View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Wits Institutional Repository on DSPACE

The Fitness Costs of Drug Resistance Mutations in Mycobacteria

Anastasia Sideris Koch

A dissertation submitted to the Faculty of Science,

University of the Witwatersrand, Johannesburg,

for the degree of Master of Science

January, 2011.

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 for any degree or examination at any other University.

Anastasia Sideris Koch

16 August 2011

Date

Abstract

The increasing emergence of drug-resistant pathogens poses a major threat to public health. Although influenced by multiple factors, resistance is often associated with mutations in drug target-encoding or associated genes. The potential fitness cost of such resistance mutations is, in turn, a key determinant of the spread of drug-resistant strains. Rifampicin (RIF) is a frontline anti-tuberculosis agent that targets the *rpoB*-encoded β-subunit of the DNA-dependent RNA polymerase (RNAP). RIF resistance (RIF^R) maps primarily to mutations in *rpoB* that might be expected to affect transcription and so the ability of the organism to cause disease. Accordingly, numerous studies have assessed the impact of RIF^R on key fitness indicators in pathogens including *Mycobacterium tuberculosis* (MTB). In contrast, the specific consequences of RIF^R for bacterial physiology remain poorly understood. Notably, previous studies of the effects of RIF^R-associated *rpoB* mutations on mycobacterial physiology have been conducted using strains generated by RIF exposure, without accounting for the potential impact of second-site mutations that may compensate for fitness costs or contribute to drug resistance. In this study, site-directed mutagenesis and allelic exchange were employed to generate a panel of *M. smegmatis* (MSM) strains containing clinically-relevant RIF^R-associated point mutations. Importantly, this methodology enables the introduction of *rpoB* mutations into defined strain backgrounds in the complete absence of RIF. Using this approach, we constructed "RIF naive" MSM *rpoB* mutant strains carrying either an S531L or H526Y mutation. The resulting mutants were 100-fold less susceptible to RIF than the isogenic, parental strain. Notably, the inclusion of selected efflux inhibitors in susceptibility assays had little impact on mutant susceptibility to RIF. In contrast, restoration of the wild-type allele returned the observed susceptibility to parental levels, thereby providing strong evidence of the sufficiency of a single *rpoB* mutation for clinical RIF^R in mycobacteria. Competitive growth assays utilizing the S531L mutant and the parental strain exposed a growth defect for the S531L mutant. However, discriminating between wild-type and mutant *rpoB* strains proved a significant technical challenge, again highlighting the difficulties associated with inferring *in vivo* fitness from *in vitro* assays conducted under a limited number of different conditions. In summary, our results suggest the benefit of a deeper exploration of the physiological and fitness implications of RIF^R-associated mutations. In addition, in coupling a system which enables an evaluation of the physiological consequences of drug resistance-associated mutations with evolutionary analyses, we provide preliminary evidence of the benefits of a multipronged approach to elucidating the physiological implications of drug resistance in MTB.

A scientist in his/her laboratory is not only a technician: (s)he is also a child placed before natural phenomena which impress him/her like a fairy tale.

- Marie Curie

It’s not courage if you’re not scared; it’s not faith if you don’t doubt

- Anonymous

Acknowledgements

Dr Digby Warner provided me with brilliant supervision and wise mentorship. For teaching me the value of rigourous scientific inquiry, while guiding me through an exploration of the field, I am deeply grateful. My sincere thanks go to Prof Valerie Mizrahi for co-supervising this project. Her insights and scientific vision have been invaluable to my project, and to my development as a young researcher.

I gratefully acknowledge the National Research Foundation, The Molecular Mycobacteriology Research Unit (MMRU), and the University of the Witwatersrand for project and bursary funding. In addition I would like to thank South African TB HIV Training (SATBAT) for funding a portion of this project, and for generous funding to attend the Johns Hopkins Summer School of Biostatistics and Epidemiology.

In the last three months Dr Pierre Durand, to whom I extend my gratitude, advised me on aspects of my project which deal with evolution. I greatly look forward to further collaboration.

I have learnt something from each of the staff and students at the MMRU. I am grateful for their support and for making themselves available to give technical and conceptual advice. Most of all I thank them for the laughter we shared and for the friendships that we have developed.

My deepest gratitude is reserved for my family. To my mom, for keeping it together against all odds, and for teaching me to always get myself up again with integrity. I'd like to thank my dad for showing me how to do this with a smile on my face and for teaching me the value of words. I would like to thank my brother for being my greatest friend (and for saving me, when things needed fixing). Finally, I would like to thank Jade, for actually listening, and for reminding me to always to stop and question why I'm doing.

Table of Contents

DECLARATION i

Abstract ii

Acknowledgements iv

Table of Contents v

List of Figures viii

List of Tables x

List of Abbreviations xi

1. Introduction 1

1.1 Tuberculosis 1

1.2 Drug resistance in mycobacteria 2

1.3 RIF: Significance and mechanism of action 4

1.4 RIF resistance: RIF^R-associated mutations in *rpoB* 7

1.5 Additional mechanisms of RIF^R 13

1.5.1 Efflux mediated drug resistance 14

1.5.2 Compensatory mutation? 15

1.5.3 Drug resistance and increased mutagenesis 18

1.6 The implications of RIF^R-associated mutations for fitness: molecular epidemiology and evolutionary analyses 21

1.7 The effect of *rpoB* mutations on global transcriptional regulation: Changes in physiology 23

1.8 Specific aims of for this project 28

1.8.1 Overall objective 28

1.8.2 Hypothesis 28

1.8.3 Specific aims 28

2. Methods 29

2.1 Selection of *rpoB* mutations to include in the study 29

2.2 Bioinformatic analysis of *rpoB* 29

2.3 General Microbiological and DNA manipulation techniques 30

2.3.1 General bacterial growth conditions 30

2.3.2 DNA extractions 35

2.3.3 Agarose electrophoresis 38

2.3.4 DNA manipulation and cloning 40

2.3.5 Transformation of cells and electroporation 42

2.3.6 Polymerase chain reaction (PCR) ... 43
2.4 Specific strategies used in this project ... 45
2.4.1 Generation of suicide vectors and plasmid constructs used in this study ... 45
2.4.2 Generation of MSM mutants containing RIF^R-associated mutations: allelic exchange ... 53
2.4.3 Construction of *rpoB* containing pOLYG plasmids and electroporation into MSM 62
2.4.5 RIF susceptibility (RIF^S) analyses ... 64
2.4.6 Evaluation of the physiological effects of RIF^R-associated mutations: competition assays ... 67
3. Results ... 68
3.1 Selection of *rpoB* mutations to study in this project ... 68
3.2 The use of MSM as a study organism ... 70
3.3 Genomic context, sequence and structural similarity of MSM RpoB and MTB RpoB: A bioinformatic analysis ... 72
3.4 Construction of suicide vector plasmids for allelic exchange ... 80
3.4.1 Modification of p2NIL to remove *Acc*65I and *Bam*HI restriction sites ... 81
3.4.2 Construction of a template for SDM ... 82
3.4.3 Introduction of RIF^R-associated point mutations: SDM ... 83
3.4.4 Introduction of the 30 bp deletion: Xover-PCR ... 84
3.4.5 AS-PCR to screen for mutations ... 87
3.5 Allelic exchange to generate MSM strains containing RIF^R-associated mutations ... 90
3.5.1 The first step of allelic exchange: Isolation of SCOs ... 91
3.5.2 Orientation of SCOs is important for *rpoB* expression ... 91
3.5.3 The second step of allelic exchange: Generation of DCOs ... 94
3.6 The impact of gene dosing for RIF^R ... 97
3.6.1 RIF^S of merodiploid strains: SCOs containing the S531L or H526Y mutation are resistant to RIF ... 97
3.6.2 The impact of gene dosing of *rpoB* on RIF^S: increased copy number of *rpoB* does not confer RIF^R ... 99
3.6.3 SCOs containing two copies of the wild-type *rpoB* gene are susceptible to RIF ... 100
3.6.4 Expression of *rpoB* on pOLYG – a multicopy episomal plasmid ... 100
3.6.5 Expression of *rpoB* from pOLYG is toxic in mycobacteria ... 103
3.7 RIF^S of MSM mutants containing single RIF^R-associated mutations ... 105
3.7.1 Mutants containing S531L or H526Y mutations are able to grow on solid media containing relatively high concentrations of RIF ... 105
3.7.2 MABA's were useful for MIC_{RIF} determination and for evaluation of physiological effects of RIF^R-associated mutations ... 107

3.7.3 MABA's in 7H9 OADC media reveal increased MIC_{RIF} for *rpoB*^{S531L}and *rpoB*^{H526Y} .109
3.7.4 The EPI's, verapamil and reserpine, do not impact MIC_{RIF} of *rpoB*^{S531L}................111
3.7.5 Restoration of the wild-type allele in *rpoB*^{S531L} restores RIF^{S} 112
3.7.6 Evaluation of potential carbon source specific effect on MIC_{RIF} of *rpoB*^{S531L} mutants 113
3.8 Competition assays reveal a modest fitness defect for *rpoB*^{S531L}...............................115
3.8.1 Concerns regarding the use of RIF^{R} to differentiate between wild-type and the *rpoB*^{S531L} mutant 118
4. Discussion..120
4.1 Advantages (and disadvantages) of the study design for this project........................122
4.2 Was MSM a good study organism for this project?..126
4.3 Merodiploidy and RIF^{R}..128
4.4 Single S531L or H526Y point mutations in the *rpoB* gene drastically decrease RIF^{S}: does this infer sufficiency for RIF^{R}? 129
4.5 Are there other contributors to RIF^{R}? 132
4.6 Expression of *rpoB* from an episomal plasmid is toxic in MSM: Does this impact RNAP subunit stoichiometry? 135
4.7 Implications of RIF^{R}-associated mutations for mycobacterial physiology..................136
4.7.1 Carbon source-specific effects for the MIC_{RIF} of wild-type but not mutant strains 137
4.7.2 A slight fitness defect for *rpoB*^{S531L} is revealed by competitive growth experiments 138
4.8 Complexity of fitness assay and physiological characterisation 140
4.9 A pathway to clinical relevance?...143
4.10 Concluding remarks and prospective outlook 147
5. Appendices..149
5.1: RIF^{R}-associated mutations included in this study 149
5.2 Bacterial growth media and supplements 151
5.3 DNA extraction solutions and DNA modifying enzymes..155
5.4 Agarose gel electrophoresis..158
5.5. Primers used in this study 159
6. References 161

List of Figures

Figure 1.1	The RRDR of a selection of pathogenic bacteria..	10
Figure 2.1	Schematic summary of the cloning strategy used in this project…………….	43
Figure 2.2	Schematic representation of allelic exchange process..............................	52
Figure 2.3	PCR strategy to determine orientation of SCO strains...............................	54
Figure 2.4	Generation MSM strains containing RIF^R-associated mutations in *rpoB*…..	58
Figure 2.5	Schematic representation of preparation of MABA 96-well plates………….	63
Figure 3.1	The genomic context of *rpoB* is conserved in mycobacteria......................	71
Figure 3.2	The structural domains of RpoB..	74
Figure 3.3	*Acc*65I and *Bam*HI sites were removed from p2NIL.................................	79
Figure 3.4	Schematic representation of the SDM strategy used to introduce point mutations into the *rpoB* gene..	81
Figure 3.4(b)	Gel images of final SDM and Xover PCR fragments generated cloning……………………………………………………………………………………….	83
Figure 3.5	Schematic representation of the Xover-PCR strategy used to introduce the 30 bp deletion into the *rpoB* gene..	83
Figure 3.6	Representative suicide vector for allelic exchange....................................	84
Figure 3.7	Schematic of AS-PCR strategy used to screen for suicide plasmids............	86
Figure 3.8	AS-PCR primers can detect the S531L mutation in a suicide construct…….	87
Figure 3.9	Allelic exchange can generate up- and down-stream SCOs........................	89
Figure 3.10	Up- and down-stream specific primer sets facilitate determination of SCO orientation..	91
Figure 3.11	Colony one contains an S531L mutation in *rpoB* and is the $rpoB^{S531L}$ DCO..	94
Figure 3.12	The $rpoBsco^{S531L/WT}$, and $rpoBsco^{WT/H526Y}$ SCO strains are resistant to RIF...	96
Figure 3.13	The $rpoBsco^{WT/WT}$ mutant is as susceptible to RIF as wild-type MSM.........	98
Figure 3.14	Reduced transformation efficiency of pG*rpoB* constructs..........................	98
Figure 3.15	Increased copy number of *rpoB* on pOLYG does not confer RIF^R...............	100
Figure 3.16	Over-expression of *rpoB* leads to plasmid re-arrangement........................	102
Figure 3.17	S531L mutants are resistant to high levels of RIF......................................	105
Figure 3.18	S531L mutants can grow on RIF_{400}..	106
Figure 3.19	Decreased RIF-susceptibility of $rpoB^{S531L}$ and $rpoB^{H526Y}$ but not$rpoB^{L521L}$	109

mutants..

Figure 3.20 Restoration of wild-type *rpoB* in *rpoB*^{S531L} restores RIF-susceptibility.......... 112
Figure 3.21 The *rpoB*^{S531L} mutant has a fitness defect when grown in competition with wild-type MSM in 7H9 OADC.. 112
Figure 3.22 Growth on RIF_{200} plates does not represent the full extent of colonies present in the competition culture... 117
Figure 4.1 Factors important in drug efficacy and the development of drug resistance... 134
Figure 5.1 Molecular weight marker III and V (MW III and MW V).............................. 158

List of Tables

Table 2.1 Bacterial strains used in this study.. 28

Table 2.2 Plasmids used in this study………………………………………………………………….. 30

Table 3.1 RIF^R-associated mutations to be examined in this project.......................... 65

Table 3.2 Transformation efficiency of pG*rpoB* constructs……………………………….….. 98

Table 3.3 MIC_{RIF} of S531L mutant strains in the presence of EPI's............................. 110

Table 3.4 $MIC_{RIF}$of S531L mutant strains in varied carbon sources............................ 113

Table 4.1 Plasmid constructs and mycobacterial mutants generated in this study.... 123

Table 5.1.1 Literature reviewed to determine the frequency of RIF^R-associated mutations in clinical and laboratory-derived MTB isolates……………………. 149

Table 5.1.2 Frequency of mutations associated with RIF^R in laboratory-derived RIF^R strain.. 150

Table 5.2.1 *Escherichia coli* growth media.. 151

Table 5.2.2 *Mycobacterium smegmatis* growth media.. 152

Table 5.2.3 Medium supplements... 153

Table 5.3.1 DNA Extraction solutions.. 155

Table 5.3.2 Restriction endonucleases.. 156

Table 5.3.3 Other DNA modifying enzymes.. 157

Table 5.3.4 DNA polymerases.. 157

Table 5.4.1 Agarose gel electrophoresis components.. 158

Table 5.5.1 Primers used in this study.. 169

List of Abbreviations

Abbreviation	Abbreviated term
AMP	Ampicillin
AS-PCR	Allele specific PCR
BLAST	Basic Local Alignment Search Tool
BLOSUM	Blocks of amino acid substitution matrix
bp	Base pairs
CFU	Colony forming units
CTAB	Cetyltrimethylammoniumbromide
DCO	Double cross over
DMSO	Dimethyl suphoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxydribonucleotide triphosphate
dsH20	Distilled sterile water
EMB	Ethambutol
EPI	Efflux pump inhibitor
ERM	Erythromycin
GS	Glucose-NaCl
HIV	Human Immunodeficiency Virus
HYG	Hygromycin
INH	Isoniazid
JCVI CMR	J. Craig Venter Institute Comprehensive Microbial Resource
KAN	Kanamycin
LTBI	Latent TB infection
MABA	Mycobacterial Alamar Blue Assay
MDR	Multidrug-resistant
MIC	Minimum inhibitory concentration
MIC_{RIF}	Minimum inhibitory concentration for RIF
MMRU	Molecular Mycobacteriology Research Unit
MSA	Multiple sequence alignment
MSM	Mycobacterium smegmatis
MT	Mutation/mutant
MTB	Mycobacterium tuberculosis
OADC	Oleic-acid-dextrose-albumin
OD_{600}	Optical density at 600nm
PCR	Polymerase chain reaction
ppGpp	Guanosine tetraphosphate
PROMALS3D	PROfile Multiple Alignment with Local Structure and 3D
PSI-BLAST	Protein specific iterative BLAST
PSIPRED	PSI-Prediction
^{R}	Resistant/resistance
RIF	Rifampicin
RNA	Ribonucleic acid
RNAP	RNA polymerase
ROS	Reactive oxygen species

RPF	Resuscitation promoting factor
RRDR	RIF Resistance Determining Region
^{S}	Susceptible/susceptibility
SCO	Single cross over
SDM	Site-directed-mutagenesis
ssDNA	Single-stranded DNA
STR	Streptomycin
TAE	Tris-acetate-EDTA buffer
TB	Tuberculosis
TBDReaMDB	TB Drug Resistance Mutation Database
TRCF	Transcription-repair coupling factor
TE	Tris-EDTA buffer
TEC	Transcriptional elongation complex
UV	Ultra violet
WHO	World Health Organisation
WT	Wild-type
XDR	Extensively drug-resistant
Xover-PCR	Cross over-PCR

1. Introduction

1.1 Tuberculosis

Tuberculosis (TB) is responsible for 1.8 million deaths annually (WHO, 2009). The causative agent of this devastating disease, *Mycobacterium tuberculosis* (MTB), is a gram positive, facultative anaerobe that is dependent on its human host for survival (Daffe & Etienne, 1999). The obligatory requirement for the human host means that physiological processes important for survival are also essentially virulence properties. Remarkably, it is estimated that up to one third of the world's population is infected by this highly successful pathogen (Zhang, 2005), however not all of these infections result in active disease. The risk of developing active disease in infected individuals is fundamentally dependent on the immune status of the individual and is distinguished by intricate interactions between MTB and host immune cells (Peyron *et al.*, 2008). Most immunocompetent individuals are able to control infection successfully, but this does not always result in complete clearance of MTB cells and the bacilli can persist in a poorly understood dormant state for extended periods (Stewart *et al.*, 2003).

MTB is inhaled into the lungs and, once in the alveoli, internalized by lung macrophages (Peyron *et al.*, 2008). An accumulation of macrophages, dendritic cells and lymphocytes is triggered, which results in granuloma formation (Peyron *et al.*, 2008). In healthy individuals, the interplay between MTB and host cellular and adaptive immune cells results in asymptomatic infection as granulomas separate MTB cells from the surrounding tissue (Peyron *et al.*, 2008), and a state of latency ensues (Stewart *et al.*, 2003). The ability of MTB to interact with – and manipulate – host immune functions allows latent infection to endure. Traditionally it was thought that granulomas were protective against TB, but a recent report indicates that MTB may actually recruit immune cells to form granulomas, wherein the bacilli are able to persist in a protected environment until conditions are favorable for reactivation (Davis & Ramakrishnan, 2009). The precise physiological state of bacilli in latent TB infection (LTBI) is unknown. Latency describes TB with no symptoms of active disease and MTB cells during this state are often referred to as dormant (Rustad, 2008), or existing in a non-replicative state. However, a recent study indicates that replication is in fact taking place during latent stages of MTB infection, at least in a mouse

model (Gill *et al.*, 2009). It is thought that host immune functions keep dormant mycobacteria in a state of slow growth and reduced metabolism (Boshoff, 2008). Conceivably, a kind of equilibrium is attained (Scanga & Flynn, 2010). Reactivation of MTB infection can occur if the host immune system is suppressed and the granuloma structure is compromised, thus allowing MTB cells to escape and begin actively replicating. Immunocompetent individuals infected with MTB have a 10 % chance of developing active TB within their lifetime however co-infection with HIV increases this risk to 10 % per year (Rustad, 2008). Developing countries bear the brunt of the TB burden, with overcrowding, malnutrition, and HIV as the biggest risk factors for MTB infection. Unfortunately, these regions have the fewest resources to manage this treatable, but deadly disease.

1.2 Drug resistance in mycobacteria

The serendipitous discovery of penicillin, and subsequently other classes of antibiotics, revolutionized treatment of infectious diseases (Cars *et al.*, 2008). Of great consequence was the ability to manage TB which was once the blight of industrialized nations (Zumla *et al.*, 2009). Microbes have developed resistance to almost every antibiotic that has reached mainstream use and in recent times, there has been a notable increase in the prevalence of antibiotic resistant pathogens (Cars *et al.*, 2008, Martinez *et al.*, 2009). Unfortunately, TB is no exception, and the sharp increase in the emergence of drug-resistant MTB strains threatens to thwart progress made by chemotherapeutic control.

According to The World Health Organisation (WHO) almost 500 000 people develop multidrug-resistant TB (MDR-TB) every year (WHO, 2009). This form of TB is highly resistant to conventional front-line drugs and places additional strain on public health care systems. MDR-TB is defined as disease resulting from MTB infection by strains that are resistant to frontline TB drugs, isoniazid (INH) and rifampicin (RIF) (Gagneux, 2006c). Treatment for drug sensitive TB is less than optimal, partly because the onerous 6 month treatment period leads to noncompliance, especially as patients start feeling better. MDR-TB treatment regimens are even longer, and the drugs used more expensive and associated with extensive side-effects. Interruption of treatment is one of the factors which allow for the development and maintenance of drug-resistant MTB strains; however, biological factors

are likely to play a significant role as well (Warner & Mizrahi, 2006). Worryingly, extensively drug-resistant TB (XDR-TB) cases are at their highest levels ever (Jassal & Bishai, 2009). These strains are resistant to INH, RIF, one of the flouroquinolones, and at least one of the injectable second line antibiotics (Jassal & Bishai, 2009). The load is especially devastating in countries where the burden of TB is highest and health care systems are already struggling to manage drug-susceptible TB, as well as other infectious diseases such as HIV and malaria. XDR-TB gained notoriety after a deadly outbreak in the Tugela Ferry region in KwaZulu-Natal, South Africa (Gandhi *et al.*, 2006). South Africa reported 16 000 cases of MDR-TB in 2007 (WHO, 2009). A recent study at a separate KwaZulu-Natal hospital revealed that 50 % of decedents had culture positive TB at the time of death, 94 % were HIV+ and, alarmingly, 17 % of TB cases were classified as MDR (Cohen *et al.*, 2010). These numbers illustrate the deadly co-association between TB and HIV and emphasize the need for a better understanding of the factors that contribute to the spread of MDR-TB with a view, ultimately, to the development of better HIV/TB management strategies (Kaufmann & Walker, 2009).

Many organisms possess intrinsic features that reduce their susceptibility to antibiotics, such as the ability to form biofilms (Donlan & Costerton, 2002) or permeability barriers (Nikaido, 1994, Nikaido, 2001, Martinez *et al.*, 2009). However, these are *tolerance mechanisms* and should be differentiated from antibiotic resistance which instead denotes the heritable acquisition (by horizontal gene transfer or mutations in target or related genes (Martinez *et al.*, 2009)) of the ability to withstand exposure to antibiotics at therapeutic concentrations. In many bacteria, drug resistance determinants are carried on mobile genetic elements. In contrast, the emergence of drug resistance in MTB is distinctive in that resistance is associated with point mutations and chromosomal rearrangements (Mariam *et al.*, 2004, Gillespie, 2002, Gagneux, 2009, Gagneux *et al.*, 2006c), a feature that results in part from the ecological isolation of MTB (within granulomas) during host infection (Gillespie, 2002). Contrary to some other bacterial and mycobacterial species (Stinear *et al.*, 2004), MTB does not possess plasmids (Derbyshire & Bardarov, 2000): very little, if any, evidence exists for horizontal gene transfer in the MTB complex (Smith *et al.*, 2009), and genetic drift is primarily responsible for diversification and adaptation of this group of organisms (Hershberg *et al.*, 2008). Although the absence of gene transfer is a defining

characteristic of MTB, the propensity to develop chromosomal mutations associated with drug resistance is not limited to mycobacteria and, in several common pathogens high level resistance is associated with mutations in target genes.

Various factors contribute to the spread of MDR disease (Borrell & Gagneux, 2009) and public health and socio-economic issues have been relatively well explored. In contrast, information on the intrinsic biological aspects is lacking. We are particularly interested in the physiological effects of drug resistance-associated mutations in mycobacterial species. In this project we focus on RIF, a frontline anti-tubercular agent whose target is a global transcriptional regulator, and so might increase the possibility of drug-independent effects of resistance-associated mutations. The discussion that follows features the structural and physiological consequences of RIF resistance (RIF^R) associated *rpoB* mutations – contextualized by what is known about RIF^R in MTB and the consequences for TB management.

1.3 RIF: Significance and mechanism of action

RIF diffuses relatively well through tissue and cell membranes which makes it a useful antibiotic in the treatment of bacterial infection (Campbell *et al.*, 2001). Many antibiotics require active growth and metabolism to kill bacteria (Lipsitch & Levin, 1997). This is especially pertinent for MTB, which has a penchant for long-term survival in the host. Within the macrophage phagosomal environment bacilli – faced by a myriad of stresses (host immune function, nutrient stress and hypoxia (Schnappinger *et al.*, 2003) – adapt to an altered, low metabolic state (Connolly *et al.*, 2007). The importance of RIF is therefore established by its efficacy in sterilizing slowly metabolising persister populations (Mitchison, 2000).

RNA transcription is an essential constituent of genetic regulation, and is disrupted by RIF. In a tight complementary fit, RIF binds to the β-subunit of RNA polymerase (RNAP), close to the DNA/RNA channel (Campbell *et al.*, 2001). Eukaryotic RNAP enzymes comprise up to 14 subunits with a molecular weight of ~500kDa (Adekambi *et al.*, 2009, Borukhov & Nudler, 2008, Lane & Darst, 2010a, Lane & Darst, 2010b). Bacteria possess a simpler variant, with a

core structure comprising α, β, β’and ω subunits; upon binding of a specific σ factor the functional holoenzyme is generated (Adekambi *et al.*, 2009, Campbell *et al.*, 2002, Zhang *et al.*, 1999). RNAPs exhibit high levels of conservation, with the large subunits (β and β’; encoded by *rpoB* and *rpoC* respectively) the most highly conserved among organisms in all divisions of life (Adekambi *et al.*, 2009, Lane & Darst, 2010a, Lane & Darst, 2010b, Werner & Grohmann, 2011). The high level of evolutionary relatedness between RNAPs indicates the importance of this enzyme for gene expression and establishes bacterial RNAP as a useful tool in the study of transcriptional processes (Campbell *et al.*, 2001, Adekambi *et al.*, 2008, Adekambi *et al.*, 2009).

