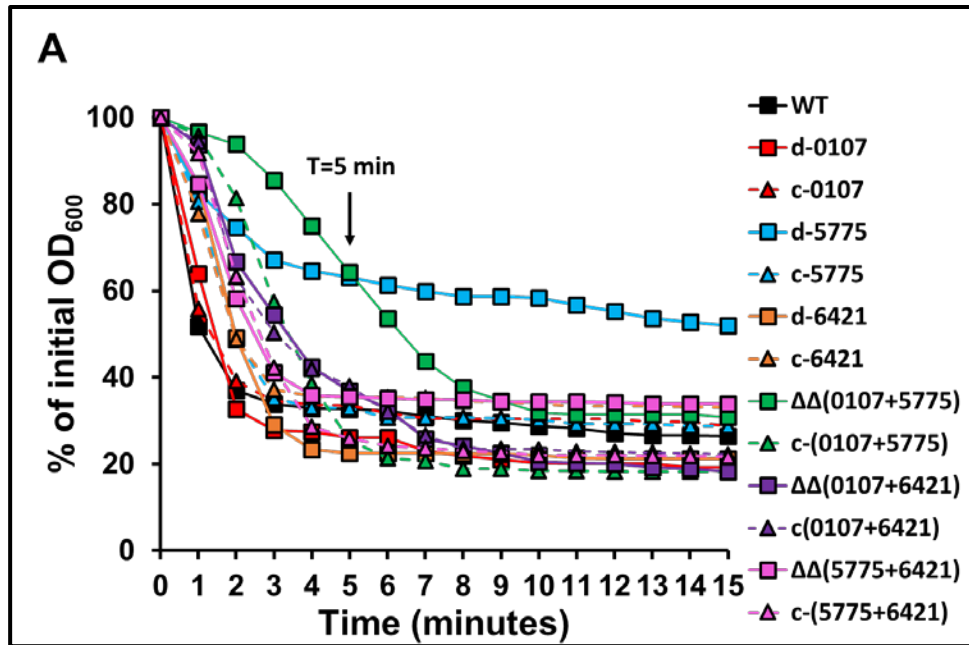
Fig 3.10. Aggregative properties of *lcp* deletion mutants. A. The aggregation plot is derived by measuring the OD_{600} of PBS-washed bacterial cells grown without Tween-80. The starting OD_{600} is adjusted to 1, and the data is normalized to percent reduction at time point zero in each well. The arrow represents the time point (T=5min) at which the image is clicked in Fig 3.10, B. **B.** Visual display of the aggregation assay using the wild strain and the mutants. The image displayed is taken at the 5th minute i.e. T=5 min. This is a representative experiment out of two independent experiments.

3.3.7. ∆∆(0107+5775) display elevated levels of glycopeptidolipids (GPLs) in thin layer chromatography

Reduction in biofilm formation and changes in other macroscopic surface properties of the double deletion mutant strain ∆∆(0107+5775), strongly indicated alterations in the lipid composition of the cell walls of these strains. Both the polar and apolar lipids were isolated from supernatant of the bacterial samples by chloroform-methanol fractionation, and subsequently analyzed by TLC. Since, the double *lcp* deletion mutant ∆∆(0107+5775) had a very slow growth rate, it was grown longer until the OD_{600} reached 2, and subsequently diluted it to an OD_{600} of 0.5 before isolating the polar and apolar lipids from this strain. It was found that, except for ∆6421, the level of GPLs seemed higher in the other two single deletion mutants, ∆0107 and ∆5775. The double *lcp* deletion mutant ∆∆(0107+6421) has not been used for this experiment, as this mutant was not constructed when this experiment was conducted. With the two available double deletion mutants for this experiment, higher GPLs in both these mutants were observed, with ∆∆(0107+5775) displaying the highest level of GPLs amongst all (Fig 3.11). The GPLs are a major class of lipids present on the outermost portion of the non-tubercle mycobacterial cell envelope, which are absent in *Mtb* (Miyamoto *et al.*, 2006). *Msmeg* produces both polar and apolar GPLs (Schorey & Sweet, 2008) also known as serovar specific and nonserovar-specific GPLs respectively (Miyamoto *et al.*, 2006, Schorey & Sweet, 2008). Elevated levels of GPLs in the double *lcp* deletion mutant ∆∆(0107+5775), led us to initially hypothesize that the cell of this mutant may be more hydrophobic. A previous study describes apolar GPLs that are present on the surface of *Msmeg* to be responsible for sliding motility and biofilm formation whereas, the presence of polar lipids such as triglycosylated GPLs have been associated with cell surface alterations resulting in reduced sliding motility as well as changes in cellular aggregation and colony appearance on Congo red (Deshayes *et al.*, 2005). The polar GPLs vary from the apolar GPLs by the addition of a rhamnosyl residue (triglycosylated GPLs) or a succinyl residue (succinylated GPLs) linked to the terminal rhamnosyl unit (Vergne *et al.*, 1995, Mills *et al.*, 2004). GPLs have also been associated with the initial attachment of mycobacterial cells to substratum like polyvinyl chloride (PVC) (Recht & Kolter, 2001). Although polystyrene plate was used for biofilm quantification in our study, the association of GPLs with PVC can be related considering it to be a similar substratum for biofilm attachment. Therefore, with an enhanced biofilm formation on polystyrene plate and elevated GPLs observed in the mutant ∆∆(0107+5775) in this preliminary experiment, it may be hypothesized that the deficiency of both *MSMEG_0107* and *MSMEG_5775* genes from *Msmeg* results in an increased GPL formation in this mutant. However, this has to be further investigated to draw a firm association of the *lcp* genes with GPLs. It may be hypothesized that

the distinct physicochemical characteristics of the GPLs might have altered the nature of interaction with the environment, resulting in the various phenotypes related to hydrophobicity observed in this study. Apart from this, it was also found that Δ6421 displayed similar intensity of the GPLs compared to the wild type in the TLC plate, whereas a reduced intensity of GPLs compared to rest of the mutants. This is congruent to the biofilm quantification results where both Δ6421 and the wild type strain display similar biofilm adherence compared to each other and reduced adherence compared to other mutants, as detected by crystal violet (Fig 3.9, B). However, more number of reproducible experiments on these mutants are required to confirm these preliminary observations. Moreover, investigating the lipid profile of these mutants at different growth phases would contribute information on the impact of *lcp* genes in lipid biosynthesis, at different stages of growth.
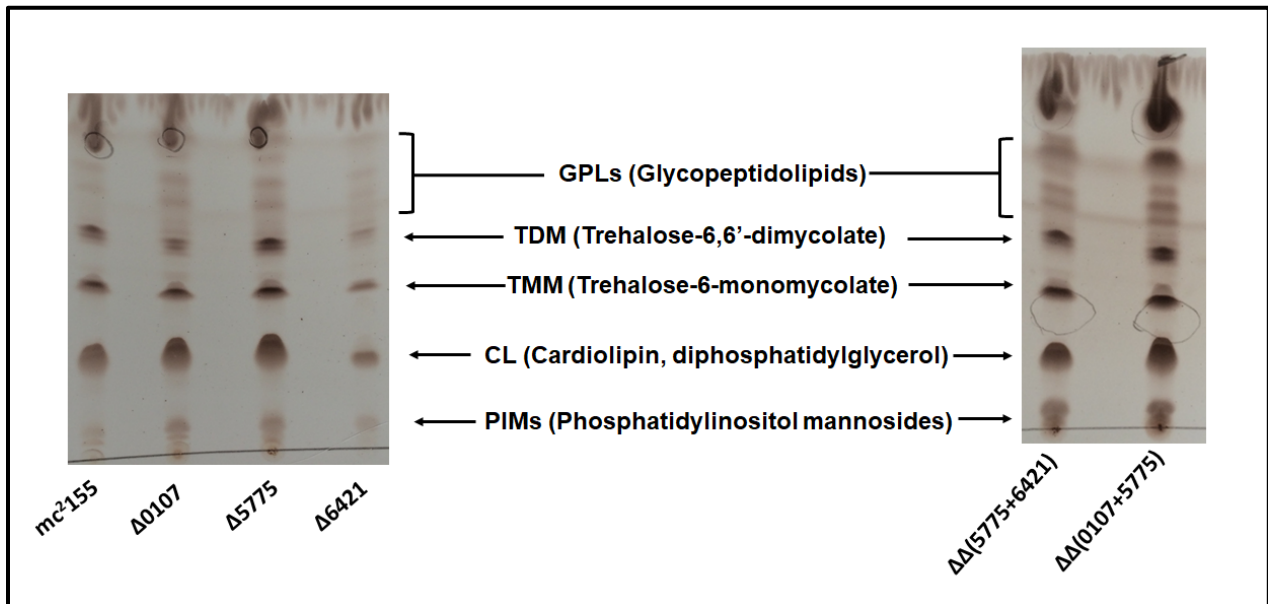
Fig 3.11. Thin Layer Chromatography of *lcp* deletion mutants. The polar and apolar lipids were extracted from the wild type strain and the *lcp* deletion mutants. The lipids were developed in the solvent system containing chloroform:methanol:water (20:4:0.5) and stained with a solution containing 8% aqueous phosphoric acid and 10% copper (II) sulfate. The TLC silica plates were subsequently charred to reveal the lipids. The different lipid fractions obtained have been labeled and compared according to the unpublished protocol by Nicholas P. West (University of Queensland). The presented data is a preliminary experiment to investigate the lipid profile in *Msmeg lcp* mutants.

3.3.8. The double *lcp* deletion mutant ΔΔ(0107+5775) is highly sensitive to antibiotics and other envelope stress conditions

The lipid rich mycobacterial cell envelope acts as an impermeable barrier to numerous chemicals and antibiotics (Jarlier & Nikaido, 1994), thus contributing to the pathogenesis of the organism (Brennan, 2003). Since loss of *lcp* genes compromise the integrity of cell envelope (Kawai *et al.*, 2011), it was thus interesting to investigate the susceptibility of these mutants to various first line anti-TB antibiotics by a resazurin based assay (Agrawal *et al.*, 2015). Resazurin is a fluorogenic redox dye, which has been used to test the antimicrobial activity of many compounds. After exposing the bacteria to different antibiotic concentration for 48 hours, resazurin was added to detect the metabolic activity of the cells. Reduction of this dye changes its color from blue with little fluorescence to highly fluorescent pink, and this conversion, which could be quantified by the fluorescence, indicates the extent of bacterial viability (Palomino *et al.*, 2002).

An 8-fold increased susceptibility to cell wall targeting antibiotics such as vancomycin and bacitracin (Table 3.5) was observed in the double deletion mutant ΔΔ(0107+5775), when compared to the wild type *Msmeg* strain. In mycobacteria, Bacitracin (BAC) binds to the decaprenyl pyrophosphate (C_{50}-PP) (Stone & Strominger, 1971) which functions as a carrier lipid in transporting PG components across the cell membrane (Valvano, 2008). In the presence of BAC, this transport is inhibited because BAC prevents the cleavage of the pyrophosphate and consequently inhibits cell wall biosynthesis (Stone & Strominger, 1971, Ming & Epperson, 2002, Economou *et al.*, 2013, Kingston *et al.*, 2014). This indicates that ΔΔ(0107+5775) which seems to have a compromised cell envelope already due to inefficient linkage of the AG and PG, further succumbs to the inhibitory effect of BAC by preventing dephosphorylation of the decaprenyl pyrophosphate and thus inhibiting the recycling of decaprenyl phosphate lipid carrier (Qi *et al.*, 2008). Although a 2-fold enhanced susceptibility to BAC is observed in Δ5775, Δ6421, ΔΔ(5775+6421) and ΔΔ(0107+6421), Δ0107 remained unaffected compared to the wild type strain. This might indicate the compensatory effects of MSMEG_1824, MSMEG_5775 and MSMEG_6421 in Δ0107 to remain uninhibited by BAC activity.

The peptidoglycan targeting antibiotic VAN binds to the **D**-alanyl–**D**-alanine carboxyl terminus of cell wall precursor molecules thus preventing the transpeptidase from acting on these newly formed blocks and ultimately preventing the cross-linking of the peptidoglycan (Walsh, 2000). This results in a less rigid and more permeable peptidoglycan layer which causes cellular contents of the bacteria to leak out and thus destroying the organism. With the deletion of both MSMEG_0107 and MSMEG_5775 in the double *lcp* deletion mutant ΔΔ(0107+5775), the already compromised

AG-PG linkage further makes it easier for the antibiotic to act efficiently. Although single *lcp* deletion mutants Δ0107 and Δ6421 showed no alterations to VAN susceptibility, but the double *lcp* deletion mutant devoid of both Δ0107 and Δ6421 displayed a 4-fold increased susceptibility compared to the wild type strain, and so does the double *lcp* deletion mutant ΔΔ(5775+6421). However, the single *lcp* deletion mutant Δ5775 displayed a 2-fold sensitivity to VAN compared to the wild type strain. This suggests that a deficiency of both MSMEG_0107 and MSMEG_5775 or, both MSMEG_5775 and MSMEG_6421 in *Msmeg* leads to an enhanced inhibitory effect of VAN on the transpeptidation of **D**-alanyl–**D**-alanine peptidoglycan precursors.

Apart from BAC and VAN that targets the cell wall biosynthesis, common β-lactam antibiotic such as PEN was also used in our study. A 4-fold sensitivity to PEN was observed in the double *lcp* deletion mutant ΔΔ(0107+5775) compared to the wild type strain. PEN targets the cell wall by disrupting the β-(1,4) linkage between *N*-acetylmuramic acid and *N*-acetylglucosamine, as well as by breaking the cross-linking peptide chains. β-lactam antibiotics which are structural mimics of the **D**-Ala-**D**-Ala terminus of the stem pentapeptide that is conserved in Lipid II stem peptides, react with the serine active-site of transpeptidases to inactivate them (Tipper & Strominger, 1965). This indicates that, ΔΔ(0107+5775) with a compromised AG-PG bonding is hyper sensitive to β-lactam antibiotics too. Apart from this, LCP proteins in *S. aureus* have also been shown to be essential for cell division, autolysis, and β-lactam resistance (Over *et al.*, 2011). For instance, in *S. aureus*, β-lactam resistance was shown to be contributed by MsrR (Rossi *et al.*, 2003), whereas both MsrR of *S. aureus* and Psr of *Enterococcus faecalis* were described to augment virulence in *Caenorhabditis elegans*, a model host organism (Bae *et al.*, 2004, Maadani *et al.*, 2007).

An 8-fold increased susceptibility of ΔΔ(0107+5775) to rifampicin and erythromycin, a 4-fold increased susceptibility to penicillin and novobiocin, and a 2-fold increased susceptibility to ethambutol (EMB), isoniazid (INH), azithromycin (AZT) and streptomycin (STR) was also observed. However, as rifampicin, erythromycin and novobiocin are not cell wall-targeting antibiotics, increased sensitivity of ΔΔ(0107+5775) to these antibiotics could be due to enhanced permeability of the drug to the mycobacterial cytoplasm to target the DNA-dependent RNA polymerase, protein synthesis and DNA gyrase respectively. Similar to the observation by Grzegorzewicz and colleagues (Grzegorzewicz *et al.*, 2016), where Rv3484 mutant and the wild type *Mtb* reported no differences in their susceptibility to rifampicin, our study also confirms this outcome with MSMEG_0107 (homologue of Rv3484) and wild type being indifferent to RIF sensitivity.

The deletion of either MSMEG_0107 or MSMEG_5775 showed resistance to INH, an inhibitor of mycolic acid synthesis, by 8-folds in Δ0107 and 2-folds in Δ5775. Additionally, Δ5775 also showed a 2-fold increased resistance to AZT. Though a deletion of *MSMEG_0107* did not show susceptibility to a wide range of antibiotics used here except STR (by 2-fold), however loss of *MSMEG_5775* had a 2-fold increased susceptibility to the cell wall targeting antibiotics such as BAC, VAN and EMB. A 2-fold enhanced susceptibility to cell wall synthesis inhibitors such as EMB and BAC was also observed by strains lacking *MSMEG_6421*. Apart from this, the complemented strain carrying both *MSMEG_0107* and *MSMEG_6421* completely restored the wild type phenotype whereas, other complemented strains of both single and double deletion mutants partly restored the phenotype. From this study, it was found that a severely compromised cell envelope in *Msmeg* with the loss of both *MSMEG_0107* and *MSMEG_5775* makes the organism permeable to antibiotics however, the peptidoglycan biosynthesis targeting drugs severely affected this mutant strain ΔΔ(0107+5775). Although the essential gene *MSMEG_1824* may prove to be a better drug target, however, the two *lcp* genes *MSMEG_0107* and *MSMEG_5775* together could be considered an alternative target for formulation of drugs.

Table 3.5: Determination of MIC of antibiotics used for *Msmeg* wild type, mutants and their complemented strains

Strain	MIC (µg/mL)											Resistance
	PEN	EMB	INH	VAN	BAC	NOV	RIF	AZT	ERY	STR		2-fold decrease
WT	300	1.25	7.81	20	250	6.25	20	1.56	12.5	0.3125		4-fold decrease
Δ0107	300	1.25	62.5	20	250	6.25	20	1.56	12.5	0.16		≥8-fold decrease
c-0107	300	1.25	62.5	20	250	3.125	20	1.56	12.5	0.16		
Δ5775	300	0.625	15.625	10	125	3.125	10	3.125	6.25	0.3125		
c-5775	300	1.25	7.81	20	250	3.125	20	1.56	6.25	0.16		
Δ6421	300	0.625	7.81	20	125	6.25	20	1.56	12.5	0.3125		
c-6421	300	1.25	15.625	20	250	6.25	20	1.56	12.5	0.16		
ΔΔ(0107+5775)	70	0.625	3.91	2.5	7.81	1.5625	1.25	0.78	1.5625	0.16		
c-(0107+5775)	300	0.625	3.91	15	15.625	3.125	15	1.56	12.5	0.3125		
ΔΔ(0107+6421)	150	0.625	7.81	5	125	6.25	15	1.56	12.5	0.16		
c-(0107+6421)	300	1.25	7.81	20	250	6.25	20	1.56	12.5	0.3125		
ΔΔ(5775+6421)	150	0.625	7.81	5	125	3.125	10	3.125	12.5	0.3125		
c-(5775+6421)	300	0.625	3.91	15	15.625	3.125	15	1.56	1.5625	0.16		

PEN, penicillin; EMB, ethambutol; INH, isoniazid, VAN, vancomycin; BAC, bacitracin; NOV, novobiocin; RIF, rifampicin; AZT, azithromycin; ERY, erythromycin; STR, streptomycin.

Based on the results described so far that suggests severe alteration of the cell envelope in the double *lcp* deletion mutant ΔΔ(0107+5775), it was investigated whether this mutant is also hypersensitive to cell wall targeting agents such as detergents like SDS. SDS is generally used to mimic surfactant stress that *Mtb* might encounter in macrophages during host infection (Johansson & Curstedt, 1997). Hence, the effect of this cationic detergent in *Msmeg* was tested and a significant 1.4-fold increased sensitiveness (P=.003) was found in the double deletion mutant ΔΔ(0107+5775) at both 8% and 10% concentration SDS (Fig 3.12, A). Like in many other experiments, the sensitivity to SDS was not lowered to wild type levels in the complemented strain of this mutant. This indicates that the expression levels of the LCP proteins need to be balanced properly to be able to form the wild type mycomembrane that can function most optimal as a permeability barrier. While Δ6421 strain displayed a slight sensitivity of 1.1-fold *to* 8% SDS (P=.01), the two double deletion mutants lacking *MSMEG_6421* showed 1.3-fold increased susceptibility to both 8% (P=.002) and 10% (P=.007) SDS.

The increased susceptibility of the double deletion *lcp* mutant ΔΔ(0107+5775) to a wide range of antibiotics suggested that deletion of both *MSMEG_0107* and *MSMEG_5775* disturbed the assembly of the *Msmeg* cell envelope. Therefore, this mutant is expected to attenuate its resistance against some external stresses. Lysozyme is such an antimicrobial factor that induces lytic stress in host phagolysosomes (Michailova *et al.*, 2000) during mycobacterial infections. It compromises the cell wall in mycobacteria by cleaving the β-1,4-glycosidic bonds of the PG backbone (Mattman, 1970). Hence, the sensitivity of wild type and all the *lcp* deletion mutants to lysozyme, was determined by resazurin reduction assay (Malm *et al.*, 2018).

All the strains in our study were found to have lost their metabolic activity in the presence of 0.5 mg/mL lysozyme (Fig 3.12, B). The fluorescence intensity of ΔΔ(0107+5775) was hardly measurable, indicating a significant (98%) decrease (P=.001) compared to the wild type in the presence of 0.0625 mg/mL lysozyme (Fig 3.12, B). Deletion of either *MSMEG_0107* or *MSMEG_6421* did not influence the susceptibility to lysozyme much as compared to the parent strain. Though the complemented strain of Δ6421 did not deviate much from the wild type, but the complemented strain of Δ0107 showed a 24%, 38% and 46% decrease in fluorescence units in the presence of 0.0625, 0.125 and 0.25 mg/mL lysozyme compared to the wild type. In contrast, *MSMEG_5775*-deficient mutant had a 44% increased susceptibility to 0.25 mg/mL of lysozyme compared to the wild type. It is also notable that complementation with a wild type copy of *MSMEG_5775* or both *MSMEG_0107* and *MSMEG_5775* in the mutants only partly restored the stress response to a level similar to the wild type. Apart from these, the most striking observation was made in the strain lacking both *MSMEG_0107* and *MSMEG_6421*. Though this strain

displayed similar phenotype to the wild type at 0.0625 mg/mL lysozyme stress, but it became gradually resistant with increased concentration of lysozyme (29% and 64% resistance at 0.125 mg/mL and 0.25 mg/mL lysozyme concentration respectively). Similar levels of resistance were also observed in the complemented strain of this mutant. On the other hand, the strain lacking both *MSMEG_5775* and *MSMEG_6421* displayed a 17% and 31% resistance to lysozyme at lower doses of 0.0625 m/mL and 0.125 mg/mL lysozyme however, with increased concentration, it became sensitive by 35% than the wild type. This indicates that the cell envelope of double *lcp* deletion mutant ΔΔ(5775+6421) is least affected with small doses of lysozyme however with increased dosage, the β-1,4-glycosidic bonds of the PG backbone is disrupted. Although the complemented strain of this mutant was similar to the wild type phenotype at lower doses of 0.0.625 mg/mL and 0.125 mg/mL lysozyme, it became sensitive by 60% than the wild type, with 0.25 mg/mL lysozyme. Thus, in our study, all the results point out at the hypersensitivity exhibited by ΔΔ(0107+5775) towards the cell wall disrupting detergents such as SDS and lysozyme as well as to various first line of antibiotics used against TB infection. This corroborates the weakened state of the cell envelope of this double *lcp* deletion mutant. Not only this, it was also observed that strains lacking *MSMEG_6421* were 20-31% resistant to lysozyme at 0.125 mg/mL concentration. Though Δ6421 and ΔΔ(5775+6421) amongst these strains became sensitive to an increased concentration of 0.25 mg/mL lysozyme, but the mutant lacking both *MSMEG_0107* and *MSMEG_6421*, as well its complemented strain became increasingly resistant by 60%-65% resistant at this concentration. This indicates that the mutants lacking *MSMEG_6421* requires higher lysozyme concentration to disrupt the PG backbone, as these mutants do not seem to disrupt the cell wall significantly. Looking at these results, we thought it to be interesting to further investigate the expression network of the *Msmeg lcp* genes under standard and stress conditions (refer to Chapter-4).
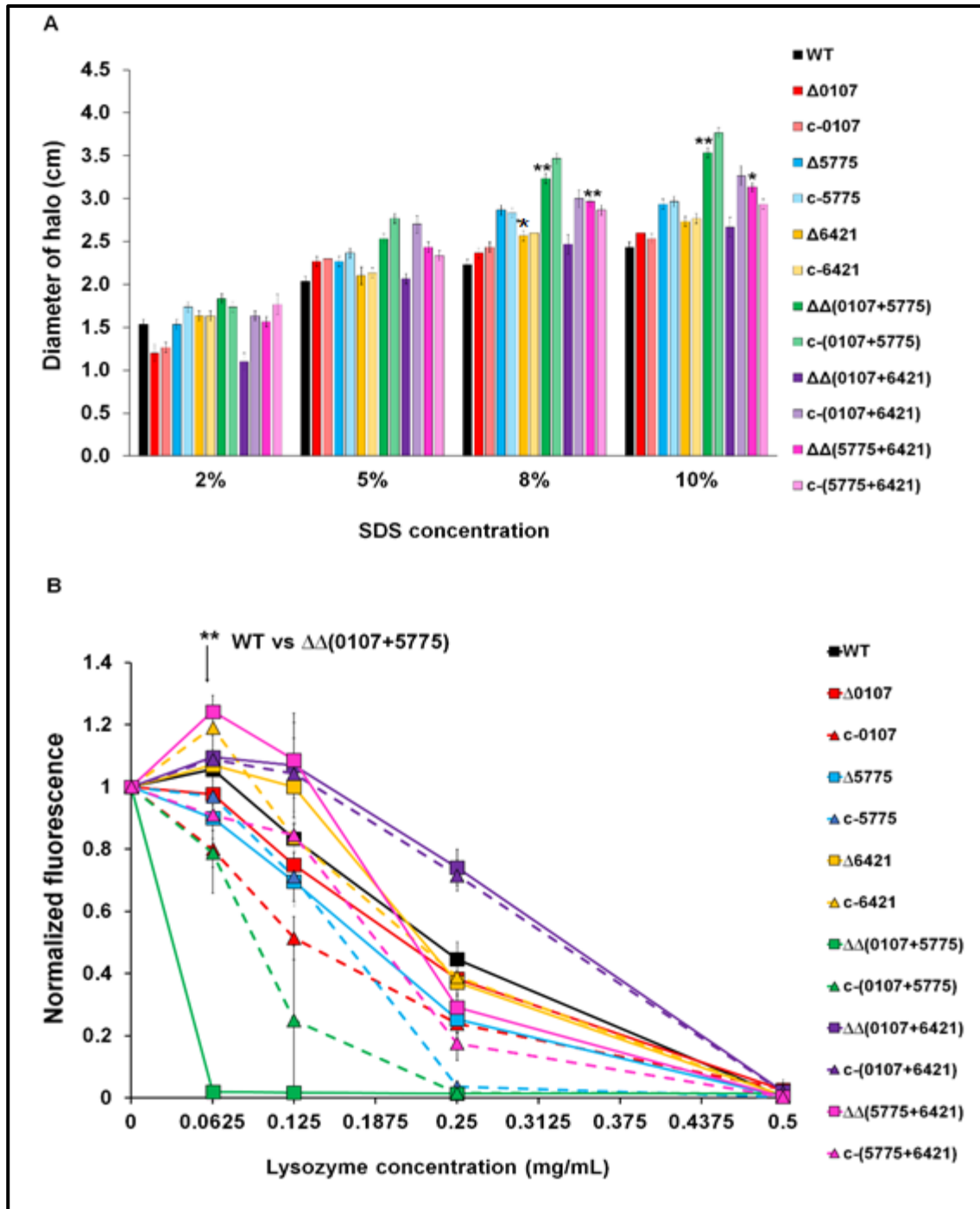
Fig 3.12. Sensitivity of the _lcp_ double deletion mutant ΔΔ(0107+5775) to stress conditions. A. Increased sensitivity of ΔΔ(0107+5775) to cationic surface detergent SDS. The diameter of the zone of complete inhibition (termed as halo) was measured and compared with the wild type strain (*, $P ≤ .05$; **, $P ≤ .005$). The representative data is an average of three independent experiments each performed in triplicate. **B.** Increased sensitivity of ΔΔ(0107+5775) to different concentrations of lysozyme. A resazurin based fluorescent assay was done and the fluorescence measured at 600 nm in a microplate reader. The double lcp deletion mutant ΔΔ(0107+5775) displayed a significant susceptibility of lysozyme at 0.0625 mg/mL concentration compared to the wild strain (**, $P = .001$). The other significant values are mentioned in the text. The representative data is an average of three independent experiments each performed in at least three replicates. In both A and B, the data is normalized to a no-stress condition of the particular strain. The error bars represent the standard deviations from three independent experiments.

