

https://dspace.org/

# Characterization of resuscitation promoting factors in Mycobacterium smegmatis

**Mapela, Lusanda Thato**

**2012-09-12**
http://hdl.handle.net/10539/11937

Downloaded from a DSpace repository. https://dspace.org/

# Characterization of Resuscitation Promoting Factors in *Mycobacterium smegmatis*

Lusanda Thato Mapela




A Dissertation submitted to the Faculty of Science, University of the Witwatersrand, Johannesburg, in fulfilment of the requirements for the degree of Master of Science.

February 2012.

“We are hard pressed on every side, but not crushed; perplexed, but not in despair; persecuted, but not abandoned; struck down, but not destroyed.”

2 Corinthians 4: 8-9

**Declaration**

I declare that this Dissertation is my own, unaided work. It is being submitted for the Degree of Master of Science at the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination at any other University.



Lusanda Thato Mapela

15 th day of May 2012

## Dedication

I dedicate this work to my late grandmother **Nompi Anna “Mme” Tshabalala**.

**Abstract**

*Mycobacterium tuberculosis*, the causative agent of tuberculosis (TB) has infected one third of the world's population and continues to claim more lives annually than any other infectious disease agent. A significant proportion of individuals carry latent TB infection (LBTI) which is characterized by the absence of clinical symptoms and it has been postulated that the tubercle bacilli are in a dormant-like state during this type of infection. Resuscitation promoting factors (Rpfs) are cell wall hydrolases which cleave glycosidic bonds within the peptidoglycan (PG), a mechanism thought to result in reactivation of bacteria from the state of dormancy. *M. tuberculosis* encodes five *rpf*-like homologues which are collectively dispensable for growth but are required for reactivation from dormancy *in vitro*, and for virulence in the mouse model of infection. LTBI thus poses a huge threat to the global burden of active disease. The purpose of this study was to further investigate the biological roles of Rpfs by assessing the effects of *rpf* gene deletion in *M. smegmatis*, a model organism used for TB research. *M. smegmatis* encodes four *rpf*-like genes designated *rpfA*, *rpfB*, *rpfC* and *rpfE*, and deletion mutants that lack one or more of these genes were constructed by allelic exchange mutagenesis. *M. smegmatis* mutant strains that lack either *rpfA* or *rpfB* display no significant differences in growth both on solid and in liquid medium when compared to wild type. However, loss of *rpfA* resulted in bacterial clumping during stationary-phase growth in broth culture and changes in cell morphology. Moreover; the Δ*rpfA* Δ*rpfB* double mutant and its derivative strain lacking *rpfC*, Δ*rpfA* Δ*rpfB* Δ*rpfC* displayed a ca. 2-4 log increase in susceptibility to erythromycin and vancomycin. Furthermore, unusual colonial morphologies with reduced serpentine cording and smooth peripheries were observed for these multiple mutants. Cell surface defects and cell distortions were evidenced as wrinkle-textured, bent cells and polar tip bulges for the abovementioned mutants. The multiple deletion strains also displayed a defect in biofilm formation revealing an inability for the mutants to form complex cell-cell interactions. Collectively, the data are suggestive of a loss of bacterial cell wall

integrity due to *rpf* gene deletion and it is proposed that *rpfA*, in combination with other genes, is largely responsible for cell wall integrity maintenance. Our data indicate that these factors play an important role in cell growth and division and therefore represent an untapped source of novel targets for anti-tubercular drug discovery.

**Acknowledgements**

I would like to acknowledge the **Almighty** in all my ways, for strength and favour; nothing would be possible without Him.

Thank you to my family at large; for encouragement, motivation and laughter! Mostly, I would love to thank my parents, **Sipho Tshabalala** and **Dieketseng Malepe** for the support and unconditional love. Through it all, you both were by my side in good and bad times. To my loving grandmother, **Mary Mapela** thank you for the support and confidence (in me). To my siblings; **Kgomotso**, **Thabang**, **Khosi**, **Lungile** and **Lebogang**, thank you for providing comic relief in my life- I appreciate your presence.

I am grateful for funding received from **NRF** (National Research Foundation), DST-NRF Centre of Excellence for Biomedical Research and the University of the Witwatersrand.

Thank you to the **MMU** (Microscopy and Microanalysis Unit) at Witwatersrand University for affording me the opportunity to work with their scanning electron microscope.

To my supervisor, **Dr Bavesh Kana**, thank you for the guidance and supervision throughout the project; I have learnt a lot.

To Dr. **Edith Machowski**, thank you for generating the necessary vectors for completion of the study.

I would like to extend my gratitude to my fellow colleagues and senior staff, at the MMRU/CBTBR, for assistance and support- this simplified many challenges.

I would like to thank all my good friends; including **Tshepo Nkosi** for his generosity and kindness and **Bongani Malaza** for support.

**Table of Contents**


Declaration ..... iii
Dedication ..... iv
Abstract ..... v
Acknowledgements ..... vi
Table of Contents ..... viii
List of Figures ..... x
List of Tables ..... xii
Nomenclature ..... xiii
1. Introduction ..... 1
1.1 Tuberculosis chemotherapy ..... 1
1.2 Drug resistance, HIV coinfection and vaccination ..... 3
1.3 *M. tuberculosis* pathogenesis and latent TB infection ..... 5
1.4 Resuscitation promoting factors ..... 8
1.4.1 *Micrococcus luteus* Rpf ..... 9
1.4.2 Mycobacterial *rpf*-like genes ..... 10
1.4.3 Rpf domain ..... 12
1.4.4. Muralytic activity ..... 12
1.4.5 The peptidoglycan ..... 21
1.4.6 Biological roles of Rpfs in *M. tuberculosis* ..... 25
1.5 Aims of this study ..... 30
2. Materials and Methods ..... 31
2.1 Bacterial strains and growth conditions ..... 31
2.2 General laboratory techniques ..... 31
2.2.1 DNA isolation ..... 31
2.2.2 DNA manipulation ..... 36
2.2.3 DNA precipitation ..... 37
2.2.4 Gel electrophoresis ..... 38
2.2.5 Gel extraction and purification ..... 38
2.2.6 DNA quantification ..... 39
2.2.7 Preparation of *E. coli* DH5α competent cells using rubidium chloride ..... 39
2.2.8 PCR screening of *E. coli* transformants ..... 40
2.2.9 Serial dilutions ..... 40
2.2.10 DNA sequencing ..... 41
2.3 Bioinformatics analysis of *rpf* genes in *M. smegmatis* ..... 41
2.4 Construction of *rpf*-knockout mutants ..... 41
2.4.1 Polymerase Chain Reaction (PCR) ..... 42
2.4.2 Construction of knockout vectors for *rpfA* and *rpfB* ..... 43
2.4.3 *M. smegmatis* $mc^2155$ competent cells preparation ..... 46
2.4.4 Screening for *M. smegmatis rpf* deletion mutants by PCR ..... 47
2.4.5 Southern blot analysis ..... 48
2.5 Growth kinetics ..... 51
2.6 Drug susceptibility assay ..... 52
2.7 Analysis of changes in colony morphology ..... 52
2.8 Anaylsis of biofilm formation ..... 52

2.9 Analysis of cell morphology by scanning electron microscopy ..... 53
3. Results ..... 54
3.1 Bioinformatics analysis of *rpf*-like genes in *M. smegmatis* ..... 54
3.1.1 *M. smegmatis* Rpf-like proteins are secreted signal peptides ..... 58
3.2 Construction of mutants defective for one or more *rpf* genes ..... 61
3.2.1 Suicide delivery vector construction ..... 63
3.2.2 Construction of single deletion mutants of *M. smegmatis* defective for *rpfA* and *rpfB* ..... 85
3.2.3 Construction of double and triple mutants of *M. smegmatis* defective for multiple *rpf*-like genes. ..... 89
3.3 Southern blot analysis of mutant strains to further confirm genotype ..... 93
3.4 Construction of complemented derivatives ..... 96
3.5 Phenotypic characterization of the *rpf* single and multiple mutants ..... 97
3.5.1 Growth kinetics of the single and multiple *rpf* mutants ..... 97
3.5.2 Bacterial clumping in liquid culture associated with *rpfA* gene deletion ..... 99
3.5.3 The multiple *rpf* deletion mutants display an increase in susceptibility to erythromycin and vancomycin ..... 101
3.5.4 Deletion of multiple *rpf* genes results in alterations in colony morphology. ..... 106
3.5.5 Biofilm formation defects associated with multiple *rpf* gene deletions ..... 108
3.5.6 Cell morphological deformities resulting from multiple *rpf* deletion ..... 110
4. Discussion ..... 113
4.1 Concluding remarks ..... 121
5. Appendices ..... 123
5.1 Appendix A: Culture media ..... 123
5.2 Appendix B: Solutions ..... 124
5.3 Appendix C: Plasmids and molecular weight markers ..... 125
6. References ..... 127

## List of Figures

Figure 1: A schematic representation of the *M. tuberculosis* Rpf-like proteins showing the presence and positioning of the common Rpf domain ..... 11
Figure 2: Crystal structure of $_{\Delta DUF}$RpfB consisting of two distinct domains: catalytic and G5 domains..... 15
Figure 3: The crystal structure of RipA catalytic domain showing the residues involved in catalysis particularly; cysteine383, histidine432 and glutamine444... 18
Figure 4: The peptidoglycan layer within the mycobacterial cell wall ..... 22
Figure 5: Peptidoglycan biosynthesis is initiated with the production of peptidoglycan precursors in the cytoplasm and transported to the periplasm by a possible “flippase”; where transglycosylation (TG) and transpeptidation (TP) reactions occur forming the peptidoglycan sacculus..... 23
Figure 6: Cleavage sites within the peptidoglycan resulting in degradation of the murein by hydrolases. Ami, N-acetylmuramyl-L-alanine amidase; LD/DD- EPase, LD/DD-endopeptidase; LD/DD-CPase, LD/DD-carboxypeptidase ..... 24
Figure 7: A schematic representation of the genomic organization of *rpf*-like genes in *M. smegmatis* (*M. smeg*) compared to *rpf*-like genes of *M. tuberculosis* (TB)..... 54
Figure 8: Comparative amino acid sequence alignment of the Rpf proteins in *M. smegmatis* and *M. tuberculosis* ..... 57
Figure 9: Signal peptide prediction and localization of signal peptide cleavage sites within the *M. smegmatis* Rpf-like protein sequences..... 60
Figure 10: A schematic representation of the *rpf* regions amplified during PCR with specifically designed primers, and the resulting PCR fragments with expected sizes..... 63
Figure 11: PCR fragments amplified for *rpfA.* (A+B) *Thermus aquaticus* (Taq) PCR ..... 64
Figure 12: PCR fragments amplified for *rpfB* homologous regions..... 66
Figure 13: pBluescript carrier plasmid vector linearized with *Sma*I enzyme. ..... 68
Figure 14: *rpfA* and *rpfB* constructs digested with appropriate enzymes releasing the cloned insert. ..... 70
Figure 15: Schematic showing three-way cloning resulting in simultaneous ligation of upstream and downstream homologous regions to p2NIL knockout vector by directional cloning. ..... 71
Figure 16: p2NIL plasmid vector digested with appropriate enzymes for cloning purposes.. ..... 72
Figure 17: Upstream and downstream homologous *rpf* fragments digested from replicating vectors. ..... 73
Figure 18: Restriction mapping of the pNΔRPFA construct shown alongside its map..... 74
Figure 19: pNΔRPFB restriction mapping ..... 76
Figure 20: *Pac*I linearized pNΔRPFA and pNΔRPFB constructs. ..... 77
Figure 21: *Pac*I cassette isolation from pGOAL19 plasmid vector..... 78

Figure 22: pSV∆RPFA restricted with various enzymes for confirmation of suicide vector integrity. Suicide delivery vector map shown alongside restriction mapping ..... 79
Figure 23: Restriction mapping of pSV∆RPFB for plasmid integrity confirmation. ..... 81
Figure 24: pGEM plasmid vector restricted with appropriate enzymes in preparation for three-way cloning ..... 82
Figure 25: Restriction mapping of pG∆RPFA construct shown alongside the knockout construct map. ..... 83
Figure 26: pG∆RPFB restriction mapping shown with the construct map ..... 84
Figure 27: Construction of knockout mutants by two step allelic exchange mutagenesis. ..... 86
Figure 28: Illustration of the screening process and identification of ∆*rpfA* mutants using *rpfA* allelic-specific screening primers following allelic exchange mutagenesis ..... 87
Figure 29: The identification of ∆*rpfB* mutants resulting from the demonstrated screening strategy using *rpfB* allelic-specific screening primers, following allelic exchange mutagenesis ..... 89
Figure 30: Identification of ∆*rpfA* ∆*rpfB* double mutants with the use of *rpfB* allelic-specific screening primers in a ∆*rpfA* genetic background, following allelic exchange mutagenesis.. ..... 90
Figure 31: The identification of *rpfC* and *rpfE* deficient genomic DNA in the ∆*rpfA* ∆*rpfB* double mutant genetic background using *rpfC* and *rpfE* allelic-specific primers following allelic exchange mutagenesis ..... 92
Figure 32: Genotyping of the upstream and/or downstream regions of the *rpf* mutants by Southern blot analysis ..... 96
Figure 33: Growth characteristics of wild type in comparison to the *rpf* mutants and a complemented derivative ..... 98
Figure 34: Culture flasks showing bacterial clumping in 33-hour stationery phase cultures of single ∆*rpfA* and multiple *rpf* mutants lacking an *rpfA* gene ..... 100
Figure 35: Drug and detergent susceptibility testing of the wild type, *rpf* mutants and a complemented derivative ..... 104
Figure 36: Colony morphological changes as a result of successive *rpf* gene deletion. ..... 108
Figure 37: Retarded biofilm formation in multiple *rpf* mutants ..... 109
Figure 38: High magnification Scanning Electron Microscopy images showing cell morphological defects associated with *rpfA* gene loss ..... 112

## List of Tables

Table 1: Strains and plasmids used in this study ..... 32
Table 2: Oligonucleotide sequences used in the construction of *rpf* knockout strains in *M. smegmatis*. Bold bases represent engineered restriction enzyme recognition sequences for cloning purposes ..... 44
Table 3: Allelic-specific oligonucleotide sequences used in this study for *rpf* deletion mutant DNA screening. F and R represent forward and reverse primers respectively. ..... 49
Table 4: Restriction mapping for pNΔRPFA ..... 74
Table 5: pNΔRPFB restriction mapping ..... 75
Table 6: Restriction mapping of Δ*rpfA* suicide vector (pSVΔRPFA) ..... 79
Table 7: Δ*rpfB* suicide vector (pSVΔRPFB) restriction mapping. ..... 80
Table 8: Mapping of pGΔRPFA construct by restriction digestion. ..... 82
Table 9: pGΔRPFB restriction mapping ..... 84

**Nomenclature**

| | |
|---|---|
| AAP | Antarctic Alkaline Phosphatase |
| Amp | Ampicillin |
| bp | Base pairs |
| CFU | Colony forming unit |
| CSPD | Disodium 2-chloro-5-(4-methoxyspiro (2-dioxetane-3,2 (2-dioxetane-3,2-(5'-chloro)-tricyclo[3.3.1.1.3,7.]decan)-.4-yl)-1-phenyl phosphate |
| CTAB | Cetyltrimethylammonium bromide |
| DCO | Double cross over |
| DIG | Digoxigenin |
| DNA | Deoxyribonucleic acid |
| dNTP | Deoxynucleotide triphosphate |
| DOTS | Directly Observed Therapy – Short Course |
| EDTA | Ethylenediaminetetraacetic acid |
| Emb | Ethambutol |
| FIB | Focused ion beam |
| g | gram |
| hr | Hours |
| HCl | Hydrochloric acid |
| HIV | Human Immunodeficiency Virus |
| *hyg* | Gene encoding hygromycin B resistance |
| Hyg | Hygromycin |
| Inh | Isoniazid |
| kb | Kilo base pair |
| Km | Kanamycin |
| LA | Luria-Bertani agar |
| *lacZ* | Gene encoding β-galactosidase |
| LB | Luria-Bertani broth |
| LTBI | Latent tuberculosis infection |

| | |
|---|---|
| MDR-TB | Multidrug-Resistant Tuberculosis |
| min | Minutes |
| ml | Millilitre |
| MOPS | 3-(N-morpholino)propane sulphonic acid |
| NaCl | Sodium chloride |
| NAG | *N*-acetylglucosamine |
| NAM | *N*-acetylmuramic acid |
| $NaNO_2$ | Sodium nitrite |
| NaOH | Sodium hydroxide |
| $OD_{600}$ | Optical density at 600 nanometre wavelength |
| PCR | Polymerase Chain Reaction |
| PG | Peptidoglycan |
| Pza | Pyrazinamide |
| R | Resistant |
| Rif | Rifampin |
| Rpf | Resuscitation promoting factor |
| rpm | Revolutions per minute |
| *sacB* | Gene encoding levansucrase |
| SCO | Single cross over |
| $sdH_2O$ | Sterile distilled water |
| SDS | Sodium dodecylsulphate |
| SEM | Scanning electron microscopy |
| Suc | Sucrose |
| TB | Tuberculosis |
| TE | Tris-EDTA |
| TG | Transglycosylation |
| TP | Transpeptidation |
| Tris | Tris(hydroxymethyl)aminomethane |
| Tween | Polyoxyethylene sorbitan monooleate |
| VBNC | Viable but non-culturable |
| v/v | Volume per volume |

| | |
|---|---|
| w/v | Weight per volume |
| WHO | World Health Organization |
| XDR-TB | Extensively Drug Resistant Tuberculosis |
| X-gal | 5-bromo-4-chloro-3-indolyl-α-D-thiogalactopyranoside |

## 1. Introduction

*Mycobacterium tuberculosis*, the causative agent of tuberculosis (TB) is a formidable human pathogen that has infected one third of the world's population and continues to cause an astounding 9.27 million new cases of TB and 1.3 million deaths annually (WHO, 2009), despite the availability of a vaccine and a chemotherapeutic regimen. The situation is further exacerbated by the presence of multi-drug (MDR-TB) and extensively-drug resistant (XDR-TB) strains, with estimated figures of MDR-TB cases reaching 0.5 million (WHO, 2009). The majority of the top 15 countries burdened with TB incident cases are African countries, an observation thought to be due to the HIV pandemic resulting in increased rates of HIV co-infection (WHO, 2009). It is estimated that 15% of the total TB incident cases are individuals who are co-infected with HIV (WHO, 2009).

### 1.1 Tuberculosis chemotherapy

Several decades ago, TB was initially treated using monotherapy with isoniazid (INH), streptomycin or pyrazinamide (PZA) which proved effective until the rise of resistant strains. Fortunately, the idea of multi-drug therapy came to the aid of the resistance problem and this form of therapy became standard for the treatment of TB (Hall*, et al.*, 2009). Currently TB therapy consists of two phases; an initial, intensive component, followed by a variable continuation phase, where the duration of the former is two months and a minimum of four months for the latter, depending on drug resistance profile (American Thoracic Society, 2003). There are a few alterations to the abovementioned conventional regimen in both the duration and drug combination of treatment due to varying circumstances presented by patients. In the initial phase; rifampin (RIF), ethambutol (EMB), INH and PZA are administered daily. Therapy then proceeds with RIF and INH only, administered thrice weekly for the final phase (American Thoracic Society, 2003).

This course pertains to adult, untreated TB, with patients that have drug susceptible strains (American Thoracic Society, 2003).

A hypothesis has been formulated from both *in vitro* and *in vivo* observations that a heterogeneous population of *M. tuberculosis* exists in infection and/or in culture consisting of actively- , slowly- and non-replicating bacilli (Mitchison, 1979). Their data suggests that active growth is seen for bacilli that establish an infection in a location where the organism has access to oxygen, in most cases the upper lobes of the lung, slower growth is observed in locations such as closed lesions under hypoxic and/or acidic conditions. They postulate that INH targets the actively-replicating population but is ineffective against the slow-growing cells; RIF is also bactericidal against the actively growing cells but is more potent than INH. PZA unlike INH and RIF targets both the actively- and slowly-replicating cells but is active in an acidic environment (Mitchison, 1979). Thus multidrug-therapy is imperative for treatment of TB. Unfortunately no drug(s) exists, as yet, that target the dormant population (Mitchison, 1979). Dormancy is a reversible state of low metabolic activity, in which bacterial cells have the inability to proliferate and hence are non-culturable without an intervening resuscitation stage (Kell & Young, 2000).

A more recent similar study in *E. coli* demonstrated the existence of this heterogeneous population and the phenomenon, **bacterial persistence** (Balaban*, et al.*, 2004). Bacterial persistence is the ability of cells to survive strong doses of antibiotic treatment thus revealing an inherent heterogeneity within the population, characterized by differing growth rates (Balaban*, et al.*, 2004). The authors' work carried out in wild type and high persistence (*hip*) mutant, *hipA7* and *hipQ*, cells showed that a few unusual cells survive and continue dividing following ampicillin treatment. The *hipA* gene has been implicated in high frequencies of persistence, toxicity and cold sensitivity. Mutations within the gene produce persisters at a significantly higher frequency (Korch*, et al.*, 2003). These were defined as persister cells and were identified to

exhibit diminished growth rates in contrast to the majority of the population prior to antibiotic exposure. Thus both *hipA7* and *hipQ* populations consist of two distinguishable subpopulations, normal cells and persisters, with differing growth rates. The persisters display the ability to assume normal fast growth thereby producing an ampicillin-sensitive population (Balaban*, et al.*, 2004). Observations from the *hipA7* and *hipQ* data led to the definitions of Type I and Type II persisters respectively. The *hipA7* persisters originate from stationary phase in an arrested growth state, while *hipQ* persisters are continuously slowly replicating and are not generated in stationary phase (Balaban*, et al.*, 2004). The *E. coli* wild type population constitutes three subpopulations; normal cells, type I and type II persisters (Balaban*, et al.*, 2004)- a similar situation described above for an *M. tuberculosis* population by (Mitchison, 1979). Persistence in bacteria is a mechanism of survival in changing environments and stress, such as antibiotic treatment (Balaban*, et al.*, 2004) - a mechanism which seems to be used by *M. tuberculosis*.

### 1.2 Drug resistance, HIV coinfection and vaccination

As mentioned earlier, World Health Organization (WHO) reported 0.5 million cases of MDR-TB in 2007 (WHO, 2009) and a minimum of one XDR-TB case in 57 countries by the end of 2009 (WHO, 2009). MDR-TB is defined as the resistance to first line drugs; particularly RIF and INH while XDR-TB is MDR-TB with additional resistance to any fluoroquinolone and either capreomycin, amikacin or kanamycin (Raviglione & Smith, 2007). The emergence of these forms of drug resistant *M. tuberculosis* strains is indicative of poor tuberculosis control within the health care system as these strains arise due to poor diagnosis, incorrect treatment, drug toxicity (American Thoracic Society, 2003, Singla*, et al.*, 2010) and patient non-compliance resulting in transmission to other individuals (Raviglione & Smith, 2007). A hindrance to the efficient identification and treatment of these strains is the lack of appropriate infrastructure, which includes laboratory capacity for culturing organisms (WHO, 2009). DOTS

(Directly Observed Treatment, Short-course) has been implemented globally in an attempt to manage the TB crisis by the identification of individuals with active disease and supervision of patients during treatment thus ensuring completion of treatment. Statistics have indicated that the programme has improved management of the TB crisis in many countries (WHO, 2006).

Bronchoalveolar antigen-specific $CD4^+$ T cells that are elicited due to the presence of bacterial or viral respiratory infections are defective in HIV+ individuals and this fact may play a role in the individuals being prone to TB (Jambo*, et al.*, 2011). It has been hypothesized that HIV exacerbates TB due to increased HIV replication at areas where *M. tuberculosis* infection has occurred thus this results in poor control of bacterial growth within the granulomas (Diedrich & Flynn, 2011).

Currently, the vaccine used to prevent TB is the Bacillus Calmette-Guerin (BCG) which was developed in 1908 and is essentially an attenuated form of *Mycobacterium bovis*. The vaccine is administered to infants in several countries around the globe and has been shown to offer between 67%-79% protection against TB meningitis and 58%-87% protection against miliary disease in young children (Swaminathan & Rekha, 2010). Experimental and observational studies have shown that the BCG vaccine affords 26% and 61% protection against leprosy respectively, although the protective efficacy was overestimated for the latter (Setia*, et al.*, 2006). Repeated vaccination, as opposed to single dose, resulted in increased protective efficacy against leprosy (Setia*, et al.*, 2006); up to 50% in Malawi (Karonga, 1996). This vaccine has proven to be inconsistent and controversial over the years more especially since the rising incidence of HIV infection. Reports have shown that the vaccine caused disseminated BCG disease in South African infants infected with HIV (Hesseling*, et al.*, 2009) and it also triggered *Lupus vulgaris,* a form of cutaneous TB in a child (Najem*, et al.*, 2009). An improved BCG vaccine or a novel TB vaccine would thus be critical to prevent TB.

### 1.3 *M. tuberculosis* pathogenesis and latent TB infection

Individuals presenting with active pulmonary TB are the most infectious within a shared space however is it important to note that not all individuals exposed to *M. tuberculosis* develop an infection and furthermore, not all of those infected develop active disease. The infection cycle begins with exposure to droplet nuclei originating from the chest and/or lungs of infected individuals during coughing. The route of infection is through inhalation of the 1-5 μm nuclei consisting of *M. tuberculosis* cells but successful transmission and type of infection is determined by various factors including; host and pathogen factors, concentration of inhaled nuclei and the duration of exposure (Dheda*, et al.*, 2010). Upon exposure; approximately 50-70% of individuals may rid the system of the infection by either innate or adaptive immunity, while the remaining will develop an infection. Of the infected individuals, 95% contain the infection in the form of latent TB infection (LTBI), and pose a possible risk of reactivation later in life, whereas the remaining 5% will present with active disease (Dheda*, et al.*, 2010).

The miniature droplet nuclei pose a size advantage which permits the *M. tuberculosis* cells to escape bronchi defenses and migrate into the alveoli where phagocytosis occurs via alveolar macrophages, dendritic cells, neutrophils and epithelial cells (Ahmad, 2010). Within the phagocytotic cells, the bacilli replicate intracellularly and are disseminated, mostly in dendritic cells, to the lymph nodes. Thus the phagocytotic cells act as protection from the host immune system and as a means of transport to other sites, which is a remarkable mechanism the tubercle bacilli have adopted. After infection, bacteria begin to grow within macrophages at the site of infections. It is only at this stage that the host adaptive immunity takes charge 2-8 weeks post-infection and replication of the bacilli is halted through the activation of cell-mediated immunity. The alveolar macrophages secrete inflammatory chemical signals indicating the presence of infecting organisms; as a result other cells of the immune system (T-lymphocytes, monocytes and neutrophils) move towards the site of infection to

destroy cells infected with the tubercle baclli but fail to completely eradicate all bacterial cells (Dheda*, et al.*, 2010). The macrophages alike are unsuccessful in the eradication of the tubercle bacilli and this ultimately leads to macrophage necrosis. The destruction of the macrophages results in the release of bacterial cells that then infect other cells and the cycle of phagocytosis is repeated. The fully developed dendritic cells containing the tubercle bacilli disseminate to the lymph nodes where the CD4+ and CD8+ T cells are primed. The primed T cells together with fibroblasts, stromal cells, dendritic cells, epithelioid macrophages, multinucleated giant cells and endothelial cells are directed to the site of infection where they form granulomas. The granulomas restrict dissemination of bacteria to other areas of the lung (Saunders & Britton, 2007, Dutta*, et al.*, 2010). Interferon-γ (IFN-γ) released by CD4+ cells identify macrophages that present antigens against the tubercle bacilli at their surface and destroy these cells. This is evidenced by the presence of granulomatous lesions and dispersed hemorrhages marked by discoloured, purple areas on the lung surface (Dutta*, et al.*, 2010).

With LTBI, some *M. tuberculosis* cells evade the host immune responses and remain within granulomas for prolonged periods (Tufariello*, et al.*, 2003, Saunders & Britton, 2007). An individual who appears healthy but presents with delayed-type hypersensitivity reaction to purified protein derivative (PPD), mycobacterial antigens, is presumed to be latently infected (Jafri*, et al.*, 2011). A more potent T-cell response is elicited in LTBI against antigens that are associated with latency (Tully*, et al.*, 2005, Schuck*, et al.*, 2009). It has been demonstrated that CD8+ T cells do not play a role in acute infection but rather in latent infection where the production of IFN-γ is observed. Furthermore, depletion of CD8+ T cells and IFN-γ during latent infection permits for a higher bacterial load in the lungs of *M. tuberculosis*-infected mice while no differences are observed in acute infection (van Pinxteren*, et al.*, 2000).

LTBI can be defined as an asymptomatic state lacking clinical or radiological signs of disease but is characterized by the presence of viable tubercle bacilli within the host (Dheda*, et al.*, 2010). This definition can be mirrored with the microbiological phenomenon of *in vitro* bacterial dormancy where these cells are said to be viable but non-culturable (VBNC) (Shleeva*, et al.*, 2002). The subject of latency is confounding and controversial in the TB literature as was previously observed in the case of Human Immunodeficiency Virus-1 (HIV-1). It is proposed that there are different responses to TB infection: innate and acquired immunity, and the degree of these responses vary with latent and active infection (Young*, et al.*, 2009). Latent infection is a consequence of an equilibrium between the host and pathogen factors resulting in controlled infection; in other words the bacilli are thought to be viable but are not actively dividing, although no evidence exists to support the idea (Young*, et al.*, 2009). It has long been hypothesized that LTBI is characterized by the presence of bacteria in a non-replicating or dormant-like state however, this has been challenged recently. A key finding in this regard is the observation that chronic, stationary infection in mice is characterized by bacilli that are continuously replicating albeit at a reduced rate (Gill*, et al.*, 2009); suggesting that latent infection may be characterized by a dynamic population of cells. However, it is important to note in this regard that the mouse is a poor model of latent TB infection.

LTBI poses a huge public health threat in the global burden of active TB thus it is a priority to comprehend the biology of *M. tuberculosis* during LTBI in order to develop interventions that target this form of infection. Critical in this regard are the factors responsible in the revival of dormant bacteria, more especially *M. tuberculosis*, *in vitro* and the best studied thus far is the Resuscitation Promoting Factor (Rpf) (Mukamolova*, et al.*, 1998, Shleeva*, et al.*, 2002). The question is “Do Rpfs have roles in human tuberculosis infection?” Indeed it has been shown that the vast majority of *M. tuberculosis* cells in smear-positive sputa are revealed only by the addition of Rpfs (Mukamolova*, et al.*, 2010), therefore this confirms that these factors do play an important role in human infection. They

also represent an untapped source of novel targets for anti-tubercular drug discovery as demonstrated by the identification of 2-nitrophenylthiocyanates (NPT) compounds which inhibit Rpf function (Demina*, et al.*, 2009).

### 1.4 Resuscitation promoting factors

Rpf is a protein first identified in *Micrococcus luteus*; a non-sporulating, Gram-positive bacterium (Mukamolova*, et al.*, 1998). This secreted 16-17 kDa protein was shown to stimulate the growth of both actively replicating and dormant *Mi. luteus* cells*;* furthermore it was responsible in the reduction of the apparent lag phase of the same bacterial cells (Mukamolova*, et al.*, 1998, Mukamolova*, et al.*, 1999). Moreover growth stimulation and revival by purified, recombinant Rpf, from *Mi luteus* was also observed with different mycobacterial species including *M. tuberculosis*, *Mycobacterium avium*, *Mycobacterium kansasii*, *Mycobacterium bovis* and *Mycobacterium smegmatis* (Mukamolova*, et al.*, 1998, Biketov*, et al.*, 2000, Shleeva*, et al.*, 2002, Shleeva*, et al.*, 2004). It has been suggested that this protein has properties of both autocrine and paracrine signaling molecules and has thus been termed a “bacterial cytokine” (Mukamolova*, et al.*, 1998). A search of *rpf* homologues in other bacteria demonstrated that *rpf* genes are present in many diverse actinobacteria, with some bacteria encoding multiple homologues (Mukamolova*, et al.*, 1998, Mukamolova*, et al.*, 2002, Ravagnani*, et al.*, 2005, Schroeckh & Martin, 2006).

In *Corynebacterium glutamicum* two *rpf*-like homologues, *rpf1* and *rpf2* have been identified and the gene products are secreted as evidenced by the presence of leader peptides. It has also been demonstrated that one of these, Rpf2, is glycosylated by a specific *pmt*-encoded glycosyl transferase (Manhe et al., 2006). Both proteins have in common the *rpf* domain which shares homology to that of *Mi. luteus*, this domain is positioned at the N- and C-terminus in Rpf1 and Rpf2 respectively (Hartmann*, et al.*, 2004). It has been demonstrated that both genes can be deleted from the genome of *C. glutamicum* without abrogation

of growth indicating that they are not essential for viability unlike the essentiality of the single *rpf* gene in *Mi. luteus* (Mukamolova*, et al.*, 2002). Interestingly, the cell surface associated Rpf2 has three conformational forms of different sizes (Hartmann*, et al.*, 2004) but these forms are not distinguished by glycosylation as previously thought (Mahne*, et al.*, 2006). The *rpf1*, *rpf2* double mutant (MH28) displays a growth defect *in vitro* and the phenotype was reversed by the simultaneous addition of both genes in the mutant background implying that the observed defect was due to loss of both *rpf* genes (Hartmann*, et al.*, 2004).

*Salmonella typhimurium* LT2 also encodes an *rpf* homologue, it has been demonstrated that the recombinant Rpf of this organism has the ability to reverse the VBNC (viable but non-culturable) state of *S. oranienburg* cells to a culturable state by recombinant Rpf supplementation to growth media (Panutdaporn*, et al.*, 2006). However, the Rpf does not possess hydrolytic activity like all Rpf-like proteins therefore it may not represent true Rpf (Kana & Mizrahi, 2009).

### 1.4.1 *Micrococcus luteus* Rpf

Contrary to what was previously thought; the genome sequence of *Mi luteus* revealed that only a single copy of the *rpf* gene exists in *Mi. luteus* which is located within a monocistronic operon (Mukamolova*, et al.*, 1998, Mukamolova*, et al.*, 2002). *In vitro* assays have shown that *rpf* transcription commences very early during bacterial growth and copious transcripts are detected during the lag phase (Mukamolova*, et al.*, 2002). A massive decrease in transcript level is observed during mid-log phase until none can be detected in late stationery phase. The presence of Rpf in culture supernatants is delayed in relation to these transcriptional changes and the protein is possibly secreted in different conformational forms (Mukamolova*, et al.*, 2002, Hartmann*, et al.*, 2004, Mahne*, et al.*, 2006). It has been reported that Rpf from environmental *Mi. luteus* strains is present in culture supernatants at low concentrations probably due to the altered structure of the protein, where the linker region that connects the protein

to the LysM domain is extended, resulting in an inability of the protein to be secreted into the medium, thus tethering it rather tightly to the cell wall (Koltunov*, et al.*, 2010). Rpf molecules are located on the surface of bacterial cells (Mukamolova*, et al.*, 1998, Mukamolova*, et al.*, 2002) and this is the most likely reason the accumulation of Rpf in culture supernatants is observed possibly explaining the disconnect between extracellular abundance and transcription (Mukamolova*, et al.*, 1998, Mukamolova*, et al.*, 2002).

### 1.4.2 Mycobacterial *rpf*-like genes

*M. tuberculosis* and *M. bovis* encode five Rpf-like genes, designated *rpfA-E* whereas two, possibly three (Kana & Mizrahi, 2009), have been uncovered in *M. leprae* and four in *M. smegmatis* (Mukamolova*, et al.*, 1998). Like *M. smegmatis*; *rpfA-C* and *rpfE* homologues are present in *M. ulcerans*, *M. avium* and *M. marinum* genomes (Kana & Mizrahi, 2009). The products of all five *M. tuberculosis rpf*-like genes have in common a conserved ca. 70 amino acid containing domain, termed the Rpf domain which shares high homology with a similar domain from *Mi. luteus* Rpf, Figure 1. All five proteins have predicted leader sequences suggesting that they are transported to the cell wall for secretion (Mukamolova*, et al.*, 1998, Mukamolova*, et al.*, 2002, Zhu*, et al.*, 2003). The Rpf-like domain of RpfA is followed by a large region of proline + alanine-rich repeats (Figure 1) consisting of APADLAPP as the consensus sequence. RpfB is the only lipoprotein predicted to be membrane bound since it consists of a prokaryotic membrane lipoprotein lipid adherence site situated at its N-terminus, thus it is the only cell-wall associated Rpf (Mukamolova*, et al.*, 2002, Rezwan*, et al.*, 2007). The Rpf-like domains of all the proteins, excluding RpfA, in *M. tuberculosis* are positioned c-terminally (Mukamolova*, et al.*, 2002). The biological activities of the *M. tuberculosis* recombinant Rpfs have been tested *in vitro* and cause a reduction in the apparent lag phase of *M. tuberculosis* (Zhu*, et al.*, 2003), *Mi. luteus* and *M. smegmatis (Mukamolova, et al., 2002)*.

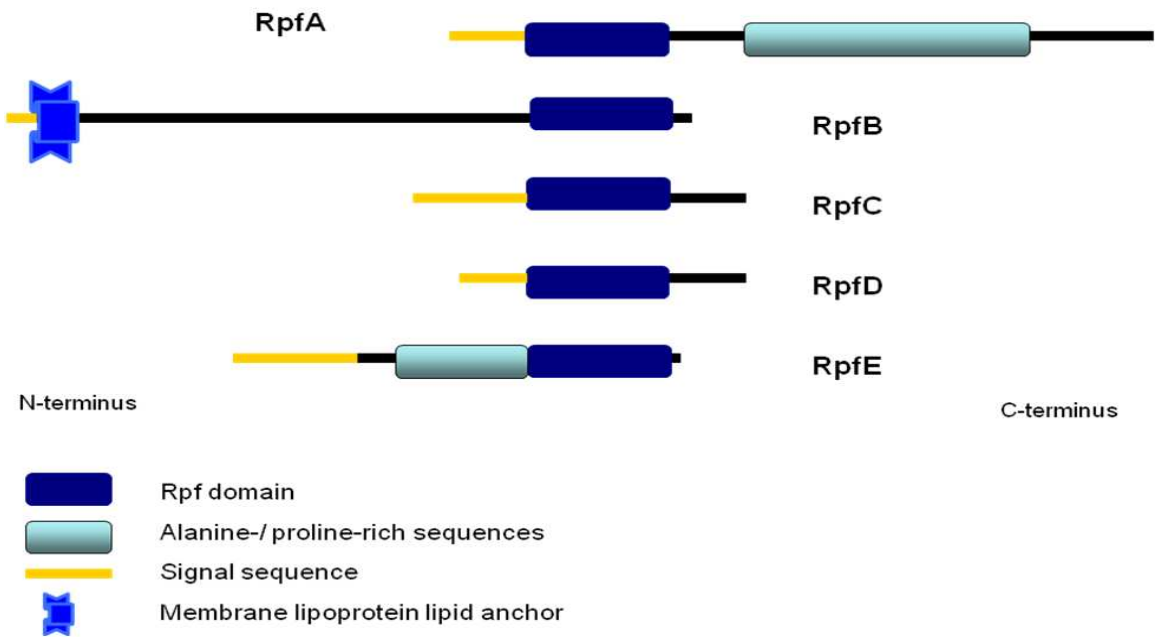



**Figure 1:** A schematic representation of the *M. tuberculosis* Rpf-like proteins showing the presence and positioning of the common Rpf domain [adapted from (Mukamolova*, et al.*, 2002)].

*M. avium* subsp. *paratuberculosis* recombinant Rpf also reduces the lag phase and has a growth-stimulatory effect on freeze-dried cells (Wu*, et al.*, 2008) or dormant-like, stationary phase cells which was further confirmed by CFU quantification (Zhu*, et al.*, 2003). In *Mi. luteus* culture grown in minimal medium, the *M. tuberculosis* RpfA and RpfC proteins display outstanding (lag phase reduction) activity whereas RpfA, RpfC, RpfD and RpfE proteins show comparable activity in *M. smegmatis* culture (Mukamolova*, et al.*, 2002). Activity profiles for Rpfs tested show that an optimal concentration range exists, where less activity is observed when the range is exceeded or completely abolished when it is sub-optimal (Mukamolova*, et al.*, 2002, Zhu*, et al.*, 2003, Koltunov*, et al.*, 2010). All five recombinant proteins of *M. tuberculosis* stimulate the growth of *M. bovis,* and the addition of anti-Rpf antibody suppress growth confirming that the Rpf-like proteins are responsible for the growth-stimulatory effect from an extracellular location (Mukamolova*, et al.*, 2002). Thus all the *M. tuberculosis*

Rpf-like proteins also stimulate growth of other organisms, which is a property shared with *Mi. luteus* Rpf (Mukamolova*, et al.*, 1998, Mukamolova*, et al.*, 2002).

### 1.4.3 Rpf domain

As mentioned earlier, all Rpf proteins have in common a conserved ca. 70 amino acid region, the **Rpf domain** (Mukamolova*, et al.*, 1998). The three dimensional structure of the mycobacterial Rpf domain is comparable to that of *c*-type lysozyme, moreover there is high sequence conservation over a stretch of 110 common sequences between Rpf, and eukaryotic *c*-type and *g*-type lysozymes (Cohen-Gonsaud*, et al.*, 2004). Lysozymes are hydrolytic enzymes which act on the peptidoglycan by cleavage of the β-(1,4)-glycosidic bond between *N*-acetylmuramic acid (NAM) and *N*-acetylglucosamine (NAG) residues; *c*-type and *g*-type refer to chicken/conventional type and goose type respectively (Callewaert & Michiels, 2010). Structural similarities are conserved between Rpfs and *c*-type lysozymes; which include residues within the catalytic site (catalytic proton donor or active site glutamate), the glycine residue which is compulsory structurally and residues residing within the saccharide-binding groove (Cohen-Gonsaud*, et al.*, 2004).

### 1.4.4. Muralytic activity

Since Rpf possesses a lysozyme-like fold, zymography confirmed that it is indeed a muralytic enzyme (Mukamolova*, et al.*, 2006). Other evidence in this regard includes the observation, *in vitro*, that bacterial lysis results from the over-expression of Rpf in the periplasm of *Escherichia coli*. In addition, an assay displaying hydrolysis of fluorescamine- labeled *Mi. luteus* cell wall due to the activity of recombinant Rpf confirmed that Rpf displays muralytic activity (Mukamolova*, et al.*, 2006). Moreover, the ability displayed by the recombinant Rpf to hydrolyse 4-methylumbelliferyl-β-D-N,N',N''-triacetylchitotrioside, a

synthetic lysozyme substrate, infers that its muralytic activity is comparable to that of conventional lysozymes (Mukamolova*, et al.*, 2006, Telkov*, et al.*, 2006). *In vitro* assays demonstrate that the conserved domains found in RpfB (Rv1009), RpfC (Rv1884c) and RpfD (Rv2389c) are alone capable of reviving dormant *Mi. luteus* cultures at low (picomolar) concentrations (Cohen-Gonsaud*, et al.*, 2005). This indicates that the muralytic domain can be held accountable for the resuscitation activity (Cohen-Gonsaud*, et al.*, 2005).

A glutamate residue (E54) (Mukamolova*, et al.*, 2006) is conserved in the Rpf domain structure which is a property also shared with *c*-type lysozymes (Cohen-Gonsaud*, et al.*, 2004). Site-directed mutagenesis experiments in *Escherichia coli* indicated that the *c*-type lysozyme glutamate residue (E35) was essential for catalysis (Malcolm*, et al.*, 1989, Callewaert & Michiels, 2010). In light of this observation, it was logical to infer that the glutamate residue in Rpf proteins may also play a critical role in catalysis. However, E54 isosteric substitutions to glutamine, alanine and lysine in Rpf confirm that the glutamate residue is significant but certainly not essential for the functioning of Rpf as the muralytic activity was still maintained although defects in resuscitation were noted (Mukamolova*, et al.*, 2006, Telkov*, et al.*, 2006). The degree to which the muralytic activity of Rpf is affected by the isosteric substitutions in the abovementioned experiments depended solely on the amino acids tested (Mukamolova*, et al.*, 2006, Telkov*, et al.*, 2006). Thus the glutamate residue serves as a hallmark feature in Rpf-like proteins. A pair of cysteine residues which have been proposed to form a disulphide bridge (Cohen-Gonsaud*, et al.*, 2004) are highly conserved in Rpf-like proteins and mutation of either one or both residues have equally noticeable impacts on muralytic activity but more dramatic effects on resuscitation activity are observed when both cysteines are mutated as opposed to only one (Mukamolova*, et al.*, 2006, Telkov*, et al.*, 2006). It is proposed that the cysteine residues not only play a role in Rpf secondary structure stabilization but may also participate in the modulation of Rpf activity (Mukamolova*, et al.*, 2006).

### 1.4.4.1 RpfB domain: NMR structure

The NMR domain structure of *M. tuberculosis* Rv1009, also referred to as RpfBc, has been determined (Cohen-Gonsaud*, et al.*, 2005) and it revealed an oligosaccharide-binding site cleft exists in the protein, which is also observed in c-type lysozymes indicating that RpfBc attaches to polysaccharide (Cohen-Gonsaud*, et al.*, 2005). Analysis of RpfBc demonstrates that the Rpf domain has structural properties of both c-type lysozymes and soluble lytic transglycosidases (SLT's) (Cohen-Gonsaud*, et al.*, 2005). Lytic transglycosylases are bacterial enzymes known to cleave the β-1,4 glycosidic bonds between N-acetylmuramoyl and N-acetylglucosaminyl resulting in the formation of 1,6-anhydromuramoyl end product (Holtje*, et al.*, 1975, Heidrich*, et al.*, 2002, Scheurwater*, et al.*, 2008). Moreover, experimental data indicate that chemical shift modifications within the binding groove occur due to bound RpfBc to *N,N',N''*-triacetylchitotriose implying that Rpf serves to cleave oligosaccharides (Cohen-Gonsaud*, et al.*, 2005). Hence, it is postulated that this action (oligosaccharide cleavage) results in bacterial resuscitation from dormancy; however the mechanism of peptidoglycan degradation in Rpfs is unclear and the chemical nature of the resulting products are unknown (Cohen-Gonsaud*, et al.*, 2005). Many peptidoglycan-degrading enzymes function in complexes with other proteins and it has been proposed that Rpfs function together with other PG degrading enzymes (Cohen-Gonsaud*, et al.*, 2005).