The crystal structure of core RNAP from *Thermus aquaticus* in complex with a RIF molecule reveals that the closest RIF gets to the active site of RNAP is 12.1 Å (Campbell *et al.*, 2001). Consequently, it appears that RIF does not inhibit transcription by directly binding to the RNAP active site. Instead, it is widely accepted that RIF inhibits initial events in transcription, although structural studies indicate that RIF does not directly obstruct specific promoter binding nor open complex formation nor formation of the first phosphodiester bond (Campbell *et al.*, 2001). Results from the *T. aquaticus* model suggest that RIF creates a steric clash between the phosphate group of the subsequent 5’ nucleotide and in this way prevents bond formation between the second and third ribonucleotide triphosphates in a growing transcript. That is, RIF effectively “plugs” the exit site of the growing RNA strand, preventing chain elongation and thereby halting transcription (Campbell *et al.*, 2001). The first phosphodiester bond can, however, be formed; therefore RIF does not impede catalytic activity, substrate binding or the translocation properties of RNAP. The initial phase of RNA synthesis is referred to as abortive transcription, whereby short RNA transcripts are synthesized, released and re-synthesized. The precise mechanisms are not well-defined, but this process is important for promoter escape and progression to the elongation phase of transcription (Werner & Grohmann, 2011, Borukhov & Nudler, 2008, Borukhov & Nudler, 2003). Notably, when an RNA chain longer than 3 – 5 nt has been formed, RIF is no longer able to inhibit transcription – possibly because longer chains prevent RIF binding (Campbell *et al.*, 2001). It is therefore possible, that by preventing bond formation in early transcription, RIF interrupts the abortive cycling process, thus scuppering another important mechanistic transcriptional process.

Biochemical evidence indicates that the binding affinity of RIF for *T. aquaticus* RNAP is much lower than to that of *Escherichia coli* (Campbell *et al.*, 2001), which results in higher RIF concentrations being required to inhibit *T. aquaticus* RNAP. Furthermore, a recent analysis of the sequence – and, in a related study, the structure – of almost all available bacterial RNAPs revealed that small, lineage-specific insertions have important consequences for RNAP regulation and function (Lane & Darst, 2010a, Lane & Darst, 2010b). Therefore while RNAP is highly conserved among prokaryotes, structural and functional comparisons of RNAPs from different organisms must be made with prudence (Kuznedelov & Severinov, 2009). Given the global increase in MDR-TB, it is critical that the precise mechanisms of interaction between RIF and MTB RNAP are identified. This may yield insights into how RIF can be derivitised to improve activity against MTB.

An RNAP holoenzyme is formed – and transcription is initiated – when an appropriate σ subunit binds the core enzyme and induces significant conformational changes (Murakami *et al.*, 2002a; Murakami *et al.*, 2002b). By directly binding to promoter elements, σ factors guide transcription initiation from appropriate DNA templates. Double stranded DNA at the σ binding site is melted, and a transcriptional bubble is formed (Murakami *et al.*, 2002a; Murakami *et al.*, 2002b; Borukhov *et al.*, 2003). To accomplish this, the housekeeping σ factor of *E. coli*, σ^{70}, uses free energy and thermal fluctuations that accumulate during RNAP conformational shifting (Borukhov *et al.*, 2008). Interestingly, eukaryotic and some alternative bacterial σ factors use energy released by ATP hydrolysis to facilitate the process. As a result of DNA unwinding, a stable open promoter complex is formed which allows initial RNA synthesis to begin (Borukhov *et al.*, 2003). Although direct interactions with core RNAP primarily occur via the β' subunit, the β subunit is also important for σ factor binding (Murakami *et al.*, 2002b; Nickels *et al.*, 2005). The initial phase of RNA synthesis is referred to as abortive transcription, whereby short RNA transcripts are synthesized, released and re-synthesized (Murakami *et al.*, 2002a; Nickels *et al.*, 2005; Borukhov *et al.*, 2008; Goldman *et al.*, 2009). Competition for chemical space between the σ subunit and the newly synthesised RNA chain facilitates σ dissociation from the core RNAP enzyme, which is generally – but not always – required for progression to elongation phase (Murakami *et al.*, 2002b; Nickels *et al.*, 2005; Borukhov *et al.*, 2008). Principle σ factors regulate the expression of housekeeping genes, and show significant homology to the *E. coli*

σ^{70} subunit – in mycobacteria this σ factor is encoded by *sigA* (Wu *et al.*, 2004). Bacteria have varying numbers of alternative σ factors, which selectively bind to promoters to express genes important for adaptation to a particular set of environmental stimuli (Campbell *et al.*, 2002; Borukhov *et al.*, 2003).

Traditionally, studies examining transcriptional processes made use of RIF as a transcriptional inhibitor. These studies revealed that mutations associated with RIF^R result in differential σ unit binding, as well as distinct promoter binding (Glass *et al.*, 1986a, Glass *et al.*, 1986b, Nomura *et al.*, 1984). Furthermore, in *E. coli* higher concentrations of RIF are needed to inhibit transcription from σ^{70}- *versus* σ^{32}- dependent promoters (Wegrzyn *et al.*, 1998). Alternative σ factors may induce different conformational changes upon RNAP binding (Scott *et al.*, 2000, Wegrzyn *et al.*, 1998), therefore it is intriguing to consider the implications of these changes for RIF binding and transcriptional inhibition. A recent investigation in MTB indicates that RIF-mediated inhibition of transcription is similar for σ^A and σ^F containing RNAP, and that over-expression of σ^F does not contribute to tolerance of RIF (Hartkoorn *et al.,* 2010). However, increased expression of σ^E was observed in a clinical MTB strain carrying an S531L mutation. Moreover, increased σ^E expression correlated with increased ability to grow in the presence of INH. Upon σ factor binding, the β subunit region containing the RRDR is proximal to several parts of the σ factor (Owens *et al.*, 1998; Murakami *et al.*, 2002a; Nickels *et al.*, 2005). It is plausible that RIF^R-associated mutations may mediate binding of alternative σ factors that increase expression of genes involved in INH^R and in this way facilitate INH^R (Evans *et al.*, 2011, submitted). This highlights the possibility that specific *rpoB* mutations may encourage INH resistance (INH^R), and may explain why RIF mono-resistance is infrequently observed in the clinical setting (Evans *et al.*, submitted).

1.4 RIF resistance: RIF^R-associated mutations in *rpoB*

RIF^R-associated mutations were originally mapped in *E. coli* and provided a means for examining mechanistic aspects of transcription (Jin & Gross, 1988). These mutations are found in cluster I, II and III of the *rpoB* gene (Campbell *et al.*, 2001). In all bacterial species examined, a large proportion of RIF^R-associated mutations occur in the RIF Resistance

Determining Region (RRDR) – a 81bp region which forms part of cluster I of the *rpoB* gene (Ramaswamy & Musser, 1998, Jin & Gross, 1988) (Figure 1). Nonetheless, mutations in other regions of the *rpoB* gene have been observed (Heep *et al.*, 2001, Heep *et al.*, 2000a). Not all RIF^R MTB strains contain mutations in *rpoB* (Telenti *et al.*, 1993, Ramaswamy & Musser, 1998), and some organisms have additional mechanisms of RIF^R (Quan *et al.*, 1997, Newell *et al.*, 2006). Hence, throughout this document, these mutations are referred to as "RIF^R-associated" mutations, or equivalent. Unambiguous genetic analysis of the role of *rpoB* mutations in RIF^R represents one of the primary aims of this project; in particular, the need to address definitively the sufficiency of *rpoB* mutations for RIF^R.

Alternative mechanisms of resistance to common anti-TB chemotherapeutics have immense impact for public health strategies to control MDR-TB. Recently, the WHO approved the use of molecular diagnostic kits for diagnosis of MDR-TB. These kits rely on detection of a limited number of common drug resistance-associated mutations, particularly those associated with RIF^R. The tests are not exhaustive, and only the most frequently observed RIF^R-associated mutations are incorporated in the test. There may be other mutations in *rpoB* (or the RRDR) that contribute to RIF^R (Heep *et al.*, 2001, Heep *et al.*, 2000b), or other physiological mechanisms that contribute to RIF^R. A strain that is resistant to RIF but does not carry the specific mutations detected by the test, could be misdiagnosed and spread in the community. Cognizance of mechanisms of drug resistance, other than common point mutations, is essential when diagnosing drug-resistant TB, and developing public health control strategies.

Examination of clinical MTB isolates from various geographical regions indicates that certain RIF^R-associated mutations in *rpoB* occur more frequently than others. Moreover, a thorough analysis of the literature reveals the most – as well as the least – common mutations associated with both laboratory-derived and clinical RIF^R MTB isolates (see Results: section 3.1). Additionally, a database has been set up to document clinically observed RIF^R MTB mutations (Sandgren *et al.*, 2009). The RRDRs of four common bacterial pathogens are illustrated in Figure 1, which also provides the opportunity to compare this region. The RRDR was originally defined in *E. coli*, therefore, as per standard practice amino acid numbering corresponds to that of the *E. coli* sequence (Telenti *et al.*, 1993). The most

common mutation associated with RIF^R in MTB is a mutation at codon 531 of the *rpoB* gene (Sandgren *et al.*, 2009). In RIF^S strains, a serine residue is encoded at position 531; in RIF^R strains, a transition mutation (t**c**g→t**t**g) results in a leucine in this site (Ramaswamy & Musser, 1998). Notably this and another common mutation, H526Y (**c**ac→**t**ac), are both cytosine to thymine transitions (Mokrousov, 2004). Spontaneous cytosine deamination to uracil occurs readily in all cells; however, MTB is thought to be especially susceptible to deamination (Warner, 2010) as a result of exposure to reactive oxygen and nitrogen intermediates (ROI and RNI) which form part of the host defence against TB infection (Mokrousov, 2004).

Of twelve amino acids that surround the RIF binding pocket, eleven mutate in association with RIF^R – albeit at different frequency. The RRDR is highly conserved between bacteria, but shows lower levels of relatedness to archaebacterial and eukaryotic RNAP; this and the fact that RIF^R-associated mutations are common in this region, imply that this region is not essential for RNAP function, but that only certain changes might be tolerated (Campbell *et al.*, 2001). The only amino acid which interacts with RIF but for which RIF^R-associated mutations have not been observed is E565, which lies outside the RRDR. This residue is highly conserved across all orders of life, therefore mutations at this site may not be viable (Campbell *et al.*, 2001). Since examination of RIF^R-associated *rpoB* sequence is often limited to the RRDR, it's possible that mutations at the E565 site may have been overlooked (Campbell *et al.*, 2001). *M. smegmatis* (MSM) strains containing a 30 bp deletion within the RRDR have been isolated on multiple occasions at the MMRU in response to UV treatment and starvation (Digby Warner, unpublished work). In a separate study, RNAP from an analogous MSM strain was found to be highly recalcitrant to RIF-mediated inhibition (Malshetty *et al.*, 2010). Viability of MSM strains with a large deletion in *rpoB* supports the notion that this region is dispensable for RNAP function. Nevertheless, there may be significant physiological consequences for the loss of this region. Interestingly, RIF^R-associated substitutions always involve a change from a residue with a compact side chain to an amino acid with a more cumbersome one – or involve the addition or removal of a proline residue (Campbell *et al.*, 2001). The frequently isolated S531 and H526 residues are directly involved with hydrogen bonding to RIF (Campbell *et al.*, 2001).

Siddiqi *et al.* (2002) identified silent mutations in a fraction of clinical strains isolated in India. These mutations are noteworthy since they do not result in amino acid substitutions but were associated with RIF^R (Siddiqi *et al.*, 2002). Only the RRDR was sequenced; therefore, it is possible that factors other than mutations in the RDDR region contributed to RIF^R in these strains. Moreover, if this mutation were also found in susceptible strains, it would indicate that it is a phylogenetic marker and is not involved in RIF^R. However, if silent mutations are in fact associated with high level RIF^R, it is a fascinating and important area that warrants further study.

Using the crystal structure of *T. aquaticus* RNAP (Campbell *et al.*, 2001) as a modelling scaffold, the structural consequences of RIF^R-associated mutations in *Staphylococcus aureus* (O'Neill *et al.*, 2006) and, more recently, MSM (Malshetty *et al.*, 2010) were elucidated. These studies reveal that structural determinants and perturbations of RNAP structure as a result of RIF^R-associated mutations are generally consistent with experimental evidence (O'Neill *et al.*, 2006, Malshetty *et al.*, 2010). However, *in silico* modelling data must be interpreted with caution, since the quality of the model is dependent on the crystal structure upon which it is built. In *E. coli* RIF^R-associated mutations are known to have pleiotropic consequences for transcriptional mechanisms of elongation, pausing and termination (Jin *et al.*, 1988, Jin & Gross, 1991). For example, in *Bacillus subtilis*, a mutation to a basic residue at the Q513 position increases auto(de)regulation of NusG, and consequently increases termination (Ingham *et al.*, 2000). NusG is an important transcriptional elongation factor and deletion of this gene in *E. coli* is lethal. Early reports revealed accelerated elongation when NusG was over-expressed in *E. coli*, leading to the suggestion that NusG acts by suppressing termination (Burova *et al.*, 1995a; Burova *et al.*, 1995b). However, NusG has been shown to enhance transcriptional pausing in *B. subtilis* (Yakhnin *et al.*, 2008). NusG can have differential effects on transcription when examined *in vitro*, and this presents a difficulty for defining the nature of NusG function and regulation (Yakhnin *et al.*, 2008). On the whole, the specific function of NusG is dependent on the organism; the specific operon under investigation and it's interaction with other transcriptional factors (Borukhov *et al.*, 2005). In *B. subtilis* transcription of the *nusG* operon is regulated by the NusG protein. Mutations in *rpoB* analogous to those found in RIF^R mycobacteria promoted NusG-mediated repression of *nusG* transcription, thus decreasing

the availability of NusG for resolution of paused RNAP complexes (Ingham *et al.*, 2000) Given the importance of all the aforementioned processes for cell functionality, the implications of *rpoB* mutations are likely to be broader than merely inhibition of RIF binding. If RIF^{R}-associated mutations trigger severe impairments to RNAP structure and function, one would be compelled to consider whether and what mutations in associated genes are alleviating these costs, given that fitness costs are negligible in some clinical isolates (Gagneux *et al.*, 2006b).

ECO	F 505	F	G	S	S	Q	L	S	Q	F	M	D	Q	N	N	P	L	S	E	I	T	H	K	R	R	I	S	A	L	G	P	P	G 527
			d	p	r		p q r	f a	l p r v p k			n v g a y	d q	d				f y	v		r	y r l p n q d			c s h l		f y	v e	p h r	s d c v a			

MTB	F 424	F	G	T	S	Q	L	S	Q	F	M	D	Q	N	N	P	L	S	G	L	T	H	K	R	R	L	S	A	L	G	P	P	G 446
			d			h	p r	t r	l k p	l	i v	y e g a v		h	k		m l	l				y d q r p l e n g t c											

SAU	F 460	F	G	S	S	Q	L	S	Q	F	M	D	Q	A	N	P	L	A	E	L	T	H	K	R	R	L	S	A	L	G	P	P	G 482
					p		s		r l			y e g			l			v d				n y d n			h				s				

BSU	F 461	F	G	S	S	Q	L	S	Q	F	M	D	Q	T	N	P	L	A	E	L	T	H	K	R	R	L	S	A	L	G	P	P	G 483
		f	g	t	s		l	s				d g v					l	s	g	l	t	r y p d c		r	r		ls	a	l	g	p	p	g

Figure 1.1 The RRDR of a selection of pathogenic bacteria.

Mutations in the RRDR of distinct bacterial species were ascertained from relevant clinical and lab based publications (indicated in brackets after organism names). Wild-type sequences were downloaded from the NCBI protein database (http://www.ncbi.nlm.nih.gov/guide/proteins/) and are indicated in caps. There were some discrepancies between codon numbering as stated in specific publications and numbering as indicated in the NBCI protein database, however sequences were aligned, and numbered according to the publication. In text amino acid numbering is referred to as defined in *E. coli* (Telenti *et al.*, 1993) however, species specific numbering is shown in the figure for comparison. Organism abbreviations and NCBI accession numbers are as follows **ECO:** *E. coli* K12 subtr. MG1655 (NP_418414.1) (Jin & Gross, 1988, Garibyan *et al.*, 2003, Reynolds, 2000); **MTB:** *M. tuberculosis* H37Rv (GenBank: AAA20242.2) (Telenti *et al.*, 1993, Ramaswamy & Musser, 1998, Williams *et al.*, 1998); **SAU:** *S. aureus* subsp. aureus ED98 (YP_003281434.1) (O'Neill *et al.*, 2006, O'Neill *et al.*, 2000, Wichelhaus *et al.*, 2002); **BSU:** *B. subtilis s*subsp. Subtilis str 168 (ZP_03589766.1) (Nicholson & Maughan, 2002, Inaoka *et al.*, 2004).

1.5 Additional mechanisms of RIF^R

A variety of distinct point mutations have been documented which, in most cases, result in single amino acid substitutions that are associated with widespread RIF^R. Culture-based susceptibility testing techniques can yield highly discordant results (Van Deun *et al.*, 2009) and susceptibility breakpoints, the concentration of a drug above which a strain is defined as drug-resistant, may be too low (Gumbo, 2010). Therefore the WHO has approved the use of rapid diagnostic tests based on molecular identification of common RIF^R-associated mutations (Uys *et al.*, 2009). Although the majority of clinical RIF^R MTB isolates contain mutations in the RRDR region, it has been estimated that 5% exhibit high level RIF^R without evidence of mutations in the RRDR or the rest of the *rpoB* gene (Ramaswamy & Musser, 1998, Telenti *et al.*, 1993). If a strain that is resistant to RIF does not have any mutations in the RRDR, its susceptibility profile will be misinterpreted, which may lead to elevated dissemination of this strain. It is crucial that we gain a better understanding of just how tightly *rpoB* point mutations are associated with RIF^R. For other anti-TB drugs, the proportion of strains without identifiable target mutations is even higher (for example, only 30 – 68 % of ethambutol resistant (EMB^R) isolates contain target-associated mutations (Safi *et al.*, 2010)) and molecular diagnostics for other anti-TB drugs may therefore be even more inefficient. The ability of isolates without

identifiable target mutations to remain viable under high concentrations of RIF calls for a deeper understanding of the mechanisms involved.

Like most bacterial species, mycobacteria have intrinsic mechanisms by which to generate drug resistance. The waxy cell wall of mycobacteria confers a barrier to influx of toxic compounds, and provides a significant challenge to the development of new TB chemotherapeutics (Louw *et al.*, 2009). One inherent resistance mechanism in MSM to RIF was shown to be a result of ribosylation of RIF (Quan *et al.*, 1997) and certain *Bacillus* and *Nocardia* species also have enzymatic mechanisms that deactivate RIF by modifying its chemical structure (Tupin *et al.*, 2010). Moreover, proteins that bind to RNAP, such as RpbA in *Streptomyces coelicolor* and MsRpbA in MSM, confer some levels of RIF^R (Malshetty *et al.*, 2010, Newell *et al.*, 2006). However, MTB is killed at 100-fold lower concentrations of RIF than MSM (Alexander *et al.*, 2003, Changsen *et al.*, 2003, Piddock *et al.*, 2000), and there is little evidence of the same drug detoxifying mechanism in this pathogen. Interestingly, a recent investigation of the mutational profile of MTB to TMC207 (one of the most promising new TB therapeutic in the last 40 years) revealed that only 28 % of resistant isolates had mutations in the known target (Huitric *et al.*, 2010). The lack of mutations in the known target may indicate that there is more than one target for this drug, or that alternative mechanisms are involved in resistance development. A better understanding of alternative mechanisms for drug resistance is crucial if the efficacy of new (and old) chemotherapeutic regimens is to be preserved.

1.5.1 Efflux-mediated drug resistance

Analysis of the mycobacterial genome indicates that there are several putative drug exporters (Louw *et al.*, 2009), none of which has been thoroughly characterized (Reviewed in (Louw *et al.*, 2009) and (Viveiros *et al.*, 2003)). A recent model which considers the development of drug resistance suggests that efflux systems may “prime” bacteria to acquire resistance-associated mutations (Fange *et al.*, 2009). Indeed, some of features of this model have been demonstrated experimentally using the macrolide antibiotic, erythromycin (ERM) (Lovmar *et al.*, 2009, Moore & Sauer, 2008). The authors of the latter report offer a different explanation, however: specifically, that erythromycin

resistance (ERM^R)-associated ribosomal mutations alter translation of cell wall proteins, thereby reducing the intracellular concentration of ERM (Moore & Sauer, 2008). Nevertheless both studies imply that the acquisition of target mutations that decrease drug-target binding affinity is not enough to confer high level drug resistance that is observed in clinical settings. Notably, a clinical MTB strain belonging to the W Beijing family was shown to have increased efflux in response to RIF and INH (Jiang *et al.*, 2008). Phenothiazine compounds, chlorpromazine and thioridazine, are suggested to have clinically relevant anti-mycobacterial possibilities. It has been proposed that these compounds function by inhibiting bacterial efflux (Amaral *et al.*, 2008, Viveiros *et al.*, 2003); however, transcriptional responses of MTB to these compounds indicate that respiratory systems are inhibited (Boshoff *et al.*, 2004, Weinstein *et al.*, 2005, Yano *et al.*, 2006). The interplay between efflux and target mutations is of great relevance to elucidating the mechanisms of drug resistance and for contemporary TB control measures. For example, recent experimental evidence demonstrates that mutations commonly observed in clinical EMB^R MTB strains to confer lower levels of resistance when introduced to susceptible strains via homologous recombination (Safi *et al.*, 2010). This recapitulates the idea that target mutations, on their own, are unable to generate the high level resistance seen in clinical environments.

1.5.2 Compensatory mutation?

A compensatory mutation could be broadly defined as a mutation which reverses the deleterious effect of a distinct mutation or restores some functional or structural characteristic (Wilke *et al.*, 2003). Genetic background can have significant effects on observed fitness of strains containing drug resistance-associated mutations (Gagneux *et al.*, 2006a, Gagneux *et al.*, 2006b). Off-site mutations that distinguish one bacterial strain from another may be neutral without any implication for physiology or resistance. If these mutations restore function or repair the deleterious effects of a separate mutation, they could be referred to as compensatory. It's important to note, that development of a compensatory mutation (as defined above) does not necessarily happen after the development of a resistance-associated mutation. Mutations may exist that predispose the development of drug resistance, by compensating for the structural or functional

consequence the target mutation might have, or by acting synergistically to increase resistance. Physiological effects of pre-existing mutations may result in the association of strain lineages with distinct drug resistance-associated mutations.

Compensatory mutation also has a bearing on the reversibility of drug resistance-associated mutation (Trindade *et al.*, 2009, Weinreich *et al.*, 2005). Streptomycin resistant (STR^R) *Salmonella typhimurium* strains develop secondary mutations that ameliorate functional disadvantages that result from the original mutation (Bjorkman *et al.*, 1998, Bjorkman *et al.*, 1999). Moreover, the type of compensatory mutation observed was subject to the infection model used; that is, growth in a mouse model produced a different spectrum of compensatory mutations than *in vitro* growth (Bjorkman *et al.*, 1998, Bjorkman *et al.*, 1999). Recently, Barrick *et al.* (2010) showed that, in *E. coli*, the advantage of a beneficial mutation is much greater for *rpoB* mutants in which fitness is already attenuated. That is, the lower the fitness, the more advantageous a beneficial mutation. Moreover, evolvability is in turn increased in *rpoB* mutants compared to an isogenic reference strain. Although this evolution does not always restore fitness to wild-type levels, some form of compensatory mutation is occurring; notably, reversion to wild-type *rpoB* was not observed after 640 generations in RIF free medium (Barrick *et al.*, 2010). These experimental findings support theoretical models by Levin *et al.* (2000) which postulate that, if compensatory mutations occur during pathogen infection, and are fixed, an “adaptive valley” will result which will make reversion to susceptibility almost impossible – even if antibiotics are completely removed. However, a more recent theoretical framework proposed by Schulz Zur Weisch *et al.* (2010) suggests it may not be so simple. This model predicts that persistence of both antibiotic resistance and compensatory mutations depends on the advantages (and disadvantages) they confer in the presence or absence of drug. Furthermore, that the rate at which bacteria are killed is a key parameter in the probability of reversion – bacteria that are more easily killed (either as a result of fitness defects or of antibiotic action) are more likely to revert to susceptibility (Schulz Zur Wiesch *et al.*, 2010).

A recent investigation into the epistatic effects of sequential antibiotic resistance mutations in *E. coli* uncovered alarming degrees of positive epistasis between genes

associated with resistance to STR, RIF and Nalidaxic acid (Trindade *et al.*, 2009). Interestingly, epistasis was allele-specific, rather than gene specific (Trindade *et al.*, 2009). Predating the abovementioned study, Chakrabarti *et al.* (1977) observed physiological effects conferred by RIF^R-associated mutations that are dependent on corresponding STR^R-associated mutations (Chakrabarti & Gorini, 1975, Chakrabarti & Gorini, 1977). Further evidence for possible epistatic effects of drug resistance-associated mutations comes from a recent study which shows that sublethal concentrations of ribosome inhibiting agents, STR and chloramphenicol, reduce fitness costs associated with RIF^R-associated mutations in *Pseudomonas aeruginosa*, while treatment with drugs that have other targets did not (Hall *et al.*, 2011). The authors suggest that transcription and translation are linearly related and therefore decreased requirements for RNAP diminishes fitness costs that may be incurred by RIF^R-associated mutations. A STR^R-associated mutation in *rps*L that diminishes ribosome functionality may have similar effects (Hall *et al.*, 2011). It would be interesting however to examine whether inhibition of RNAP lessens fitness costs that may be associated with STR^R-associated mutations. Epistasis has significant consequences for drug design, particularly because enzymes in these key pathways are common targets for drug development. If the target of a novel drug is known, and the spectrum of mutations that may cause resistance are modelled (Cao *et al.*, 2005), the level of epistasis between these alleles and existing resistance alleles could be examined. If high levels of positive epistasis are observed, the usefulness of the new drug in treating resistant disease is diminished. Compensation of the fitness costs associated with a RIF^R-associated mutation by, for example, a STR^R-associated mutation would have devastating consequences for MDR-TB control. If this is occurring in natural populations, mutant combinations of alleles that are positively epistatic should be more frequently observed in clinical settings.