3.3.9. Inflammatory responses to *Msmeg lcp* mutants in THP-1 macrophages

Based on our results on lysozyme susceptibility of ΔΔ(0107+5775) and resistance observed in ΔΔ(0107+6421) at all concentrations used in our study, we next ventured into seeing the effect of these mutants on cytokines produced in macrophage-like THP-1 cells. As macrophages are known to release lysozymes as the first defense strategy against pathogens, it would be interesting to see if these mutants produce the same effect in macrophages. For this study, the PMA differentiated THP-1 cells were infected with the wild type and the mutants at an MOI of 10, and lysed the macrophages at different time points (2h, 4h and 8h) to extract mRNA, and used it to investigate the transcription levels of cytokines by quantitative PCR.

This is a preliminary experiment and the first thing to be noticed was the overall heightened upregulation of IL-1β of all the strains under study when compared to other pro-inflammatory cytokines (about 50-fold higher than TNF-α and 400 to 600-fold higher than IL-6) (Fig 3.13). This inflammatory cytokine expression pattern of IL-1β may be related to the persistence of *Msmeg* in the THP-1 cell line. The second thing to be noticed is the decrease in IL-1β expression levels of the *Msmeg* wild type strain and the *lcp* mutants at 4 hours of infection (T=4h). This pattern was found to be consistent in the LPS control too. Infact, in comparison to the LPS control, all the *Msmeg* strains showed decreased transcription at T=4h. Though, in the mutant lacking *MSMEG_6421*, a 1.3-fold decrease in transcription level was observed at this time, when compared to that in wild type. Although after T=4h, LPS showed similar transcription levels to this time point at the 8th hour (T=8h), however, the *Msmeg* wild type and mutants showed a heightened expression again at T=8h. Given that there are no extra cellular bacteria to have caused this, we speculate that, majority of the bacteria escapes the immune response and therefore, there is an increase in the pro-inflammatory cytokines again, at T=8h, and this observation was find in almost all strains. This data is contradictory to our idea on the double *lcp* deletion mutant, ΔΔ(0107+5775) as it shows similar transcription level of IL-1β with the wild type at T=4h and 8h, whereas a 1.4-fold increase than the wild type in the cytokine release at T=2h. Thirdly, elevated levels of expression of IL-1β in the *MSMEG_5775*-deleted strain was observed at both T=2h and T=8h in comparison to other *Msmeg* strains. However, at T=4h, the expression of this cytokine was similar to that of the wild type. Not only IL-1β, but the *MSMEG_5775*-deleted strain showed higher transcription levels for the other pro-inflammatory cytokines TNF-α as well as IL-6. This is in contrast to our result on lysozyme sensitivity of this strain at all concentrations used in our study. Lastly, our anticipation of ΔΔ(0107+6421) being resistant to lytic stress produced by the macrophage-like cells, are in congruence with the result obtained for this strain at T=8h post infection. A 1.2-fold increase in IL-1β production compared to the wild type suggests this mutant

to evade the lytic cycle of the THP-1 macrophages. However, our data needs to be reproducible to draw a firm conclusion hence, we aim to pursue further experiments.
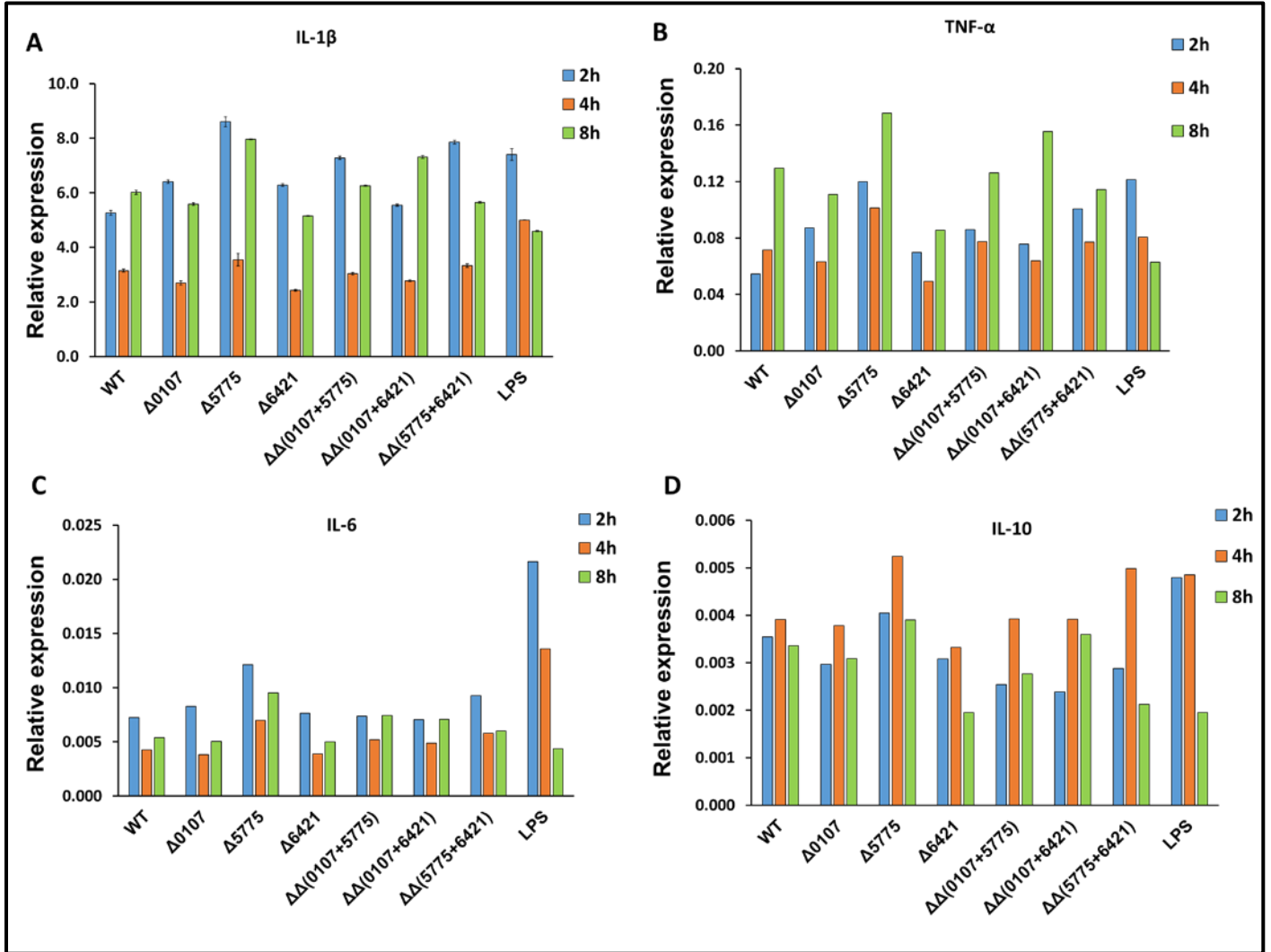
Fig 3.13. Determination of the production of different cytokines in THP-1 macrophages infected with *Msmeg lcp* mutants. PMA-differentiated THP-1 macrophages were infected at an MOI of 10. Culture supernatants were collected at indicated intervals, and the transcription of **A.** IL-1β, **B.** TNF-α, **C.** IL-6 and **D.** IL-10 mRNA were detected by real-time PCR. The preliminary data presented here is of one independent experiment performed in triplicate. The error bars in A represent standard deviation between the triplicates of the single experiment.

3.3.10. *Msmeg* double *lcp* deletion mutants showed exposed AG moieties in the cell envelope

The localization of polysaccharides and glycoconjugates in the cell wall and membrane of a bacterium serves as an interface between the cell and its environment, which facilitates cell-cell interactions and cell signaling mechanisms (Dwek, 1996). AG is one such major cell wall polysaccharide in mycobacteria, and it is composed of galactose and arabinose sugar residues, in the furanose (*f*) ring form collectively called Gal*f* and Ara*f* (McNeil *et al.*, 1987). Gal*f* and Ara*f* are galactan monomers that are embedded in the mAGP core of the mycobacterial cell wall. In mammals, galactose which is a common monosaccharide, is found prominently in the pyranose form (Gal*p*) whereas, Gal*f* is exclusive to non-mammalians (Tefsen *et al.*, 2012). The galactan component of Gal*f* is a linear chain of about 30 alternating 5- and 6-linked *β*-D-Gal*f* residues (Daffe *et al.*, 1990) which is interconnected with three highly branched arabinan chains, consisting of approximately 30 Ara*f* residues (Besra *et al.*, 1995). The non-reducing end of these arabinan chains serve as an attachment site for mycolic acids, succinyl and galactosamine (**D**-GalN) moieties (Draper *et al.*, 1997, Bhamidi *et al.*, 2008).

Since, Gal*f* in *Aspergillus niger* could be detected by anti-Gal*f* EB-A2 monoclonal antibody (mAB) (Chiodo *et al.*, 2014), we believed that it could also detect the Gal*f* moieties in *Msmeg*. Hence, an attempt was made to identify these moieties by EB-A2 mAB. Though it is believed that a galactomannan disaccharide fragment, Gal*f*-β-(1→5)-Gal*f* acts as the epitope for EB-A2 detection, however a recent study in *Aspergillus fumigatus* suggested that the mAB EB-A2 used in the kit detects multiple epitopes of circulating galactomannan. Thus, it is perceived to be a major drawback for the detection of Aspergillosis, due to increased occurrence of false positives (Krylov *et al.*, 2019). However, for the sole purpose of potentially detecting exposed galactan moieties in the *lcp* mutants, this antibody proved to be useful.

Mtb and related pathogenic mycobacteria also contain mannose-capped lipoarabinomannan (ManLAM) that binds to cell surface receptors of macrophages and dendritic cells. This results in several immunomodulatory responses, downregulation of cell-mediated immunity and effective invasion of the host cells (Briken *et al.*, 2004). Apart from ManLAM, the surface of the *Mtb* cell is also rich in other mannose-containing biomolecules like lipomannan (LM), phosphatidyl-myo-inositol mannosides (PIMs), arabinomannan, mannan, and manno-glycoproteins. PIMs, LM, and ManLAM are exposed on the *Mtb* cell surface, and they act as ligands for host cell receptors and contribute to the pathogenesis of *Mtb* (Torrelles & Schlesinger, 2010). Lectins have been extensively used to detect these mannosylated glycoconjugates in *Mtb* (Gonzalez-Zamorano *et*

al., 2009, Mohanty *et al.*, 2015). The use of small bioreceptors (based on lectin-carbohydrate binding) have also been recently shown as potent molecules for pathogen detection (Saucedo *et al.*, 2018). However, in the non-pathogenic *Msmeg*, ManLAM is partially replaced by inositol phosphate units (PILAM) (Khoo *et al.*, 1995, Gilleron *et al.*, 1997). The presence of mannosylated glycans like LM and PIMs in *Msmeg* cell envelope led us to determine the ConA lectin binding efficiency of these exposed units of mannose-containing biomolecules in the *lcp* deletion mutants of *Msmeg*.

Given that the double *lcp* deletion mutant ΔΔ(0107+5775) has a weakened cell envelope, it was hypothesized that the Gal*f* moieties present in the cell wall core (Fig 1.5) of this mutant, might be exposed. Thus, we were intrigued to detect these elements by Gal*f* specific EB-A2 mAB, in a direct ELISA method, and used multivalent gold nanoparticles carrying Gal*f* (Gal*f*-GNPs) as a positive control (Chiodo *et al.*, 2014). In this study, EB-A2 mAB was used to detect the exposed galactan in the whole bacterial cells of the wild type and the *lcp* mutants. Most strains did not show a signal deviating from the signal that the wild type generated in this experiment. Interestingly, the single deletion mutant Δ0107, showed a significant 1.8-fold decrease (*P*=.03) in detectable Gal*f* whereas, the *lcp* double deletion mutant ΔΔ(0107+5775) displayed a highly significant 5-fold increase (*P*=.0002) of detectable Gal*f* moieties (Fig 3.14, A). This indicates that the Gal*f* in the cell envelope that is normally shielded in the wild type is widely accessible by the EB-A2 antibody in the mutant ΔΔ(0107+5775). However, there are no experimental ELISA data for the complemented strains currently, and we plan to use them in the near future.

ConA, is a plant derived terminal α-D-mannose and α-D-glucose-binding lectin that recognizes mannosylated *N*-glycans (Cummings & Etzler, 2009). The mycobacterial cell wall is composed of many immunomodulatory mannose containing glyco-conjugated complexes, such as the LAM, LM and PIMs (Fukuda *et al.*, 2013, Gilleron *et al.*, 2003). The difference between the LAMs in *Mtb* and *Msmeg* is that, *Mtb* has mannose-capped LAM (ManLAM) whereas, *Msmeg* has phosphomyoinositol-capped LAM (PILAM) (Nigou *et al.*, 2003). ConA does not bind PILAM in *Msmeg*, rather it has higher affinity for terminal mannosyl units of LAM in *Mtb* (Zhang *et al.*, 2019),. However, not only mannosides, but ConA has also been reported to bind glucosides (Saucedo *et al.*, 2018). It was thus hypothesized ConA to detect the exposed glucosides or other forms of mannosylated molecules in the putatively cell envelope-altered structure of *lcp* deletion mutants. This implicates potential altered levels of mannosylated glycans in the mutants, which are either embedded in the plasma membrane or in the outer membrane in *Msmeg*, with the help of their lipid moiety (Pitarque *et al.*, 2008). Unexpectedly, deletion of either *MSMEG_0107* or *MSMEG_5775*, led to a 2-fold (*P*=.02) and 2.3-fold (*P*=.001) decrease in the levels of mannose-

exposed moieties compared to that of the wild type strain as detected by ConA in our setup (Fig 3.14, B). Since this is a preliminary report on the abundance of mannosylated glycoconjugates in the *lcp* deletion mutants, the nature of these mannosylated complexes detected by ConA is still unclear. This study could be a foundation to understand the underlying mechanism that involves the impact of *lcp* genes on the mannosylated glycolipids.
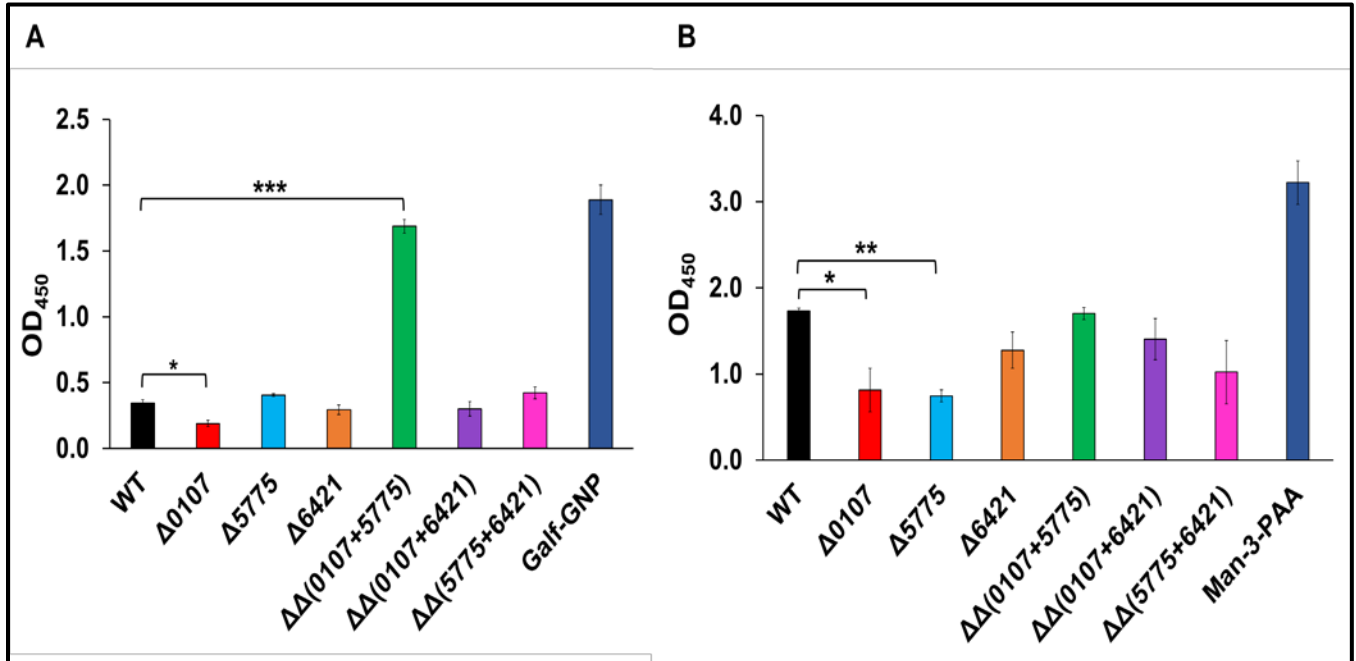
Fig 3.14. Detection of exposed galactofuranose and mannose moieties in *lcp* mutants. Direct ELISA was performed on a 96-well plate for both the experiments. The detection of monosaccharides present on the cell envelope of mycobacteria is relative to the amount of coated bacterial cells (1.5 X 10^8cells/well). **A.** The arabinofuranose-bearing molecules present on the cell envelope were detected by EB-A2 Mab. Gal*f* coated with gold nano-particle (20 µg/mL) was used as the positive control. **B.** The exposed mannose residues were detected by ConA lectin. Man-3-PAA (Mannotriose-polyacrylamide) at a concentration of 2 µg/mL was used as the positive control. The data represented here is an average of three independent experiments each performed at least in duplicate. The error bars represent standard deviation and asterisks indicate significant differences (* $P \leq$ .05, ** $P \leq$ .005, *** $P \leq$ .0005).

3.4. Conclusions

In this chapter, the impact of three *lcp* genes in *Msmeg* viz., *MSMEG_0107*, *MSMEG_5775* and *MSMEG_6421* was studied, by developing single and double deletion mutants of these genes. The complemented strains for all the mutants were used in experiments such as growth curve, colony morphology, aggregation, biofilm formation in 96-well plate, as well as stress responses to various antibiotics, detergent and lysozyme. The higher expression levels of these complemented strains as evident from the real-time PCR results, and their partial restoration to the wild type phenotype, suggests that native promoter of *lcp* genes is required for complete restoration to wild type expression level.

The effect of the *lcp* mutants on the physiological characteristics of *Msmeg* were explored and it was found that disruption of two of these *lcp* genes viz., *MSMEG_0107* and *MSMEG_5775* proved to be detrimental. This double *lcp* deletion mutant ΔΔ(0107+5775) had altered bacterial cell surface properties compared to the wild type strain. These include smooth colony morphology, slower aggregation rate, diminished biofilm formation on air-liquid interface and increased susceptibility to antibiotics as well as stress conditions. However, opposite effects for this mutant was observed in the crystal violet detection of biofilms. Since, higher levels of biofilm were not observed to adhere to the 96-well plate in ΔΔ(0107+5775) compared to the wild type, it is assumed that the crystal violet dye might be more reactive to this mutant thus showing elevated levels. The exposed galactan moieties on this mutant's cell envelope could be detected by Gal*f* specific EB-A2 mAB. This further strengthens our belief that the cell wall of this mutant might have disrupted thus, revealing the Gal*f* motifs that are otherwise embedded below the mycolic acid layer. Since the experiment design in this study included the use of whole bacterial cells, our conclusions are confined to the exposed Gal*f* motifs. However, a future direction could be detecting these motifs in the culture supernatants to identify if the double *lcp* deletion mutant ΔΔ(0107+5775) releases un-ligated AG into the culture medium. The outcomes from this chapter thus suggest that the loss of both *MSMEG_0107* and *MSMEG_5775* severely compromises the *Msmeg* cell wall integrity and indicate that, although these genes are not essential, they contribute to building a healthy cell envelope.

Chapter-4

Differential regulation of *Mycobacterium smegmatis* LCP proteins under *in vitro* stress conditions

Abhipsa Sahu, Ziwen Xie and Boris Tefsen

Department of Biological Sciences, Xi'an Jiaotong-Liverpool University, Suzhou, Jiangsu, 215123, PR China

4.1. Introduction

Mycobacterium tuberculosis (Mtb), the causative agent of TB is a global health threat, and its success as a pathogen partially owes to its ability to evade the host immune system, and stays dormant until the host becomes immunocompromised. With the elusive growth of the organism as multidrug resistant strains, it is necessary to understand the underlying mechanisms of the mycobacterial immune system evasion to combat the disease. All the species of mycobacteria are predominantly obligate aerobes, which includes the pathogenic *Mtb* and the non-pathogenic *Msmeg* (Cumming *et al.*, 2014). The hydrophobic cell surface of mycobacteria permits tenacity and survival under various adverse environmental conditions and interfaces (Parker *et al.*, 1983, Collins *et al.*, 1984, Santos *et al.*, 2015, Yang *et al.*, 2018). Adaptation to different environmental stress conditions, encountered during various stages of the life cycle, is a key survival strategy of this human intracellular pathogen (Dogra *et al.*, 2015). Not only *Mtb*, but also the saprophytic and model organism *Msmeg*, is capable of existing in several ecological setups, thereby showcasing its ability to adapt to different environmental cues.

Redox reactions are essential in regulating the cellular, biochemical and physiological processes (Kondo *et al.*, 2006, Du *et al.*, 2012) and, also play a crucial role in both aerobic and anaerobic respiration. In aerobic bacteria, the oxidative and reactive species are neutralized by antioxidants to maintain redox homeostasis (Cumming *et al.*, 2014). This equilibrium is important to efficiently prevent damage caused by oxidative stress, which otherwise ensues impairment of cellular functions, such as lipid peroxidation, protein degradation, DNA damage, and subsequently, disease, senility, and cell death (Berg *et al.*, 2004, Kondo *et al.*, 2006, Du *et al.*, 2012). During infection, *Mtb* is exposed to various redox stressors such as reactive oxygen species (ROS), hydrogen peroxide (H_2O_2), superoxide (O_2-) and hydroxyl radicals (OH). The released superoxide gets converted into hydrogen peroxide by the antioxidant, superoxide dismutase. The

decomposition of this hydrogen peroxide to water and oxygen, is then catalysed by catalase (Piddington *et al.*, 2001). *Mtb* also encounters reactive nitrogen species (RNS) during infection, which is primarily produced by NADPH oxidase (NOX2), the main producer of ROS inside the host. This especially includes nitric oxide (NO), which is produced largely by inducible nitric oxide synthase (iNOS) (Bhat *et al.*, 2012). Higher oxidative stress marker levels and lower antioxidant capacity have been regarded as the main culprit for pathogenesis of TB (Vidhya *et al.*, 2019). In order to maintain the redox state, almost all microorganisms including *Mtb* use antioxidant enzymes (Fukumori & Kishii, 2001, Bryk *et al.*, 2002, Dalle-Donne *et al.*, 2008, Lu & Holmgren, 2014, Si *et al.*, 2015, Watanabe *et al.*, 2016). This indicates the importance of ROS in monitoring mycobacterial infections.

A study in the non-pathogenic *Msmeg* has also reported generation of ROS in phagocytic models (Ghosh *et al.*, 2019). In the natural habitat, *Msmeg* is known to use a wide range of antioxidant enzymes such as superdioxide dismutase, catalase, and peroxidase to detoxify ROS (Rizvi *et al.*, 2019). Acidity is also generated in the phagocytic cells during infection, and *Mtb* has the capability to resist this stress condition. Similarly, the prevalence of acidic conditions in soil and aquatic habitats where the saprophytic mycobacteria reside, indicates the acid sensing mechanisms of *Msmeg*. It has been reported that in *in vitro* conditions, *Msmeg* initially withstands the acid insult, and subsequently delays the acidification of phagosomes (Anes *et al.*, 2006). Acid resistance in *Msmeg* has also been related to downregulation of transmembrane proteins involved in transport mechanisms, and upregulation of fatty acid metabolism encoding enzymes (Roxas & Li, 2009). Apart from these, lysozyme also acts as a common and crucial stress giver to mycobacterial cells. Lysozyme as an antimicrobial factor, is released from phagocytic cells in response to mycobacterial infections *in vivo*. It is the first enzymatic weapon used by the host defense mechanism against pathogens, which has a strong bactericidal effect. Under *in vitro* conditions, lysozyme induces the conversion of bacillary to defective forms of cell wall in mycobacteria (Mattman, 1970). *In vivo*, these defective cell wall type of mycobacteria develop, when encountered with certain antimicrobial drugs, or, in response to lytic stress that is regulated by host phagolysosomes (Mattman, 1970, Michailova *et al.*, 2000). Lysozyme is such an enzyme that induces lytic stress by hydrolyzing 1,4-β-linkages between *N*-acetylmuramic acid and *N*-acetyl-D-glucosamine residues in PG layer. Transcriptional regulation is one of the important mechanisms controlling the gene expression of such stress responses. Transcription units, promoters and sigma factors are all part of this transcriptional machinery.