### 1.4.4.2 $\Delta_{DUF}$RpfB: Crystal structure

The crystal structure of a truncated form of RpfB, $_{\Delta DUF}$RpfB, revealed that it consists of a **catalytic domain** and a **G5 domain**, assuming a "comma-like shape" (Figure 2) (Ruggiero*, et al.*, 2009). The catalytic domain bears a likeness to that of the NMR structure (Cohen-Gonsaud*, et al.*, 2005), and as previously described, possesses certain properties of c-type lysozymes (Cohen-Gonsaud*,*

*et al.*, 2005). In addition to that, structurally it is comparable to invertebrate- type (i-type) lysozymes (Ruggiero*, et al.*, 2009).

The G5 domain presents a unique super-secondary structure consisting of two three-stranded β-sheets, β1 and β2 (Ruggiero*, et al.*, 2009) which are connected by cross-linking strands (Figure 2) (Ruggiero*, et al.*, 2009). The G5 domain family, comprising 35 members, was initially observed in the Pfam-B database and the G5 refers to the glycine residues that are conserved within the structure (Bateman*, et al.*, 2005). The domain is present in various domain designs in bacterial proteins that play a role in pathogenesis and antimicrobial resistance; furthermore it is generally affiliated with an enzymatic domain (Bateman*, et al.*, 2005) like the catalytic domain in the case of RpfB (Ruggiero*, et al.*, 2009). It has been proposed that it is probable that the G5 domain adheres to N-acetylglucosamine (NAG), one of two monomers making up the glycan strands of the peptidoglycan (Vollmer*, et al.*, 2008), which may facilitate cell wall degradation by the catalytic domain in Rpfs (Bateman*, et al.*, 2005).

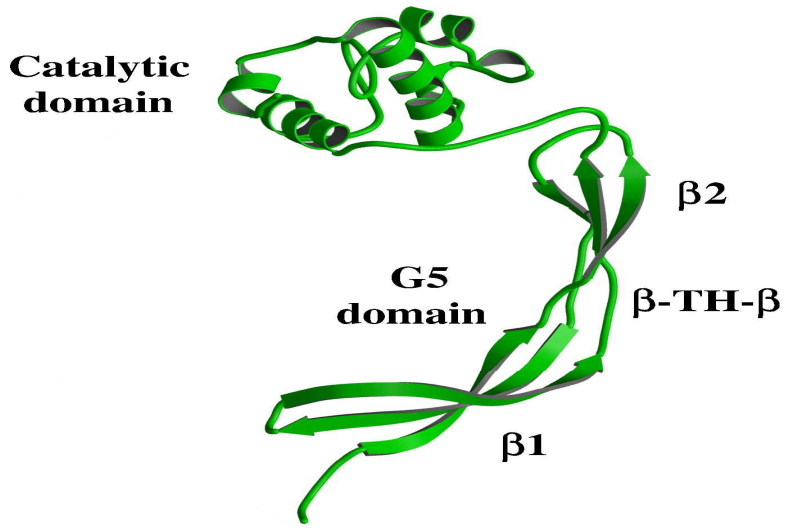



**Figure 2**: Crystal structure of $_{\Delta DUF}$RpfB consisting of two distinct domains: catalytic and G5 domains (taken from Ruggiero *et al*, 2009).

In accordance to the name given to the domain, not all five glycine residues are conserved in RpfB; only Gly245, Gly269, Gly229, Gly232 and the last one is replaced by a serine residue (Ser221). Gly229 and Gly232 are situated within the DP**G**VP**G** sequence, a sequence comparable to collagen sequence, found in one of the cross-linking strands connecting the triple-stranded β-sheets (Ruggiero*, et al.*, 2009). The two glycine residues function to stabilize the β-TH-β motif (Figure 2) and this is the most probable reason they are most conserved (Ruggiero*, et al.*, 2009). The lengthy G5 domain is designed in such a way that ample β-sheet backbone atoms are exposed (Ruggiero*, et al.*, 2009) and this structural property has been implicated in protein aggregation and possibly association with the cell wall (Srisailam*, et al.*, 2002, Esposito*, et al.*, 2006). In light of the structural properties of the G5 domain described above and the fact that it probably binds NAG (Bateman*, et al.*, 2005), implies that the domain serves an adherence function to the cell surface of bacteria (Ruggiero*, et al.*, 2009). The presence of both catalytic and G5 domains in RpfB provides convincing evidence for the possible mechanism of action of the protein.

#### 1.4.4.3 Synergistic proteins: RpfB and RipA

It was hypothesized that in light of the Rpf-based reactivation function initially observed in *Mi. luteus*, a second protein which is secreted or surface-associated could be present that works in synergy with Rpf (Hett*, et al.*, 2007). Rpf-interacting protein A (RipA) has been identified by yeast two-hybrid screen assay, which showed that it interacts with RpfB and RpfE (Hett*, et al.*, 2007). The 472 amino-acid protein is encoded by the Rv1477 gene in *M. tuberculosis* and its predicted domain, which is situated at its C-terminus and consists of 105 amino acids, shares 40% homology to the p60 protein from *Listeria monocytogenes* (Hett*, et al.*, 2007). It was previously shown that a *L. monocytogenes* strain lacking this gene (∆p60) displayed atypical cell division (Pilgrim*, et al.*, 2003) and furthermore, the absence of autolytic bands in zymograms in the same strain provided evidence that p60 protein cleaves the bacterium's cell wall (Lenz*, et al.*,

2003). Thus p60 can be considered a cell wall hydrolase. It has also been demonstrated that RipA has the ability to cleave *Mi. luteus, M. smegmatis* and *Streptomyces* cell wall thus confirming that the protein can degrade peptidoglycan (Hett*, et al.*, 2007, Hett*, et al.*, 2008). The peptidoglycan, also referred to as murein, is a large complex molecule that envelopes the cytoplasmic membrane and is defined by glycan strands inter-connected by small peptides. It provides protection from cell rupture due to turgor pressure (Vollmer & Bertsche, 2008). RipA can therefore also be regarded a hydrolase or an endopeptidase (Hett*, et al.*, 2007) like the p60 protein.

#### 1.4.4.4 RipA$_{263-472}$: Crystal structure

RipA consists of four defined areas according to its sequence analysis; the N-terminal signal sequence, the predicted NlpC/P60 catalytic domain positioned at its C-terminus and two as yet uncharacterized domains, PB07342 and PB015164 (Hett*, et al.*, 2007, Ruggiero*, et al.*, 2010). Recombinant forms of RipA were constructed; one lacking the signal sequence (RipA$_{40-472}$), the other consisting only of the predicted catalytic domain and PB015164 (RipA$_{263-472}$) and the last retaining only the catalytic domain (RipA$_{332-472}$) (Ruggiero*, et al.*, 2010). The crystal structure of RipA$_{263-472}$ reveals a very close association between the catalytic and PB015164 domains; moreover the PB015164 domain forms an obstruction over the RipA catalytic site cleft suggesting that RipA is produced in an inactive form in a manner akin to zymogens (Ruggiero*, et al.*, 2010). A zymogen, also referred to as proenzyme or proprotein, is an inactive form of an enzyme that has to undergo biochemical processing in order to mature and perform its function (Basak, 2005, Mallorqui-Fernandez*, et al.*, 2008). RipA is active as an enzyme only when it is processed; this is evidenced by the *in vitro* observation that RipA$_{332-472}$ displays the ability to cleave *Micrococcus lysodeikticus* cell wall material whereas RipA$_{263-472}$ does not (Ruggiero*, et al.*, 2010).

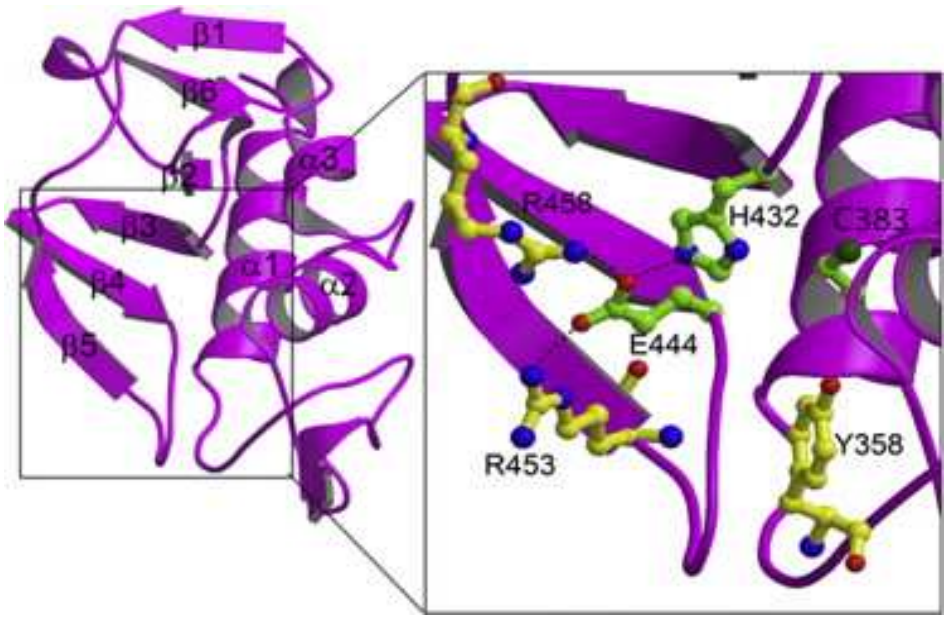



**Figure 3:** The crystal structure of RipA catalytic domain showing the residues involved in catalysis particularly; cysteine383, histidine432 and glutamine444 [Taken from (Ruggiero*, et al.*, 2010)].

The catalytic domain structure, Figure 3, consists of two β sheet types; the major sheet is composed of six β strands arranged in an antiparallel fashion and the other, minor sheet has two β strands all connected to six helices (Ruggiero*, et al.*, 2010). Three residues; cysteine383, histidine432 and glutamate444, shown in Figure 3, are involved in catalysis- an observation atypical of NlpC/P60 domains. Substitution of cysteine to alanine abrogates RipA hydrolytic activity thereby confirming that cysteine is essential for RipA activity (Ruggiero*, et al.*, 2010). RipA therefore belongs to a class of cysteine proteases (Ruggiero*, et al.*, 2010). The loop region of the PB015164 domain, is responsible for creating a physical barrier across the catalytic cleft making the residues inaccessible for activity (Ruggiero*, et al.*, 2010). This prodomain has been shown to be dispensable for enzyme folding as observed *in vitro* from recombinant $RipA_{332\text{-}472}$, thus indicating that it serves a regulatory role for catalysis (Ruggiero*, et al.*, 2010).

RipA is a secreted protein and like RpfB, it is cell-associated; furthermore it has been microscopically demonstrated in *M. smegmatis* that both proteins localize to the septa of replicating cells with the use of red fluorescent protein (RFP) fused to the individual proteins (Hett*, et al.*, 2007). As mentioned earlier, RipA was identified as a partnering protein for RpfB and experimental evidence- both *in vivo* and *in vitro* supports the observation (Hett*, et al.*, 2007). RipA-RpfB interaction has been proven *in vitro* by co-precipitation with the use of fusion proteins and these proteins were also produced and co-purified in *M. smegmatis* providing evidence for interaction (Hett*, et al.*, 2007). Since the proteins have been observed to indeed interact, it was necessary to map the defined area of interaction. The RipA C-terminus was identified by yeast three- and two-hybrid assays to be the region of interaction and further experiments showed that 25 amino acids at the C-terminal end are sufficient for interaction (Hett*, et al.*, 2007). Moreover, deletion of amino acids within the C-terminus results in weakened interaction and deletion of the N-terminus has no effect on interaction confirming that the RipA C-terminus is necessary for interaction (Hett*, et al.*, 2007).

#### 1.4.4.5 RipA-RpfB complex: Mechanism of action

As mentioned earlier, RipA and RpfB have been demonstrated to co-localize at the septum of dividing cells (Hett*, et al.*, 2007). The above observations may indicate the mechanism of action of the synergizing proteins resulting in the degradation of peptidoglycan possibly allowing bacteria to exit a state of dormancy and consequently in bacterial revival (Hett*, et al.*, 2007).

To comprehend the role that each partnering protein plays in the RipA-RpfB complex, an investigation of *M. smegmatis* strains depleted of *ripA* and *rpfB* has been undertaken (Hett*, et al.*, 2008). Deletion of the *ripA* gene (*MSMEG_3153*) proved impossible which indicates that it is an essential gene in *M. smegmatis*; however the knockdown of *ripA*, under the control of a tetracycline promoter, was successful (Hett*, et al.*, 2008). The *ripA*-depleted strain is growth defective and

bacterial clumps are evident in culture, furthermore this strain forms branching filaments or chain-like bacteria with the presence of nucleoids and septa at intervals along the filament (Hett*, et al.*, 2008). Conversely, the *rpfB* depleted strain displays normal growth and morphology; therefore these observations suggest that *ripA* is important for typical cell division and growth whereas the opposite is true for *rpfB* (Hett*, et al.*, 2008). Reintroducing a copy of the *M. tuberculosis ripA* gene into the *ripA*-depleted *M. smegmatis* strain reverses the observed phenotype, implying that the biological activities of both *M. tuberculosis* and *M. smegmatis* RipA proteins are functionally alike. RipA is therefore indispensable for the formation of daughter cells arising from cell division (Hett*, et al.*, 2008).

The *ripA*-depleted *M. smegmatis* strain displays an increased susceptibility to β-lactam drugs, specifically carbenicillin (Hett*, et al.*, 2008), although the molecular basis for this increase in susceptibility is unknown. The data suggests that the biological activity of RipA is to degrade the peptidoglycan by the hydrolysis of the peptide cross-linkages; lack of this activity may change PG structure/stability and thus predispose cells to antibiotic activity (Hett*, et al.*, 2008). The proposed mechanism of action of RipA and the activity of RpfB provides a synergistic effect in the degradation of the peptidoglycan (Hett*, et al.*, 2008).

#### 1.4.4.6 PBP1 and RipA interaction

Further evidence for a peptidoglycan degrading complex of proteins was provided by the demonstration that the C-terminus of the mycobacterial **Penicillin-binding protein 1 (PBP1),** which retains the penicillin binding transpeptidase domain, interacts with RipA (Hett*, et al.*, 2010). The interaction was confirmed *in vivo* and *in vitro* as exemplified by RipA-RpfB interaction. The signal sequence and a transglycosylase domain are both positioned at the N-terminus. Both the penicillin binding transpeptidase and transglycosylase domains show homology to *Esherichia coli* (*E. coli*) PBP1 (Hett*, et al.*, 2010).

PBP1a and PBP1b serve different functions in *E. coli*; the former is responsible for synthesis of the basic framework of the peptidoglycan by the assembly of NAM and NAG monomers that build the glycan strands by transglycosylation, and catalyzing the reaction by transpeptidation creating cross-linkages between stem peptides (Born*, et al.*, 2006, Vollmer & Bertsche, 2008). The latter is responsible for synthesis of linear peptidoglycan chains comprising *N*-acetylglucosaminyl-*β*-1,4-*N*-acetylmuramoyl-$_{L}$-Ala-*γ*-$_{D}$-Glu-*meso*-$A_2$pm-$_{D}$-Ala-$_{D}$-Ala or lipid II units (Fraipont*, et al.*, 2006, Vollmer & Bertsche, 2008). They therefore have almost similar functions (Born*, et al.*, 2006) and are bound at the cytoplasmic membrane but the catalytic domains are located in the periplasm (Born*, et al.*, 2006, Fraipont*, et al.*, 2006, Vollmer & Bertsche, 2008).

It has been proposed that the synergistic hydrolysis rendered by RipA-RpfB complex is suppressed by PBP1; however PBP1 does not interfere with RipA functioning as observed through *in vitro* experiments that measure degradation of Streptomyces cell wall material by the enzymes, implying that PBP1 plays a regulatory role in the function of the RipA-RpfB complex (Hett*, et al.*, 2010). Cellular localization studies indicate that PBP1 is not only confined to the septa of cells but also to the poles, which makes biological sense since it binds RipA. A PBP1-depleted *M. smegmatis* strain displays growth deficiency which is in line with observed morphological alterations and these depleted cells are devoid of septa and display bulb-like ends with diminished cell lengths. These data are indicative of cell elongation and septation roles served by PBP1 in mycobacteria especially since cell wall growth, for this family of bacteria, occurs at the poles (Hett*, et al.*, 2010).

.

### 1.4.5 The peptidoglycan

The mycobacterial cell wall consists of two sections; the outer and inner layers that enclose the cell membrane (Brennan, 2003).

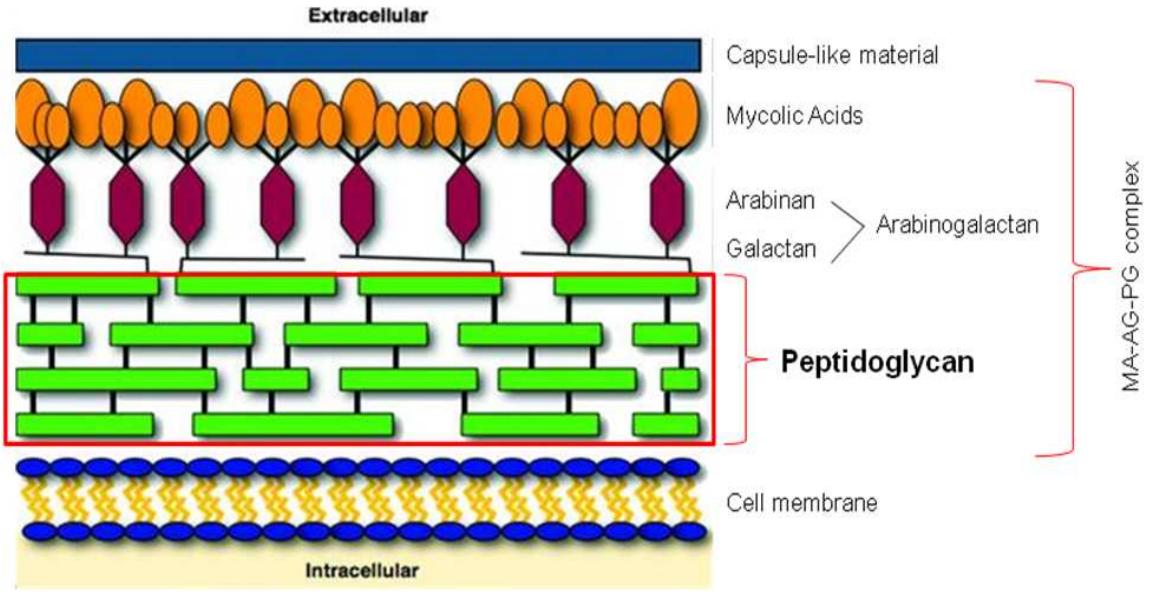



**Figure 4:** The peptidoglycan layer within the mycobacterial cell wall. [Taken from (Hett & Rubin, 2008)].

The cell wall core is composed of the peptidoglycan covalently bonded to the arabinogalactan, and mycolic acids thus making up the mycolyl arabinogalactan-peptidoglycan (mAGP) (Brennan, 2003) or MA-AG-PG complex (MAPc) (Hett & Rubin, 2008) as demonstrated in Figure 4. The cell wall core constitutes the inner layer, while the outer layer consists of lipids and proteins (Brennan, 2003) forming capsule-like material (Hett & Rubin, 2008). Cell wall disturbance results in the solubilization of dissociated lipids and proteins (outer section), while the inner core remains insoluble. This observation indicates the essentiality of the MAPc to bacterial cell viability (Brennan, 2003).

The peptidoglycan, also referred to as murein, is a large complex molecule that envelopes the cytoplasmic membrane and is defined by glycan strands inter-connected through peptide cross-links (Vollmer & Bertsche, 2008).

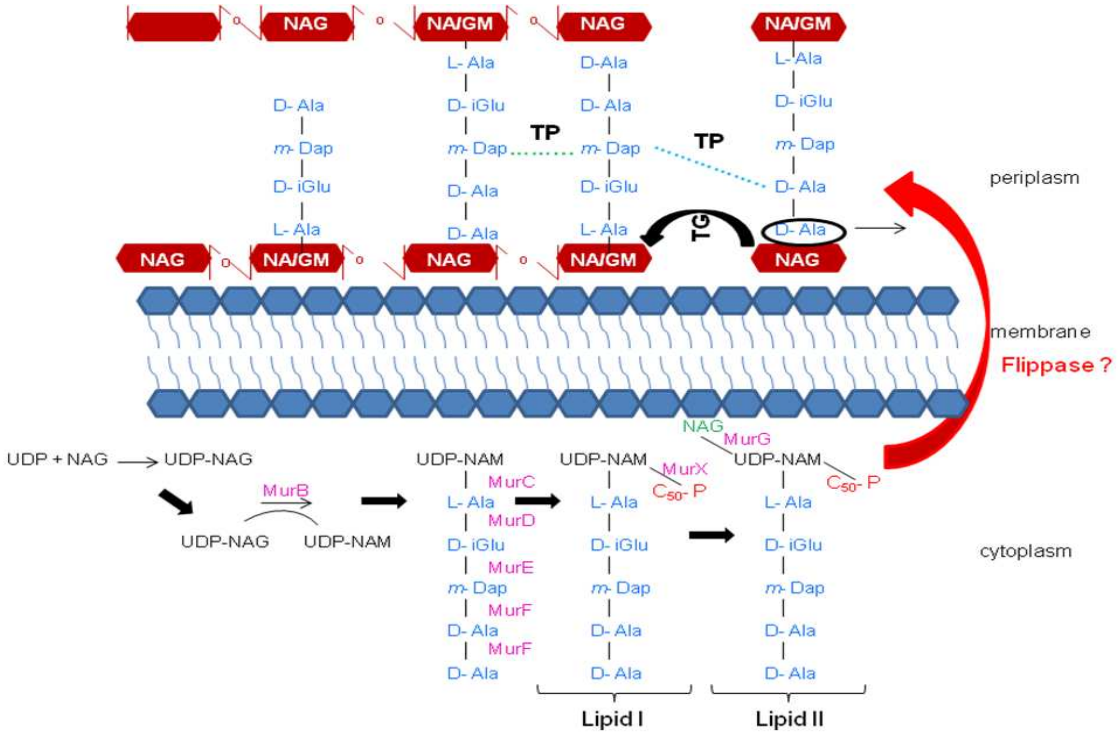



**Figure 5:** Peptidoglycan biosynthesis is initiated with the production of peptidoglycan precursors in the cytoplasm and transported to the periplasm by a possible “flippase”; where transglycosylation (**TG**) and transpeptidation (**TP**) reactions occur forming the peptidoglycan sacculus. [Adapted from (Hett & Rubin, 2008, Kana & Mizrahi, 2009)].

### 1.4.5.1 Peptidoglycan biosynthesis

Peptidoglycan precursors are synthesized within the cytoplasm and transported to the periplasm, where they are cross-linked to form the peptidoglycan sacculus (Hett & Rubin, 2008). Synthesis is initiated with the formation of UDP-NAG; which is converted to UDP-NAM (muramic acid) by a transferase and reductase, MurA and MurB respectively. In mycobacteria, it has been experimentally demonstrated that the muramic acid consists of a combination of *N*-acetyl and *N*-glycolyl residues (Mahapatra*, et al.*, 2005). A suite of enzymes catalyze the sequential addition of L-alanine, D-glutamine, DAP, and D-alanyl-D-alanine amino acids to UDP-NAM as depicted in Figure 5. The UDP-NAM-stem peptide

is displaced to a carrier lipid, decaprenyl phosphate ($C_{50}$-P) by MurX thus forming Lipid I. MurG catalyzes the conversion of Lipid I to Lipid II by a further addition of NAG to Lipid I (Hett & Rubin, 2008). A proposed 'flippase' protein is said to exist which functions to turn Lipid II from the cytoplasmic side to the periplasmic membrane side, although little knowledge is available on the exact mode of action (Bouhss*, et al.*, 2008). A further set of reactions occur in the periplasm; where a transglycosylase reaction takes place resulting in extension of the glycan strand (Hett & Rubin, 2008) by glycosidic linkage (Vollmer & Bertsche, 2008). Peptide cross-links are formed by a transpeptidase reaction forming bonds between the acceptor peptide DAP to either a DAP or the fourth-positioned D-alanine of the incoming peptide. The terminal D-alanine is released, as shown in Figure 5, thereby providing energy for the transpeptidation reaction (Hett & Rubin, 2008).

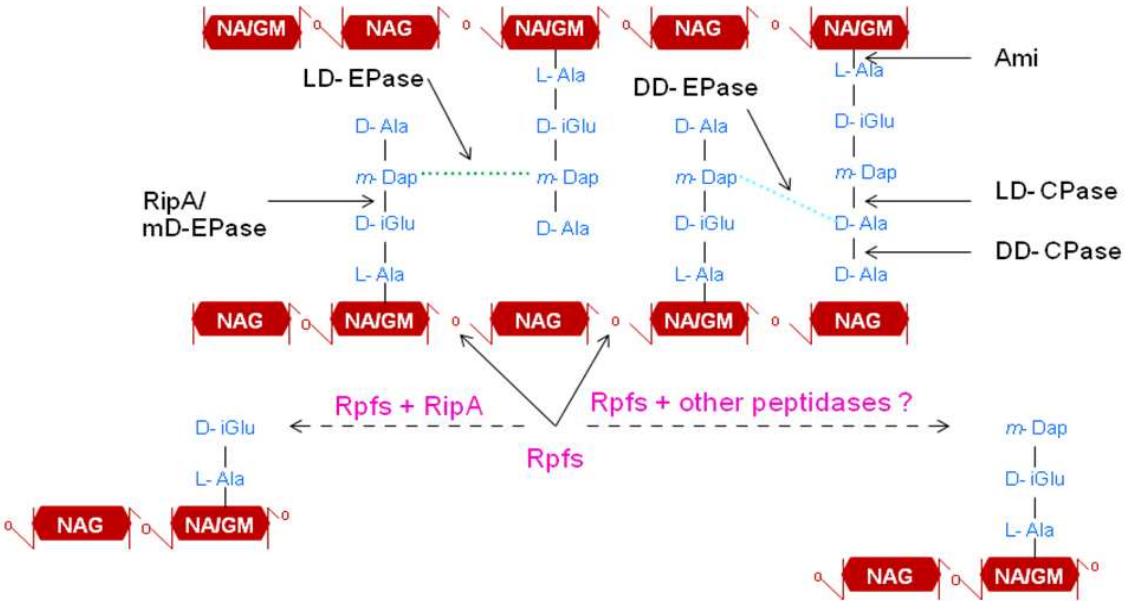



**Figure 6:** Cleavage sites within the peptidoglycan resulting in degradation of the murein by hydrolases. Ami, N-acetylmuramyl-L-alanine amidase; LD/DD- EPase, LD/DD-endopeptidase; LD/DD-CPase, LD/DD-carboxypeptidase. Taken from (Kana & Mizrahi, 2009).

### 1.4.5.2 Peptidoglycan degradation

Degradation of the peptidoglycan results from a host of hydrolases which include; amidases, endopeptidases, carboxypeptidases, glycosidases, *N*-acetylglucosaminidases, lytic transglycosylases and lysozymes (Vollmer*, et al.*, 2008) as shown in Figure 6. The *N*-acetylmuramyl-L-alanine amidases hydrolyse the link that joins the peptide stem to the glycan strand by acting on the amide bond between the L-alanine residue and NAM, as the name suggests (Vollmer*, et al.*, 2008). The endopeptidases and carboxypeptidases act on the peptide stem; cleaving different linkages at different positions within the peptide stem for the former, while removing the C-terminally positioned amino acid from the peptide stem for the latter. Peptidases also act on amide bonds but the hydrolysis also depends on the type of amino acid bond, either between *D*-amino acids or between *L*- and *D*-amino acids (Vollmer*, et al.*, 2008). Enzymes that hydrolyze the glycan strands act on the β-1,4 glycosidic bond between NAM and NAG like lysozymes and lytic transglycosylases, although lytic transglycosylases produce a 1,6-anhydromuramoyl end product (Holtje*, et al.*, 1975). Similarly, as mentioned earlier, the Rpfs are proposed to act on the β-1,4 glycosidic bond. The synergistic interaction of RipA and RpfB has been established (Hett*, et al.*, 2007) but it is speculative that Rpfs may also interact with other peptidases as a mechanism for muropeptide signaling (Kana & Mizrahi, 2009).

### 1.4.6 Biological roles of Rpfs in *M. tuberculosis*

To discern the biological roles of the five *rpf*-like genes in *M. tuberculosis*, gene disruptions have been made for each of the *rpf* homologues by two-step allelic exchange mutagenesis and/or specialized phage transduction (Downing*, et al.*, 2004, Tufariello*, et al.*, 2004). Growth and survival of all the individual *rpf* mutants during logarithmic and stationary phases is comparable to that of the parental wild type strain in liquid culture (Downing*, et al.*, 2004, Tufariello*, et al.*, 2004). It was noted that the Δ*rpfA* mutant displays clumping in liquid culture when this

gene was deleted in the H37Rv background (Downing*, et al.*, 2004). Hence, all five *rpf*-like genes are individually not required for survival in *M. tuberculosis* (Downing*, et al.*, 2004, Tufariello*, et al.*, 2004). The colony morphology of four of the five single mutants is comparable to that of the wild type however, the Δ*rpf*B mutant (Rv1009 deletion) displays a distinct phenotype where the size of the colony is reduced, in contrast to wild type and this observation is explained to be as a result of polar effects conferred by the downstream *ksgA* gene (Tufariello*, et al.*, 2004). The reduced colony size phenotype in Δ*rpfB* is not observed in a H37Rv background from a distinct study, due to the different nature of the deletion, instead fragile, dry colonies on solid media are seen (Downing*, et al.*, 2004). Interestingly all other mutants, except Δ*rpf*A, exhibit alterations in colony morphology which may be indicative of the function that the gene products serve in cell wall biology (Downing*, et al.*, 2004).

Growth of the mutants *in vivo* (murine aerosol infection model) shows similar observations to those seen *in vitro*. The growth kinetics of the individual mutants in the organs of mice is comparable to that of wild type, moreover the histopathology of the organs indicate that both mutant and wild type cells behave similarly. Thus, growth and persistence of cells is unaffected by single *rpf* gene disruptions (Tufariello*, et al.*, 2004, Downing*, et al.*, 2005, Tufariello*, et al.*, 2006). However, other *in vivo* data report an extended duration of survival of mice challenged with the ΔRv1009 (Δ*rpfB*) mutant strain in chronic infection, furthermore the bacterial loads are conspicuously lower in hepatic and pulmonary systems compared to wild type and other single mutant strains (Tufariello*, et al.*, 2006). The observation indicates the role of RpfB in virulence.

The absence of significant growth/virulence phenotypes led to the hypothesis that the individual *rpf* deletions probably have complete or partially redundant functions (Tufariello*, et al.*, 2004). This notion was further supported by gene expression analyses which revealed differential and coincidental expression patterns of all *rpf* genes in wild type, *in vitro,* and possible compensatory

increases in gene expression in mutant strains (Downing*, et al.*, 2004, Tufariello*, et al.*, 2004). In wild type *M. tuberculosis*, the highest and lowest expression levels are seen for *rpfC* and *rpfE* respectively (Downing*, et al.*, 2004, Tufariello*, et al.*, 2004) and it has also been shown that *rpfB* is relatively up-regulated in a persistence rabbit model (Kesavan*, et al.*, 2009). Expression of all the *rpf* genes in a clinical strain background is also evident *in vivo* during acute infection in mice, although expression of *rpfC* was not detected in another independent study (Tufariello*, et al.*, 2004, Tufariello*, et al.*, 2006). Another study revealed that *M. tuberculosis rpf* expression has relevance to human infection, although the expression of specific *rpf* genes is not established (Davies*, et al.*, 2008). It is likely that each Rpf has a distinct role during different stages of infection (Tufariello*, et al.*, 2004). With regards to expression levels in mutant strains, in the background of the Δ*rpfC* mutant, a sizeable amount of genes are observed to be differentially expressed in contrast to other *rpf* mutants. Sress response genes, including heat shock, and NADH dehydrogenase genes are up- and down-regulated respectively in the Δ*rpfC* mutant and this is proposed to be due to alterations in cell wall metabolism (Downing*, et al.*, 2004).

In light of the fact that all of the *rpf* genes can be individually depleted from the *M. tuberculosis* genome without impairment or abrogation of growth *in vitro* and *in vivo* due, possibly, to functional redundancy (Downing*, et al.*, 2004, Tufariello*, et al.*, 2004); multiple *rpf* genes have been serially deleted in *M. tuberculosis* (Downing*, et al.*, 2005, Biketov*, et al.*, 2007, Kana*, et al.*, 2008, Russell-Goldman*, et al.*, 2008) in efforts to decipher the biological roles of the gene products.

Of all the double mutant combinations, the Δ*rpfA* Δ*rpfB* mutant displays severe phenotypes (Russell-Goldman*, et al.*, 2008). Colony morphology is modified with reduced serpentine cording at the surface of the colony with the periphery being less dense than the central area thus creating a halo appearance (Russell-Goldman*, et al.*, 2008). All double mutant strains that lack *rpf*B gene, Δ*rpfA* Δ*rpfB* and Δ*rpfB* Δ*rpfD*, display a retarded reactivation phenotype in a mouse model

following immune suppression in chronic infection as initially observed in the single mutant Δ*rpfB* (Tufariello*, et al.*, 2006, Biketov*, et al.*, 2007, Russell-Goldman*, et al.*, 2008). The defect is even more pronounced in the Δ*rpfA* Δ*rpfB* mutant; furthermore the strain lacks a persistence phenotype as measured by the lower bacterial loads in the organs of mice during chronic infection in contrast to those observed for wild type (Russell-Goldman*, et al.*, 2008).

Two triple mutants constructed in the Δ*rpfA* Δ*rpfC* genetic background, with additional deletion of *rpfB* and *rpfD* genes have been assessed for growth and survival in stationary phase *in vitro* (Downing*, et al.*, 2005). The mutants, Δ*rpfA* Δ*rpfC* Δ*rpfB* and Δ*rpfA* Δ*rpfC* Δ*rpfD,* survive in stationary phase and growth is comparable to wild type (Downing*, et al.*, 2005). Moreover; the multiple mutants display a defect in spontaneous resuscitation from a non-culturable state in broth medium, unlike the wild type and single *rpfB* mutant (Downing*, et al.*, 2005). Furthermore, growth of the triple mutants was attenuated in a mouse model where lower lung bacillary loads were observed as compared to wild type. The multiple mutants also displayed an inability to persist in the spleens of the mice, unlike wild type and single mutant cells thus indicating reduced virulence associated with multiple *rpf* gene deletion (Downing*, et al.*, 2005).

Deletion of four *rpf*-like genes in different combinations does not alter the *in vitro* growth phenotype observed in the triple mutants (Kana*, et al.*, 2008). Contrary to what is seen in liquid medium, a delayed growth phenotype is observed for a quadruple mutant that retains *rpf*D (Δ*rpfA* Δ*rpfC* Δ*rpfB* Δ*rpfE*) on solid media. Two other quadruple mutants; one that retains *rpfE* (Δ*rpfA* Δ*rpfC* Δ*rpfB* Δ*rpfD*), and another containing *rpfB* (Δ*rpfA* Δ*rpfC* Δ*rpfD* Δ*rpfE*) lack the abovementioned colony formation phenotype on solid media. Returning a functional copy of *rpfE* in Δ*rpfA* Δ*rpfC* Δ*rpfB* Δ*rpfE* reverses the phenotype to normal, wild type-like colony formation (Kana*, et al.*, 2008). Moreover, quadruple mutants lacking these genes, *rpfE* and *rpfB*, are highly sensitive to sodium dodecyl sulfate (SDS), a detergent (Kana*, et al.*, 2008).

An inverse relationship between gradual depletion of *rpf*-like genes and gene expression is observed with multiple mutants. There seems to be a decline in the expression of the remaining *rpf*-like genes in triple and quadruple mutants due to progressive deletion of these genes (Kana*, et al.*, 2008). Growth of the Δ*rpfA* Δ*rpfC* Δ*rpfB* Δ*rpfE* mutant was attenuated in the lungs in a mouse model and reduced growth was observed for the Δ*rpfA* Δ*rpfC* Δ*rpfB* Δ*rpfD* quadruple mutant, to levels significantly lower to those of wild type (H37Rv) (Kana*, et al.*, 2008). Dissemination of the Δ*rpfA* Δ*rpfC* Δ*rpfB* Δ*rpfD* and Δ*rpfA* Δ*rpfC* Δ*rpfB* Δ*rpfE* mutants to other organs, the spleen, was impaired as evidenced by low bacterial loads for the former and undetectable levels for the latter. Moreover, 90% of mice infected with wild type cells succumbed to the disease while all mice survived infection with the quadruple mutants (Kana*, et al.*, 2008).

The observation that all five *rpf*-like genes can be deleted from the genome of *M. tuberculosis* indicates that none of these genes are essential for viability (Kana*, et al.*, 2008). The *in vitro* phenotypes of the quintuple mutant (Δ*rpfA* Δ*rpfC* Δ*rpfB* Δ*rpfE* Δ*rpfD*) were comparable to those of its quadruple mutant (Δ*rpfA* Δ*rpfC* Δ*rpfB* Δ*rpfE*) progenitor. Similarly, complementation with *rpfE* and *rpfB* and not *rpfC* and *rpfD* genes in the Δ*rpfA* Δ*rpfC* Δ*rpfB* Δ*rpfE* Δ*rpfD* mutant restored the plating defect as observed in the quadruple mutant (Kana*, et al.*, 2008). Surprisingly deletion of all *rpf* genes from *M. tuberculosis* does not result in increased susceptibility to most drugs tested; however, the mutant does show increased susceptibility to erythromycin and vancomycin (Kana*, et al.*, 2010).
A certain level of functional organization exists within the Rpfs where RpfB and RpfE seem to play the most important functions (Kana*, et al.*, 2008)

## 1.5 Aims of this study

To further study the biological function of Rpf-like proteins by assessing the effects of *rpf* gene deletion in *M. smegmatis*. *M. smegmatis* is genetically closely related to *M. tuberculosis* but is non-pathogenic and has a faster dividing time, making it a more tractable organism to study Rpf function in mycobacteria. However, unlike *M. tuberculosis*, *M. smegmatis* encodes four (as opposed to five) *rpf*-like genes, thus making it a simpler system to assess *rpf*-like gene function in bacterial growth, cell wall remodeling and dormancy. The specific aims are outlined below.

1. Bioinformatics analysis of Rpfs in *M. smegmatis*.
2. Construction of mutants defective for one or more *rpf* genes.
3. Evaluate the effect of *rpf* gene depletion on growth and survival of the organism.
4. Evaluate the effect of *rpf* gene loss on susceptibility to different cell wall stresses.
5. Assess the role of Rpfs on cell morphology.

.

.

## 2. Materials and Methods

### 2.1 Bacterial strains and growth conditions

All strains and plasmids used in this study are listed below in Table 1. *Escherichia coli* strains were stored at -70˚C in 66% glycerol and *M. smegmatis* strains were stored at -80˚C in growth medium (details below).

*E. coli* DH5α cells, and their transformants, were grown at 37˚C in Luria Bertani (LB) broth or agar (LA) containing appropriate selection antibiotics (ampicillin 100 µg/ml, hygromycin 200 µg/ml or kanamycin 50 µg/ml) for all plasmid DNA manipulation purposes. *M. smegmatis* $mc^2155$ was grown at 37˚C in Middlebrook 7H10 agar or 7H9 broth supplemented with 0.2% glucose, 0.85% salt and 0.05% Tween or 0.02% Tyloxapol. Hygromycin (Hyg) 50 µg/ml and/or kanamycin (Km) 25 µg/ml were added where appropriate. *M. smegmatis rpf* deletion mutant strains were grown as wild type, with the exception of the complemented derivative which required addition of Hyg to a final concentration of 50 µg/ml for maintenance of the complementing vector (see Appendix C, Figure A2).

### 2.2 General Laboratory Techniques

#### 2.2.1 DNA isolation

##### 2.2.1.1 Isolation of genomic DNA from *M. smegmatis*

In all cases, freezer stocks for the relevant strains were thawed and 100 µl of liquid culture was removed and placed on 7H10 solid media. This was incubated for 3 days. The cells were scraped off the plate with an inoculating loop and re-suspended in 500 µl TE buffer. The cell suspension was heat killed at 70˚C for 35 min followed by brief cooling on ice.

**Table 1:** Strains and plasmids used in this study

| **Name** | **Description** | **Source** |
|---|---|---|
| **Strains** | | |
| *M. smegmatis* $mc^2155$ | *M. smegmatis* ATCC 607 high frequency transformation mutant | (Snapper*, et al.*, 1990) |
| *E. coli* DH5α | *supE44 ΔlacU169 hsdR17 recA1 endA1 gyrA96 thi-1 relA1* | Promega, Madison, WI |
| Δ*rpfA* | Derivative of *M. smegmatis* $mc^2155$, carrying an in-frame deletion in *rpfA* | This work |
| Δ*rpfB* | Derivative of *M. smegmatis* $mc^2155$ carrying an in-frame deletion in *rpfB* | This work |
| Δ*rpfA* Δ*rpfB* | Derivative of *M. smegmatis* Δ*rpfA* carrying an in-frame deletion in *rpfB* | This work |
| Δ*rpfA* Δ*rpfB* Δ*rpfC* | Derivative of *M. smegmatis* Δ*rpfA* Δ*rpfB* carrying an in-frame deletion in *rpfC* | This work |
| Δ*rpfA* Δ*rpfB::*pMVRPFAB | Derivative of *M. smegmatis* Δ*rpfA* Δ*rpfB* carrying pMRPFAB integrated at the *attB* phage attachment site | This work |
| **Plasmids** | | |
| pGEM3Z(+)f | *E. coli* cloning vector; $Amp^R$ | Promega |
| pGEM-T Easy | *E. coli* PCR TA cloning vector; $Amp^R$ | Promega |

| | | |
|---|---|---|
| pBluescript | *E. coli* cloning vector; $Amp^R$ | Promega |
| P2NIL | *E. coli* cloning vector; $Km^R$ | (Parish & Stoker, 2000) |
| pGOAL19 | Plasmid consisting of a *Pac*I cassette carrying *hyg*$^R$, *lacZ* and *SacB* genes; $Amp^R$ | (Parish & Stoker, 2000) |
| pOLYG | *E. coli-Mycobacterium* shuttle vector; $Hyg^R$ | (Ã“ Gaora*, et al.*, 1997) |
| pGUΔ*rpfA* | pGEM3Z(+)f cloning vector carrying a homologous region upstream of *rpfA* which includes 46 bp of the 5' end of *rpfA* and a further 853 bp of upstream sequence encoding for *moaA* and *MSMEG*_5699 genes; $Amp^R$ | This work |
| pGDΔ*rpfA* | pGEM3Z(+)f cloning vector carrying a homologous region downstream of *rpfA* which includes 56 bp of the 3' end of *rpfA* and a further 832 bp of downstream sequence encoding for *moaE* and *MSMEG*_5702 genes; $Amp^R$ | This work |
| pBUΔ*rpfB* | pBluescript cloning vector carrying a homologous region upstream of *rpfB* which includes 49 bp of the 5' end of *rpfB* and a further 772 bp of upstream sequence encoding for *ksgA* gene; $Amp^R$ | This work |
| pBDΔ*rpfB* | pBluescript cloning vector carrying a homologous region downstream of *rpfB* which includes 4 bp of the 3' end of *rpfB* and a further 693 bp of downstream sequence encoding for | This work |

| | | |
|---|---|---|
| | *MSMEG*_5440 gene; $Amp^R$ | |
| pGΔRPFA | Derivative of pGEM3Z(+)f carrying the in-frame *rpfA* deletion allele created by the ligation of the abovementioned upstream and downstream regions; $Amp^R$ | This work |
| pGΔRPFB | Derivative of pGEM3Z(+)f carrying the in-frame *rpfB* deletion allele created by the ligation of the upstream and downstream regions; $Amp^R$ | This work |
| pNΔRPFA | Derivative of p2NIL carrying the in-frame Δ*rpfA* deletion allele from pGΔRPFA, intermediate knockout construct; $Km^R$ | This work |
| pNΔRPFB | Derivative of p2NIL carrying the in-frame Δ*rpfB* deletion allele from pGΔRPFB, intermediate knockout construct; $Km^R$ | This work |
| pSVΔRPFA | Derivative of pNΔRPFA carrying the *lacZ*, *sacB* and *hyg*$^R$ marker gene cassette from pGOAL19, final knockout construct; $Hyg^R$, $Km^R$ | This work |
| pSVΔRPFB | Derivative of pNΔRPFB carrying the *lacZ*, *sacB* and *hyg*$^R$ marker gene cassette from pGOAL19, final knockout construct; $Hyg^R$, $Km^R$ | This work |
| pSVΔRPFCE | Knockout construct for deletion of *rpfC* and *rpfE*, carrying the *lacZ* and *sacB* marker gene cassette from pGOAL17, $Km^R$ | E. Machowski |
| pMRPFAB | Derivative of pMV306H carrying *rpfA* and *rpfB*; $Hyg^R$ | E. Machowski |

$Amp^R$, $Hyg^R$, $Km^R$: ampicillin-, hygromycin-, kanamycin resistance.

The heat-killed cells were treated with 50 µl lysozyme (10 mg/ml) and incubated at 37˚C for an hr; thereafter 70 µl of 10% SDS and 6 µl proteinase K (10 mg/ml) were added. This was further incubated at 65˚C for 1.5 - 2 hrs. Following incubation; 100 µl of 5 M NaCl was added, mixed and 80 µl CTAB/NaCl solution (10% CTAB in 0.7 M NaCl) was added to the mixture followed by gentle, thorough mixing through careful inversion of the tubes. The mixture was incubated at 65˚C for 10 min which was followed by the addition of 750 µl of chloroform: isoamyl alcohol (24:1 v/v). The aqueous and organic phases were mixed by inversion and separated by centrifugation at 13000 rpm for 5 min. The aqueous layer was collected into sterile Eppendorf tubes and DNA was precipitated by addition of 450 µl isopropanol and incubation on ice for 30 min. DNA was harvested by centrifugation for 20 min at 13000 rpm. The pellet was washed with ice cold 70% ethanol and dried under a vacuum. The pellet was re-constituted in 50 -100 µl sterile distilled water, depending on the size of the pellet.