Recently, two independent groups sequenced and comparatively analyzed MTB strains implicated in the infamous Tugela Ferry XDR-TB outbreak (Motiwala *et al.*, 2010, Ioerger *et al.*, 2009). The results of one study (Ioerger *et al.*, 2009) show that the XDR transmission in KZN is due to clonal strain expansion. Even with mutations in multiple essential genes, this strain remained highly virulent and transmissible (Ioerger *et al.*, 2009). Compensatory evolution may be one mechanism by which fitness costs of the

aforementioned mutations could be ameliorated. However, the subsequent study compared XDR strains from the Tugela Ferry region to a phylogenetically comprehensive panel of drug-resistant and susceptible MTB strains, and found little evidence for compensatory mutations (Motiwala *et al.*, 2010). This poses critical questions surrounding the ability of this and other drug-resistant MTB strains to successfully infect and cause disease, while harbouring mutations in more than one essential gene. One possibility is that in a host population with compromised immunity, the physiological requirements for MTB are somewhat different and therefore fitness of resistant strains less attenuated (Motiwala *et al.*, 2010). Since TB is the biggest killer of HIV positive individuals (WHO, 2009), these findings are especially important in developing countries where the burden HIV/TB co-infection is highest.

1.5.3 Drug resistance and increased mutagenesis

Bacterial populations are not homogeneous, and it has been proposed that exposure to antibiotics selects for resistant clones that exist in a population before antibiotic contact (Livermore, 2003). However, there is significant debate surrounding the ability of bacteria to induce a state of adaptive mutagenesis under stressful conditions (Wrande *et al.*, 2008, Roth, 2010). One of the mechanisms for bacterial adaptation – including development of drug resistance – is the induction of error prone DNA repair systems (Boshoff *et al.*, 2004, Friedberg *et al.*, 2001). The enzymes involved are best characterized in *E. coli* [Reviewed in (Friedberg *et al.*, 2001)] however, analogous pathways are thought to contribute to the emergence of drug resistance in an assortment of pathogens (Oliver *et al.*, 2000, Bjorkholm *et al.*, 2001, Boshoff *et al.*, 2003, Boshoff *et al.*, 2004). Moreover, it is likely that exposure to antimicrobials induces reactive oxygen species (ROS) production, and oxidative damage is likely to be involved in antibiotic mediated cell death (Dwyer *et al.*, 2009). Furthermore, ROS production leads to activation of the SOS response, and recently Kohanski *et al.* (2010) suggested that, as a consequence of ROS-mediated mutagenesis, treatment with one antibiotic can encourage resistance to other unrelated antibiotics (Kohanski *et al.*, 2010). However, opinions diverge when it comes to the reasons for the observed increase in mutants in stressful microenvironments – can harsh conditions select for a heightened state of mutability? Alternatively, a small pool of mutants that

already exist in the population may be selected for by the conditions to which they confer beneficial capabilities (Roth, 2010, Wrande *et al.*, 2008). Some researchers have proposed that an increase in *rpoB* mutations under slow or non-growing conditions indicates that some organisms have the ability to increase their mutational rate under these conditions (Karunakaran & Davies, 2000, Bjedov *et al.*, 2003, Cairns *et al.*, 1988, Rosenberg & Hastings, 2004). Wrande *et al.* (2008) offer another explanation based on classic selection. Specifically, these authors suggest that pre-existing *rpoB* mutants are selected during stationary phase growth because the *rpoB* mutation confers some physiological benefit, based on the following evidence. A logarithmic increase in *rpoB* mutations was observed in ageing *E. coli* colonies: this suggests expansion of a single clone. For increased mutagenesis to be responsible, mutagenic rates would have to have increased 10^4-fold, which has never been observed even for lab-induced mutagenesis. Moreover, this kind of increase in mutagenesis would likely induce lethal mutation and lead to loss of viability. Clones containing *rpoB* mutations were clustered within sectors of a plate, and moreover, strains within these clusters contained the same *rpoB* mutation. Increased mutagenesis would have resulted in spatially distributed colonies with a range of *rpoB* mutations. Finally, competition assays between strains carrying RIF^R-associated *rpoB* mutations, and the wild-type strain revealed that mutants had increased growth by day seven of the assay (Wrande *et al.*, 2008). Not only do these findings have important implications for adaptive mutagenesis but they also support the idea that *rpoB* mutations have considerable consequences for the physiological outcomes of RIF^R populations.

Lesions in DNA which hinder RNAP progress pose a risk to both RNA and DNA fidelity. Transcription coupled repair (TCR) is an important subclass of nucleotide excision repair, and functions to maintain DNA fidelity at genes which are actively transcribed (Roberts *et al.*, 2004). Transcription-repair coupling factor (TRCF) plays a pivotal role in this process and is encoded by the *mfd* (mutation frequency decline) gene in *E. coli* (Park *et al.*, 2002). When sufficient nucleotides are present, TRCF reactivates stalled RNAP complexes (Borukhov *et al.*, 2005). However, RNAPs that are stuck at a lesion can also be dislodged by TRCF pushing forward in an ATP dependent manner (Park *et al.*, 2002). In this case DNA repair is initiated by TRCF binding to UvrA (Park *et al.*, 2002; Roberts *et al.*, 2004; Westblade *et al.*, 2010). The RNAP interaction domain (RID) of TRCF interacts with the N-

terminus of the β subunit of RNAP. In *E. coli* residues 19-142 within the β1-lobe are sufficient for *mfd* interaction and point mutations in this region prevent Mfd-RNAP interaction (Trautinger *et al.*, 2002). This region is not commonly associated with RIF^R, but lies next to the RRDR in RNAP tertiary structure (Ganesan *et al.*, 2007). An investigation of *E. coli* strains exposed to RIF and selected for both RIF^R and a defect in TCR revealed that only one mutation, R529C, diminished TCR. Neither TRCF-RNAP interaction nor the stability of the RNAP-DNA complex were affected, therefore the loss of TCR activity in these strains was as a result of an undefined mechanism (Ganesan *et al.*, 2007).

Collisions between RNAP and DNA replication complexes also present an obstacle for DNA replication and threaten DNA fidelity. Although not a formal component of TCR, other replisome-associated proteins play a key role in resolution of replisome-protein blocks – of which RNAPs constitute the most common type (Guy *et al.*, 2009). In *E. coli* UvrD, Rep and DinG helicases share some functional redundancy in that they are all able to remove protein blocks during replication (Guy *et al.*, 2009; Boubakri *et al.*, 2010; Baharoglu *et al.*, 2011). Mutants that lack either UvrD or Rep are viable however deletion of both proteins is lethal. Increased levels of (p)ppGpp and mutations in *rpoB* which mimic the stringent response, can suppress the viability defect of *rep uvrD* double mutants (Guy *et al.*, 2009). Some stringent response mutations are found in the RRDR, and may also be associated with decreased levels of RIF susceptibility. Mutations in this region, specifically H447R and D444G, minimise replisome blockage by stalled RNAP and diminish the need for the aforementioned helicases (Baharoglu *et al.*, 2011). Moreover, mutations in the RRDR that mimic the stringent response also decrease susceptibility to UV damage in cells lacking functional RecBCD and RuvABC enzyme complexes, and in strains lacking *mfd* (Trautinger *et al.*, 2002). These mutations lie next to the DNA-RNAP channel, and decrease stability of RNAP-DNA interactions (Trautinger *et al.*, 2002; Baharoglu *et al.*, 2011). These findings are important for maintenance of DNA fidelity in strains which lack important DNA repair-associated enzymes and for understanding the function of these enzymes. However, the effect of these mutations on DNA fidelity in wild-type cells has not been full characterised. Moreover, decreased stability of transcriptional complexes could have significant consequences for gene expression. Amelioration of DNA repair by RIF^R-associated *rpoB* mutations provides an advantage for

cells containing these mutations. If DNA repair is disrupted, the extent of disruption and how other DNA repair mechanisms compensate for this disruption would be important in determining the outcome for physiology. Increased mutation may provide a favourable environment for compensatory mutation or for the development of resistance to other drugs. However, if DNA repair is severely compromised, there may be an accumulation of mutations that is lethal to the cell.

1.6 The implications of RIF^R-associated mutations for fitness: molecular epidemiology and evolutionary analyses

In combination, the high level of evolutionary conservation between RNAPs and the importance of this enzyme for expression, suggests that changes in this gene would result in fitness costs. However, several studies that have investigated the fitness cost of RIF^R-associated mutations in the *rpoB* gene indicate that this may not be the case (Reviewed in (Borrell & Gagneux, 2009) and (Andersson, 2006)). Studies have generally reported that the fitness of strains with distinct *rpoB* mutations varies relative to their isogenic, drug-susceptible parents according to the specific strain background, as well as the nature of the *in vitro* assay (Mariam *et al.*, 2004, Gagneux *et al.*, 2006a, Gagneux *et al.*, 2006b, Billington *et al.*, 1999). Considering the structural implications for amino acid changes in the β-subunit of RNAP, these data provide evidence that different *rpoB* point mutations influence fitness in different ways. The frequency of RIF^R-associated mutations at certain loci of the RRDR differs between different organisms, despite high similarity in the surrounding sequence (Kim *et al.*, 2004, Garibyan *et al.*, 2003, Zeibell *et al.*, 2007). The spectrum of observed mutations in *rpoB* may represent the functional or structural constraints for *rpoB* mutability. However, Zeibell *et al.* (2007) show that at nucleotide position C1576 (which is the base where a change occurs to confer an H526Y mutation (**c**ac → **t**ac) *E. coli* has a mutation rate (μ – mutation rate per replication) of 0.91, while *B. anthracis* only has a mutation rate of 0.47. This indicates that within the limits of *rpoB* mutability, hotspots occur at diverse sites in different organisms. Distinct physiological and therefore fitness requirements for individual species may result in a different set of mutations being selected for (Jenkins *et al.*, 2009).

Sequence variation between MTB strains is conventionally thought of as being exceptionally low. However, recent studies provide compelling evidence that MTB sequences differ more than initially thought (Hershberg *et al.*, 2008, Gagneux & Small, 2007, Nicol & Wilkinson, 2008, Niemann *et al.*, 2009). Distinct MTB strains may have varying virulence properties as a consequence of relative clonality of the MTB complex (Niemann *et al.*, 2009, Gagneux & Small, 2007). This is supported by several indications that distinct MTB strains behave differently in models of disease (Gagneux & Small, 2007). Furthermore, Hershberg *et al.* (2008) suggest the evolution of distinct phylogeographic lineages has occurred in close alliance with the human populations they infect. Intriguingly, subsequent work from the same group indicates that in contrast to the newly described MTB strain diversity, genomic regions involved in recognition of the bacillus by the host immune system are highly conserved (Comas *et al.,* 2010). This conservation of antigenic regions is counter-intuitive to the antigen variation found in other common pathogens. Dynamics between various biological factors, which translate into a complex network of evolutionary pressures, play an enormous role in the adaptation and fitness of MTB. In addition to providing tools to investigate outbreak of TB disease, molecular genotyping and epidemiological studies provide a means to investigate the implications of particular mutations for fitness. This is especially useful when examining the fitness implications of drug resistance-associated mutations. Strains containing mutations with higher fitness costs should, in principle, result in fewer secondary cases. Although this approach is useful, biases – such as selective conditions imposed during strain isolation – may be present, and therefore extrapolation of fitness effects of certain mutations from epidemiological data may not be straight-forward. From a phylogenetic analysis of INH^R-associated mutations, Hazbon *et al.* (2008) suggested that assessment of convergent evolution may be a tool to determine levels of selective advantage conferred by SNP's. Phylogenetic and further evolutionary analysis of drug resistance-associated mutations, and other SNP's, may therefore provide a vehicle to examine the biological significance of drug resistance mutations.

As molecular genotyping techniques became available, an association between W Beijing strains and drug-resistant TB outbreaks were observed (reviewed in (Parwati *et al.*, 2010)). This could indicate that these strains have a greater propensity to adapt to

environmental conditions (van der Spuy *et al.*, 2009) or that there is a general increase in dissemination of both susceptible and resistant variants of this strain. A rare variant of the W Beijing lineage is differentiated by a small change in its IS*6110* fingerprint (Strauss *et al.*, 2008). These atypical strains are thought to be less fit, since they are infrequently observed in clinical settings. Interestingly, a RIF^R clinical isolate of this variation sustained transmission in a community with high levels of HIV. Accordingly, a recent article describes a RIF^R MTB isolate with uncommon *rpoB* mutations, and unusually high virulence and transmission in an HIV positive background (Malhotra *et al.*, 2010). This indicates that, against a background of elevated HIV prevalence, RIF^R-associated mutations can be tolerated without compromising community transmissibility even in a strain with already attenuated fitness, (Strauss *et al.*, 2008). Recent whole-genome sequencing and subsequent phylogenetic analysis of drug resistance isolates from the Western Cape indicate that drug resistance is acquired rather than transmitted (Ioerger *et al.,* 2010). However, similar analysis of the XDR strains in the Tugela Ferry region of KZN, shows that strains exhibit a definite pattern of clonality (Ioerger *et al.*, 2009). These contrasting patterns of spread of drug-resistant strains within the same country affirm the importance of considering environmental conditions when evaluating fitness.

1.7 The effect of *rpoB* mutations on global transcriptional regulation: Changes in physiology

It has been postulated that the natural antibiotics function to limit the growth of competitor organisms in resource-poor environments (Goh *et al.*, 2002). Contemporary models propose a subtle modification of this hypothesis, holding that antibiotics operate instead as low-concentration effector molecules in other ecological processes such as signalling and transport mediation (Davies, 2009, Martinez *et al.*, 2009). It follows, therefore, that mutations which arise in response to excessive concentrations of signalling or regulatory molecules might have profound consequences for bacterial physiology. In support of this idea, transcription in *Mycobacteria* and *Salmonella* species is modulated by sub-inhibitory concentrations of RIF and STR and, furthermore, that resistance-associated mutations weaken the level of transcriptional regulation induced by each drug (Goh *et al.*, 2002, Yim *et al.*, 2006, Boshoff *et al.*, 2004).

In combination with small proteins such as DksA in *E. coli* (Potrykus & Cashel, 2008) and CarD in mycobacteria (Stallings *et al.*, 2009), guanosine tetraphosphate (ppGpp) is a central modulator of the bacterial response to environmental stress (Chatterji & Ojha, 2001, Traxler *et al.*, 2006, Traxler *et al.*, 2008). In addition, it is well established that antibiotic production by *Streptomyces* sp. is fundamentally dependent on ppGpp-activated pathways (Xu *et al.*, 2002, Hu *et al.*, 2002, Ochi, 1987). In *S. coelicolor* antibiotic producer strains, , for example, RIF^R-associated mutations corresponding to those found in clinical MTB isolates are associated with increased antibiotic production (Xu *et al.*, 2002, Hu *et al.*, 2002, Hosaka *et al.*, 2009), as well as the inducing the production of novel antibiotics in defined non-producer strains (Hosaka *et al.*, 2009). Moreover, mutations in *rpoB* increase secondary metabolism in *relA/relC* mutants that are defective in antibiotic production owing to an inability to generate ppGpp (Hu *et al.*, 2002). In addition to their enormous utility in the pharmaceutical industry, these findings imply that *rpoB* mutations have significant consequences for secondary metabolism in *Streptomyces* species. Interestingly, some *Nonomuraea* and *Nocardia* species encode two paralogues of the *rpoB* gene. In these strains, the alternative *rpoB* gene confers RIF^R and its activation induces the expression of dormant genes (Tala *et al.*, 2009, Vigliotta *et al.*, 2005, Ishikawa *et al.*, 2006). Notably, the sequence variations that differentiate the alternative *rpoB* gene from the "wild-type" correspond to common *rpoB* mutations in clinically isolated RIF^R MTB strains. Moreover, in *S. lividans* expression of the *Nonomuraea rpoB(R)* gene induces antibiotic biosynthesis (Tala *et al.*, 2009, Vigliotta *et al.*, 2005). In *B. subtilis*, dormant antibiotic production is activated by RIF^R-associated *rpoB* mutations (Inaoka *et al.*, 2004), while RIF^R-associated mutations in the *E. coli rpoB* gene resulted in RNAP behaving as if it was under ppGpp regulation (Zhou & Jin, 1998). Together, these data point to the ability of RIF^R-associated *rpoB* mutations to mimic functionally the actions of ppGpp (Hu *et al.*, 2002, Xu *et al.*, 2002), possibly by modifying promoter specificity. Considering these observations, as well as the complexity of transcriptional regulation, it is likely that RIF^R-associated mutations alter mechanisms of transcription (Tupin *et al.*, 2010).

More than two decades ago, researchers began investigating the effect of *rpoB* mutations on transcriptional pausing and termination (Jin & Gross, 1991, Landick *et al.*, 1990, Singer *et al.*, 1993). In *E. coli*, certain mutations were found to increase termination while others

resulted in decreased termination; moreover compensatory mutations within *rpoB* itself were found to restore the wild-type phenotype (Singer *et al.*, 1993). Jin *et al.* (1988) observed that RIF^{R-}associated mutations in *E. coli* resulted in a range of phenotypic effects including temperature sensitivity and phage susceptibility which they suggested might result from the effect of RIF^{R}-associated mutations on transcriptional pausing and termination. In a 2002 report, Nicholson *et al.* (2002) showed that the spectrum of RIF^{R}-associated mutations in *B. subtilis* differs significantly, depending on whether cells are in a vegetative state or not (Nicholson & Maughan, 2002). Notably, this study demonstrated that *B. subtilis* spores exhibit mutations in the RDDR that are similar to those commonly found in clinically isolated RIF^{R} MTB strains, prompting the suggestion that the DNA in dormant MTB cells may be in a state analogous to that of DNA in *B. subtilis* spores (Nicholson & Maughan, 2002). Further investigation of the effect of *rpoB* mutations on *B. subtilis* physiology revealed that single amino acid changes in *rpoB* resulted in major changes in physiological processes such as sporulation, germination and competence. For example, a mutant carrying a H526Y was not able to form spores efficiently at 48 ˚C, while a mutant carrying the S531L had enhanced spore forming capacity (Maughan *et al.*, 2004). Moreover, RIF^{R}-associated mutations altered the metabolic capability of *B. subtilis* – allowing the organism to utilize substrates that had not previously been identified as suitable for *B. subtilis* (Perkins & Nicholson, 2008). Analogously, an *S. aureus* strain with low level RIF^{R} was better able to establish biofilms on catheters in mice than parental strains (Yu *et al.*, 2005). This RIF^{R}-associated mutation, S509P, was most commonly isolated during murine infection, whereas during *in vitro* biofilm growth, the mutations occurred at position 516 more frequently (Yu *et al.*, 2005). This phenomenon has recently been investigated in MTB: particularly, adjusting the pH of a chemostat culture of MTB resulted in a widely varied spectrum of *rpoB* mutations. At pH 7.0, 89 % of all strains isolated contained the S531L mutation. Although the majority of strains still contained the S531L mutation when the pH was lowered to pH 6.0, the absolute number had dropped to 50 % (Jenkins *et al.*, 2009). This finding is particularly important given that a key environmental niche for MTB is the acidic macrophage phagosomal compartment (Schnappinger *et al.*, 2003).

Mutations in the *rpoB* gene that are associated with RIF^R are likely to have drug-independent consequences for bacterial physiology. The biological implications of drug resistance-associated mutations have a substantial role to play in determining the fitness of resistant bacteria (Borrell & Gagneux, 2009), which has critical bearing on the type of treatment and control measures for MDR disease. The assay system used to measure these effects is not trivial. For MTB in particular, most assays examining the relative fitness of RIF^R mutants are conducted under conditions using a sugar as the primary carbon source (Billington *et al.*, 1999, Gagneux *et al.*, 2006b). The MTB genome encodes a sophisticated repertoire of enzymes to facilitate lipid metabolism when glucose is limited (Savvi *et al.*, 2008, McKinney *et al.*, 2000) and microarray studies have shown that many of these genes are up-regulated during infection (Schnappinger *et al.*, 2003). These (and other) studies suggest mycobacteria favour lipid metabolism during infection, where nutrient levels may be low (Pandey & Sassetti, 2008, Brzostek *et al.*, 2009, Yang *et al.*, 2009, Yam *et al.*, 2009). Recent studies in *E. coli* indicate that adaptation to growth in nutrient limited environments is facilitated by *rpoB* mutation, some of which are analogous to those found in clinical MTB strains (Conrad *et al.*, 2010). Moreover, fitness of *P. aeruginosa* strains containing RIF^R-associated mutations varies depending on the carbon source applied (Hall *et al.*, 2011). These studies support the idea that carbon source (and other environmental conditions) can have a major impact on the observed fitness of strains containing RIF^R-associated mutations, and that these factors must be considered during fitness assay.

In an effort to understand the fitness consequences of RIF^R in MTB, many lab-based studies have isolated spontaneous mutants derived from exposure to RIF. A potential limitation of this approach is that it almost certainly fails to replicate adequately the complex conditions that MTB is expected to encounter *in vivo*. In addition, it does not take into account the potential for (undetected) compensatory mutations to ameliorate fitness costs. Clinical strains harbouring mutations that confer the lowest fitness cost in lab-derived mutants containing the identical *rpoB* mutation are most commonly isolated from patients with MDR-TB disease (Borrell & Gagneux, 2009, Gillespie *et al.*, 2002, Shcherbakov *et al.*, 2010). However, caution must be exercised when making these kinds of inferences (Mariam *et al.*, 2004, Andersson, 2006). There may be selective pressures

which act during the isolation of clinical strains, from sputum collection through strain propagation, storage, and re-growth *in vitro.* Although the S531L mutation is frequently observed in clinical isolates, mutations at other sites in *rpoB* may not be observed because they negatively impact the “culturability” of the strain at each of these stages. These pressures may be imposed by the kind of media used during culture of clinical strains – which may not always be representative of conditions during host infection. Usually similar media are used during fitness assays, and in this way a “selection bias” could be imposed. This has potentially far reaching implications for the reliability of these data to inform about the fitness cost of clinical resistance.

A thorough understanding of the impact that chemotherapeutic measures (as well as resistance to these measures) have on microbial physiology is a key component in answering the many questions surrounding microbial drug resistance. In addition to preventing spread of MDR and XDR disease, the global effects of resistance mechanisms provide an exciting space for a deeper understanding of microbial physiology.

1.8 Specific aims of for this project

1.8.1 Overall objective

The overall objective of this project is to determine the extent to which common mutations associated with RIF^R in clinical MTB isolates impact mycobacterial physiology, primarily carbon source metabolism. Given that physiology and fitness cannot be uncoupled an understanding of the impact of these mutations for physiology will be especially useful for evaluating the fitness, and ultimately transmissibility, of MTB. In this dissertation, we utilize MSM as surrogate for MTB: specifically, we describe the construction of a panel of MSM mutant strains containing RIF^R-associated point mutations that are equivalent to the corresponding mutations in clinically RIF^R MTB strains. Importantly, our methodology ensures that *rpoB* mutants are generated without exposure to RIF, such that the ability of single point mutations to confer high-level RIF resistance could be evaluated.

1.8.2 Hypothesis

The central hypothesis informing this project is that clinically-relevant *rpoB* mutations impact the capacity of MTB to metabolize different carbon sources and, further, that such mutations arise, and are maintained, in the population in the absence of RIF exposure.

1.8.3 Specific aims

a) Construction of a panel of MSM mutants containing point mutations in *rpoB* that are equivalent to the corresponding mutations in clinically RIF^R MTB strains

b) Evaluation of the fitness of the MSM mutants in comparison with the otherwise isogenic parental strain during growth on various carbon sources.

2. Methods

2.1 Selection of *rpoB* mutations to include in the study

An analysis of the literature and preliminary data generated at the MMRU informed the selection of RIF^R-associated mutations to be included in the project. Recent literature was broadly analyzed to determine the most commonly observed *rpoB* mutations in clinically isolated and laboratory-derived RIF^R MTB strains. The terms "*rpoB*" AND "mutations" AND "*Mycobacterium tuberculosis*" were included in a search on the NCBI Pubmed database (http://www.ncbi.nlm.nih.gov/pmc/). The analysis was by no means exhaustive, nor was it intended to be a complete meta-analysis of the literature. Rather, the purpose of the literature search was to gain a sense of common *rpoB* mutations in RIF^R isolates from a variety of geographical regions. Therefore, restrictions relating to study size or the technique applied in strain isolation, were not imposed. A list of papers included in the analysis can be found in Appendix 5.1.

The observation that a silent mutation (L521L) was associated with relatively high levels of RIF^R in a clinical isolate (Siddiqi *et al.*, 2002) suggested this allele for inclusion in the study. The remaining three mutations were included owing to their potential significance as established by previous work at the MMRU (Table 3.1). The TH526TY mutation is commonly associated with error prone DNA polymerase activity, in particular DnaE2 (Boshoff *et al.*, 2003). In preliminary experiments at the MMRU, a MSM *rpoB* mutant containing a Q513V mutation displayed differential growth on OADC (oleic acid-albumin-dextose-catalase) vs. Glucose-NaCl (Glucose Salt; GS) containing 7H10 media (D Warner and Susan Puckett, unpublished results). The 30 bp deletion has been isolated at the MMRU in response to DNA damaging agents and nutrient stress, and has also been studied in an independent study (Malshetty *et al.*, 2010).