DNA-protein interactions that take place during transcription initiation play a crucial role in regulating the gene expression. Transcription factors (TFs), the DNA binding proteins that are

capable of activating or repressing transcription of particular genes, regulate the bacterial metabolic adaptations at the transcriptional level (Browning & Busby, 2004). TFs function by activating their respective regulons, which are a collection of operons that are transcriptionally co-regulated (Liu *et al.*, 2016). Bacterial TFs include sigma (σ) factors, which aids promoter recognition and specificity to the RNA polymerase holoenzyme under various stress conditions (Ishihama, 2010). The σ factors first form a complex with the core enzyme of RNA polymerase, and subsequently escorts the RNA polymerase to promoter DNA, resulting in the formation of a transcription bubble with the opening of the double-stranded DNA, thus facilitating the synthesis of initial short RNA transcripts. This is followed by the assistance of σ factors in promoter escape (Saecker *et al.*, 2011). These σ factors along with two-component signal transduction systems in bacteria are well known to coordinate several stress responses by differential expression of transcriptional regulators (Feng *et al.*, 2016). *Msmeg* has been reported to have 28 σ factors which is twice the number that are present in *Mtb*, and therefore the regulatory repertoires of *Msmeg* are found to be far more complex than in *Mtb* (Waagmeester *et al.*, 2005). The σ factors that are unique to *Msmeg* are believed to be involved in regulating the genes that are specific to this non-pathogen. For instance, an enhanced form of SigJ, SigL and SigH subfamily are found in this species, with SigH containing seven paralogs, which is believed to be acquired by gene duplication followed by speciation (Waagmeester *et al.*, 2005). SigF and SigH are alternative σ factors in *Msmeg* that plays a vital role in oxidative stress (Gebhard *et al.*, 2008, Provvedi *et al.*, 2008). While additionally, SigH plays a crucial role in heat stress response in *Msmeg* (Fernandes *et al.*, 1999). SigF has been demonstrated to be involved in regulating the genes involved in carotenoid biosynthesis as well as other cell wall components (Provvedi *et al.*, 2008). With this background on mycobacterial regulatory elements, it will thus be interesting to predict putative transcriptional motifs and regulators responsible for regulating the gene expression of *lcp* genes.

The LCP proteins, a family of enzymes responsible for ligating the AG and PG in mycobacteria, has four *lcp* homologues in *Msmeg*, which might have occurred due to gene duplication based on its large genome size (Hubscher *et al.*, 2008). Under standard conditions, all the four *lcp* genes would be expected to have a certain transcription level. However, when encountered with a specific stress, some of these genes might be differentially regulated, as a survival mechanism of the organism. Moreover, the transcriptional profile that may exist in *Msmeg* during various stress situations in comparison to the standard phase, has remained elusive. Our results on stress responses of the *lcp* mutants towards the detergent SDS and antimicrobial compounds such as lysozyme and antibiotics in chapter-3 intrigued us to investigate the differential expression of these cell envelope proteins under three stress conditions, namely, lysozyme, acidic and oxidative

stress. These three stress conditions are encountered by *Msmeg*, both as a soil bacterium and as an opportunistic infection inside phagocytic cells of immunocompromised hosts (Pierre-Audigier *et al.,* 1997). Hence, in this study, the induction of these stressors were simulated in *Msmeg* wild type strain and the *lcp* deletion mutants *in vitro*, in order to identify the transcription profiles of these genes during such adverse situations. This investigation will contribute to answering some questions regarding the lifestyle of *Msmeg* with respect to these cell envelope proteins, such as: i) What is the need of four *lcp* homologues in *Msmeg*? ii) Is the gene expression of the four *lcp* homologues in *Msmeg* controlled by a single or multiple regulator(s)? This study will therefore help gain new insights into the role of *Msmeg lcp* genes when encountered with environmental stress that have not been reported earlier. Apart from this, the potential regulatory elements in *Msmeg* that govern the expression of all these four *lcp* genes would also be investigated. This study along with further investigation on the regulatory motifs might relay information on different pathways involved in the cell envelope biogenesis in *Msmeg*.

4.2. Materials and methods

4.2.1. Bacterial strains and growth conditions

The bacterial strains were grown as mentioned in chapter-3 (section 3.2.1). Briefly, for every experiment, fresh seed culture of *Msmeg* wild type, mutant and complemented strains were grown in 7H9 medium (Difco) supplemented with 10% (v/v) ADS, 0.5% glycerol (v/v), 0.05% Tween-80 (v/v) (Solarbio, China) and appropriate antibiotics wherever necessary (section 3.2.1). The overnight culture was prepared by inoculating the seed culture at a dilution of 1:100 in 7H9 medium, supplemented with Tween-80 and antibiotics. The marked double deletion mutants ΔΔ(0107+5775), ΔΔ(0107+6421) and ΔΔ(5775+6421), as well as complemented strains of single deletion mutants viz., c-0107, c-5775 and c-6421 were supplemented with 0.25 µg/mL of hygromycin (Hyg). The complemented strains of double deletion mutants, c-(0107+5775), c-(0107+6421) and c-(5775+6421) were supplemented with a final concentration of 0.25 µg/mL of kanamycin (Kan).

4.2.2. Stress conditions induced in *Msmeg*

In this study, three stressors were used viz., oxidative (H_2O_2), acidic pH (HCl) and lysozyme.

4.2.2.1. Acidic stress

Strains were grown to late-log phase (OD_{600} of 1.5 -2.0) in 7H9 medium (Difco) supplemented with 10% (v/v) ADS, 0.5% glycerol (v/v), 0.05% Tween-80 (v/v) (Solarbio, China), and diluted further to OD_{600} of 0.5 in 7H9 medium. The treated samples were prepared by adjusting the culture medium to pH 6.7 and pH 5.7 by adding HCl. Then the strains with and without stress were incubated at 37°C, while shaking at 200 rpm for 15 minutes before isolating mRNA.

4.2.2.2. Lysozyme stress

Strains were grown to late-log phase (OD_{600} of 1.5 -2.0) in 7H9 medium (Difco) supplemented with 10% (v/v) ADS, 0.5% glycerol (v/v), 0.05% Tween-80 (v/v) (Solarbio, China), and diluted further to OD_{600} of 0.5 in 7H9 medium. The diluted cultures were supplemented with lysozyme at a final concentration of 0 mg/mL, 0.125 mg/mL and 0.5 mg/mL. All the strains with and without lysozyme were incubated at 37°C, shaking at 200 rpm for 1 hour.

4.2.2.3. Oxidative stress

Strains were grown to late-log phase (OD_{600} of 1.5 -2.0) in 7H9 medium (Difco) supplemented with 10% (v/v) ADS, 0.5% glycerol (v/v), 0.05% Tween-80 (v/v) (Solarbio, China), and diluted further to OD_{600} of 0.5 in 7H9 medium. The diluted cultures were treated with H_2O_2 at a final concentration of 5 mM and 7 mM. All the treated and untreated strains were incubated at 37°C, shaking at 200 rpm for 15 minutes.

4.2.3. mRNA extraction and RT-qpCR

Bacterial culturing conditions for RNA extraction was done as in chapter-3 (section 3.2.5). Reverse transcription was performed using the manufacturer's instructions on the kit (Promega, China). The RT-qPCR was performed using the manufacturer's instructions of the kit (Promega, China). The RT-qPCR primers used are listed in Table 3.2 in chapter-3. Apart from these, RT-

qPCR primers used for determining the expression of *MSMEG_1824* are: 1824-qPCR-F (5' ATCTCGAGCAAGACCCTGTC 3') and 1824-qPCR-R (5' ATCCATGATGTCCCAACCGT 3'). SigA (rpoD) was used as the reference gene for the real-time PCR. The relative expression analyses are based on three replicates of two independent experiments.

4.2.4. *In silico* analysis of regulatory elements of the LCP proteins

The operon and the promoter sequence for each of the four *Msmeg lcp* genes was predicted by DMINDA[2] server (http://bmbl.sdstate.edu/DMINDA2/). The DNA binding transcriptional motifs in the promoter region were subsequently predicted in the MEME-suite server (Bailey *et al.*, 2009) using default settings. The classic discovery mode was employed to optimize the *E*-value of the motif information content. The site distribution was set to "Any number of repetitions" of a contributing motif site per sequence. A default motif width between 6 and 50 was used. The predicted motifs ranked on the basis of their lowest statistical significance (*E*-value) were compared with the available transcriptional motifs in the TOMTOM database of the MEME suite, and were aligned for each significant match. The predicted motifs were highlighted in the sequence map scaled in Snapgene software.

4.3. Results and discussion

Based on our results on the pyrophosphatase activity exhibited by the *lcp* genes in *Msmeg* (refer section 2.3.4), and the fact that *Msmeg* has four *lcp* homologues in its genome, we were intrigued to study the differential expression of these four genes in the wild *Msmeg* strain, as well as in the deletion mutants that were generated during this study. In this study, three stressors were included, namely, oxidative (H_2O_2), acidic pH (HCl) and lysozyme, as these are commonly encountered by *Msmeg* in its natural habitat, as well as *in vitro* during macrophage infection. As mentioned in Chapter-3, a strain devoid of *MSMEG_1824* has not been included in this work, due to unavailability of resources to construct a conditional mutant for this gene.

4.3.1. Differential expression of *Msmeg* wild type *lcp* genes under standard conditions

Genes in an organism will show variations in expression at different growth stages and *Msmeg* is no exception (Singh & Singh, 2009). Here, we wanted to investigate the differential expression of the four *lcp* genes under standard laboratory conditions, and so, cultured the strain until late-log phase i.e., until OD_{600} of 1.5 - 2.0 (Singh & Singh, 2009). The average expression level of *MSMEG_1824* in three independent experiments was 2.1 to 1.6-fold lower than the other three *lcp* homologues, (P = .012 for *MSMEG_0107*, .065 for *MSMEG_5775* and .219 for *MSMEG_6421*) (Fig 4.1). Both *MSMEG_5775* and *MSMEG_6421* seemed slightly higher expressed than *MSMEG_0107*, although more independent experiments would be needed to confirm that. However, this study focusses on the expression levels of the *lcp* genes at late log phase where the strains exhibit negligible or no cell division at all. Hence, a future perspective is evident in this perspective where early log phase cultures would be used to study the transcription profile of the *lcp* genes.
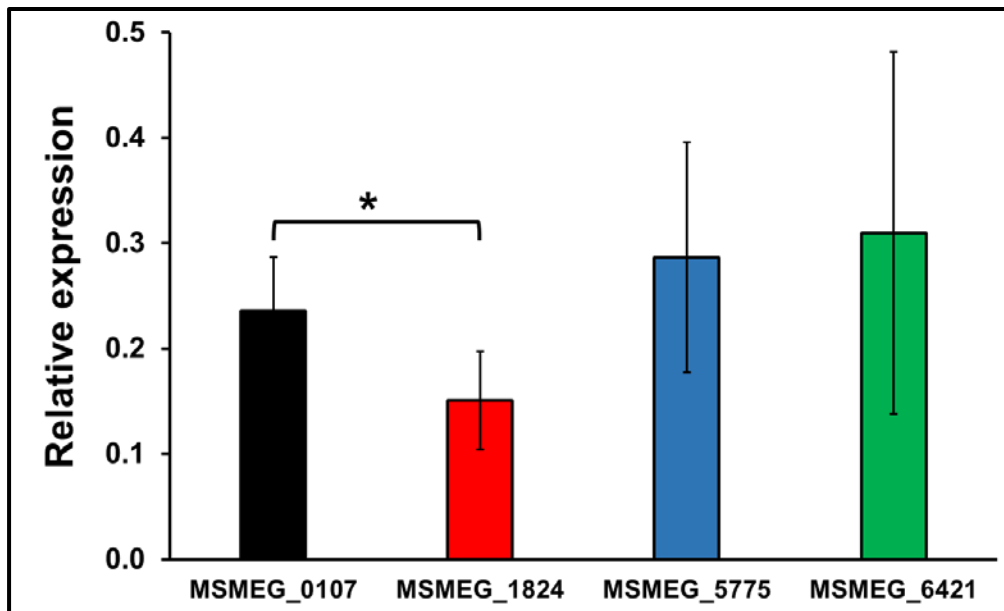
Fig 4.1. Differential expression of *Msmeg lcp* homologs in the wild type strain. The relative expression of *lcp* genes in the wild type strain mc[2]155 was determined with respect to the reference gene *rpoD* (SigA). The presented data is an average of three independent experiments each performed in triplicate. The error bars represent standard deviation between the three experiments and asterisk represents significant difference (*, $P \leq .05$).

4.3.2. Differential expression of *Msmeg lcp* genes in the wild type strain under various stress conditions

Next, it was interesting to see whether certain stress conditions, would lead to a change in the overall expression pattern of the four *lcp* genes. Therefore, three different stressors we induced to the wild type strain and the relative expression levels of the four *lcp* genes were compared by real-time PCR to those in a wild type strain under normal, non-stressed conditions. A 1.3 to 1.5-fold decrease in expression levels of all the *lcp* genes was observed under acidic stress of pH 5.7 compared to the wild type (Fig 4.2, A). This pH was chosen with regard to the low end of pH 5.0 that *Mtb* could possibly encounter in a macrophage during early infection (Pethe *et al.*, 2004). As the LCP family members are transmembrane proteins our finding is consistent with a study where the downregulation of transmembrane and transporter proteins under acidic stress in *Msmeg* is described (Roxas & Li, 2009). Overall, all the four *lcp* genes was observed to be downregulated in a similar manner and the expression pattern to be persistent. Amongst the four, *MSMEG_1824* displayed the lowest and *MSMEG_6421* the highest transcription level.

When subjected to 0.5 mg/mL lysozyme, all the *lcp* genes displayed about 1.5 to 1.9-fold reduction in expression levels however, when induced with a lower concentration of 0.125 mg/mL lysozyme, *MSMEG_0107* and *MSMEG_6421* displayed about 1.6 and 1.7-fold increase in transcription levels (Fig 4.2, B). This was also observed in our resazurin-based assay, where the single deletion mutant of *MSMEG_6421* showed about 20% resistance at 0.125 mg/mL lysozyme concentration (Fig 3.12, B). Additionally, the strain lacking both *MSMEG_0107* and *MSMEG_6421* showed about 29% and 64% at 0.125 mg/mL and 0.25 mg/mL lysozyme concentration (Fig 3.12, B). This suggests that lysozyme at 0.5 mg/mL might be toxic to *Msmeg*, resulting in the degradation of cell wall. But, at lower concentrations, the non-essential *lcp* genes are upregulated to withstand cell damage in the wild type *Msmeg*.

When induced with peroxide stress, *MSMEG_1824* was neither up nor downregulated, whereas rest all three non-essential genes are upregulated with increase in H2O2 stimulation. This seems to be an interesting observation, as this indicates the differential regulation between this essential gene and the rest of the non-essential genes.
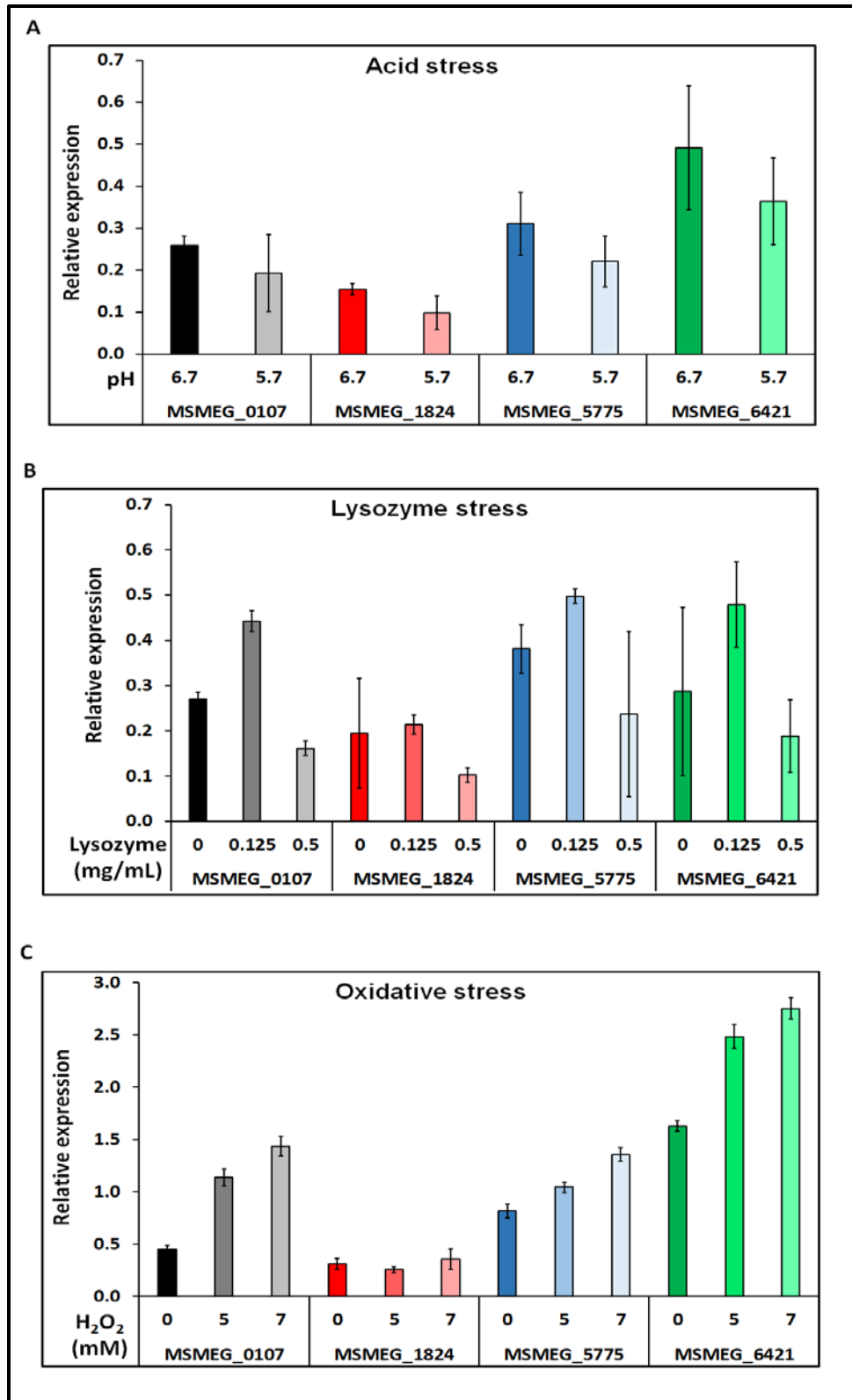
Fig 4.2. Differential expression of _Msmeg lcp_ homologs in the wild type strain under various _in vitro_ stress conditions. Determination of relative expression of the wild type strain with respect to _rpoD_ (SigA) gene as the reference, under **A.** acidic stress, **B.** lysozyme stress and **C.** oxidative stress. The data shown here is a representative of two independent experiments each performed in triplicate. Error bars represent standard error of triplicate values.

4.3.3. Differential expression of *Msmeg lcp* mutants under standard conditions

The results with the wild type strain, made us wonder whether deletion of one or two non-essential *lcp* genes would have an influence on the expression of the other *lcp* genes. Therefore, the single and double *lcp* deletion mutants that we constructed during our study were used, to investigate the differential expression of the *lcp* genes under standard conditions. Strikingly, in the single deletion mutant Δ5775, all the three genes were overexpressed by 4.6-fold (P = .02), 3-fold and 6.8-fold (P = .02) for *MSMEG_0107*, *MSMEG_1824* and *MSMEG_6421* respectively, which might be needed for that strain to compensate the loss of *MSMEG_5775* and still survive (Fig 4.3). This shows that *MSMEG_5775* is a very important entity amongst the four *lcp* genes in *Msmeg*. In a strain lacking *MSMEG_5775*, other three *lcp* genes compensate its loss. However, with a loss of any of the other two non-essential genes, there is no such compensatory effect seen.
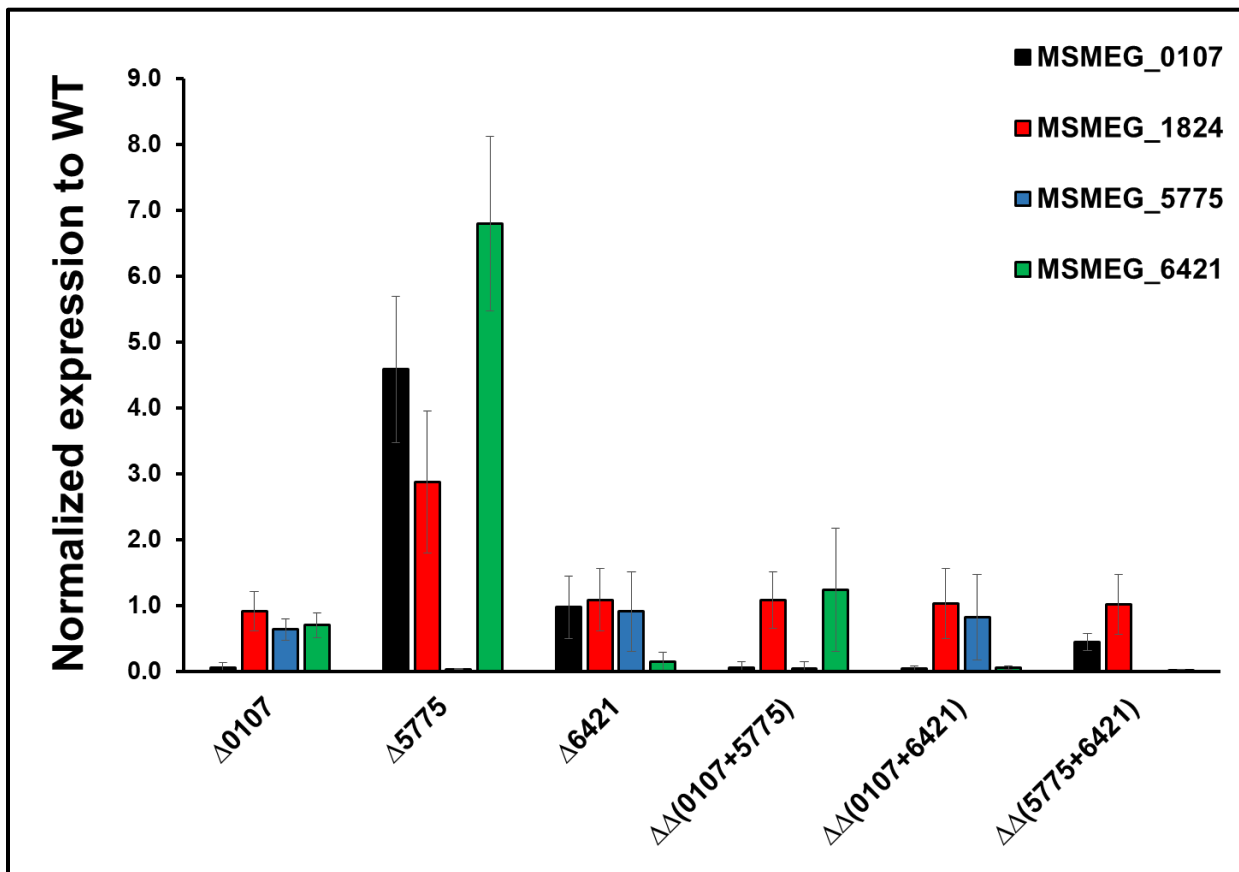
Fig 4.3. Differential expression of *Msmeg lcp* homologs in the deletion mutants under standard stress-free conditions. The relative expression of the *lcp* deletion mutants is determined with respect to *rpoD* (SigA) gene as the reference. The presented data is an average of four independent experiments each performed in triplicate, and is normalized to the relative expression of the wild type strain. Asterisks represent significant difference (*, $P \leq .05$). The error bars represent ±SD of triplicate values.

4.3.4. Differential expression of *Msmeg lcp* mutants under acid stress

In order to investigate whether stress conditions change expression patterns of the remaining *lcp* genes in the deletion mutants, a 15-minute acidic stress (pH 5.7) was applied, as described for the wild type in the previous section. Interestingly, the expression levels of the *lcp* genes under acid stress change depending on which of the non-essential *lcp* genes is lacking in the deletion mutant (Fig 4.4, A-F).

Acid stress did not have much effect on the expression of *MSMEG_1824*, *MSMEG_5775* and *MSMEG_6421* in Δ0107 compared to the standard pH (Fig 4.4, A). However, in the Δ5775 mutant, the expression levels of the other three *lcp* genes that were dramatically upregulated under standard conditions compared to the wild type strain, were found to retain its wild type transcription levels at the respective pH, on exposure to acid stress (Fig 4.4, B). However, in Δ6421 mutant, all the available *lcp* genes that expressed lower than the wild type strain under standard conditions, were found to be upregulated towards the wild type transcription levels at the respective pH (Fig 4.4, C). On a striking note, the mutant devoid of both *MSMEG_0107* and *MSMEG_5775* showed increased expression by 3-folds for *MSMEG_1824* and 2.5-folds for *MSMEG_6421*, than that of standard pH conditions (Fig 4.4, D). Interestingly, from this, it was also observed that *MSMEG_1824* which had no altered expression under normal pH conditons, showed an elevated expression under acid stress. On the other hand, at normal pH, *MSMEG_6421* was downregulated compared to the wild type but, under acid stress, *MSMEG_6421* upregulates towards the wild type transcription levels. Similar phenomenon was seen in the strain that lacks both *MSMEG_0107* and *MSMEG_6421* (Fig 4.4, E). In ΔΔ(5775+6421), *MSMEG_0107* that was downregulated in this mutant under standard conditions compared to that in wild type, was seen to be upregulated beyond the wild type transcription levels by almost 2-folds, whereas, *MSMEG_1824* that had similar transcription to the wild type, had a 2.6-fold increased transcription (Fig 4.4, F).