#### 2.2.1.2 *E. coli* plasmid DNA extraction (Mini-prep)

Overnight (1.5 ml) cultures were briefly centrifuged (13000 rpm) and the cell pellets were re-suspended in 100 µl cold solution I (0.05 M glucose, 0.01 M EDTA, 0.025 M Tris-HCl, distilled $H_20$). A volume of 200 µl room-temperature solution II (0.2 M NaOH, 1% SDS, distilled $H_20$) was added to the cell suspension, followed by gentle mixing and incubation on ice for 5 min. A further addition of 150 µl chilled solution III (3 M potassium acetate, 11.5% glacial acetic acid, distilled $H_20$) was made to the suspension followed by mixing and incubation on ice for 10 min. The suspension was then centrifuged at 13000 rpm for 10 min and supernatants were collected into sterile Eppendorf tubes containing 1 µg/ml RNAseA. RNA was digested at 42˚C for 30-60 min. DNA was precipitated by the addition of an equal volume (~ 450 µl) of isopropanol and the pellets were washed with 70% ethanol, dried under vacuum at 45˚C for 20 min and re-suspended in 10-20 µl sd$H_20$.

#### 2.2.1.3 *E. coli* plasmid DNA purification (Maxi-prep)

100 ml cultures grown overnight were harvested by centrifugation (4500 rpm) for 10 min at 4˚C. Purified DNA was extracted with the use of the NucleoBond kit (Macherey-Nagel) in accordance with manufacturer's instructions.

### 2.2.2 DNA manipulation

#### 2.2.2.1 Restriction endonuclease digestion

All restriction endonucleases, and their buffers, used were from Roche Diagnostics (Mannheim, Germany) or New England Biolabs. The total digestion reaction volume varied, depending on the concentration or amount of DNA used and application. Concentration, as quantified by UV spectroscopy (absorbance at 260 nm on the Nanodrop spectrophotomer – NanoDrop Technologies see section 2.2.6), of plasmid DNA used was within the range of 1-2 µg in a total reaction volume of up to 30 µl. A one-tenth volume of buffer and 1 µl restriction enzyme were added to the reaction with sterile distilled water making up the rest of the total volume. For restriction digestions with two enzymes, both enzymes were added in a single reaction with a compatible restriction buffer. Reactions were incubated at room temperature, 30˚C or 37˚C for 1 hr as per manufacturer's instructions for individual enzymes.

#### 2.2.2.2 Dephosphorylation of plasmid DNA

Antarctic phosphatase (New England BioLabs) was employed for removal of 5' phosphate groups from DNA to prevent re-circularization of cloning vectors. A one-tenth volume of 10x Antarctic phosphatase buffer was added to heat-inactivated restriction digestion reaction, mixed and briefly centrifuged. Antarctic phosphatase enzyme (1 µl) was added, mixed thoroughly and incubated at 37˚C

for 1 hr. Subsequent to incubation; the DNA was visualized on agarose gels (section 2.2.4) and purified by gel extraction (section 2.2.5) to ensure that linear vector, free of any phosphatase, was used for subsequent cloning steps.

### 2.2.2.3 Ligations

All ligations were carried out with the use of a Fast-Link DNA Ligation Kit (Epicentre Biotechnologies; Madison, Wisconsin), according to instructions supplied in the manufacturers' manual. Plasmid vector and insert were prepared by digestion (section 2.2.2.1), gel purification (2.2.5) and quantification (2.2.6). A standard of 50 ηg of vector was used for all ligation reactions. The molar amount of insert DNA to be added to 50 ηg of vector DNA to obtain equivalent numbers of vector and insert molecules in the ligation reaction was calculated as follows:

$$m = \frac{50ng \times bpI}{bpV} \qquad Vol_{insert} = \frac{mass}{concentration}$$

m – amount of insert DNA for 50 ηg of vector, I- insert size, V- vector size, bp- size of insert or vector in base pairs.

The calculated insert volume was added to the ligation reaction for every 50 ng of vector in different Vector: Insert ratios, with an increment of insert added for each ratio. The standard ratios used for all ligations were 1:1 Vector: Insert, 1:2 Vector: Insert and 2:1 Vector: Insert. The reaction was incubated at 25˚C for 1 hr and heat inactivated at 65˚C for 10 min.

### 2.2.3 DNA precipitation

For ethanol salt precipitation of DNA, one-tenth volume of 3 M Sodium acetate (pH 5.2) was added to DNA and briefly inverted. This was followed by the addition of 2.5 volume ice-cold 100% ethanol and mixing by inversion. The

mixture was incubated at -20˚C for 20 min and the DNA pellet was harvested by centrifugation at 13000 rpm for 20 min. The supernatant was decanted and the DNA pellet was washed with ice-cold 70% ethanol. DNA was dried under vacuum and re-constituted in sterile distilled water.

### 2.2.4 Gel electrophoresis

All DNA samples were separated and viewed in 1% agarose gels and these were visualized with the use of a gel fluorescence imaging system; G:BOX and GeneSnap image acquisition software (Syngene). The gels were prepared by dissolving agarose powder (Invitrogen, Life Technologies), by heating to boiling point, in 1X TAE buffer (40 mM Tris-acetate, 1 mM EDTA) to a final concentration of 1%. For every 30 ml of dissolved agarose suspension, 2 µl of ethidium bromide (10 mg/ml) was added to the cooled agarose solution which was then poured into gel trays, consisting of gel combs, balanced on gel casters. Once set; the gels were placed in electrophoresis tanks (Biorad Laboratories) submerged in 1X TAE buffer and the DNA samples, combined with gel loading dye (0.04 g bromophenol blue, 30% glycerol, 70% sd$H_20$), were loaded in wells. The gels were generally run at 0.4 A, 90 V but in other instances the voltage was lower, depending on application. Molecular weight markers (Roche Diagnostics; Mannheim, Germany) were run simultaneously with DNA samples to aid in determining the size of DNA fragments.

### 2.2.5 Gel extraction and purification

Following gel electrophoresis, the gel was placed on a UV transilluminator (Fotodyne Incorporated) and DNA was visualized with UV light. DNA fragments of interest were excised out of the gel using a sterile scalpel. The mass of the gel slice was weighed and the DNA was extracted from the gel with the use of a PCR clean-up Gel extraction kit (NucleoSpin Extract II, Macherey-Nagel), in

accordance to instructions supplied by manufacturers' manual. The DNA was quantified as described below.

### 2.2.6 DNA quantification

Following DNA and/or gel extraction, the DNA was quantified by NanoDrop ND-1000 Spectrophotometer (NanoDrop Technologies) used in conjunction with software (Coleman Technologies) provided by the manufacturer. In other instances, DNA was quantified by estimation of DNA amounts on gels. For this, gel electrophoresis (section 2.2.4) was conducted for the DNA of interest and a comparison was made between the intensity of bands to that of similar size bands in the molecular weight marker, the concentration of which are provided by the manufacturer.

### 2.2.7 Preparation of *E. coli* DH5α competent cells using rubidium chloride

*E. coli* DH5α cells were grown from a pre-culture in 100 ml Psi broth at 37˚C, with aeration, to $OD_{600nm}$ ~ 0.48 - 0.52. The culture was placed on ice for 15 min and the bacterial cells were harvested by centrifugation (3500 rpm, Allegra X-22R, Beckman Coulter) for 5 min. The supernatant was discarded, the cell pellet was re-suspended in 0.4 volume TfbI (15% v/v glycerol, 30 mM potassium acetate, 100 mM rubidium chloride, 10 mM calcium chloride, 50 mM manganese chloride) and the cell suspension was incubated on ice for 15 min. Centrifugation was repeated and the pellet was re-suspended in 0.04 volume TfbII (10 mM MOPS, 75 mM calcium chloride) with subsequent incubation on ice for another 15 min. The cells were used immediately or frozen in ethanol at -70˚C for later use. In all cases, the transformation efficiency was determined to be between 1 $X10^8$ -1 $X10^9$ CFU/µg DNA.

#### 2.2.7.1 Transformation of *E. coli* competent cells

The competent *E. coli* DH5α cells were thawed on ice and 4 µl of ligation reaction was added in 100 µl aliquots of cells. This was incubated on ice for 25 min, followed by heat shocking for 90 sec at 42˚C. The DNA-cell suspension was returned to ice for 2 min and thereafter, rescued with 800 µl LB. A subsequent incubation period was carried out for 45-60 min at 37˚C. Following incubation, the cells were centrifuged (5000 rpm) for 10 min and re-suspended in 100 µl fresh LB. The cell suspension was spread on appropriate antibiotic-containing solid media and incubated overnight at 37˚C.

### 2.2.8 PCR screening of *E. coli* transformants

*E. coli* clones selected on appropriate antibiotic-containing solid media were picked and re-suspended in 20 µl broth. The cell suspension (5 µl) was used as template DNA in a Taq PCR (Faststart Taq DNA Polymerase, dNTPack kit, Roche Diagnostics) with the appropriate primers and cycling conditions. The PCR was carried out as stipulated in the manufacturers' manual. The remaining cell suspension was used to start cultures for any positive clones.

### 2.2.9 Serial dilutions

A series of ten-fold dilutions were performed by adding 100 µl undiluted culture, $10^0$, to an Eppendorf tube containing 900 µl broth followed by mixing with gentle vortexing. Thereafter, 100 µl of the 10X diluted culture, $10^1$, was then added (100 µl) to the next tube containing 900 µl broth and mixed as before. This procedure was repeated to obtain the entire dilution series, usually up to $10^7$ for most applications.

### 2.2.10 DNA sequencing

The constructs consisting of upstream and downstream *rpf* PCR fragments cloned separately to high-copy plasmid vectors (pBluescript and pGEM) were sent for sequencing at Inqaba Biotec (Pretoria, RSA). The purified plasmids were sequenced with the use of M13 universal primers and the sequencing data was analyzed with the SeqMan module from DNASTAR program.

## 2.3 Bioinformatics analysis of *rpf* genes in *M. smegmatis*

The bioinformatics program ACT (Artemis Comparison Tool) was used to compare the *M. tuberculosis rpf* genes with their *M. smegmatis* homologues within their relative genomic context. A linear, genomic map of the region upstream and downstream of each *rpf*-like gene in *M. smegmatis* was created using the abovementioned tool.

*M. tuberculosis* and *M. smegmatis rpf* protein sequences were taken from Artemis or NCBI, http://www.ncbi.nlm.nih.gov/genomes//proks.cgi, and the overall homology between the *rpf* gene products in both organisms was analyzed using BLAST http://blast.ncbi.nlm.nih.gov/ or http://blast.jcvi.org/cmr-blast/. The prediction of signal peptides within Rpf amino acid sequences was carried out by submission of Rpf protein sequences to SignalP 4.0 server, http://www.cbs.dtu.dk/services/signalp.

## 2.4 Construction of *rpf*-knockout mutants

The method of choice for gene deletion was allelic exchange using homologous recombination and a suicide plasmid (Parish & Stoker, 2000). For this method homologous regions located upstream and downstream of the gene of interest are required. These regions are then fused by cloning into the suicide vector to create a deletion allele. For this, primers were designed to amplify the regions

upstream and downstream of *rpfA* and *rpfB*. These primers were designed with engineered restriction sites to facilitate cloning into the suicide vector to produce the deletion allele, Table 2. The primers were designed with the use of an online program, Primer 3 http://frodo.wi.mit.edu/primer3/. *Asp*718 and *Bgl*II recognition sequences were included in forward and reverse primers respectively for all upstream region amplicons. Similarly, the forward and reverse primers for all downstream regions contained *Bgl*II and *Hind*III recognition sequences respectively. As mentioned above, these restriction enzyme recognition sequences were engineered for cloning purposes and were included in the generation of knockout constructs for *MSMEG*_5700 (*rpfA*) and *MSMEG*_5439 (*rpfB*).

### 2.4.1 Polymerase Chain Reaction (PCR)

In order to generate upstream and downstream *rpf* DNA fragments, PCR was carried out with *rpfA*/*rpfB* forward and reverse primers shown in Table 2. The cycling conditions were performed as follows: the first step, denaturation, proceeded at 95˚C for 5 min, followed by a further 35 cycles of denaturation at 95˚C for 30 sec, an annealing step at 60˚C for 30 sec and elongation at 72˚C for 150 sec. A final step of elongation at 72˚C for 10 min was included and the reaction concluded with a cooling and holding temperature step set at 16˚C.

The PCR was initially carried out using Faststart Taq DNA Polymerase, dNTPack kit (Roche Diagnostics; Mannheim, Germany) according to manufacturer's instructions for optimization PCR cycling conditions. Following this optimization step, Phusion High-Fidelity DNA Polymerase (Finnzymes) and Expand High-Fidelity Polymerase (Roche Diagnostics) PCR was performed according to manufacturer's instructions. In the case of Phusion PCR, the annealing temperature was changed to 62˚C, to reduce non-specific amplification, while the remaining reaction conditions remained unaltered. Genomic DNA from *M.*

*smegmatis* $mc^2155$ (section 2.2.1.1) was used as a template in these PCR reactions.

### 2.4.2 Construction of knockout vectors for *rpfA* and *rpfB*

#### 2.4.2.1 Cloning of upstream and downstream homologous regions

The Phusion PCR products corresponding to the *rpfA* and *rpfB* upstream and downstream homologous regions were separated and viewed on gels by agarose gel electrophoresis (section 2.2.4). These fragments were then purified by gel extraction as previously described (section 2.2.5). A high-copy plasmid vector, pBluescript was prepared by plasmid extraction (section 2.2.1.3) and linearized by digestion (section 2.2.2.1) with *Sma*I. The linearized plasmid vector was viewed on agarose gels subsequent to gel electrophoresis and purified by gel extraction. The purified linear plasmid and PCR amplicons were quantified (section 2.2.6) and the *rpfB* upstream and downstream regions were cloned separately into pBluescript to yield pBUΔ*rpfB* and pBDΔ*rpfB* respectively. Similarly, the *rpfA* upstream and downstream regions were cloned separately by ligation (section 2.2.2.3) into pGEM3Z(+)f with the use of a pGEM-T Easy Vector system kit (Promega, Madison) according to instructions supplied by the manufacturer. Ligation reactions were transformed (section 2.2.7.1) into *E. coli* DH5α cells followed by selection of white, $Amp^R$ clones on LA plates containing 100 µg/ml ampicillin (Amp) and 2% 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal). These clones were picked and screened by PCR to confirm the correct cloned insert.

**Table 2:** Oligonucleotide sequences used in the construction of *rpf* knockout strains in *M. smegmatis*. Bold bases represent engineered restriction enzyme recognition sequences for cloning purposes.

| Gene | Mutant allele | Internal deletion size | Knockout vector used | Oligonucleotide pairs used in the construction of vectors | |
|---|---|---|---|---|---|
| | | | | 5'-3' sequence | Region amplified |
| *rpfA* | Δ*rpfA* | 1359 bp | pSVΔRPFA | GGGG**GGTACC**GATCTGCGGCACCTGCTG<br>GGGG**AGATCT**TGGCAGATGACGAAGAAG | 46 bp of the 5' end of *rpfA* and 853 bp of homologous upstream sequence. |
| | | | | GGGG**AGATCT**GCCCCGGTGGGCGCACCGG<br>GGGG**AAGCTT**CCCGACGGCGACGCCGTGG | 56 bp of 3' end of *rpfA* and 832 bp of homologous downstream sequence. |
| *rpfB* | Δ*rpfB* | 1075 bp | pSVΔRPFB | GGGG**GGTACC**TGGACGCGGCACATTGGG<br>GGGG**AGATCT**CGAAGCAACGGCGAGCGGG | 49 bp of 5' end of *rpfB* and 772 bp of homologous upstream sequence. |
| | | | | GGGG**AGATCT**GACTATTCGACTGCTCGGG<br>GGGG**AAGCTT**GCGGCGCTGCGCGAAGCC | 4 bp of the 3' end of *rpfB* and 693 bp of downstream sequence. |

Small scale plasmid DNA extraction (section 2.2.1.2) was carried out from positive clones and the resulting DNA was digested with enzymes to release cloned upstream and downstream inserts. *Asp*718 and *Bgl*II enzymes were used to release the upstream fragment, while *Bgl*II and *Hind*III enzymes were used to release the downstream fragment; the digested constructs were viewed on agarose gels following gel electrophoresis. The correct clones were picked and grown in LB containing 100 µg/ml Amp and a large scale plasmid DNA extraction was performed (section 2.2.1.3). The purified constructs were sent for sequencing (section 2.2.10) to confirm that no mutations were inadvertently introduced during PCR amplificaiton.

### 2.4.2.2 Three-way cloning

The upstream and downstream *rpfA* and *rpfB* fragments were digested with appropriate enzymes from the high-copy plasmid vectors and simultaneously ligated to purified p2NIL (Parish & Stoker, 2000) plasmid vector by three-way cloning (illustrated in section 3.2.1.3, Figure 15). The p2NIL vector was prepared by DNA extraction from *E. coli* DH5α cells and digested with *Asp*718 and *Hind*III. White, $Km^R$ clones were selected on 50 µg/ml kanamycin-containing LA, following *E. coli* transformation. The clones were screened by PCR (section 2.2.8) and restriction analysis of extracted plasmid DNA, from putative correct clones was carried out. Three-way cloning was also carried out with pGEM3ZF(+) as the vector backbone using the same restriction enzymes and cloning strategy.

### 2.4.2.3 Cloning of *Pac*I cassette

The p2NIL intermediate knockout constructs (pNΔRPFA and pNΔRPFB) and pGOAL19 vector were digested with *Pac*I restriction enzyme and the relevant fragments were purified from gels for cloning. The *Pac*I cassette (containing the

selectable – *lacZ* – and counter selectable – *sacB* – markers) from pGOAL19 was cloned into linearized, dephosphorylated (section 2.2.2.2) pNΔRPFA and pNΔRPFB constructs. Ligation reactions were transformed into *E. coli* followed by selection of blue, $Km^R$ clones on LA containing 50 µg/ml Km and X-gal at 30˚C. This step of cloning completed the construction of the final suicide delivery vectors. The clones were grown for 3 days in broth containing Km at 30˚C and tested for sucrose sensitivity on 5% sucrose. The cultures were serially diluted (section 2.2.9) from $10^1$ to $10^7$, plated on LA plates containing 50 µg/ml Km, with and without 5% sucrose and incubated at 30˚C for 3 days. Putative positive clones were further analyzed by extensive restriction analysis.

### 2.4.3 *M. smegmatis* $mc^2155$ competent cells preparation

For transformation of *M. smegmatis* ($mc^2155$ or mutant strains), a pre-culture was grown overnight from freezer stocks to stationary phase in 7H9 broth at 37˚C, with 100 rpm rotary shaking. This culture was then diluted 1:100 in fresh media and grown overnight at 37˚C to O.D. 0.8. The cells were harvested by centrifugation (4500 rpm) for 10 min and the cell pellets were re-suspended in 30 ml 10% ice-cold glycerol, cells were kept on ice at all times. Centrifugation was carried out again and the cells were repeatedly washed with 10% glycerol four times. The cells were kept on ice, in glycerol suspension, until they were electroporated with the relevant constructs.

#### 2.4.3.1 Electroporations

Electro-competent cells were added to Gene Pulser 2 mm cuvettes (Bio-Rad Laboratories) in 400 µl aliquots followed by addition of plasmid DNA (knockout constructs/suicide vectors, complementing vectors and pOLYG) at varying concentrations (1-5 µg) with gentle mixing to prevent the formation of air bubbles. The Gene Pulser Xcell (Bio-Rad Laboratories) was used for electroporations with

the following settings: the voltage was set at 2500 V, capacitance at 25 µF, resistance at 1000 Ω and cuvette size at 2 mm. Cuvettes were placed into the shock pod and the cells were electroporated for delivery of plasmid DNA into cells. The cells were rescued immediately with 800 µl 7H9 broth with brief incubation on ice followed by incubation at 37˚C for 3 hrs or overnight. Blue $Hyg^R$, $Km^R$ *M. smegmatis* clones were selected on 7H10 solid media containing X-gal, 50 µg/ml Hyg and 20 µg/ml Km or 50 µg/ml Hyg only for serially diluted pOLYG control. A 'No DNA' control was included in the electroporations.

Blue colonies selected from these plates were grown overnight in 5 ml 7H9 broth containing 20 µg/ml Km at 37˚C; thereafter 200 µl of the above culture was used as inoculum in 10 ml of 7H9 broth without antibiotic. Cells from the 10 ml culture were harvested by centrifugation (4500 rpm) for 10 min and the pellet was re-suspended in 400 µl broth. A 100 µl fraction of the cell suspension was spread on 7H10 plates containing 2% sucrose and this was done in duplicate while another was spread on solid media without sucrose. The remaining 100 µl fraction was serially diluted from $10^1$ to $10^7$. A 100 µl volume of each dilution was spread on solid media containing X-gal, 2% sucrose and without sucrose. The plates were incubated at 37˚C for 3-5 days. White, $Suc^R$ colonies, possible *rpf* deletion mutants, were picked and screened by PCR as described below.

### 2.4.4 Screening for *M. smegmatis rpf* deletion mutants by PCR

White $Suc^R$ colonies were re-suspended in 50 µl sterile water and 10 µl of the cell suspension was used to make cell spots on 7H10 plates for future use. The rest of the cell suspension was boiled at 95˚C for 5 min, 40 µl phenol chloroform was added and briefly vortexed. The contents were centrifuged (13000 rpm) for 5 min and the aqueous layer was removed and placed into a sterile Eppendorf tube. The extracted genomic DNA was used as template DNA in the Taq PCR reaction using the primers described in Table 3 and following manufacturer's instructions.

All three allelic-specific primers, shown below in Table 3, were used in a reaction for screening each *rpf* mutant (illustrated in section 3.2.2, Figure 28 and 29). The cycling conditions were performed as follows: the denaturation step proceeded at 95˚C for 5 min and was followed by 30 cycles of further denaturation at 95˚C for 30 sec, an annealing step at 60˚C for 30 sec and extension at 72˚C for 40 sec. Final extension proceeded at 72˚C for 7 min and the reaction concluded with a cooling and holding temperature step set at 16˚C. The PCR products were viewed on 1% agarose.

### 2.4.5 Southern blot analysis

A concentration range of between 1-2 µg extracted *M. smegmatis*, wild type and *rpf* mutant genomic DNA was digested with appropriate enzymes and incubated overnight at the appropriate temperature. The fragmented genomic DNA was separated by gel electrophoresis at 80 V in 1% agarose gels (section 2.2.4). The gels were submerged in 0.2 M HCl for 15 min with continuous shaking for depurination of DNA. The acid solution was discarded and the gels were rinsed twice in distilled $H_20$ and placed in a solution of 0.5 M NaOH /1.5 M NaCl. The gels were left submerged in the solution for 30 min with continuous shaking to allow for DNA denaturation. The solution was discarded and gels were equilibrated with 1X TBE buffer (89 mM Tris-Borate, 2 mM EDTA, pH 8.3) for 5 min.

The DNA was transferred from gels to membranes by sandwiching the gels and membrane (Hybond-N, Amersham), in 1X TBE buffer, between layers of thick and thin foam sponges, and blotting paper. The “sandwiches” were encased in gel cassettes (Amersham) and transferred to a tank transfer unit (Hoefer TE 22, Amersham) containing the same buffer. DNA was transferred at 0.6 A, 0.5 V for 2 hrs at 4˚C.

**Table 3:** Allelic-specific oligonucleotide sequences used in this study for *rpf* deletion mutant DNA screening. F and R represent forward and reverse primers respectively.

| Oligonucleotide name | Sequence 5' – 3' | Region targeted |
|---|---|---|
| rpfA- F | GAGGACCTCGACGCTTATGA | 4 bp of the 5' end of *rpfA* and 296 bp of internal *rpfA* sequence or 1524 bp of internal *rpfA* sequence including 47 bp of downstream homologous sequence |
| rpfA- R1 | TCCTTGGTGGCCATGTACGC | |
| rpfA- R2 | ACCGACGAGTGGGTCAACTC | |
| | | |
| rpfB- F | AGGACCACACGACAGCCGTG | 3 bp of the 5' end of *rpfB* and 296 bp of internal *rpfB* sequence or 1240 bp of internal *rpfB* sequence including 96 bp of downstream homologous sequence |
| rpfB- R1 | CGGCTGCGGCGCAACACGAT | |
| rpfB- R2 | CATCGTGGACGAAGTTCTGT | |
| | | |
| rpfC- F | GAACGTTTGTACGGCAGTGA | 4 bp of the 5' end of *rpfC* and 297 bp of internal *rpfC* sequence or 560 bp of internal *rpfC* sequence including 82 bp of downstream homologous sequence |
| rpfC- R1 | GCCAGACCCTGGGTGCGCAG | |
| rpfC- R2 | GCGCGACCGTGCTGGTGCCC | |
| | | |
| rpfE- F | TGGCGAAAGGAACGAAGTGA | 4 bp of the 5' end of *rpfE* and 296 bp of internal *rpfE* sequence or 422 of internal *rpfE* sequence including 82 bp of downstream homologous sequence |
| rpfE- R1 | TGGGTACGCAGCACGTTCTC | |
| rpfE- R2 | GCAGAAGGGCCTGGAGAGGA | |

After electroblotting, the membranes were removed from “sandwiches” with forceps, placed on blotting paper and allowed to air-dry. Membranes were then placed in a pre-warmed UV crosslinker (UVC 500, Amersham Biosciences) with the transferred DNA at the membrane surface and the DNA was crosslinked at 120 000 μJ/cm$^2$ for 2 min. This was repeated twice.

A DIG High Prime DNA Labeling and Detection Starter Kit (Roche Diagnostics; Mannheim, Germany) was used for hybridization, labeling and detection. Probe hybridization and washes were performed in the Hybaid (Thermo Scientific) system with associated bottles. The hybridization temperature was calculated to be between 52-54˚C according to the formula given in the manufacturers’ manual. Membranes were pre-hybridized in roller bottles containing DIG Easy Hyb Granules solution (Roche Diagnostics; Mannheim, Germany) for 2 hrs in the oven by rotation. A DIG-labeled DNA probe was synthesized using a PCR DIG Probe Synthesis Kit according to instructions provided in the manufacturers’ manual. Constructs listed in Table 1 (pGDΔ*rpfA*, pBUΔ*rpfB*, pBDΔ*rpfB*, pSVΔRPFCE) were used as template DNA in the above PCR reactions and Primers listed in Table 2 were also included in the reaction. The probes were added to the membranes and hybridized overnight. The membranes were washed twice at room temperature in 25 ml solution containing 2X SSC (300 mM NaCl, 30 mM Sodium citrate) and 0.1% SDS for 5 min. This was followed by another two washes at 68˚C in 25 ml solution containing 0.5X SSC (75 mM NaCl, 7.5 mM Sodium citrate) and 0.1% SDS.

#### 2.4.5.1 Digoxigenin (DIG)-dUTP labeling and detection

Membranes were transferred to a glass dish and equilibrated with 20 ml 1X wash buffer containing 1X Maleic acid buffer (10 mM Maleic acid, 15 mM NaCl, pH 7.5) and 0.3% Triton X-100 for 5 min. The buffer was discarded and 120 ml 1X blocking buffer consisting 1X Maleic acid buffer and 1X blocking solution (DIG High Prime DNA Labeling and Detection Starter Kit, Roche Diagnostics) was

added and allowed to shake at room temperature for 30 min. The blocking buffer was discarded and another 25 ml fresh blocking buffer and 25 µl Anti-Digoxigenin- AP Fab fragments (DIG High Prime DNA Labeling and Detection Starter Kit, Roche Diagnostics) were added to membranes. This was allowed to shake at room temperature for 30 min and discarded. Another two washes in 120 ml 1X wash buffer were conducted at room temperature for 15 min. The membranes were equilibrated with 25 ml detection buffer (50 mM $Mg_2Cl$, 0.1 M Tris-HCl, 1 M NaCl, and pH 9).

Membranes were inserted in hybridization bags (Roche Diagnostics; Mannheim, Germany), and chloro-5-substituted adamantyl-1, 2-dioxetane phosphate (CSPD) substrate was added to membranes for chemiluminescent detection of probe. Probe-complementary DNA sequence hybrids were revealed on X-ray film following membrane exposure to the film and incubated at room temperature for 30 min-2 hrs.

### 2.5 Growth kinetics

Overnight pre-cultures of *M. smegmatis* wild type, Δ*rpfA*, Δ*rpfB*, Δ*rpfA* Δ*rpfB*, Δ*rpfA* Δ*rpfB* Δ*rpfC* and Δ*rpfA* Δ*rpfB*::pMVRPFAB strains were inoculated in 100 ml 7H9 broth to a starting $OD_{600nm}$ (optical density at 600nm) of 0.01. The cultures were placed in a 37˚C shaking incubator (115 rpm; IncoCool, Labotec) over a period of 33 hrs. Aliquots were taken every 3 hrs and $OD_{600nm}$ readings were determined on a Shimadzu UV1601 spectrophotometer. At these time points, colony forming units (CFUs) were enumerated by plating ten-fold serial dilutions on solid media. Plates were incubated at 37˚C for 5-7 days. These experiments were repeated three times to confirm biological reproducibility.

## 2.6 Drug susceptibility assay

Cultures of the six different strains were grown at 37˚C in a shaking incubator from overnight pre-cultures to an $OD_{600nm}$ of 0.4. The cultures were serially diluted from $10^0$ to $10^6$; and 10 µl of each dilution was spotted on detergent (sodium dodecyl sulfate - SDS) and antibiotic-containing 7H11 solid media. A no treatment control was included in the assay. Antibiotics (Sigma-Aldrich) used were 2000 µg/ml carbenicillin, 150 µg/ml D-cycloserine, 10 µg/ml erythromycin, 6 µg/ml vancomycin and SDS (Sigma-Aldrich) was used at concentrations of 0.001% and 0.005%. CFUs were also determined by plating ten-fold serial dilutions on antibiotic containing and antibiotic free media. All plates were incubated at 37˚C for 4 days. Three biological replicates were performed to ensure reproducibility of results.

## 2.7 Analysis of changes in colony morphology

*M. smegmatis* wild type, *rpf* deletion mutants and complemented strain were grown in culture to an $OD_{600nm}$ of 0.5 from an overnight pre-culture. A ten-fold dilution series was carried out for all cultures, $10^4$-$10^6$ dilutions were plated on 7H10 solid media for each strain. The plates were incubated at 37˚C for 5 days. Colony morphology differences were analyzed on a Vacutec G-Box Gel documentation/photography system. Photographs of colonies were taken at several magnifications. The experiment was performed several times to ensure reproducibility of different colony morphotypes.

## 2.8 Anaylsis of biofilm formation

Cultures were grown to an $OD_{600nm}$ ~ 0.5, as described above for most experiments, and were diluted 10X. A volume of 5 ml 7H9 broth was placed in

each well of a 6-well cell culture multidish (Nunclon Delta, Thermo Fisher Scientific). Each well was inoculated with 20 µl diluted culture of each strain and incubated at 30˚C, without shaking, for 7 days. Biofilm formation was determined by monitoring the formation of a thick pellicle of bacterial growth at the liquid-air interface as previously described (Ojha*, et al*., 2008). The experiment was performed at least three times to ensure biological reproducibility.

### 2.9 Analysis of cell morphology by scanning electron microscopy

*M. smegmatis* strains were grown to an $OD_{600nm}$ ~ 0.5, as described above, and the cells were harvested by centrifugation (4500 rpm; Allegra X-22, Beckman Coulter) for 10 min. The cell pellets were washed once in Phosphate buffer saline (PBS - 137 mM NaCl, 2.7 mM KCl, 10 mM Phosphate buffer solution, pH 7.4) by re-suspension and centrifugation. The cells were then re-suspended and fixed in the same buffer containing 2.5% glutaraldehyde overnight at 4˚C. The cells were washed again in PBS and dehydrated with increasing concentrations of ethanol for 10 min. Ethanol concentration used were 30%, 50%, 70%, 90% and 100%. A few drops of the absolute ethanol-cell suspension was placed on stubs and allowed to air-dry. The dried samples were coated with carbon and gold sputtering and analyzed with a Nova-200 FIB scanning electron microscope.

## 3. Results

### 3.1 Bioinformatics analysis of *rpf*-like genes in *M. smegmatis*

Several studies report the presence of multiple *rpf*-like genes in a variety of different bacterial and Firmicute species (Mukamolova*, et al.*, 1998, Mukamolova*, et al.*, 1999, Mukamolova*, et al.*, 2002, Ravagnani*, et al.*, 2005, Mukamolova*, et al.*, 2006), suggesting that the multiplicity of Rpf homologues may play an important role in bacterial physiology.

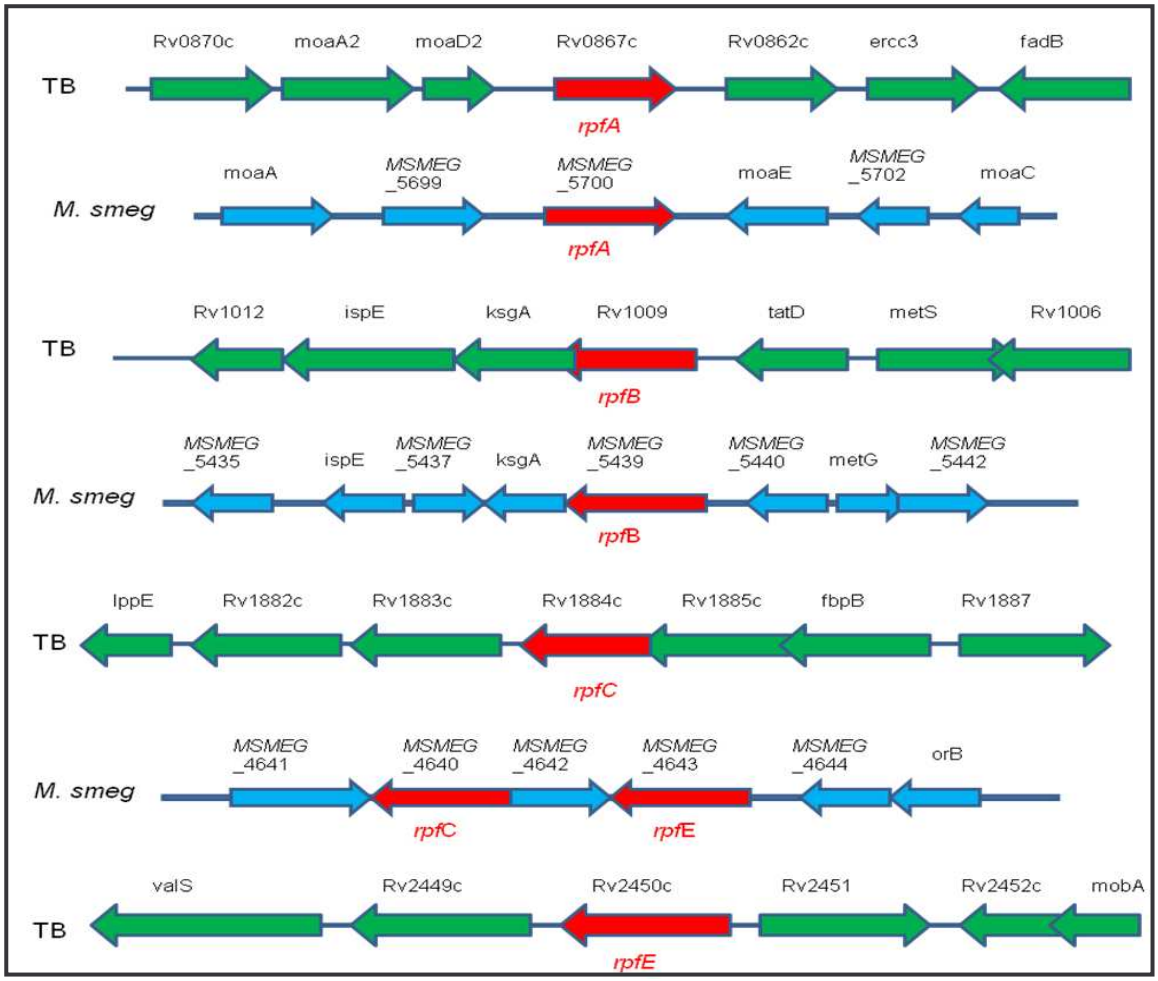



**Figure 7**: A schematic representation of the genomic organization of *rpf*-like genes in *M. smegmatis* (*M. smeg*) compared to *rpf*-like genes of *M. tuberculosis* (TB).

In this study, bioinformatics analysis was conducted to determine the number of *rpf*-like genes in *M. smegmatis* and the homology to their *M. tuberculosis* counter parts, Figure 7. Bioinformatics analysis, with the use of ACT (Artemis Comparison Tool), revealed that four *rpf* genes exist in *M. smegmatis* (Figure7) as opposed to the five identified in *M. tuberculosis* (Mukamolova*, et al.*, 1998). *M. smegmatis* lacks a homologue for *rpfD* from *M. tuberculosis*. In terms of the arrangement of genes, both *M. smegmatis* and *M. tuberculosis rpfA* and *rpfB* genes are in the same genomic context (Figure7). The *rpfA* gene in both organisms is flanked by various Molybdenum cofactor biosynthesis genes. Similarly, genes upstream and downstream of *rpfB* are comparably arranged in both organisms. However, the *rpfC* and *rpfE* genes lie in different contexts in *M. smegmatis* and *M. tuberculosis*, Figure 7. In *M. tuberculosis*, *rpfC* and *rpfE* are located at distinct, separate chromosomal loci. However, in *M. smegmatis* these genes are proximal to each other and separated by only a short fragment (566 bp) of intervening sequence that encodes a putative hypothetical protein with no known homologues.

In *M. smegmatis*; it is probable that *rpfA* is transcribed from a monocistronic mRNA message as previously noted for *M. tuberculosis* (Tufariello*, et al.*, 2004). In contrast, *rpfB* is part of a polycistronic message together with *ksgA* and *MSMEG_5440* which encode a dimethyladenosine transferase and a deoxyribonuclease respectively. This has also been demonstrated in *M. tuberculosis* (Tufariello*, et al.*, 2004). It is unlikely that the *rpfC* and *rpfE* genes are under the control of one promoter since there is an intermediate gene, MSMEG_4642, which is transcribed in the opposite direction to these *rpf* homologues. The *MSMEG_4642* open reading frame extends into *rpfC* by 145 bp, this gene is predicted to encode for a protein of unknown function. BlastP analysis of *MSMEG_4642* does not reveal any significant hits against the non-redundant database. Hence, we currently cannot ascribe a function to this protein nor confirm that the gene annotation is correct.

The overall amino acid sequence homology between *M. smegmatis* and *M. tuberculosis* was high for all Rpf proteins, Figure 8. RpfE ranked highest with 80% sequence identity between the two organisms, while 65% sequence similarity was observed for RpfC. A 67% and 78% sequence identity was observed for RpfA and RpfB respectively.

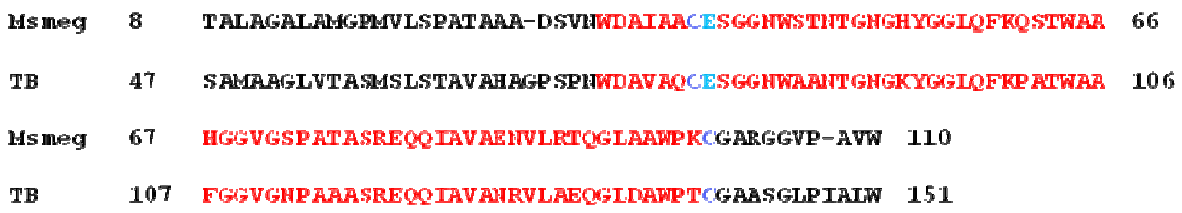
**RpfA**

```
Msmeg   1    MSGRHRKPTSSSSAKNVAKIAFTGAVLGGGGLGMAGQAMAATDGEWDQVARCESGGNWAI  60
TB      1    MSGRHRKPTTSNVS--VAKIAFTGAVLGGGGIAMAAQATAATDGEWDQVARCESGGNWSI  58
Msmeg   61   NTGNGYHGGLQFSPSTWRAHGGTEFAPAAYMATKEEQIAVAERVLASQGKGAWPTCGRGL  120
TB      59   NTGNGYLGGLQFTQSTWAAHGGGEFAPSAQLASREQQIAVGERVLATQGRGAWPVCGRGI  118
Msmeg   121  SGATPRNVV--AEPEPAPLDAPGVNGE--------LPPPAPVDMLAAPMPAPEG  164
TB      119  SNATPREVLPASAAMDAPLDAAAVNGEPAPIAPPPADPAPPVELAANDLPAPLG  172
```

**RpfB**

```
Msmeg   29   AGSYAVASHKTVTLSVDGAPMTVTTMKSRVIDIVEENGFEVSDRDDLFPAATETVHQSDT  88
TB      16   AGGYAVAACKTVTLTVDGTAMRVTTMKSRVIDIVEENGFSVDDRDDLYPAAGVQVHDADT  75
Msmeg   89   IVLRRSRPLEVSLDGEGTKQVWTTASTVDEALAQLRMTDKAPAAASRGSRVPLAGMALPV  148
TB      76   IVLRRSRPLQISLDGHDAKQVWTTASTVDEALAQLAMTDTAPAAASRASRVPLSGMALPV  135
Msmeg   149  VSPKTVQIDDGGKVSIVHLAAPNVAGLLEAAGVPLEQNDKVVPAASSPVVDGMQIQVTRM  208
TB      136  VSAKTVQLNDGGLVRIVHLPAPNVAGLLSAAGVPLLQSDHVVPAATAPIVEGMQIQVTRN  195
Msmeg   209  RIEKVTERVPLAPGNERIEDPTMMMSRQVVEDPGAPGLQDVTFAIAKVNGVETGRLPVAN  268
TB      196  RIKKVTERLPLPPNARRVEDPEMMMSREVVEDPGVPGTQDVTFAVAEVNGVETGRLPVAN  255
Msmeg   269  QVIEPARNGVLRVGAKPGIEVPPVSNGHTWDALAQCEAGGHWAINTGNGFYGGVQFDQNT  328
TB      256  VVVTPAMEAVVRVGTKPGIEVPPVIDGSIWDAIAGCEAGGHWAINTGNGYYGGVQFDQGT  315
Msmeg   329  WERNGGLRYAPRADLATREEQIAIATVTQSRQGWGAWPVCSGRIGAR  375
TB      316  WEANGGLRYAPRADLATREEQIAVAEVTRLRQGWGAWPVCAARAGAR  362
```

**RpfC**




**RpfE**

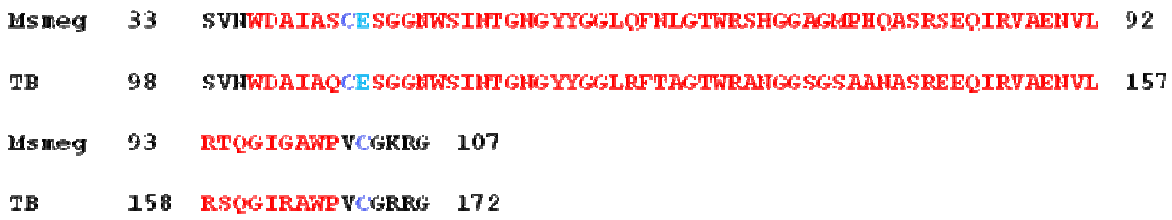



**Figure 8: Comparative amino acid sequence alignment of the Rpf proteins in *M. smegmatis* and *M. tuberculosis*.** The protein sequences shown above for RpfA have been omitted, in the Blast search, for the last few hundred amino acid residues probably due to the inability of the sequences to align in the correct manner for the organisms in question, in this case the conserved Rpf domain is shown. Colour-coded sequences: Rpf domain (*red*), conserved cysteine residue (*blue*), conserved catalytic glutamate residue (*light blue*), DUF 348 Superfamily domain (*purple*), G5 Superfamily domain (*green*), glycine residues (*yellow*).