2.2 Bioinformatic analysis of *rpoB*

Nucleic acid and protein sequence of the MSM and MTB *rpoB* gene were downloaded from The Comprehensive Microbial Resource database of the J. Craig Venter (JCVI CMR; http://cmr.jcvi.org/tigr-scripts/CMR/CmrHomePage.cgi) (Peterson *et al.*, 2001). Analysis

of the similarity of the nucleotide and amino acid sequences of the *rpoB* genes from these two organisms was accomplished using the Blast2seq tool available on the NCBI website (http://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE_TYPE=BlastSearch&PROG_DEF=blastn&BLAST_PROG_DEF=megaBlast&BLAST_SPEC=blast2seq). Fasta format files containing the respective *rpoB* gene sequences were uploaded onto the online program and the option to optimise the program for highly similar sequences was selected. Algorithm parameters were maintained as default settings. Default settings for scoring parameters were as follows: mismatch/match scores: 1,-2; gap costs: linear. The algorithm used for the blast2seq are similar to those used for standard BLASTn queries (Altschul *et al.*, 1990). ClustalX 2.0.12 software allowed for additional analysis and visualisation of the similarity of RpoB protein sequence MSM and MTB (Thompson *et al.*, 1997). Finally, genomic context information for the *rpoB* gene in mycobacterial species was obtained from the JCVI CMR (Peterson *et al.*, 2001).

2.3 General Microbiological and DNA manipulation techniques

All DNA manipulations were performed as described previously (Sambrook *et al.*, 1989; Sambrook & Russell, 2001). All mycobacterial culture and molecular cloning techniques were performed as described in Jacobs *et al.* (1991), Larsen (2000), and Parish & Stoker (2001).

2.3.1 General bacterial growth conditions

2.3.1.1 Growth media, reagents and solutions

Growth media, reagents, solutions and buffers used in this study are described in detail in Appendices 5.2.

2.3.1.2 MSM growth conditions

Unless otherwise described, MSM strains were cultured on Middlebrook 7H9 (liquid medium) or Middlebrook 7H10 medium (solid medium) (Difco™, USA) supplemented with oleic acid-albumin-dextrose-catalase (Beckton Dickinson, USA), 0.05 % tween80

(Sigma Aldrich, USA), 0.02 % glycerol (Merck, Germany) for liquid media (0.05 % glycerol for solid media), supplemented with antibiotics as required. Liquid cultures were incubated in Erlenmeyer flasks at 37 °C, with shaking at 120 rpm, in a New Brunswick Scientific Innova 400 incubator shaker (New Brunswick Scientific Co; Inc, USA). Strains cultured on solid media were incubated at 37 ˚C in an Incotherm incubator (Labotec, South Africa). MSM strains to be cultured from freezer stocks were cultured in liquid media for 48 hours to allow for sufficient recovery. Subsequently, an aliquot was sub-cultured into fresh media and strains were grown to the appropriate optical density (OD) for further analysis.

2.3.1.3 *E. coli* growth conditions

The *E. coli* DH5α (Promega, USA) host strain was used in this study. *E. coli* strains were grown in Luria Bertani (LB) broth or on Luria Bertani agar plates (LA), supplemented with appropriate antibiotics (Appendix 5.2). Liquid cultures were incubated in Erlenmeyer flasks with shaking at 120 rpm, at 37 °C or 30 °C in the New Brunswick Scientific Innova 400 incubator shaker. Strains cultured on solid media were incubated at 37 ˚C or at 30 ˚C in an Incotherm incubator.

2.3.1.4 Medium supplements

Appropriate compounds were included in growth media to select for strains containing the desired plasmids. Details of plasmid and strain selection conditions can be found in Tables 2.1 and 2.2. Antibiotics were used at the following concentrations: (1) *E. coli*: ampicillin (AMP), 200 µg/ml (AMP_{200}); hygromycin (HYG), 200 µg/ml (HYG_{200}); kanamycin (KAN) 200 µg/ml (KAN_{200}) (2) MSM: HYG, 50 µg/ml (HYG_{50}); KAN 20 µg/ml (KAN_{20}). RIF concentrations ranging from 50 µg/ml to 4000 µg/ml were included in media to evaluate RIF^R of MSM strains. Selection of double cross over (DCO) strains required sucrose supplementation which was at 5 % (w/v) for *E. coli*, and 2 % (w/v) for MSM. Where plasmids or strains contained the *lacZ* gene, 5-bromo-4-chloro-3-indolyl- β-galactoside (X-gal) was included at 40 µg/ml for *E. coli* and MSM. Details of supplements can be found in Appendix 5.2.

2.3.1.5. Bacterial strains and plasmids

Strains and plasmids used in this study are listed in Table 2.1. Strains were stored at -70 °C in 66 % glycerol (v/v). Strains for storage were prepared as follows. Strains were grown to logarithmic phase (OD of 600 nm (OD_{600}) = 0.6 – 1.0), and 500 µl of the strain added to 1.5ml microfuge tubes (Whitehead Scientific, South Africa) containing 500 µl 66 % (v/v) glycerol. Glycerol was prepared by diluting 66 ml of glycerol in 44 ml of sterile distilled water (dsH_20); and autoclaved for 15 min using the Speedy Autoclave (Huxley, Taiwan).

Table 2.1 Bacterial strains used in this study

Strain	**Details**	**Reference**
E. coli DH5α	Host *E. coli* strain used in this study.	Promega (USA)
Mycobacterium smegmatis mc^2155	High frequency transformation strain.	(Snapper *et al.*, 1990)
rpoB^{S531L}	MSM mutant containing S531L mutation in the *rpoB* gene.	This study
rpoB$^{S531L \rightarrow WT}$	MSM mutant wherein the wild-type allele was restored in the S531L mutant strain.	This study
rpoB^{H526Y}	MSM mutant containing H526Y mutation in the *rpoB* gene.	This study
rpoB^{L521L}	MSM mutant containing L521L mutation in the *rpoB* gene.	This study
*rpoB*sco$^{S531L/WT}$	MSM up-stream SCO mutant containing *rpoB* with S531L mutation and wild-type *rpoB*; contains KAN^R; HYG^R; *Lac*Z and *Sac*B markers.	This study
*rpoB*sco$^{WT/H526Y}$	MSM SCO (down-stream) containing *rpoB* with H526Y mutation and wild-type *rpoB*; contains KAN^R; HYG^R; *Lac*Z and *Sac*B markers.	This study
*rpoB*sco$^{L521L/WT}$	MSM SCO (up-stream) mutant containing *rpoB* with L521L mutation and wild-type *rpoB*; contains KAN^R; HYG^R; *lac*Z and *sac*B markers.	This study

*rpoB*sco^{Q513V}	MSM SCO mutant (orientation not determined) containing *rpoB* with Q513V mutation and wild-type *rpoB*; contains KAN^R; HYG^R; *lac*Z and *sac*B markers.	This study
*rpoB*sco$^{WT/WT}$	MSM SCO mutant containing two copies of wild-type *rpoB*; contains KAN^R; HYG^R; *lac*Z and *sac*B markers.	This study
*rpoB*sco^{TH526TY}	MSM SCO mutant (orientation not determined) containing *rpoB* with TH526TY mutation and wild-type *rpoB*; contains KAN^R; HYG^R; *lac*Z and *sac*B markers.	This study
*rpoB*sco^{30del}	MSM SCO mutant (orientation not determined) containing *rpoB* with a 30 bp deletion and wild-type *rpoB*; contains KAN^R; HYG^R; *lac*Z and *sac*B markers.	This study

Table 2.2 Plasmids used in this study

Strain	**Details**	**Reference**
p2NIL	Core vector used in this study. *E. coli*-*Mycobacterium* shuttle vector, KAN^R	(Parish & Stoker, 2000)
p2NILAB	Modified version of p2NIL where *Acc*65I and *Bam*HI sites had been removed; KAN^R	This study
p2*rpoB*	Template for SDM: p2NILAB with *rpoB* gene inserted; KAN^R	This study
	p2*rpoB*^{MT} constructs	
p2*rpoB*^{S531L}	Suicide vector: p2NILAB with *rpoB* containing an S531L mutation inserted; KAN^R	This study
p2*rpoB*^{H526Y}	Suicide vector: p2NILAB with *rpoB* containing an H526Y mutation inserted; KAN^R	This study
p2*rpoB*^{L521L}	Suicide vector: p2NILAB with *rpoB* containing an L521L mutation inserted; KAN^R	This study

p2*rpoB*^{Q513V}	Suicide vector: p2NILAB with *rpoB* containing an Q513V mutation inserted; KAN^{R}	This study
p2*rpoB*^{TH526TY}	Suicide vector: p2NILAB with *rpoB* containing an TH526TY mutation inserted; KAN^{R}	This study
p2*rpoB*^{30del}	Suicide vector: p2NILAB with *rpoB* containing a 30bp del inserted; KAN^{R}	This study
pOLYG	*E. coli-Mycobacterium* shuttle vector; Replicates episomally; HYG^{R}	(O'Gaora *et al.*, 1997)
rpoB containing pOLYG constructs (pG*rpoB*)		
pG*rpoB*^{WT}	This derivative of pOLYG replicates episomally in MSM and contains the wild-type *rpoB* gene; HYG^{R}	This study
pG*rpoB*^{S531L}	This derivative of pOLYG replicates episomally in MSM and contains the *rpoB* gene with a S531L mutation; HYG^{R}	This study
pG*rpoB*^{H526Y}	This derivative of pOLYG replicates episomally in MSM and contains the *rpoB* gene with a H526Y mutation; HYG^{R}	This study
pGOAL 19	Marker gene cassette for allelic exchange was isolated from this vector; AMP^{R} for replication in *E. coli*	(Parish & Stoker, 2000)

2.3.2 DNA extractions

Large scale plasmid extraction from *E. coli* (Maxi-prep)

2.3.2.1 Extraction of DNA

An *E. coli* culture (100 ml) was grown to stationary phase (8 – 12 hours). Cells were harvested by centrifugation (3500 rpm; 10 min) in a Beckmann J2-21 centrifuge. Pellets were re-suspended in 1 ml Solution I by aspiration and vortexing. EDTA in Solution I inhibits any DNase activity that may be present in the unpurified sample, and glucose maintains osmotic pressure. An aliquot (2 ml) Solution II was added, and tubes were inverted twice to allow for cell lysis to occur before the addition of 1.5 ml Solution III. Solution II contains NaOH which increases the pH of the sample, and facilitates lysis. By adding Solution III, which contains sodium acetate, the pH of the sample is restored to approximately 7.0. Details of the composition of Solutions I, II and III can be found in Appendix 5.3 Mixing by vortexing was avoided from this stage onwards to prevent any shearing of DNA. Aliquots of 800 µl each of the lysed cell mixture were decanted into 1.5 ml microfuge tubes, and centrifuged in the Eppendorf 5415 benchtop centrifuge (Eppendorf, Germany) at 13000 rpm for 10 min. The supernatant was decanted into fresh microfuge tubes, and the pellet containing protein and genomic DNA was discarded. An aliquot of 100 µg RNase A (Roche Applied Science, Germany) was added to each microfuge tube and samples were incubated at 42 °C for 1 hour to allow for RNA degradation. After incubation, 700 µl of 100 % isopropanol (Merck, Germany) was added, and microfuge tubes were left to stand at room temperature for 10 min to allow for DNA precipitation. Thereafter, tubes were centrifuged at 13000 rpm for 10 min, to sediment the DNA. The DNA pellet was washed with 70 % ethanol (Merck, Germany) and excess ethanol was removed by drying in a vacuum centrifuge (SpeedVac, USA) at 45 ˚C for 20 min. DNA pellets were re-suspended in 100 µl dsH_20. From an initial 100 ml culture of *E. coli*, ten samples of 100 µl each should be obtained. The compositions of the relevant solutions are provided in Appendix 5.3.

2.3.2.2 Plasmid DNA purification

Sodium acetate, which provides sufficient salt for subsequent precipitation of DNA, was added to each sample to a final concentration of 0.05 M. An aliquot (700 µl) phenol-

chloroform solution (Appendix 5.3) was added to solution, and inverted repeatedly. DNA is water soluble and remains in the aqueous phase, while proteins accumulate in the organic phase. Therefore, after centrifugation for 10 min at 13000 rpm, the supernatant was removed cautiously, and added to a fresh microfuge tube. Upon addition of 350 µl of chloroform-isoamyl alcohol solution (Appendix 5.2), the mixture was centrifuged for another 10 min. Chloroform-isoamyl alcohol provides an initial clean up step, and minimises phenol carry over. The supernatant was removed, and placed in fresh microfuge tubes.

2.3.2.3 DNA precipitation

An aliquot (1 ml) of cold (-20 °C) 100 % ethanol was added to each sample. Samples were incubated at -20 °C for 1 hour to allow for DNA precipitation. After precipitation a stringy mass of DNA was often visible in the sample. Subsequently, samples were centrifuged for 20 min to allow for sedimentation of DNA. The supernatant was discarded and the DNA pellet was washed with 70 % ethanol, and dried for 20 min at 45 °C for 20 min. DNA was re-suspended in 30 µl dsH_20.

2.3.2.4 Quantification and analysis of DNA purity

The NanoDrop ND-1000 Spectrophotometer (Thermo Scientific, USA) was used to assess the concentration and quality of plasmid DNA. Pure DNA is required for all molecular biological techniques as impurities in the DNA can interfere with down-stream processes such as enzymatic treatment and PCR. Although different nucleotides comprising the DNA sample have slightly different absorbance, optimal absorbance of light for nucleic acids is at 260 nm. Owing to the presence of the aromatic amino acids, tryptophan and tyrosine, protein predominately absorbs light at 280 nm. Therefore, spectrophotometric measurement of the extracted DNA sample is taken at 260 nm and 280 nm. The 260:280 ratio gives some indication of the relative levels of DNA and protein in the sample. If a ratio of 1.8 is obtained DNA is considered pure, if the ratio is 2.0 DNA is considered very pure. Given that EDTA, carbohydrates and phenol have absorbance close to 230 nm the absorbance of the sample at this wavelength is also measured. There should be little of these compounds in the sample, and therefore a 260:230 ratio of 2.0 – 2.2 is considered

acceptable. Nanodrop measurement of DNA concentration and purity are generally reliable, but can vary depending on the composition of DNA (linear versus plasmid, chromosomal versus PCR fragment, *etc*.), and the calibration status of the machine.

2.3.2.5 Small-scale plasmid extraction from *E. coli* (Mini-prep)

Culture and reagent volumes were reduced appropriately for small-scale plasmid extraction. This method was especially useful when large numbers of colonies had to be screened. An *E. coli* culture of 2 ml was grown to stationary phase in 2 ml microfuge tubes. The cultures were centrifuged for 10 min at 13000 rpm, and the supernatant was discarded. Pellets were re-suspended in 100 µl Solution I. An aliquot (200 µl) of Solution II was added, samples were inverted, and 150 µl Solution III was added. Samples were centrifuged for 5 min at 13000 rpm, and the supernatant decanted to fresh microfuge tubes. Thereafter 10 µl of a 10 mg/ml solution of RNase was added, and samples incubated for 15 min at 42 °C. After incubation, 350 µl of 100 % isopropanol was added to each sample, which were left to stand at room temperature for 5 min. Samples were centrifuged at 13000 rpm for 10 min, and the supernatant discarded. The pellet containing the DNA was washed with 70 % ethanol, and dried in the vacuum centrifuge. The dried DNA was re-suspended in 20 µl of $dsH_2 0$.

Chromosomal DNA extraction from MSM

2.3.2.6 Cetyltrimethylammoniumbromide (CTAB) isolation of chromosomal DNA

To obtain pure, concentrated chromosomal DNA from mycobacterial cells, the CTAB method of DNA extraction was employed. MSM cultures were grown to logarithmic phase, and decanted into 2 ml microfuge tubes. Samples were centrifuged for 5 min at 13000 rpm, and the supernatant discarded. Cell pellets were re-suspended in 500 µl tris-EDTA (TE) (Appendix 5.3) buffer. Cells were heat killed by incubation at 65 °C for 20 min, 500 µg of lysozyme (Roche Applied Sciences, Germany) was added to samples, and samples were incubated overnight at 37 °C with shaking. Following incubation, 70 µl 10 % sodium dodecyl sulphate (SDS) (Fluka, Switzerland) and 10 µl of proteinase K (10 mg/ml) (Roche Applied Sciences, Germany) were added to samples. SDS and lysozyme facilitate breakage of the mycobacterial cell wall and proteinase K degrades protein contaminants.

An aliquot (100 µl) of 5 M NaCl (Merck, Germany) was added to each sample, and inverted repeatedly to mix. Thereafter, 80 µl pre-warmed CTAB was added to samples, which were then incubated 65 ^{o}C for 10 min. CTAB binds proteins, polysaccharides and other impurities in the sample, thereby facilitating purification. An equal volume of chloroform-isoamyl alcohol was added, and samples were centrifuged for 5 min at 13000 rpm. The supernatant was removed to fresh microfuge tubes, and precipitation continued as described in Section 2.3.2.3.

2.3.2.7 Colony boil

The colony boil method was applied to facilitate genotypic screening of large numbers of MSM transformants following electroporation. Colonies were picked and re-suspended in 100 µl TE buffer. Samples were incubated for 10 min at 65 °C to kills cells and allow disruption of the cell wall. An aliquot (100 µl) of chloroform-isoamyl alcohol was added for limited purification of samples, which were then left to stand at room temperature for 10 min, with periodic mixing by inverting. Samples were centrifuged for 10 min at 13000 rpm. DNA contained in the supernatant was used for subsequent PCR screening or processing.

2.3.2.8 DNA storage

DNA is generally stable at room temperature; however, to prevent degradation or damage, extracted chromosomal and plasmid DNA were stored in dsH_2O at -20 °C.

2.3.3 Agarose electrophoresis

Agarose is a seaweed-based polymer that forms a gel that inhibits passage of DNA through a polymer matrix when an electric current is applied. DNA is negatively charged, and will therefore migrate from the negative to the positive electrode in an electric field. Depending on the concentration of agarose used in preparation, different sized “pores” are obtained which allow for separation of DNA of different sizes. The more concentrated the agarose, the smaller the cross-linked pores, and the harder it is for large DNA fragments to migrate across the gel. In this way, larger fragments of DNA become trapped closer to the negative electrode, while smaller fragments migrate further down

the gel as they move towards the positive electrode. Addition of ethidium bromide, which intercalates between nucleotides and fluoresces under UV light, allows for visualisation of DNA. Agarose gel electrophoresis was invaluable for separation, visualisation and purification of DNA fragments.

2.3.3.1 Preparation of agarose gels

For general DNA separation, 1 % agarose gels were employed. However, the size of allele specific PCR (AS-PCR) products required the higher resolution of small (100 – 200 bp) PCR fragments; therefore, 2 % gels were utilised. Appropriate amounts of agarose were added to 1 X tris-acetate-EDTA (TAE) buffer (Appendix 5.4). The mixture was weighed and dsH_2O equivalent to 10 g was added. The mixture was heated, to allow for melting of agarose powder, until the weight of the solution was the same as before dsH_2O was added. Addition of dsH_2O, and subsequent weight measurement maintained TAE buffer concentration, and ensured that only excess water evaporated during heating. The mixture was allowed to cool, and 0.5 µl of ethidium bromide per 30 ml geL was added. Gels were allowed to set in casting trays at 4 °C for 1 hour.

2.3.3.2 Electrophoresis

Cold (4 °C) 1 X TAE buffer was added to tanks prior to running. The buffer stabilises the pH and provides ions for conductivity. A 2 µl sample of loading dye was added for every 4 µl DNA sample to be run. Loading dye was added to all DNA samples including DNA ladders, and contains bromophenol blue which is negatively charged, and therefore runs in the same direction as DNA. This facilitates evaluation of the migration distance of DNA. Additionally, the loading dye mixture contains glycerol which facilitates loading of the DNA into liquid-filled wells. For most gels, Molecular Weight Marker III was used (Roche Applied Science, Germany); however, to quantify accurately the smaller fragments generated by AS-PCR, Molecular Weight Marker V (Roche Applied Science, Germany) was used. Tanks were attached to a power pack (BIO-RAD, UK) and gels were run at 100 mV for the appropriate time.

2.3.3.3 Purification of DNA from agarose gels

Following separation of DNA by agarose gel electrophoresis, the required bands were excised from gels, and added to 1.5 ml microfuge tubes. To extract DNA, the QIAquick Gel Extraction Kit^{TM} (QIAGEN, USA) was used according to manufacturer's instructions. Appropriate amounts of the QG Buffer^{TM} were added to the gel slices, and samples were incubated at 65 °C until agarose had completely melted. The QG Buffer^{TM} contains chaotropic agents which facilitate dissociation of DNA strands. Single-stranded DNA (ssDNA) binds to silica based columns in the presence of high salt concentrations, and in this way DNA is purified. The solution was then added to the QIAquik Spin Column^{TM}, and centrifuged for 1 min at 13000 rpm. The flow-through was discarded and 800 µl of PE Buffer^{TM} added to facilitate flow-through of impurities. After addition of the PE buffer the sample was centrifuged for 1 min at 13000 rpm, and the flow through discarded. The column was centrifuged again to remove any excess PE Buffer^{TM}. The column was then placed in a fresh 1.5 ml microfuge tube, and 30 µl pre-warmed dsH_20 was then added to column and DNA was eluted by centrifugation for 1 min at 13000 rpm. Before DNA elution the sample was left to stand for 2 min. It was observed that warming of dsH_20 significantly improved final DNA yield from the column.

2.3.4 DNA manipulation and cloning

Construction of vectors was accomplished using the *E. coli* DH5α host strain, and final vector products were used to transform MSM strains. Broadly, vector construction and cloning involved three core steps. The appropriate plasmid vector was used to facilitate insertion of the gene of interest, for this study the *rpoB* gene or gene fragments were inserted. After extraction, plasmid DNA, or PCR product were digested with the appropriate restriction endonuclease. If the same enzyme is used to cleave two distinct DNA fragments, the ends of the fragments will have the same nucleotide overhangs, and will therefore be complimentary. This facilitates the ligation step, in which the two fragments are covalently bound. Once the required construct was generated, *E. coli* cells were transformed with the plasmid and extracted DNA was used to genetically manipulate MSM.

2.3.4.1 Restriction endonuclease digests

For construct screening, and generation of appropriate fragments for cloning, restriction endonuclease digestion was utilized. Restriction enzymes have high levels of specificity, and cleave defined sequences of DNA at the sugar phosphate backbone. These enzymes are found in bacterial and archaeal species, and form part of an innate defence system protecting against bacteriophage attack. Restriction endonucleases commonly recognise 4 – 8 bp palindromic nucleotide sequences, and often cleave close to (or within) the recognition site. The high specificity of restriction enzyme recognition sites make them very useful for molecular cloning techniques. One unit of the appropriate enzyme was added to 10 µg of required DNA, and incubated overnight without shaking. Unless otherwise described, optimal incubation temperature for restriction digests was 37 °C. Different restriction enzymes require different buffer conditions, and some enzymes require bovine serum albumin (BSA) to stabilise the enzyme during reaction. In general, 1 X buffer/BSA is included in the reaction mixture. A list of restriction enzymes that were frequently used in this study, and the corresponding buffer conditions, is provided in Appendix 5.3. Fragments containing the required DNA were isolated from agarose gels as described in Section 2.3.3.

2.3.4.2 Removal of 5’ ends from vector DNA

To minimise vector re-ligation during cloning, 5’ phosphate groups are removed after restriction digestion. The ligase enzyme requires a 5’ phosphate group to join two DNA molecules. Removal of these groups from the vector means that only the insert DNA fragment can provide the necessary chemical scaffold, and therefore minimises vector re-ligation. After digestion, 1 unit of Antarctic Phosphatase (New England Biolabs, USA) was added to digested DNA and the sample was incubated for 1 hour at 37 °C. The sample was removed and purified using the QIAquick PCR Purification Kit^{TM} (QIAGEN, USA) purification kit according to manufacturer’s instructions. The principle of DNA purification is similar to that of the QIAquick Gel Extraction Kit^{TM} (described in Section 2.3.3.3) and the protocol is essentially the same, except that melting of agarose is not necessary.

2.3.4.3 Ligations

T4 DNA ligase (Roche Applied Sciences, Germany) was used to join the vector and insert DNA of interest to generate the desired plasmid. Depending on DNA fragment sizes, different amounts of vector and insert are required. Furthermore, to maximise the likelihood of legitimate ligation events, ligation reactions are set up using various vector:insert ratios. After estimation of DNA concentration using an agarose gel, the amount of insert and vector DNA required is calculated using the following formula:

[(ng vector x kb size of insert) / kb size of vector] x desired ratio = ng insert

One unit of T4 DNA ligase was added to the mixture of vector and insert DNA together with 1 X ligation buffer. The reaction buffer contains ATP, which is an essential cofactor for the ligation reaction. The sample was incubated at 4 °C overnight, for sticky ended ligations, and at room temperature for blunt ended ligations.

2.3.5 Transformation of cells and electroporation

2.3.5.1 Transformation of E. coli

Transformation-competent *E. coli* DH5α cells were used to efficiently introduce plasmids into the *E. coli* cell. Upon culture and replication of *E. coli*, larger amounts of the plasmid of interest could be extracted. After the ligation reaction had been allowed to proceed, the T4 DNA ligase was inactivated by heating at 65 °C for 20 min. Transformation-competent *E. coli* cells were thawed on ice, and 100 µl of cells were added to each ligation reaction. The mixture was heat shocked at 42 °C for 90 s to allow for DNA uptake, and placed on ice. An aliquot of 500 µl of LB broth was added, and cells were incubated at 37 °C to allow for replication and recovery. After 1 hour of incubation, cells were spread on LA plates containing appropriate antibiotics, and incubated to allow for colony growth. All vectors contain defined antibiotic-resistance genes; therefore, colonies that grew on selective media were likely to contain the vector (Table 2.2)

2.3.5.2 Electroporation of MSM

To transform MSM with appropriate vector, the electroporation technique as described in (Larson, 2000) was employed. A 100 ml culture of MSM cells was grown to logarithmic

phase. Cells were chilled at 4 °C and centrifuged at 3500 rpm for 10 min to sediment cells. Cells were re-suspended in 30 ml 10 % glycerol that had been chilled to 4 °C and filter sterilised. Cells were maintained on ice, and washed three times by centrifugation, each time re-suspending in the 10 % glycerol solution. After the third centrifugation step, cells were re-suspended in 7 ml 10 % glycerol, and used immediately. DNA was precipitated as described in 2.3.2.3 and a 1 µg sample of DNA was added to 0.2 cm electroporation cuvettes (BIO-RAD, UK). Electroporations were carried out using the BIORAD gene pulser, using the following settings: 2.5 kV; 25 µF and 1000 Ω. After electroporation, cells were removed from the cuvette and added to fresh 2 ml microfuge tubes. An aliquot (800 µl) 7H9 OADC medium was added, and samples were incubated overnight at 37 °C to allow for recovery. Thereafter, cells were plated on 7H10 OADC media containing relevant selective compounds.