Thus, this study shows that a strain lacking *MSMEG_5775* under standard pH conditions, shows elevated levels of transcription of the other three *lcp* genes. However, when under acidic stress of pH 5.7, these upregulated *lcp* genes restore back to the wild type levels. In contrast, Δ6421 that showed lower transcription levels of the available *lcp* genes compared to the wild type under standard pH conditions, showed an upregulation under acid stress.
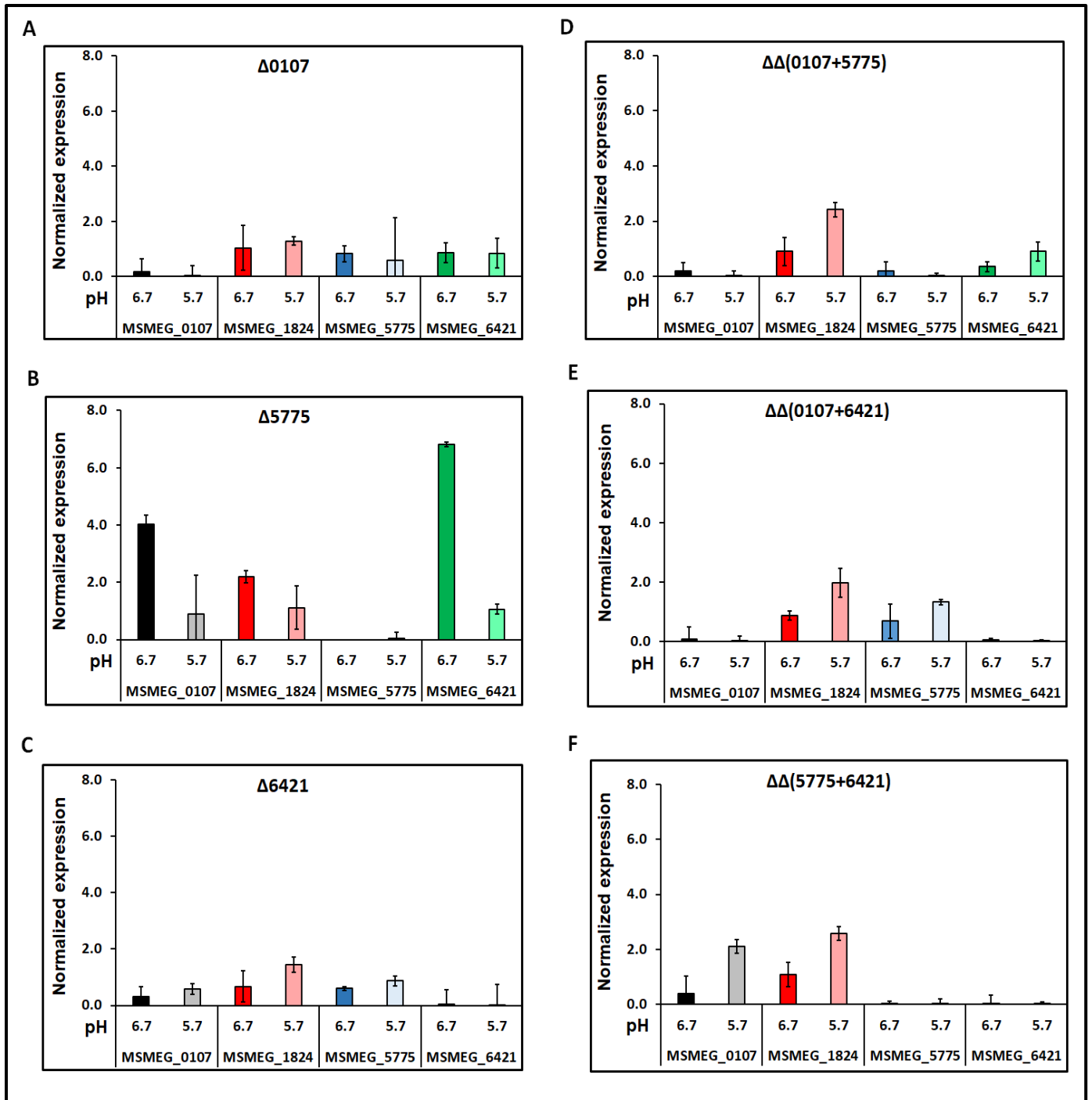
Fig 4.4. Differential expression of *lcp* genes in *Msmeg lcp* deletion mutants under acid stress. The relative expression of *lcp* genes was determined with treated and untreated samples of the deletion mutants. The untreated cultures have a standard pH of 6.7 and the treated cultures have a pH of 5.7. The data presented here are relative expression normalized to that of the wild type strain mc^{2}155, at respective pH. This is a representative experiment out of two independent experiments each done in triplicate. Error bars represent standard error of triplicate values.

4.3.5. Differential expression of *Msmeg lcp* mutants under lysozyme stress

In order to investigate whether lytic stress conditions change expression patterns of the remaining *lcp* genes in the deletion mutants, an hour lysozyme stress was applied at varied concentrations (0.125 mg/mL and 0.5 mg/mL) (Fig 4.5, A-F) as described for the wild type in the previous section in two independent experiments. It was found that, Δ0107 itself had lower transcription levels at standard conditions compared to the wild type which remained the same at 0.125 mg/mL lysozyme stimulation (Fig 4.5, A). But, when administered with 0.5 mg/mL lysozyme, all the rest three *lcp* genes were upregulated compared to the wild type transcription levels at the given dose (1.6-fold, 1.7-fold and 2-fold upregulation of *MSMEG_1824*, *MSMEG_5775* and *MSMEG_6421* respectively). As described in previous sections, Δ5775 had elevated transcription levels than the wild type strain. However, at a lower lysozyme dose of 0.125 mg/mL, both *MSMEG_0107* and *MSMEG_6421* almost restored back to the wild type transcription levels at this given concentration, whereas, the expression of *MSMEG_1824* almost reduced to half than the wild type at 0.125 mg/mL concentration (Fig 4.5, B). When administered with even a higher dose of 0.5 mg/mL, a 3-fold, 1.3-fold and 2.7 fold upregulation was observed in *MSMEG_0107*, *MSMEG_1824* and *MSMEG_6421* respectively. The most striking of all was the strain lacking *MSMEG_6421* that showed increased transcription levels of all the available *lcp* genes with increase in lysozyme concentration Fig 4.5, C). Amongst the three *lcp* genes, *MSMEG_0107* showed the highest upregulation by 5.4-fold and 15-fold at 0.125 mg/mL and 0.5 mg/mL of lysozyme respectively, compared to the wild type, at the given concentration. On the other hand, *MSMEG_1824* and *MSMEG_5775* showed an upregulation by 2.4-fold and 1.6-fold at 0.125 mg/mL concentration, as well as a 5-fold and 3.2-fold respective upregulation of these two genes at 0.5 mg/mL lysozyme. It was initially assumed that the weakened cell wall of ΔΔ(0107+5775) might witness a decreased expression levels of the *lcp* genes with increased lysozyme concentration. Though the transcription levels dropped at 0.125 mg/mL concentration, but surprisingly at 0.5 mg/mL dose, *MSMEG_6421* showed a 2.2-fold upregulation than the wild type at this concentration (Fig 4.5, D). *MSMEG_1824* on the other hand, restored its expression to the wild type levels for this mutant at this concentration. In ΔΔ(0107+6421), a similar phenomenon was observed for *MSMEG_1824*, where at standard conditions, the transcription levels were lower than the wild type by 1.3 folds but with a lower dose of 0.125 mg/mL concentration of lysozyme, there was a downregulation of this gene and finally at 0.5 mg/mL, there was an upregulation again by 1.4-fold compared to the wild type at this concentration (Fig 4.5, E). But, that was not the case for *MSMEG_5775* in this mutant. A constant upregulation of this gene was observed with increased lysozyme concentration and at 0.5 mg/mL, which was quite near to the

wild type expression level at this concentration (about 1 fold lower than the wild type level). This phenomenon was seen even in the mutant that lacked both *MSMEG_5775* and *MSMEG_6421*. Here, in the mutant ΔΔ(5775+6421), both the available *lcp* genes showed increased transcription with increasing dose of lysozyme. Hence, about 1.5-1.6 fold upregulation of both *MSMEG_0107* and *MSMEG_1824* genes was observed at 0.125 mg/mL lysozyme, and about 3.4-fold and 2.4-fold upregulation of *MSMEG_0107* and *MSMEG_1824* respectively at 0.5 mg/mL concentration (Fig 4.5, F). Therefore, from this study, it was found that, a deletion of either *MSMEG_5775* or *MSMEG_6421* or both, leads to an upregulation of the other two available *lcp* genes at 0.5 mg/mL lysozyme concentration, but at 0.125 mg/mL lysozyme concentration, downregulates the available *lcp* genes in a mutant that lacks *MSMEG_5775*.

These results showing enhanced transcription of all the available *lcp* genes in the mutants Δ6421 and ΔΔ(5775+6421) than the wild type levels, is comparable to our data on lysozyme resistance conferred upon by these two mutants (Fig 3.12, B), that showed increased resistance to lysozyme at 0.125 mg/mL concentration than the wild type strain. However, the mutants appeared more sensitive to 0.5 mg/mL lysozyme concentration in the resazurin-based assay, probably because they were induced for a longer period (1 hour), than the real-time PCR-based experiment where the mutants were induced for a shorter time period of 30 minutes only.

Based on our results on differential expression of the *lcp* genes, it would thus be interesting to see the complete transcription profile for these mutants, to determine what other genes are affected similarly in those mutants, as this might point to particular regulators. A subsequent *in silico* analysis was thus performed to predict putative regulatory elements that might govern the expression of these genes.
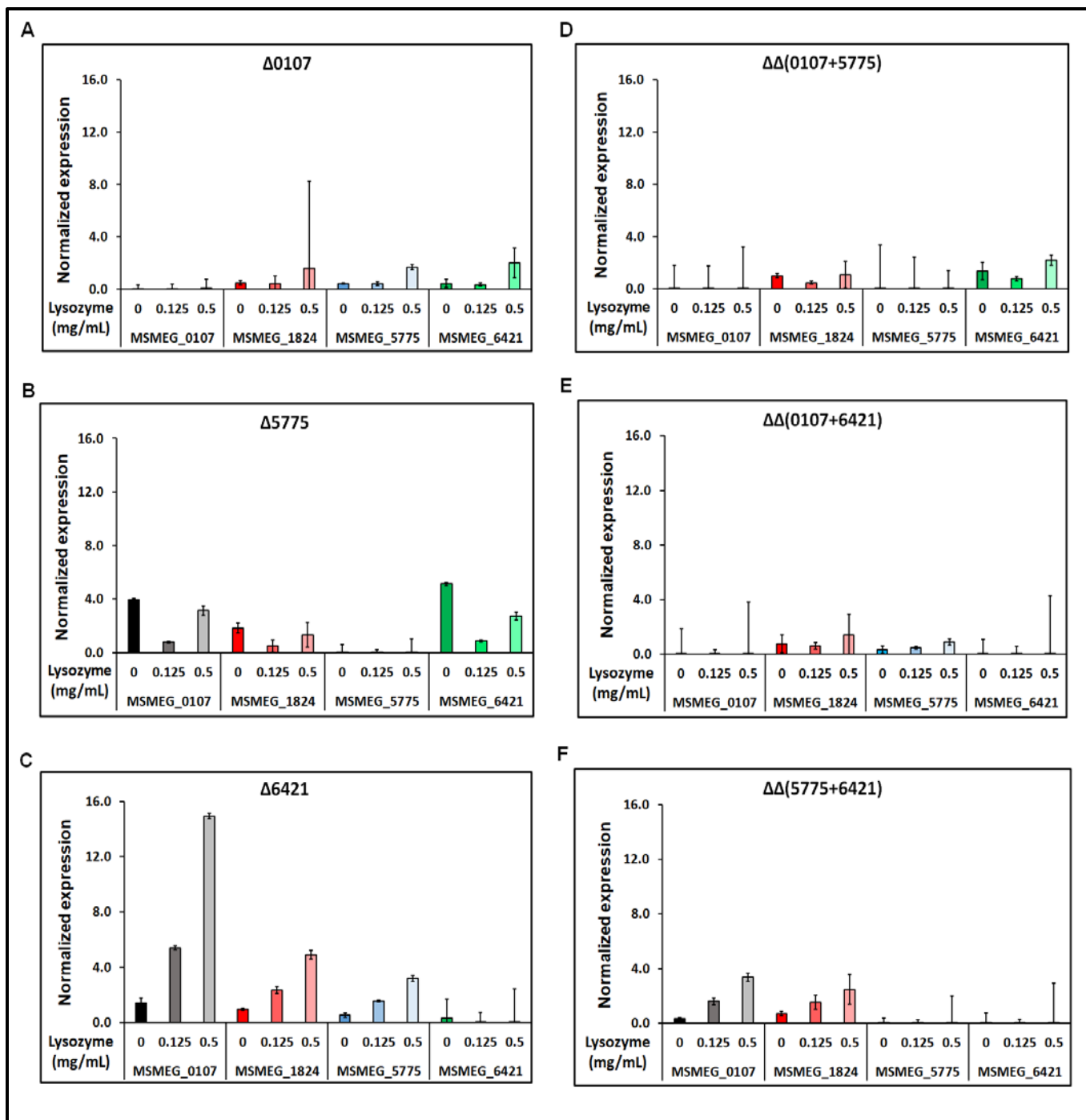
Fig 4.5. Differential expression of *lcp* genes in *Msmeg lcp* deletion mutants under lysozyme stress. The relative expression of *lcp* genes was determined with treated and untreated samples of the deletion mutants. Two different lysozyme concentrations (0.125 mg/mL and 0.5 mg/mL) were used for the treated samples. The data presented here are relative expression normalized to that of the wild type strain mc²155, at the given lysozyme concentration. This is a representative experiment out of two independent experiments each done in triplicate. Error bars represent standard error of triplicate values.

4.3.6. Prediction of regulatory elements of the *lcp* genes

The ability of *Msmeg* to flourish and survive under stress conditions largely depends on the regulation of gene expression. In order to predict regulatory elements responsible for regulating the expression of *lcp* genes in *Msmeg*, first the operons encompassing each of the *lcp* genes were determined by DMINDA webserver (Fig 4.6, A-D). Subsequently, MEME-suite was used to find potential transcriptional motifs. Since the regulatory elements are usually present in the upstream region of the target gene, therefore, the sequences upstream of the start of the target *lcp* gene until the positioned start site of the adjacent open reading frame (ORF) were selected. In this manner the sequences selected were of 1560 bp, 982 bp, 1606 bp and 1061 bp upstream region of *MSMEG_0107*, *MSMEG_1824*, *MSMEG_5775* and *MSMEG_6421*. Several DNA-binding sites were predicted for each *lcp* gene, but the site with the lowest *E*-value was considered to be the putative binding site for the regulation of the target *lcp* gene. This binding site has many motifs on it that matches the transcriptional motif of an organism, which is retrieved from the database. These transcriptional motifs (Fig 4.6) are sorted by their *P*-values. *P*-values are probability values that a random motif of the same width as the target would have an optimal alignment with a match score as good as or better than the target's. In all the *lcp* genes except *MSMEG_6421*, the transcriptional motifs were identified to align significantly with AmrZ motif of *Pseudomonas aeruginosa* (Fig 4.6, A-C). This bi-functional transcriptional factor that acts either as a repressor or activator on a wide range of gene targets is involved in *P. aeruginosa* virulence. Furthermore, it is considered as a key global regulator involved in environmental sensing and adaption (Jones *et al.*, 2014, Martinez-Granero *et al.*, 2014). TF that was predicted for *MSMEG_6421* with the lowest *E*-value was LexA, which is a transcriptional repressor protein involved in the SOS response in *Mtb*. It has a coordinated activation of a cohort of genes required for DNA repair and mutagenesis in response to DNA damage in *Mtb* (Smollett *et al.*, 2012) and *C. glutamicum* (Jochmann *et al.*, 2009). In all, motifs were found that maximally matched *P. aeruginosa* motifs. The motifs that were more related to mycobacteria or actinobacteria were GlxR in *C. glutamicum*, DosR in *Mtb*, EspR in *Mtb*.

Though the regulatory elements involved in the regulation of *lcp* genes have not been studied before, this preliminary *in silico* analysis can serve as a starting point to understand the regulatory mechanisms of these cell envelope proteins. Such information can help understand the differential regulation of the multiple homologues much better.

A

pntA | pntB | MSMEG_0108 | MSMEG_0107 [LCP] | MSMEG_0106

Motif-3
Motif-5
Motif-1
Motif-2
Motif-4

Motif-1 CGATGATCGTCACGCCGTTGTCGGTCACGATGGCCTGGTCCTTGACCGT
(P-value: 6.57E-19)

Motif-2 CGCTGCACACCCCGCTGATGTCGGTGACGAATGCGATCTCGGGCATCAT
(P-value: 1.01E-18)

Motif-3 CCTCGACCACCGCGCGGTAACCGGCGATGTTGGCCATCGACGACAACAC
(P-value: 1.20E-17)

Motif-4 CGTTGCCGATCACGTAGTACCCGATCACGATCGCCAACGCGAACACCAC
(P-value: 1.67E-17)

Motif-5 TCATCACCGCCGAGATGGTGGCGTCGATGAAGGCCGGCAGCGTGATCGT
(P-value: 7.13E-16)

Transcription motif match	Organism match	E-value
AmrZ	Pseudomonas aeruginosa PAO1	2.95E-03
MatP	Escherichia coli str. K-12 substr. MG1655	7.54E-02
VqsM	Pseudomonas aeruginosa PAO1	3.31E-01
IHF	Pseudomonas putida	7.45E-01
Fur	Listeria monocytogenes 10403S	2.46e+00
GlxR	Corynebacterium glutamicum ATCC 13032	2.98e+00
AlgR	Pseudomonas aeruginosa PAO1	3.82e+00
BldD	Streptomyces coelicolor A3(2)	5.26e+00
PvdS	Pseudomonas aeruginosa PAO1	5.40e+00
Fur	Salmonella enterica subsp. enterica serovar Typhimurium str. 14028S	6.04e+00
HrpX	Xanthomonas oryzae pv. oryzae MAFF 311018	6.23e+00

B

MSMEG_1824 [LCP] | MSMEG_1823 | MSMEG_1822

rfbD

Motif-2
Motif-5
Motif-3
Motif-4
Motif-1

Motif-1 CCGCACCGCATCCTCGTCGGTGATGTC
(P-value: 4.98e-12)

Motif-2 CTGCTGCAGATCGTCGACGGCGGTGTC
(P-value: 4.98e-12

Motif-3 CTGATCCACATCTCCACCGACTATGTA
(P-value: 9.30e-12)

Motif-4 CCGCTCGGGCTCGTCCTCGGCCTTGTC
(P-value: 1.27e-10)

Motif-5 CTGATCGGCCACGACGTCGATGGCGCC
(P-value: 5.86e-10)

Transcription motif match	Organism match	E-value
AmrZ	Pseudomonas aeruginosa PAO1	8.52E-01
VqsM	Pseudomonas aeruginosa PAO1	1.20E+00
RutR	Escherichia coli str. K-12 substr. MG1655	1.43E+00
PvdS	Pseudomonas aeruginosa PAO1	1.63E+00
Vfr	Pseudomonas aeruginosa PAO1	1.76E+00
Fur	Caulobacter crescentus	3.00e+00
IHF	Pseudomonas putida KT2440	4.03e+00
LasR	Pseudomonas aeruginosa PAO1	4.42e+00
AlgR	Pseudomonas aeruginosa PAO1	4.62e+00
AdpA	Streptomyces coelicolor A3(2)	5.39e+00
DosR	Mycobacterium tuberculosis H37Rv	7.36e+00
PvdS	Pseudomonas syringae pv. tomato str. DC3000	8.70e+00
Fur	Pseudomonas syringae pv. tomato str. DC3000	9.21e+00
HrpX	Xanthomonas oryzae pv. oryzae MAFF 311018	9.57e+00

C

Transcription motif match	Organism match	E-value
AmrZ	*Pseudomonas aeruginosa PAO1*	1.66e-02
MatP	*Escherichia coli str. K-12 substr. MG1655*	1.70e-01
AdpA	*Streptomyces coelicolor A3(2)*	2.05e-01
PvdS	*Pseudomonas aeruginosa PAO1*	6.50e-01
ArgR	*Pseudomonas aeruginosa PAO1*	1.06e+00
EspR	*Mycobacterium tuberculosis H37Rv*	1.20e+00
IHF	*Pseudomonas putida*	1.88e+00
AlgR	*Pseudomonas aeruginosa PAO1*	1.96e+00
DosR	*Mycobacterium tuberculosis H37Rv*	2.89e+00
BldD	*Streptomyces coelicolor A3(2)*	3.34e+00
HrpX	*Xanthomonas oryzae pv. oryzae MAFF 311018*	3.53e+00
VqsM	*Pseudomonas aeruginosa PAO1*	3.79e+00
IHF	*Pseudomonas putida KT2440*	4.13e+00
Vfr	*Pseudomonas aeruginosa PAO1*	5.75e+00
AmpR	*Pseudomonas aeruginosa PAO1*	7.76e+00
PvdS	*Pseudomonas syringae pv. tomato str. DC3000*	8.02e+00
LasR	*Pseudomonas aeruginosa PAO1*	8.82e+00
RutR	*Escherichia coli str. K-12 substr. MG1655*	9.87e+00

Motif-1 TTCGAGATCATCCCGTCATCGGCTGGGACTGGCTGGCTCCAAGGTTGT
(P-value: 4.80e-20)

Motif-2 TCGCCGACCACATCGACATGAACTTCGGCTGCCCGGTGCCCAAGGTGA
(P-value: 2.96e-16)

Motif-3 ATGTTGATCTTGCCGTGGTCGGCTCCGACTGGCGGAGTCTCACGGTGT
(P-value: 3.23e-16)

Motif-4 ATCCGGTCGCGGCCATCATGGCCAGCGCGTCGTCGGCTTCGAAGATGT
(P-value: 5.70e-16)

Motif-5 AGCGCCATCCCGTCACCATGCACATGACGACGTTCGCGCCCGAGGAGT
(P-value: 1.49e-14)

D

Transcription motif match	Organism match	E-value
LexA	*Mycobacterium tuberculosis H37Rv*	4.45e-01
IHF	*Pseudomonas putida*	7.65e-01
IHF	*Pseudomonas putida KT2440*	8.91e-01
Crp	*Escherichia coli str. K-12 substr. MG1655*	1.24e+00
H-NS	*Vibrio cholerae O395*	1.85e+00
RutR	*Escherichia coli str. K-12 substr. MG1655*	1.99e+00
LexA	*Corynebacterium glutamicum ATCC 13032*	2.38e+00
CRP	*Vibrio vulnificus CMCP6*	2.53e+00
AmrZ	*Pseudomonas aeruginosa PAO1*	4.55e+00
LuxR	*Vibrio campbellii ATCC BAA-1116*	5.10e+00
LexA	*Bacillus subtilis subsp. subtilis str. 168*	6.59e+00
HapR	*Vibrio cholerae MJ-1236*	6.84e+00
Vfr	*Pseudomonas aeruginosa PAO1*	6.86e+00
AlgR	*Pseudomonas aeruginosa PAO1*	8.40e+00

Motif-1 TCGAGTTCTCGATC
(P-value: 1.04e-8)

Motif-2 TCGCCTTCACGATC
(P-value: 5.94e-8)

Motif-3 CCGAGTTCTCGATC
(P-value: 1.46e-7)

Motif-4 TTAACTGCATGATC
(P-value: 1.48e-7)

Motif-5 TCGACATCATGGCC
(P-value: 1.06e-6)

Fig 4.6. Schematic representation of the locations of the predicted transcriptional motifs of *lcp* genes. The target *lcp* genes **A. *MSMEG_0107* B. *MSMEG_1824* C. *MSMEG_5775*** and **D. *MSMEG_6421*** have been labeled along with other adjacent genes that are a part of the single operon inside thick black arrows. The adjacent gene(s) that are not a part of the operon are labeled inside thick grey arrows. The direction of gene expression is represented by the direction of arrow. The promoter sequence as predicted in DMINDA^2 server is shown as a yellow rectangular box. The initiation site for gene transcription is represented by a bent arrow over the gene. The transcriptional motifs for the expression of *lcp* genes that were predicted in MEME suite webserver are drawn to scale using Snapgene software. The motifs numbering are done in the ascending order of *P*-values of the given motif as predicted in the MEME analysis are in parentheses. The *P*-value in MEME analysis is the probability that a random motif of the same width as the target would have an optimal alignment with a match score as good as or better than the target's. The motifs predicted on the sense and anti-sense strands are represented by right and left red arrow respectively. The transcription motif sites and their source organism that matched the *Msmeg* motifs in MEME analysis, have been described in the box with their *E*-values. The *E*-values represent the statistical significance of the motif and are displayed in ascending order. The lowest *E*-value are the most statistically significant.