As is observed in Figure 8, the Rpf domain (highlighted in red), which is a characteristic feature for all the Rpf proteins in both organisms, retains one of a pair of cysteine residues (blue) and the catalytic glutamate residue (light blue). The other cysteine residue is located outside the Rpf domain and is proposed to be involved in the formation of a disulphide bridge (Mukamolova*, et al.*, 2006). Interestingly, the Rpf domain is positioned C-terminally in all the Rpfs with the exception of RpfA, a feature originally observed in *M. tuberculosis* (Mukamolova*, et al.*, 2002). The significance of this is unknown but may be due to the presence

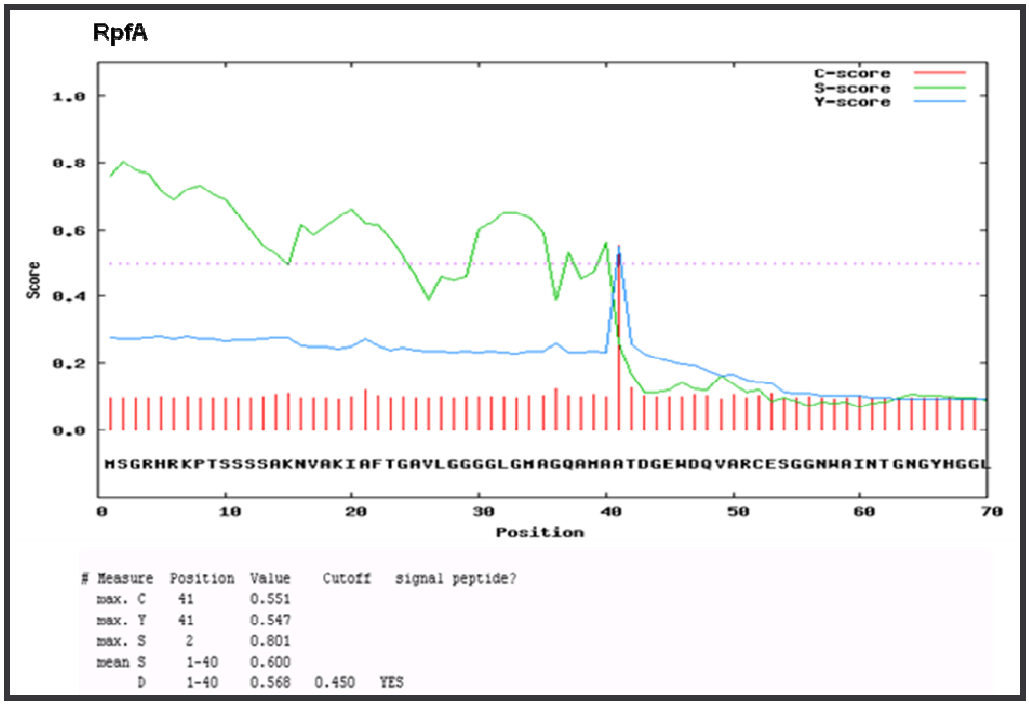
of a signal sequence (with associated post-translational cleavage) at the N-terminus. *M. smegmatis* RpfC and RpfE are 153 and 107 amino acid residue proteins respectively, which are the smallest of all the Rpfs. Moreover, approximately 62% of *M. smegmatis* RpfE is comprised of the Rpf domain. One of the largest and more complex Rpfs is RpfB, which consists of 375 amino acid residues in *M. smegmatis* and retains other domains in addition to the conserved Rpf domain. Most noteworthy is the DUF 348 superfamily domain; DUF's (Domains of Unknown Function) are essentially large groups of protein families with unknown function available in the Pfam database. These families may represent one or more protein domain(s) and are numbered in order of submission to the database (Bateman*, et al.*, 2010).

### 3.1.1 *M. smegmatis* Rpf-like proteins are secreted signal peptides

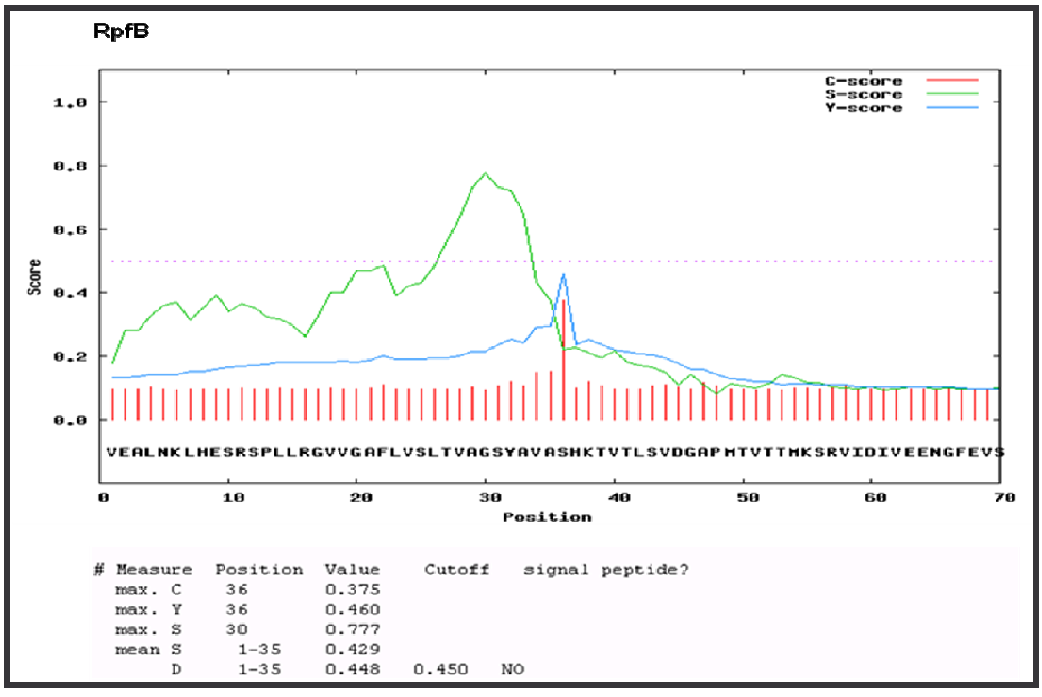
The *M. tuberculosis* Rpfs are predicted to be secreted peptides, as evidenced by the presence of secretory signal sequences at the N-terminus (Mukamolova*, et al.*, 1998). It is highly likely that the *M. smegmatis* Rpf-like proteins are also secreted, especially since a high protein sequence similarity is observed between the homologues in both organisms (Figure 8). This hypothesis was further explored with bioinformatics analysis using SignalP, a program that is able to identify secreted signals.

The *M. smegmatis* Rpf-like proteins are secreted as predicted by SignalP 4.0 server. SignalP 4.0 server predicts the presence of signal peptide sequences and cleavage sites within the amino acid sequence; furthermore it predicts the cleavage site position. Three different scores (C, S, Y) are used in the graphical output from SignalP following submission of protein sequences to the server. These scores determine whether the protein is a signal peptide and also localizes the cleavage site. All three scores are reported for every amino acid position in the sequence; the C-score (cleavage site score) should have a

significantly high value at the cleavage site, while the S-score should display high values at every amino acid that constructs the signal peptide.




| # Measure | Position | Value | Cutoff | signal peptide? |
|---|---|---|---|---|
| max. C | 41 | 0.551 | | |
| max. Y | 41 | 0.547 | | |
| max. S | 2 | 0.801 | | |
| mean S | 1-40 | 0.600 | | |
| D | 1-40 | 0.568 | 0.450 | YES |




| # Measure | Position | Value | Cutoff | signal peptide? |
|---|---|---|---|---|
| max. C | 36 | 0.375 | | |
| max. Y | 36 | 0.460 | | |
| max. S | 30 | 0.777 | | |
| mean S | 1-35 | 0.429 | | |
| D | 1-35 | 0.448 | 0.450 | NO |

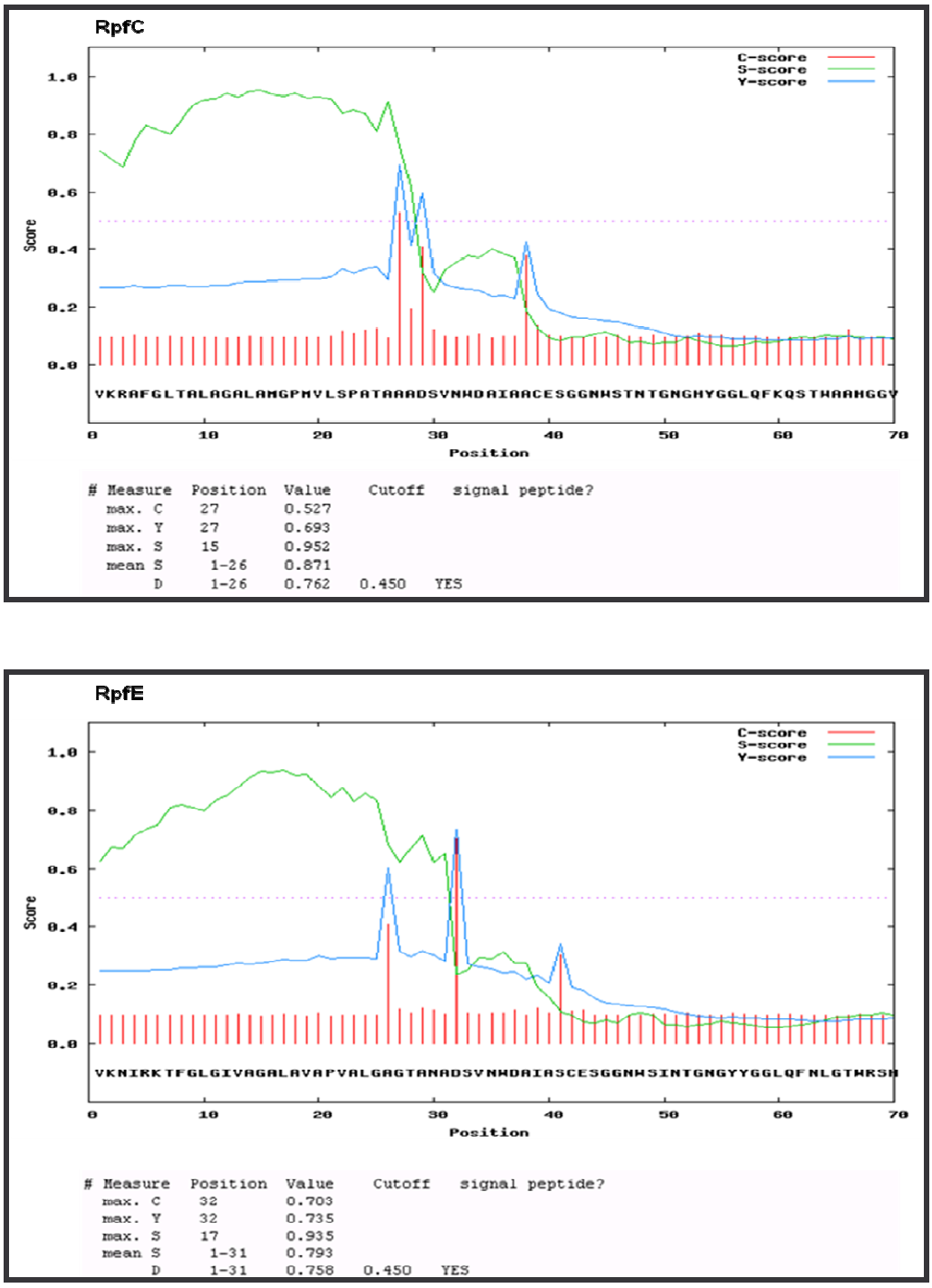



**Figure 9**: Signal peptide prediction and localization of signal peptide cleavage sites within the *M. smegmatis* Rpf-like protein sequences.

A few high-peaking C-scores (red vertical lines) can be reported in a single sequence, as depicted above in Figure 9, and it is thus essential to determine which one is the actual cleavage site. The Y-score is assigned for the above task; it is a combination of the C- and S-scores resulting in a more accurate cleavage site prediction. Low values (below the threshold line) for all scores represent a non-secretory protein.

In our analysis shown in Figure 9, three of the four *M. smegmatis* Rpf-like proteins are predicted to be signal peptides, whereas RpfB is not. The sole reason for a negative signal peptide prediction for RpfB is due to the fact that the protein is a lipoprotein and prokaryotic lipoproteins cleavage sites cannot be predicted with the use of SignalP. Lipoproteins are processed by specific signal peptidases, *Lsp* or signal peptidase II (Sander*, et al.*, 2004, Denham*, et al.*, 2008), whereas the SignalP server reports predictions based on signal peptides processed by signal peptidase I. A more appropriate server for prokaryotic lipoproteins is LipoP; however it is only designed for Gram-negative bacteria and thus we conducted no further analysis on RpfB with this program.

### 3.2 Construction of mutants defective for one or more *rpf* genes

To assess the function of *rpf*-like genes in *M. smegmatis*, defined knockout mutants that lack one or more of these genes were constructed by two-step allelic mutagenesis using homologous recombination (Gordhan & Parish, 2001). The bioinformatics analysis provided the relevant insight into gene distribution and organization which allowed for the design of a suitable knockout strategy to delete *rpfA* and *rpfB* individually and in combination. These two particular genes were chosen for the following reasons: (I) the regulation of *rpfA* has been linked to the cAMP repressor in *M. tuberculosis* (Rickman*, et al.*, 2005). cAMP is an important regulator of bacterial metabolism and this combined with the fact that RpfA is the largest of all Rpfs necessitated the inclusion of RpfA in our analysis (II) As mentioned previously, RpfB has been shown to interact with an essential

endopeptidase RipA and co-localizes with this protein to the cell poles. It was further demonstrated that RpfB synergizes with RipA in peptidoglycan degradation suggesting that RpfB plays an important role in cell growth hence its inclusion in our analysis.

For gene deletion, primers were designed, using the program PRIMER 3, in such a way to create in-frame deletions of the *rpf* genes. The reading frame of the homologous upstream and downstream regions should not be disrupted in order to create an in-frame mutated copy of the target gene. The amplified homologous upstream and downstream regions flanking the deletion are directionally cloned into the multiple cloning sites (MCS) of p2NIL. A marker gene cassette from pGOAL19 (containing the *lacZ* gene– for blue colour selection, the *sacB* gene– for sucrose counter selection and the hygromycin resistance gene) is subsequently inserted into the *PacI* site of the p2NIL vector, thus completing the suicide plasmid construction. During allelic exchange mutagenesis, a mutated copy of the gene of interest is introduced into the chromosome with the use of suicide vectors. A recombination event between wild type and mutant allele gives rise to a strain with both gene copies, while two recombination events result in the replacement of the wild type gene with the mutant copy. This two-step strategy for constructing *M. smegmatis deletion* mutants involves selection of blue, $Hyg^R$ and $Km^R$ transformants on solid media containing those antibiotics and X-gal. These have undergone a single recombination event are referred to as single cross-over (SCO) strains. A second recombination event is permitted to occur by growth of the abovementioned transformants under the same conditions. The cells are further grown and selected on solid media containing sucrose and X-gal; and the white $sucrose^R$ clones represent potential double cross-over strains (DCO), the deletion mutants.

### 3.2.1 Suicide delivery vector construction

#### 3.2.1.1 Polymerase chain reaction (PCR)

With the use of the designed primers, PCR was performed in order to amplify upstream and downstream regions of the *rpf* genes. This is illustrated below in Figure 10.

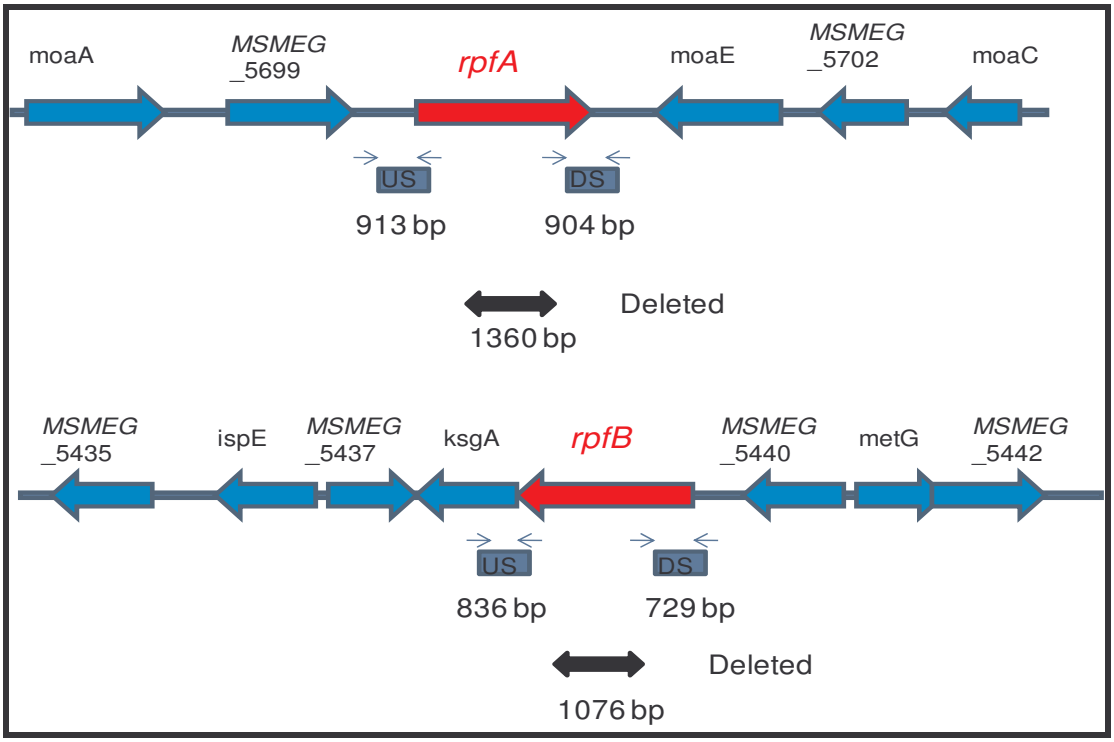



**Figure 10: A schematic representation of the *rpf* regions amplified during PCR with specifically designed primers, and the resulting PCR fragments with expected sizes.** Small arrows represent forward and reverse primers and the gene of interest is highlighted and labeled in red. US-upstream and DS-downstream regions.

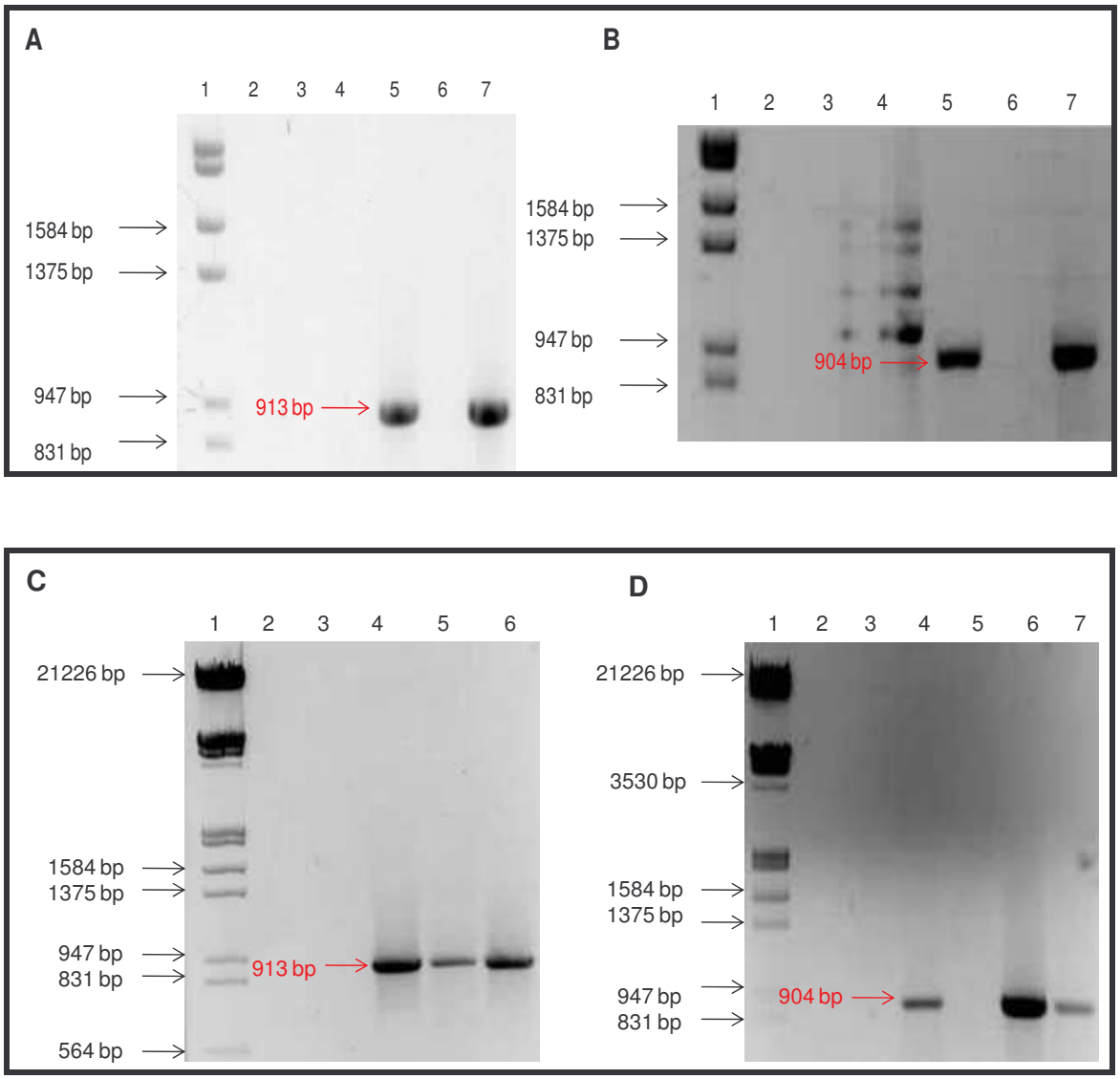



**Figure 11: PCR fragments amplified for *rpfA*. (A+B)** *Thermus aquaticus* (Taq) PCR. **(A)** Upstream PCR products: **Lane 1**: Molecular weight marker λIII, **lane 2**: no sample loaded, **lane 3**: forward primer control, **lane 4**: reverse primer control, **lane 5**: Test reaction with upstream PCR product of interest (1 µl template - *M. smegmatis* genomic DNA), **lane 6**: no DNA control, **lane 7**: Test reaction with upstream PCR product (5 µl template DNA). **(B)** Downstream PCR products: **Lane 1**: λIII, **lane 2**: no sample loaded, **lane 3**: forward primer control, **lane 4**: reverse primer control, **lane 5**: Test reaction with downstream PCR product of interest (5µl template - *M. smegmatis* genomic DNA), **lane 6**: no DNA control, **lane 7**: Test reaction with downstream PCR product (10µl template DNA). **(C+D)** Expand High Fidelity PCR amplification of upstream and downstream

regions for *rpfA*. **(C)** Upstream *rpfA* PCR products: **Lane 1**: λIII, **lane 2**: no sample loaded, **lane 3**: no DNA control, **lane 4-6**: Test reaction with upstream PCR products run in triplicate. **(D)** Downstream *rpfA* PCR products: **Lane 1**: λIII, **lane 2**: no sample loaded, **lane 3**: no DNA control, **lane 4-7**: Test reaction with downstream PCR products.

The PCR products for the upstream and downstream homologous regions for deletion of ***rpfA***; seen above in 1% agarose gels, Figure 11, were of the correct expected sizes. The upstream and downstream fragments were expected to be 913 bp and 904 bp respectively, thus the PCR proved successful. All control reactions, including no DNA, forward primer only and reverse primer only, yielded the expected result with the exception of the reverse only control for the downstream fragment, gel B. However, in this case there was marginal amplification of fragments with similar sizes to the expected PCR product, suggesting that the majority of these bands were non-specific. Initially, the PCR was carried out with Taq enzyme, seen in Figure 11 (A+B) in order to optimize conditions. Thereafter PCR proceeded with Expand High Fidelity enzyme, as the name suggests, preventing error-prone PCR or point mutations. In this regard, both upstream and downstream homologous fragments for *rpfA* were amplified successfully with the Expand enzyme.

Similarly, the fragment sizes of both upstream and downstream ***rpfB*** PCR products were correct as is seen in Figure 12. The upstream and downstream fragments were expected to be 836 bp and 729 bp respectively, as observed in the gels below. PCR optimization was carried out with Taq enzyme; thereafter the PCR was repeated with Phusion enzyme to reduce any chances of possible point mutations introduced within the sequence of the PCR products.

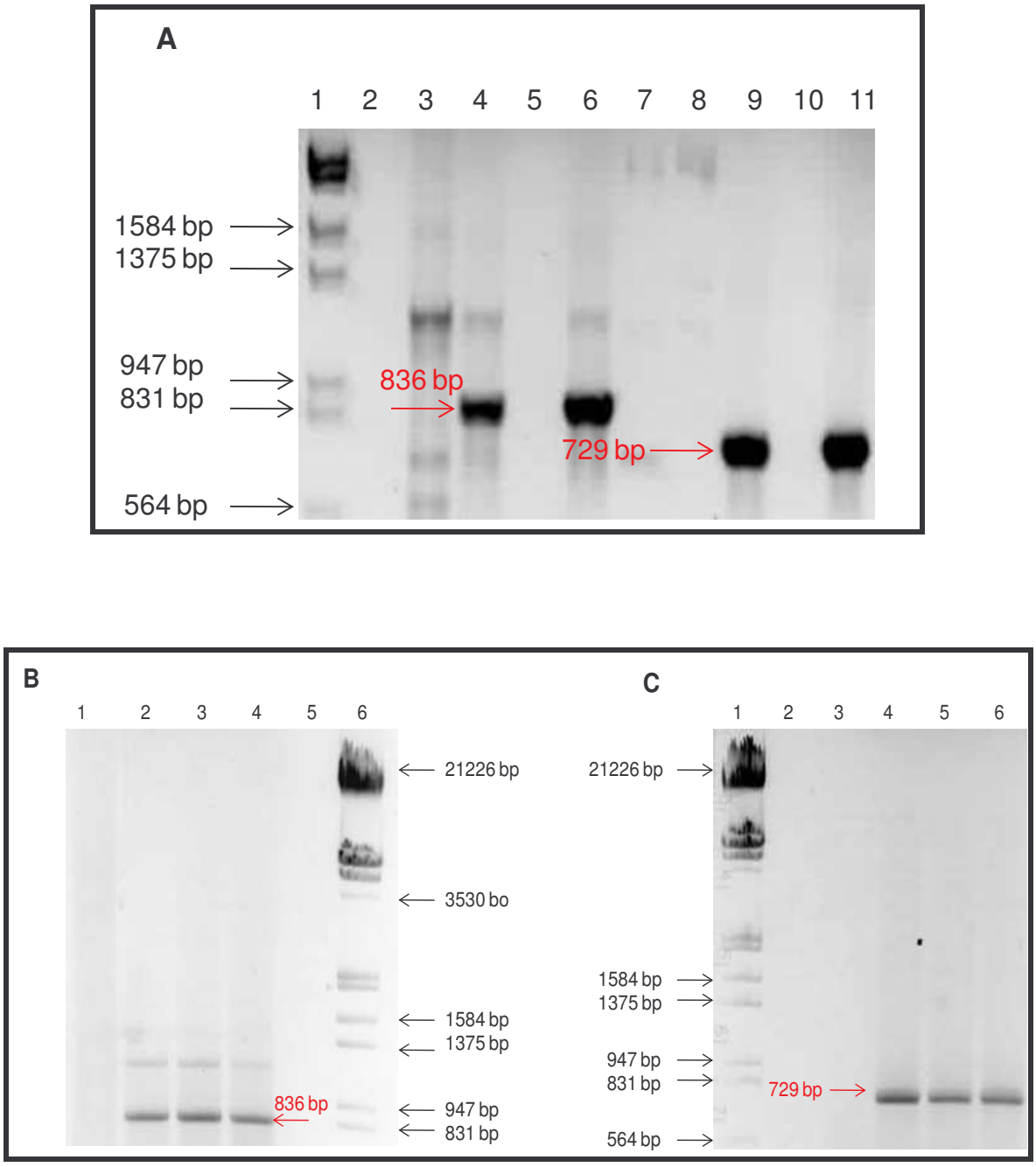



**Figure 12: PCR fragments amplified for *rpfB* homologous regions**. **(A)** Taq PCR: **Lane 1**: Molecular weight marker λIII, **lane 2**; forward primer control, **lane 3**: reverse primer control, **lane 4**: Test reaction with upstream PCR product of interest (1µl template DNA), **lane 5**: no DNA control, **lane 6**: Test reaction with upstream PCR product of interest (5µl template DNA), **lane 7**: forward primer control, **lane 8**: reverse primer control, **lane 9**: Test reaction with downstream PCR product of interest (1µl template

DNA), **lane 10**: no DNA control, **lane 11**: Test reaction with downstream PCR product of interest (5µl template DNA). **(B+C)** PCR products amplified with Phusion PCR. (B)Upstream PCR fragments: **Lane 1**: no DNA control, **lane 2-4**: Test reaction with upstream PCR products, **lane 5**: no sample loaded, **lane 6**: λIII. **(C)** Downstream PCR-amplified fragments: **Lane 1**: λIII, lane 2: no sample loaded, **lane 3**: no DNA control, **lane 4-6**: Test reaction with downstream PCR products.

Non-specific bands were observed for the reverse primer control during optimization; Figure 12 A, lane 3, and were also present in the test reactions (gel A, lane 4 and 6) for upstream *rpfB* fragment amplification. Further optimization reduced and removed amplification of these non-specific bands, gel B and gel C respectively. All other control and test reactions yielded the expected results for both upstream and downstream *rpfB* fragments. The addition of 5 µl template *M. smegmatis* DNA provided marginally better yields of PCR product. Amplification of both upstream and downstream *rpfB* regions was thus successful.

#### 3.2.1.2 Cloning of upstream and downstream homologous regions

The abovementioned upstream and downstream PCR fragments, purified from gels, were cloned separately into carrier plasmid vectors. The carrier plasmid vectors used in this study were pBluescript and pGEM3Zf (+); where pBluescript was used as carrier vector for *rpfB* upstream and downstream fragments, whereas pGEM3Zf (+) was used for *rpfA* fragment cloning.

Expand High Fidelity enzyme produces T-overhangs in the PCR fragments thus this was the reason pGEM-TEazy kit, which utilizes these overhangs to facilitate cloning, was suitable for cloning. The pGEM vector in this kit came ready for use and no further preparation was required unlike pBluescript. Phusion enzyme produces blunt ends in its PCR products and pBluescript was appropriate for facilitating such cloning. The different High Fidelity enzymes were used for different *rpf* gene region cloning due to availability of either enzyme at the time of

cloning. Both these vectors allow for blue/white colony selection, due to the presence of a *lacZ* gene within the plasmid, which is useful in determining whether cloning is successful or not. The *lacZ* gene expresses β-galactosidase and in the presence of 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal), *E. coli* colonies turn blue. In the case of pBluescript, *Sma*I, a blunt cutter, was the appropriate choice of enzyme for cloning, and the *Sma*I site is located within the multiple cloning sites (MCS) within the *lacZ* gene sequences. Cloned fragments at this site disrupt the *lacZ* gene, thus rendering the resulting protein inactive. Hence, successful cloning, in this case, results in white colonies.

For cloning into pBluescript, the vector was digested with *Sma*I and separated on a 1% agarose gel, shown in Figure 13.

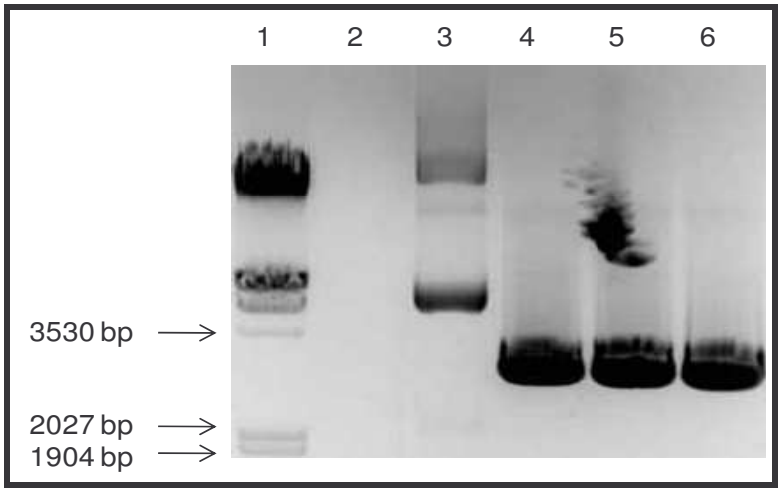



**Figure 13**: **pBluescript carrier plasmid vector linearized with *Sma*I enzyme**. **Lane 1**: Molecular weight marker λIII, **lane 2**: no sample loaded, **lane 3**: uncut plasmid, **lane 4-6:** *Sma*I- linearized plasmid.

The vector size is 2964 bp which corresponded to the size observed on the gel shown in Figure 13, lanes 4-6. This band was excised and purified and purified *rpfB* upstream and downstream homologous PCR fragments (Figure 12 B+C) were ligated to the vector by blunt cloning and transformed into *E. coli* DH5 α cells, followed by the selection and screening of white colonies.

Similarly, the *rpfA* upstream and downstream homologous regions were cloned into pGEM (preparation of the plasmid was not required as explained earlier) in an identical manner and *E. coli* cells were transformed with the construct.

In both cases screening of numerous white clones, selected on Amp 100µg/ml, was initially undertaken by PCR (gels not shown) to confirm whether cloning was successful for both *rpfA* and *rpfB* intermediate constructs. Both white and blue colonies selected from LA plates were re-suspended in LB and the cell suspension was used as template DNA for the PCR reaction. Indeed the cloning was successful as was evidenced by the presence of fragments generated by PCR with observed sizes corresponding to expected sizes of upstream and downstream homologous regions for both genes. Selected blue colonies were used as a control, since they represented clones in which cloning was unsuccessful. Therefore no PCR product was expected from blue colonies as they did not contain the construct. Positive clones were selected for further screening and cloning purposes.

As shown in Table 2 (section 2.4.1), the primers used to amplify the upstream and downstream homologous regions were designed in such a way to include restriction enzyme recognition sequences; where *Asp*718 and *Bgl*II restriction sites were present on either ends of upstream PCR fragments, likewise *Bgl*II and *Hind*III restriction sites on either sides of downstream PCR DNA fragments for cloning purposes.

Plasmid DNA was extracted from PCR-screened, white colonies, and digested with the restriction enzymes mentioned above to ensure that the correct fragments and expected restriction sites were present in these vectors. As shown in Figure 14, for one clone from each cloning, all sizes observed corresponded to those expected for the upstream and downstream homologous regions. This served as further confirmation for the success of these cloning steps.

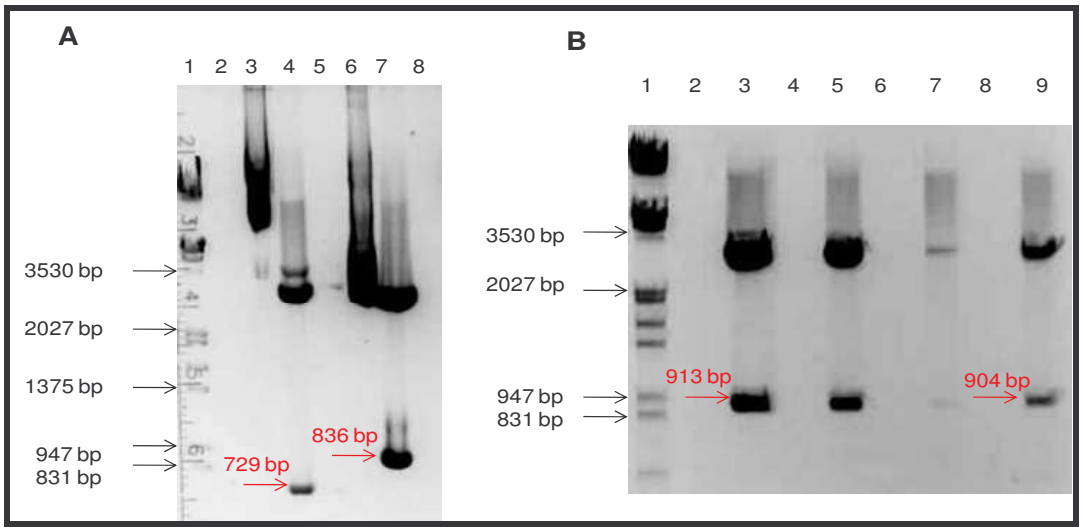



**Figure 14: *rpfA* and *rpfB* constructs digested with appropriate enzymes releasing the cloned insert. (A)** pBUΔ*rpfB* and pBDΔ*rpfB* constructs: **Lane 1**: Molecular weight marker λIII, **lane 2**: no sample loaded, **lane 3**: undigested plasmid, **lane 4**: *Bgl*II/*Hind*III double digestion releasing the downstream *rpfB* homologous fragment, **lane 5**: no sample loaded, **lane 6**: undigested plasmid, **lane 7**: *Asp*718/*Bgl*II-digested construct releasing upstream fragment, lane 8: no sample loaded. **(B)** pGUΔ*rpfA* and pGDΔ*rpfA* constructs (loaded on gel in duplicate): **Lane 1**: λIII, **lane 2**: no sample loaded, **lane 3** & 5: *Asp*718/*Bgl*II-digested construct releasing upstream fragment, **lane 4 & 6**: no sample loaded, **lane 7 & 9**: *Bgl*II/*Hind*III-digested plasmid releasing downstream fragment, **lane 8**: no sample loaded.

Fragments shown in Figure 14, indicated by red arrows with corresponding sizes confirmed the integrity of the clones. The DNA concentration yielded in gel B, lane 7 was low, which was not a matter of concern since the same DNA extracted from another clone yielded a good concentration as observed in lane 9. Sequencing of the above constructs confirmed their identity and indicated that no point mutations had occurred in the *rpfA* and *rpfB* upstream and downstream homologous sequences during PCR amplification. Hence these fragments could be used for the next phase of suicide vector construction.

### 3.2.1.3 Three-way cloning to create deletion alleles

Three-way cloning involves simultaneous cloning of both upstream and downstream fragments into the p2NIL vector, a vector specifically developed for gene knockout in mycobacteria (Parish & Stoker, 2000). The partial sequence for this vector has been provided by the researchers who have generated it, however, it has been noted that there are several unmapped restriction sites.

The vector was prepared and digested with appropriate enzymes for directional cloning, outlined in Figure 15. In this particular case, the vector was linearized with *Asp*718/*Hind*III to create the linear vector required for cloning, Figure 16. As a control for restriction digestion, the vector was also digested individually with these enzymes to ensure that each enzyme gave complete digestion of the same preparation of vector. These controls confirm that the enzymes would digest the vector in the double digest reaction.

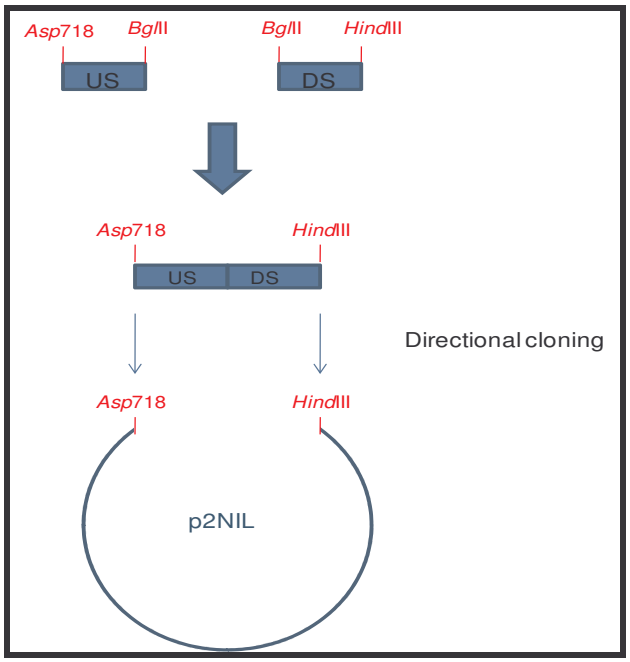



**Figure 15**: Schematic showing three-way cloning resulting in simultaneous ligation of upstream and downstream homologous regions to p2NIL knockout vector by directional cloning.

The cloning was directional, as the name suggests where the *Asp*718 end of the upstream fragment and the *Hind*III end of the downstream fragment simultaneously ligate to the compatible ends of the *Asp*718/*Hind*III double-digested p2NIL vector. The fragments ligate to each other since they both have *Bgl*II compatible ends, as illustrated in Figure 15.

The size of the p2NIL vector is 4753 bp and this corresponded to the linearized vector observed in Figure 16, in lane 3 and 4. The circular, undigested p2NIL plasmid migrated ahead of the linear plasmid due to the fact that it was supercoiled and thus migrated faster. The vector digested with both enzymes was expected to yield a 4436 bp vector backbone and a 317 bp fragment which is released from the vector.

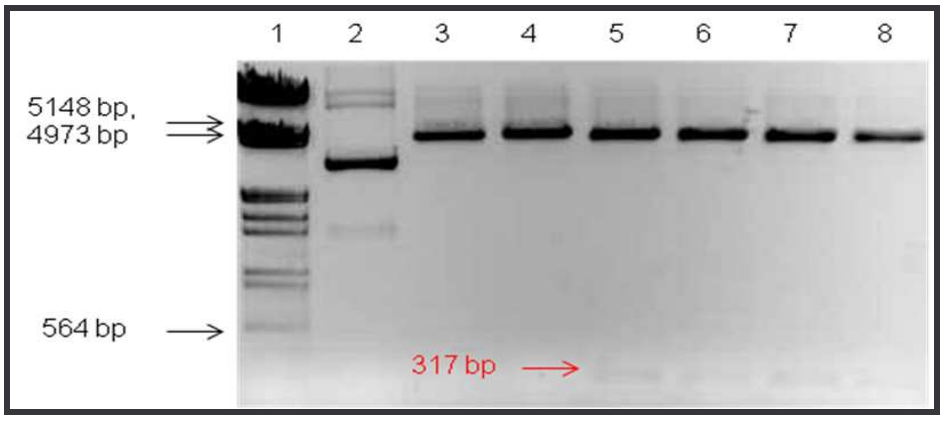



**Figure 16: p2NIL plasmid vector digested with appropriate enzymes for cloning purposes**. **Lane 1**: Molecular weight marker λIII, **lane 2**: undigested vector, **lane 3**: *Hind*III-digested vector, **lane 4**: *Asp*718-digested plasmid, **lane 5-8**: *Asp*718/*Hind*III double-digested p2NIL vector.

As observed in Figure 16, lane 5-8, the vector backbone DNA bands were slightly smaller to those seen in lane 3 and 4 and therefore corresponded to the size expected. Furthermore, small faint bands were released from the digestions

as indicated by the red arrow and the expected band sizes were 317 bp in accordance to what was observed from the gel. These digests confirm the integrity of the vector used for cloning. With regards to the inserts, the upstream and downstream homologous fragments were digested from intermediate vectors with the restriction enzymes mentioned above, Figure 17. The bands corresponding to the expected sizes (shown in Figure 14 and in Figure 17) were purified and simultaneously cloned into the purified 4436 bp p2NIL fragment.

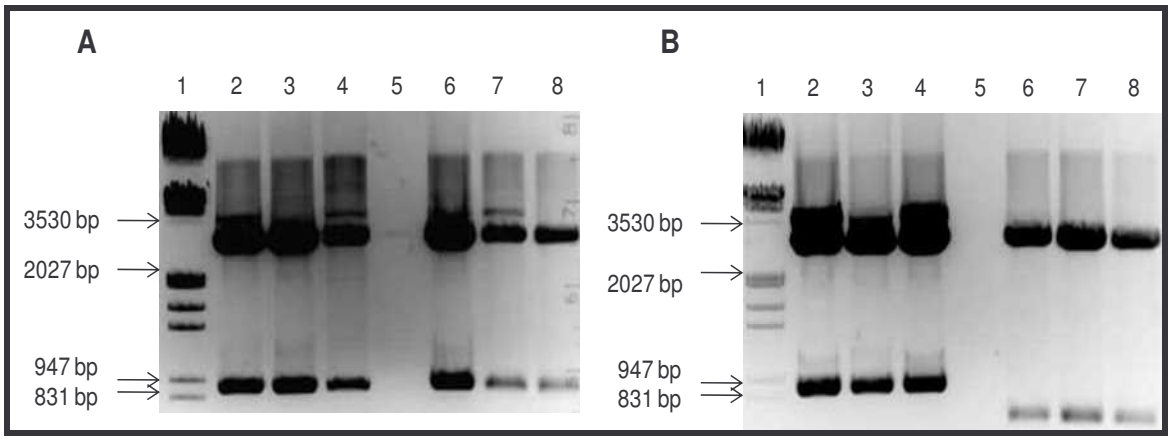



**Figure 17: Upstream and downstream homologous *rpf* fragments digested from replicating vectors. (A)** *rpfA* homologous regions released from pGEM: **Lane 1**: Molecular weight marker λIII, **lane 2-4**: Upstream fragment released by *Asp*718/*Bgl*II digestions, **lane 5**: no sample loaded, **lane 6-8**: *Bgl*II/*Hind*III digestions releasing downstream fragment. **(B)** pBluescript-released *rpfB* homologous fragments: **Lane 1**: λIII, **lane 2-4**: *Asp*718/*Bgl*II digestions releasing the upstream fragment, **lane 5**: no sample loaded, **lane 6-8**: Downstream fragment released by *Bgl*II/*Hind*III digestions.

*E. coli* DH5α cells were transformed with the ligation and clones were selected on LA containing Km 50 ug/ml. Selected clones were initially screened by PCR (gels not shown), using primers to amplify the upstream homologous regions for both *rpfA* and *rpfB*. Each selected colony was re-suspended in LB and the cell suspension was used as template DNA in the PCR reaction. The presence of

PCR products with the expected upstream fragment size inferred that the three-way cloning was successful, since it would be impossible to clone the upstream region only without the downstream region due to compatible ends.

A single PCR-positive clone was selected for each p2NIL knockout construct (pNΔRPFA or pNΔRPFB) for further restriction mapping, to confirm whether the three-way cloning was indeed correct.

**Table 4:** Restriction mapping for pNΔRPFA

| Enzyme | Expected fragment sizes (bp) |
|---|---|
| *Pvu*I | 2901, 1022, 1683 |
| *Mlu*I | 4045, 1561 |
| *Bgl*II/*Hind*III | 890, 4716 |
| *Cla*I | 2166, 3440 |
| *Asp*718/*Bgl*II | 899, 4707 |
| *Asp*718/*Hind*III | 1789, 3817 |

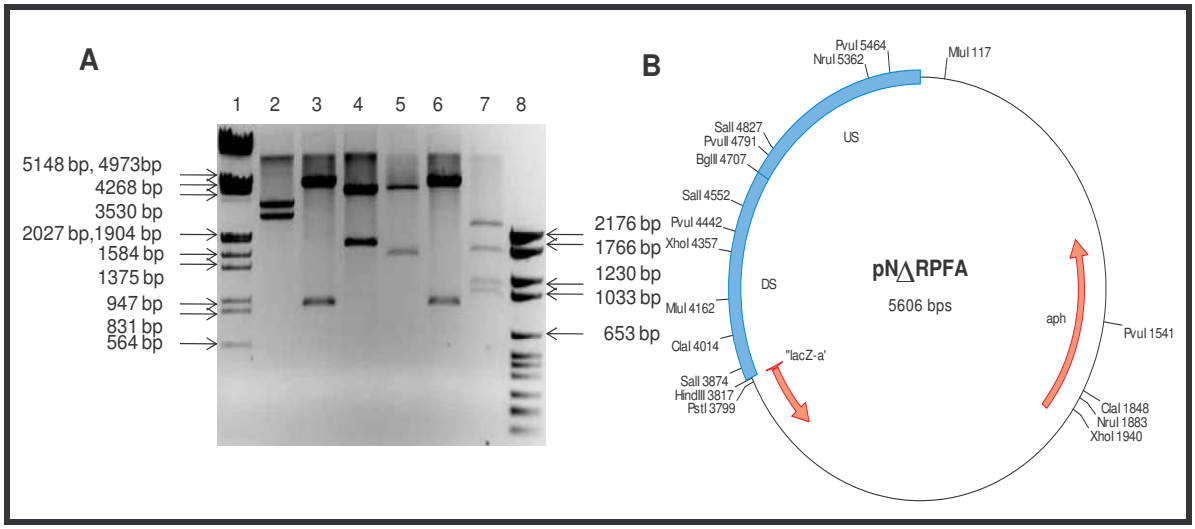



**Figure 18: Restriction mapping of the pNΔRPFA construct shown alongside its map. (A)** Restriction digestions with various enzymes: **Lane 1**: Molecular weight marker

λIII, **lane 2**: ClaI, **lane 3**: *Asp*718/*Bgl*II, **lane 4**: *Asp*718/*Hind*III, **lane 5**: *Mlu*I, **lane 6**: *Bgl*II/*Hind*III, **lane 7**: *Pvu*I, **lane 8**: λVI. **(B)** pNΔRPFA map showing cloned upstream and downstream *rpfA* regions into p2NIL.