2.3.6 Polymerase chain reaction (PCR)

PCR was employed to isolate the *rpoB* gene from the MSM chromosome to construct a template for site-directed-mutagenesis (SDM), to introduce point mutations via SDM, and to screen for point mutations using AS-PCR. During template construction, as well as SDM, it was essential that no additional point mutations were introduced into the *rpoB* gene. The DNA polymerase often used in PCR reactions is isolated from *T. aquaticus* (Taq) and has no 3'→ 5' proofreading ability. The variant used for this project, FastStart Taq DNA Polymerase™ (Roche Applied Sciences, Germany) results in approximately 1.3 errors every 10^5 bases. To minimise the likelihood of erroneous bases being incorporated during PCR, Phusion High Fidelity DNA Polymerase™ (New England Biolabs, USA) was employed. This enzyme has a reduced error rate of approximately 4 errors every 10^7 bases. During PCR screening, large numbers of colonies were screened, and the intrinsic error-rate of FastStart Taq DNA Polymerase™ was acceptable.

2.3.6.1 PCR using Phusion High Fidelity DNA polymerase™

PCR reactions contained the following reaction components: 1 X Phusion High Fidelity PCR buffer™, 1 X Roche GC solution™, 1 µM forward primer, 1 µM reverse primer, 100 ng template DNA and 200 µM of each deoxydribonucleotide phosphate (dNTP): dATP;

dCTP; dGTP and dTTP (New England Biolabs, USA). Reactions were made up to 20 µl or 50 µl with dsH_2O. Cycling conditions were modified according to the specific fragment being amplified, and are described in relevant Sections.

2.3.6.2 PCR using FastStart Taq DNA Polymerase™

PCR reactions contained the following reaction components: 1 X Roche PCR reaction buffer™, 1 X Roche GC solution™, 1 µM forward primer, 1 µM reverse primer, 100 ng template DNA and 200 µM of each dNTP: dATP; dCTP; dGTP and dTTP. Reactions were made up to 20 µl or 50 µl using dsH20. Cycling conditions were modified according to the specific fragment being amplified, and will be described in relevant Sections.

2.3.6.3 Primer design

Clone Manager version 6.00 (Scientific & Educational Software, USA) was used to facilitate planning of the cloning strategy and primer design. Clone manager allowed prediction of product sizes for restriction digestion as well as PCR. As far as possible, oligonucleotide primers were designed with the following specifications: 50 % GC content, 18 – 20 bp, and an annealing temperature of 60 ˚C. Primers were obtained from Integrated DNA Technologies (USA). A list of primers and primer sequences can be found in Appendix 5.5.

2.3.6.4 Sequencing

All sequencing was conducted at the Stellenbosch Sequencing Facility (Stellenbosch, South Africa). Plasmids; purified chromosomal and unpurified/purified plasmid DNA at a concentration of at least 100 ng/µl was delivered to the facility, together with sequencing primers (Appendix 5.5) at a concentration of at least 1.1 µM that had been ordered from Integrated DNA technologies.

2.4 Specific strategies used in this project

2.4.1 Generation of suicide vectors and plasmid constructs used in this study

2.4.1.1 Cloning strategy

A cloning strategy was designed to introduce a selection of RIF^R-associated mutations into the *rpoB* gene of MSM via SDM and allelic exchange. The core vector used in this study was p2NIL (Parish & Stoker, 2000), which contains KAN^R genes for selection in *E. coli* and MSM. This vector also has a multiple cloning site, which includes a single *Pac*I site for convenient addition of a marker cassette, which facilitates selection of SCO and DCO strains during allelic exchange. A template containing the full length *rpoB* gene was constructed so that SDM could be performed. In Figure 2.1, the full length *rpoB* gene is represented in green. Specific *rpoB* mutations were introduced into the appropriate region in the *rpoB* gene via SDM. SDM was performed on a smaller fragment of the *rpoB* gene (shown in red in Figure 2.1), which was inserted back into the template to yield a vector containing a mutated version of the *rpoB* gene. For five out of six mutations, SDM was used. Cross over (Xover) PCR was used to generate the 30 bp deletion. The term for these vectors was p2*rpoB*^{MT}; where reference is made to a specific mutation, for example the S531L mutation, the vector would be called p2*rpoB*^{S531L} (Table 2.2). The *Pac*I marker cassette was added to facilitate selection during allelic exchange.

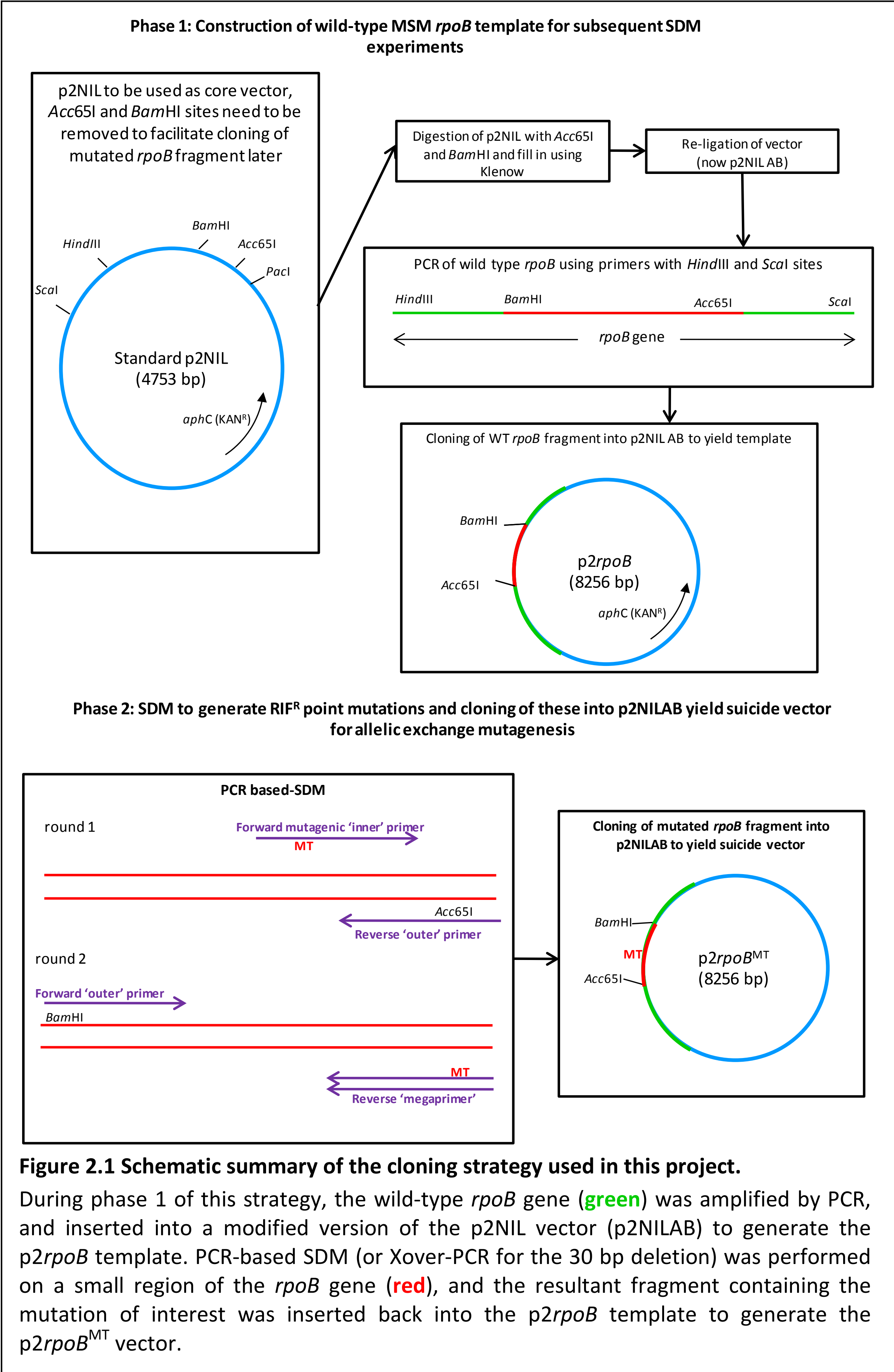


Figure 2.1 Schematic summary of the cloning strategy used in this project.
During phase 1 of this strategy, the wild-type *rpoB* gene (**green**) was amplified by PCR, and inserted into a modified version of the p2NIL vector (p2NILAB) to generate the p2*rpoB* template. PCR-based SDM (or Xover-PCR for the 30 bp deletion) was performed on a small region of the *rpoB* gene (**red**), and the resultant fragment containing the mutation of interest was inserted back into the p2*rpoB* template to generate the p2*rpoB*^{MT} vector.

2.4.1.2 Suicide vector construction

Generation of template for SDM

a) Modification of p2NIL vector backbone: fill-in using DNA polymerase Large Fragment I (Klenow)

The core vector used to generate the p2*rpoB* template was p2NIL (Parish & Stoker, 2000). However, as indicated in Figure 2.1, *Acc*65I (New England Biolabs, USA) and *Bam*HI (New England Biolabs, USA) restriction sites were used to facilitate cloning of the mutated fragment of *rpoB* back into the full length *rpoB* gene. Since these sites occur in the *rpoB* gene, and in the p2NIL vector, they had to be removed from the vector to prevent digestion at two sites during cloning. Restriction digestion was carried out as described in Section 2.3.4.1. After digestion with *Acc*65I, Klenow enzyme (New England Biolabs, USA) was used to fill in the overhangs at each end. The Klenow fragment retains polymerase activity as well as 3'→ 5' exonuclease activity, but has lost 5'→ 3' exonuclease activity. The following components were included in the reaction mixture: 1 µg of DNA; 200 µM of each dNTP; 1 X Klenow buffer. The sample was made up to 15 µl with dsH_20. The reaction was incubated for 13 min at 37 °C. Incubation time was found to be an important parameter and it was found that longer incubation times induced excessive 3' → 5' exonuclease activity which resulted in failed ligation reactions. After Klenow fill in of the *Acc*65I site, the vector was re-ligated using T4 DNA ligase with reaction conditions as described in 2.3.4.3. *E. coli* was transformed with this modified version of p2NIL (p2NILA) and, after screening for removal of the *Acc*65I site, the *Bam*HI site was removed by the same process. The resultant vector had both the *Acc*65I and *Bam*HI sites removed (p2NILAB). The p2NIL vector contains the *aph*C gene which confers KAN^R. Removal of *Acc*65I and *Bam*HI sites did not disrupt this gene, and therefore transformed *E. coli* colonies were selected on LA plates containing KAN_{200}.

b) Isolation of full length *rpoB* gene from MSM chromosome

The full length wild-type MSM *rpoB* gene with 300 bp up-stream flanking sequence was amplified from wild-type MSM by PCR. It was important to include 300 bp up-stream sequence to ensure the inclusion in the final construct of the (unknown) *rpoB* promoter region. Phusion High Fidelity DNA Polymerase™ was employed according to

manufacturer's specifications as described in Section 2.3.6.1. Primers used were Whole*rpoB*F and Whole*rpoB*R (Appendix 5.5). The following cycling conditions were used: Initial denaturation at 98°C for 2 min to facilitate denaturation of double stranded DNA molecule, 30 cycles of [denaturation at 98 °C for 30 s; annealing of 60 °C for 30 s, and extension at 72 °C for 2 min] and a final extension at 72 °C for 7 min to ensure all DNA molecules had been replicated.

The manufacturer suggests that, when using Phusion High Fidelity DNA Polymerase™, extension times of 15 – 30 s for every kb of DNA is employed. Given that the *rpoB* gene is 3510 bp in length and 300 bp up-stream sequence was included, 2 min extension time was used.

c) Cloning of full length *rpoB* into p2NILAB

The Whole*rpoB*F and Whole*rpoB*R primers were designed so that *Hind*III and *Sca*I restriction sites were introduced into the PCR product (Appendix 5.5). The *rpoB* PCR product and the p2NILAB vector were digested with *Hind*III and *Sca*I these enzymes as described in Section 2.3.4.1. Fragments were separated and purified from an agarose gel as described in Section 2.3.3 Digestion of the p2NILAB vector with *Hind*III and *Sca*I yielded two fragments of 398 bp and 4363 bp. The 4363 bp fragment was required for cloning, and was isolated from the agarose gel accordingly. Digestion of the *rpoB* PCR product with *Hind*III and *Sca*I shaved a few base pairs off either end of the fragment. A ligation reaction was set up as described in 2.3.4.3 and, after overnight incubation, competent *E. coli* cells were transformed as described in 2.3.5.1 and selected on LA plates containing KAN_{200}. Colonies were screened for the presence of the correct vector and restriction analysis of the correct vector ensured vector integrity. The following enzymes were use to check the vector integrity *Hind*III, *Sca*I, *Acc*65I *Bam*HI and *Bfa*I (New England Biolabs, USA). In this way a template for SDM was generated (p2*rpoB*).

2.4.1.4 SDM

To introduce point mutations into the *rpoB* gene, a PCR-based SDM method was used (Smith & Klugman, 1997). This technique involved two rounds of PCR (Figure 3.4). In the

first round, a non-mutagenic reverse "outer" primer was used in combination with a forward primer designed to introduce the desired point mutation, and amplify the down-stream sequence. A non-mutagenic forward "outer" primer was used in the second round of PCR, while the PCR product of the first round was then used as the reverse primer. In this way the entire region of interest was amplified, while introducing a single desired point mutation. This method is often referred to as the "megaprimer" method owing to the large size of the primer used in the second round of PCR (Smith & Klugman, 1997). Following mutagenic PCR, fragments were inserted back into the p2*rpoB* vector and AS-PCR was used to confirm the SDM process was successful in introducing the correct point mutation. Constructs were also sequenced to ensure that unwanted point mutations had not been introduced during PCR.

a) Round one of SDM

In the first round of PCR for SDM, the megaprimer containing the point mutation of interest was generated. A forward mutagenic inner primer was used to introduce the point mutation of interest. The SDM outer R primer was used as the reverse primer however the forward mutagenic primer for each mutation would be different (Appendix 5.5). The final size of the megaprimer varied slightly for each mutation. Specific sizes were as follows: S531L) 537 bp; H526Y) 553 bp L521L) 568 bp Q513V) 592 bp and TH526Y) 554 bp. In this PCR reaction it was of utmost importance that unwanted, off-site mutations were not introduced into the PCR product. Therefore, the Phusion High Fidelity DNA Polymerase™ enzyme was employed using reaction component concentrations as described in 2.3.6.1. The following cycling conditions were utilised: Initial denaturation at 98 ˚C for 2 min; 30 cycles of [98 ˚C for 30 s; 60 ˚C for 30 s, and 72 ˚C for 90 s] and a final extension at 72 ˚C for 7 min. Resultant product was visualised, and purified from a 1 % agarose gel to be used as the megaprimer in the next round of PCR.

b) Round two of SDM

The megaprimer contained the point mutation of interest, and its use in the second round of PCR facilitated introduction of the mutation into the full-length product (Figure 3.4). A common outer forward primer (SDM outer F) was used in this reaction (Appendix 5.5). For the first five cycles of PCR, only the megaprimer, and not the forward primer,

was included in the reaction. All other reaction components necessary for High Fidelity DNA Polymerase™ were included in the reaction mixture. As a result of the unusual length of the megaprimer, inclusion of only this primer for the first five rounds of PCR increased binding of the megaprimer to the template. The cycling conditions of the first five rounds of PCR were as follows: Initial denaturation: 98 ˚C for 2 min; 5 cycles of [98 ˚C for 30 s and 72 ˚C for 120 s]. Only denaturation and extension were included so that the megaprimer was allowed to bind to fragments and extend fragments. Thereafter, the forward primer was added and amplification continued for another 25 cycles. The cycling conditions for the second half of the reaction were: Initial denaturation: 98 ˚C for 2 min; 25 cycles of [98 ˚C for 30 s; 60 ˚C for 30 s, and 72˚C for 120 s]. This strategy produced a 1578 bp fragment containing the mutation of interest. Fragments were visualised and purified from a 1 % agarose gel.

2.4.1.5 Xover-PCR

The "megaprimer" method could not be used to generate the 30 bp deletion, therefore Xover-PCR was used (Link *et al.*, 1997). Two rounds of PCR were necessary to generate the final product containing the 30 bp deletion. Moreover, during the first round of PCR two separate reactions are set up. To amplify the region up-stream from the deletion, reaction A was set up and reaction B was set up to amplify the region down-stream from the deletion. The reverse primer used in reaction A is unique in that it contains a "tail" at the 3' end, which is homologous to the 5' end of the region down-stream from the deletion (Figure 3.5 and Appendix 5.5). Products from the first set of PCR were purified on a 1 % agarose gel, and used as templates in the second round of PCR. The product from reaction A contains a region that is homologous to the product of reaction B. This facilitates homologous recombination between the two products. For this reason it was important to allow for five rounds of PCR without addition of the outer primers. This allowed for recombination and for "fill-in" of the ends by the DNA polymerase. After outer primer addition, the full length product was amplified. In this reaction, two general outer primers were used (SDM outer F and SDM outer R; Appendix 5.5).

a) Round one of Xover-PCR

During this the first round of Xover-PCR, two separate reactions were set up. In reaction A, the region up-stream of the deletion was amplified. This was accomplished using a non-mutagenic forward primer (SDM outer F), and a unique reverse primer (Appendix 5.5). The reverse primer (Xover R inner) was unique in that the 3' region contained a portion that was homologous to the 5' region of the *rpoB* region down-stream of the deletion (Figure 3.5). In this way, this homologous region was introduced into the product of reaction A (2237 bp). In reaction B, both the forward (Xover F inner) and the reverse primer (SDM outer R) were non-mutagenic and generated a 1596 bp product. Phusion High Fidelity DNA Polymerase™ enzyme was employed using reaction component concentrations as described in 2.3.6.1. The following cycling conditions (for both reaction A and B) were utilized: Initial denaturation at 98 ˚C for 2 min; 30 cycles of [98 ˚C for 30 s; 63 ˚C for 30 s, and 72 ˚C for 90 s] and a final extension at 72 ˚C for 7 min. Resultant product was visualised, and purified from a 1 % agarose gel to be used as the template in the next round of PCR.

b) Round two of Xover-PCR

The products of reaction A and B from the first round of Xover-PCR, were used as a template in the second round. The reverse primer from reaction B contained a tail that was homologous to the 5' region of the product of reaction A, and in this way this region was introduced into the product of reaction A (Figure 3.5). All reaction components necessary for High Fidelity DNA Polymerase™ were included in the reaction mixture. Due to homologous regions in the fragments generated in round 1, recombination should occur. Primers were excluded for the first ten rounds of this PCR to maximize recombination, and to allow the DNA polymerase to "fill in" ends. Thereafter, non-mutagenic outer primers were added, and the PCR continued. The cycling conditions of the first ten rounds of PCR were as follows: Initial denaturation: 98 ˚C for 2 min; 10 cycles of [98 ˚C for 30 s; 60 ˚C for 30 s and 72 ˚C for 55 s]. Thereafter, the forward (SDM outer F) and reverse primers (SDM outer R) were added and amplification continued for another 20 cycles. The cycling conditions for the second part of the reaction were: Initial denaturation: 98 ˚C for 2 min; 25 cycles of [98 ˚C for 30 s; 60 ˚C for 30 s, and 72˚C for 155

s]. This strategy produced a 1578 bp fragment containing the mutation of interest. Fragments were visualised and purified from a 1 % agarose gel.

2.4.1.5 Cloning of SDM and Xover-PCR products into the template

Once fragments containing the desired *rpoB* mutation had been successfully generated via PCR-based SDM, they were inserted back into the p2*rpoB* vector to generate p2*rpoB*^{MT} vectors (Figure 2.1). *Acc*65I and *Bam*HI were used to digest SDM (or Xover-PCR) fragments, and the template, using conditions as described in 2.3.4.1. Digested SDM (or Xover-PCR) fragments and vector were visualised and isolated from a 1 % agarose gel. Digestion of the p2*rpoB* vector with *Acc*65I and *Bam*HI yielded fragments of 6676 bp and 1555 bp. The 6676 bp fragment was required, and was purified from the agarose gel as described in 2.3.3. The phosphate groups were removed from the vector as described in 2.3.4.2, and ligation proceeded as in 2.3.4.3. Vector insert molar ratios of 1:1; 1:2 and 1:3 were employed. The resultant ligation reactions were used to transform competent *E. coli* cells as described in 2.3.5.1, and colonies were screened for the presence of the correct vector. AS-PCR was used to screen for the presence of the specific mutation.

2.4.1.6 *Pac*I cloning

Upon successful generation of p2*rpoB* vectors containing the desired mutation, a marker gene cassette was added to the vectors to complete suicide vector construction. The marker gene cassette contained genes which facilitate selection during allelic exchange in MSM. The selection strategy is described in detail in Section 2.4.2. The marker cassette was obtained from pGOAL19 (Parish & Stoker, 2000). Both the pGOAL19 vector, and p2*rpoB*^{MT} vectors were digested with *Pac*I under conditions as described in 2.3.4.1. Digestion of pGOAL19 vector with *Pac*I resulted in 7939 bp and 2496 bp fragments. The 7939 bp fragment was isolated from the agarose gel. Digestion of the p2*rpoB*^{MT} vectors with the *Pac*I enzyme linearised the vector, which was purified from an agarose gel to ensure any uncut vector was no present. After removal of 5’ phosphate groups from the vector, ligation of was accomplished as described in 2.3.4.3, and *E. coli* cells were transformed. Plasmids were extracted using the large scale extraction protocol as described in 2.3.2.1, and used for subsequent allelic exchange.

2.4.2 Generation of MSM mutants containing RIF^{R}-associated mutations: allelic exchange

Allelic exchange mutagenesis allows introduction of the required mutation into the MSM chromosome and is essentially dependent on homologous recombination. The p2NIL vector series, which was used for the construction of all suicide vectors, contains an *oriE* site for replication in *E. coli*, and the *aph*C gene which confers KAN^{R} for selection in both *E. coli* and mycobacteria (Parish & Stoker, 2000). As described in Section 2.4.1.6, the marker cassette was added to the p2*rpoB*^{MT} constructs using the *Pac*I cloning strategy. The marker gene cassette comprises Ag85p-l*acZ*; *hsp60p-sacB*; and the HYG^{R} gene (Gordhan and Parish, 2001). Using the Ag85 promoter, mycobacteria are able to constitutively express the *lacZ* operon and form blue colonies on X-gal (Gordhan and Parish, 2001), while cells expressing the *sacB* gene, express levansucrase and are unable to grow in media containing sucrose (Machowski *et al.*, 2005). The use of HYG^{R} as well as KAN^{R} decreases the likelihood of false positives due to spontaneous mutation (Gordhan and Parish, 2001). The allelic exchange process involves two selection steps. In the primary selection step cells are plated onto solid media containing KAN, HYG and X-gal. Blue colonies that are able to grow on KAN and HYG indicate that in those cells a single cross over (SCO) event has occurred and that the mycobacterial chromosome now includes the artificial vector as well as the native gene (Gordhan and Parish, 2001). Blue colonies from the primary selection step are streaked onto solid media to allow the second, DCO event to occur, and then plated on solid media containing sucrose as well as X-gal. White colonies that are sucrose tolerant indicate two possible events. The desirable outcome is that a DCO has occurred that has, in essence, “looped out” the vector containing the marker genes. If the *sacB* gene was still present, cells would not be tolerant of sucrose, and blue colonies would indicate that *lacZ* was still present. The second possibility is that cells are wild-type revertants that have eliminated the inserted DNA through the same process of homologous recombination. The protocol that was used was a slight departure from the standard allelic mutagenesis technique in that the original was developed to create gene deletion mutants (Gordhan and Parish, 2001). For our purposes, it was of utmost importance that the *rpoB* gene was kept exactly the same as (isogenic with) the wild-type, except for the specific RIF^{R} mutation. For this reason, an unmarked mutant was engineered. HYG^{R} genes could not be used to mark allele

replacement, nor could they be used to selectively drive a DCO event. DNA extracted from “white, sucrose resistant, second selection step colonies”, was screened via AS-PCR to ensure the point mutation has been successfully cloned into the *rpoB* gene. This system was developed for use in mycobacterial strains by Dr. Bhavna Gordhan, a long-time member of the MMRU, in collaboration with Dr. Tanya Parish (Gordhan and Parish, 2001).