4.4. Conclusions

Prokaryotic cells respond to environmental perturbations by acclimatizing their gene expression or by modifying the stability of their existing proteins. A study performed previously in *Msmeg* to study its adaptation to different stresses for survival, has been associated with different growth stages (Drapal *et al.*, 2014). Since in our study, cultures from a late log phase were used therefore, it is likely to be expected that the differential expression profile in between the *Msmeg lcp* homologues might vary from that in a different growth stage. In this study, differential expression of the *lcp* homologues in *Msmeg* wild type strain as well as the mutants under standard as well as stress conditions were observed. The *Msmeg* strain lacking *MSMEG_5775* was seen to display a 3-fold to 7-fold upregulation of all the *lcp* genes. Under both acidic and lysozyme stress, loss of *MSMEG_5775* downregulates all the other *lcp* genes available in the *Msmeg* genome, whereas, loss of *MSMEG_6421* upregulates them. Thus the participation of regulator(s) for such kind of differential regulation comes into picture. From *in silico* analysis, putative transcriptional motifs in *Mtb* and *C. glutamicum* were retrieved, however experimental evidences are necessary to confirm their role in *Msmeg*. The experimental data from this study indicates the role of all four *lcp* homologues in *Msmeg* that differentially gets regulated to act upon stress situations. Such regulation might help the organism to survive the adversity. Overall, a bigger understanding of such stress response profiles incurred by the *Msmeg lcp* genes will gain more insights into the differences between pathogenic *Mtb* and nonpathogenic *Msmeg*, and will help us to understand the pathogenesis of *Mtb* much better.

Chapter-5

General Discussion

LCP proteins are often present as multiple homologues in almost all Gram-positive bacteria including mycobacteria (Hubscher *et al.*, 2008). They play a crucial role in the cell wall assembly by catalyzing the transfer of glycopolymers from a polyprenyl pyrophosphate carrier to numerous cell envelope acceptors including PG (Kawai *et al.*, 2011, Chan *et al.*, 2013, Liszewski Zilla *et al.*, 2015, Wang *et al.*, 2015, Baumgart *et al.*, 2016, Grzegorzewicz *et al.*, 2016, Harrison *et al.*, 2016). In mycobacteria, these glycopolymers are transferred from AG onto the C-6 hydroxyl of MurNAc residues in PG. This coupling process between these two major cell envelope heteropolysaccharides, results in hydrolysis of the pyrophosphate group to release inorganic phosphate, and thus LCP proteins are said to have pyrophosphatase activity (Grzegorzewicz *et al.*, 2016). Many studies have suggested the participation of more than one LCP homologue to exhibit this enzymatic activity. For instance, in *Mtb*, three out of four LCP proteins show pyrophosphatase activity (Grzegorzewicz *et al.*, 2016) whereas, in *C. glutamicum*, only one essential protein was reported to exhibit this function (Baumgart *et al.*, 2016). In our study, all the four *Msmeg* LCP homologues were found to demonstrate this activity at varied levels (Fig 2.10), suggesting multiple LCP proteins to be involved in the cell wall assembly in *Msmeg*. This may account for the genome size of *Msmeg* (6.99 MB) which is larger than most of the *Mycobacterium* spp. (Prasanna & Mehra, 2013). While MSMEG_5775 displayed a 2-fold higher pyrophosphatase activity than MSMEG_1824, MSMEG_0107 showed about 1.4-fold higher pyrophosphatase activity than this essential protein. On the other hand, MSMEG_6421 showed a 2.2-fold reduced activity. In addition to this, differential expression of *lcp* genes in the wild *Msmeg* strain under standard no-stress conditions at a late log phase (Fig 4.1), revealed the essential gene *MSMEG_1824* to have 1.6-fold to 2.1-fold reduced expression compared to the rest of the three *lcp* homologues. Such coexistence in the variation of pyrophosphatase activity and higher transcription levels of *MSMEG_0107* and *MSMEG_5775* compared to the essential gene *MSMEG_1824*, leads us to speculate that higher number of mRNA transcripts of both *MSMEG_0107* and *MSMEG_5775* results in higher pyrophosphatase activity. In order to rule out the possibility of different expression levels of the *lcp* genes at different growth phases, a post-study was done to determine the differential expression of *lcp* homologues in *Msmeg* wild type strain at an early log phase (OD_{600} of 0.5-0.8) (data not shown here). There was a 1.6 to 3.6-fold

reduced expression of *MSMEG_1824* compared to the rest three non-essential *lcp* genes at early log phase, and a 2-fold reduced expression of this essential gene at early log phase compared to that of the late log phase. Thus it seems that, though *MSMEG_1824* is essential in linking AG to PG in *Msmeg*, the rest three non-essential *lcp* genes play a crucial role in aiding this essential gene. Finally, the results of pyrophosphatase activity obtained in our study does not conclude the ligation of AG to PG by the *lcp* homologues in *Msmeg*. Besides, it was previously demonstrated that decaprenyl-1-monophosphate co-purifies with the LCP protein of *Mtb* and is likely to be bound within the hydrophobic cavity of the protein (Harrison *et al.*, 2016). This leads us to speculate if the inorganic phosphate released in our study was obtained after GPP was cleaved by the *Msmeg* LCPΔTM proteins, or from the co-purified decaprenyl-1-monophosphate, or both. This can be elucidated by determining the ligation activity between AG and PG, an approach that has been previously established using a cell-free biochemical functional assay in *Mtb* where the *in vitro* covalent attachment of AG intermediates to PG was investigated by a [14C] carbohydrate radiolabeling assay (Harrison *et al.*, 2016) that monitored the synthesis of Dec-P-P-linked intermediates of mycobacterial cell wall biosynthesis (Birch *et al.*, 2009).

To study the effects of LCP proteins on their morphology, ability to form biofilms, growth patterns and other such physiological features, a panel of single and double *lcp* deletion mutants were generated. Unfortunately, we have had to deal with certain limitations in our study. Since, *MSMEG_1824* is an essential gene, and the resources to make a conditional mutant were unavailable, the non-essential *lcp* genes were only focused, in our study. An almost complete deletion of *MSMEG_0107* was possible, however the total deletion of the other two non-essential genes i.e., *MSMEG_5775* and *MSMEG_6421* could not be generated, either due to the presence of overlapping genes, or due to cloning incompatibilities. Thus, there were some remnants of these genes remaining in the deletion mutants of *MSMEG_5775* and *MSMEG_6421*. This could be reasoned for the complemented strains to not restore the phenotype of the wild type *Msmeg*. The other reason could be the use of a non-native promoter of the *lcp* gene and using the endogenous Phsp60 promoter of the pSMT3 vector that induces higher expression of genes in mycobacteria (Movahedzadeh & Bitter, 2009).

LCP family proteins have shown to be associated with cell surface properties (Bender *et al.*, 2003, Wang *et al.*, 2015, Baumgart *et al.*, 2016). When the *Msmeg lcp* mutants were subjected to such physiological assays, it was found that the strain lacking both *MSMEG_0107* and *MSMEG_5775* was the most compromised of all mutants (Table 5.1).

Table 5.1: Overview of the significant cell physiology alterations exhibited by *lcp* deletion mutants and their complemented strains compared to the wild type strain

Strain	Growth	Colony morphology	Pellicle formation	Biofilm adherence	Cellular aggregation	MIC	SDS stress	Lysozyme stress	Detection of Gal*f* moiety
WT	■	■	■	■	■	■	■	■	■
Δ0107									
C-0107								Sensitive	■
Δ5775		Altered			Slow			Sensitive	
C-5775								Sensitive	■
Δ6421									
C-6421									■
ΔΔ(0107+5775)	Slow	Size and morphology altered	Thin layer	Higher	Slow	More sensitive	More sensitive	Highly sensitive	Highly significant
C-(0107+5775)	Slower								■
ΔΔ(0107+6421)		Altered						Resistant	
C-(0107+6421)		Altered						Resistant	■
ΔΔ(5775+6421)		Altered	Thick layer	Higher			More sensitive		
C-(5775+6421)	Slower							Sensitive	■

The mutant ΔΔ(0107+5775), showed altered morphology not only under SEM (Fig 3.7) but also on Congo red plates with an erose texture of the colony (Fig 3.8, A). Apart from this, this strain also displayed slower growth rate (Fig 3.6), diminished biofilm formation on the air-liquid interface (Fig 3.9, A) and slower aggregation rate (Fig 3.10), all of which indicates that the cell envelope of this mutant exhibits decreased surface hydrophobicity. However, a preliminary TLC experiment displayed increased level of outermost cell envelope GPLs in ΔΔ(0107+5775) and ΔΔ(5775+6421) (Fig 3.11). The surface hydrophobicity of the wild type, mutants and the complemented strains were quantified using a crystal violet method. The double *lcp* deletion mutant ΔΔ(0107+5775) displayed reduced pellicle formation (biofilm on air-liquid interface) but significantly higher biofilm adherence to the walls of polystyrene plate. This is due to the association of pellicles with lipids

in the mycolic acid layer (Ojha *et al.*, 2005), and biofilm adherence on polystyrene 96-well plate with GPLs present in *Msmeg* cell wall (Recht *et al.*, 2000, Recht & Kolter, 2001, Schorey & Sweet, 2008).

The lipid rich mycolic acid layer of the mycobacterial cell envelope acts as an impermeable barrier to numerous antimicrobial compounds (Jarlier & Nikaido, 1994). The LCP proteins are known to aid the transfer of AG precursors to uncrosslinked PG (Hancock *et al.*, 2002, Schaefer *et al.*, 2018). With the compromised AG-PG ligase function of the double *lcp* deletion mutant $\Delta\Delta$(0107+5775), the hyper susceptibility to BAC could therefore be due to an additional shortage of the lipid carriers. Since, BAC targets the carrier lipid decaprenyl pyrophosphate (C_{50}-PP) (Stone & Strominger, 1971), the transportation of PG precursors across the cell membrane is inhibited (Valvano, 2008) by preventing dephosphorylation of the decaprenyl pyrophosphate (Stone & Strominger, 1971, Ming & Epperson, 2002, Economou *et al.*, 2013, Kingston *et al.*, 2014) (Fig 5.1). Therefore, $\Delta\Delta$(0107+5775) which already has a compromised cell envelope due to inefficient AG-PG coupling, becomes highly susceptible to BAC by inhibiting lipid carrier recycling (Qi *et al.*, 2008). Additionally, in this study, $\Delta\Delta$(0107+5775) was also found to be hyper sensitive to VAN which is known to inhibit the crosslinking of PG precursors **D**-alanyl–**D**-alanine present at the carboxy terminus, by inhibiting the transpeptidase activity on these precursor molecules (Walsh, 2000). Thus, synergistic effects between MSMEG_0107-MSMEG_5775 deficiency and peptidoglycan-targeting antibiotics is speculated in *Msmeg*. Not only the peptidoglycan-targeting drugs, but $\Delta\Delta$(0107+5775) owing to its ruptured cell envelope and increased permeability, was also found to be sensitive to the cationic detergent SDS and the PG targeting lysozyme.

Unlike *M. marinum*, where loss of CpsA exhibited altered cell morphology and enhanced cell wall permeability (Wang *et al.*, 2015), the homologue in *Msmeg* i.e., Δ0107 did not produce significant alterations in colony morphology on Congo red plate (Fig 3.8, A). Some organisms have been known to survive under adverse conditions and become extremely resistant to antimicrobial factors, despite of having an altered cell wall (Rosu *et al.*, 2013). In our study, a single deletion of either *MSMEG_0107* or *MSMEG_5775* showed similar phenomenon of about 8-fold and 2-fold resistance to INH respectively (Table 3.5), which is a mycolic acid targeting first line drug against TB. Such resistance to INH was also observed in two complemented strains, c-0107 (8-fold) and c-6421 (2-fold). An increased resistance to INH to the mutants may be associated with a downregulation in the *kat*G gene or a mutation in the *kat*G gene of the mutant strain (Alekshun & Levy, 2007, Ando *et al.*, 2011, Warner & Mizrahi, 2013). However, this needs further investigation by sequencing the *kat*G gene in these mutant strains (Δ0107 and Δ5775) and complemented strains (c-0107 and c-6421).
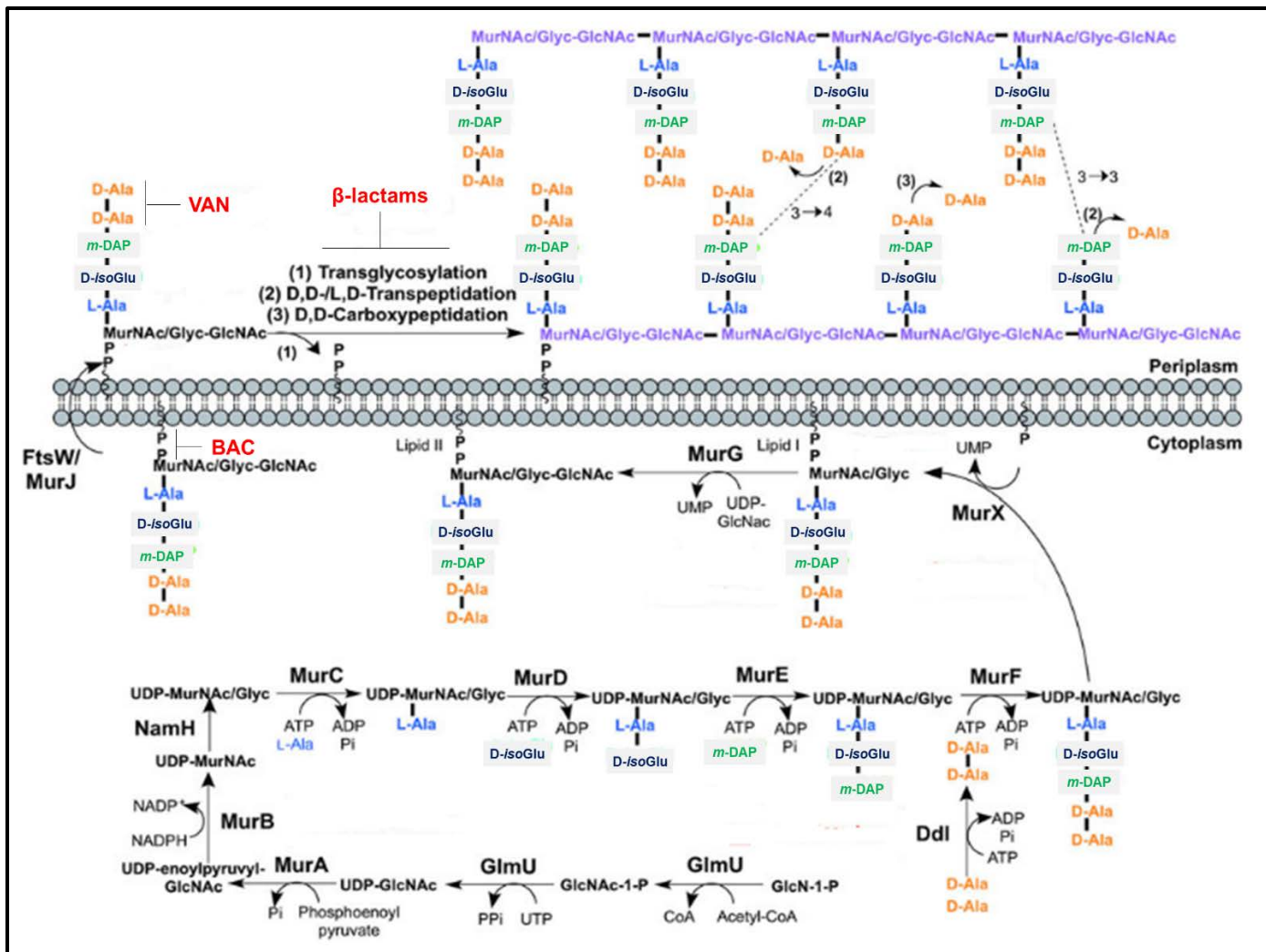
Fig 5.1. Inhibitors used in this study that target peptidoglycan biosynthesis. The roles of the key enzymes involved in peptidoglycan biosynthesis are illustrated. Reported inhibitors are shown in red. **Adapted from (Abrahams & Besra, 2016).**

A novel approach to study the LCP proteins in this thesis, was assessing the galactan moieties in the wild type and the deletion mutants by using of EB-A2 mAB. The Gal*f* moiety has been previously detected in *Aspergillus niger* by anti-Gal*f* EB-A2 mAB (Chiodo *et al.*, 2014) and was hypothesized to be detected in *Msmeg* mutants owing to the altered cell envelope. Though the galactomannan disaccharide fragment, Gal*f*-β-(1→5)-Gal*f* acts as the epitope for EB-A2 detection in *Aspergillus* species, however mAB EB-A2 used in the Platelia kit was recently found to detect multiple epitopes of circulating galactomannan (Krylov *et al.*, 2019). The mAB EB-A2 was however found to be useful in our study for the sole purpose of detecting exposed galactan moieties in the *lcp* mutants, and a significant 5-fold increase of detectable Gal*f* moieties was observed in the double *lcp* deletion mutant ΔΔ(0107+5775). This indicates the exposure of these galactan moieties as a result of a leaky cell envelope of this mutant (Fig 5.2).
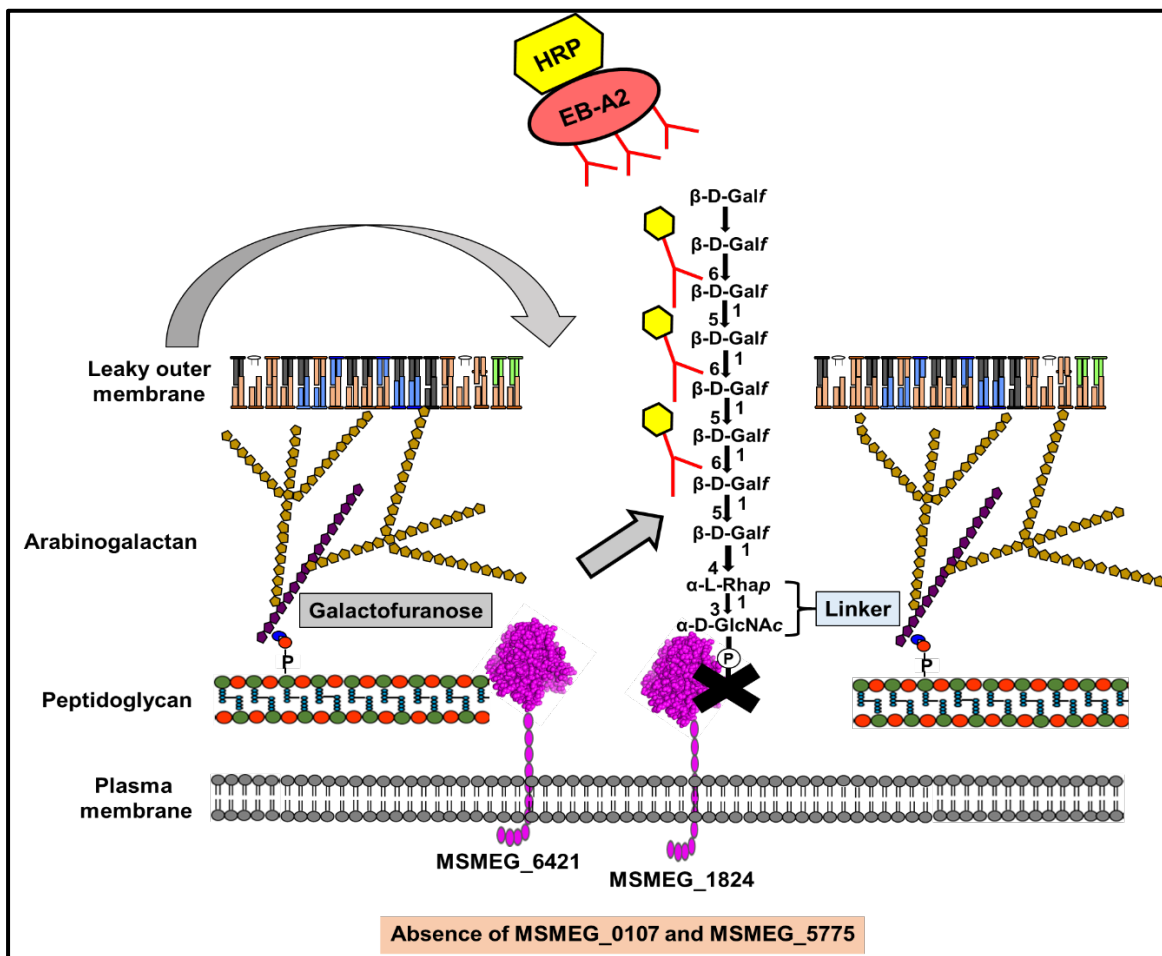
Fig 5.2. Schematic representation of the detection of Gal*f*-β-(1→5)-Gal*f* epitope by EB-A2 Mab in the double *lcp* deletion mutant ΔΔ(0107+5775). With the loss of MSMEG_0107 and MSMEG_5775 in *Msmeg*, phosphotransferase activity is lost from these two LCP proteins. This results in a partial linkage of the AG to PG leading to the formation of a compromised cell envelope, with a leaky outer membrane. Such a leaky cell envelope results in an exposure of the Gal*f* moieties of the AG which are otherwise embedded in the cell mycobacterial core. On addition of HRP-conjugated EB-A2 mAB, the Gal*f* moiety is detected.

The transcriptional profiling of the *lcp* genes in the wild type and mutant strains were studied in late log phase, in this thesis. The significance of the *lcp* gene MSMEG_5775 in *Msmeg* under standard non-stress condition, is associated with the display of 3-fold to 7-fold upregulation of the other three *lcp* genes in Δ5775, compared to that in the wild type strain (Fig 4.2). Under both acid and lysozyme stress conditions, Δ5775 downregulated all other *lcp* genes whereas Δ6421 upregulated these genes. Thus, this study suggests the role of multiple *lcp* homologues in *Msmeg* genome to differentially regulate under various environmental conditions as a survival strategy of the organism. Moreover, such stress response profiles incurred by the *Msmeg lcp* genes during different time points of the log phase might be useful to understand the mechanism of the bacterium to survive under normal and stress conditions. Such studies might also help us to distinguish the stress mechanisms between pathogenic *Mtb* and nonpathogenic *Msmeg*, thus contributing to comprehend the pathogenesis of *Mtb* much better.

Finally, *in silico* analysis to predict putative transcriptional motifs provides a scope to explore the regulatory elements of this family of proteins. Based on MEME analysis, a bi-functional regulator AmrZ was found to regulate the gene expression of *MSMEG_0107*, *MSMEG_1824* and *MSMEG_5775*, whereas a transcriptional repressor protein LexA was found to regulate *MSMEG_6421*. However, these *in silico* findings need experimental validation to trace putative regulatory mechanisms conferred on *Msmeg lcp* genes.

On a different note, Rv2700, an essential gene in *Mtb* required for optimal growth *in vitro* (Ballister *et al.*, 2019) encodes a conserved putative secreted alanine rich protein with a predicted lipid-binding domain related to TraT, LytR and CpsA (Samanovic *et al.*, 2015). This gene has been found to co-occur with Rv3484, which encodes CpsA in string-db webserver (https://string-db.org/cgi/network.pl?taskId=SoOHkscmS8te). This gene with an LytR_C domain was proposed to contribute to cell wall integrity and virulence, suggesting that this domain may be a significant contribution to the cell envelope functions (Ballister *et al.*, 2019). The presence of five conserved residues in this domain was also found in our study. This in addition to the presence of a CXN conserved motif across species, indicates a hidden significance of this domain. Hence, investigating the role of this domain and its contribution to LCP domain could be an added advantage to understand the LCP family of proteins much better.

Finally, with the knowledge that LCP proteins catalyze the phosphotransferase reaction where galactan intermediates of AG coupled with decaprenyl phosphate are transferred to the 6'-OH of MurNAc residues present in PG (Grzegorzewicz *et al.*, 2016, Harrison *et al.*, 2016), this study hypothesizes the low levels of mature AG formation with the absence of MSMEG_0107 and

MSMEG_5775 in *Msmeg* mutant ΔΔ(0107+5775). LCP proteins first recognize the translocation of the galactan intermediates across the plasma membrane and then prepares itself to bind to the LU (Harrison *et al.*, 2016). In our study, the mutant strain ΔΔ(0107+5775), is believed to possess a partial ability to bind to this LU thus, partly coupling these galactan intermediates to the C6-hydroxyl of MurNAc in PG. This partial coupling is likely to be conferred upon by the presence of MSMEG_6421 and the essential gene MSMEG_1824 in this mutant, thus leading to inefficient cleavage of the pyrophosphate group, which possibly releases limited amount of free decaprenyl phosphate. Subsequently, the uncleaved or partially-cleaved pyrophosphate group accumulates in the periplasm thus, limiting the formation of mature mAGP complex (Fig 5.3). This in turn is likely to affect the final anchoring of the terminal β-**D**-Ara*f* and the penultimate 2-α-**D**-Ara*f* to the mycolic acids (McNeil *et al.*, 1991). The cell envelope of ΔΔ(0107+5775) is thus expected to have a leaky outer membrane because of the inefficient attachment of the anchoring Ara*f* residues present in AG to the mycolic acids.
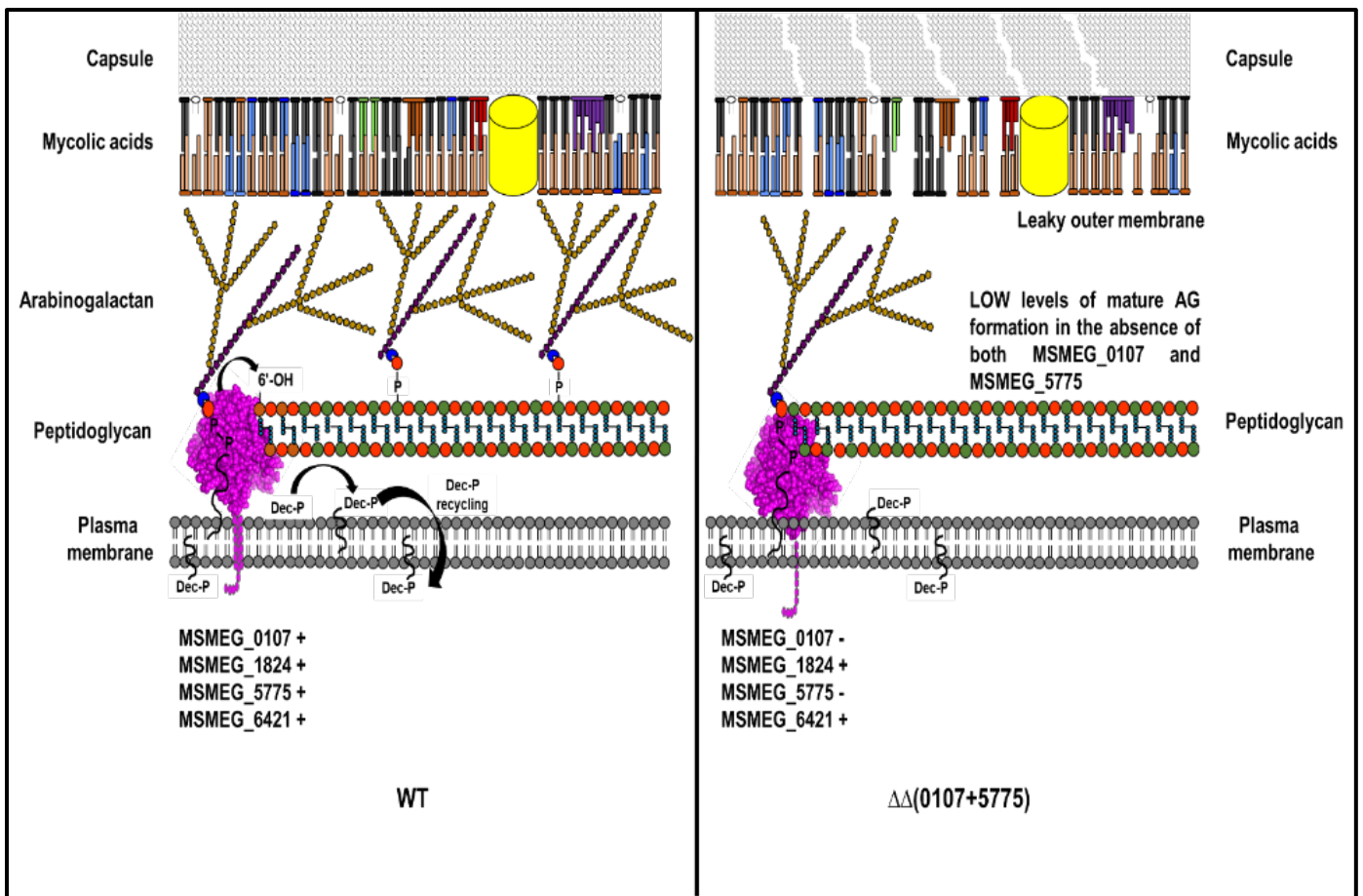

Fig 5.3. Schematic representation of the impact of presence and absence of both *MSMEG_0107* and *MSMEG_5775 lcp* genes in *Msmeg*.