The restriction mapping of pNΔRPFA (Figure 18) with *Cla*I, *Asp*718/*Bgl*II, *Asp*718/*Hind*III, *Mlu*I, and *Bgl*II/HindIII all gave the expected fragments. One digestion gave unexpected results; three fragments were expected from the *Pvu*I digestion and four fragments were observed. In this case the expected 1022 bp and 1683 bp fragments (which are internal to the homologous regions cloned) were observed however, the 2901 bp vector-containing band was digested into two smaller fragments suggesting the presence of an unmapped *Pvu*I site in the vector or new site inadvertently introduced during the cloning through vector backbone rearrangement.

Table 5**: pNΔRPFB restriction mapping**

| Enzyme | Expected fragment sizes (bp) |
|---|---|
| *Nru*I | 3321, 2033 |
| *Bgl*II/*Hind*III | 715, 4639 |
| *Pst*I | 4254, 1100 |
| *Stu*I | 748, 4606 |
| *Pvu*II | 953, 4401 |
| *Pvu*I | 2722, 938, 1694 |

Restriction mapping of the pNΔRPFB clone with *Stu*I, *Pvu*II, *Pvu*I, *Nru*I and *Bgl*II/*Hind*III all yielded the expected fragment sizes, Figure 19. In some cases, digestions were repeated to ensure reproducibility (Figure 19B). The *Pvu*II digest in lane 6 was incomplete, however, a complete digest for this enzyme with pNΔRPFB is shown in Figure 19A, lane 5– which yielded the expected sizes. The *Pst*I digest however, gave an inconsistent result, with three fragments instead of the expected two bands.

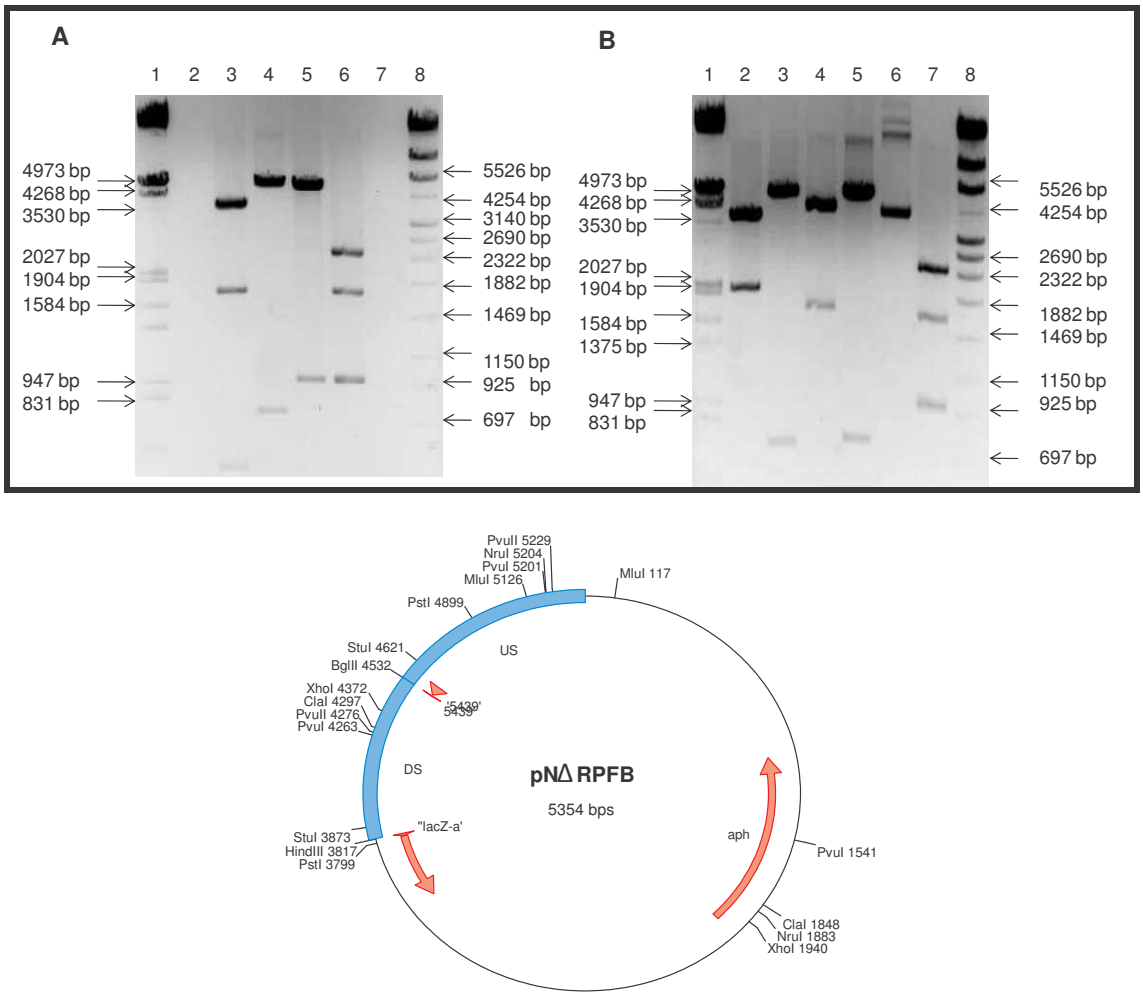



**Figure 19: pNΔRPFB restriction mapping. (A) Lane 1**: Molecular weight marker λIII, **lane 2**: no sample loaded, **lane 3**: *Pst*I, **lane 4**: *Stu*I, **lane 5**: *Pvu*II, **lane 6**: *Pvu*I, **lane 7**: no sample loaded, **lane 8**: λIV. **(B) Lane 1**: λIII, **lane 2**: *Nru*I, **lane 3**: *Bgl*II/*Hind*III, **lane 4**: *Pst*I, **lane 5**: *Stu*I, **lane 6**: *Pvu*II, **lane 7**: *Pvu*I, **lane 8**: λIV. **(C)** pNΔRPFB map.

In this case, the additional fragment would be due to an unmapped *Pst*I site at the vector-insert junction which most likely seems to have been introduced through the cloning process. The restriction mapping of both constructs yielded the expected fragments confirming the integrity of these clones, which were then used for the next step of suicide vector construction.

### 3.2.1.4 Completion of the suicide delivery vector: Cloning of selectable and counter selectable markers- *Pac*I cassette cloning

The final stage in the construction of the suicide vector involves cloning of the *Pac*I cassette from pGOAL19 plasmid vector into the p2NIL constructs (pNΔRPFA and pNΔRPFB). The p2NIL knockout construct (pNΔRPFA and pNΔRPFB) clones were re-grown and extracted plasmid was linearized with *Pac*I enzyme. At the same time, pGOAL19 was prepared and digested with *Pac*I enzyme which resulted in the release of the *Pac*I cassette. The *Pac*I cassette carries the *sacB*, *lacZ* and hygromycin resistant genes, enabling for ease of selection and counter-selection.

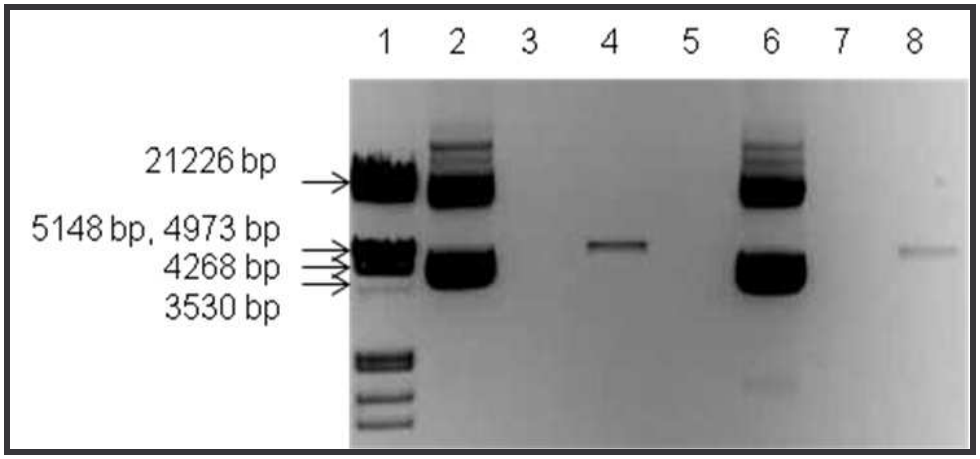



**Figure 20: *Pac*I linearized pNΔRPFA and pNΔRPFB constructs**. **Lane 1**: Molecular weight marker λIII, **lane 2**: uncut pNΔRPFA, **lane 3**: no sample loaded, **lane 4**: *Pac*I digested pNΔRPFA, **lane 5**: no sample loaded, **lane 6**: uncut pNΔRPFB, **lane 7**: no sample loaded, **lane 8**: *Pac*I digested pNΔRPFB.

The expected sizes of the linear pNΔRPFA and pNΔRPFB were 5606 bp and 5354 bp respectively and this corresponded to what was observed in the gel; Figure 20, lane 4 and 8.

The *Pac*I cassette is a large fragment carried in pGOAL19 plasmid vector which is 7939 bp and according to what was observed in Figure 21 below, the larger fragment corresponded to the *Pac*I cassette fragment size. The linear pNΔRPFA and pNΔRPFB constructs and the *Pac*I cassette were purified from the gels in preparation for cloning.

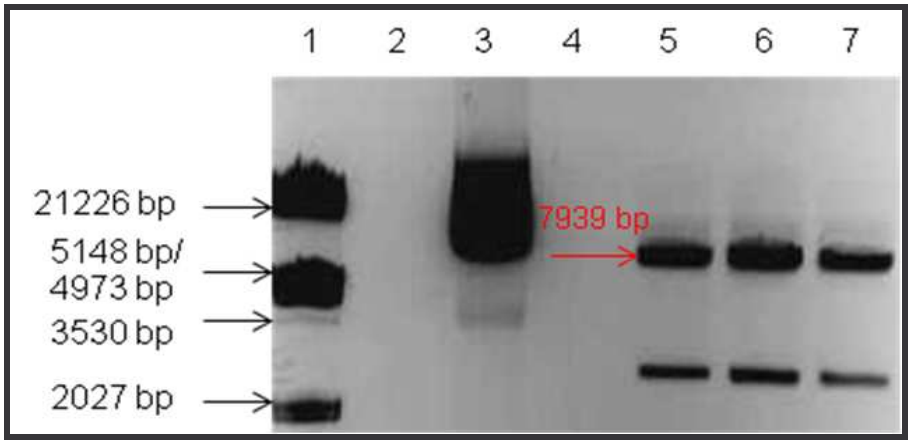



**Figure 21: *Pac*I cassette isolation from pGOAL19 plasmid vector. Lane 1**: Molecular weight marker λIII, **lane 2:** no sample loaded, **lane 3**: uncut plasmid vector, **lane 4**: no sample loaded, **lane 5-7**: *Pac*I digested pGOAL19.

The *Pac*I cassette was cloned into the p2NIL knockout constructs and transformed into *E. coli* DH5α cells. Colonies were selected on Km at 50 µg/ml LA containing X-gal, and grown slowly to prevent rearrangements since this would be a large construct. Selected blue colonies grown in broth containing Km at 50 µg/ml were serially diluted and further screened on Km at 50 µg/ml LA with and without 5% sucrose. The *sacB* gene within the *Pac*I cassette expresses an enzyme; levansucrase which is responsible for the conversion of sucrose to levan which is toxic to bacteria thereby resulting in cell death. It was thus expected to observe killing on sucrose-containing LA and normal growth on LA lacking sucrose. The observation confirms that the integrity of suicide vectors

had been maintained during growth. All selected clones were observed to be sucrose sensitive as described above, providing confirmation that gross rearrangements in marker genes had not occurred. Furthermore, restriction mapping was carried out to ascertain that the cloning was correct thus extracted plasmid was restricted with various enzymes and viewed on 1% agarose gels.

**Table 6:** Restriction mapping of Δ*rpfA* suicide vector (pSVΔRPFA).

| Enzyme | Expected fragment sizes (bp) |
|---|---|
| *Cla*I | 3105, 337, 1386, 1958, 2166, 4593 |
| *Mlu*I | 425, 780, 3567, 4045, 4728 |
| *Bgl*II/*Hind*III | 7048, 739, 27, 3906, 890, 935 |
| *Pst*I | 5696, 2031, 3847, 1971 |
| *Sal*I | 6547, 3975, 678, 275, 2070 |
| *Xho*I | 2293, 340, 7175, 2417, 1320 |

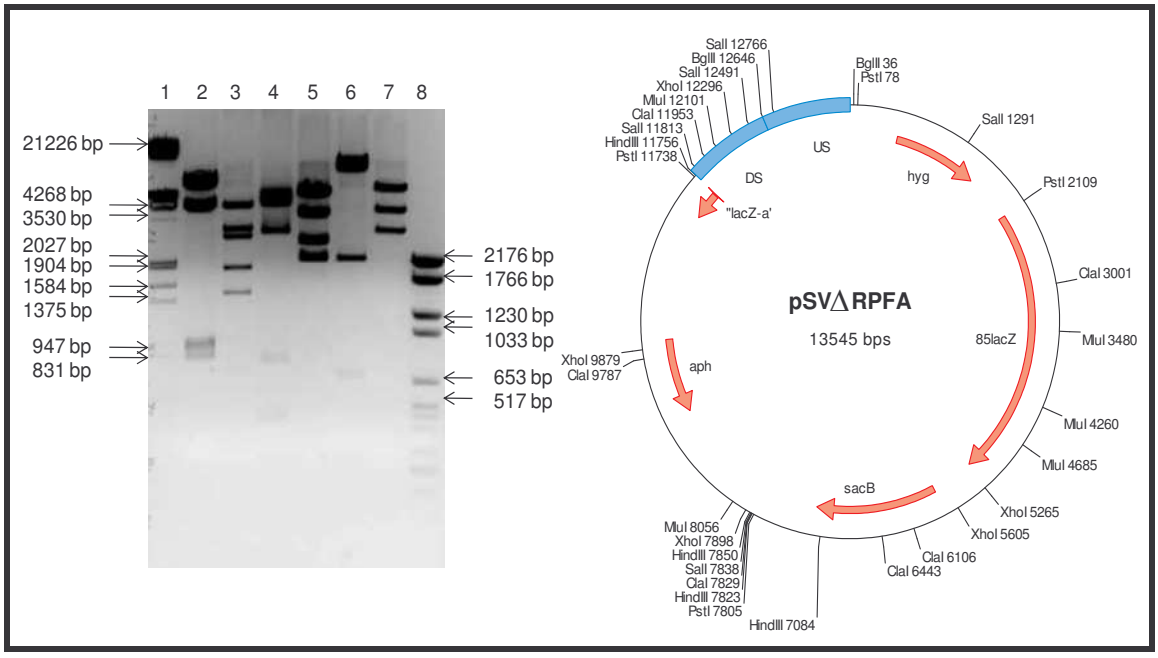



**Figure 22: pSVΔRPFA restricted with various enzymes for confirmation of suicide vector integrity.** Suicide delivery vector map shown alongside restriction mapping.

Digestions: **Lane 1**: Molecular weight marker λIII, **lane 2**: *Bgl*II/*Hind*III, **lane 3**: *Cla*I, **lane 4**: *Mlu*I, **lane 5**: *Pst*I, **lane 6**: *Sal*I, **lane 7**: *Xho*I, **lane 8**: λVI.

Restriction analysis, seen above in Figure 22, with *Bgl*II/*Hind*III, *Cla*I and *Mlu*I yielded the correct sizes. However, cutting with *Sal*I, *Xho*I and *Pst*I indicated that some vector rearrangements had occurred. We further analyzed this by checking the functionality of the markers and were able to select $km^R$ blue colonies confirming that *lacZ* and $km^R$ genes were intact. Furthermore the vector was sucrose sensitive in *E.coli* (data not shown) confirming that the *sacB* gene was functional. Extensive sequencing revealed that the fragments for the deletion allele were correct. Hence, this vector was used for mutant construction since the final allelic exchange mutant would not have any vector sequence.

**Table 7:** Δ*rpfB* suicide vector (pSVΔRPFB) restriction mapping.

| Enzyme | Expected fragment sizes (bp) |
|---|---|
| *Nru*I | 4954, 4356, 3321, 662 |
| *Pst*I | 2031, 5696, 3933, 1100, 533 |
| *Cla*I | 3105, 337, 1386, 1958, 2449, 4058 |
| *Sal*I | 6547, 6746 |
| *Bgl*II/*Hind*III | 7048, 739, 27, 3906, 715, 858 |
| *Xho*I | 340, 2293, 1981, 2432, 6247 |

Similarly, the digestions observed in Figure 23 indicated rearrangement events within the construct. As in the case with pSVΔRPFA restriction mapping, Figure 22, fragments in some of the digestions were either slightly larger, smaller or absent. Generally in all cases both for pSVΔRPFA and pSVΔRPFB, digestions that gave unexpected results always mapped to a p2NIL vector site. The suicide vectors were used for mutant construction since the inconsistency mapped to the vector backbone and not the deletion allele. Extensive restriction and sequence analysis of the mutant allele revealed no problems in this region.

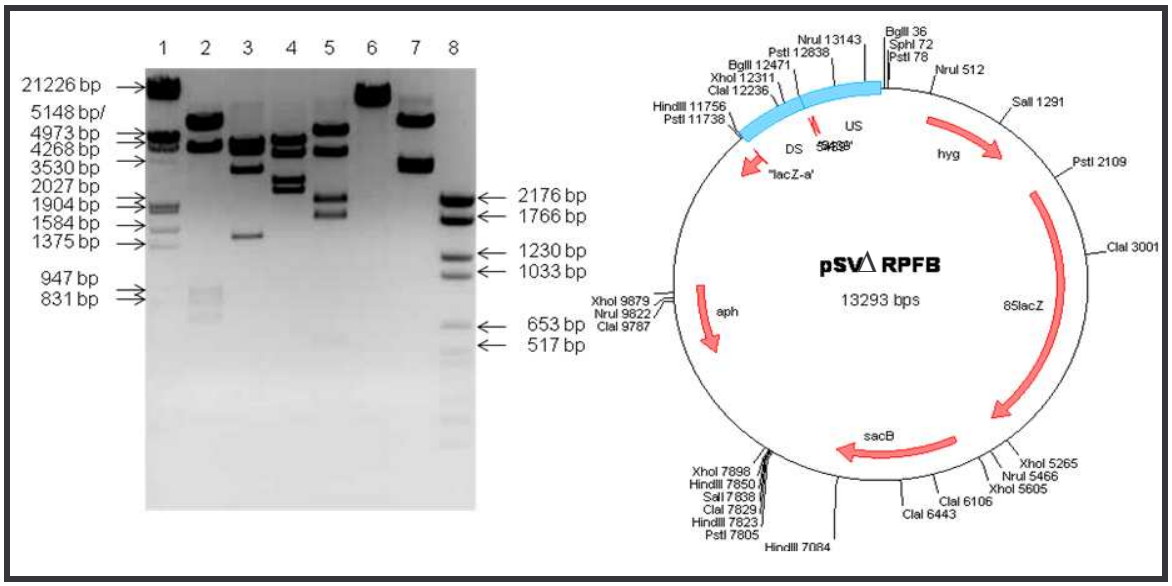



**Figure 23: Restriction mapping of pSVΔRPFB for plasmid integrity confirmation.**

pSVΔRPFB map shown alongside. Digestions: **Lane 1**: Molecular weight marker λIII, **lane 2**: *Bgl*II/*Hind*III, **lane 3**: *Cla*I, **lane 4**: *Nru*I, **lane 5**: *Pst*I, **lane 6**: *Sal*I, **lane 7**: *Xho*I, **lane 8**: λVI.

In parallel, a new strategy was therefore devised in efforts to avoid rearrangements; pGEM was employed as vector in the construction of the suicide vectors thereby replacing p2NIL. The pGEM vector was prepared and digested with *Asp*718/*Hind*III simultaneously for three-way cloning as described earlier for p2NIL, section 3.2.1.3. The expected pGEM vector size digested with *Asp*718 and *Hind*III was 3160 bp, which corresponded to what was observed on the gel as indicated with the red arrow, Figure 24. A 39 bp fragment was released from the abovementioned digestion and was visible as faint bands at the base of the gel. The upstream and downstream *rpf* regions were cloned into the *Asp*718 and *Hind*III sites of digested pGEM as previously described and illustrated in Figure 15, and the pGEM knockout constructs were mapped by restriction digestion to confirm that the cloning was indeed correct.

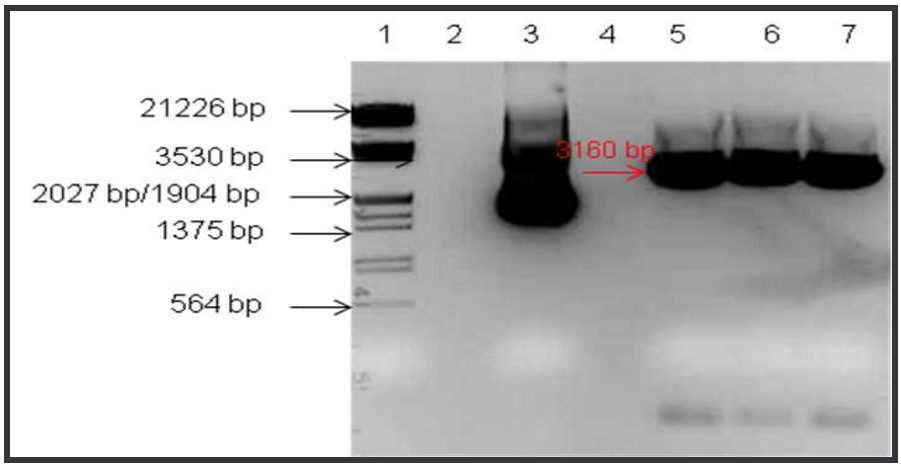



**Figure 24: pGEM plasmid vector restricted with appropriate enzymes in preparation for three-way cloning**. **Lane 1**: Molecular weight marker λIII, **lane 2**: no sample loaded, **lane 3**: uncut pGEM, **lane 4**: no sample loaded, **lane 5-7**: *Asp*718/*Hind*III digested vector.

**Table 8:** Mapping of pGΔRPFA construct by restriction digestion.

| Enzyme | Expected fragment sizes (bp) |
|---|---|
| *Pvu*I | 1022, 2274, 1352, 301 |
| *Pvu*II | 1185, 2820, 944 |
| *Sal*I | 275, 678, 3996 |
| *Nae*I | 538, 10, 3463, 938 |
| *Sac*I | 273, 33, 353, 722, 3568 |
| *Asp*718/*Hind*III | 1789, 3160 |

The integrity of both pGΔRPFA and pGΔRPFB constructs was confirmed by restriction mapping as seen in Figures 25 and 26; all digestions were correct although one of the fragments observed in the *Nae*I digestion in Figure 25, lane 3

gave a slightly smaller band than expected. It was surprising that the restriction mapping made sense in pGEM whereas it did not in p2NIL, considering that the same upstream and downstream *rpf* regions were used in both cloning events. This confirmed previous suspicions that the p2NIL vector was unstable.

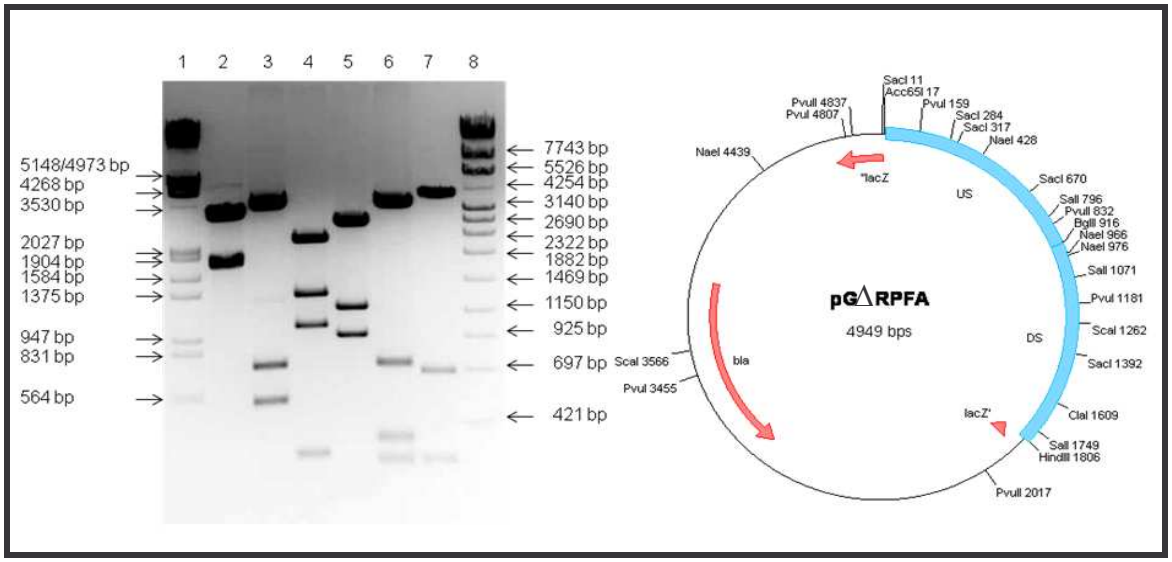



**Figure 25: Restriction mapping of pGΔRPFA construct shown alongside the knockout construct map.** Digestions: **Lane 1**: Molecular weight marker λIII, **lane 2**: *Asp7*18/*Hind*III, **lane 3**: *Nae*I, **lane 4**: *Pvu*I, **lane 5**: *Pvu*II, **lane 6**: *Sac*I, **lane 7**: *Sal*I, **lane 8**: λIV.

Construction of suicide vectors in pGEM was an attractive option rather than in p2NIL since the vector was stable however, several attempts were made to clone the *Pac*I cassette in pGΔRPFA and pGΔRPFB but proved unsuccessful. The pGEM strategy was abandoned and the p2NIL strategy was re-adopted. The cloned *rpf* regions in pGEM were digested out from the vector and cloned into p2NIL again, furthermore the *Pac*I cassette was cloned into those constructs.

Restriction mapping for the newly-made p2NIL knockout constructs (pNΔRPFA and pNΔRPFB) and suicide vectors (pSVΔRPFA and pSVΔRPFB) remained unchanged from previous cloning events.

**Table 9:** pGΔRPFB restriction mapping.

| Enzyme | Expected fragment sizes (bp) |
|---|---|
| *Pvu*II | 953, 670, 2820, 254 |
| *Pvu*I | 938, 2095, 1352, 312 |
| *Sac*II | 570, 416, 3711 |
| *Aat*II | 2279, 2418 |
| *Asp*718/*bgl*II | 822, 3875 |
| *Asp*718/*Hind*III | 1537, 3160 |

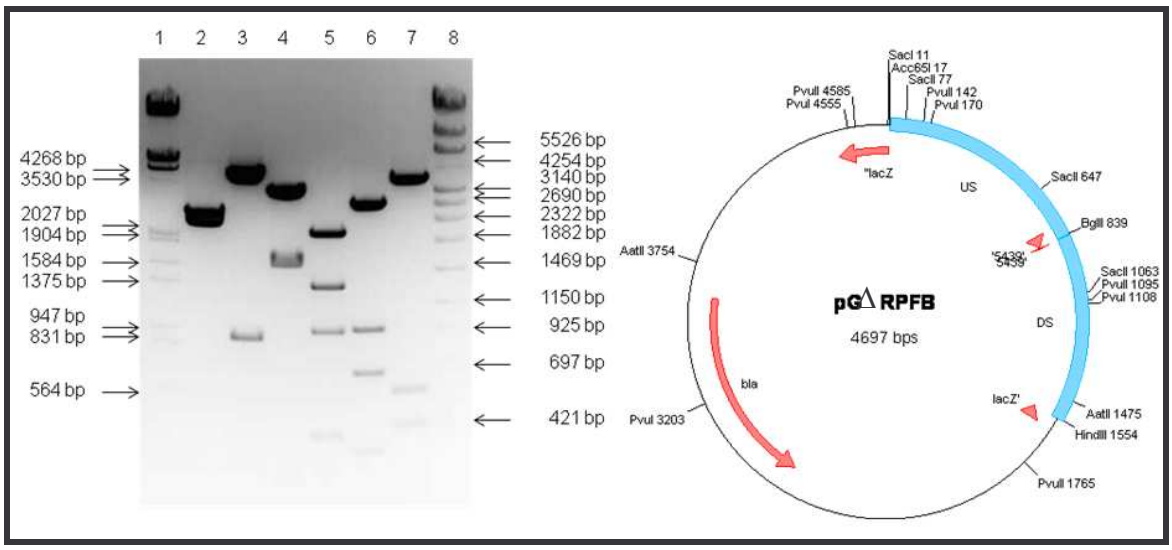



**Figure 26: pGΔRPFB restriction mapping shown with the construct map**. Digestions: **Lane 1**: Molecular weight marker λIII, **lane 2**: *Aat*II, **lane 3**: *Asp*718/*Bgl*II, **lane 4**: *Asp*718/*Hind*III, **lane 5**: *Pvu*I, **lane 6**: *Pvu*II, **lane 7**: *Sac*II, **lane 8**: λIV.

Construction of a suicide delivery vector for *rpf*CE (*MSMEG_4643/4643*), where both genes are deleted simultaneously, proved impossible with the use of the

same strategy utilized in the construction of suicide vectors for *rpfA* (*MSMEG_5700*) and *rpfB* (*MSMEG_5439*). Construction of another suicide vector for *rpfCE* (Appendix C, Figure A1) using a different method was undertaken (by another member of the laboratory), which proved to be successful and this construct was generously provided for use in this study.

### 3.2.2 Construction of single deletion mutants of *M. smegmatis* defective for *rpfA* and *rpfB.*

In this study, gene knockout was achieved by two step allelic exchange mutagenesis using a suicide vector carrying selectable and counter selectable markers. For this process, the vector is electroporated in *M. smegmatis* followed by the selection of blue, $Hyg^R$ and $Km^R$ colonies. These transformants represent single cross-over integrants which are then grown in the presence of sucrose, a counter selectable marker, to facilitate the second cross-over event. The *sacB* gene, present in the vector, encodes for levanesucrose which converts sucrose into a toxic metabolite resulting in bacterial death. Hence, in the presence of sucrose, bacterial death facilitates the second cross-over event to expel the vector and associated toxicity. It is important to note that this second cross-over can result in incorporation of the mutant allele, or reconstitution of the wild type allele, Figure 27. Hence, further screening of sucrose-resistant white colonies is imperative.

The pSVΔRPFA and pSVΔRPFB suicide delivery vectors were integrated into the genome of *M. smegmatis* $mc^2155$ by electroporation. The blue, $Hyg^R$ and $Km^R$ clones were selected post electroporation and grown in 10 ml of 7H9 broth to an OD600nm of ~1 and plated in the presence of sucrose for counter-selection. White, sucrose resistant colonies were selected and the success of allelic exchange mutagenesis was determined by PCR using allelic-specific primers. As shown in Figure 28 (for *rpfA*), three primers were used and added to the PCR reaction together. The mutant allele was amplified by primers F1 and R2 (since

R1 sequence is deleted in this genetic background) to yield a 160 bp fragment and the wild type allele was amplified by F1 and R1 to yield a 300 bp fragment.

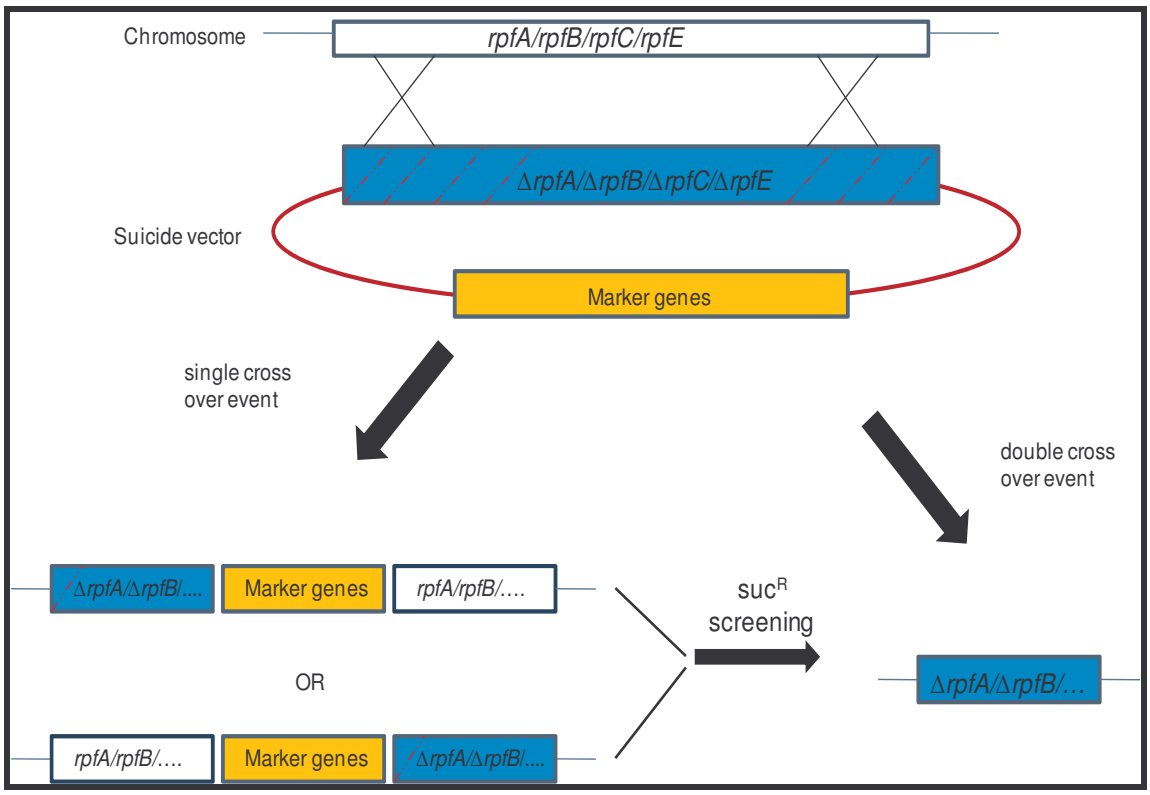
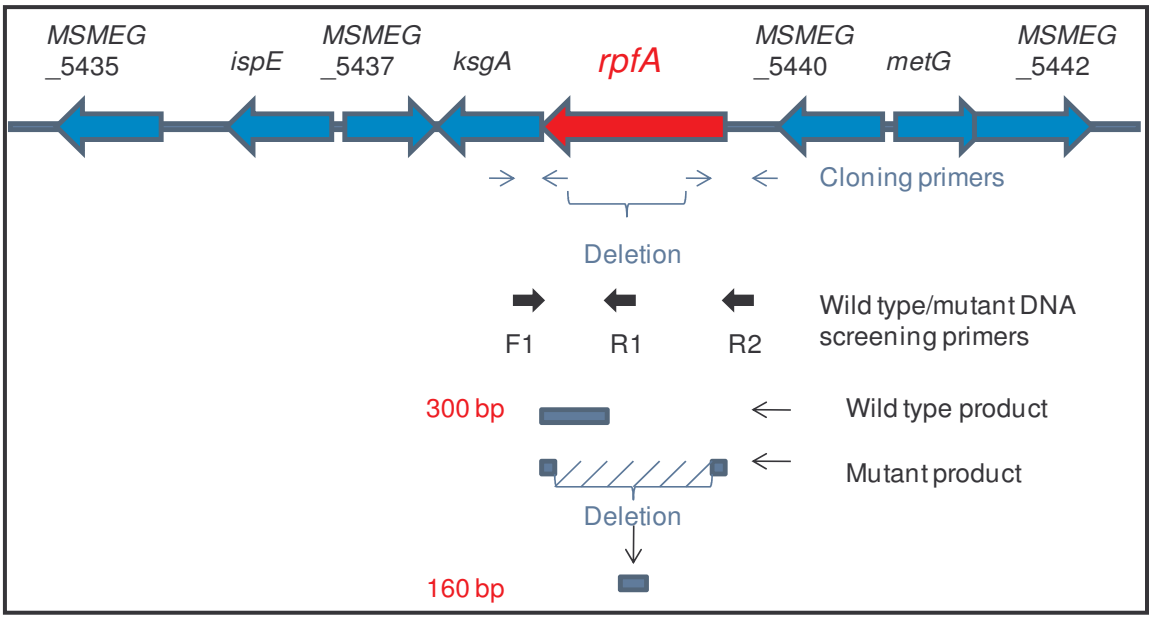



**Figure 27:** Construction of knockout mutants by two step allelic exchange mutagenesis.

As shown in Figure 28, the allelic exchange mutagenesis yielded several mutant strains, with no wild type band visible. The controls in gel A of Figure 28 made it clear which bands were wild type or mutant DNA, and in some instances both wild type and mutant bands were observed for a sample (gel A, lane 10 and gel B, lane 7,8,11,12) which was indicative of single cross-over strains. The mutant band was expected to be smaller since the screening primers would only amplify regions flanking the deletion. Five Δ*rpfA* mutants were identified from the screen.



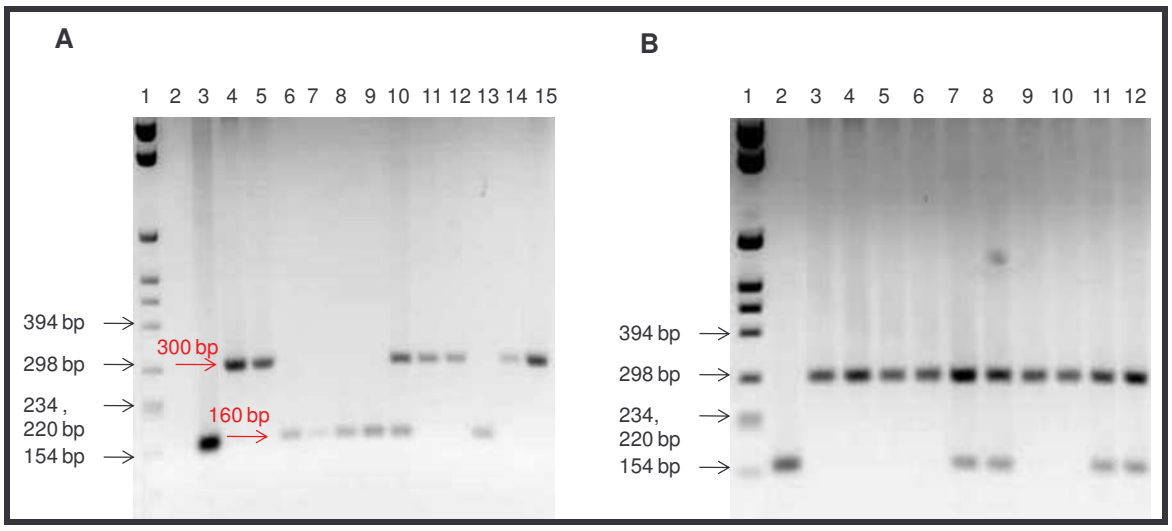
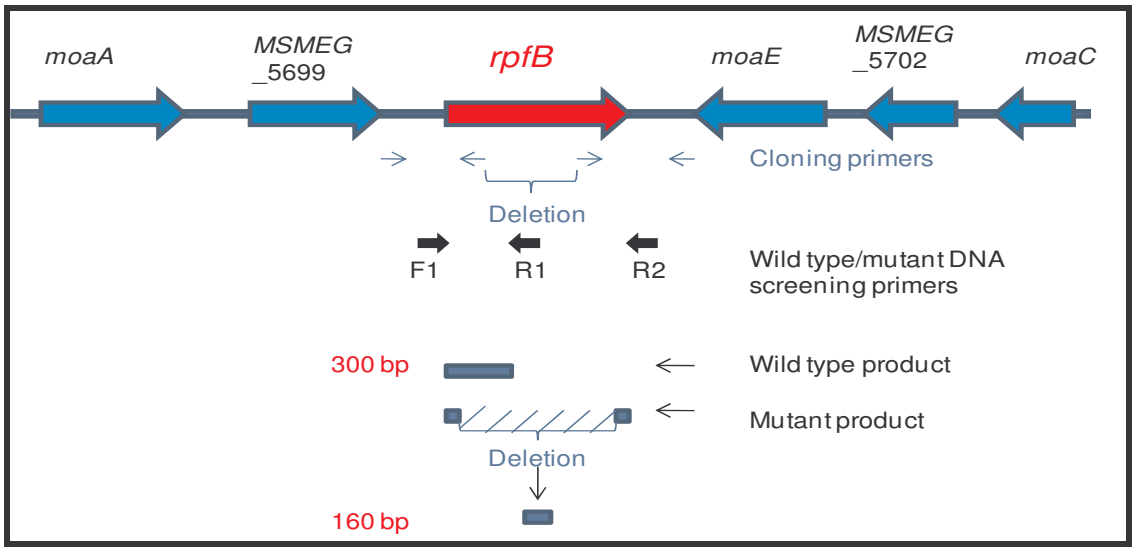




**Figure 28: Illustration of the screening process and identification of Δ*rpfA* mutants using *rpfA* allelic-specific screening primers following allelic exchange mutagenesis. (A) Lane 1:** Molecular weight marker λVI, **lane 2**: No DNA control, **lane 3**: pSVΔRPFA control, **lane 4**: wild type DNA control, **lane 5-15**: suc$^R$ *M. smegmatis* clone DNA. **(B**) Lane 1: λVI, **lane 2-12**: suc$^R$ *M. smegmatis* clone DNA.

A similar screening strategy used to screen for Δ*rpfA* mutants was applied for Δ*rpfB* mutants as explained above. The mutant and wild type allele DNA bands were expected to be 160 bp and 300 bp respectively; the former being amplified by F1 and R2 primers, whereas the F1 and R1 primers amplifying the latter. Screening for Δ*rpfB* mutants was performed twice, where only a single Δ*rpfB* mutant was identified in the initial screen (gels not shown), further screening yielded two additional mutants (Figure 29,gel B lane 7 and 9). The clones corresponding to both Δ*rpfA* and Δ*rpfB* mutant DNA identified, shown in Figures 28 and 29, were grown for future phenotypic characterization.



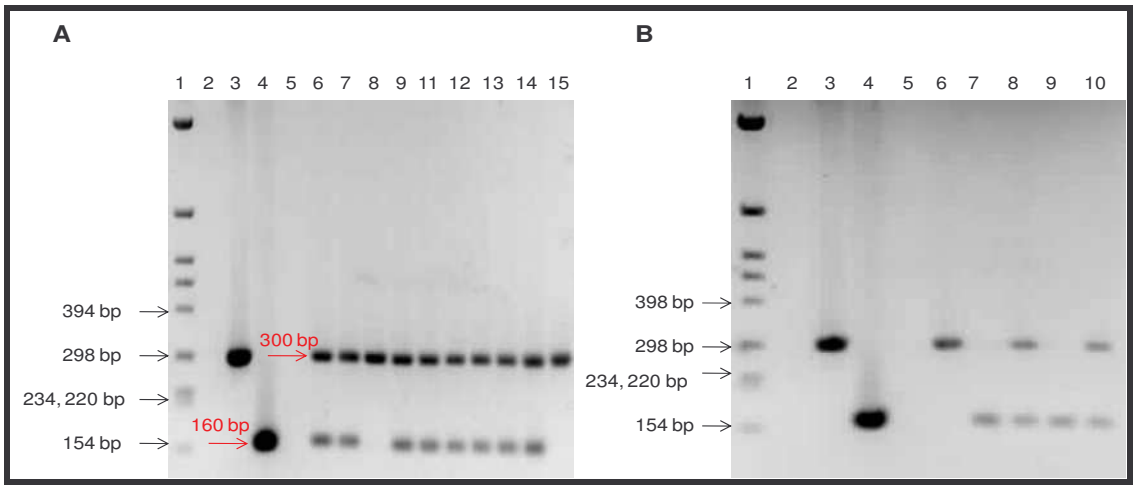



**Figure 29: The identification of Δ*rpfB* mutants resulting from the demonstrated screening strategy using *rpfB* allelic-specific screening primers, following allelic exchange mutagenesis. (A) Lane 1**: Molecular weight marker λVI, **lane 2**: No DNA control, **lane 3**: wild type DNA control, **lane 4**: pSVΔRPFB control, **lane 5**: no sample loaded, **lane 6-15**: $suc^R$ *M. smegmatis* clone DNA. **(B) Lane 1**: λVI, **lane 2**: No DNA control, **lane 3**: wild type DNA control, **lane 4**: pSVΔRPFB control, **lane 5**: no sample loaded, **lane 6-10**: $suc^R$ *M. smegmatis* clone DNA.