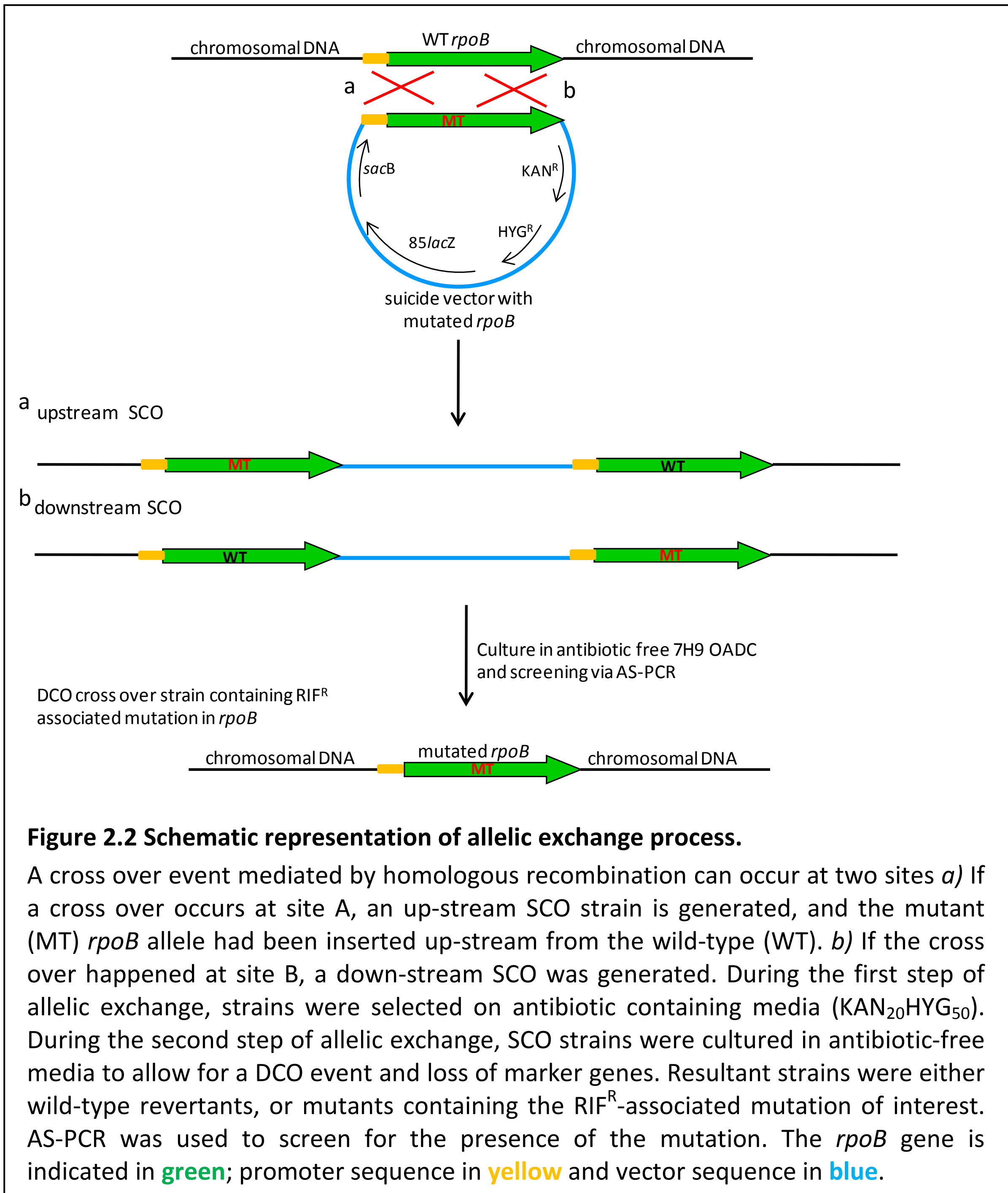


Figure 2.2 Schematic representation of allelic exchange process.
A cross over event mediated by homologous recombination can occur at two sites *a)* If a cross over occurs at site A, an up-stream SCO strain is generated, and the mutant (MT) *rpoB* allele had been inserted up-stream from the wild-type (WT). *b)* If the cross over happened at site B, a down-stream SCO was generated. During the first step of allelic exchange, strains were selected on antibiotic containing media ($KAN_{20}HYG_{50}$). During the second step of allelic exchange, SCO strains were cultured in antibiotic-free media to allow for a DCO event and loss of marker genes. Resultant strains were either wild-type revertants, or mutants containing the RIF^R-associated mutation of interest. AS-PCR was used to screen for the presence of the mutation. The *rpoB* gene is indicated in **green**; promoter sequence in **yellow** and vector sequence in **blue**.

2.4.2.1 First step of allelic exchange: Generation of SCO strains

Each suicide vector was introduced into MSM by electroporation as described in Section 2.3.5.2. After electroporation and cell recovery, cells were plated on 7H10 OADC media containing KAN_{20}; HYG_{50} and X-gal. Blue colonies that were able to grow on 7H10 OADC $KAN_{20}HYG_{50}$X-gal plates indicated that a probable SCO event had occurred. These colonies

were picked, and cultured overnight in 7H9 OADC media containing $KAN_{20}HYG_{50}$. DNA was extracted using the CTAB extraction method and the orientation and presence of the mutation was verified as described in 2.4.2.4.

2.4.2.2 PCR strategy to determine orientation of SCOs

During the first step of allelic exchange, the gene of interest can be inserted into the mycobacterial chromosome in one of two orientations. The mutated version of the *rpoB* gene may have been inserted up-stream of the native *rpoB* gene, or it may have been inserted down-stream of the native *rpoB* gene (Figure 2.3). A PCR strategy, using two primer sets, was designed to differentiate the up-stream and down-stream *rpoB* genes. The forward primer for the primer set one, UpOriF (F1 in Figure 2.3), was designed so that it would bind chromosomal sequence up-stream of the 300 bp *rpoB* promoter region. The reverse primer for primers set one, UpOriR (R1 in Figure 2.3), was designed to bind within the *rpoB* gene. In this way, only the up-stream gene would be amplified. For the primer set two, the forward primer, DownOriF (F2 in Figure 2.3), was designed to bind within the *rpoB* gene, while the reverse primer, DownOriR (R2 in Figure 2.3) (Appendix 5.5), was designed to bind in the chromosomal region down-stream of the *rpoB* gene. Given that the two *rpoB* genes would be separated in the chromosome by vector sequence, primer set one would be specific for the up-stream *rpoB* gene, and primer set two would be specific for the down-stream *rpoB* gene. The presence of the mutation was then checked using AS-PCR. If the mutation was present in the up-stream gene, then an up-stream SCO had been generated. Conversely, if the mutation was present in the down-stream *rpoB* gene, a down-stream SCO had been generated. This technique served the further purpose of confirming that the mutation had, been introduced into SCO strains.

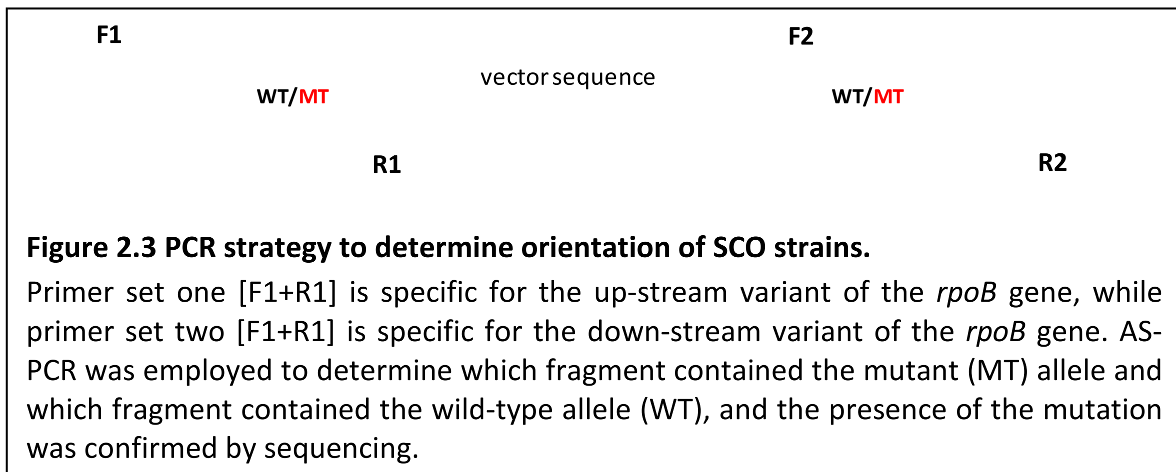


Figure 2.3 PCR strategy to determine orientation of SCO strains.
Primer set one [F1+R1] is specific for the up-stream variant of the *rpoB* gene, while primer set two [F1+R1] is specific for the down-stream variant of the *rpoB* gene. AS-PCR was employed to determine which fragment contained the mutant (MT) allele and which fragment contained the wild-type allele (WT), and the presence of the mutation was confirmed by sequencing.

To generate up-stream and down-stream *rpoB* fragments from SCO strains, DNA was extracted from SCO colonies using the CTAB method. Given that this was essentially a screening technique, and that DNA would not be used for further processing, the potential introduction of point mutations into the product was not that important, and therefore FastStart Taq DNA Polymerase™ was used using reaction components as described in Section 2.3.6.2. The cycling conditions for this PCR were as follows: Initial denaturation at 95 ˚C for 2 min; 30 cycles of [95 ˚C for 30 s; 62 ˚C for 30 s, and 72 ˚C for 90 s] and a final extension at 72 ˚C for 7 min. The product of the up-stream *rpoB* gene, using primer set one was 3259 bp; and the product for the down-stream *rpoB* gene, using primer set two was 2770 bp. These two product sizes were easily distinguishable on a 1 % agarose gel. Products were purified from an agarose gel, and AS-PCR was used to check for the presence of the mutations in either of the two fragments.

2.4.2.3 Step two of allelic exchange: Generation of DCOs

After the orientation of SCO strains and the presence of the mutation of these strains had been verified, DCO strains were generated. SCO strains were grown in antibiotic-free 7H9 OADC media overnight at 37 ˚C with shaking at 120 rpm, to allow for a DCO event to occur. Subsequently, a serial dilution series of cells was prepared, and plated in duplicate on 7H10 OADC containing 2 % sucrose, and standard 7H10 OADC media. White colonies that were able to grow on sucrose indicated that the marker genes (*lac*Z and s*ac*B) had been lost. DNA was extracted from these colonies using the colony boil methodology as

described in Section 2.3.2.7, and the presence (or absence) of the mutation was determined by AS-PCR.

2.4.2.4 Screening for the presence of a point mutation: AS-PCR

During the initial planning stages of this project, an allele-specific PCR (AS-PCR) strategy was designed to screen both suicide vectors and DCOs. This technique relies on the specificity of a particular primer for its complementary template in a PCR reaction, and forms the basis of commercially available molecular diagnostic kits for MDR-TB (Makinen *et al.*, 2006). In this project, two sets of primers were designed for each mutation – one that is specific for the wild-type allele, and another that is specific for the mutant allele. Primers were designed so that the variable nucleotide was at the last position at the 3' end of the primer. If the wild-type-specific primer is used in a reaction with a template containing the wild-type allele, a PCR product of the correct size should be formed. If the wild-type primer is included in a reaction containing a template with a mutant allele, no product should be formed. The same principle applies for mutant alleles and mutant specific primers (Figure 3.7). The AS-PCR methodology used to screen for mutations in plasmid DNA differed slightly from the methodology used to screen for mutation in DCO chromosomal DNA in that it required two rounds of PCR. During the first round, general outer primers were used to generate a 311 bp fragment that contained the RRDR. In the second of PCR, the AS-primer was used, together with a general reverse primer.

a) AS-PCR: Screening for the presence of mutations in suicide vectors

Once suicide vectors had been successfully generated, AS-PCR was employed to check for the presence of a mutation. To screen plasmid DNA two PCR reactions were set up for each sample. In each reaction, either wild-type-specific or mutant specific primer was used. Naming, and sequence, of mutant- or wild-type-specific primers are dependent on the mutation being screened (Appendix 5.5). In addition, in each reaction, a general forward and a general reverse primer were used (SDM outer F and SDM outer R). The addition of two general primers results in one band being formed for every reaction, which acts as a control to ensure successful amplification. If a wild-type-specific primer was used with a sample that contained the wild-type allele, an extra band would be

visible on the agarose gel. If a wild-type-specific primer was used in a sample containing a mutant allele, no extra product would be formed. For screening purposes, FastStart Taq DNA Polymerase™ was employed using components as described in Section 2.3.6.2. The product of the general primers, SDM outer F and SDM outer R, was visible as a 1578 bp band on a 1 % agarose gel. Expected sizes of the additional (diagnostic) bands were as follows: S531L) 546 bp H526Y) 559 bp L521L) 576 bp Q513V) 597 bp TH526TY) 559bp 30bp del) 597 bp. Annealing temperature of primers is important for AS-PCR therefore a gradient PCR was performed to ensure that binding of primers was specific (Figure 3.8). If the annealing temperature was too low, primers would bind to sequence non-specifically. If the primer specificity was reduced, a wild-type specific primer could bind to a mutant allele, and vice versa. This would have generated false results. If the annealing temperature was too high, primers would not have bind at all. Gradient AS-PCR conditions were as follows: Initial denaturation: 95 ˚C for 2 min; 30 cycles of [95 ˚C for 30 s; (55 °C; 56.9 °C; 61.1°C or 64.2°C) for 30 s, and 72 ˚C for 2 min] and a final extension at 72 ˚C for 7 min. Based on results from gradient AS-PCR an annealing temperature of 60 °C was selected for subsequent AS-PCR based screening of suicide vectors.

2.4.2.5 AS-PCR: screening for the presence of mutations in DCOs

The strategy to screen for the presence of a mutation in chromosomal DNA from possible MSM DCOs required considerable optimisation. Only MSM mutants containing S531L, H526Y, and L521L alleles were generated in this study; therefore, AS-PCR was only adapted for these mutations. In this version of the AS-PCR, two separate PCR steps were required. Instead of using three primers in one reaction, a fragment of the *rpoB* gene was amplified from the MSM genome using general primers. This PCR product was used in the next PCR as a template, with an AS-forward primer and a general reverse primer. If a band was formed in the second reaction, the correct AS-primer had been used with the correct allele. Positive controls used for this PCR included wild-type MSM chromosomal DNA, and chromosomal DNA from MSM strains with known *rpoB* mutations (generated by UV exposure, Digby Warner, unpublished results). FastStart Taq DNA Polymerase™ was used for this screening technique with components as described in Section 2.3.6.2. SeqF RRDR and SeqR RRDR primers were used to 311 bp fragment from MSM

chromosomal DNA (Appendix 5.5). Cycling conditions for the first round of PCR were as follows: Initial denaturation 95 ˚C for 2 min; 20 cycles [95 ˚C for 30 s; 62 ˚C for 30 s, and 72 ˚C for 30 s] and final extension at 72 ˚C for 7 min.

The 311 bp product from reaction one was used as a template in the second round of PCR, where the AS-primer was used as the forward primer, and the Whole R3 primer was used as the reverse primer. Cycling conditions for this PCR were as follows. Initial denaturation: 95 ˚C for 2 min; 20 cycles of [95 ˚C for 15 s; 62 ˚C for 15 s, and 72 ˚C for 15 s] and a final extension at 72 ˚C for 7 min. Given the short product formed by this round of PCR, only 15 s of extension time was required. Specific sizes of mutant bands are as follows: S531L) 109 bp H526Y) 122 bp L521L) 139 bp. A 2 % agarose gel was required to view these products.

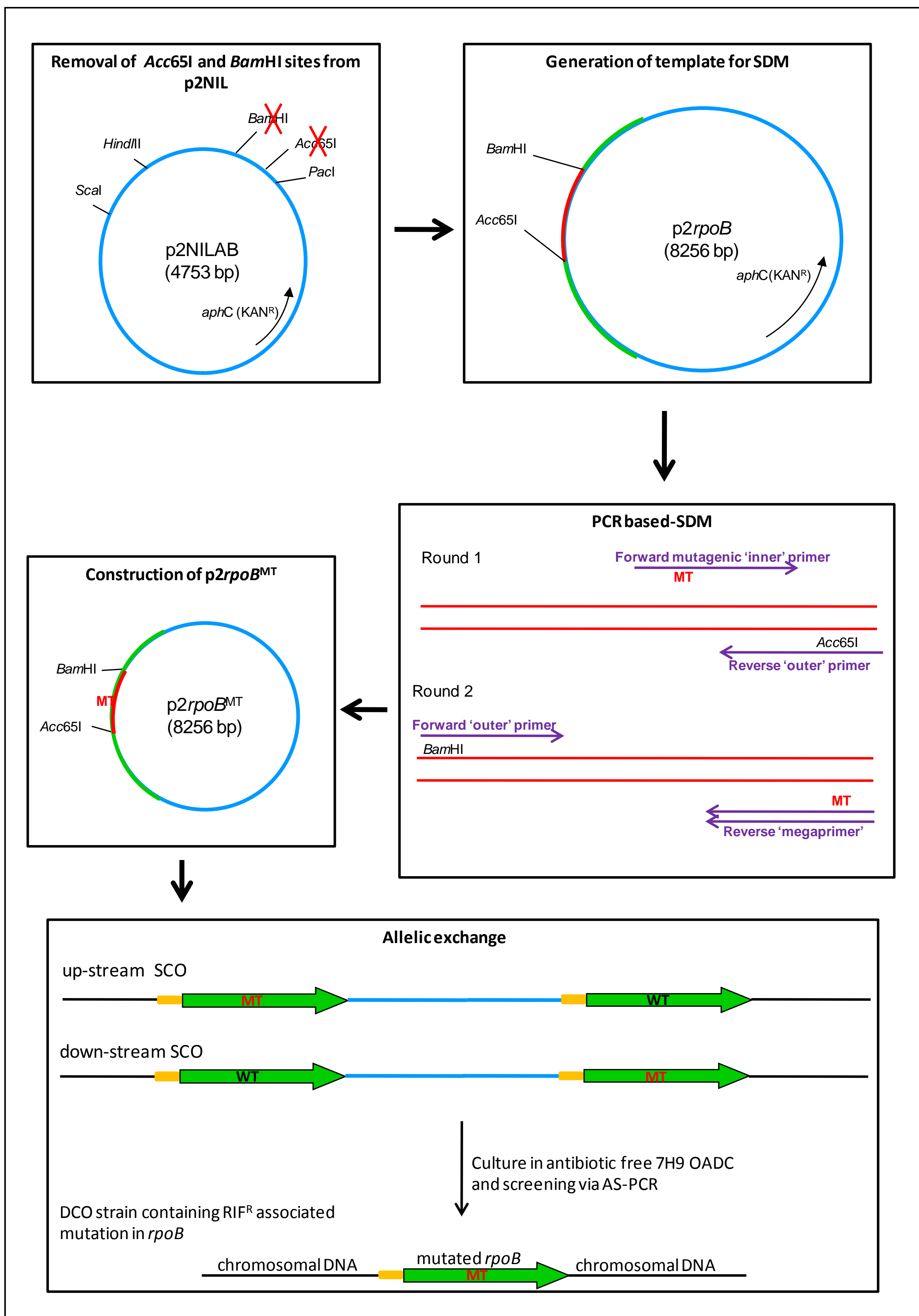

Removal of Acc65I and BamHI sites from p2NIL
BamHI
HindIII
Acc65I
PacI
ScaI
p2NILAB
(4753 bp)
aphC (KANR)
Generation of template for SDM
BamHI
Acc65I
p2rpoB
(8256 bp)
aphC(KANR)
PCR based-SDM
Round 1
Forward mutagenic 'inner' primer
MT
Acc65I
Reverse 'outer' primer
Round 2
Forward 'outer' primer
BamHI
MT
Reverse 'megaprimer'
Construction of p2rpoBMT
BamHI
MT
Acc65I
p2rpoBMT
(8256 bp)
Allelic exchange
up-stream SCO
MT
WT
down-stream SCO
WT
MT
Culture in antibiotic free 7H9 OADC
and screening via AS-PCR
DCO strain containing RIFR associated
mutation in rpoB
chromosomal DNA
mutated rpoB
MT
chromosomal DNA

Figure 2.4 Generation MSM strains containing RIF^R-associated mutations in *rpoB*.
During the first phase, *Acc*65I, and *Bam*HI sites were removed from the p2NIL vector (Section 2.4.1.2) to generate p2NILAB. The wild-type *rpoB* gene was isolated from MSM via PCR, and cloned into p2NILAB to generate p2*rpoB* (Section 2.4.1.2). SDM was performed on a portion of the wild-type *rpoB* gene to introduce RIF^R-associated point mutations, (or Xover-PCR for the 30 bp deletion) into the *rpoB* gene. This fragment was cloned back into the p2*rpoB* vector, to generate p2*rpoB*^{MT} (Section 2.4.1). These suicide vectors were electroporated into MSM, to allow for allelic exchange (Section 2.4.2). During the first step of SCO strains were selected on $KAN_{20}HYG_{50}$X-gal containing 7H10 OADC plates. SCO strains were cultured in antibiotic free media to antibiotic free media to allow for generate DCO strains. The presence of a mutation was determined using AS-PCR (Section 2.4. 2. 5) and confirmed by sequencing. The *rpoB* gene is indicated in **green**; promoter sequence in **yellow** and vector sequence in **blue**.

2.4.3 Construction of *rpoB* containing pOLYG plasmids and electroporation into MSM

The pOLYG vector is a HYG^R–containing vector that replicates episomally in MSM (O'Gaora *et al.*, 1997) and exists in multicopy at up to 23 copies per cell (Huff *et al.*, 2010). Insertion of the *rpoB* gene into this vector allowed evaluation of the implication of over-expression of *rpoB* in MSM. Wild-type and S531L and H526Y mutant versions of *rpoB* were inserted on the pOLYG vector. The *rpoB* containing pOLYG constructs were generated by digesting the pOLYG vector and the p2*rpoB* template with *Acl*I and *Hind*III under conditions described in 2.3.4.1. The enzymes have the same buffer conditions and could therefore be used simultaneously in a single reaction. Digestion of the pOLYG backbone with *Acl*I and *Hind*III resulted in 55 bp and 5260 bp fragments; the 5260 bp fragment was purified from the gel. Digestion of p2*rpoB* with *Acl*I and *Hind*III yielded 4107 bp and 4124 bp fragments, and the 4124 bp fragment was isolated from the gel. A ligation reaction was set up under conditions described in 2.3.4.3, and molar ratios of vector:insert of 1:1, 1:2 and 1:3 were used. The resultant ligation was transformed into competent *E. coli* cells, and plated on LA media containing HYG_{200}. The pOLYG backbone contains resistance genes for HYG, and therefore only cells containing the vector grow on HYG_{200} LA plates. DNA was extracted using the small-scale extraction protocol, and screened for the presence of the correct vector. Cells containing the correct vector were inoculated into a larger scale culture, and a large-scale DNA extraction performed. AS-PCR was used to confirm the presence of the mutation, where mutated versions of *rpoB* had been inserted into the pOLYG vector. Resulting constructs were transferred into MSM via electroporation as described in 2.3.5.2. Cells were selected on 7H10 OADC

media containing HYG_{50}. The HYG^R gene is expressed at higher levels in *E. coli* than it is in MSM, therefore higher concentrations of HYG are used to select for *E. coli* cells containing the pOLYG constructs than MSM cells.

2.4.3.1 Electroduction

The pOLYG constructs containing *rpoB* were isolated from putative MSM transformants after electroporation and selection by electroduction into *E. coli*. A single MSM colony containing the construct was re-suspended in 50 µl cold (4 ˚C) 10 % glycerol and incubated on ice for 10 min, and 70 µL of the electrocompetent *E. coli* cells were added. The sample was transferred to a 0.1 cm electroporation cuvette and cells were pulsed with the following conditions: 1.8 kV, 25 µF and 200Ω. A 500 µl aliquot of LB media was added to electroporated cells, which were incubated for 1 hour at 37 °C to allow for recovery. Cells were plated on LA plates containing HYG_{200}. DNA was isolated from resultant colonies as described in 2.3.2.1.

2.4.3.2 Toxicity evaluation and restriction analysis

Re-arrangement of a plasmid after electroporation suggests that expression of the plasmid is toxic to the host cell. A restriction analysis of *rpoB* containing pOLYG plasmids was conducted to check for re-arrangement of the plasmid. DNA from cells obtained during electroduction experiments was extracted using the large-scale DNA extraction protocol (Section 2.3.2.1). The following restriction enzymes were applied to check the integrity of the plasmids *Aat*II, *Eco*RI, *Nae*I and *Nsp*I (New England Biolabs, USA) under conditions described in 2.3.4.1 with buffers as described in Appendix 5.3. Standard pOLYG, as well as samples of the plasmids before electroporation, were also digested as controls.

2.4.5 RIF susceptibility (RIF^S) analyses

2.4.5.1 RIF^S on solid media

Strains to be tested on RIF-containing solid media were grown to logarithmic phase and serial dilutions were prepared in un-supplemented 7H9 liquid medium. Aliquots (20 µl) of serial log-fold dilutions (10^0 – 10^{-6}) were spotted on 7H10 OADC media containing 50 µg/ml; 100 µg/ml; 200 µg/ml or 400 µg/ml RIF (RIF_{50}; RIF_{100}; RIF_{200}; RIF_{400} respectively). Strains were also plated on standard OADC to measure growth, as well as 7H10 $KAN_{20}HYG_{50}$Xgal plates to determine cross over status

2.4.5.2 Mycobacterial Alamar Blue Assay (MABA)

To assess the sensitivity of MSM strains containing RIF^R associated mutations, the MABA technique was employed. This assay has been shown to detect drug resistance in clinical MTB isolates, and was modified from published standards to assay RIF^S in MSM strains (Collins & Franzblau, 1997, Franzblau *et al.*, 1998, Palomino *et al.*, 2002). The utility of this assay resides largely in its use of 96-well plates which allow evaluation of MSM growth across a variety of MSM concentrations.

The functionality of the microtitre plate derives from its ability to allow different concentrations of a compound(s) to be tested to be generated *down* the rows, while different kinds of growth analysis can be accomplished *across* the plates in the columns. A 50 µl sample of appropriate liquid media was added to all wells, except the upper-most wells. The compounds to be tested were then added to the upper-most row. The outer-most columns served as controls to ensure that normal growth had occurred; therefore 100 µl of media was added to those wells at the top of the outer-most columns. The next two inner columns served as controls to assay whether the RIF diluent had any effect on cell growth; therefore, 100 µL of the dimethyl sulphoxide (DMSO) diluent was added to the top most wells in the second outer-most column. The 4 innermost columns contained decreasing concentrations of RIF, and therefore 100 µL of an appropriate starting concentration of RIF was added to the upper-most well of the four inner columns of the plate. Initially a starting concentration of 1000 µg/ml of RIF was added, however when growth in higher concentrations of RIF were to be assayed a starting concentration of

4000 µg/ml of RIF was used. Using a Research Pro 8 multichannel pipette (Eppendorf, Germany) 50 µL from the upper-most well was added to the second row, and then 50 µl was removed from this row and added to the next row. In this way a serial dilution of the contents in the upper-most wells (media, diluents and RIF) was achieved.

Strains to be tested were grown to logarithmic phase in 7H9 OADC media, and diluted to and OD_{600} of 0.05 in un-supplemented 7H9 media. An aliquot (50 µl) of cells were added to all wells except the upper-most well. It is worth mentioning here that this halved the concentration of drug or diluents present in the wells after serial dilution. For example, if a 1000 µg/ml starting concentration of RIF was used, and serially diluted, the next row would contain a 50 µl of RIF at 500 µg/ml. However, when 50 µL of cells were added the concentration of RIF in this row was halved to 250 µg/ml. Moreover, by added cells to wells already containing 50 µl of liquid, the concentration of cells was halved to an OD_{600} of 0.025.