Final remarks and future perspectives

In this thesis, the role of all the four *lcp* homologues in *Msmeg* were studied in a comprehensive manner. All of these four homologues displayed pyrophosphatase activity at varied levels by cleaving the pyrophosphate group of the commercial substrate GPP, to release free inorganic phosphate. However, a single concentration of GPP and the proteins were used in our study to determine the pyrophosphatase activity. A future work aims at using various concentrations of the sample proteins and the substrate to study enzyme kinetics of the LCP proteins. A limitation of the $P_i$per method to determine the pyrophosphatase activity is its highly sensitive protocol to determine the presence of inorganic phosphate. Since, inorganic phosphate is abundantly found in nature, the probability of contamination is high in such sensitive assays. Besides, the inorganic phosphate estimation in our study requires further investigation, if it was obtained after GPP-cleavage by the LCP proteins, or from the co-purified decaprenyl-1-monophosphate, or both, as has been previously demonstrated in *Mtb* (Harrison *et al.*, 2016). Apart from this, further investigations are required to study which amino acid residues of the LCP proteins in *Msmeg* play crucial roles in binding and coordinating the sugar residues of both Dec-P-P-GlcNAc-Rha-galactose (natural substrate) and GPP (artificial substrate used in this study).

The role of LCP proteins in this thesis were further analyzed by studying their impact on cell physiology. This was attained by the construction of single and double deletion mutants for the non-essential *Msmeg lcp* genes, and the double *lcp* deletion mutant ΔΔ(0107+5775) was found to be the most compromised mutant of all. This was demonstrated by slower growth rate, altered colony morphology, reduced biofilm formation on air-liquid interface and enhanced biofilm adherence to polystyrene plate, slower aggregation rate and higher sensitivity to antibiotics used in this study. However, the compromised cell envelope of ΔΔ(0107+5775) is also hypothesized to contain altered capsular polysaccharides resulting in increased permeability. Therefore, studying the impact of *lcp* deletion mutants on the capsular α-glucan production by a method previously established (van de Weerd *et al.*, 2016) would be an interesting approach to understand the underlying mechanism of LCP proteins in cell envelope biogenesis. Since, this thesis hypothesizes the reduced levels of mature AG in the double *lcp* deletion mutant ΔΔ(0107+5775), estimating the saccharides content in AG is possibly an interesting approach to elucidate this theory. A novel approach in this thesis was the detection of Gal*f* entities of AG by EB-A2 mAB. Since the experiment design in this study included the use of whole bacterial cells, our conclusions are confined to the exposed Gal*f* motifs only. However, a future direction lies in detecting these Gal*f* motifs in the culture supernatants to identify if the double *lcp* deletion mutant ΔΔ(0107+5775)

releases un-ligated AG into the culture medium. A primary result in our study included the lipid profiling of the wild type strain and the *lcp* deletion mutants except ΔΔ(0107+6421), where a high content of GPLs was observed in ΔΔ(0107+5775). A future perspective lies in reproducibility of this TLC result and studying the effect of GPLs in the *lcp* deletion mutants by transcriptional profiling and promoter analysis. The primary SEM results displayed alterations in the surface morphology of ΔΔ(0107+5775). An immediate future approach is to reproduce the results and quantify the altered size in this double *lcp* deletion mutant. Finally, creating a conditional mutant of the *Msmeg* essential *lcp* gene i.e., *MSMEG_1824* and a conditional triple deletion mutant of the three non-essential *lcp* genes also remains a high priority to study its impact in cell envelope biogenesis, which can be attained by a method previously established in *Msmeg* (Harrison *et al.*, 2016).

It was found that the four *lcp* homologues in *Msmeg* differentially express when encountered with different external stimuli. Thus it seems that, though *MSMEG_1824* is essential in linking AG to PG in *Msmeg*, the rest three non-essential *lcp* genes play a crucial role in aiding this essential gene during various stress conditions. However, the results obtained from this study were confined to the transcriptional profiling at a late log phase of the bacterial culture under stress or stress-free conditions, where the bacterial cells stop dividing. Therefore, transcriptional profiling from bacterial cultures grown at early- and mid-log phase would provide us a better insight of the differential expression of the *Msmeg lcp* homologues. Additionally, identifying the key regulatory elements of the *Msmeg lcp* genes would contribute to understanding the LCP family of proteins in the non-pathogenic *Mycobacterium* exclusively.

References

Abrahams, K.A., and Besra, G.S. (2016) Mycobacterial cell wall biosynthesis: a multifaceted antibiotic target. *Parasitology* **145**: 116-133.

Agrawal, P., Miryala, S., and Varshney, U. (2015) Use of *Mycobacterium smegmatis* deficient in ADP-ribosyltransferase as surrogate for *Mycobacterium tuberculosis* in drug testing and mutation analysis. *PLoS One* **10**: e0122076.

Alderwick, L.J., Birch, H.L., Krumbach, K., Bott, M., Eggeling, L., and Besra, G.S. (2018) AftD functions as an alpha1-->5 arabinofuranosyltransferase involved in the biosynthesis of the mycobacterial cell wall core. *Cell Surf* **1**: 2-14.

Alderwick, L.J., Dover, L.G., Seidel, M., Gande, R., Sahm, H., Eggeling, L., and Besra, G.S. (2006a) Arabinan-deficient mutants of *Corynebacterium glutamicum* and the consequent flux in decaprenylmonophosphoryl-D-arabinose metabolism. *Glycobiology* **16**: 1073-1081.

Alderwick, L.J., Dover, L.G., Veerapen, N., Gurcha, S.S., Kremer, L., Roper, D.L., Pathak, A.K., Reynolds, R.C., and 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**: 332-341.

Alderwick, L.J., Lloyd, G.S., Lloyd, A.J., Lovering, A.L., Eggeling, L., and Besra, G.S. (2011) Biochemical characterization of the *Mycobacterium tuberculosis* phosphoribosyl-1-pyrophosphate synthetase. *Glycobiology* **21**: 410-425.

Alderwick, L.J., Radmacher, E., Seidel, M., Gande, R., Hitchen, P.G., Morris, H.R., Dell, A., Sahm, H., Eggeling, L., and Besra, G.S. (2005) Deletion of Cg-*emb* in *Corynebacterianeae* leads to a novel truncated cell wall arabinogalactan, whereas inactivation of Cg-*ubi*A results in an arabinan-deficient mutant with a cell wall galactan core. *J Biol Chem* **280**: 32362-32371.

Alderwick, L.J., Seidel, M., Sahm, H., Besra, G.S., and Eggeling, L. (2006b) Identification of a novel arabinofuranosyltransferase (AftA) involved in cell wall arabinan biosynthesis in *Mycobacterium tuberculosis*. *J Biol Chem* **281**: 15653-15661.

Alekshun, M.N., and Levy, S.B. (2007) Molecular mechanisms of antibacterial multidrug resistance. *Cell* **128**: 1037-1050.

Amar, C., and Vilkas, E. (1973) Isolation of arabinose phosphate from the walls of *Mycobacterium tuberculosis H37Ra*. *C R Acad Sci Hebd Seances Acad Sci D* **277**: 1949-1951.

Ando, H., Kitao, T., Miyoshi-Akiyama, T., Kato, S., Mori, T., and Kirikae, T. (2011) Downregulation of *kat*G expression is associated with isoniazid resistance in *Mycobacterium tuberculosis*. *Mol Microbiol* **79**: 1615-1628.

Anes, E., Peyron, P., Staali, L., Jordao, L., Gutierrez, M.G., Kress, H., Hagedorn, M., Maridonneau-Parini, I., Skinner, M.A., Wildeman, A.G., Kalamidas, S.A., Kuehnel, M., and Griffiths, G. (2006) Dynamic life and death interactions between *Mycobacterium smegmatis* and J774 macrophages. *Cell Microbiol* **8**: 939-960.

Angala, S.K., Belardinelli, J.M., Huc-Claustre, E., Wheat, W.H., and Jackson, M. (2014) The cell envelope glycoconjugates of *Mycobacterium tuberculosis*. *Crit Rev Biochem Mol Biol* **49**: 361-399.

Babaoglu, K., Page, M.A., Jones, V.C., McNeil, M.R., Dong, C., Naismith, J.H., and Lee, R.E. (2003) Novel inhibitors of an emerging target in *Mycobacterium tuberculosis*; substituted thiazolidinones as inhibitors of dTDP-rhamnose synthesis. *Bioorg Med Chem Lett* **13**: 3227-3230.

Bae, T., Banger, A.K., Wallace, A., Glass, E.M., Aslund, F., Schneewind, O., and Missiakas, D.M. (2004) *Staphylococcus aureus* virulence genes identified by *Bursa aurealis* mutagenesis and nematode killing. *Proc Natl Acad Sci U S A* **101**: 12312-12317.

Bailey, T.L., Boden, M., Buske, F.A., Frith, M., Grant, C.E., Clementi, L., Ren, J., Li, W.W., and Noble, W.S. (2009) MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res* **37**: W202-208.

Ballister, E.R., Samanovic, M.I., and Darwin, K.H. (2019) *Mycobacterium tuberculosis* Rv2700 contributes to cell envelope integrity and virulence. *J Bacteriol* **201**: e00228-00219.

Barreteau, H., Kovac, A., Boniface, A., Sova, M., Gobec, S., and Blanot, D. (2008) Cytoplasmic steps of peptidoglycan biosynthesis. *FEMS Microbiol Rev* **32**: 168-207.

Barry, C.E., Crick, D.C., and McNeil, M.R. (2007) Targeting the formation of the cell wall core of *M. tuberculosis*. *Infect Disord Drug Targets* **7**: 182-202.

Baumgart, M., Schubert, K., Bramkamp, M., and Frunzke, J. (2016) Impact of LytR-CpsA-Psr proteins on cell wall biosynthesis in *Corynebacterium glutamicum*. *J Bacteriol* **198**: 3045-3059.

Beiter, K., Wartha, F., Hurwitz, R., Normark, S., Zychlinsky, A., and Henriques-Normark, B. (2008) The capsule sensitizes *Streptococcus pneumoniae* to alpha-defensins human neutrophil proteins 1 to 3. *Infect Immun* **76**: 3710-3716.

Belanova, M., Dianiskova, P., Brennan, P.J., Completo, G.C., Rose, N.L., Lowary, T.L., and Mikusova, K. (2008) Galactosyl transferases in mycobacterial cell wall synthesis. *J Bacteriol* **190**: 1141-1145.

Bender, M.H., Cartee, R.T., and Yother, J. (2003) Positive correlation between tyrosine phosphorylation of CpsD and capsular polysaccharide production in *Streptococcus pneumoniae*. *Journal of Bacteriology* **185**: 6057-6066.

Bendinger, B., Rijnaarts, H.H., Altendorf, K., and Zehnder, A.J. (1993) Physicochemical cell surface and adhesive properties of coryneform bacteria related to the presence and chain length of mycolic acids. *Appl Environ Microbiol* **59**: 3973-3977.

Benkert, P., Biasini, M., and Schwede, T. (2011) Toward the estimation of the absolute quality of individual protein structure models. *Bioinformatics* **27**: 343-350.

Berg, D., Youdim, M.B., and Riederer, P. (2004) Redox imbalance. *Cell Tissue Res* **318**: 201-213.

Berg, S., Kaur, D., Jackson, M., and Brennan, P.J. (2007) The glycosyltransferases of *Mycobacterium tuberculosis* - roles in the synthesis of arabinogalactan, lipoarabinomannan, and other glycoconjugates. *Glycobiology* **17**: 35-56R.

Besra, G.S., Khoo, K.H., McNeil, M.R., Dell, A., Morris, H.R., and 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**: 4257-4266.

Bhakta, S., and Basu, J. (2002) Overexpression, purification and biochemical characterization of a class A high-molecular-mass penicillin-binding protein (PBP), PBP1* and its soluble derivative from *Mycobacterium tuberculosis*. *Biochem J* **361**: 635-639.

Bhamidi, S., Scherman, M.S., Rithner, C.D., Prenni, J.E., Chatterjee, D., Khoo, K.H., and 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**: 12992-13000.

Bharati, B.K., Sharma, I.M., Kasetty, S., Kumar, M., Mukherjee, R., and Chatterji, D. (2012) A full-length bifunctional protein involved in c-di-GMP turnover is required for long-term survival under nutrient starvation in *Mycobacterium smegmatis*. *Microbiology* **158**: 1415-1427.

Bhat, S.A., Singh, N., Trivedi, A., Kansal, P., Gupta, P., and Kumar, A. (2012) The mechanism of redox sensing in *Mycobacterium tuberculosis*. *Free Radic Biol Med* **53**: 1625-1641.

Bhat, Z.S., Rather, M.A., Maqbool, M., Lah, H.U., Yousuf, S.K., and Ahmad, Z. (2017) Cell wall: A versatile fountain of drug targets in *Mycobacterium tuberculosis*. *Biomed Pharmacother* **95**: 1520-1534.

Bhatt, A., and Jacobs, W.R., Jr. (2009) Gene essentiality testing in *Mycobacterium smegmatis* using specialized transduction. *Methods Mol Biol* **465**: 325-336.

Birch, H.L., Alderwick, L.J., Bhatt, A., Rittmann, D., Krumbach, K., Singh, A., Bai, Y., Lowary, T.L., Eggeling, L., and Besra, G.S. (2008) Biosynthesis of mycobacterial arabinogalactan: identification of a novel alpha(1-->3) arabinofuranosyltransferase. *Mol Microbiol* **69**: 1191-1206.

Birch, H.L., Alderwick, L.J., Rittmann, D., Krumbach, K., Etterich, H., Grzegorzewicz, A., McNeil, M.R., Eggeling, L., and Besra, G.S. (2009) Identification of a terminal rhamnopyranosyltransferase (RptA) involved in *Corynebacterium glutamicum* cell wall biosynthesis. *J Bacteriol* **191**: 4879-4887.

Bitoun, J.P., Liao, S., Yao, X., Ahn, S.J., Isoda, R., Nguyen, A.H., Brady, L.J., Burne, R.A., Abranches, J., and Wen, Z.T. (2012) BrpA is involved in regulation of cell envelope stress responses in *Streptococcus mutans*. *Appl Environ Microbiol* **78**: 2914-2922.

Borrego, S., Niubo, E., Ancheta, O., and Espinosa, M.E. (2000) Study of the microbial aggregation in *Mycobacterium* using image analysis and electron microscopy. *Tissue Cell* **32**: 494-500.

Brennan, P.J. (2003) Structure, function, and biogenesis of the cell wall of *Mycobacterium tuberculosis*. *Tuberculosis (Edinb)* **83**: 91-97.

Brennan, P.J., and Nikaido, H. (1995) The envelope of mycobacteria. *Annu Rev Biochem* **64**: 29-63.

Briken, V., Porcelli, S.A., Besra, G.S., and Kremer, L. (2004) Mycobacterial lipoarabinomannan and related lipoglycans: from biogenesis to modulation of the immune response. *Mol Microbiol* **53**: 391-403.

Brosch, R., Gordon, S.V., Marmiesse, M., Brodin, P., Buchrieser, C., Eiglmeier, K., Garnier, T., Gutierrez, C., Hewinson, G., Kremer, K., Parsons, L.M., Pym, A.S., Samper, S., van Soolingen, D., and Cole, S.T. (2002) A new evolutionary scenario for the *Mycobacterium tuberculosis* complex. *Proc Natl Acad Sci U S A* **99**: 3684-3689.

Brown-Elliott, B.A., Griffith, D.E., and Wallace, R.J., Jr. (2002) Newly described or emerging human species of nontuberculous mycobacteria. *Infect Dis Clin North Am* **16**: 187-220.

Browning, D.F., and Busby, S.J. (2004) The regulation of bacterial transcription initiation. *Nat Rev Microbiol* **2**: 57-65.

Bryk, R., Lima, C.D., Erdjument-Bromage, H., Tempst, P., and Nathan, C. (2002) Metabolic enzymes of mycobacteria linked to antioxidant defense by a thioredoxin-like protein. *Science* **295**: 1073-1077.

Cangelosi, G.A., Palermo, C.O., Laurent, J.P., Hamlin, A.M., and Brabant, W.H. (1999) Colony morphotypes on Congo red agar segregate along species and drug susceptibility lines in the *Mycobacterium avium-intracellulare* complex. *Microbiology* **145 (Pt 6)**: 1317-1324.

Chackerian, A.A., Alt, J.M., Perera, T.V., Dascher, C.C., and Behar, S.M. (2002) Dissemination of *Mycobacterium tuberculosis* is influenced by host factors and precedes the initiation of T-cell immunity. *Infect Immun* **70**: 4501-4509.

Chan, Y.G., Frankel, M.B., Dengler, V., Schneewind, O., and Missiakas, D. (2013) *Staphylococcus aureus* mutants lacking the LytR-CpsA-Psr family of enzymes release cell wall teichoic acids into the extracellular medium. *J Bacteriol* **195**: 4650-4659.

Chan, Y.G., Kim, H.K., Schneewind, O., and Missiakas, D. (2014) The capsular polysaccharide of *Staphylococcus aureus* is attached to peptidoglycan by the LytR-CpsA-Psr (LCP) family of enzymes. *J Biol Chem* **289**: 15680-15690.

Chiodo, F., Marradi, M., Park, J., Ram, A.F., Penades, S., van Die, I., and Tefsen, B. (2014) Galactofuranose-coated gold nanoparticles elicit a pro-inflammatory response in human monocyte-derived dendritic cells and are recognized by DC-SIGN. *ACS Chem Biol* **9**: 383-389.

Churchward, M.A., Brandman, D.M., Rogasevskaia, T., and Coorssen, J.R. (2008) Copper (II) sulfate charring for high sensitivity on-plate fluorescent detection of lipids and sterols: quantitative analyses of the composition of functional secretory vesicles. *J Chem Biol* **1**: 79-87.

Cieslewicz, M.J., Kasper, D.L., Wang, Y., and Wessels, M.R. (2001) Functional analysis in type Ia group B *Streptococcus* of a cluster of genes involved in extracellular polysaccharide production by diverse species of streptococci. *J Biol Chem* **276**: 139-146.

Collins, C.H., Grange, J.M., and Yates, M.D. (1984) Mycobacteria in water. *J Appl Bacteriol* **57**: 193-211.

Cooper, A.M., Mayer-Barber, K.D., and Sher, A. (2011) Role of innate cytokines in mycobacterial infection. *Mucosal Immunol* **4**: 252-260.

Cumming, B.M., Lamprecht, D.A., Wells, R.M., Saini, V., Mazorodze, J.H., and Steyn, A.J.C. (2014) The physiology and genetics of oxidative stress in mycobacteria. *Microbiol Spectr* **2**: MGM2-0019-2013.

Cummings, R.D., and Etzler, M.E., (2009) Antibodies and lectins in glycan analysis. In: Essentials of Glycobiology. Varki, A., Cummings, R.D., Esko, J.D., Freeze, H.H., Stanley, P., Bertozzi, C.R., Hart, G.W. & Etzler, M.E. (eds). Cold Spring Harbor (NY).

Daffe, M., Brennan, P.J., and McNeil, M. (1990) Predominant structural features of the cell wall arabinogalactan of *Mycobacterium tuberculosis* as revealed through characterization of oligoglycosyl alditol fragments by gas chromatography/mass spectrometry and by 1H and 13C NMR analyses. *J Biol Chem* **265**: 6734-6743.

Daffe, M., and Draper, P. (1998) The envelope layers of mycobacteria with reference to their pathogenicity. *Adv Microb Physiol* **39**: 131-203.

Dalle-Donne, I., Milzani, A., Gagliano, N., Colombo, R., Giustarini, D., and Rossi, R. (2008) Molecular mechanisms and potential clinical significance of S-glutathionylation. *Antioxid Redox Signal* **10**: 445-473.

Deshayes, C., Laval, F., Montrozier, H., Daffe, M., Etienne, G., and Reyrat, J.M. (2005) A glycosyltransferase involved in biosynthesis of triglycosylated glycopeptidolipids in *Mycobacterium smegmatis*: impact on surface properties. *J Bacteriol* **187**: 7283-7291.

Dinadayala, P., Sambou, T., Daffe, M., and Lemassu, A. (2008) Comparative structural analyses of the alpha-glucan and glycogen from *Mycobacterium bovis*. *Glycobiology* **18**: 502-508.

Dogra, N., Arya, S., Singh, K., and Kaur, J. (2015) Differential expression of two members of Rv1922-LipD operon in *Mycobacterium tuberculosis*: Does *rv1923* qualify for membership? *Pathog Dis* **73**.

Donlan, R.M. (2002) Biofilms: microbial life on surfaces. *Emerg Infect Dis* **8**: 881-890.

Dover, L.G., Cerdeno-Tarraga, A.M., Pallen, M.J., Parkhill, J., and Besra, G.S. (2004) Comparative cell wall core biosynthesis in the mycolated pathogens, *Mycobacterium tuberculosis* and *Corynebacterium diphtheriae*. *FEMS Microbiol Rev* **28**: 225-250.

Drapal, M., Perez-Fons, L., Wheeler, P.R., and Fraser, P.D. (2014) The application of metabolite profiling to *Mycobacterium spp*.: determination of metabolite changes associated with growth. *J Microbiol Methods* **106**: 23-32.

Draper, P. (1984) Wall biosynthesis: a possible site of action for new antimycobacterial drugs. *Int J Lepr Other Mycobact Dis* **52**: 527-532.

Draper, P., Khoo, K.H., Chatterjee, D., Dell, A., and Morris, H.R. (1997) Galactosamine in walls of slow-growing mycobacteria. *Biochem J* **327 (Pt 2)**: 519-525.

Du, Y., Zhang, H., Lu, J., and Holmgren, A. (2012) Glutathione and glutaredoxin act as a backup of human thioredoxin reductase 1 to reduce thioredoxin 1 preventing cell death by aurothioglucose. *J Biol Chem* **287**: 38210-38219.

Dwek, R.A. (1996) Glycobiology: Toward Understanding the Function of Sugars. *Chem Rev* **96**: 683-720.

Economou, N.J., Cocklin, S., and Loll, P.J. (2013) High-resolution crystal structure reveals molecular details of target recognition by bacitracin. *Proc Natl Acad Sci U S A* **110**: 14207-14212.

Ernst, W.A., Maher, J., Cho, S., Niazi, K.R., Chatterjee, D., Moody, D.B., Besra, G.S., Watanabe, Y., Jensen, P.E., Porcelli, S.A., Kronenberg, M., and Modlin, R.L. (1998) Molecular interaction of CD1b with lipoglycan antigens. *Immunity* **8**: 331-340.

Espinal, M.A., Laszlo, A., Simonsen, L., Boulahbal, F., Kim, S.J., Reniero, A., Hoffner, S., Rieder, H.L., Binkin, N., Dye, C., Williams, R., and Raviglione, M.C. (2001) Global trends in resistance to antituberculosis drugs. World Health Organization-International Union against Tuberculosis and Lung Disease Working Group on Anti-Tuberculosis Drug Resistance Surveillance. *N Engl J Med* **344**: 1294-1303.

Feng, L., Chen, Z., Wang, Z., Hu, Y., and Chen, S. (2016) Genome-wide characterization of monomeric transcriptional regulators in *Mycobacterium tuberculosis*. *Microbiology* **162**: 889-897.

Fernandes, N.D., Wu, Q.L., Kong, D., Puyang, X., Garg, S., and Husson, R.N. (1999) A mycobacterial extracytoplasmic sigma factor involved in survival following heat shock and oxidative stress. *J Bacteriol* **181**: 4266-4274.

Flynn, J.L., Chan, J., and Lin, P.L. (2011) Macrophages and control of granulomatous inflammation in tuberculosis. *Mucosal Immunol* **4**: 271-278.

Fukuda, T., Matsumura, T., Ato, M., Hamasaki, M., Nishiuchi, Y., Murakami, Y., Maeda, Y., Yoshimori, T., Matsumoto, S., Kobayashi, K., Kinoshita, T., and Morita, Y.S. (2013) Critical roles for lipomannan and lipoarabinomannan in cell wall integrity of mycobacteria and pathogenesis of tuberculosis. *MBio* **4**: e00472-00412.