### 3.2.3 Construction of double and triple mutants of *M. smegmatis* defective for multiple *rpf*-like genes.

Allelic exchange mutagenesis with wild type *M. smegmatis*, $mc^2155$, yielded Δ*rpfA* and Δ*rpfB* single mutants. The multiplicity of *rpf* genes in *M. smegmatis* poses significant problems to dissecting individual function using gene deletion. Hence, strains defective for multiple *rpf* genes are required to understand collective or synergistic function. For this, the Δ*rpfB* deletion mutation was introduced into the Δ*rpfA* progenitor strain using the pSVΔRPFB suicide vector described above. The subsequent selection and PCR screening was conducted as described above. PCR screening, shown in Figure 30, of white $suc^R$ colonies to detect deletion of *rpfB*, in the Δ*rpfA* deletion strain, indicated that several double mutants were generated using this method, these are clones in gel A lanes 6,8,11,12,15 and gel B lanes 7, 8, 9,12 and 13. Since the electroporations were performed in the background of an *M. smegmatis* strain lacking *rpfA*; it was expected that the PCR analysis of all the positive clones would reveal a deletion in *rpfA*. All the positive clones identified in Figure 30A and Figure 30B were hence analyzed for deletion of *rpfA*. As expected, all of these clones retain a deletion in *rpfA*, Figure 30C.

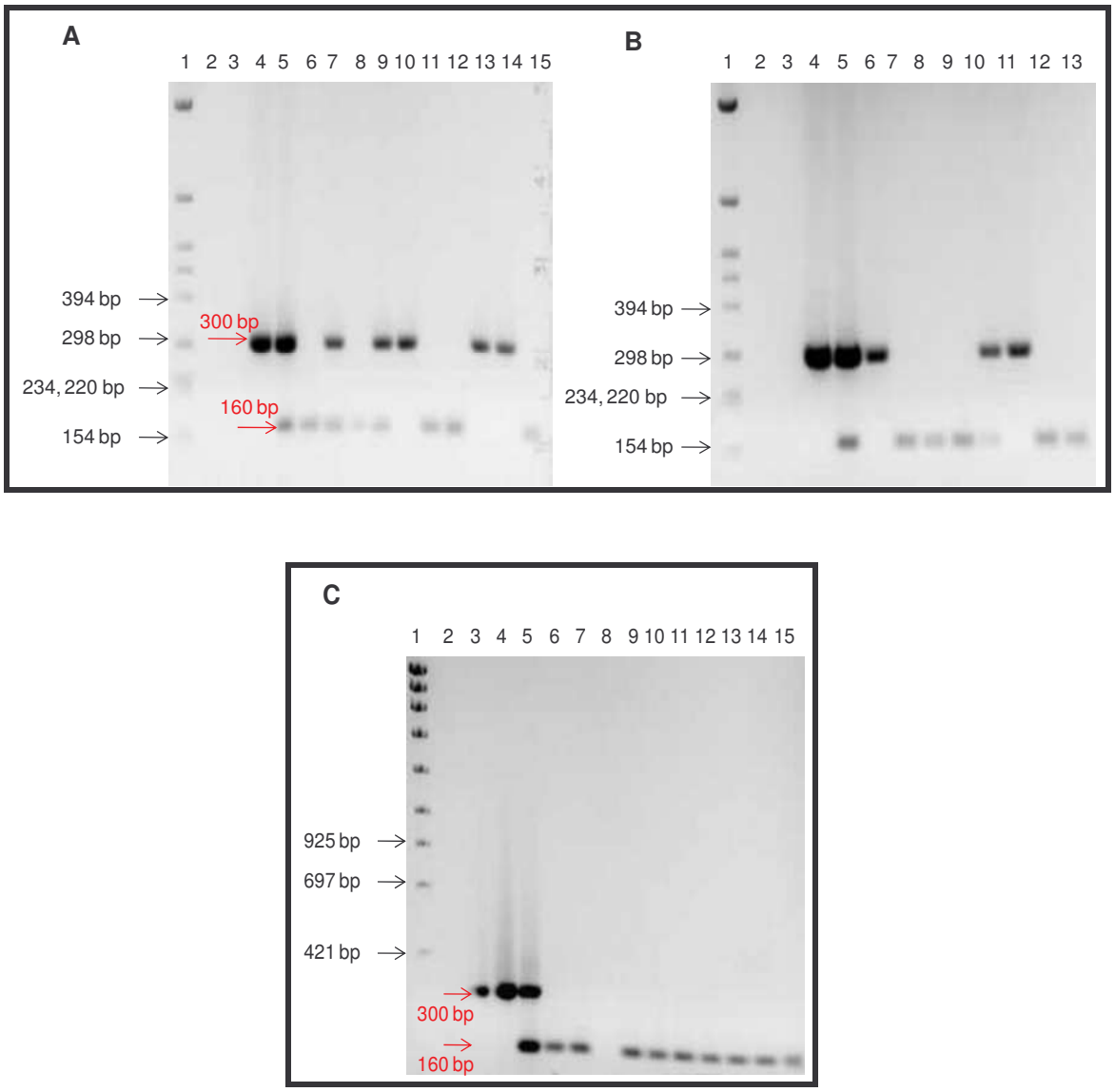



**Figure 30: Identification of Δ*rpfA* Δ*rpfB* double mutants with the use of *rpfB* allelic-specific screening primers in a Δ*rpfA* genetic background, following allelic exchange mutagenesis. (A) Lane 1:** Molecular weight marker λVI, **lane 2**: no sample loaded, **lane 3**: No DNA control, **lane 4**: wild type DNA control, **lane 5**: SCO (single cross-over) DNA control, **lane 6-15**: $suc^R$ *M. smegmatis* clone DNA. **(B) Lane 1**: λVI, **lane 2**: no sample loaded, **lane 3**: No DNA control, **lane 4**: wild type DNA control, **lane 5:** SCO DNA control, **lane 6-13**: $suc^R$ *M. smegmatis* clone DNA. **(C)** Confirmation of the

Δ*rpfA* genetic background using *rpfA* allelic-specific screening primers from identified clones with the *rpfB* gene deletion: **Lane 1:** molecular weight marker λIV, **lane 2**: No DNA control, **lane 3-4**: wild type DNA control, **lane 5**: SCO DNA control, **lane 6-15**: $suc^R$ *M. smegmatis* clone DNA.

Attempts to create suicide vectors for deletion of *rpfC* and *rpfE* were unsuccessful. A few problems were encountered in the attempted construction of the above suicide vectors; which included amplification of non-specific bands. Optimisation of PCR and re-designing primers for the abovementioned genes proved futile in efforts to resolve the amplification problem. It was, however, surprising that the amplicons obtained for both gene fragments were of the expected size yet the sequencing data did not validate the genes. However, a suicide delivery vector for *rpfC* and *rpfE* (Appendix C, Figure A1) was later constructed with the use of a different strategy, by Dr. Edith Machowski at the MMRU-CBTBR and was generously provided for use in this study to assist with creating further combinatorial deletion mutants. The above construct was introduced later in the study and integration of the vector into the wild type genetic background which would result in single *rpfC* and *rpfE* mutants was unsuccessful. No further attempts were made to obtain the single *rpfC* and *rpfE* mutants due to time constraints. Other combinatorial deletion mutants lacking *rpfC* and *rpfE* were successful and included in the study. This suicide vector was integrated in the Δ*rpfA* Δ*rpfB* double mutant followed by two step allelic exchange mutagenesis and screening as described above. PCR screening primers for *rpfC* and *rpfE* were designed in a manner similar to that for *rpfA* and *rpfB* and used for screening white, sucrose resistant colonies. The result of this screen, shown in Figure 31, indicate that six clones were identified to be *rpfC* deficient, Figure 31A, lane 8-12,15. Only 2 clones were found to be *rpfE* deficient as seen in Figure 31B, lane 7 and 13. Thus 8 triple *rpf* mutants were identified from the screen, 6 being Δ*rpfA* Δ*rpfB* Δ*rpfC* and 2 being Δ*rpfA* Δ*rpfB* Δ*rpfE* mutants, since these

genes were deleted in the background of a double Δ*rpfA* Δ*rpfB M. smegmatis* mutant strain.

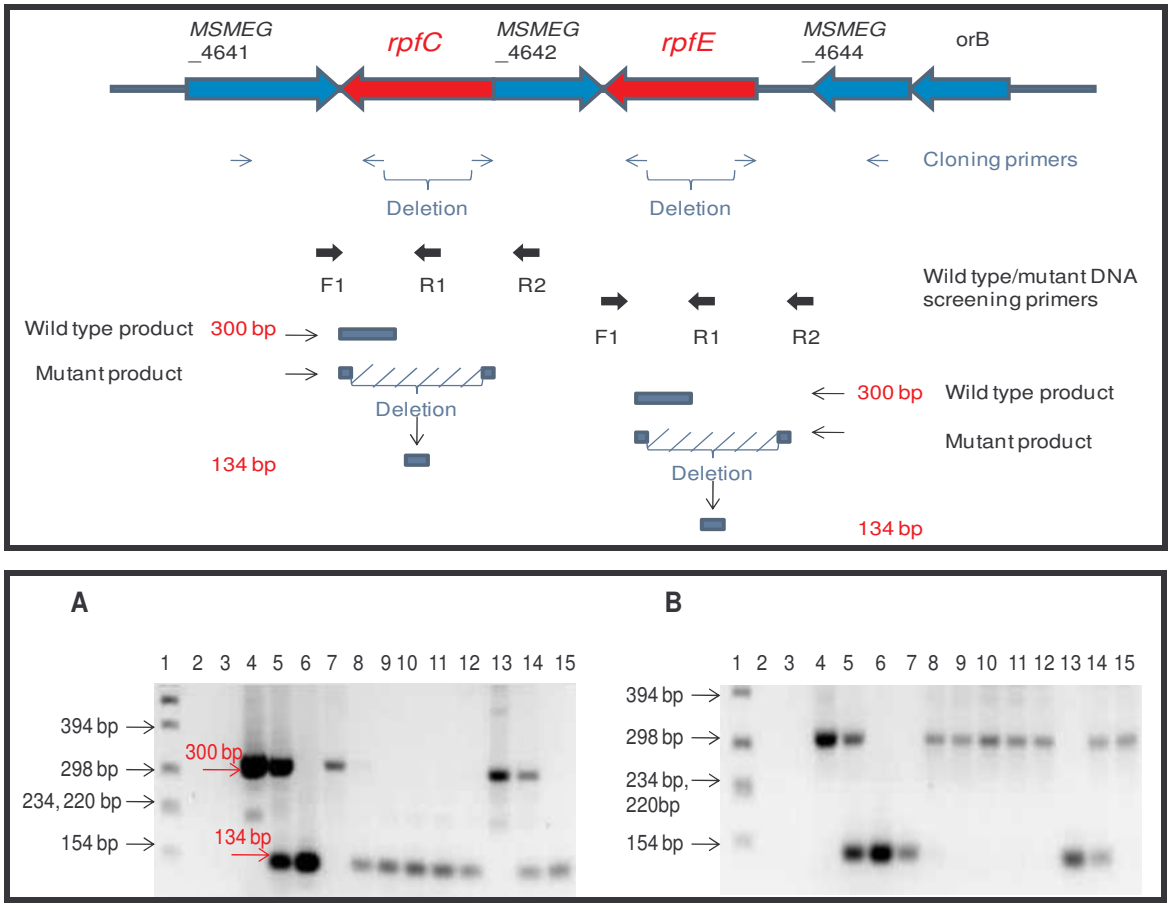


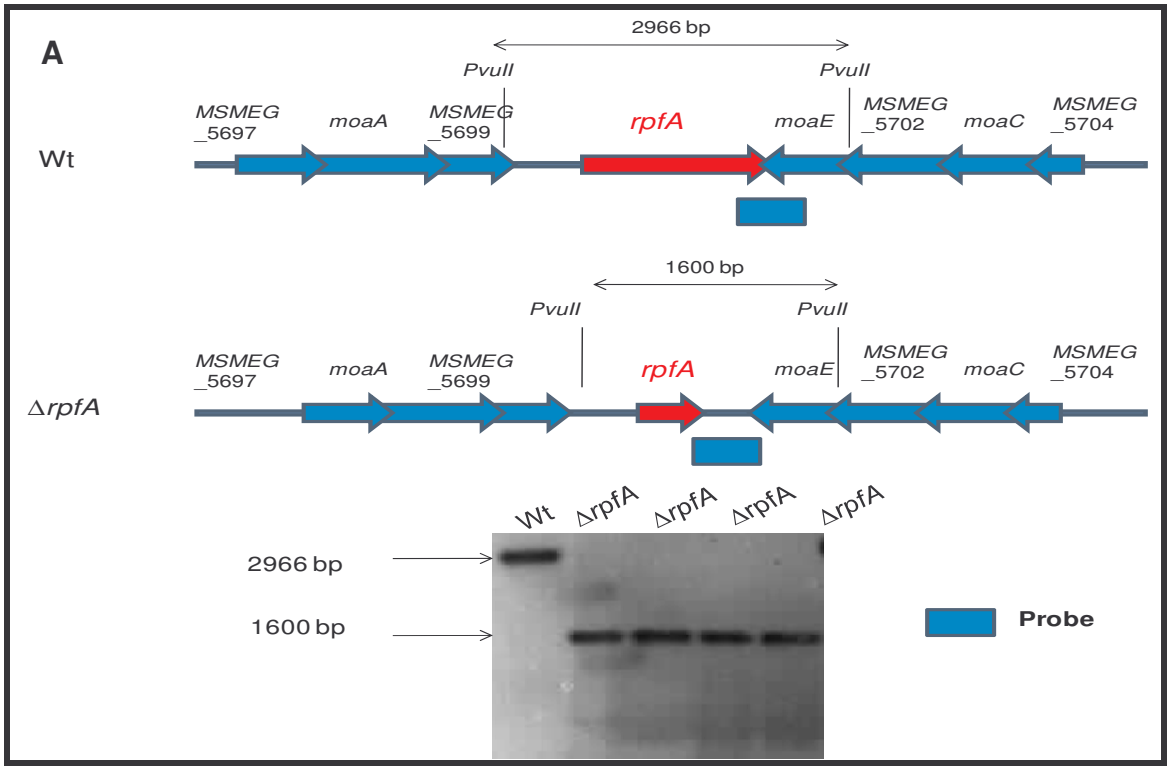

**Figure 31: The identification of *rpfC* and *rpfE* deficient genomic DNA in the Δ*rpfA* Δ*rpfB* double mutant genetic background using *rpfC* and *rpfE* allelic-specific primers following allelic exchange mutagenesis. (A)** PCR screen with *rpfC* allelic-specific primers**: Lane 1:** Molecular weight marker λVI, **lane 2**: no sample loaded, **lane 3**: No DNA control, **lane 4**: wild type DNA control, **lane 5**: SCO DNA control, **lane 6**: pSVΔRPFCE control, **lane 7-15**: $suc^R$ *M. smegmatis* clone DNA. **(B)** PCR screen with *rpfE* allelic-specific primers: **Lane** 1: λVI, **lane 2**: no sample loaded, **lane 3**: No DNA control, **lane 4**: wild type DNA control, **lane 5**: SCO DNA control, **lane 6**: pSVΔRPFCE control, **lane 7-15**: $suc^R$ *M. smegmatis* clone DNA.

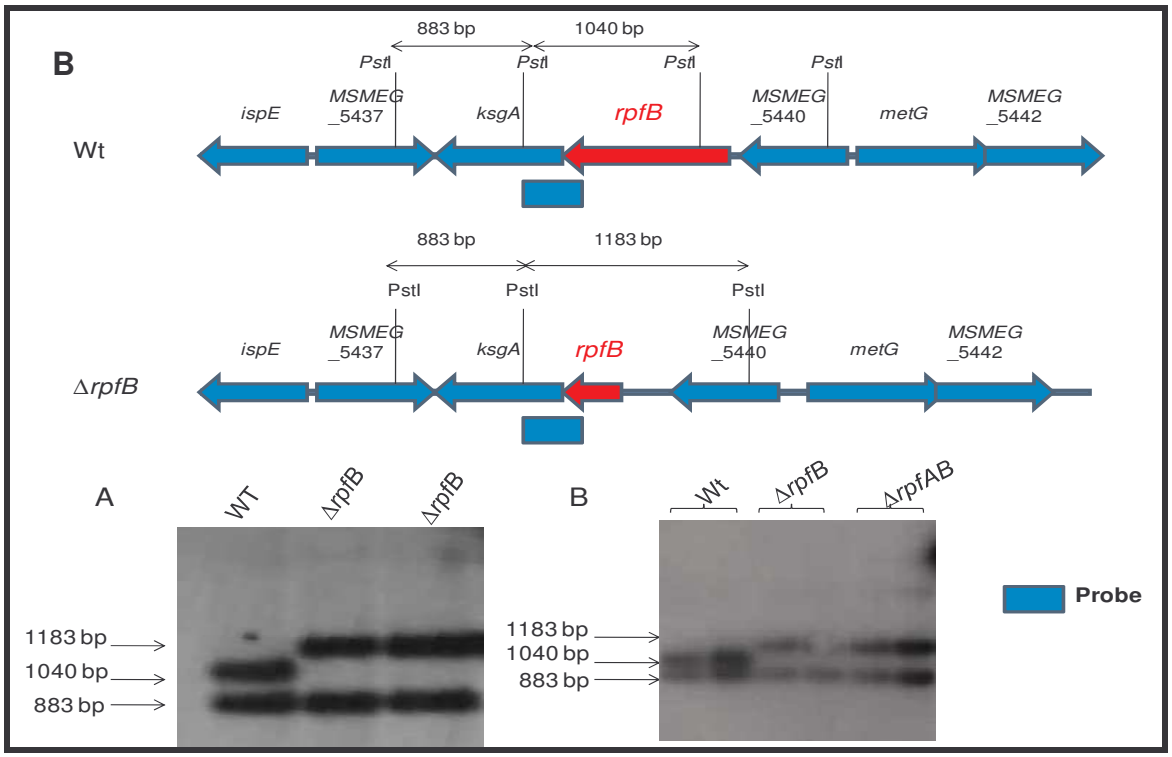
### 3.3 Southern blot analysis of mutant strains to further confirm genotype

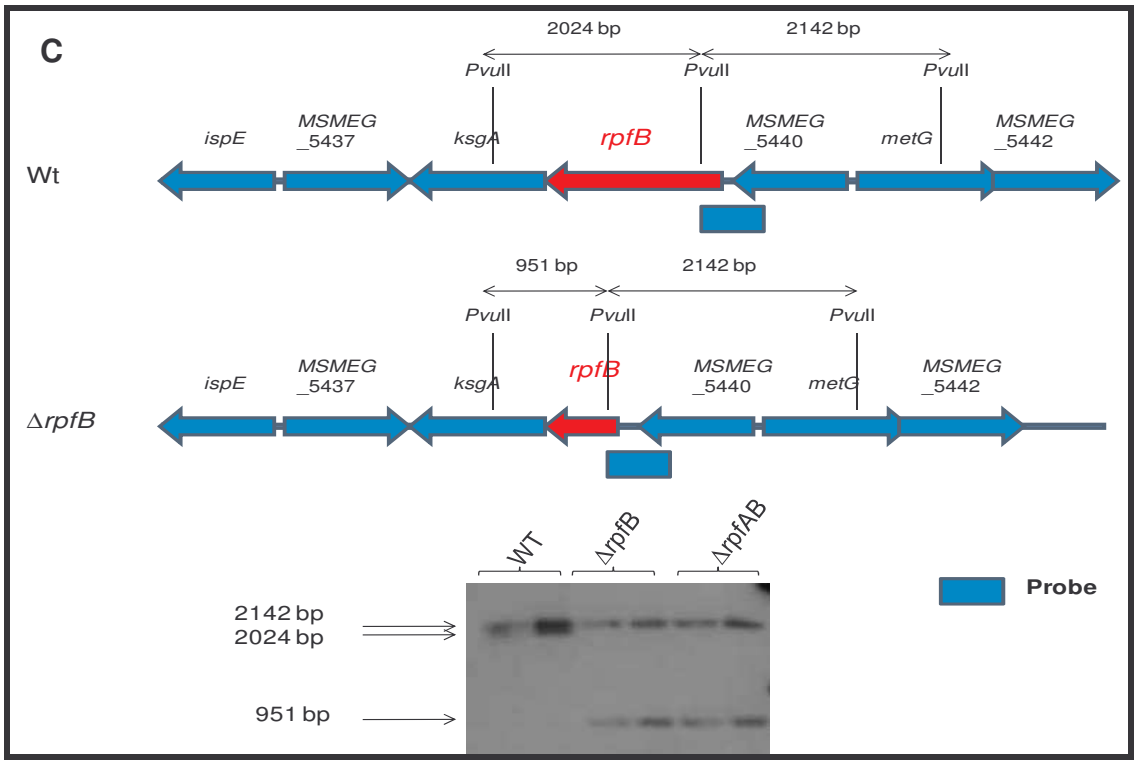
Southern blotting was necessary to validate the mutants identified in the PCR screening as these screens were insufficient alone to confirm that the mutants were indeed correct. It was therefore imperative to perform Southern blotting analysis using restriction sites outside of the regions of homology used for deletion mutagenesis to ensure site-specific integration of these mutational events. These analyses would confirm the integrity of the genomic region in question. Appropriate enzymes, with one of the sites being present outside the region of homology, were chosen and both wild type and the *rpf* mutant DNA was digested with the chosen enzymes. All regions, upstream and/or downstream, were analyzed and were confirmed to be indeed correct as all the expected fragment sizes corresponded to the observed fragments on the blots for both wild type and *rpf* mutants.

For *rpfA*, as shown in Figure 32A, in the wild type background the *Pvu*II sites are positioned 2966 bp apart at this locus and a 2966 bp fragment was expected from wild type genomic DNA, whereas a 1600 bp fragment was expected from the Δ*rpfA* mutant DNA since 1366 bp internal sequences were deleted from *rpfA*. Southern blotting with a downstream probe confirmed this genotype, Figure 32A. An upstream probe was used for *rpfB* genotyping; as shown in Figure 32B the enzyme chosen for this region was *Pst*I, which had 3 sites within the region. Two fragments were thus anticipated for wild type; 883 bp and 1040 bp fragments. However, one of the *Pst*I sites was positioned within the deleted *rpfB* internal sequences and was therefore lost in the *rpfB* deletion strain. In the *rpfB* mutant background; the 883 bp fragment was therefore expected since those sites were unaffected by the deletion but a slightly larger fragment (1183 bp) was generated due to loss of the abovementioned site, which extended over to a fourth downstream *Pst*I site in the DNA – southern blotting with the upstream region confirmed this genotype, Figure 32B. The *rpfB* deletion allele was further

genotyped with *Pvu*II, for the downstream *rpfB* region. As shown in Figure 32C, a 2142 bp common fragment was expected for wild type and *rpfB* mutant, while the 2024 bp fragment observed for the wild type was reduced in size in the mutant due to a 1073 bp deletion of internal sequences resulting in a 951 bp fragment. Hence, despite some vector rearrangement, the suicide vectors created in this study were successfully used to construct site-specific, allelic exchange mutants, with no disruption/rearrangement of upstream and downstream regions in the chromosome. Similarly, *Mlu*I was used for genotyping the *rpfC* downstream region, Figure 32D. In the wild type background, a 1468 bp fragment was expected since two *Mlu*I sites were positioned 1468 bp apart in this gene region. The fragment size was reduced in Δ*rpfC* mutant background due to 425 bp internal sequences being deleted.




B
883 bp
1040 bp
PstI
PstI
PstI
PstI
ispE
MSMEG _5437
ksgA
rpfB
MSMEG _5440
metG
MSMEG _5442
Wt
883 bp
1183 bp
PstI
PstI
PstI
ispE
MSMEG _5437
ksgA
rpfB
MSMEG _5440
metG
MSMEG _5442
ΔrpfB
A
WT
ΔrpfB
ΔrpfB
1183 bp
1040 bp
883 bp
B
Wt
ΔrpfB
ΔrpfAB
1183 bp
1040 bp
883 bp
Probe


C
2024 bp
2142 bp
PvuII
PvuII
PvuII
ispE
MSMEG _5437
ksgA
rpfB
MSMEG _5440
metG
MSMEG _5442
Wt
951 bp
2142 bp
PvuII
PvuII
PvuII
ispE
MSMEG _5437
ksgA
rpfB
MSMEG _5440
metG
MSMEG _5442
ΔrpfB
WT
ΔrpfB
ΔrpfAB
2142 bp
2024 bp
951 bp
Probe

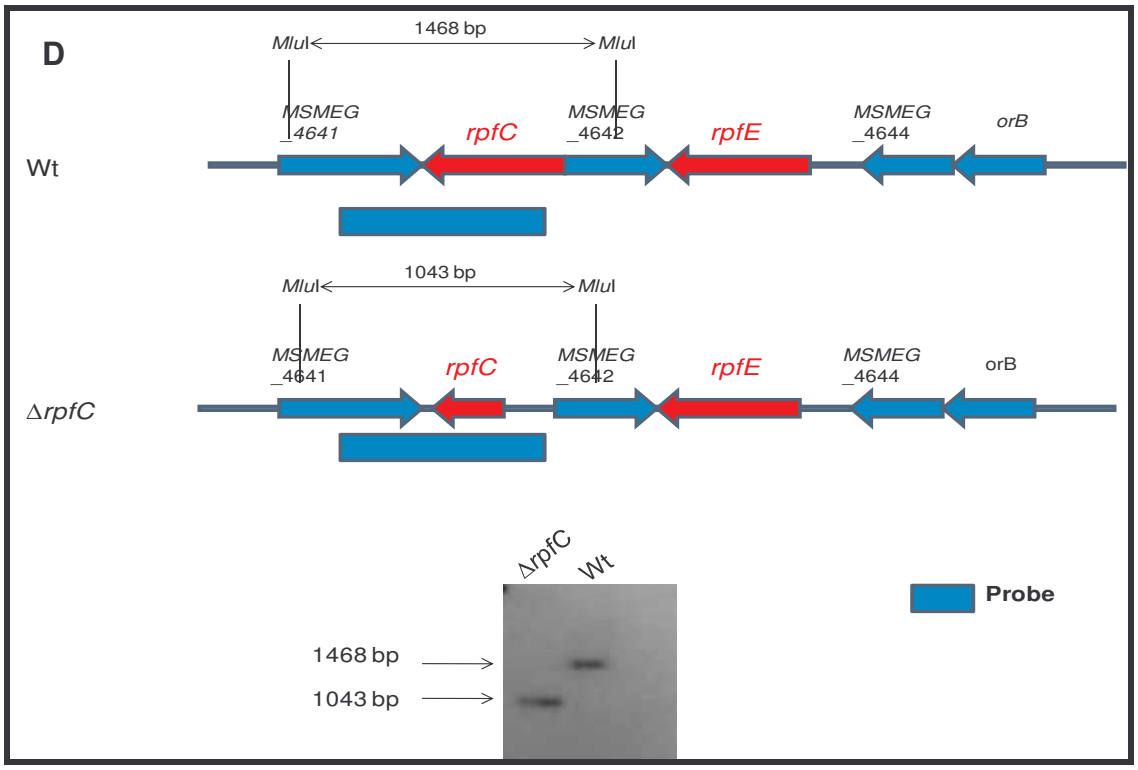



**Figure 32: Genotyping of the upstream and/or downstream regions of the *rpf* mutants by Southern blot analysis.** Genotyping**: (A)** Δ*rpfA* downstream region, **(B** and **C)** Δ*rpfB* upstream and downstream regions respectively, **(D)** Δ*rpfC* downstream region. Figures not drawn to scale.

### 3.4 Construction of complemented derivatives

A previously constructed complementation vector by a member of the MMRU was included for use in this study. The pMRPFAB construct (Appendix C, Figure A2) is essentially an integration vector carrying both *M. tuberculosis rpfA* and *rpfB* genes with their native promoter regions. The construct was introduced into the Δ*rpfA* Δ*rpfB* double mutant by electroporation and the transformants were selected on Hyg 50 µg/ml. A complemented derivative strain was thus constructed, Δ*rpfA* Δ*rpfB*::pMVRPFAB, with an anticipated wild type phenotype. Reverse-engineering the double mutant genotype to expected wild type

phenotypes served as confirmation that the observed mutant phenotypes were due to loss of these *rpf* genes.

### 3.5 Phenotypic characterization of the *rpf* single and multiple mutants

#### 3.5.1 Growth kinetics of the single and multiple *rpf* mutants

There is extensive evidence in the literature that deletion of *rpf* genes results in significant growth defects in other organisms. Hence, a comparative analysis of the growth characteristics of the wild type, *rpf* mutants and a complemented derivative was carried out. The different strains were grown at 37˚C in Middlebrook 7H9 broth supplemented with 0.02% tyloxapol to prevent clumping. Growth was monitored by optical density at 600nm ($OD_{600nm}$) readings and colony forming unit (CFU) enumeration, every 3 hrs over a 33 hr period. No significant differences in growth were observed between wild type and the single mutants as observed by $OD_{600nm}$ and CFU counts on solid media. However, the double Δ*rpfA* Δ*rpfB* and triple Δ*rpfA* Δ*rpfB* Δ*rpfC* mutants did display a marginal, statistically significant, extension of the lag phase as measured by $OD_{600nm}$. These differences observed early in the growth curve were also noted in the CFU analysis, suggesting that the Rpfs are required for growth stimulation in early axenic culture. However, it should be noted that these differences were marginal and only restricted to early growth since no significant differences were noted after 24 hrs of growth in culture.

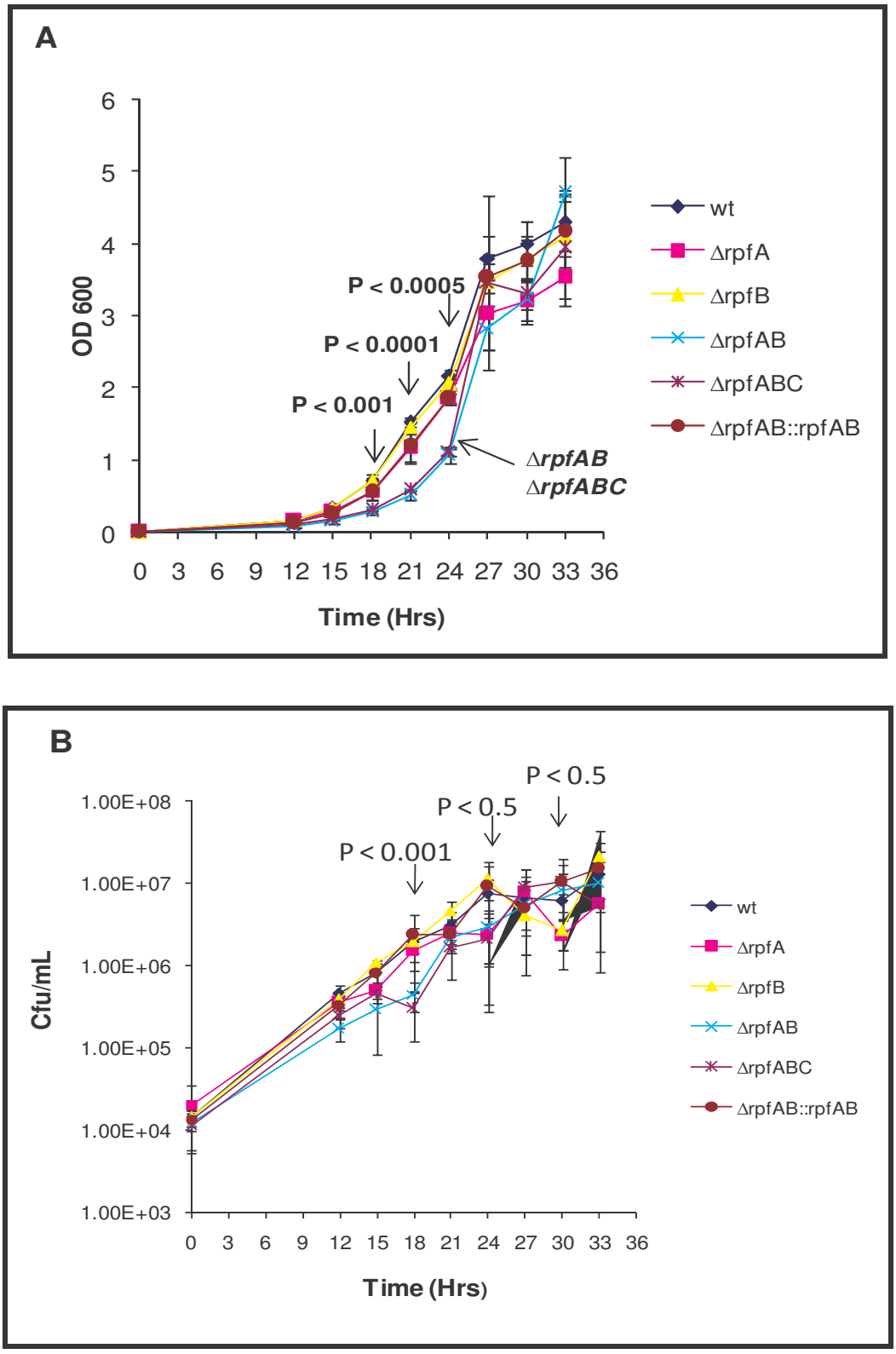



**Figure 33: Growth characteristics of wild type in comparison to the *rpf* mutants and a complemented derivative.** (**A**) Growth as determined by optical density

readings. (**B**) Cfu counts as a measure of growth. P-values were determined by the Student's *t*-test. The results shown are an average of three independent experiments.

The differences in growth attenuation obtained from the $OD_{600nm}$ readings did not correlate well with the CfU data reported in Figure 33B. In the CFU analysis, growth of the multiple deletion mutants was significantly different to wild type only at the eighteenth hr of growth as indicated on the graph. In general the CFU analysis gave a higher standard error when compared across experiments. This suggests that even though every attempt was made to standardize the experimental protocol and reagents, subtle differences between experiments may have affected the outcome. The complemented derivative (Δ*rpfA* Δ*rpfB::*pMVRPFAB), where functional copies of both *rpfA* and *rpfB* from *M. tuberculosis* were introduced into the Δ*rpfA* Δ*rpfB* mutant, grew comparably to wild type. This suggests that the marginal difference in growth obtained was due to deletion of *rpf* genes. In general, there was no observable growth phenotype for both the single *rpfA* and *rpfB* deletion mutants and the multiple mutants. What was observed to be a slightly prolonged lag phase by $OD_{600nm}$ readings, could have been an "apparent lag" although certain experimentation work is required to explore and confirm this.

### 3.5.2 Bacterial clumping in liquid culture associated with *rpfA* gene deletion

The slightly prolonged lag could have been due to bacterial clumping since it was noted that all strains lacking *rpfA*, displayed clumping in liquid cultures in stationary phase, Figure 34. The clumping (indicated by red arrows) observed for Δ*rpfA* Δ*rpfB* and Δ*rpfA* Δ*rpfB* Δ*rpfC* cultures was pronounced when compared to the single Δ*rpfA* mutant strain suggesting that progressive deletion of *rpf* genes in the Δ*rpfA* mutant background reinforces this phenotype. The Δ*rpfA* clumps were larger than those of the multiple mutants, which were observed smaller fine-textured clumps resembling bread crumbs. It is possible that micro-clumps develop very early in the growth curve but these were not detectable with our

standard methods. Aggregation in several actinomycetes is an intrinsic response to increased stress, which may be exogenous such as external stimuli or endogenous as exemplified in this study by loss of important genes. However, a comparison is made to wild type in the extent of clumping due to loss of *rpf* genes in the mutants.

The observed bacterial clumping phenotype was reversed back to wild type phenotype in the complemented derivative (Δ*rpfA* Δ*rpfB*::pMVRPFAB) as shown in Figure 34, indicating that the phenotype observed was indeed due to loss of the *rpf* genes.

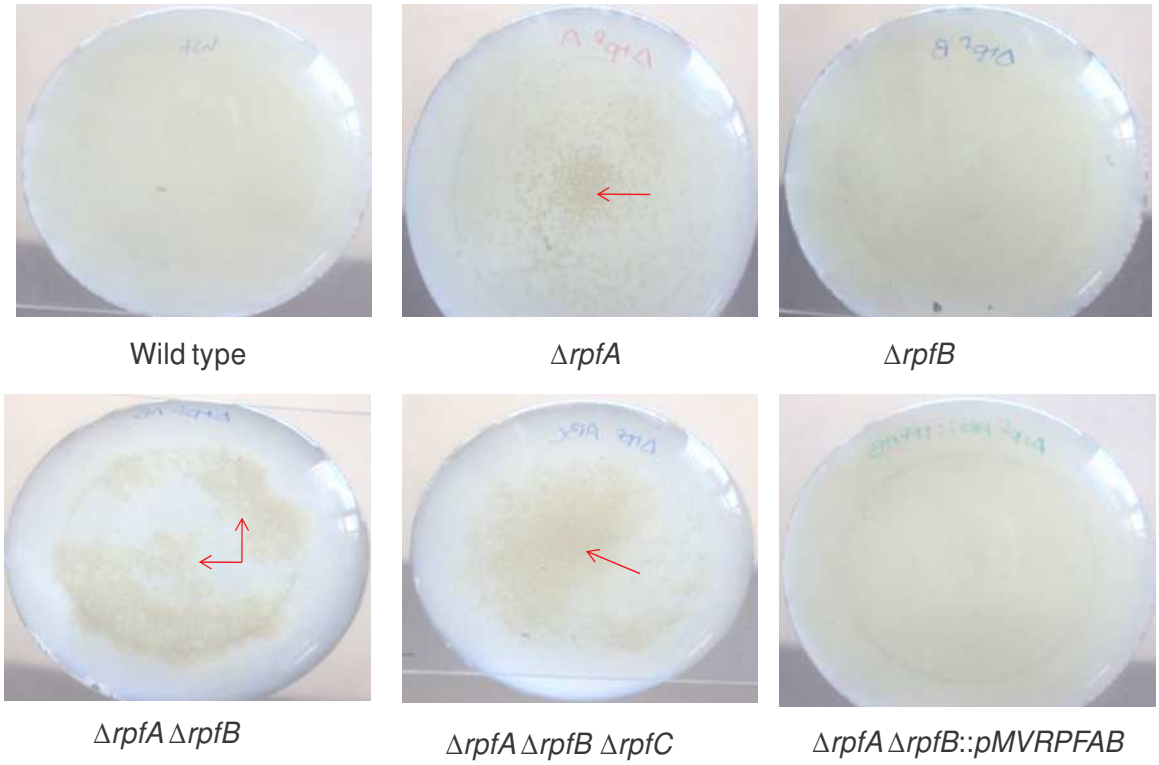
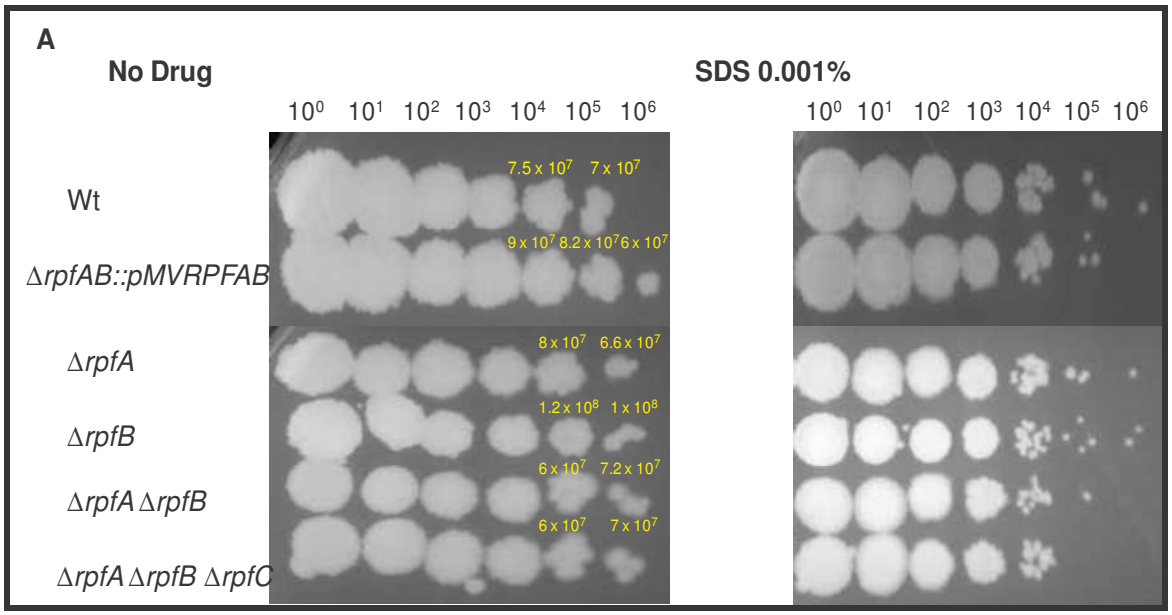
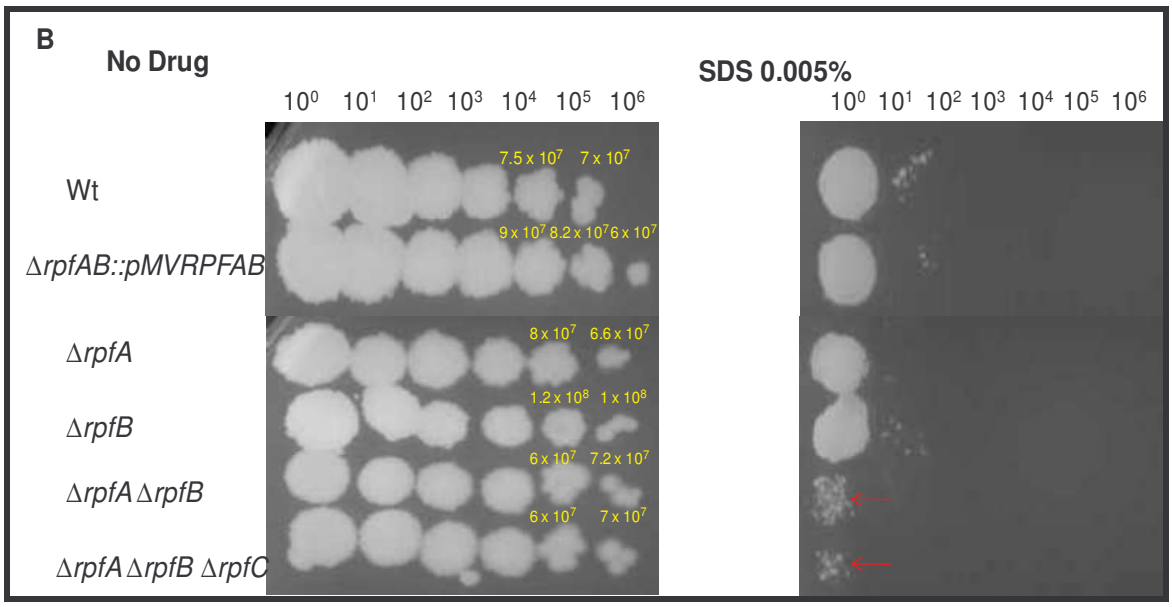


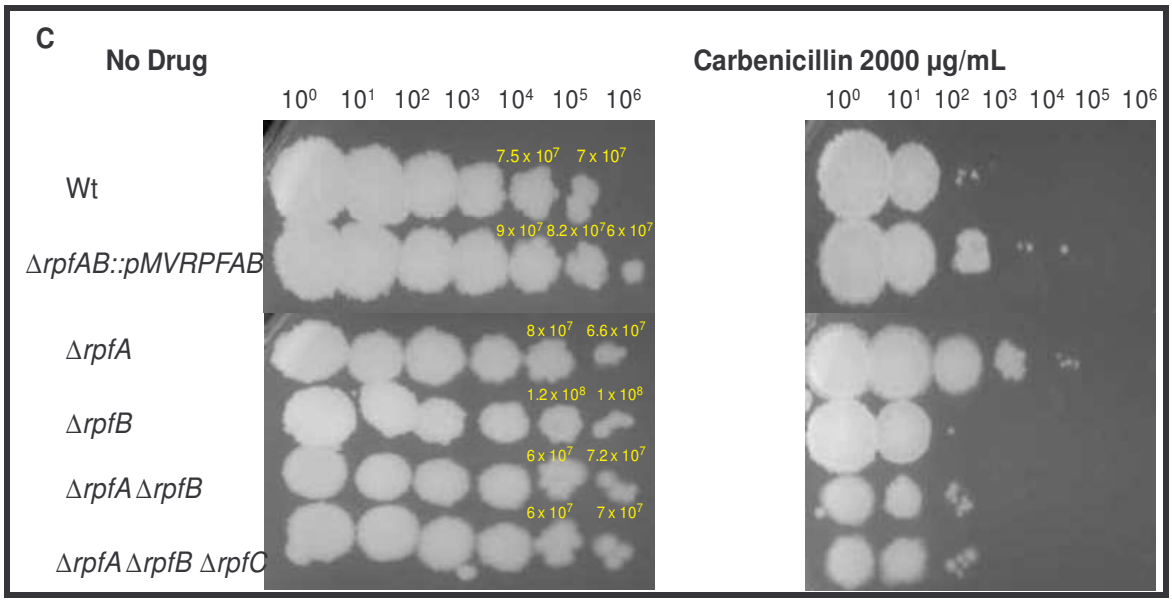

**Figure 34:** Culture flasks showing bacterial clumping in 33-hour stationery phase cultures of single Δ*rpfA* and multiple *rpf* mutants lacking an *rpfA* gene. This phenotype was observed consistently in all cultures grown in this study.