Plates were incubated for 48 hours at 37 ˚C without shaking. After 48 hours of incubation, 10 µL Alamar Blue (Invitrogen, USA) was added to each well, and plates were incubated for a further 24 hours. The active compound in Alamar Blue, resazurin, changes colour upon reduction. A change from blue to pink indicates that growth has occurred. Given the deep red colour of RIF, there were some instances where a resazurin colour change could not be used to evaluate growth. In these cases, pellets from a selection of wells were plated on standard 7H10 OADC media to ensure that pellets were representative of cell growth, and not of cell debris due to cell death.

Step 1: Media added to each well except uppermost well

Step 2: Diluent/ drug added to topmost w

Step 3: Dilution series prepared

Step 4: Cells added to every well except uppermost well

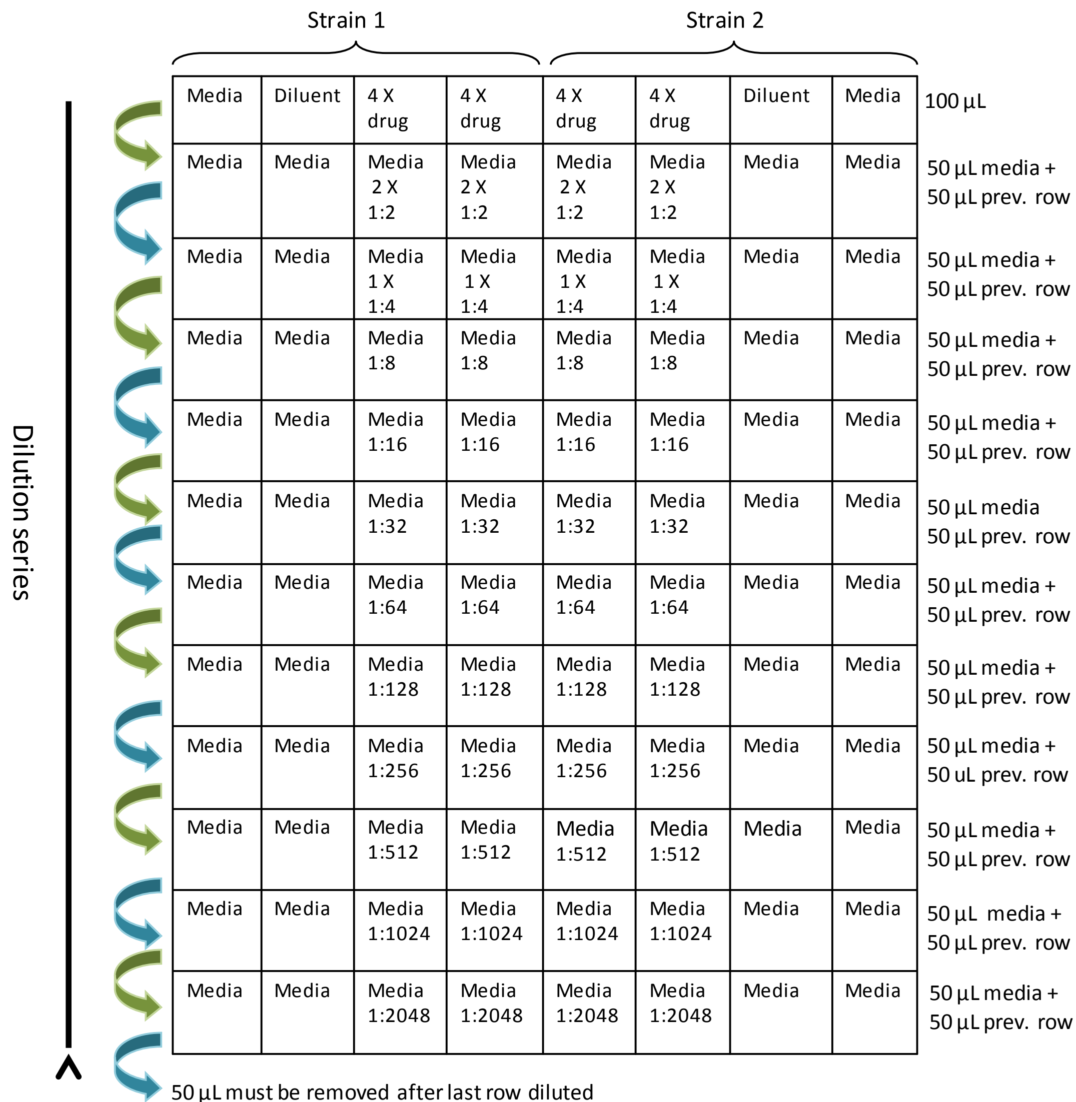


Figure 2.5 Schematic representation of preparation of MABA 96-well plates. Media is added to every well except the uppermost well. Diluent or RIF is added to the uppermost well, and a dilution series is prepared. At this point each well contains a 100 µL solution. Thereafter, cells are added to every well, except the topmost well, which serves as a positive control. The final volume is each well is 150 µL

2.4.6 Evaluation of the physiological effects of RIF^R-associated mutations: competition assays

Competition assays were used to evaluate competitive growth between MSM strains containing an S531L mutation, and parental MSM strains. Each strain was grown in parallel in axenic culture as a control. Competition experiments were performed using 7H9 OADC media, based on a modified version of Gagneux *et al.* (2006b). Strains were grown to logarithmic phase, and diluted to an OD_{600} of 0.05. Equal amounts of cells were added into a 250 ml Erlenmeyer flask. Cultures were incubated at 37 ˚C with 120 rpm shaking. For the first day, 1 ml samples were taken at 0 hours; 4.5 hours and 9 hours. For the next 6 days, samples were taken at 24 hour intervals, and every seven days thereafter. Samples (1 ml) were serially diluted in 0.5 % tween80. For measurements from the first day, 400 µl aliquots from dilutions 10^{-4} – 10^{-7} were plated on standard 7H10 plates, and 400 µl were plated on plates containing RIF_{200}. For all following measurements dilutions of 10^{-6} – 10^{-8} were plated. Plates were incubated for seven days, and colonies enumerated. An average of CFU/ml from at least two plates was used, and experiments were conducted in triplicate up until week 3. Two sets of cultures became contaminated after week three, therefore subsequent results describe a single experiment and so constitute preliminary data. The number of colony forming units (CFU/ml) was calculated using the following formula:

[(number of colonies) X (dilution factor) X 2.50] = CFU/ml

The relative fitness of strains was determined using the formula:

$$W = \ln (R^F/R^0)/\ln (S^F/S^0)$$

R^0 and S^0 indicate cell concentration of resistant (R) and susceptible (S) strains at the beginning of the experiment, and R^F and S^F indicate the final cell concentration of the same strains (Gagneux *et al.*, 2006b).

3. Results

3.1 Selection of *rpoB* mutations to study in this project

A number of different *rpoB* mutations have been identified in clinical MTB strains that are resistant to RIF. These mutations have been associated with varied levels of RIF^R and have been found to impose heterogeneous fitness costs (Gagneux *et al.*, 2006b, Mariam *et al.*, 2004, Gillespie *et al.*, 2002). At the onset of this study, we conducted an analysis of published literature to generate an inventory of RIF^R-associated *rpoB* alleles identified in MTB in both clinical and *in vitro* studies. Our analysis revealed a set of mutations that are most commonly identified in both clinical and laboratory-generated RIF^R MTB mutants (Table 3.1). More recently, Sandgren *et al.* (2009) established the TBDReaMDB to document RIF^R mutations in MTB (Sandgren *et al.*, 2009), that became available subsequent to the completion of our analysis. Comparison of their results with ours reveals minor differences in reported mutation frequencies. However, these discrepancies are readily ascribed to the fact that Sandgren *et al.* (2009) used very specific inclusion criteria (for example, only clinical studies with more than 100 samples, that describe the proportion of both sensitive and resistant strains with an associated mutation were included) and were far more exhaustive in their analysis (Sandgren *et al.*, 2009). The TBDReaDB was not available at the start of this project, and therefore our literature analysis was intended to give some indication of commonly occurring mutations in clinical and lab-derived MTB isolates. Consequently, our literature analysis was by no means exhaustive, and we did not impose strict inclusion criteria.

Based on the results of our analysis (Appendix 5.1), we selected five *rpoB* alleles for further study. Table 3.1 provides a list of the alleles selected for introduction into MSM, as well as the rationale for the investigation of these particular RIF^R-associated mutations. It should be noted that, since the RRDR was initially defined in *E. coli* (Jin & Gross, 1988), mutations are conventionally referred to using the numbering of the *E.coli rpoB* gene (Telenti *et al.*, 1993).

The observation that a silent mutation (L521L) was associated with relatively high levels of RIF^R in a clinical isolate (Siddiqi *et al.*, 2002) suggested this allele for inclusion in the study. The remaining three mutations were included owing to their potential significance as established by previous work at the MMRU (Table 3.1). The TH526TY mutation is commonly associated with error prone DNA polymerase activity, in particular DnaE2 (Boshoff *et al.*, 2003). In preliminary experiments at the MMRU, a MSM *rpoB* mutant containing a Q513V mutation displayed differential growth on OADC (oleic acid-albumin-dextose-catalase) vs. Glucose-NaCl (Glucose Salt; GS) containing 7H10 media (D Warner and Susan Puckett, unpublished results). The 30 bp deletion has been isolated at the MMRU in response to DNA damaging agents and nutrient stress, and has also been studied in an independent study (Malshetty *et al.*, 2010).

Table 3.1 RIF^R-associated mutations to be examined in this project

Mutation[a]	**Reason for study**
S531L (t**c**g→t**t**g)	Most frequently isolated RIF^R mutation in both clinical isolates as well as lab-derived RIF^R strains (Appendix 5.1)
H526Y (**c**ac→**t**ac)	Another common RIF^R conferring mutation (Appendix 5.1)
L521L (**c**tg→**t**tg)	Silent mutation observed in clinical RIF^R MTB strains (Yang *et al.*, 1998, Siddiqi *et al.*, 2002)
Q513V (**ca**g→**gt**g)	In preliminary research at the MMRU, the equivalent mutation incurs no obvious growth phenotype in MSM cells cultures in different media *in vitro*; for this reason, the Q513V allele was identified as a potentially useful negative control.
T-H526T-Y (ac**cc**ac→ac**tt**ac)	A double transition mutation that is commonly associated with DnaE2 activity in both MTB and MSM. DnaE2 is an important error-prone DNA repair enzyme (Boshoff *et al.*, 2003).

A 30 bp deletion in the *rpoB* gene (Q512 – S522) (cagttcatggacca gaacaacccgctgtcg)	During studies at the MMRU, this deletion was isolated on more than one occasion, in response to exposure to UV irradiation, starvation, and treatment with replication inhibitors (Digby Warner, unpublished). This deletion has also been isolated in a separate study (Malshetty *et al.*, 2010).

[a]Mutations have been numbered according to the standard system based on the *E. coli* genome as per standard practise (Telenti *et al.*, 1993)

3.2 The use of MSM as a study organism

There has been significant debate regarding the use of MSM as a model for answering questions pertinent to control of TB (Barry, 2001, Reyrat & Khan, 2001). There are significant differences in genome composition and biology between these two organisms (Barry, 2001) and moreover, there are important disparities in RIF metabolism. MSM has intrinsic RIF detoxifying mechanisms, and can tolerate much higher concentrations of RIF than MTB (Quan *et al.*, 1997, Alexander *et al.*, 2003, Dey *et al.*, 2010). In spite of these important differences, several factors make MSM a useful study organism. The feasibility of introducing point mutations conducting this study in MSM or MTB was evaluated and it was clear that the use of MSM was preferable.

Several physiologically significant and clinically relevant mycobacterial functions have been uncovered while studying MSM. For example, the characterisation of two proteins that form part of a damage-induced mutagenesis system in MTB was accomplished using an MSM based system (Warner *et al.*, 2010). This enzymatic machinery plays an important role in MTB adaptation and the development of drug resistance, and is therefore an attractive drug target. The ability of MTB to modulate its growth is a key factor for virulence and survival under stress. Much of this growth modulation is dependent on rRNA synthesis (Stallings *et al.*, 2009). Using MSM, Stallings *et al.* (2009) revealed by binding to RpoB, CarD regulates rRNA synthesis, and went on to show that CarD was essential for both replicative and persistent phases of MTB infection in mice. In a study more directly relevant to this project, genetic engineering of MSM was used to suggest a model for evolution of aminoglycoside resistance in MTB (Shcherbakov *et al.*,

2010). Moreover, TMC207 – a promising new chemotherapeutic that inhibits both drug sensitive and drug-resistant TB – was initially isolated in a screen against MSM (Andries *et al.*, 2005). These studies exemplify the usefulness of MSM for elucidating important physiological mechanisms, and how this information can be applied to clinical problems.

Several practical considerations also informed the use of MSM. MSM is non-pathogenic, which allows use of this organism in a BLS2 laboratory, whereas any work involving MTB culture requires strict BSL3 conditions. MTB has an approximate generation time of 24 hours as established in our lab, and colonies can take up to a month to be visible on a plate. In contrast, MSM has a generation time of 3 – 4 hours (Snapper *et al.*, 1990), and growth can be elucidated after 2 – 3 days. Given the time constraints for a Masters research project, both biosafety and culture characteristics favoured MSM as a study organism. In addition, MSM is more amenable to genetic manipulation than MTB. SDM and allelic exchange were the methods of choice to introduce *rpoB* mutations into the chromosome of MSM. Although this kind of genetic manipulation was more time consuming, it was preferred as mutations could be introduced in the complete absence of RIF. We were firm on this, because any exposure to RIF may have introduced second-site mutations which would confound any assay of drug resistance, or physiological change. It is worth mentioning here, that this strategy fashioned the dual nature of this project. Generation of strains containing RIF^R-associated mutations in the absence of RIF provided an opportunity to investigate the role of these mutations in RIF^R, without second-site mutations. Had this been the primary aim, MTB might have been a more suitable model. However, the driving aim of this project was to investigate the implications of point mutations in *rpoB* for mycobacterial physiology, so for this purpose MSM was suitable.

Recombineering has been used to introduce point mutations into the mycobacterial genes without the need for selection (van Kessel & Hatfull, 2008, van Kessel & Hatfull, 2007). However, when this technique was applied to RIF^R-associated mutations, RIF was used to select for mutants (van Kessel & Hatfull, 2008). Although it was technically feasible to use recombineering to generate point mutants without selection (van Kessel & Hatfull, 2008), mutagenesis and screening techniques would still have required

optimisation. Moreover, our laboratory has extensive experience in generating mutants via allelic exchange. For these reasons the combinatorial process of SDM and allelic exchange was therefore the preferred method for this project. Until this study, allelic exchange had never been used to introduce RIF^R-associated mutations into the chromosome of a mycobacterial strain. Given the essentiality of *rpoB* for bacterial function, there was no certainty that this kind of genetic manipulation was viable in mycobacteria. The allelic exchange procedure requires two selection steps and the only difference between the wild-type gene, and the mutated version of *rpoB* (*rpoB*^{MT}) was a point mutation. Therefore, screening for mutants after allelic exchange was accomplished using AS-PCR. The use of MSM provided the opportunity to assess the feasibility of this process in mycobacteria, within the time constraints of an MSc. After consideration of the aforementioned factors, we were confident that study of MSM would not only facilitate a proof-of-concept experiment that could be repeated in MTB, but also inform important issues relevant to mycobacterial physiology.

3.3 Genomic context, sequence and structural similarity of MSM RpoB and MTB RpoB: A bioinformatic analysis

The use of MSM to investigate drug resistance is expected to yield interesting insights into mycobacterial physiology. However, the clinical importance (and therefore clinical relevance) of this project demanded that all “proof-of-principle” work be applicable, as far as possible, to MTB in terms of the techniques developed as well as the implications for fundamental mycobacterial physiology. For this reason it was important to examine the similarities between the MSM and MTB *rpoB* genes – at the sequence level, and in terms of genomic context. Moreover, this information was essential for designing efficient cloning and allelic exchange strategies. The JCVI CMR (http://cmr.jcvi.org/tigr-scripts/CMR/CmrHomePage.cgi) (Peterson *et al.*, 2001) was used to examine the *rpoB* gene in both organisms.

As shown in Figure 3.1, the genomic context of *rpoB* is conserved in mycobacteria: although the genes occurring up-stream of *rpoB* vary between species, in all cases *rpoC* is

located down-stream of *rpoB*. For allelic exchange, it was important to include sufficient promoter sequence in the suicide construct, (as discussed in Section 3.5.2) however, the sequence and co-ordinates of the *rpoB* promoter region in mycobacteria have not been defined. The gene map from the JCVI CMR indicates that in both MTB and MSM the *rpoB* gene is separated from the up-stream gene (*Rv0666* in MTB; *MSMEG_1366* (a putative ATP binding protein) in MSM,) by a large intergenic region. In MTB, the region is 498 bp in size, in MSM the region is 600 bp (Figure 3.1). In *E. coli*, *rpoB* occurs in a gene cluster that includes the *rplKALJ* operon which encodes the ribosome large-subunit (Passador & Linn, 1989). Expression of *rpoB* and *rpoC* in *E. coli* is driven by strong promoters up-stream of *rpLK* and *rpLJ*. In contrast, the genetic organisation and transcriptional regulation of *rpoB* differs in *B. subtilis* (Boor *et al.*, 1995). In this gram positive organism, an additional open reading frame (ORF) is present, which encodes a 22513 Da protein that is essential for *B. subtilis* viability (Boor *et al.*, 1995). The *rpoBC* promoter element overlaps with the 3' end of this ORF therefore expression of *rpoBC* is not dependent on promoter sequence up-stream of the *rplK* and *rplJ* genes. If the mycobacterial *rpoB* gene arrangement is comparable to *B. subtilis* (Boor *et al.*, 1995), the region up-stream of the *rpoB* gene is expected to contain the *rpoB* promoter. For this reason, the conservative decision was taken to retain at least 300 bp up-stream sequence in all MSM *rpoB* mutation alleles. Subsequent results (see Section: 3.5.2) provided strong evidence to suggest that the promoter was located in this region.

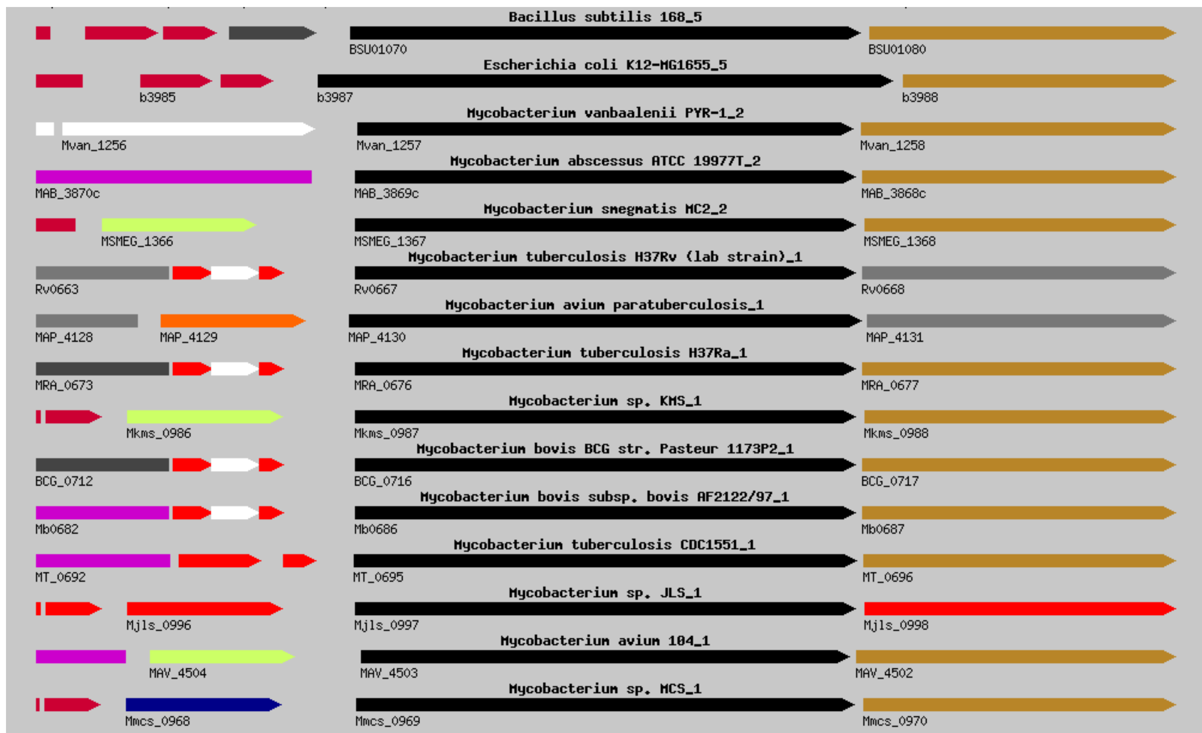


Figure 3.1 The genomic context of *rpoB* is conserved in mycobacteria.
The genomic context of *rpoB* in all mycobacterial genomes available on the JCVI CMR database (http://cmr.jcvi.org/tigr-scripts/CMR/CmrHomePage.cgi) (Peterson *et al.*, 2001) was retrieved on 03 January 2010. The *E. coli* and *B. subtilis rpoB* gene context is included at the top of the figure. The *rpoB* gene is indicated in the centre of the figure in black, and genes occurring immediately up- or down-stream are shown on either side.

When MSM and MTB *rpoB* genes were compared using the BLASTn algorithm (Altschul *et al.*, 1990) they were found to be 87 % similar. When just the RRDR region was compared 90 % similarity was reported. At the amino acid level the full length *rpoB* gene from these two organisms was 91 % similar. Most notably, the RRDR regions of both MTB and MSM RpoB proteins display perfect identity (Figure 3.2). To gain a sense of the position of important structural domains in RpoB, and how *rpoB* mutations may affect them, structural domains were mapped onto a multiple sequence alignment (MSA) generated in PROMALS3D (http://prodata.swmed.edu/promals3d/promals3d.php) (Pei *et al.*, 2008). The PSI-BLAST (Altschul *et al.*, 1997), and PSIPRED (Jones, 1999) databases are then scanned for secondary structure determination, and Hidden Markov Models are employed to examine the probabilities of the matches (Pei & Grishin, 2007). The conservation index is then determined, which indicates the level of functional or structural significance of conservation at a particular site (Pei & Grishin, 2001).

Structural domains as defined by Lane and Darst (2010) were mapped onto the PROMALS3D alignment (Figure 3.2). Although it is beyond the scope of this discussion to describe structure-function relationships in detail, particularly interesting and pertinent features are summarized below. It is important to note that structural information comes from crystal structures of *T. thermophilus* (Vassylyev *et al.*, 2007) and *T. aquaticus* (Zhang *et al.*, 1999), and structural domains may not be precisely the same in other organisms. The β-lobe 1 domain contains the active site channel and covers the RNA/DNA hybrid (Lane & Darst, 2010b), and makes extensive interactions with the β' subunit (Zhang *et al.*, 1999). In addition, β-lobe 1 contains a site important for interactions with transcription-repair coupling factor (TRCF), and is therefore important for resolution of stalled RNAP (Westblade *et al.*, 2010). The β-lobe 2 domain covers down-stream DNA during transcription and a large insertion in the *E. coli* sequence of this domain alters promoter specificity (Lane & Darst, 2010b). Together β-lobe 1 and 2 may have a role for opening and closing of the main RNA channel, which is important for maintenance of the transcription bubble during elongation, and for σ factor binding and release (Borukhov & Nudler, 2008). The fork loop 2 region is important for unwinding DNA during elongation, and maintaining the transcription bubble in the transcriptional elongation complex (TEC) (Lane & Darst, 2010b), and may play an important role in RNAP translocation (Borukhov & Nudler, 2008). Domains β-a10 – 14 contain key components of the RNAP active site and the β-flap contained within these domains is an important constituent of the RNA exit channel during elongation (Lane & Darst, 2010b), and for interactions with the β' subunit (Darst, 2001). This region also undergoes significant alteration during σ factor binding (Vassylyev *et al.*, 2002), and directly interacts with RNA during exit, which has implications for pausing and termination (Borukhov & Nudler, 2008). Within the flap tip, a short region termed the flap tip helix (not shown in Figure 3.2, but coordinates are listed in the Figure legend) is important for hairpin induced pausing (Lane & Darst), and for σ factor interaction (Vassylyev *et al.*, 2002). The β-clamp region continues in the β' primer subunit, and forms an important structural unit for elongation (Lane & Darst, 2001b).