Fukumori, F., and Kishii, M. (2001) Molecular cloning and transcriptional analysis of the alkyl hydroperoxide reductase genes from *Pseudomonas putida KT2442*. *J Gen Appl Microbiol* **47**: 269-277.

Gale, R.T., Li, F.K.K., Sun, T., Strynadka, N.C.J., and Brown, E.D. (2017) *B. subtilis* LytR-CpsA-Psr enzymes transfer wall teichoic acids from authentic lipid-linked substrates to mature peptidoglycan *in vitro*. *Cell Chem Biol* **24**: 1537-1546 e1534.

Gebhard, S., Humpel, A., McLellan, A.D., and Cook, G.M. (2008) The alternative sigma factor SigF of *Mycobacterium smegmatis* is required for survival of heat shock, acidic pH and oxidative stress. *Microbiology* **154**: 2786-2795.

Ghosh, C., Sarkar, A., Anuja, K., Das, M.C., Chakraborty, A., Jawed, J.J., Gupta, P., Majumdar, S., Banerjee, B., and Bhattacharjee, S. (2019) Free radical stress induces DNA damage response in RAW264.7 macrophages during *Mycobacterium smegmatis* infection. *Arch Microbiol* **201**: 487-498.

Gilleron, M., Himoudi, N., Adam, O., Constant, P., Venisse, A., Riviere, M., and Puzo, G. (1997) *Mycobacterium smegmatis* phosphoinositols-glyceroarabinomannans. Structure and localization of alkali-labile and alkali-stable phosphoinositides. *J Biol Chem* **272**: 117-124.

Gilleron, M., Quesniaux, V.F., and Puzo, G. (2003) Acylation state of the phosphatidylinositol hexamannosides from *Mycobacterium bovis*, Bacillus Calmette Guerin and *Mycobacterium tuberculosis H37Rv* and its implication in Toll-like receptor response. *J Biol Chem* **278**: 29880-29889.

Goffin, C., and Ghuysen, J.M. (2002) Biochemistry and comparative genomics of SxxK superfamily acyltransferases offer a clue to the mycobacterial paradox: presence of penicillin-susceptible target proteins versus lack of efficiency of penicillin as therapeutic agent. *Microbiol Mol Biol Rev* **66**: 702-738, table of contents.

Gonzalez-Zamorano, M., Mendoza-Hernandez, G., Xolalpa, W., Parada, C., Vallecillo, A.J., Bigi, F., and Espitia, C. (2009) *Mycobacterium tuberculosis* glycoproteomics based on ConA-lectin affinity capture of mannosylated proteins. *J Proteome Res* **8**: 721-733.

Goude, R., Amin, A.G., Chatterjee, D., and Parish, T. (2008) The critical role of *emb*C in *Mycobacterium tuberculosis*. *J Bacteriol* **190**: 4335-4341.

Griffin, J.E., Gawronski, J.D., Dejesus, M.A., Ioerger, T.R., Akerley, B.J., and Sassetti, C.M. (2011) High-resolution phenotypic profiling defines genes essential for mycobacterial growth and cholesterol catabolism. *PLoS Pathog* **7**: e1002251.

Grzegorzewicz, A.E., de Sousa-d'Auria, C., McNeil, M.R., Huc-Claustre, E., Jones, V., Petit, C., Angala, S.K., Zemanova, J., Wang, Q., Belardinelli, J.M., Gao, Q., Ishizaki, Y., Mikusova, K., Brennan, P.J., Ronning, D.R., Chami, M., Houssin, C., and Jackson, M. (2016) Assembling of the *Mycobacterium tuberculosis* cell wall core. *J Biol Chem* **291**: 18867-18879.

Gupta, K.R., Kasetty, S., and Chatterji, D. (2015) Novel functions of (p)ppGpp and cyclic di-GMP in mycobacterial physiology revealed by phenotype microarray analysis of wild-type and isogenic strains of *Mycobacterium smegmatis*. *Appl Environ Microbiol* **81**: 2571-2578.

Gupta, R.S., Lo, B., and Son, J. (2018) Phylogenomics and comparative genomic studies robustly support division of the genus *Mycobacterium* into an emended genus *Mycobacterium* and four novel genera. *Front Microbiol* **9**: 67.

Hancock, I.C., Carman, S., Besra, G.S., Brennan, P.J., and 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**: 3059-3067.

Hanson, B.R., Runft, D.L., Streeter, C., Kumar, A., Carion, T.W., and Neely, M.N. (2012) Functional analysis of the CpsA protein of *Streptococcus agalactiae*. *J Bacteriol* **194**: 1668-1678.

Harrison, J., Lloyd, G., Joe, M., Lowary, T.L., Reynolds, E., Walters-Morgan, H., Bhatt, A., Lovering, A., Besra, G.S., and Alderwick, L.J. (2016) Lcp1 Is a phosphotransferase responsible for ligating arabinogalactan to peptidoglycan in *Mycobacterium tuberculosis*. *MBio* **7**: e00972-00916.

Hingley-Wilson, S.M., Sambandamurthy, V.K., and Jacobs, W.R., Jr. (2003) Survival perspectives from the world's most successful pathogen, *Mycobacterium tuberculosis*. *Nat Immunol* **4**: 949-955.

Hmama, Z., Pena-Diaz, S., Joseph, S., and Av-Gay, Y. (2015) Immunoevasion and immunosuppression of the macrophage by *Mycobacterium tuberculosis*. *Immunol Rev* **264**: 220-232.

Hoang, T.T., Ma, Y., Stern, R.J., McNeil, M.R., and 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**: 361-371.

Huang, H., Berg, S., Spencer, J.S., Vereecke, D., D'Haeze, W., Holsters, M., and McNeil, M.R. (2008) Identification of amino acids and domains required for catalytic activity of DPPR synthase, a cell wall biosynthetic enzyme of *Mycobacterium tuberculosis*. *Microbiology* **154**: 736-743.

Huang, H., Scherman, M.S., D'Haeze, W., Vereecke, D., Holsters, M., Crick, D.C., and 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**: 24539-24543.

Hubscher, J., Luthy, L., Berger-Bachi, B., and Stutzmann Meier, P. (2008) Phylogenetic distribution and membrane topology of the LytR-CpsA-Psr protein family. *BMC Genomics* **9**: 617.

Ikeda, M., Wachi, M., Jung, H.K., Ishino, F., and Matsuhashi, M. (1991) The *Escherichia coli mra*Y gene encoding UDP-*N*-acetylmuramoyl-pentapeptide: undecaprenyl-phosphate phospho-*N*-acetylmuramoyl-pentapeptide transferase. *J Bacteriol* **173**: 1021-1026.

Ishihama, A. (2010) Prokaryotic genome regulation: multifactor promoters, multitarget regulators and hierarchic networks. *FEMS Microbiol Rev* **34**: 628-645.

Ishikawa, E., Mori, D., and Yamasaki, S. (2017) Recognition of mycobacterial lipids by immune receptors. *Trends Immunol* **38**: 66-76.

Jackson, M. (2014) The mycobacterial cell envelope-lipids. *Cold Spring Harb Perspect Med* **4**.

Jackson, M., McNeil, M.R., and Brennan, P.J. (2013) Progress in targeting cell envelope biogenesis in *Mycobacterium tuberculosis*. *Future Microbiol* **8**: 855-875.

Jager, W., Schafer, A., Puhler, A., Labes, G., and Wohlleben, W. (1992) Expression of the *Bacillus subtilis sac*B gene leads to sucrose sensitivity in the Gram-positive bacterium *Corynebacterium glutamicum* but not in *Streptomyces lividans*. *J Bacteriol* **174**: 5462-5465.

Jankute, M., Cox, J.A., Harrison, J., and Besra, G.S. (2015) Assembly of the mycobacterial cell wall. *Annu Rev Microbiol* **69**: 405-423.

Jankute, M., Grover, S., Birch, H.L., and Besra, G.S. (2014) Genetics of mycobacterial arabinogalactan and lipoarabinomannan assembly. *Microbiol Spectr* **2**: MGM2-0013-2013.

Jankute, M., Grover, S., Rana, A.K., and Besra, G.S. (2012) Arabinogalactan and lipoarabinomannan biosynthesis: structure, biogenesis and their potential as drug targets. *Future Microbiol* **7**: 129-147.

Jankute, M., Nataraj, V., Lee, O.Y., Wu, H.H.T., Ridell, M., Garton, N.J., Barer, M.R., Minnikin, D.E., Bhatt, A., and Besra, G.S. (2017) The role of hydrophobicity in tuberculosis evolution and pathogenicity. *Sci Rep* **7**: 1315.

Jarlier, V., and Nikaido, H. (1990) Permeability barrier to hydrophilic solutes in *Mycobacterium chelonei*. *J Bacteriol* **172**: 1418-1423.

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

Jiang, T., He, L., Zhan, Y., Zang, S., Ma, Y., Zhao, X., Zhang, C., and Xin, Y. (2011) The effect of *MSMEG_6402* gene disruption on the cell wall structure of *Mycobacterium smegmatis*. *Microb Pathog* **51**: 156-160.

Jin, Y., Xin, Y., Zhang, W., and Ma, Y. (2010) *Mycobacterium tuberculosis* Rv1302 and *Mycobacterium smegmatis* MSMEG_4947 have WecA function and MSMEG_4947 is required for the growth of *M. smegmatis*. *FEMS Microbiol Lett* **310**: 54-61.

Jochmann, N., Kurze, A.K., Czaja, L.F., Brinkrolf, K., Brune, I., Huser, A.T., Hansmeier, N., Puhler, A., Borovok, I., and Tauch, A. (2009) Genetic makeup of the *Corynebacterium glutamicum* LexA regulon deduced from comparative transcriptomics and *in vitro* DNA band shift assays. *Microbiology* **155**: 1459-1477.

Johansson, J., and Curstedt, T. (1997) Molecular structures and interactions of pulmonary surfactant components. *Eur J Biochem* **244**: 675-693.

Jones, C.J., Newsom, D., Kelly, B., Irie, Y., Jennings, L.K., Xu, B., Limoli, D.H., Harrison, J.J., Parsek, M.R., White, P., and Wozniak, D.J. (2014) ChIP-Seq and RNA-Seq reveal an AmrZ-mediated mechanism for cyclic di-GMP synthesis and biofilm development by *Pseudomonas aeruginosa*. *PLoS Pathog* **10**: e1003984.

Kantardjieff, K.A., Kim, C.Y., Naranjo, C., Waldo, G.S., Lekin, T., Segelke, B.W., Zemla, A., Park, M.S., Terwilliger, T.C., and Rupp, B. (2004) *Mycobacterium tuberculosis* RmlC epimerase (Rv3465): a promising drug-target structure in the rhamnose pathway. *Acta Crystallogr D Biol Crystallogr* **60**: 895-902.

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., and Errington, J. (2011) A widespread family of bacterial cell wall assembly proteins. *EMBO J* **30**: 4931-4941.

Khoo, K.H., Dell, A., Morris, H.R., Brennan, P.J., and Chatterjee, D. (1995) Inositol phosphate capping of the nonreducing termini of lipoarabinomannan from rapidly growing strains of *Mycobacterium*. *J Biol Chem* **270**: 12380-12389.

Kiehn, T.E., Edwards, F.F., Brannon, P., Tsang, A.Y., Maio, M., Gold, J.W., Whimbey, E., Wong, B., McClatchy, J.K., and Armstrong, D. (1985) Infections caused by *Mycobacterium avium* complex in immunocompromised patients: diagnosis by blood culture and fecal examination, antimicrobial susceptibility tests, and morphological and seroagglutination characteristics. *J Clin Microbiol* **21**: 168-173.

Kingston, A.W., Zhao, H., Cook, G.M., and Helmann, J.D. (2014) Accumulation of heptaprenyl diphosphate sensitizes *Bacillus subtilis* to bacitracin: implications for the mechanism of resistance mediated by the BceAB transporter. *Mol Microbiol* **93**: 37-49.

Klepp, L.I., Forrellad, M.A., Osella, A.V., Blanco, F.C., Stella, E.J., Bianco, M.V., Santangelo Mde, L., Sassetti, C., Jackson, M., Cataldi, A.A., Bigi, F., and Morbidoni, H.R. (2012) Impact of the deletion of the six *mce* operons in *Mycobacterium smegmatis*. *Microbes Infect* **14**: 590-599.

Kondo, N., Nakamura, H., Masutani, H., and Yodoi, J. (2006) Redox regulation of human thioredoxin network. *Antioxid Redox Signal* **8**: 1881-1890.

Koster, S., Upadhyay, S., Chandra, P., Papavinasasundaram, K., Yang, G., Hassan, A., Grigsby, S.J., Mittal, E., Park, H.S., Jones, V., Hsu, F.F., Jackson, M., Sassetti, C.M., and Philips, J.A. (2017)

Mycobacterium tuberculosis is protected from NADPH oxidase and LC3-associated phagocytosis by the LCP protein CpsA. *Proc Natl Acad Sci U S A* **114**: E8711-E8720.

Kremer, L., and Besra, G.S. (2002) Re-emergence of tuberculosis: strategies and treatment. *Expert Opin Investig Drugs* **11**: 153-157.

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., and Besra, G.S. (2001) Galactan biosynthesis in *Mycobacterium tuberculosis*. Identification of a bifunctional UDP-galactofuranosyltransferase. *J Biol Chem* **276**: 26430-26440.

Krylov, V.B., Solovev, A.S., Argunov, D.A., Latge, J.P., and Nifantiev, N.E. (2019) Reinvestigation of carbohydrate specificity of EB-A2 monoclonal antibody used in the immune detection of *Aspergillus fumigatus* galactomannan. *Heliyon* **5**.

Kumar, P., Arora, K., Lloyd, J.R., Lee, I.Y., Nair, V., Fischer, E., Boshoff, H.I., and Barry, C.E., 3rd (2012) Meropenem inhibits D,D-carboxypeptidase activity in *Mycobacterium tuberculosis*. *Mol Microbiol* **86**: 367-381.

Kumar, S., Stecher, G., and Tamura, K. (2016) MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol Biol Evol* **33**: 1870-1874.

Kurosu, M., Mahapatra, S., Narayanasamy, P., and 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.

Lai, L.Y., Lin, T.L., Chen, Y.Y., Hsieh, P.F., and Wang, J.T. (2018) Role of the *Mycobacterium marinum* ESX-1 secretion system in sliding motility and biofilm formation. *Front Microbiol* **9**: 1160.

Lavollay, M., Arthur, M., Fourgeaud, M., Dubost, L., Marie, A., Veziris, N., Blanot, D., Gutmann, L., and Mainardi, J.L. (2008) The peptidoglycan of stationary-phase *Mycobacterium tuberculosis* predominantly contains cross-links generated by L,D-transpeptidation. *J Bacteriol* **190**: 4360-4366.

Lazarevic, V., Margot, P., Soldo, B., and Karamata, D. (1992) Sequencing and analysis of the *Bacillus subtilis lytRABC* divergon: a regulatory unit encompassing the structural genes of the N-acetylmuramoyl-L-alanine amidase and its modifier. *J Gen Microbiol* **138**: 1949-1961.

Lederer, E., Adam, A., Ciorbaru, R., Petit, J.F., and Wietzerbin, J. (1975) Cell walls of mycobacteria and related organisms; chemistry and immunostimulant properties. *Mol Cell Biochem* **7**: 87-104.

Lee, R.E., Brennan, P.J., and Besra, G.S. (1997) Mycobacterial arabinan biosynthesis: the use of synthetic arabinoside acceptors in the development of an arabinosyl transfer assay. *Glycobiology* **7**: 1121-1128.

Lemassu, A., and Daffe, M. (1994) Structural features of the exocellular polysaccharides of *Mycobacterium tuberculosis*. *Biochem J* **297 (Pt 2)**: 351-357.

Lewin, A., and Sharbati-Tehrani, S. (2005) Slow growth rate of mycobacteria. Possible reasons and significance for their pathogenicity. *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz* **48**: 1390-1399.

Li, W., Xin, Y., McNeil, M.R., and Ma, Y. (2006) *rml*B and *rml*C genes are essential for growth of mycobacteria. *Biochem Biophys Res Commun* **342**: 170-178.

Ligozzi, M., Pittaluga, F., and Fontana, R. (1993) Identification of a genetic element (*psr*) which negatively controls expression of *Enterococcus hirae* penicillin-binding protein 5. *J Bacteriol* **175**: 2046-2051.

Liszewski Zilla, M., Chan, Y.G., Lunderberg, J.M., Schneewind, O., and Missiakas, D. (2015) LytR-CpsA-Psr enzymes as determinants of *Bacillus anthracis* secondary cell wall polysaccharide assembly. *J Bacteriol* **197**: 343-353.

Liu, B., Zhou, C., Li, G., Zhang, H., Zeng, E., Liu, Q., and Ma, Q. (2016) Bacterial regulon modeling and prediction based on systematic *cis* regulatory motif analyses. *Sci Rep* **6**: 23030.

Lu, J., and Holmgren, A. (2014) The thioredoxin antioxidant system. *Free Radic Biol Med* **66**: 75-87.

Ma, Y., Mills, J.A., Belisle, J.T., Vissa, V., Howell, M., Bowlin, K., Scherman, M.S., and 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)**: 937-945.

Ma, Y., Pan, F., and McNeil, M. (2002) Formation of dTDP-rhamnose is essential for growth of mycobacteria. *J Bacteriol* **184**: 3392-3395.

Ma, Y., Stern, R.J., Scherman, M.S., Vissa, V.D., Yan, W., Jones, V.C., Zhang, F., Franzblau, S.G., Lewis, W.H., and 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**: 1407-1416.

Maadani, A., Fox, K.A., Mylonakis, E., and Garsin, D.A. (2007) *Enterococcus faecalis* mutations affecting virulence in the *Caenorhabditis elegans* model host. *Infect Immun* **75**: 2634-2637.

Mahapatra, S., Scherman, H., Brennan, P.J., and Crick, D.C. (2005a) N Glycolylation of the nucleotide precursors of peptidoglycan biosynthesis of *Mycobacterium spp.* is altered by drug treatment. *J Bacteriol* **187**: 2341-2347.

Mahapatra, S., Yagi, T., Belisle, J.T., Espinosa, B.J., Hill, P.J., McNeil, M.R., Brennan, P.J., and Crick, D.C. (2005b) Mycobacterial lipid II is composed of a complex mixture of modified muramyl and peptide moieties linked to decaprenyl phosphate. *J Bacteriol* **187**: 2747-2757.

Malm, S., Maass, S., Schaible, U.E., Ehlers, S., and Niemann, S. (2018) *In vivo* virulence of *Mycobacterium tuberculosis* depends on a single homologue of the LytR-CpsA-Psr proteins. *Sci Rep* **8**: 3936.

Marques, M.B., Kasper, D.L., Pangburn, M.K., and Wessels, M.R. (1992) Prevention of C3 deposition by capsular polysaccharide is a virulence mechanism of type III group B streptococci. *Infect Immun* **60**: 3986-3993.

Martinez-Granero, F., Redondo-Nieto, M., Vesga, P., Martin, M., and Rivilla, R. (2014) AmrZ is a global transcriptional regulator implicated in iron uptake and environmental adaption in *P. fluorescens F113*. *BMC Genomics* **15**: 237.

Masjedi, M.R., Farnia, P., Sorooch, S., Pooramiri, M.V., Mansoori, S.D., Zarifi, A.Z., Akbarvelayati, A., and Hoffner, S. (2006) Extensively drug-resistant tuberculosis: 2 years of surveillance in Iran. *Clin Infect Dis* **43**: 841-847.

Mathew, R., Mukherjee, R., Balachandar, R., and Chatterji, D. (2006) Deletion of the *rpo*Z gene, encoding the omega subunit of RNA polymerase, results in pleiotropic surface-related phenotypes in *Mycobacterium smegmatis*. *Microbiology* **152**: 1741-1750.

Mattman, L.H. (1970) Cell wall-deficient forms of mycobacteria. *Ann N Y Acad Sci* **174**: 852-861.

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

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

McNeil, M., Wallner, S.J., Hunter, S.W., and 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., and Brennan, P.J. (1994) Enzymatic evidence for the presence of a critical terminal hexa-arabinoside in the cell walls of *Mycobacterium tuberculosis*. *Glycobiology* **4**: 165-173.

Mella-Herrera, R.A., Neunuebel, M.R., and Golden, J.W. (2011) *Anabaena sp. strain PCC 7120 con*R contains a LytR-CpsA-Psr domain, is developmentally regulated, and is essential for diazotrophic growth and heterocyst morphogenesis. *Microbiology* **157**: 617-626.

Meng, L., Tong, J., Wang, Q., Niu, C., and Gao, Q. (2017) Diverse effects of mycobacterial proline-proline-glutamic acid proteins upon interaction with host macrophages. *FEMS Microbiol Lett* **364**.

Mengin-Lecreulx, D., Texier, L., Rousseau, M., and van Heijenoort, J. (1991) The *mur*G gene of *Escherichia coli* codes for the UDP-*N*-acetylglucosamine: N-acetylmuramyl-(pentapeptide) pyrophosphoryl-undecaprenol N-acetylglucosamine transferase involved in the membrane steps of peptidoglycan synthesis. *J Bacteriol* **173**: 4625-4636.

Michailova, L., Stoitsova, S., Markova, N., Kussovski, V., Jordanova, M., and Dimova, I. (2000) Interaction of alveolar macrophages with *Staphylococcus aureus* and induction of microbial L-forms during infection in rats. *Int J Med Microbiol* **290**: 259-267.

Mikusova, K., Belanova, M., Kordulakova, J., Honda, K., McNeil, M.R., Mahapatra, S., Crick, D.C., and Brennan, P.J. (2006) Identification of a novel galactosyl transferase involved in biosynthesis of the mycobacterial cell wall. *J Bacteriol* **188**: 6592-6598.

Mikusova, K., Huang, H., Yagi, T., Holsters, M., Vereecke, D., D'Haeze, W., Scherman, M.S., Brennan, P.J., McNeil, M.R., and 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**: 8020-8025.

Mikusova, K., Mikus, M., Besra, G.S., Hancock, I., and Brennan, P.J. (1996) Biosynthesis of the linkage region of the mycobacterial cell wall. *J Biol Chem* **271**: 7820-7828.

Mikusova, K., Yagi, T., Stern, R., McNeil, M.R., Besra, G.S., Crick, D.C., and Brennan, P.J. (2000) Biosynthesis of the galactan component of the mycobacterial cell wall. *J Biol Chem* **275**: 33890-33897.

Mills, J.A., Motichka, K., Jucker, M., Wu, H.P., Uhlik, B.C., Stern, R.J., Scherman, M.S., Vissa, V.D., Pan, F., Kundu, M., Ma, Y.F., and McNeil, M. (2004) Inactivation of the mycobacterial rhamnosyltransferase, which is needed for the formation of the arabinogalactan-peptidoglycan linker, leads to irreversible loss of viability. *J Biol Chem* **279**: 43540-43546.

Ming, L.J., and Epperson, J.D. (2002) Metal binding and structure-activity relationship of the metalloantibiotic peptide bacitracin. *J Inorg Biochem* **91**: 46-58.

Minnikin, D.E., and Goodfellow, M. (1980) Lipid composition in the classification and identification of acid-fast bacteria. *Soc Appl Bacteriol Symp Ser* **8**: 189-256.

Mio, T., Yabe, T., Arisawa, M., and Yamada-Okabe, H. (1998) The eukaryotic UDP-*N*-acetylglucosamine pyrophosphorylases. Gene cloning, protein expression, and catalytic mechanism. *J Biol Chem* **273**: 14392-14397.

Miyamoto, Y., Mukai, T., Nakata, N., Maeda, Y., Kai, M., Naka, T., Yano, I., and Makino, M. (2006) Identification and characterization of the genes involved in glycosylation pathways of mycobacterial glycopeptidolipid biosynthesis. *J Bacteriol* **188**: 86-95.

Mohammadi, T., van Dam, V., Sijbrandi, R., Vernet, T., Zapun, A., Bouhss, A., Diepeveen-de Bruin, M., Nguyen-Disteche, M., de Kruijff, B., and Breukink, E. (2011) Identification of FtsW as a transporter of lipid-linked cell wall precursors across the membrane. *EMBO J* **30**: 1425-1432.

Mohanty, S., Jagannathan, L., Ganguli, G., Padhi, A., Roy, D., Alaridah, N., Saha, P., Nongthomba, U., Godaly, G., Gopal, R.K., Banerjee, S., and Sonawane, A. (2015) A mycobacterial phosphoribosyltransferase promotes bacillary survival by inhibiting oxidative stress and autophagy pathways in macrophages and zebrafish. *J Biol Chem* **290**: 13321-13343.

Movahedzadeh, F., and Bitter, W. (2009) Ins and outs of mycobacterial plasmids. *Methods Mol Biol* **465**: 217-228.

Munshi, T., Gupta, A., Evangelopoulos, D., Guzman, J.D., Gibbons, S., Keep, N.H., and Bhakta, S. (2013) Characterisation of ATP-dependent Mur ligases involved in the biogenesis of cell wall peptidoglycan in *Mycobacterium tuberculosis*. *PLoS One* **8**: e60143.

Nachega, J.B., and Chaisson, R.E. (2003) Tuberculosis drug resistance: a global threat. *Clin Infect Dis* **36**: S24-30.

Nath, H., and Ryoo, S., (2013) First– and Second–Line Drugs and Drug Resistance. *Tuberculosis - Current Issues in Diagnosis and Management*. Mahboub, B.H. and Vats, M.G. *IntechOpen.* Available at

https://www.intechopen.com/books/tuberculosis-current-issues-in-diagnosis-and-management/first-and-second-line-drugs-and-drug-resistance.

Nei, M., and Kumar, S. (2000) Molecular evolution and phylogenetics. *Oxford University Press, New York*.

Nelson, A.L., Roche, A.M., Gould, J.M., Chim, K., Ratner, A.J., and Weiser, J.N. (2007) Capsule enhances pneumococcal colonization by limiting mucus-mediated clearance. *Infect Immun* **75**: 83-90.