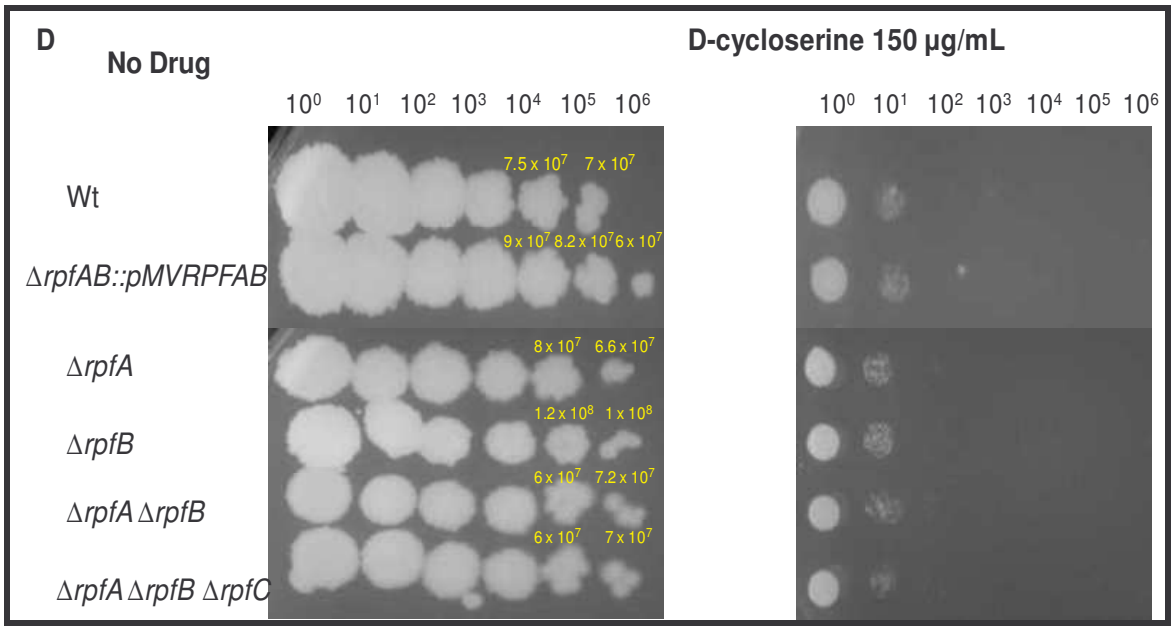
### 3.5.3 The multiple *rpf* deletion mutants display an increase in susceptibility to erythromycin and vancomycin

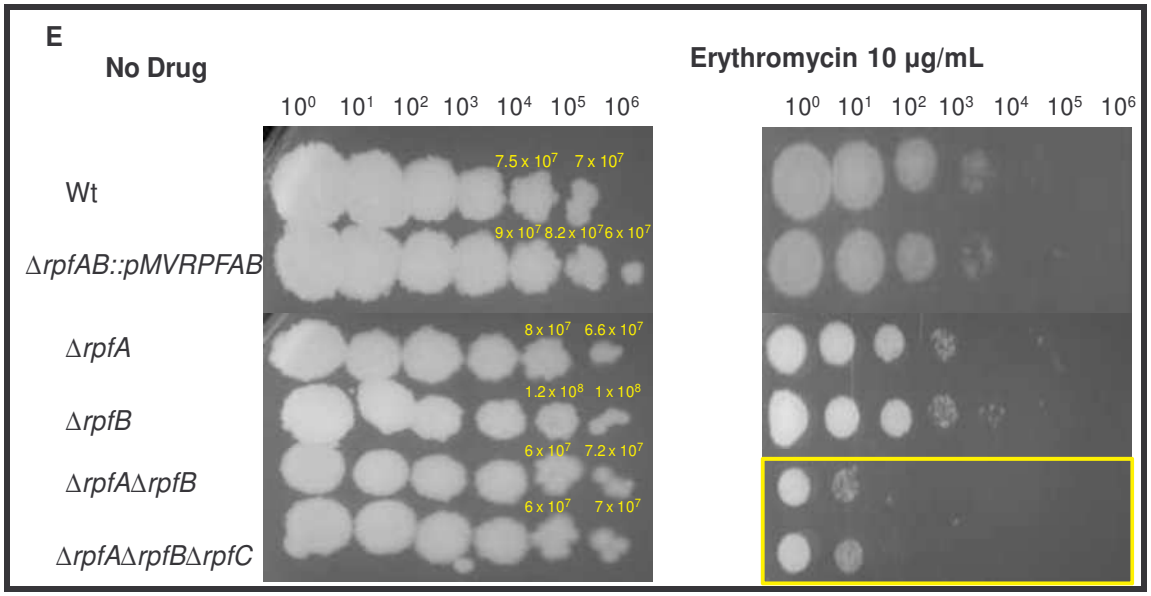
Previous analysis of *rpf* deletion mutants in *M. tuberculosis* indicated that deletion of multiple *rpf* genes leads to increase in susceptibility to select drugs, including vancomycin, erythromycin and detergent (SDS) (Kana*, et al.*, 2008, Kana*, et al.*, 2010). Hence, it was anticipated that deletion of one or more *rpf* genes in *M. smegmatis*, may render the resulting strains susceptible to similar drugs and detergents. Other cell wall-targeting drugs (carbenicillin, D-cycloserine) were also tested, since the *rpf* gene products act on the bacterial cell wall and deletion thereof may affect susceptibility to these cell wall-targeting antibiotics. Cultures of the six different strains, grown at $OD_{600nm}$ ~ 0.4, were tested for detergent and antibiotic sensitivity by serially diluting the culture and spotting 10 ul of each dilution on solid media containing drug/detergent, Figure 35.




B
No Drug
SDS 0.005%
10⁰ 10¹ 10² 10³ 10⁴ 10⁵ 10⁶
Wt
ΔrpfAB::pMVRPFAB
ΔrpfA
ΔrpfB
ΔrpfA ΔrpfB
ΔrpfA ΔrpfB ΔrpfC
7.5 x 10⁷ 7 x 10⁷
9 x 10⁷ 8.2 x 10⁷ 6 x 10⁷
8 x 10⁷ 6.6 x 10⁷
1.2 x 10⁸ 1 x 10⁸
6 x 10⁷ 7.2 x 10⁷
6 x 10⁷ 7 x 10⁷


C
No Drug
Carbenicillin 2000 µg/mL
10⁰ 10¹ 10² 10³ 10⁴ 10⁵ 10⁶
Wt
ΔrpfAB::pMVRPFAB
ΔrpfA
ΔrpfB
ΔrpfA ΔrpfB
ΔrpfA ΔrpfB ΔrpfC
7.5 x 10⁷ 7 x 10⁷
9 x 10⁷ 8.2 x 10⁷ 6 x 10⁷
8 x 10⁷ 6.6 x 10⁷
1.2 x 10⁸ 1 x 10⁸
6 x 10⁷ 7.2 x 10⁷
6 x 10⁷ 7 x 10⁷


D
No Drug
D-cycloserine 150 µg/mL
10⁰ 10¹ 10² 10³ 10⁴ 10⁵ 10⁶
Wt
ΔrpfAB::pMVRPFAB
ΔrpfA
ΔrpfB
ΔrpfA ΔrpfB
ΔrpfA ΔrpfB ΔrpfC
7.5 x 10⁷ 7 x 10⁷
9 x 10⁷ 8.2 x 10⁷ 6 x 10⁷
8 x 10⁷ 6.6 x 10⁷
1.2 x 10⁸ 1 x 10⁸
6 x 10⁷ 7.2 x 10⁷
6 x 10⁷ 7 x 10⁷


E
No Drug
Erythromycin 10 µg/mL
10⁰ 10¹ 10² 10³ 10⁴ 10⁵ 10⁶
Wt
ΔrpfAB::pMVRPFAB
ΔrpfA
ΔrpfB
ΔrpfAΔrpfB
ΔrpfAΔrpfBΔrpfC
7.5 x 10⁷ 7 x 10⁷
9 x 10⁷ 8.2 x 10⁷ 6 x 10⁷
8 x 10⁷ 6.6 x 10⁷
1.2 x 10⁸ 1 x 10⁸
6 x 10⁷ 7.2 x 10⁷
6 x 10⁷ 7 x 10⁷

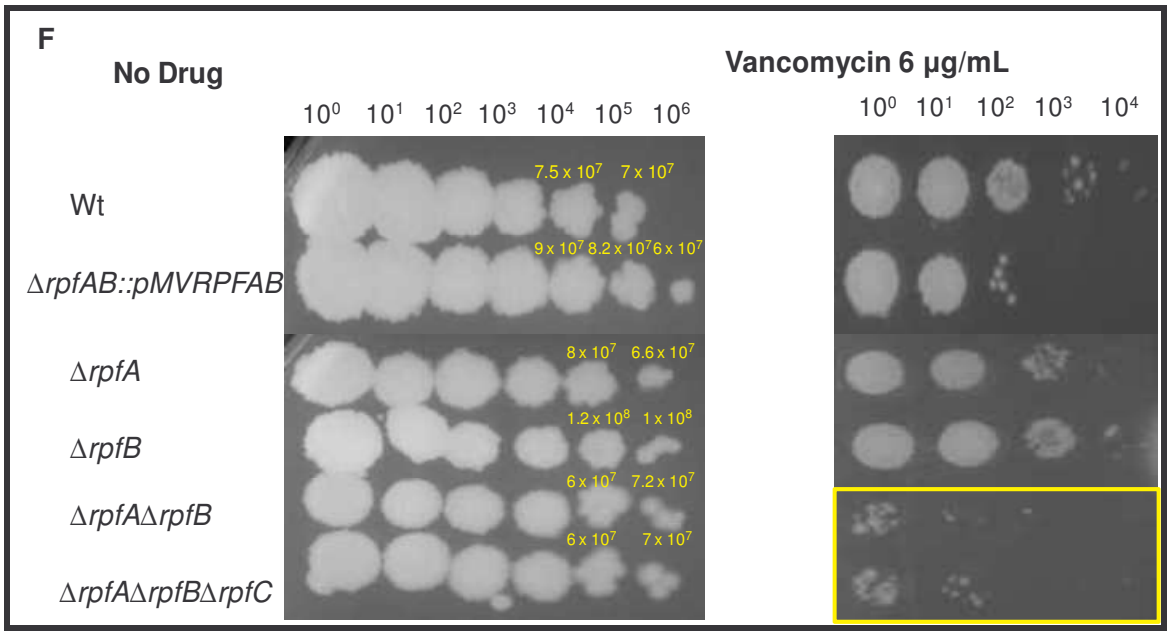



**Figure 35: Drug and detergent susceptibility testing of the wild type, *rpf* mutants and a complemented derivative**. **(A)** 0.001% SDS, **(B)** 0.005% SDS, **(C)** 2000 µg/ml carbenicillin, **(D)** 150 µg/ml D-cycloserine, **(E)** 10 µg/ml erythromycin, **(F)** 6 µg/ml vancomycin. Images shown after four days of incubation. The left and right panels represent the "no drug" control and drug or detergent testing respectively.

All *rpf* mutant and wild type cells grew comparably in SDS at low concentrations as seen in Figure 35A, thus it was decided to increase the concentration. SDS is an anionic detergent which inhibits bacterial growth, like all detergents, by protein denaturation and insertion into the cell membrane. The concentration was increased five-fold in order to create a slightly more stressful environment to assess differences in growth between wild type and the *rpf* mutants. The 0.005% SDS concentration proved sufficient to result in significant death of the wild type strain, Figure 35B. Under these conditions, the double and triple *rpf* deletion mutants *(ΔrpfA ΔrpfB and ΔrpfA ΔrpfB ΔrpfC*) displayed enhanced susceptibility, as indicated by red arrows in Figure 35B. In the undiluted spot, sparse growth

was observed for multiple mutants while a lawn was seen for wild type and single mutant cells; however this phenotype was abrogated with further incubation. The image was taken 4 days post-incubation and the observed phenotype disappeared following 2 days of further incubation.

No significant differences in susceptibility were observed between wild type and *rpf* mutants with 2000 µg/ml carbenicillin as seen in Figure 35C. Carbenicillin is a β-lactam antibiotic and acts to inhibit the process of cross-linking in the cell wall by binding to enzymes responsible for cross-linkage- a process important for cell wall synthesis. The concentration used was astoundingly high since mycobacteria in general are not highly susceptible to β-lactam antibiotics due the presence of β-lactamases (Hugonnet*, et al.*, 2009, Sala*, et al.*, 2009). Similarly, no significant differences were observed between all strains when tested on D-cycloserine, all the mutants grew comparably to wild type as shown in Figure 35D. D-cycloserine antibiotic acts by inhibiting cell wall synthesis but at very different levels to carbenicillin. The mode of action for this antibiotic is specific to inhibition of the formation and linking of D-alanine amino acid residues at the Lipid II peptide terminal end, Lipid II being a peptidoglycan precursor.

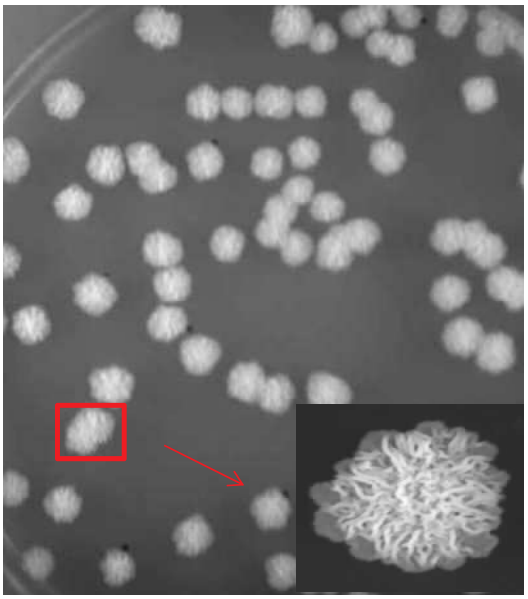
An approximate 2 log increase in susceptibility to 10 µg/ml erythromycin and 6 µg/ml vancomycin was observed for Δ*rpfA* Δ*rpfB* and Δ*rpfA* Δ*rpfB* Δ*rpfC*, as indicated by the yellow rectangles in Figure 35E and 35F respectively when compared to wild type. Erythromycin, a macrolide, is bacteriostatic unlike vancomycin and it generally inhibits protein synthesis by binding in the polypeptide exit tunnel of the large subunit of the ribosome. Vancomycin is a glycopeptide antibiotic which prevents cell wall synthesis due to cross-linking interference in the peptidoglycan layer of the bacterial cell wall (Schilling*, et al.*, 2011). A reduction in susceptibility to both erythromycin and vancomycin was observed in the complemented double mutant (Δ*rpfA* Δ*rpfB::pMVRPFAB*), to levels similar to wild type, confirming that the increased susceptibility in this case was due to loss of *rpf* genes. Under normal conditions; the multiple mutants grew

throughout to the $10^5$ dilution factor whereas under stressful conditions (antibiotics present), growth was only observed up to the $10^1$. Thus growth of the multiple mutants was inhibited by at least four logs under stressful conditions. This indicates the weakened cell wall of the cells due to multiple *rpf* loss.

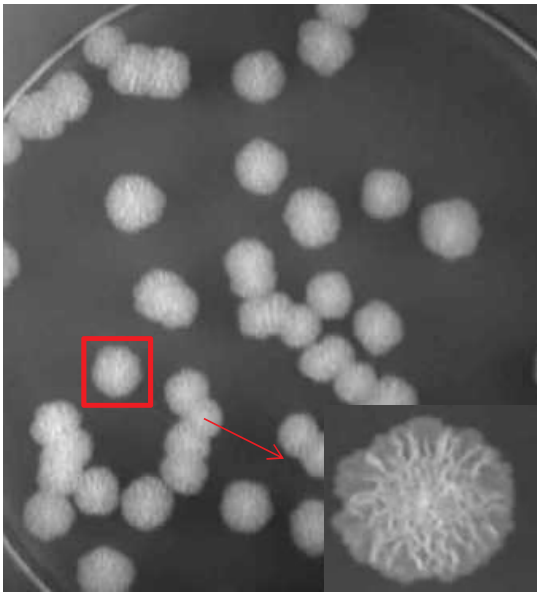
### 3.5.4 Deletion of multiple *rpf* genes results in alterations in colony morphology.

Since the *rpf* gene products act on the cell wall at the peptidoglycan layer, we postulated that deletion of *rpf* genes would result in changes in colony morphology. To test this, serially diluted cultures of all strains were spread on solid media and incubated for 5 days. The single mutant colonies, Δ*rpfA* and Δ*rpfB*, were comparable to wild type colonies since they retained the serpentine cording at the surface of the colony, which was marginally reduced in the Δ*rpfA* mutant, Figure 36. However, significant differences in cording and colony morphology were observed for the multiple mutants; Δ*rpfA* Δ*rpfB* and Δ*rpfA* Δ*rpfB* Δ*rpfC*. The cording effect was greatly reduced at the colony surface of Δ*rpfA* Δ*rpfB* and even more so with Δ*rpfA* Δ*rpfB* Δ*rpfC* as shown in the inserts of Figure 36. These colonies retained a smoother appearance with a halo-like periphery, and papillae at the centre.

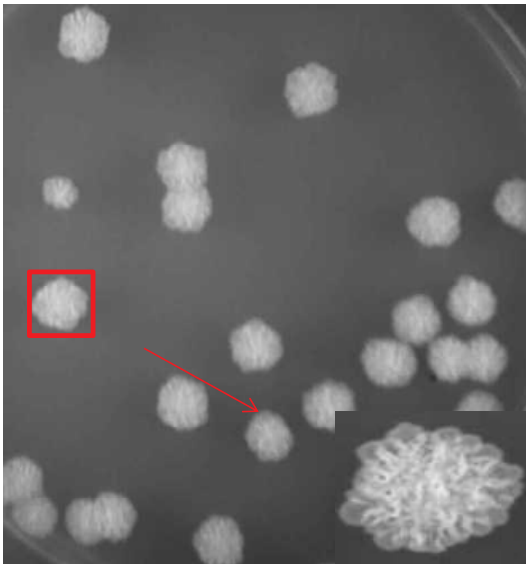
Wt



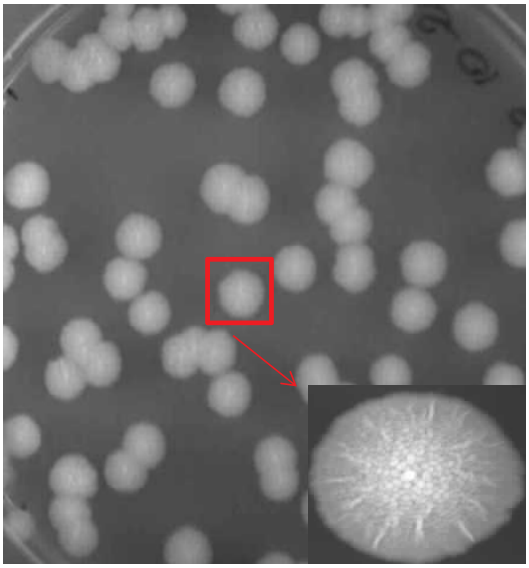
Δ*rpfA*



Δ*rpfB*



Δ*rpfA* Δ*rpfB*



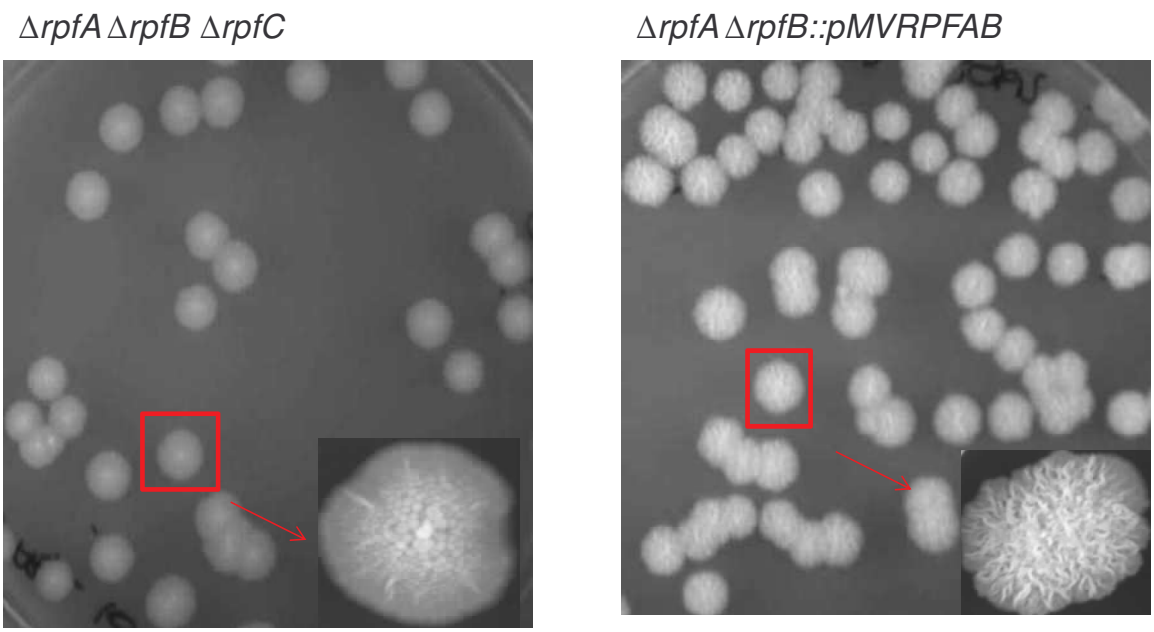



**Figure 36: Colony morphological changes as a result of successive *rpf* gene deletion. These phenotypes were consistently observed throughout the study**. Logarithmic phase cultures were serially diluted and plated on Middlebrook 7H10 agar and incubated for 4 days.

When Δ*rpfA* Δ*rpfB* was complemented with both *rpfA* and *rpfB M. tuberculosis* genes, Δ*rpfA* Δ*rpfB*::*pMVRPFAB*, the observed phenotype was reversed to wild type phenotype confirming that the phenotype was due solely to deletion of *rpfA* and *rpfB* genes. These results confirm that the *rpf* gene products have specific biological roles within the bacterial cell wall.

### 3.5.5 Biofilm formation defects associated with multiple *rpf* gene deletions

The changes in colony morphology, combined with the clumping of some strains in media prompted us to assess if there were any defects in biofilm formation for *rpf* deletion mutants. Biofilm formation requires complex interactions and signaling between the bacterial cells, thus the cell wall is an important factor in

the process. Biofilms were set up in tissue culture multi-well plates consisting six large wells, and each well contained 10X diluted culture at $OD_{600nm}$~ 0.5 for each strain. All strains were standardized for inoculation by ensuring that equivalent starting OD600nm readings were accurate. Biofilm formation in this assay is measured by the formation of lush growth at the liquid-surface interface as seen in Figure 37 for the wild type strain (Ojha*, et al.*, 2008).

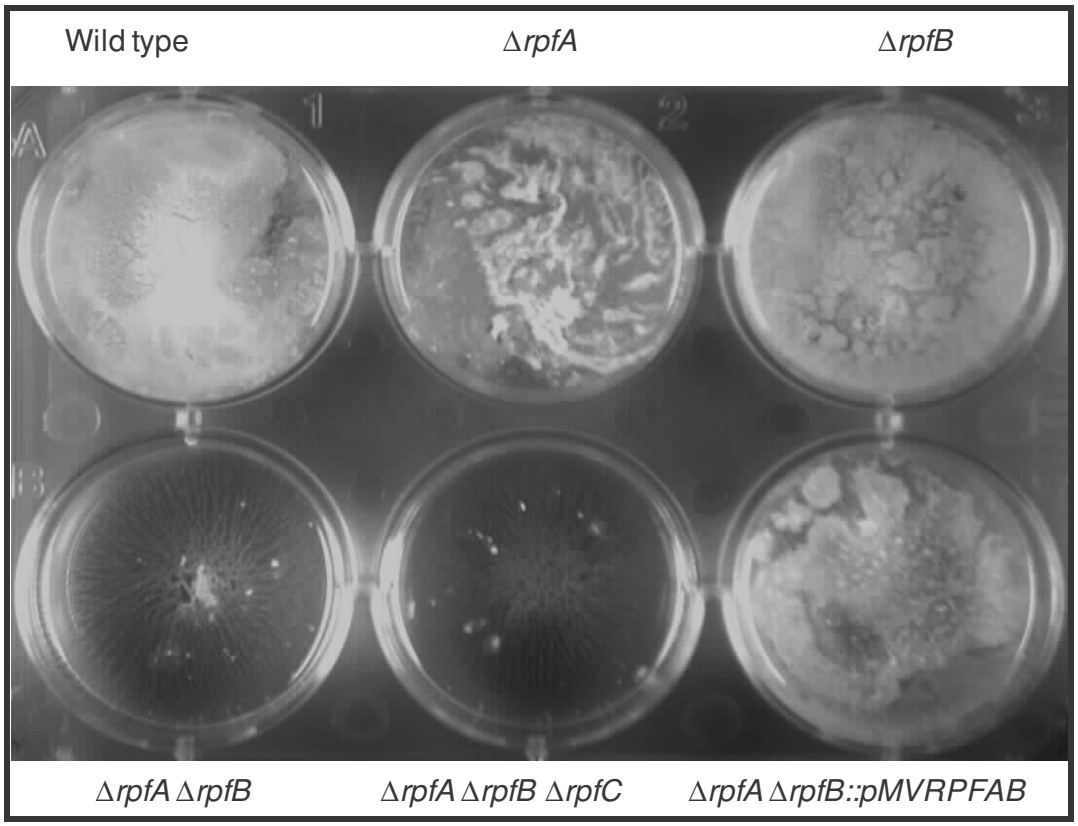



**Figure 37: Retarded biofilm formation in multiple *rpf* mutants.** A 1:10 diluted log-phase culture of the above strains was inoculated into fresh Middlebrook 7H9 medium and images were taken 7 days post-inoculation. The results observed represent three independent experiments.

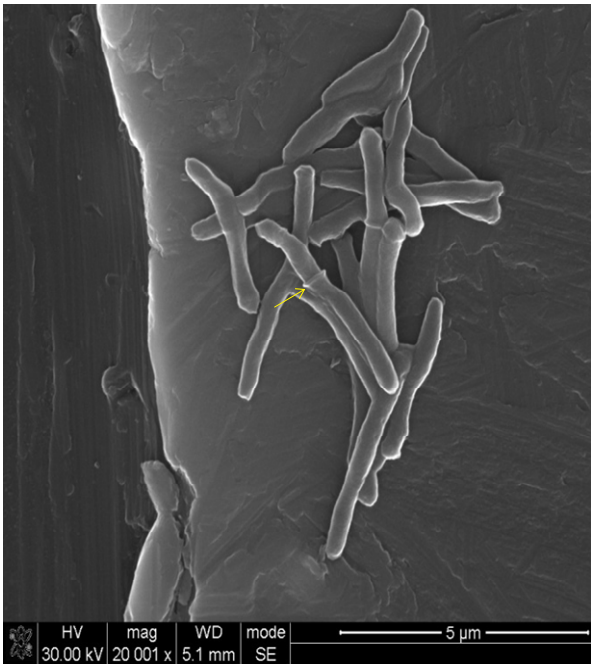
As seen in Figure 37, biofilm formation in the single mutants *(*Δ*rpfA* and Δ*rpfB*) was comparable to that observed for wild type, whereas Δ*rpfA* Δ*rpfB* and Δ*rpfA* Δ*rpfB* Δ*rpfC* mutants displayed significantly retarded biofilm formation. The

complemented derivative, Δ*rpfA* Δ*rpfB::pMVRPFAB*, displayed a phenotype similar to that of the wild type, confirming the observed double mutant phenotype to be due solely to deletion of *rpfA* and *rpfB* genes.

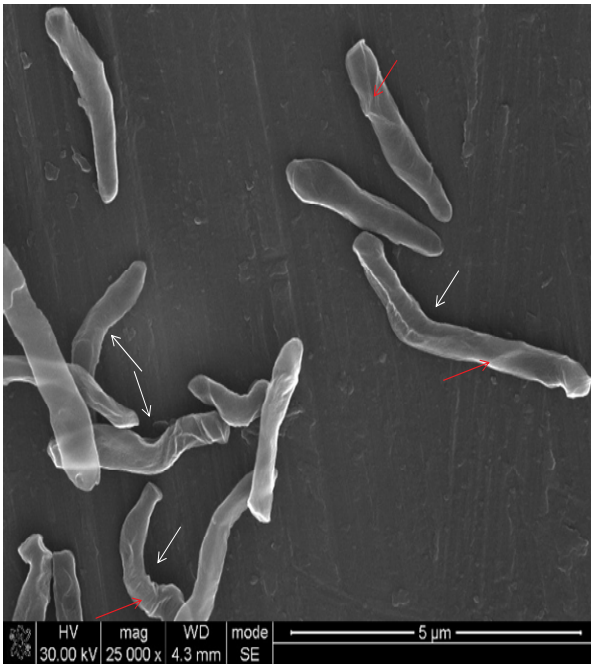
### 3.5.6 Cell morphological deformities resulting from multiple *rpf* deletion

Previous observations from colony morphology alterations in the multiple *rpf* mutants, Figure 36, led us to speculate that such changes could be evidenced at the cellular level. Furthermore, the observed clumping in select (*rpfA* deficient strains) cultures also prompted the examination of all strains at the cell surface level using Scanning Electron Microscopy (SEM). It was initially suspected that all *rpfA*-depleted strains displayed a filamentous morphology since clumping in culture is often associated with filamentation (Hett*, et al.*, 2008, Nguyen*, et al.*, 2010); however this was not the case as seen in Figure 38. Cultures were grown to $OD_{600nm}$ ~ 0.5 from overnight pre-cultures and subsequently prepared for SEM. *M. smegmatis* $mc^2155$ (wild type) cells appear on electron micrographs as straight rods with a septum located at an asymmetrical position along the cell length. The septum is a protruding ridge on the cell's surface and functions as a site of cell division (Dahl, 2004). The Δ*rpfB* mutant resembled wild type cells but also appeared slightly longer than wild type. However, the Δ*rpfA* mutant displayed a rough-like or even "wrinkled" appearance at its surface as indicated by red arrows, furthermore these cells were bent (white arrows) and had slight bulges at a single pole. These phenotypes were exacerbated in the multiple mutants; moreover Δ*rpfA* Δ*rpfB* and Δ*rpfA* Δ*rpfB* Δ*rpfC* mutant cells were fragile as was evidenced by the presence of broken cells (green arrows). Preparation of these cells could have compromised the cell walls during the dehydration process and made them brittle; however, since all strains were prepared under identical conditions, the increased breakage in multiple *rpf* deletion mutants suggests changes in cell wall stability. The septa in these strains were not well defined like wild type cells but this is probably due to cell surface defects.

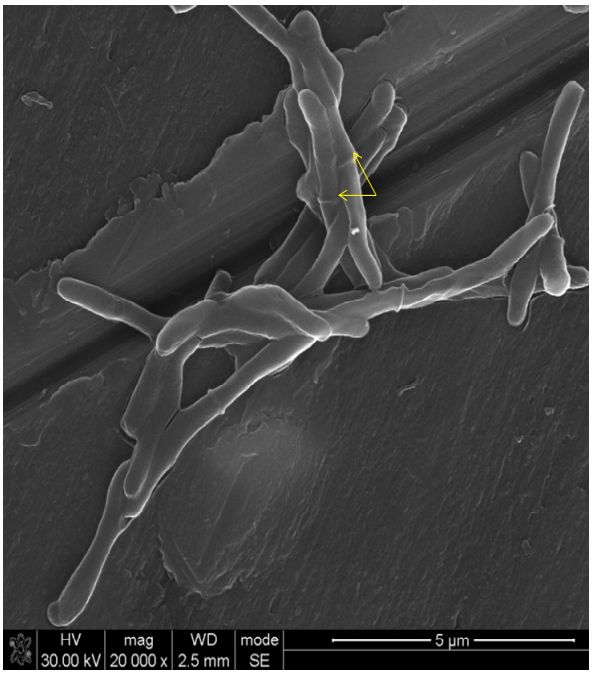
Wild type



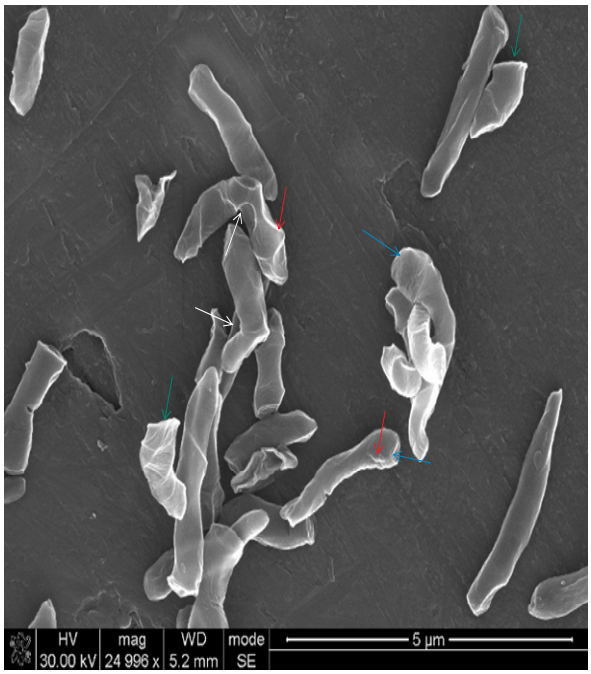

Δ*rpfA*




Δ*rpfB*




Δ*rpfA*Δ*rpfB*



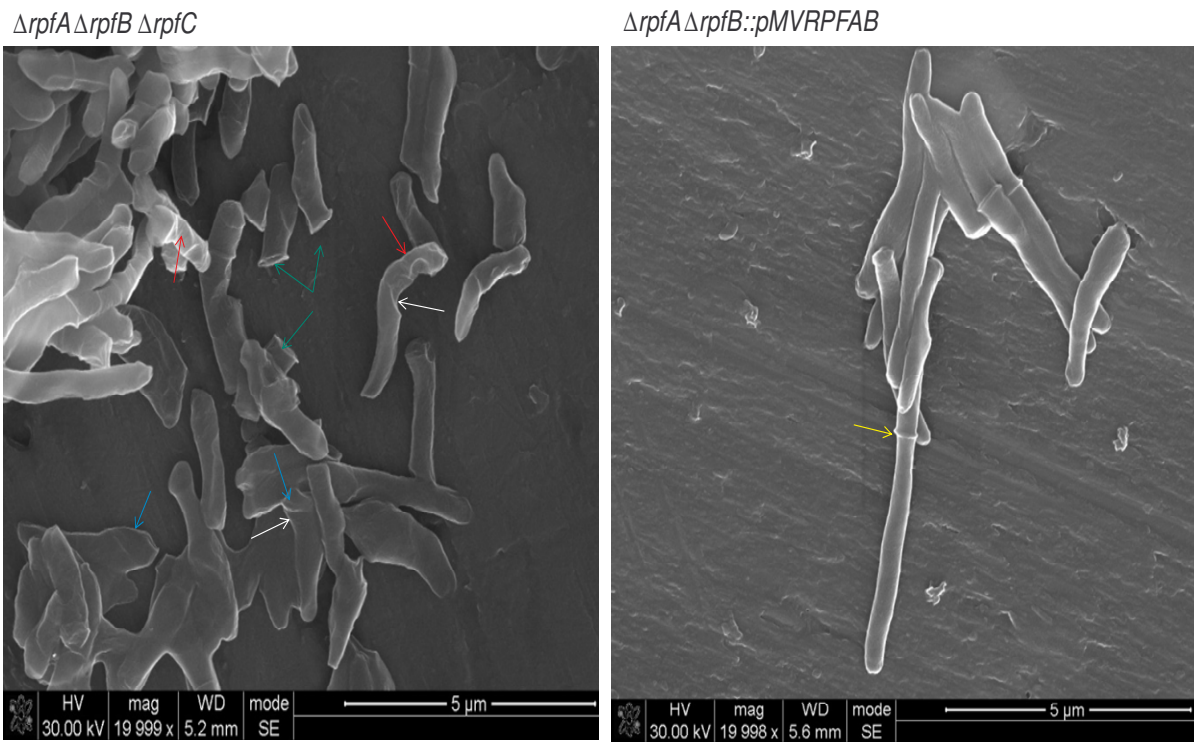



**Figure 38: High magnification Scanning Electron Microscopy images showing cell morphological defects associated with *rpfA* gene loss.** Log-phase cultures were dehydrated by increasing concentrations of ethanol and coated with gold sputtering. Yellow arrows indicate septa; white arrows indicate bending of cells; red arrows indicate cell surface defects; blue arrows indicate bulging at a single pole and green arrows indicate broken cells.

Heterologous expression of *M. tuberculosis rpfA* and *rpfB* genes in the *M. smegmatis* double mutant (Δ*rpfA* Δ*rpfB*) restores wild type phenotype, as seen in the complemented derivative Δ*rpfA* Δ*rpfB::pMVRPFAB*, although the cells seem slightly longer than wild type. This observation indicates similar biological functions of the Rpf-like proteins but slight differences in specificities between the two organisms.

## 4. Discussion

A significant proportion of individuals infected with *Mycobacterium tuberculosis*, harbor LTBI. In these cases the tubercle bacilli survive in the human host, without presentation of clinical symptoms, for as long as three decades or more with a defined risk of subsequent endogenous reactivation (Lillebaek*, et al.*, 2002). Recently, it has been suggested that the bacilli in LTBI undergo limited replication as evidenced by the absence of genomic content changes following analysis of DNA sequences from *M. tuberculosis* during three decades of LTBI (Yang*, et al.*, 2011). The above findings may also indicate that very little interaction between the host immune system and the bacilli in LTBI occurs, thus this corroborates the hypothesis that the pathogen is in a genuinely dormant, immune subversive state, during LTBI (Yang*, et al.*, 2011).

Rpf, first identified in *Mi. luteus* (Mukamolova*, et al.*, 1998), the progenitor of the Rpf-like protein family, was observed to stimulate growth of dormant, non-dividing cells of the organism (Mukamolova*, et al.*, 1998). Several Gram-positive bacteria, including *M. tuberculosis* and *M. smegmatis*, encode *rpf*-like genes within their genomes (Mukamolova*, et al.*, 1998). A 70 residue, lysozyme-like domain is conserved in all Rpf-like proteins (Kell & Young, 2000, Mukamolova*, et al.*, 2006) and it is this domain that is necessary for the muralytic activity displayed by these proteins (Mukamolova*, et al.*, 2006). *Mi. luteus* Rpf has been shown to hydrolyze cell wall material *in vitro* (Mukamolova*, et al.*, 2006, Telkov*, et al.*, 2006) and this compared with the observation that the *M. tuberculosis* RpfB domain structure is comparable to glycoside hydrolases (Cohen-Gonsaud*, et al.*, 2005), known to hydrolyze the peptidoglycan (Heidrich*, et al.*, 2002), confirms that the Rpf-like proteins, including *Mi. luteus* Rpf, are peptidoglycan hydrolases (Telkov*, et al.*, 2006). It is this mechanism of action, peptidoglycan cleavage, that promotes growth and more importantly permits for the bacteria to exit the dormant state (Keep*, et al.*, 2006). However, the precise nature of the

processes following peptidoglycan cleavage, that leads to reactivation of dormant organisms, remain to be unraveled.

Here, we report the characterization of the *rpf*-like genes in *M. smegmatis* by assessing the effects of gene deletion on survival, growth and cell wall homeostasis. *M. smegmatis* encodes four *rpf*-like genes as opposed to the five in *M. tuberculosis* (Figure 7), where the former lacks *rpf*D. The precise biological functions of individual Rpf-like proteins have yet to be elucidated; however it is currently known that a functional hierarchy exists within the *M. tuberculosis* Rpf protein family, with RpfB and RpfE ranking higher than RpfD (Kana*, et al.*, 2008). Furthermore, expression analysis of the *rpf* genes in *M. tuberculosis* suggests the individual Rpf-like proteins have distinct biological roles during different stages of growth and under different conditions of stress (Gupta*, et al.*, 2010). It has been suggested that the presence of DUF (Domains of Unknown Function) families within protein sequences is indicative of biological functions that are specific to certain (environmental) conditions (Bateman*, et al.*, 2010). This may provide evidence that RpfB possibly has a distinct biological role under different conditions, since it is the sole Rpf which includes such domains in its protein sequence Figure 8. Knowledge of the exact specificity of *M. tuberculosis rpfD* gene product is key in providing an explanation to the absence of an orthologue in *M. smegmatis*. The difference in *rpf* gene complement between *M. tuberculosis* and *M. smegmatis* could be attributed to the fact that these organisms have different lifestyles and different requirements to support growth in their respective environmental niches (Howard & Byrd, 2000, Abramovitch*, et al.*, 2011, Russell, 2011). Moreover the pathogenic nature of *M. tuberculosis* may demand a greater *rpf* gene repertoire as evidenced by the observation that all pathogenic mycobacterial genomes encode multiple *rpf*-like genes (Kana & Mizrahi, 2009). These data strongly suggest that the *rpf*-like orthologues may have differing biological roles in their respective organisms.

The organization of *rpf*A and *rpf*B genes is comparable in both *M. tuberculosis* and *M. smegmatis*, however a dissimilar situation is observed for the arrangement of the *rpfC* and *rpfE* genes. From the above observations; it is likely that RpfA and RpfB proteins in both organisms have similar biological functions, further supported by the significant protein sequence identity (65% and 78% in RpfA and RpfB respectively) which is observed for the orthologues (Figure 8). Moreover, returning functional copies of *M. tuberculosis rpfA* and *rpfB* genes simultaneously in the *M. smegmatis* Δ*rpfA* Δ*rpfB* double mutant background restores wild type phenotypes as observed in this study.

It is possible that *rpfC* and *rpfE* orthogues in the abovementioned organisms have dissimilar functions. The *rpfC* and *rpfE* genes are positioned in the same *M. smegmatis* chromosomal location separated by a single gene (*MSMEG*_4642) whereas in *M. tuberculosis* they are placed in different regions of the chromosome. The function of the intermediate gene (*MSMEG*_4642) is unclear since the gene product is annotated as a hypothetical protein, thus it is unknown how the gene affects expression of *rpfC* or *rpfE* or both, if at all. Numerous genes upstream and downstream *rpfC* and *rpfE* are not annotated in both organisms making it difficult to ascribe function by gene synteny comparisons.

Loss of up to three of four *M. smegmatis rpf*-like genes does not result in loss of viability, an observation seen in other Gram-positives including *M. tuberculosis* (Hartmann*, et al.*, 2004, Tufariello*, et al.*, 2004, Downing*, et al.*, 2005, Kana*, et al.*, 2008). These data indicate that the *rpf*-like genes are not essential for growth in these organisms, contrary to what has been observed in *Mi. luteus* (Mukamolova*, et al.*, 2002). However, the multiple Δ*rpfA* Δ*rpfB* and Δ*rpfA* Δ*rpfB* Δ*rpfC* mutants seem to display slightly reduced growth rates than the single Δ*rpfA* and Δ*rpfB* mutants or wild type during lag phase (Figure 33A) when cell density is measured by light absorption at 600 nm. Surprisingly, growth of all *rpf* mutants is indistinguishable from wild type when assessed by CFU enumeration (Figure 33B) with bacterial clumps forming in late stationary phase for Δ*rpfA*,

Δ*rpfA* Δ*rpfB* and Δ*rpfA* Δ*rpfB* Δ*rpfC* mutants (Figure 34). Clumping of the Δ*rpfA* mutant strain culture has also been observed in *M. tuberculosis* (Downing*, et al.*, 2004). A possible explanation for the above observation is that the apparent reduced growth rates for multiple *rpf*-like mutants could be as a result of inaccurate cell density readings due to dispersed micro-clumps in culture and would be consistent with the propensity of mutant strains, defective for two or more *rpf* genes, to clump. Bacterial clumping is documented in numerous other bacteria where genes associated with cell wall metabolism have been disrupted (Parish*, et al.*, 1997, Stapleton*, et al.*, 2007, Hett*, et al.*, 2008, Nguyen*, et al.*, 2010); and in some cases the observed clumping in culture is associated with bacterial filamentation (Gomez & Bishai, 2000, Hett*, et al.*, 2008, Nguyen*, et al.*, 2010). However, this is not the case in the multiple *rpf*-like mutants; filamentation is not observed in these strains as assessed by scanning electron microscopy (SEM) (Figure 38). In contrast, mutant strains of *M. smegmatis* defective for *rpfA*, *rpfB* and *rpfC* in double or triple gene deletion combinations displayed abnormally reduced cell size, severe clumping, and bulging at the poles.

Bacterial filamentation is a growth phenotype that is evident as elongated cells forming a chain-like appearance; in most cases due to defects in septum formation or septum inhibition and defects in cell division (Samaluru*, et al.*, 2007, Hett*, et al.*, 2008, Goley*, et al.*, 2010, England*, et al.*, 2011, Perez-Nunez*, et al.*, 2011, Shippy*, et al.*, 2011). In other instances, elongation is due to responses to various internal or external stresses for some bacteria. The lack of filamentation in *rpf* mutants may indicate that the Rpf-like proteins are not septum specific hydrolases and/or solely responsible for cell fragmentation or division since cells depleted of these genes have the ability to divide without forming chains. However, it has been shown that *M. tuberculosis* and *M. smegmatis* RpfB and RpfE proteins co-localize to the septa with an endopeptidase, RipA (Hett*, et al.*, 2007); it has also been demonstrated that RpfB is not compulsory for RipA functioning and possibly the interaction is conditional, probably under specific stresses (Hett*, et al.*, 2008). These data suggest that RpfB may not be an

essential septum hydrolase. However, the SEM images show that the septa are not well-defined and seem to be positioned towards a single pole of the cell in Δ*rpfA* Δ*rpfB* and Δ*rpfA* Δ*rpfB* Δ*rpfC* mutants (Figure 38). Whist the possibility that there are actual defects in septum formation or positioning cannot be ruled out, these observations may be due to general/mechanical cell distortions, forcing the septa to be displaced to another position, closer to a pole, of the cell.

Analysis of cell morphology by SEM also reveals distorted cells for Δ*rpfA*, Δ*rpfA* Δ*rpfB* and Δ*rpfA* Δ*rpfB* Δ*rpfC* mutants. The Δ*rpfA* cells have a wrinkle-like appearance, are irregularly shaped and display a slightly bulbous pole which is bent at the tip (Figure 38). Mycobacteria grow in a remarkably unipolar manner (Aldridge*, et al.*, 2012), and this property could explain the bulbous pole defect being seen at a single pole and not both. The abovementioned phenotype is exacerbated in the multiple mutants; whereas straight rods, with well-positioned septa are observed in wild type and Δ*rpfB* mutant cells. The wild type phenotype is restored in the complemented double mutant (Δ*rpfA* Δ*rpfB*::pMVRPFAB), although the cells seem slightly longer in this strain. A bulbous pole phenotype has been observed in mycobacteria as a result of disruptions to an essential cell wall gene, *wag31*, with its gene product localizing to the poles (Nguyen*, et al.*, 2007, Kang*, et al.*, 2008). The *wag31* gene encodes a cell division protein, Antigen 84, which is responsible for cell shape and cell wall synthesis (Nguyen*, et al.*, 2007, Kang*, et al.*, 2008). The observed phenotype could be as a result of changes in localization of the peptidoglycan biosynthetic or degradative proteins (Kang*, et al.*, 2008), which results in disrupted coordination of the two processes leading to asymmetrically shaped cells.

It is interesting that the cell morphology phenotype, to some degree, is associated with *rpfA* deletion, individually and in combination, thus it is possible that *rpfA* localizes to the pole(s) and regulates polar growth. A *Staphylococcus aureus* 8325/4 autolysin *atl* gene product plays a role in cell lysis and separation; insertional inactivation of the gene displays a rough outer surface due to absence

in peptidoglycan release at its surface (Foster, 1995) possibly this is the case with Δ*rpfA*, Δ*rpfA* Δ*rpfB* and Δ*rpfA* Δ*rpfB* Δ*rpfC* mutant cells. These *rpf*-like multiple mutants seemed "fragile" as evidenced by breakages in the cells during careful preparation for SEM, which is probably the result of weakening of the cell wall due to changes in the peptidoglycan stability.

Although the above cell morphology phenotypes are general observations, there are some mutant cells in certain microscopic images that do not display such defects (data not shown), indicating heterogeneity within the population. Growth and cell division in mycobacteria is asymmetric, giving rise to daughter cells with differing elongation rates and sizes thus creating heterogeneity within a clonal population consisting of physiologically distinct subpopulations (Aldridge*, et al.*, 2012). It is speculative that the introduction of mutated copies of *rpf* genes, and/or combinations thereof, into possibly physiologically distinct cells could have further exacerbated these effects and resulted in a phenotypically heterogeneous population. A heterogeneous population has been observed in *S. aureus* mutant strains defective in PBP1, a cell wall transpeptidase, and was attributed to be a result of disrupted cell division caused by PBP1 inactivation in cells at dissimilar stages of the cell cycle (Pereira*, et al.*, 2007). A similar situation may prevail in our *rpf* mutant strains.

The microscopic analysis of cell morphology suggests that the observed clumping of Δ*rpfA*, Δ*rpfA* Δ*rpfB* and Δ*rpfA* Δ*rpfB* Δ*rpfC* mutant cultures, Figure 34, is not as a result of filamentation. It is clear; however, that the clumping phenotype is linked to deletion of *rpfA* and becomes progressively worse with the deletion of further genes in the Δ*rpfA* mutant background. A similar situation has been observed in *S. aureus*; where deletion of genes encoding lytic transglycosylases, *isaA* and *sceD*, where the *sceD* mutant alone displays clumping and the phenotype is exacerbated by deletion of both genes (Stapleton*, et al.*, 2007). IsaA and SceD are autolysins like the Rpf-like proteins and share similar hydrolytic bond specificities in peptidoglycan, the β-1,4 glycosidic bond

cleavage between NAM and NAG residues. It is therefore possible that the cause for clumping seen in the *isaA* and *sceD* mutants, which is not associated with filamentation, is the same to that of *rpf*-like gene deletion. Defects in cell separation and surface properties are proposed as explanations to the observed clumping in broth (Foster, 1995, Kajimura*, et al.*, 2005, Stapleton*, et al.*, 2007); the latter sounds more probable. Hence it is also possible that deletion of multiple *rpf* genes results in changes in peptidoglycan composition leading to overall differences in cell hydrophobicity or charge. In this context, there may also be differences in mycolic acid profiles. These possibilities are currently being investigated.

Although there are no observable differences in growth between wild type and *rpf*-like mutants in liquid culture under those conditions tested (Figure 33); growth defects are observed, as assessed by biofilm formation (Figure 37). Biofilm formation is retarded for the multiple *rpf*-like mutants, whereas the single deletion mutants exhibit the formation of mature biofilms. Bacterial biofilms are surface-associated cell aggregates, forming communities which are then encompassed by an extracellular matrix, the composition of which can differ between organisms (Donlan, 2002). Biofilm formation involves a few stages of development, which include surface attachment, microcolony formation and maturation (Stanley & Lazazzera, 2004). Commencement of the abovementioned stages is influenced and directed by environmental signals and bacterial regulatory pathways (Stanley & Lazazzera, 2004), ensuring that the bacteria initiate and form successful biofilm structures in appropriate environments (Stanley & Lazazzera, 2004). The occurrence of bacterial biofilms varies, from natural aquatic systems to implanted medical devices, and each type is well adapted to its environment (Donlan, 2002). Microorganisms within biofilms are dissimilar to their planktonic forms at the genetic level, with respect to gene transcription. Bacterial cells in a biofilm have several advantages over their planktonic counterparts; since the biofilm ecology provides a suitable environment for cell-cell signaling which maintains

structure, exchange of genetic material, formation of nutrient gradients and the community is well protected against antibiotic treatment (Donlan, 2002).

Biofilms are relevant in the clinical setting and human disease; several pathogens responsible for chronic infections have been associated with biofilm formation (Hall-Stoodley & Stoodley, 2009). Bacteria in biofilms are highly recalcitrant to antibiotic treatment and such infections persist irrespective of host immune responses (Hall-Stoodley & Stoodley, 2009). The observed "resistance" to antibiotics in *M. tuberculosis* is due to the presence of a heterogeneous population within the biofilm consisting of slow-growing cells, cells in stationary phase and persisters (Lewis, 2007, Ojha*, et al.*, 2008). Persisters survive antibiotic treatment irrespective of drug concentration used, an *in vitro* observation seen in *M. tuberculosis (Ojha, et al., 2008)*; moreover the extra-cellular matrix guards against immune cells (Lewis, 2007).

The Δ*rpfA* Δ*rpfB* and Δ*rpfA* Δ*rpfB* Δ*rpfC* mutants display under-developed biofilms with a few loose colony islands floating on the liquid surface with most of the bacterial cells settling at the bottom of the wells. It may seem that the multiple mutants had an inability to form biofilms but it may be more likely that these mutants were growing at a much reduced rate thus the initial stages of biofilm formation may have been disrupted. A defect in biofilm formation has been linked to colony morphology alterations (Bourai*, et al.*, 2011) due to disrupted cell wall genes such as *lpqB* and *M. tuberculosis* Rv2911; the former gene product being responsible for processes in cell wall homeostasis and the latter encoding DacB2, which plays a role in peptidoglycan synthesis (Nguyen*, et al.*, 2010). The above observation is consistent with what is seen in this study (Figure 36). A reduction in serpentine cording is observed for Δ*rpfA* Δ*rpfB* and Δ*rpfA* Δ*rpfB* Δ*rpfC* mutant colonies, thus colonies appear smoother with regular edges. A similar phenotype is seen in *M. tuberculosis* Δ*rpfA* Δ*rpfB* mutant (Russell-Goldman*, et al.*, 2008). It is suggested that a smoother colony phenotype is likely due to alterations in cell surface properties caused by a reduction in

hydrophobicity (Dahl & Lau Bonilla, 2011). It is probable that such changes in cell surface properties might impact on cell-cell interactions within a community, in the case of biofilms, thus explaining the observed defect in biofilm formation with the multiple mutants (Bourai*, et al.*, 2011).

It was most surprising that the *rpf*-like mutants, more especially the multiple mutants, lacked an SDS susceptibility phenotype since it is observed that several mycobacterial mutants that display a defect in cell wall construction are sensitive to this detergent (Kocincova*, et al.*, 2004). Moreover, it has been shown that *M. tuberculosis* quadruple *rpf*-like mutants that are deficient in *rpfB* and *rpfE* display hypersensitivity to the detergent (Kana*, et al.*, 2008). This may be indicative of the differences in Rpf-like protein specificities between the two organisms or other mechanisms of resistance to detergent in *M. smegmatis*. However; the multiple *rpf*-like mutants and not the single mutants display a 2-4 log increase in susceptibility to erythromycin and vancomycin when compared to the wild type (Figure 35E & F), a phenotype also observed in *M. tuberculosis* (Kana*, et al.*, 2010). In *E. coli*; it has also been shown that multiple hydrolase mutants and not single mutants, display a similar antibiotic susceptibility phenotype (Korsak*, et al.*, 2005). Generally, increased susceptibility to antibiotics is due to compromised cell wall integrity (Korsak*, et al.*, 2005, Nguyen*, et al.*, 2010), however it is unclear why the mutants display a high degree of susceptibility specifically to erythromycin and vancomycin. Heterologous expression of *M. tuberculosis rpfA* and *rpfB* genes in *M. smegmatis* Δ*rpfA* Δ*rpfB* mutant background restored wild type phenotypes in all cases, indicating similarities in biological functions of Rpf-like proteins between the two organisms.

### 4.1 Concluding remarks

Taken together, these results are indicative of a loss of bacterial cell wall integrity in *rpf* mutants and our data suggest that *rpfA* is likely to be largely responsible for bacterial cell wall integrity maintenance. Some of the phenotypes mentioned

above are associated with the loss of *rpfA* further indicating that the Rpf-like proteins may indeed have distinct biological roles and are not functionally redundant as previously suggested (Downing*, et al.*, 2004).

Further experimental work needs to be undertaken in order to prove or disprove the above speculations and to explain some of the observed phenotypes. Localization of Rpf proteins within cells will provide insight as to where the proteins exert their biological roles and this may reveal specificities of each of the Rpf-like proteins.

In summary, we have demonstrated that the *rpf* genes play an important role in some essential cell division processes and are likely to regulate bacterial growth through multiple mechanisms. The mechanism through which these proteins stimulate growth remains a mystery but is likely to be some essential function associated with remodeling of the bacterial cell wall. These and other such questions are currently being investigated.

## 5. Appendices

### 5.1 Appendix A: Culture media

Luria-Bertani Agar (LA)

5g yeast extract, 10g tryptone, 10g NaCl, 1.5g agar, 1000 ml sd$H_2O$

Luria-Bertani Broth (LB)

5g yeast extract, 10g tryptone, 10g NaCl, 1000 ml sd$H_20$

2TY

10g yeast extract, 16g tryptone, 5g NaCl, 1000 ml sd$H_2O$

Middlebrook 7H9

4.7 g Difco Middlebrook 7H9 powder, 2 ml glycerol, 900 ml d$H_2O$, 100 ml 10X glucose-salt, 2 ml Tween80 or 2 ml Tyloxapol

Middlebrook 7H10

19 g Difco Middlebrook 7H10 powder, 5 ml glycerol, 900 ml d$H_2O$, 100 ml 10X glucose-salt

Middlebrook 7H11

21 g Difco Middlebrook 7H11 powder, 5 ml glycerol, 900 ml d$H_2O$, 100 ml 10X glucose-salt

Glucose salt 10 g glucose, 4.25 g NaCl, 500 ml d$H_2O$

Media was sterilized by autoclaving at 121℃ for 20 min

## 5.2 Appendix B: Solutions

Plasmid DNA preparation solutions

| | |
|---|---|
| 1 M Tris-HCl | 60.56 g Tris, pH 8.0 with HCl, 500 ml $dH_2O$ final volume |
| 0.5 M EDTA | 18.6 g powder, pH 8.0 with NaOH, $dH_2O$ |
| Soln I | 50 nM glucose, 25 mM Tris-HCl, 10 mM EDTA, pH 8.0 |
| Soln II | 0.2 M NaOH, 1.0% SDS |
| Soln III | 5 M potassium acetate, 11.5 ml glacial acetic acid, 88.5 $dH_20$, pH 4.8 |
| TE | 1 ml 1 M Tris-HCl, 2 ml 0.5 M EDTA |
| TAE buffer | 242 g Tris base, 100 ml 0.5 M EDTA, 57.1 ml glacial acetic acid, pH 8.0, 1000 ml final volume $dH_2O$ |

Southern Blotting solutions

| | |
|---|---|
| 20X SSC | 175.3 g NaCl, 88.2g sodium citrate, make up to 1000 ml, pH 7.0 |
| Soln I | 10 ml 20X SSC, 1 ml 10% SDS, 89 ml $dH_2O$ |
| Soln II | 2.5 ml 20X SSC, 1 ml 10% SDS, 96.5 ml |
| 10% SDS | 10 g powder, 100 ml $sdH_20$ |
| Maleic acid buffer | 116.1 g maleic acid powder, 87. 66 g NaCl, pH 7.5 with NaOH pellets, make 1000 ml final volume with $dH_2O$ |

## 5.3 Appendix C: Plasmids and molecular weight markers

Plasmids

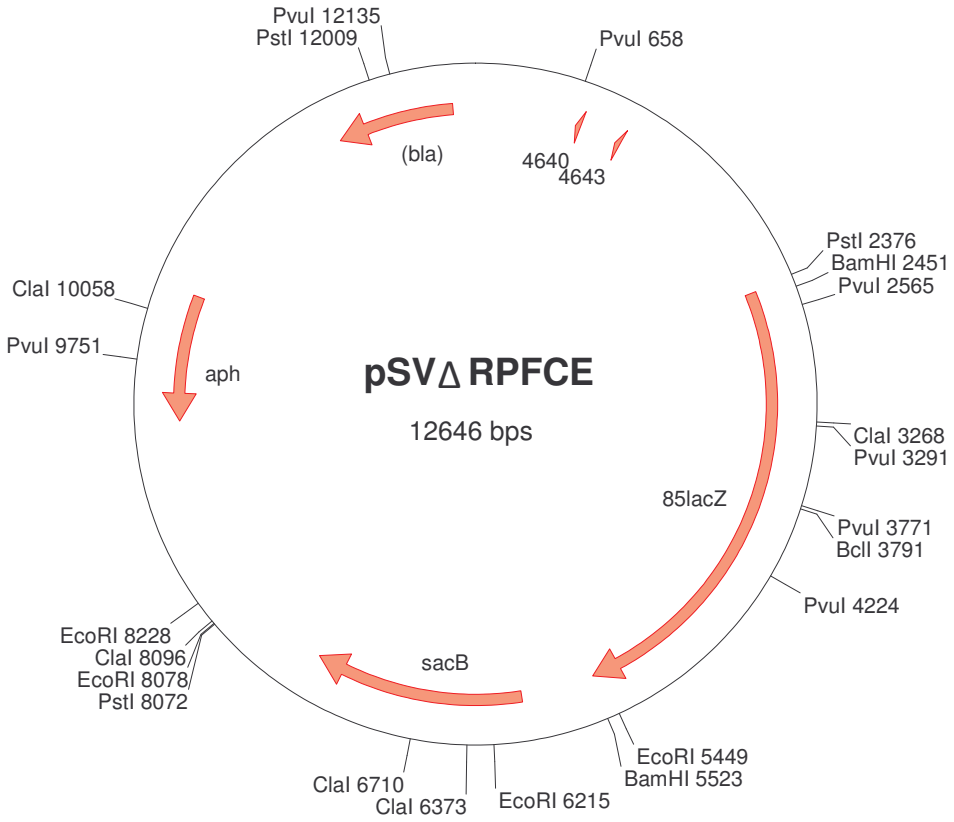



**Figure A1:** Knockout construct, pSVΔRPFA, carrying *rpfC* and *rpfE* deletion alleles for construction of Δ*rpfC* and Δ*rpfE* mutants.

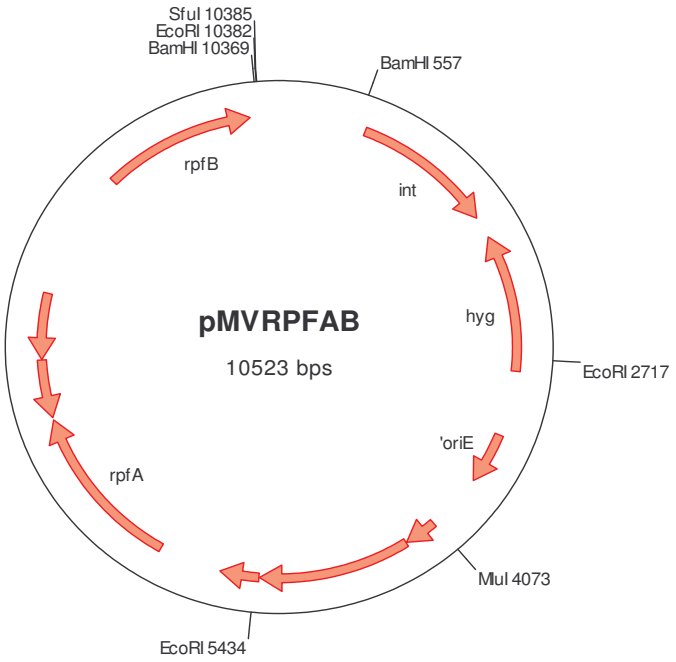
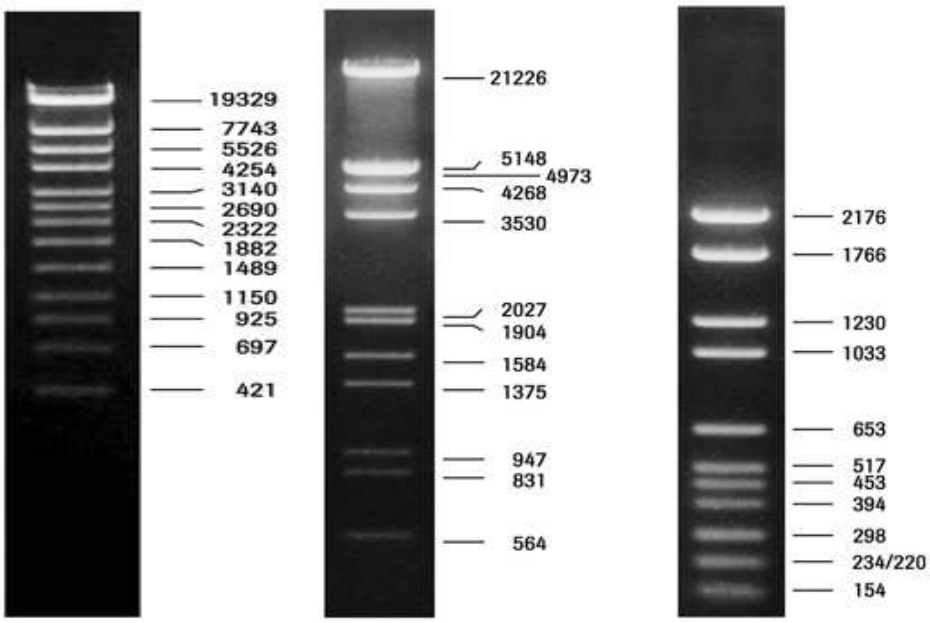



**Figure A2**: Integrating vector carrying *M. tuberculosis rpfA* and *rpfB* genes under the control of their native promoter regions, for complementation purposes.

Lambda DNA molecular weight markers




DNA molecular weight marker **III**, **IV** and **VI** supplied by Roche Applied Science

## 6. References


Ã“ Gaora P, Barnini S, Hayward C, Filley E, Rook G, Young D & Thole J (1997) Mycobacteria as Immunogens: Development of Expression Vectors for Use in Multiple Mycobacterial Species. *Medical Principles and Practice* **6**: 91-96.

Abramovitch RB, Rohde KH, Hsu FF & Russell DG (2011) aprABC: a *Mycobacterium tuberculosis* complex-specific locus that modulates pH-driven adaptation to the macrophage phagosome. *Mol Microbiol* **80**: 678-694.

Ahmad S (2010) Pathogenesis, immunology, and diagnosis of latent *Mycobacterium tuberculosis* infection. *Clin Dev Immunol* **2011**: 814943.

Aldridge BB, Fernandez-Suarez M, Heller D, Ambravaneswaran V, Irimia D, Toner M & Fortune SM (2012) Asymmetry and aging of mycobacterial cells lead to variable growth and antibiotic susceptibility. *Science* **335**: 100-104.

American Thoracic Society CfDCaP, Infectious Diseases Society of America (2003) Treatment of tuberculosis. *MMWR Recomm Rep* **52**: 1-77.

Balaban NQ, Merrin J, Chait R, Kowalik L & Leibler S (2004) Bacterial persistence as a phenotypic switch. *Science* **305**: 1622-1625.

Basak A (2005) Inhibitors of proprotein convertases. *J Mol Med* **83**: 844-855.

Bateman A, Holden MT & Yeats C (2005) The G5 domain: a potential N-acetylglucosamine recognition domain involved in biofilm formation. *Bioinformatics* **21**: 1301-1303.

Bateman A, Coggill P & Finn RD (2010) DUFs: families in search of function. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **66**: 1148-1152.

Biketov S, Potapov V, Ganina E, Downing K, Kana BD & Kaprelyants A (2007) The role of resuscitation promoting factors in pathogenesis and reactivation of

*Mycobacterium tuberculosis* during intra-peritoneal infection in mice. *BMC Infect Dis* **7**: 146.

Biketov S, Mukamolova GV, Potapov V*, et al.* (2000) Culturability of *Mycobacterium tuberculosis* cells isolated from murine macrophages: a bacterial growth factor promotes recovery. *FEMS Immunol Med Microbiol* **29**: 233-240.

Born P, Breukink E & Vollmer W (2006) In vitro synthesis of cross-linked murein and its attachment to sacculi by PBP1A from *Escherichia coli*. *J Biol Chem* **281**: 26985-26993.

Bouhss A, Trunkfield AE, Bugg TD & Mengin-Lecreulx D (2008) The biosynthesis of peptidoglycan lipid-linked intermediates. *FEMS Microbiol Rev* **32**: 208-233.

Bourai N, Jacobs WR, Jr. & Narayanan S (2011) Deletion and overexpression studies on DacB2, a putative low molecular mass penicillin binding protein from *Mycobacterium tuberculosis* H(37)Rv. *Microb Pathog* **52**: 109-116.

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

Callewaert L & Michiels CW (2010) Lysozymes in the animal kingdom. *J Biosci* **35**: 127-160.

Cohen-Gonsaud M, Keep NH, Davies AP, Ward J, Henderson B & Labesse G (2004) Resuscitation-promoting factors possess a lysozyme-like domain. *Trends Biochem Sci* **29**: 7-10.

Cohen-Gonsaud M, Barthe P, Bagneris C, Henderson B, Ward J, Roumestand C & Keep NH (2005) The structure of a resuscitation-promoting factor domain from *Mycobacterium tuberculosis* shows homology to lysozymes. *Nat Struct Mol Biol* **12**: 270-273.

Dahl JL (2004) Electron microscopy analysis of *Mycobacterium tuberculosis* cell division. *FEMS Microbiol Lett* **240**: 15-20.

Dahl JL & Lau Bonilla D (2011) The wag31 gene of *Mycobacterium tuberculosis* is positively regulated by the stringent response. *FEMS Microbiol Lett* **319**: 153-159.

Davies AP, Dhillon AP, Young M, Henderson B, McHugh TD & Gillespie SH (2008) Resuscitation-promoting factors are expressed in *Mycobacterium tuberculosis*-infected human tissue. *Tuberculosis (Edinb)* **88**: 462-468.

Demina GR, Makarov VA, Nikitushkin VD*, et al.* (2009) Finding of the low molecular weight inhibitors of resuscitation promoting factor enzymatic and resuscitation activity. *PLoS One* **4**: e8174.

Denham EL, Ward PN & Leigh JA (2008) Lipoprotein signal peptides are processed by Lsp and Eep of *Streptococcus uberis*. *J Bacteriol* **190**: 4641-4647.

Dheda K, Schwander SK, Zhu B, van Zyl-Smit RN & Zhang Y (2010) The immunology of tuberculosis: from bench to bedside. *Respirology* **15**: 433-450.

Diedrich CR & Flynn JL (2011) HIV-1/*Mycobacterium tuberculosis* Coinfection Immunology: How Does HIV-1 Exacerbate Tuberculosis? *Infect Immun* **79**: 1407-1417.

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

Downing KJ, Betts JC, Young DI, McAdam RA, Kelly F, Young M & Mizrahi V (2004) Global expression profiling of strains harbouring null mutations reveals that the five rpf-like genes of *Mycobacterium tuberculosis* show functional redundancy. *Tuberculosis (Edinb)* **84**: 167-179.

Downing KJ, Mischenko VV, Shleeva MO*, et al.* (2005) Mutants of *Mycobacterium tuberculosis* lacking three of the five *rpf*-like genes are defective for growth in vivo and for resuscitation in vitro. *Infect Immun* **73**: 3038-3043.

Dutta NK, Mehra S, Didier PJ*, et al.* (2010) Genetic requirements for the survival of tubercle bacilli in primates. *J Infect Dis* **201**: 1743-1752.

England K, Crew R & Slayden RA (2011) *Mycobacterium tuberculosis* septum site determining protein, Ssd encoded by rv3660c, promotes filamentation and elicits an alternative metabolic and dormancy stress response. *BMC Microbiol* **11**: 79.

Esposito L, Pedone C & Vitagliano L (2006) Molecular dynamics analyses of cross-beta-spine steric zipper models: beta-sheet twisting and aggregation. *Proc Natl Acad Sci U S A* **103**: 11533-11538.

Foster SJ (1995) Molecular characterization and functional analysis of the major autolysin of *Staphylococcus aureus* 8325/4. *J Bacteriol* **177**: 5723-5725.

Fraipont C, Sapunaric F, Zervosen A*, et al.* (2006) Glycosyl transferase activity of the *Escherichia coli* penicillin-binding protein 1b: specificity profile for the substrate. *Biochemistry* **45**: 4007-4013.

Gill WP, Harik NS, Whiddon MR, Liao RP, Mittler JE & Sherman DR (2009) A replication clock for *Mycobacterium tuberculosis*. *Nat Med* **15**: 211-214.

Goley ED, Comolli LR, Fero KE, Downing KH & Shapiro L (2010) DipM links peptidoglycan remodelling to outer membrane organization in *Caulobacter*. *Mol Microbiol* **77**: 56-73.

Gomez JE & Bishai WR (2000) whmD is an essential mycobacterial gene required for proper septation and cell division. *Proc Natl Acad Sci U S A* **97**: 8554-8559.

Gordhan BG & Parish T (2001) Gene Replacement using Pretreated DNA. *Methods in Molecular Medicine,* Vol. **54**: 77-91. Humana Press Inc., Totowa NJ.

Gupta RK, Srivastava BS & Srivastava R (2010) Comparative expression analysis of *rpf*-like genes of *Mycobacterium tuberculosis* H37Rv under different physiological stress and growth conditions. *Microbiology* **156**: 2714-2722.

Hall-Stoodley L & Stoodley P (2009) Evolving concepts in biofilm infections. *Cell Microbiol* **11**: 1034-1043.

Hall RG, Leff RD & Gumbo T (2009) Treatment of active pulmonary tuberculosis in adults: current standards and recent advances. Insights from the Society of Infectious Diseases Pharmacists. *Pharmacotherapy* **29**: 1468-1481.

Hartmann M, Barsch A, Niehaus K, Puhler A, Tauch A & Kalinowski J (2004) The glycosylated cell surface protein Rpf2, containing a resuscitation-promoting factor motif, is involved in intercellular communication of *Corynebacterium glutamicum*. *Arch Microbiol* **182**: 299-312.

Heidrich C, Ursinus A, Berger J, Schwarz H & Holtje JV (2002) Effects of multiple deletions of murein hydrolases on viability, septum cleavage, and sensitivity to large toxic molecules in *Escherichia coli*. *J Bacteriol* **184**: 6093-6099.

Hesseling AC, Johnson LF, Jaspan H*, et al.* (2009) Disseminated bacille Calmette-Guerin disease in HIV-infected South African infants. *Bull World Health Organ* **87**: 505-511.

Hett EC & Rubin EJ (2008) Bacterial growth and cell division: a mycobacterial perspective. *Microbiol Mol Biol Rev* **72**: 126-156, table of contents.

Hett EC, Chao MC & Rubin EJ (2010) Interaction and modulation of two antagonistic cell wall enzymes of mycobacteria. *PLoS Pathog* **6**: e1001020.

Hett EC, Chao MC, Deng LL & Rubin EJ (2008) A mycobacterial enzyme essential for cell division synergizes with resuscitation-promoting factor. *PLoS Pathog* **4**: e1000001.

Hett EC, Chao MC, Steyn AJ, Fortune SM, Deng LL & Rubin EJ (2007) A partner for the resuscitation-promoting factors of *Mycobacterium tuberculosis*. *Mol Microbiol* **66**: 658-668.

Holtje JV, Mirelman D, Sharon N & Schwarz U (1975) Novel type of murein transglycosylase in *Escherichia coli*. *J Bacteriol* **124**: 1067-1076.

Howard ST & Byrd TF (2000) The rapidly growing mycobacteria: saprophytes and parasites. *Microbes Infect* **2**: 1845-1853.

Hugonnet JE, Tremblay LW, Boshoff HI, Barry CE, 3rd & Blanchard JS (2009) Meropenem-clavulanate is effective against extensively drug-resistant *Mycobacterium tuberculosis*. *Science* **323**: 1215-1218.

Jafri SM, Singal AG, Kaul D & Fontana RJ (2011) Detection and management of latent tuberculosis in liver transplant patients. *Liver Transpl* **17**: 306-314.

Jambo KC, Sepako E, Fullerton DG*, et al.* (2011) Bronchoalveolar CD4+ T cell responses to respiratory antigens are impaired in HIV-infected adults. *Thorax* **66**: 375-382.

Kajimura J, Fujiwara T, Yamada S*, et al.* (2005) Identification and molecular characterization of an N-acetylmuramyl-L-alanine amidase Sle1 involved in cell separation of *Staphylococcus aureus*. *Mol Microbiol* **58**: 1087-1101.

Kana BD & Mizrahi V (2009) Resuscitation-promoting factors as lytic enzymes for bacterial growth and signaling. *FEMS Immunol Med Microbiol* **58**: 39-50.

Kana BD, Mizrahi V & Gordhan BG (2010) Depletion of resuscitation-promoting factors has limited impact on the drug susceptibility of *Mycobacterium tuberculosis*. *J Antimicrob Chemother* **65**: 1583-1585.

Kana BD, Gordhan BG, Downing KJ*, et al.* (2008) The resuscitation-promoting factors of *Mycobacterium tuberculosis* are required for virulence and

resuscitation from dormancy but are collectively dispensable for growth in vitro. *Mol Microbiol* **67**: 672-684.

Kang CM, Nyayapathy S, Lee JY, Suh JW & Husson RN (2008) Wag31, a homologue of the cell division protein DivIVA, regulates growth, morphology and polar cell wall synthesis in mycobacteria. *Microbiology* **154**: 725-735.

Karonga (1996) Randomised controlled trial of single BCG, repeated BCG, or combined BCG and killed *Mycobacterium leprae* vaccine for prevention of leprosy and tuberculosis in Malawi. Karonga Prevention Trial Group. *Lancet* **348**: 17-24.

Keep NH, Ward JM, Cohen-Gonsaud M & Henderson B (2006) Wake up! Peptidoglycan lysis and bacterial non-growth states. *Trends Microbiol* **14**: 271-276.

Kell DB & Young M (2000) Bacterial dormancy and culturability: the role of autocrine growth factors. *Curr Opin Microbiol* **3**: 238-243.

Kesavan AK, Brooks M, Tufariello J, Chan J & Manabe YC (2009) Tuberculosis genes expressed during persistence and reactivation in the resistant rabbit model. *Tuberculosis (Edinb)* **89**: 17-21.

Kocincova D, Sonden B, de Mendonca-Lima L, Gicquel B & Reyrat JM (2004) The Erp protein is anchored at the surface by a carboxy-terminal hydrophobic domain and is important for cell-wall structure in *Mycobacterium smegmatis*. *FEMS Microbiol Lett* **231**: 191-196.

Koltunov V, Greenblatt CL, Goncharenko AV, Demina GR, Klein BY, Young M & Kaprelyants AS (2010) Structural changes and cellular localization of resuscitation-promoting factor in environmental isolates of *Micrococcus luteus*. *Microb Ecol* **59**: 296-310.

Korch SB, Henderson TA & Hill TM (2003) Characterization of the hipA7 allele of *Escherichia coli* and evidence that high persistence is governed by (p)ppGpp synthesis. *Mol Microbiol* **50**: 1199-1213.

Korsak D, Liebscher S & Vollmer W (2005) Susceptibility to antibiotics and beta-lactamase induction in murein hydrolase mutants of *Escherichia coli*. *Antimicrob Agents Chemother* **49**: 1404-1409.

Lenz LL, Mohammadi S, Geissler A & Portnoy DA (2003) SecA2-dependent secretion of autolytic enzymes promotes *Listeria monocytogenes* pathogenesis. *Proc Natl Acad Sci U S A* **100**: 12432-12437.

Lewis K (2007) Persister cells, dormancy and infectious disease. *Nat Rev Microbiol* **5**: 48-56.

Lillebaek T, Dirksen A, Baess I, Strunge B, Thomsen VO & Andersen AB (2002) Molecular evidence of endogenous reactivation of *Mycobacterium tuberculosis* after 33 years of latent infection. *J Infect Dis* **185**: 401-404.

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

Mahne M, Tauch A, Puhler A & Kalinowski J (2006) The *Corynebacterium glutamicum* gene pmt encoding a glycosyltransferase related to eukaryotic protein-O-mannosyltransferases is essential for glycosylation of the resuscitation promoting factor (Rpf2) and other secreted proteins. *FEMS Microbiol Lett* **259**: 226-233.

Malcolm BA, Rosenberg S, Corey MJ, Allen JS, de Baetselier A & Kirsch JF (1989) Site-directed mutagenesis of the catalytic residues Asp-52 and Glu-35 of chicken egg white lysozyme. *Proc Natl Acad Sci U S A* **86**: 133-137.

Mallorqui-Fernandez N, Manandhar SP, Mallorqui-Fernandez G*, et al.* (2008) A new autocatalytic activation mechanism for cysteine proteases revealed by Prevotella intermedia interpain A. *J Biol Chem* **283**: 2871-2882.

Mitchison DA (1979) Basic mechanisms of chemotherapy. *Chest* **76**: 771-781.

Mukamolova GV, Kormer SS, Kell DB & Kaprelyants AS (1999) Stimulation of the multiplication of *Micrococcus luteus* by an autocrine growth factor. *Arch Microbiol* **172**: 9-14.

Mukamolova GV, Kaprelyants AS, Young DI, Young M & Kell DB (1998) A bacterial cytokine. *Proc Natl Acad Sci U S A* **95**: 8916-8921.

Mukamolova GV, Turapov O, Malkin J, Woltmann G & Barer MR (2010) Resuscitation-promoting factors reveal an occult population of tubercle Bacilli in Sputum. *Am J Respir Crit Care Med* **181**: 174-180.

Mukamolova GV, Turapov OA, Young DI, Kaprelyants AS, Kell DB & Young M (2002) A family of autocrine growth factors in *Mycobacterium tuberculosis*. *Mol Microbiol* **46**: 623-635.

Mukamolova GV, Turapov OA, Kazarian K, Telkov M, Kaprelyants AS, Kell DB & Young M (2002) The *rpf* gene of *Micrococcus luteus* encodes an essential secreted growth factor. *Mol Microbiol* **46**: 611-621.

Mukamolova GV, Murzin AG, Salina EG, Demina GR, Kell DB, Kaprelyants AS & Young M (2006) Muralytic activity of *Micrococcus luteus* Rpf and its relationship to physiological activity in promoting bacterial growth and resuscitation. *Mol Microbiol* **59**: 84-98.

Najem NM, Zadeh VB, Al-Abdulrazzaq AH, Al-Otaibi SR, Kadyan S & Joneja M (2009) Bacillus Calmette-Guerin vaccine- induced lupus vulgaris in a child. *Acta Dermatovenerol Alp Panonica Adriat* **18**: 195-197.

Nguyen HT, Wolff KA, Cartabuke RH, Ogwang S & Nguyen L (2010) A lipoprotein modulates activity of the MtrAB two-component system to provide intrinsic multidrug resistance, cytokinetic control and cell wall homeostasis in Mycobacterium. *Mol Microbiol* **76**: 348-364.

Nguyen L, Scherr N, Gatfield J, Walburger A, Pieters J & Thompson CJ (2007) Antigen 84, an effector of pleiomorphism in *Mycobacterium smegmatis*. *J Bacteriol* **189**: 7896-7910.

Ojha AK, Baughn AD, Sambandan D*, et al.* (2008) Growth of *Mycobacterium tuberculosis* biofilms containing free mycolic acids and harbouring drug-tolerant bacteria. *Mol Microbiol* **69**: 164-174.

Panutdaporn N, Kawamoto K, Asakura H & Makino SI (2006) Resuscitation of the viable but non-culturable state of Salmonella enterica serovar Oranienburg by recombinant resuscitation-promoting factor derived from *Salmonella typhimurium* strain LT2. *Int J Food Microbiol* **106**: 241-247.

Parish T & Stoker NG (2000) Use of a flexible cassette method to generate a double unmarked *Mycobacterium tuberculosis* tlyA plcABC mutant by gene replacement. *Microbiology* **146 ( Pt 8)**: 1969-1975.

Parish T, Liu J, Nikaido H & Stoker NG (1997) A *Mycobacterium smegmatis* mutant with a defective inositol monophosphate phosphatase gene homolog has altered cell envelope permeability. *J Bacteriol* **179**: 7827-7833.

Pereira SF, Henriques AO, Pinho MG, de Lencastre H & Tomasz A (2007) Role of PBP1 in cell division of *Staphylococcus aureus*. *J Bacteriol* **189**: 3525-3531.

Perez-Nunez D, Briandet R, David B*, et al.* (2011) A new morphogenesis pathway in bacteria: unbalanced activity of cell wall synthesis machineries leads to coccus-to-rod transition and filamentation in ovococci. *Mol Microbiol* **79**: 759-771.

Pilgrim S, Kolb-Maurer A, Gentschev I, Goebel W & Kuhn M (2003) Deletion of the gene encoding p60 in *Listeria monocytogenes* leads to abnormal cell division and loss of actin-based motility. *Infect Immun* **71**: 3473-3484.

Ravagnani A, Finan CL & Young M (2005) A novel firmicute protein family related to the actinobacterial resuscitation-promoting factors by non-orthologous domain displacement. *BMC Genomics* **6**: 39.

Raviglione MC & Smith IM (2007) XDR tuberculosis--implications for global public health. *N Engl J Med* **356**: 656-659.

Rezwan M, Grau T, Tschumi A & Sander P (2007) Lipoprotein synthesis in mycobacteria. *Microbiology* **153**: 652-658.

Rickman L, Scott C, Hunt DM*, et al.* (2005) A member of the cAMP receptor protein family of transcription regulators in *Mycobacterium tuberculosis* is required for virulence in mice and controls transcription of the *rpfA* gene coding for a resuscitation promoting factor. *Mol Microbiol* **56**: 1274-1286.

Ruggiero A, Tizzano B, Pedone E, Pedone C, Wilmanns M & Berisio R (2009) Crystal structure of the resuscitation-promoting factor (DeltaDUF)RpfB from *M. tuberculosis*. *J Mol Biol* **385**: 153-162.

Ruggiero A, Marasco D, Squeglia F, Soldini S, Pedone E, Pedone C & Berisio R (2010) Structure and functional regulation of RipA, a mycobacterial enzyme essential for daughter cell separation. *Structure* **18**: 1184-1190.

Russell-Goldman E, Xu J, Wang X, Chan J & Tufariello JM (2008) A *Mycobacterium tuberculosis* Rpf double-knockout strain exhibits profound defects in reactivation from chronic tuberculosis and innate immunity phenotypes. *Infect Immun* **76**: 4269-4281.

Russell DG (2011) *Mycobacterium tuberculosis* and the intimate discourse of a chronic infection. *Immunol Rev* **240**: 252-268.

Sala C, Haouz A, Saul FA, Miras I, Rosenkrands I, Alzari PM & Cole ST (2009) Genome-wide regulon and crystal structure of BlaI (Rv1846c) from *Mycobacterium tuberculosis*. *Mol Microbiol* **71**: 1102-1116.

Samaluru H, SaiSree L & Reddy M (2007) Role of SufI (FtsP) in cell division of *Escherichia coli*: evidence for its involvement in stabilizing the assembly of the divisome. *J Bacteriol* **189**: 8044-8052.

Sander P, Rezwan M, Walker B*, et al.* (2004) Lipoprotein processing is required for virulence of *Mycobacterium tuberculosis*. *Mol Microbiol* **52**: 1543-1552.

Saunders BM & Britton WJ (2007) Life and death in the granuloma: immunopathology of tuberculosis. *Immunol Cell Biol* **85**: 103-111.

Scheurwater E, Reid CW & Clarke AJ (2008) Lytic transglycosylases: bacterial space-making autolysins. *Int J Biochem Cell Biol* **40**: 586-591.

Schilling A, Neuner E & Rehm SJ (2011) Vancomycin: a 50-something-year-old antibiotic we still don't understand. *Cleve Clin J Med* **78**: 465-471.

Schroeckh V & Martin K (2006) Resuscitation-promoting factors: distribution among actinobacteria, synthesis during life-cycle and biological activity. *Antonie Van Leeuwenhoek* **89**: 359-365.

Schuck SD, Mueller H, Kunitz F*, et al.* (2009) Identification of T-cell antigens specific for latent *Mycobacterium tuberculosis* infection. *PLoS One* **4**: e5590.

Setia MS, Steinmaus C, Ho CS & Rutherford GW (2006) The role of BCG in prevention of leprosy: a meta-analysis. *Lancet Infect Dis* **6**: 162-170.

Shippy DC, Heintz JA, Albrecht RM, Eakley NM & Fadl AA (2011) Deletion of glucose-inhibited division (gidA) gene alters the morphological and replication characteristics of *Salmonella* enterica Serovar *typhimurium*. *Arch Microbiol*.

Shleeva M, Mukamolova GV, Young M, Williams HD & Kaprelyants AS (2004) Formation of 'non-culturable' cells of *Mycobacterium smegmatis* in stationary phase in response to growth under suboptimal conditions and their Rpf-mediated resuscitation. *Microbiology* **150**: 1687-1697.

Shleeva MO, Bagramyan K, Telkov MV, Mukamolova GV, Young M, Kell DB & Kaprelyants AS (2002) Formation and resuscitation of "non-culturable" cells of Rhodococcus rhodochrous and *Mycobacterium tuberculosis* in prolonged stationary phase. *Microbiology* **148**: 1581-1591.

Singla R, Sharma SK, Mohan A*, et al.* (2010) Evaluation of risk factors for antituberculosis treatment induced hepatotoxicity. *Indian J Med Res* **132**: 81-86.

Snapper SB, Melton RE, Mustafa S, Kieser T & Jacobs WR, Jr. (1990) Isolation and characterization of efficient plasmid transformation mutants of *Mycobacterium smegmatis*. *Mol Microbiol* **4**: 1911-1919.

Srisailam S, Kumar TK, Srimathi T & Yu C (2002) Influence of backbone conformation on protein aggregation. *J Am Chem Soc* **124**: 1884-1888.

Stanley NR & Lazazzera BA (2004) Environmental signals and regulatory pathways that influence biofilm formation. *Mol Microbiol* **52**: 917-924.

Stapleton MR, Horsburgh MJ, Hayhurst EJ*, et al.* (2007) Characterization of IsaA and SceD, two putative lytic transglycosylases of *Staphylococcus aureus*. *J Bacteriol* **189**: 7316-7325.

Swaminathan S & Rekha B (2010) Pediatric tuberculosis: global overview and challenges. *Clin Infect Dis* **50 Suppl 3**: S184-194.

Telkov MV, Demina GR, Voloshin SA*, et al.* (2006) Proteins of the Rpf (resuscitation promoting factor) family are peptidoglycan hydrolases. *Biochemistry (Mosc)* **71**: 414-422.

Tufariello JM, Chan J & Flynn JL (2003) Latent tuberculosis: mechanisms of host and bacillus that contribute to persistent infection. *Lancet Infect Dis* **3**: 578-590.

Tufariello JM, Jacobs WR, Jr. & Chan J (2004) Individual *Mycobacterium tuberculosis* resuscitation-promoting factor homologues are dispensable for growth in vitro and in vivo. *Infect Immun* **72**: 515-526.

Tufariello JM, Mi K, Xu J*, et al.* (2006) Deletion of the *Mycobacterium tuberculosis* resuscitation-promoting factor Rv1009 gene results in delayed reactivation from chronic tuberculosis. *Infect Immun* **74**: 2985-2995.

Tully G, Kortsik C, Hohn H*, et al.* (2005) Highly focused T cell responses in latent human pulmonary *Mycobacterium tuberculosis* infection. *J Immunol* **174**: 2174-2184.

van Pinxteren LA, Cassidy JP, Smedegaard BH, Agger EM & Andersen P (2000) Control of latent *Mycobacterium tuberculosis* infection is dependent on CD8 T cells. *Eur J Immunol* **30**: 3689-3698.

Vollmer W & Bertsche U (2008) Murein (peptidoglycan) structure, architecture and biosynthesis in *Escherichia coli*. *Biochim Biophys Acta* **1778**: 1714-1734.

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

Vollmer W, Joris B, Charlier P & Foster S (2008) Bacterial peptidoglycan (murein) hydrolases. *FEMS Microbiol Rev* **32**: 259-286.

WHO (2006) The stop TB strategy.

WHO (2009) Global tuberculosis control: A short update to the 2009 report.

WHO (2009) Global tuberculosis control.

Wu X, Yang Y, Han Y*, et al.* (2008) Effect of recombinant Rv1009 protein on promoting the growth of *Mycobacterium tuberculosis*. *J Appl Microbiol* **105**: 1121-1127.

Yang Z, Rosenthal M, Rosenberg NA*, et al.* (2011) How dormant is *Mycobacterium tuberculosis* during latency? A study integrating genomics and molecular epidemiology. *Infect Genet Evol* **11**: 1164-1167.

Young DB, Gideon HP & Wilkinson RJ (2009) Eliminating latent tuberculosis. *Trends Microbiol* **17**: 183-188.

Zhu W, Plikaytis BB & Shinnick TM (2003) Resuscitation factors from mycobacteria: homologs of *Micrococcus luteus* proteins. *Tuberculosis (Edinb)* **83**: 261-269.