Nigou, J., Gilleron, M., and Puzo, G. (2003) Lipoarabinomannans: from structure to biosynthesis. *Biochimie* **85**: 153-166.

North, E.J., Jackson, M., and Lee, R.E. (2014) New approaches to target the mycolic acid biosynthesis pathway for the development of tuberculosis therapeutics. *Curr Pharm Des* **20**: 4357-4378.

Nunes-Alves, C., Booty, M.G., Carpenter, S.M., Jayaraman, P., Rothchild, A.C., and Behar, S.M. (2014) In search of a new paradigm for protective immunity to TB. *Nat Rev Microbiol* **12**: 289-299.

Ofer, N., Wishkautzan, M., Meijler, M., Wang, Y., Speer, A., Niederweis, M., and Gur, E. (2012) Ectoine biosynthesis in *Mycobacterium smegmatis*. *Appl Environ Microbiol* **78**: 7483-7486.

Ojha, A., Anand, M., Bhatt, A., Kremer, L., Jacobs, W.R., Jr., and Hatfull, G.F. (2005) GroEL1: a dedicated chaperone involved in mycolic acid biosynthesis during biofilm formation in mycobacteria. *Cell* **123**: 861-873.

Ojha, A.K., Baughn, A.D., Sambandan, D., Hsu, T., Trivelli, X., Guerardel, Y., Alahari, A., Kremer, L., Jacobs, W.R., Jr., and Hatfull, G.F. (2008) Growth of *Mycobacterium tuberculosis* biofilms containing free mycolic acids and harbouring drug-tolerant bacteria. *Mol Microbiol* **69**: 164-174.

Ortalo-Magne, A., Dupont, M.A., Lemassu, A., Andersen, A.B., Gounon, P., and Daffe, M. (1995) Molecular composition of the outermost capsular material of the tubercle bacillus. *Microbiology* **141 (Pt 7)**: 1609-1620.

Over, B., Heusser, R., McCallum, N., Schulthess, B., Kupferschmied, P., Gaiani, J.M., Sifri, C.D., Berger-Bachi, B., and Stutzmann Meier, P. (2011) LytR-CpsA-Psr proteins in *Staphylococcus aureus* display partial functional redundancy and the deletion of all three severely impairs septum placement and cell separation. *FEMS Microbiol Lett* **320**: 142-151.

Palomino, J.C., Martin, A., Camacho, M., Guerra, H., Swings, J., and Portaels, F. (2002) Resazurin microtiter assay plate: simple and inexpensive method for detection of drug resistance in *Mycobacterium tuberculosis*. *Antimicrob Agents Chemother* **46**: 2720-2722.

Parker, B.C., Ford, M.A., Gruft, H., and Falkinham, J.O., 3rd (1983) Epidemiology of infection by nontuberculous mycobacteria. IV. Preferential aerosolization of *Mycobacterium intracellulare* from natural waters. *Am Rev Respir Dis* **128**: 652-656.

Pavelka, M.S., Jr., Mahapatra, S., and Crick, D.C. (2014) Genetics of peptidoglycan biosynthesis. *Microbiol Spectr* **2**: MGM2-0034-2013.

Pelicic, V., Reyrat, J.M., and Gicquel, B. (1996) Expression of the *Bacillus subtilis sac*B gene confers sucrose sensitivity on mycobacteria. *J Bacteriol* **178**: 1197-1199.

Pethe, K., Swenson, D.L., Alonso, S., Anderson, J., Wang, C., and Russell, D.G. (2004) Isolation of *Mycobacterium tuberculosis* mutants defective in the arrest of phagosome maturation. *Proc Natl Acad Sci U S A* **101**: 13642-13647.

Petit, J.F., Adam, A., Wietzerbin-Falszpan, J., Lederer, E., and Ghuysen, J.M. (1969) Chemical structure of the cell wall of *Mycobacterium smegmatis*. I. Isolation and partial characterization of the peptidoglycan. *Biochem Biophys Res Commun* **35**: 478-485.

Piddington, D.L., Fang, F.C., Laessig, T., Cooper, A.M., Orme, I.M., and Buchmeier, N.A. (2001) Cu,Zn superoxide dismutase of *Mycobacterium tuberculosis* contributes to survival in activated macrophages that are generating an oxidative burst. *Infect Immun* **69**: 4980-4987.

Pierre-Audigier, C., Jouanguy, E., Lamhamedi, S., Altare, F., Rauzier, J., Vincent, V., Canioni, D., Emile, J.F., Fischer, A., Blanche, S., Gaillard, J.L., and Casanova, J.L. (1997) Fatal disseminated *Mycobacterium*

smegmatis infection in a child with inherited interferon gamma receptor deficiency. *Clin Infect Dis* **24**: 982-984.

Pitarque, S., Larrouy-Maumus, G., Payre, B., Jackson, M., Puzo, G., and Nigou, J. (2008) The immunomodulatory lipoglycans, lipoarabinomannan and lipomannan, are exposed at the mycobacterial cell surface. *Tuberculosis (Edinb)* **88**: 560-565.

Prasanna, A.N., and Mehra, S. (2013) Comparative phylogenomics of pathogenic and non-pathogenic *Mycobacterium*. *PLoS One* **8**: e71248.

Provvedi, R., Kocincova, D., Dona, V., Euphrasie, D., Daffe, M., Etienne, G., Manganelli, R., and Reyrat, J.M. (2008) SigF controls carotenoid pigment production and affects transformation efficiency and hydrogen peroxide sensitivity in *Mycobacterium smegmatis*. *J Bacteriol* **190**: 7859-7863.

Puffal, J., Garcia-Heredia, A., Rahlwes, K.C., Siegrist, M.S., and Morita, Y.S. (2018) Spatial control of cell envelope biosynthesis in mycobacteria. *Pathog Dis* **76**.

Qi, Z.D., Lin, Y., Zhou, B., Ren, X.D., Pang, D.W., and Liu, Y. (2008) Characterization of the mechanism of the *Staphylococcus aureus* cell envelope by bacitracin and bacitracin-metal ions. *J Membr Biol* **225**: 27-37.

Radmacher, E., Stansen, K.C., Besra, G.S., Alderwick, L.J., Maughan, W.N., Hollweg, G., Sahm, H., Wendisch, V.F., and Eggeling, L. (2005) Ethambutol, a cell wall inhibitor of *Mycobacterium tuberculosis*, elicits L-glutamate efflux of *Corynebacterium glutamicum*. *Microbiology* **151**: 1359-1368.

Rana, A.K., Singh, A., Gurcha, S.S., Cox, L.R., Bhatt, A., and Besra, G.S. (2012) Ppm1-encoded polyprenyl monophosphomannose synthase activity is essential for lipoglycan synthesis and survival in mycobacteria. *PLoS One* **7**: e48211.

Raymond, J.B., Mahapatra, S., Crick, D.C., and Pavelka, M.S., Jr. (2005) Identification of the *nam*H gene, encoding the hydroxylase responsible for the N-glycolylation of the mycobacterial peptidoglycan. *J Biol Chem* **280**: 326-333.

Recht, J., and Kolter, R. (2001) Glycopeptidolipid acetylation affects sliding motility and biofilm formation in *Mycobacterium smegmatis*. *J Bacteriol* **183**: 5718-5724.

Recht, J., Martinez, A., Torello, S., and Kolter, R. (2000) Genetic analysis of sliding motility in *Mycobacterium smegmatis*. *J Bacteriol* **182**: 4348-4351.

Reiley, W.W., Calayag, M.D., Wittmer, S.T., Huntington, J.L., Pearl, J.E., Fountain, J.J., Martino, C.A., Roberts, A.D., Cooper, A.M., Winslow, G.M., and Woodland, D.L. (2008) ESAT-6-specific CD4 T cell responses to aerosol *Mycobacterium tuberculosis* infection are initiated in the mediastinal lymph nodes. *Proc Natl Acad Sci U S A* **105**: 10961-10966.

Riley, R.L., Mills, C.C., Nyka, W., Weinstock, N., Storey, P.B., Sultan, L.U., Riley, M.C., and Wells, W.F. (1959) Aerial dissemination of pulmonary tuberculosis. A two-year study of contagion in a tuberculosis ward. 1959. *Am J Hyg* **142**: 3-14.

Rizvi, A., Yousf, S., Balakrishnan, K., Dubey, H.K., Mande, S.C., Chugh, J., and Banerjee, S. (2019) Metabolomics studies to decipher stress responses in *Mycobacterium smegmatis* point to a putative pathway of methylated amine biosynthesis. *J Bacteriol* **201**.

Rose, N.L., Completo, G.C., Lin, S.J., McNeil, M., Palcic, M.M., and Lowary, T.L. (2006) Expression, purification, and characterization of a galactofuranosyltransferase involved in *Mycobacterium tuberculosis* arabinogalactan biosynthesis. *J Am Chem Soc* **128**: 6721-6729.

Rossi, J., Bischoff, M., Wada, A., and Berger-Bachi, B. (2003) MsrR, a putative cell envelope-associated element involved in *Staphylococcus aureus sar*A attenuation. *Antimicrob Agents Chemother* **47**: 2558-2564.

Rosu, V., Bandino, E., and Cossu, A. (2013) Unraveling the transcriptional regulatory networks associated to mycobacterial cell wall defective form induction by glycine and lysozyme treatment. *Microbiol Res* **168**: 153-164.

Roxas, B.A., and Li, Q. (2009) Acid stress response of a mycobacterial proteome: insight from a gene ontology analysis. *Int J Clin Exp Med* **2**: 309-328.

Ruiz, N. (2015) Lipid flippases for bacterial peptidoglycan biosynthesis. *Lipid Insights* **8**: 21-31.

Saecker, R.M., Record, M.T., Jr., and Dehaseth, P.L. (2011) Mechanism of bacterial transcription initiation: RNA polymerase - promoter binding, isomerization to initiation-competent open complexes, and initiation of RNA synthesis. *J Mol Biol* **412**: 754-771.

Saitou, N., and Nei, M. (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* **4**: 406-425.

Samanovic, M.I., Tu, S., Novak, O., Iyer, L.M., McAllister, F.E., Aravind, L., Gygi, S.P., Hubbard, S.R., Strnad, M., and Darwin, K.H. (2015) Proteasomal control of cytokinin synthesis protects *Mycobacterium tuberculosis* against nitric oxide. *Mol Cell* **57**: 984-994.

Sambou, T., Dinadayala, P., Stadthagen, G., Barilone, N., Bordat, Y., Constant, P., Levillain, F., Neyrolles, O., Gicquel, B., Lemassu, A., Daffe, M., and Jackson, M. (2008) Capsular glucan and intracellular glycogen of *Mycobacterium tuberculosis*: biosynthesis and impact on the persistence in mice. *Mol Microbiol* **70**: 762-774.

Sangiliyandi, G., Kannan, T.R., Chandra Raj, K., and Gunasekaran, P. (1999) Separation of levan-formation and sucrose-hydrolysis catalyzed by levansucrase of *Zymomonas mobilis* using *in vitro* mutagenesis. *Brazilian Archives of Biology and Technology* **42**: 375-379.

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., and 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.

Santos, R., de Carvalho, C.C., Stevenson, A., Grant, I.R., and Hallsworth, J.E. (2015) Extraordinary solute-stress tolerance contributes to the environmental tenacity of mycobacteria. *Environ Microbiol Rep* **7**: 746-764.

Sapunaric, F., Franssen, C., Stefanic, P., Amoroso, A., Dardenne, O., and Coyette, J. (2003) Redefining the role of *psr* in beta-lactam resistance and cell autolysis of *Enterococcus hirae*. *J Bacteriol* **185**: 5925-5935.

Saucedo, N.M., Gao, Y., Pham, T., and Mulchandani, A. (2018) Lectin- and saccharide-functionalized nano-chemiresistor arrays for detection and identification of pathogenic bacteria infection. *Biosensors (Basel)* **8**.

Sauvage, E., Kerff, F., Terrak, M., Ayala, J.A., and Charlier, P. (2008) The penicillin-binding proteins: structure and role in peptidoglycan biosynthesis. *FEMS Microbiol Rev* **32**: 234-258.

Schaefer, K., Matano, L.M., Qiao, Y., Kahne, D., and Walker, S. (2017) *In vitro* reconstitution demonstrates the cell wall ligase activity of LCP proteins. *Nat Chem Biol* **13**: 396-401.

Schaefer, K., Owens, T.W., Kahne, D., and Walker, S. (2018) Substrate preferences establish the order of cell wall assembly in *Staphylococcus aureus*. *J Am Chem Soc* **140**: 2442-2445.

Schorey, J.S., and Sweet, L. (2008) The mycobacterial glycopeptidolipids: structure, function, and their role in pathogenesis. *Glycobiology* **18**: 832-841.

Seidel, M., Alderwick, L.J., Birch, H.L., Sahm, H., Eggeling, L., and 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**: 14729-14740.

Senthilkumar, V., Busby, S.J., and Gunasekaran, P. (2003) Serine substitution for cysteine residues in levansucrase selectively abolishes levan forming activity. *Biotechnol Lett* **25**: 1653-1656.

Sham, L.T., Butler, E.K., Lebar, M.D., Kahne, D., Bernhardt, T.G., and Ruiz, N. (2014) Bacterial cell wall. MurJ is the flippase of lipid-linked precursors for peptidoglycan biogenesis. *Science* **345**: 220-222.

Shi, T., Fu, T., and Xie, J. (2011) Polyphosphate deficiency affects the sliding motility and biofilm formation of *Mycobacterium smegmatis*. *Curr Microbiol* **63**: 470-476.

Si, M., Xu, Y., Wang, T., Long, M., Ding, W., Chen, C., Guan, X., Liu, Y., Wang, Y., Shen, X., and Liu, S.J. (2015) Functional characterization of a mycothiol peroxidase in *Corynebacterium glutamicum* that uses both mycoredoxin and thioredoxin reducing systems in the response to oxidative stress. *Biochem J* **469**: 45-57.

Singh, A.K., and Singh, B.N. (2009) Differential expression of *sig*H paralogs during growth and under different stress conditions in *Mycobacterium smegmatis*. *J Bacteriol* **191**: 2888-2893.

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., and Jackson, M. (2010) Biosynthetic origin of the galactosamine substituent of arabinogalactan in *Mycobacterium tuberculosis*. *J Biol Chem* **285**: 41348-41355.

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., and Jackson, M. (2009) AftD, a novel essential arabinofuranosyltransferase from mycobacteria. *Glycobiology* **19**: 1235-1247.

Smith, I. (2003) *Mycobacterium tuberculosis* pathogenesis and molecular determinants of virulence. *Clin Microbiol Rev* **16**: 463-496.

Smollett, K.L., Smith, K.M., Kahramanoglou, C., Arnvig, K.B., Buxton, R.S., and Davis, E.O. (2012) Global analysis of the regulon of the transcriptional repressor LexA, a key component of SOS response in *Mycobacterium tuberculosis*. *J Biol Chem* **287**: 22004-22014.

Snapper, S.B., Melton, R.E., Mustafa, S., Kieser, T., and Jr, W.R.J. (1990) Isolation and characterization of efficient plasmid transformation mutants of *Mycobacterium smegmatis*. *Mol Microbiol* **4**: 1911-1919.

Sousa, S., Bandeira, M., Carvalho, P.A., Duarte, A., and Jordao, L. (2015) Nontuberculous mycobacteria pathogenesis and biofilm assembly. *Int J Mycobacteriol* **4**: 36-43.

Stanley, S.A., and Cox, J.S. (2013) Host-pathogen interactions during *Mycobacterium tuberculosis* infections. *Curr Top Microbiol Immunol* **374**: 211-241.

Stern, R.J., Lee, T.Y., Lee, T.J., Yan, W., Scherman, M.S., Vissa, V.D., Kim, S.K., Wanner, B.L., and McNeil, M.R. (1999) Conversion of dTDP-4-keto-6-deoxyglucose to free dTDP-4-keto-rhamnose by the *rml*C gene products of *Escherichia coli* and *Mycobacterium tuberculosis*. *Microbiology* **145 (Pt 3)**: 663-671.

Stone, K.J., and Strominger, J.L. (1971) Mechanism of action of bacitracin: complexation with metal ion and C 55 -isoprenyl pyrophosphate. *Proc Natl Acad Sci U S A* **68**: 3223-3227.

Takayama, K., and Goldman, D.S. (1970) Enzymatic synthesis of mannosyl-1-phosphoryl-decaprenol by a cell-free system of *Mycobacterium tuberculosis*. *J Biol Chem* **245**: 6251-6257.

Tefsen, B., Ram, A.F., van Die, I., and Routier, F.H. (2012) Galactofuranose in eukaryotes: aspects of biosynthesis and functional impact. *Glycobiology* **22**: 456-469.

Telenti, A., Philipp, W.J., Sreevatsan, S., Bernasconi, C., Stockbauer, K.E., Wieles, B., Musser, J.M., and Jacobs, W.R., Jr. (1997) The *emb* operon, a gene cluster of *Mycobacterium tuberculosis* involved in resistance to ethambutol. *Nat Med* **3**: 567-570.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., and Higgins, D.G. (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* **25**: 4876-4882.

Thompson, J.D., Higgins, D.G., and Gibson, T.J. (1994) Clustal-W - Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22**: 4673-4680.

Tipper, D.J., and Strominger, J.L. (1965) Mechanism of action of penicillins: a proposal based on their structural similarity to acyl-D-alanyl-D-alanine. *Proc Natl Acad Sci U S A* **54**: 1133-1141.

Torrelles, J.B., and Schlesinger, L.S. (2010) Diversity in *Mycobacterium tuberculosis* mannosylated cell wall determinants impacts adaptation to the host. *Tuberculosis (Edinb)* **90**: 84-93.

Valvano, M.A. (2008) Undecaprenyl phosphate recycling comes out of age. *Mol Microbiol* **67**: 232-235.

van de Weerd, R., Boot, M., Maaskant, J., Sparrius, M., Verboom, T., van Leeuwen, L.M., Burggraaf, M.J., Paauw, N.J., Dainese, E., Manganelli, R., Bitter, W., Appelmelk, B.J., and Geurtsen, J. (2016) Inorganic phosphate limitation modulates capsular polysaccharide composition in mycobacteria. *J Biol Chem* **291**: 11787-11799.

van Heijenoort, J. (2007) Lipid intermediates in the biosynthesis of bacterial peptidoglycan. *Microbiol Mol Biol Rev* **71**: 620-635.

Vergne, I., Gilleron, M., and Nigou, J. (2014) Manipulation of the endocytic pathway and phagocyte functions by *Mycobacterium tuberculosis* lipoarabinomannan. *Front Cell Infect Microbiol* **4**: 187.

Vergne, I., Prats, M., Tocanne, J.F., and Laneelle, G. (1995) Mycobacterial glycopeptidolipid interactions with membranes: a monolayer study. *FEBS Lett* **375**: 254-258.

Vidhya, R., Rathnakumar, K., Balu, V., and Pugalendi, K.V. (2019) Oxidative stress, antioxidant status and lipid profile in pulmonary tuberculosis patients before and after anti-tubercular therapy. *Indian J Tuberc* **66**: 375-381.

Vollmer, W., Blanot, D., and de Pedro, M.A. (2008) Peptidoglycan structure and architecture. *FEMS Microbiol Rev* **32**: 149-167.

Waagmeester, A., Thompson, J., and Reyrat, J.M. (2005) Identifying sigma factors in *Mycobacterium smegmatis* by comparative genomic analysis. *Trends Microbiol* **13**: 505-509.

Walsh, C. (2000) Molecular mechanisms that confer antibacterial drug resistance. *Nature* **406**: 775-781.

Wang, Q., Zhu, L., Jones, V., Wang, C., Hua, Y., Shi, X., Feng, X., Jackson, M., Niu, C., and Gao, Q. (2015) CpsA, a LytR-CpsA-Psr family protein in *Mycobacterium marinum*, is required for cell wall integrity and virulence. *Infect Immun* **83**: 2844-2854.

Warner, D.F., and Mizrahi, V. (2013) Complex genetics of drug resistance in *Mycobacterium tuberculosis*. *Nat Genet* **45**: 1107-1108.

Watanabe, A., Kaneko, C., Hamada, Y., Takeda, K., Kimata, S., Matsumoto, T., Abe, A., Tanaka, N., Okada, S., Uchino, M., Satoh, J., Nakagawa, J., and Niimura, Y. (2016) Isolation of lactic acid bacteria exhibiting high scavenging activity for environmental hydrogen peroxide from fermented foods and its two scavenging enzymes for hydrogen peroxide. *J Gen Appl Microbiol* **62**: 75-82.

West, N.P., Cergol, K.M., Xue, M., Randall, E.J., Britton, W.J., and Payne, R.J. (2011) Inhibitors of an essential mycobacterial cell wall lipase (Rv3802c) as tuberculosis drug leads. *Chem Commun (Camb)* **47**: 5166-5168.

WHO (2018) Global Tuberculosis Report 2018. *World Health Organization*.

Wolf, A.J., Desvignes, L., Linas, B., Banaiee, N., Tamura, T., Takatsu, K., and Ernst, J.D. (2008) Initiation of the adaptive immune response to *Mycobacterium tuberculosis* depends on antigen production in the local lymph node, not the lungs. *J Exp Med* **205**: 105-115.

Wolucka, B.A., and de Hoffmann, E. (1995) The presence of beta-D-ribosyl-1-monophosphodecaprenol in mycobacteria. *J Biol Chem* **270**: 20151-20155.

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

Wu, C., Huang, I.H., Chang, C., Reardon-Robinson, M.E., Das, A., and Ton-That, H. (2014) Lethality of sortase depletion in *Actinomyces oris* caused by excessive membrane accumulation of a surface glycoprotein. *Mol Microbiol* **94**: 1227-1241.

Xin, Y., Lee, R.E., Scherman, M.S., Khoo, K.H., Besra, G.S., Brennan, P.J., and McNeil, M. (1997) Characterization of the *in vitro* synthesized arabinan of mycobacterial cell walls. *Biochim Biophys Acta* **1335**: 231-234.

Yagi, T., Mahapatra, S., Mikusova, K., Crick, D.C., and Brennan, P.J. (2003) Polymerization of mycobacterial arabinogalactan and ligation to peptidoglycan. *J Biol Chem* **278**: 26497-26504.

Yamada, H., Yamaguchi, M., Igarashi, Y., Chikamatsu, K., Aono, A., Murase, Y., Morishige, Y., Takaki, A., Chibana, H., and Mitarai, S. (2018) *Mycolicibacterium smegmatis*, basonym *Mycobacterium smegmatis*, expresses morphological phenotypes much more similar to *Escherichia coli* than *Mycobacterium tuberculosis* in quantitative structome analysis and cryoTEM examination. *Front Microbiol* **9**: 1992.

Yang, Y., Richards, J.P., Gundrum, J., and Ojha, A.K. (2018) GlnR activation induces peroxide resistance in mycobacterial biofilms. *Front Microbiol* **9**: 1428.

Yang, Y., Thomas, J., Li, Y., Vilcheze, C., Derbyshire, K.M., Jacobs, W.R., Jr., and Ojha, A.K. (2017) Defining a temporal order of genetic requirements for development of mycobacterial biofilms. *Mol Microbiol* **105**: 794-809.

Yokoyama, K., Miyashita, T., Araki, Y., and 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**: 479-489.

Zanfardino, A., Migliardi, A., D'Alonzo, D., Lombardi, A., Varcamonti, M., and Cordone, A. (2016) Inactivation of *MSMEG_0412* gene drastically affects surface related properties of *Mycobacterium smegmatis*. *BMC Microbiol* **16**: 267.

Zhang, N., Torrelles, J.B., McNeil, M.R., Escuyer, V.E., Khoo, K.H., Brennan, P.J., and Chatterjee, D. (2003) The Emb proteins of mycobacteria direct arabinosylation of lipoarabinomannan and arabinogalactan via an N-terminal recognition region and a C-terminal synthetic region. *Mol Microbiol* **50**: 69-76.

Zhang, S., Wu, Q., Lei, H., Zheng, H., Zhou, F., Sun, Z., Zhao, J., Yu, X., and Zhang, S. (2019) Mannosylated structures of mycobacterial lipoarabinomannans facilitate the maturation and activation of dendritic cells. *Cell Immunol* **335**: 85-92.

Zhang, Z., Bulloch, E.M., Bunker, R.D., Baker, E.N., and Squire, C.J. (2009) Structure and function of GlmU from *Mycobacterium tuberculosis*. *Acta Crystallogr D Biol Crystallogr* **65**: 275-283.

APPENDIX: Enlarged images of pellicle formation (biofilm on air-liquid interface) by the *Msmeg* wild type strain and *lcp* deletion mutants
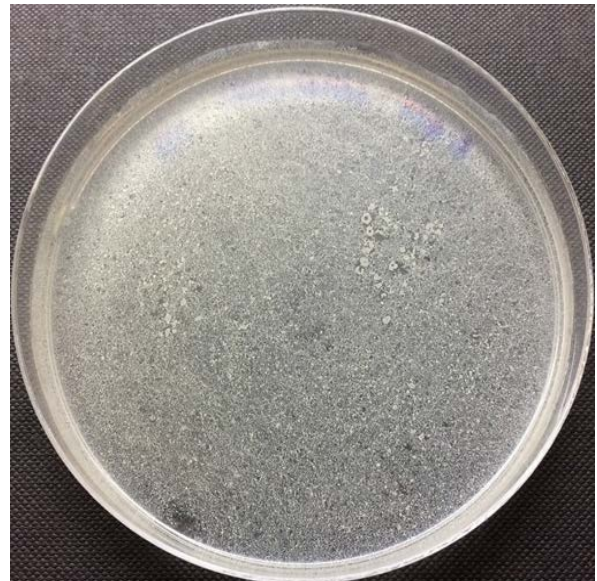

Δ5775

Δ6421

ΔΔ(0107+5775)

ΔΔ(0107+6421)

ΔΔ(5775+6421)

The biofilm observed on the surface of 7H9 medium supplemented with ADS but without Tween-80, is after incubation of the plates at 37°C for 4 days. The data is a representative of one of the three independent experiments.