The RRDR region is mapped, and residues that directly interact with RIF are indicated in Figure 3.2 (hydrogen bonds: orange; van der Waals interactions: purple). The crystal

structure of core *T. aquaticus* RNAP enzyme, in complex with RIF, indicates that RIF binds close to the DNA/RNA channel, but only comes within 12 Å of the active site (Campbell *et al.*, 2001). Six amino acids are able to form hydrogen bonds with RIF, and another six are likely to partake in van der Waals interactions. Of these 12 amino acids, 11 mutate in association with RIF^R (I572, *E. coli* numbering, is just outside the RRDR). The 12th (E565, *E. coli* numbering) is highly conserved, and mutations at this site might be lethal (Campbell *et al.*, 2001). However, this residue falls outside of the RRDR and therefore, mutations here may not have been detected in studies of RIF^R isolates. Studies have shown that RIF^R-associated mutations can have pleiotropic effects for transcriptional mechanisms (Jin & Gross, 1989, Jin & Gross, 1991, Landick *et al.*, 1990). Looking at the map in Figure 3.2, it would not seem that the RRDR sits completely in any structural domain. However folding to form tertiary structures can greatly influence domain interactions, and therefore the RRDR is likely to intereact with structurally important regions. RIF^R-associated mutations have long been studied for their effect of mechanistic aspects of transcription (Jin & Gross, 1989). Given the importance of RpoB for genetic expression, it is likely that the potential functional consequences of *rpoB* mutations have physiological implications.

```
                                                                                β-lobe 1
Conservation:            9                                    6  666  666 69 9  69  96
TTH                   1  M-----------------------------------EIKRFGRIREVIPLPPLTEIQVESYRRALQADVP   35
TAQ                   1  M-----------------------------------EIKRFGRIREVIPLPPLTEIQVESYKKALQADVP   35
ECO                   1  M----VYSYTEKKR----------------------IRKDFGKRPQVLDVPYLLSIQLDSFQKFIEQDPE   44
SAU                   1  M-AGQVVQYGRHR-----------------------KRRNYARISEVLELPNLIEIQTKSYEWFLRE---   43
BSU                   1  M-TGQLVQYGRHR-----------------------QRRSYARISEVLELPNLIEIQTSSYQWFLDE---   43
MSM                   1  MLEGCILAVSSQSKSNA---------ITNNSVPGAPNRVSFAKLREPLEVPGLLDVQTDSFEWLVGSDRW   61
MTB                   1  MLEGCILADSRQSKTAASPSPSRPQSSSNNSVPGAPNRVSFAKLREPLEVPGLLDVQTDSFEWLIGSPRW   70
Consensus_aa:            M...................................p..p@t+lpEsl.lPsLh-lQhcS@pbhl.....
Consensus_ss:                                                    hhh  hhh    hhhhhhhhhhhhh    h
                                                     β-lobe 1
Conservation:                          96 6 6   6996      9   9 6  66 666     96  66 99 999 6
TTH                  36  PEKRENV-------GIQAAFRETFPIEEEDKGKGGLVLDFLEYRLGEPPFPQDECREKDLTYQAPLYARL   98
TAQ                  36  PEKRENV-------GIQAAFKETFPIEEGDKGKGGLVLDFLEYRIGDPPFSQDECREKDLTYQAPLYARL   98
ECO                  45  GQY-----------GLEAAFRSVFPIQSY---SGNSELQYVSYRLGEPVFDVQECQIRGVTYSAPLRVKL  100
SAU                  44  --------------GLIEMFRDISPIEDF---TGNLSLEFVDYRLGEPKYDLEESKNRDATYAAPLRVKV   96
BSU                  44  --------------GLREMFQDISPIEDF---TGNLSLEFIDYSLGEPKYPVEESKERDVTYSAPLRVKV   96
MSM                  62  RQAAIDRGEENPVGGLEEVLAELSPIEDF---SGSMSLSFSDPRFDEVKASVDECKDKDMTYAAPLFVTA  128
MTB                  71  RESAAERGDVNPVGGLEEVLYELSPIEDF---SGSMSLSFSDPRFDDVKAPVDECKDKDMTYAAPLFVTA  137
Consensus_aa:            ..............Glp.hh.-h.PIE-b...pGshsLpF.-.Rhs-s.hs.-Et+p+DhTY.APL.hph
Consensus_ss:            hhhh           hhhhhh     ee      eeeeeeeeeeee     hhhhhh    eeeeeeeee
                                   β-lobe 1                                   β-lobe 2
Conservation:             66          66 69 6 9669  9 996 969699966996999669996969
TTH                  99  QLIHKD-------TGLIKEDEVFLGHIPLMTEDGSFIINGADRVIVSQIHRSPGVYFTPD-----PARPG  156
TAQ                  99  QLIHKD-------TGLIKEDEVFLGHLPLMTEDGSFIINGADRVIVSQIHRSPGVYFTPD-----PARPG  156
ECO                 101  RLVIYEREAPEGTVKDIKEQEVYMGEIPLMTDNGTFVINGTERVIVSQLHRSPGVFFDSDKGKTHSSGKV  170
SAU                  97  RLIIKE-------TGEVKEQEVFMGDFPLMTDTGTFVINGAERVIVSQLVRSPSVYFNEKIDK--NGREN  157
BSU                  97  RLINKE-------TGEVKDQDVFMGDFPIMTDTGTFIINGAERVIVSQLVRSPSVYFSGKVDK--NGKKG  157
MSM                 129  EFINNN-------TGEIKSQTVFMGDFPMMTEKGTFIINGTERVVVSQLVRSPGVYFDETIDK--STEKT  189
MTB                 138  EFINNN-------TGEIKSQTVFMGDFPMMTEKGTFIINGTERVVVSQLVRSPGVYFDETIDK--STDKT  198
Consensus_aa:            phI.pp.......TGblKpppVFhGchPhMT-pGoFlINGh-RVlVSQlhRSPtVYFs.p.....ssc.s
Consensus_ss:            eeee           eeeeeeeeeeeeeee   eeeeee   eeeeeehhh   eeeee          e
                                                     β-lobe 2
Conservation:               6 69  9 99 966 6 6  6 6 66666 96 9 6 696696   66   66
TTH                 157  RYIASIIPLPKRGPWIDLEVEPNGVVSMKVNK-RKFPLVLLLRVLGYDQETLARELG-------------  212
TAQ                 157  RYIASIIPLPKRGPWIDLEVEASGVVTMKVNK-RKFPLVLLLRVLGYDQETLVRELS-------------  212
ECO                 171  LYNARII--PYRGSWLDFEFDPKDNLFVRIDRRRKLPATIILRALNYTTEQ-ILDLFFEKVIFEIRDNKL  237
SAU                 158  -YDATII--PNRGAWLEYETDAKDVVYVRIDRTRKLPLTVLLRALGFSSDQEIVDLL-------------  211
BSU                 158  -FTATVI--PNRGAWLEYETDAKDVVYVRIDRTRKLPVTVLLRALGFGSDQEILDLI-------------  211
MSM                 190  LHSVKVI--PGRGAWLEFDVDKRDTVGVRIDRKRRQPVTVLLKALGWTNEQ-IVERF-------------  243
MTB                 199  LHSVKVI--PSRGAWLEFDVDKRDTVGVRIDRKRRQPVTVLLKALGWTSEQ-IVERF-------------  252
Consensus_aa:            .@.hplI..P.RGsWl-h-h-.pshV.h+ls+.R+bPlhlLL+hLG@sp-p.l.-...............
Consensus_ss:            eeeeeeeee    eeeeeee    eeeeeeee ee  eeeeeee    hhhhhhhh
                                                     β-lobe 2
Conservation:
TTH                      ----------------------------------------------------------------------
TAQ                      ----------------------------------------------------------------------
ECO                 238  QMELVPERLRGETASFDIEANGKVYVEKGRRITARHIRQLEKDDVKLIEVPVEYIAGKVVAKDYIDESTG  307
SAU                      ----------------------------------------------------------------------
BSU                      ----------------------------------------------------------------------
MSM                      ----------------------------------------------------------------------
MTB                      ----------------------------------------------------------------------
Consensus_aa:            ......................................................................
Consensus_ss:
                                                     β-lobe 2
Conservation:                                                      66           96  66  6999699
TTH                 213  ----------------------------------AYGELVQGLMDESVFAMRPEEALIRLFTLLRPGDPP  248
TAQ                 213  ----------------------------------AYGDLVQGLLDEAVLAMRPEEAMVRLFTLLRPGDPP  248
ECO                 308  ELICAANMELSLDLLAKLSQSGHKRIETLFTNDLDHGPYISETLRVDPTNDR-LSALVEIYRMMRPGEPP  376
SAU                 212  ----------------------------------GDNEYLRNTLEKDGTENT-EQALLEIYERLRPGEPP  246
BSU                 212  ----------------------------------GENEYLRNTLDKDNTENS-DKALLEIYERLRPGEPP  246
MSM                 244  ----------------------------------GFSEIMMGTLEKDTTSGT-DEALLDIYRKLRPGEPP  278
MTB                 253  ----------------------------------GFSEIMRSTLEKDNTVGT-DEALLDIYRKLRPGEPP  287
Consensus_aa:            ..................................t.s-hhpshL-cssh..p.-pALlcl@p.LRPG-PP
Consensus_ss:                                                hhhhhhhhhh      hhhhhhhhhhhh
                         β-lobe
                           2
Conservation:            6 6 9  6   6 6 6 9999  69969 6 696             6     6    6
TTH                 249  KRDKAVAYVYGLIADPRRYDLGEAGRYKAEEKLGIR------LSGRTLARFEDGEF--KDEVFL------  304
TAQ                 249  KKDKALAYLFGLLADPKRYDLGEAGRYKAEEKLGVG------LSGRTLVRFEDGEF--KDEVFL------  304
ECO                 377  TREAAESLFENLFFSEDRYDLSAVGRMKFNRSL----------LREEI---EGSGILSKDDII-------  426
SAU                 247  TVENAKSLLYSRFFDPKRYDLASVGRYKTNKKLHLKHRLFNQKLAEPIVNTETGEIVVEEGTVLDRRKID  316
BSU                 247  TVENAKSLLDSRFFDPKRYDLANVGRYKINKKLHIKNRLFNQRLAETLVDPETGEILAEKGQILDRRTLD  316
MSM                 279  TKESAQTLLENLFFKEKRYDLARVGRYKVNKKLGLN-------AGKPI----TSSTLTEEDVV-------  330
MTB                 288  TKESAQTLLENLFFKEKRYDLARVGRYKVNKKLGLH-------VGEPI----TSSTLTEEDVV-------  339
Consensus_aa:            p.-pAbshlbs.hhc.+RYDLtphGRYKhpcKL.lp.......tcsl....stph..c-.hh........
Consensus_ss:             hhhhhhhhhhh          hhhhhhhhhhhhhh      hhhhhh           hhhhh
```

```
Conservation:                      6      6               969
TTH           305  -------------------PTLRYLFALT---------------AGVPGH--------------------  320
TAQ           305  -------------------PTLRYLFALT---------------AGVPGH--------------------  320
ECO           427  -------------------DVMKKLIDIR----------------------------------------  436
SAU           317  EIMDVLESNANSEVFELHGSVIDEPVEIQSIKVYVPNDDEGRT-TTVIGNAFPDSEVKCITPADIIASMS  385
BSU           317  KVLPYLENGIGFRKLYPNGGVVEDEVTLQSIKIFAPTDQEGEQVINVIGNAYIEEEIKNITPADIISSIS  386
MSM           331  -------------------ATIEYLVRLH----------EGQTSMTVPGGVEVPV---------------  356
MTB           340  -------------------ATIEYLVRLH----------EGQTTMTVPGGVEVPV---------------  365
Consensus_aa:      ...................shlcbbh.lp...............hsV.G.....................
Consensus_ss:                         hhhhhhhhhh                hh
```

β-lobe 1

```
Conservation:                 6999996999996969996  6969696 9 69 96996  66   6696 669 69  9
TTH           321  ----------EVDDIDHLGNRRIRTVGELMTDQFRVGLARLARGVRERMLMGSEDSLTPAKLVNSRPLEA  380
TAQ           321  ----------EVDDIDHLGNRRIRTVGELMADQFRVGLARLARGVRERMVMGSPDTLTPAKLVNSRPLEA  380
ECO           437  ------NGKGEVDDIDHLGNRRIRSVGEMAENQFRVGLVRVERAVKERLSLGDLDTLMPQDMINAKPISA  500
SAU           386  YFFNLLSGIGYTDDIDHLGNRRLRSVGELLQNQFRIGLSRMERVVRERMSIQDTESITPQQLINIRPVIA  455
BSU           387  YFFNLLHGVGDTDDIDHLGNRRLRSVGELLQNQFRIGLSRMERVVRERMSIQDTNTITPQQLINIRPVIA  456
MSM           357  ----------EVDDIDHFGNRRLRTVGELIQNQIRVGLSRMERVVRERMTTQDVEAITPQTLINIRPVVA  416
MTB           366  ----------ETDDIDHFGNRRLRTVGELIQNQIRVGMSRMERVVRERMTTQDVEAITPQTLINIRPVVA  425
Consensus_aa:      ..........-hDDIDHhGNRRlRoVGELhpsQhRlGLtRh.RsVRERMsh.s.-slTP.pLlN.RPl.A
Consensus_ss:                     hhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhh     hhh       hhh
```

β-lobe 1 | RRDR | Fork loop 2

```
Conservation:      6 69996 999999696 9996 669999699999999969999 699999 69999 996999996999
TTH           381  AIREFFSRSQLSQFKDETNPLSSLRHKRRISALGPGGLTRERAGFDVRDVHRTHYGRICPVETPEGANIG  450
TAQ           381  ALREFFSRSQLSQFKDETNPLSSLRHKRRISALGPGGLTRERAGFDVRDVHRTHYGRICPVETPEGANIG  450
ECO           501  AVKEFFGSSQLSQFMDQNNPLSEITHKRRISALGPGGLTRERAGFEVRDVHPTHYGRVCPIETPEGPNIG  570
SAU           456  SIKEFFGSSQLSQFMDQANPLAELTHKRRLSALGPGGLTRERAQMEVRDVHYSHYGRMCPIETPEGPNIG  525
BSU           457  SIKEFFGSSQLSQFMDQTNPLAELTHKRRLSALGPGGLTRERAGMEVRDVHYSHYGRMCPIETPEGPNIG  526
MSM           417  AIKEFFGTSQLSQFMDQNNPLSGLTHKRRLSALGPGGLSRERAGLEVRDVHPSHYGRMCPIETPEGPNIG  486
MTB           426  AIKEFFGTSQLSQFMDQNNPLSGLTHKRRLSALGPGGLSRERAGLEVRDVHPSHYGRMCPIETPEGPNIG  495
Consensus_aa:      tl+EFFtpSQLSQFbDpsNPLt.LpHKRRlSALGPGGLoRERAGh-VRDVH.oHYGRhCPlETPEGsNIG
Consensus_ss:      hhhh       eeee     hhhhhhhhh          eeeeee           eeee        ee
```

```
Conservation:      99 996 99 66  9966999969     6 66   6 9669 996 6669999    6     6     6  9
TTH           451  LITSLAAYARVDELGFIRTPYRRV--VGGVVTDEVVYMTATEEDRYTIAQANTPLEGN-RIAAERVVARR  517
TAQ           451  LITSLAAYARVDALGFIRTPYRRV--KNGVVTEEVVYMTASEEDRYTIAQANTPLEGD-RIATDRVVARR  517
ECO           571  LINSLSVYAQTNEYGFLETPYRKV--TDGVVTDEIHYLSAIEEGNYVIAQANSNLDEEGHFVEDLVTCRS  638
SAU           526  LINSLSSYARVNEFGFIETPYRKVDLDTHAITDQIDYLTADEEDSYVVAQANSKLDENGRFMDDEVVCRF  595
BSU           527  LINSLSSYAKVNRFGFIETPYRRVDPETGKVTGRIDYLTADEEDNYVVAQANARLDDEGAFIDDSIVARF  596
MSM           487  LIGSLSVYARVNPFGFIETPYRKV--ENGVVTDQIDYLTADEEDRHVVAQANSPTDENGRFTEDRVMVRK  554
MTB           496  LIGSLSVYARVNPFGFIETPYRKV--VDGVVSDGIVYLTADEEDRHVVAQANSPIDADGRFVEPRVLVRR  563
Consensus_aa:      LIsSLtsYA+Vs.hGFIcTPYR+V...sGhVT-plsYhTAsEEDp@hlAQANo.l-.p.+hhp-pVhhR.
Consensus_ss:      eehhhhhhhh          hhhe  eeeeee eeeee hhhh     eee        hh hh       eee
```

β-a10

```
Conservation:      6 96         66699996 9669    6   99999699699999969999699996      69 969969
TTH           518  K-GEPVIVSPEEVEFMDVSPKQVFSVNTNLIPFLEHDDANRALMGSNMQTQAVPLIRAQAPVVMTGLEER  586
TAQ           518  R-GEPVIVAPEEVEFMDVSPKQVFSLNTNLIPFLEHDDANRALMGSNMQTQAVPLIRAQAPVVMTGLEER  586
ECO           639  K-GESSLFSRDQVDYMDVSTQQVVSVGASLIPFLEHDDANRALMGANMQRQAVPTLRADKPLVGTGMERA  707
SAU           596  R-GNNTVMAKEKMDYMDVSPKQVVSAATACIPFLENDDSNRALMGANMQRQAVPLMNPEAPFVGTGMEHV  664
BSU           597  R-GENTVVSRNRVDYMDVSPKQVVSAATACIPFLENDDSNRALMGANMQRQAVPLMQPEAPFVGTGMEYV  665
MSM           555  KGGEVEFVSADQVDYMDVSPRQMVSVATAMIPFLEHDDANRALMGANMQRQAVPLVRSEAPLVGTGMELR  624
MTB           564  KAGEVEYVPSSEVDYMDVSPRQMVSVATAMIPFLEHDDANRALMGANMQRQAVPLVRSEAPLVGTGMELR  633
Consensus_aa:      +.GEs.hht.ppV-@MDVSP+QhhShsTshIPFLEpDDtNRALMGtNMQpQAVPLhpspAPhV.TGhE..
Consensus_ss:              ee   hheeeeeeeee   eee     eeeee            eeeee          eee           hhhh
```

```
Conservation:         966           9 6   9     6 6               6     6 6 66  999 696 69 9 9   9
TTH           587  VVRDSLAALYAEEDGEVAKVDGNRIVV--------RYEDGRLVEYPLRRFYRSNQGTALDQRPRVVVGQR  648
TAQ           587  VVRDSLAALYAEEDGEVVKVDGTRIAV--------RYEDGRLVEHPLRRYARSNQGTAFDQRPRVRVGQR  648
ECO           708  VAVDSGVTAVAKRGGVVQYVDASRIVI-KVNEDEMYPGEAGIDIYNLTKYTRSNQNTCINQMPCVSLGEP  776
SAU           665  AARDSGAAITAKHRGRVEHVESNEILVRRLVEENGVEHEGELDRYPLAKFKRSNSGTCYNQRPIVAVGDV  734
BSU           666  SGKDSGAAVICKHPGIVERVEAKNVWVRRYEEVDGQKVKGNLDKYSLLKFVRSNQGTCYNQRPIVSVGDE  735
MSM           625  AAIDAGDVVVADKTGVIEEVSADYITV--------MADDGTRQSYRLRKFARSNHGTCANQRPIVDAGQR  686
MTB           634  AAIDA-ATSSSQESGVIEEVSADYITV--------MHDNGTRRTYRMRKFARSNHGTCANQCPIVDAGDR  694
Consensus_aa:      hh.Dt.hhhhtccsG.l.cVpts.IhV..........cGp..pY.L.+@hRSNpGTthsQ.P.VshGp.
Consensus_ss:    hh      ee        eeeee   eeee          ee     eeeeee                  eeee h
```

β-a11 | β-flap

```
Conservation:      9   6  69999      966999 9   966966 96969966   96666   9   9699966 6   6999999
TTH           649  VRKGDLLADGPASENGFLALGQNVLVAIMPFDGYNFEDAIVISEELLKRDFYTSIHIERYEIEARDTKLG  718
TAQ           649  VKKGDLLADGPASEEGFLALGQNVLVAIMPFDGYNFEDAIVISEELLKRDFYTSIHIERYEIEARDTKLG  718
ECO           777  VERGDVLADGPSTDLGELALGQNMRVAFMPWNGYNFEDSILVSERVVQEDRFTTIHIQELACVSRDTKLG  846
SAU           735  VEYNEILADGPSMELGEMALGRNVVVGFMTWDGYNYEDAVIMSERLVKDDVYTSIHIEEYESEARDTKLG  804
BSU           736  VVKGEILADGPSMELGELALGRNVMVGFMTWDGYNYEDAIIMSERLVKDDVYTSIHIEEYESEARDTKLG  805
MSM           687  VEAGQVIADGPCTQNGEMALGKNLLVAIMPWEGHNYEDAIILSNRLVEEDVLTSIHIEEHEIDARDTKLG  756
MTB           695  VEAGQVIADGPCTDDGEMALGKNLLVAIMPWEGHNYEDAIILSNRLVEEDVLTSIHIEEHEIDARDTKLG  764
Consensus_aa:      Vc.GpllADGPt.-.GbhALGpNlhVthMs@-G@N@EDAIlhSpcLlccDhhTSIHIEc@E.-ARDTKLG
Consensus_ss:      h                        eeeeeee             eeeeeehhhhh    eeeeeeeeeeeeeee
```

```
                                       β-a12
                       β-flap                         β-flap tip
Conservation:           6969969996 66 69  996 966699999  9969996666996 6666996996699999669696
TTH              719    PERITRDIPHLSEAALRDLDEEGVVRIGAEVKPGDILVGRTSFKGESEPTPEERLLRSIFGEKARDVKDT   788
TAQ              719    PERITRDIPHLSEAALRDLDEEGIVRIGAEVKPGDILVGRTSFKGEQEPSPEERLLRSIFGEKARDVKDT   788
ECO              847    PEEITADIPNVGEAALSKLDESGIVYIGAEVTGGDILVGKVTPKGETQLTPEEKLLRAIFGEKASDVKDS   916
SAU              805    PEEITRDIPNVSESALKNLDDRGIVYIGAEVKDGDILVGKVTPKGVTELTAEERLLHAIFGEKAREVRDT   874
BSU              806    PEEITRDIPNVGEDALRNLDDRGIIRIGAEVKDGDLLVGKVTPKGVTELTAEERLLHAIFGEKAREVRDT   875
MSM              757    AEEITRDIPNVSDEVLADLDERGIVRIGAEVRDGDILVGKVTPKGETELTPEERLLRAIFGEKAREVRDT   826
MTB              765    AEEITRDIPNISDEVLADLDERGIVRIGAEVRDGDILVGKVTPKGETELTPEERLLRAIFGEKAREVRDT   834
Consensus_aa:           sEcITRDIPplt-.hL.sLD-cGIV.IGAEV+sGDILVG+ho.KG.oE.TsEERLL+tIFGEKAR-V+DT
Consensus_ss:             hh         hhhh                  eee    eeeeee          hhhhhhhh          eeee
                          β-a13
                                                   β-flap
Conservation:           99699 9 69 6     66 9
TTH              789    SLRVPPGEGGIVVRTVRLRRG-------------------------------------------------   809
TAQ              789    SLRVPPGEGGIVVGRLRLRRG-------------------------------------------------   809
ECO              917    SLRVPNGVSGTVIDVQVFTRDGVEKDKRALEIEEMQLKQAKKDLSEELQILEAGLFSRIRAVLVAGGVEA   986
SAU              875    SLRVPHGAGGIVLDVKVFNRE-------------------------------------------------   895
BSU              876    SLRVPHGGGGIIHDVKVFNRE-------------------------------------------------   896
MSM              827    SLKVPHGESGKVIGIRVFSRE-------------------------------------------------   847
MTB              835    SLKVPHGESGKVIGIRVFSRE-------------------------------------------------   855
Consensus_aa:           SL+VP.G.tGbVlsh..hpR..................................................
Consensus_ss:           eeee      eeeeeeeeee
                                                   β-flap
Conservation:                                                          6   9 699    9696 6
TTH              810    ---------------------------------------------------DPGVELKPGVREVVRVYVA   828
TAQ              810    ---------------------------------------------------DPGVELKPGVREVVRVFVA   828
ECO              987    EKLDKLPRDRWLELGLTDEEKQNQLEQLAEQYDELKHEFEKKLEAKRRKITQGDD-LAPGVLKIVKVYLA  1055
SAU              896    ---------------------------------------------------EGDDTLSPGVNQLVRVYIV   914
BSU              897    ---------------------------------------------------DGDE-LPPGVNQLVRVYIV   914
MSM              848    ---------------------------------------------------DDDE-LPAGVNELVRVYVA   865
MTB              856    ---------------------------------------------------DEDE-LPAGVNELVRVYVA   873
Consensus_aa:           ...................................................-ss..L.sGVpplVRVYlh
Consensus_ss:                                                                        eeeeeeeeee
                                                 β-a14
Conservation:           69966  99966699999996 99 9 9999 6 6969696 996699969996996969   6
TTH              829    QKRKLQVGDKLANRHGNKGVVAKILPVEDMPHLPDGTPVDVILNPLGVPSRMNLGQILETHLGLAGYF--   896
TAQ              829    QKRKLQVGDKLANRHGNKGVVAKILPVEDMPHLPDGTPVDVILNPLGVPSRMNLGQILETHLGLAGYF--   896
ECO             1056    VKRRIQPGDKMAGRHGNKGVISKINPIEDMPYDENGTPVDIVLNPLGVPSRMNIGQILETHLGMAAKGIG  1125
SAU              915    QKRKIHVGDKMCGRHGNKGVISKIVPEEDMPYLPDGRPIDIMLNPLGVPSRMNIGQVLELHLGMAAKN--   982
BSU              915    QKRKISEGDKMAGRHGNKGVISKILPEEDMPYLPDGTPIDIMLNPLGVPSRMNIGQVLELHMGMAARY--   982
MSM              866    QKRKISDGDKLAGRHGNKGVIGKILPVEDMPFLPDGTPVDIILNTHGVPRRMNIGQILETHLGWVAKA-G   934
MTB              874    QKRKISDGDKLAGRHGNKGVIGKILPVEDMPFLADGTPVDIILNTHGVPRRMNIGQILETHLGWCAHS-G   942
Consensus_aa:           QKRKlpsGDKhAsRHGNKGVltKIlP.EDMP@LsDGTPlDlhLNshGVPpRMNlGQlLEhHLGhtt....
Consensus_ss:           ee     hhhhhhhhh    eeeeeeeehhh         eeeeee          hhhhhhhhhhhhhhh

Conservation:                 6                                                 6969999 6 6
TTH              897    ------LGQR----------------------------------------------YISPIFDGAKEPEI   914
TAQ              897    ------LGQR----------------------------------------------YISPVFDGATEPEI   914
ECO             1126    DKINAMLKQQQEVAKLREFIQRAYDLGADVRQKVDLSTFSDEEVMRLAENLRKGMPIATPVFDGAKEAEI  1195
SAU              983    ------LGIH----------------------------------------------VASPVFDGANDDDV  1000
BSU              983    ------LGIH----------------------------------------------IASPVFDGAREEDV  1000
MSM              935    WNIDVAAGVPDWASKLPEELYSA----------------------------PADSTVATPVFDGAQEGEL   976
MTB              943    WKVDAAKGVPDWAARLPDELLEA----------------------------QPNAIVSTPVFDGAQEAEL   984
Consensus_aa:           ......hG................................................hhoPVFDGApEs-l
Consensus_ss:                    e                                              e        hhhh

Conservation:             66                                                   66  9699969966
TTH              915    KELLAQAFEVYFGKRKGEGFGVDKREVEVLRRAEKLGLVTPGKTPEEQLKELFLQGKVVLYDGRTGEPIE   984
TAQ              915    KELLAEAFNLYFGKRQGEGFGVDKREKEVLARAEKLGLVSPGKSPEEQLKELFDLGKVVLYDGRTGEPFE   984
ECO             1196    KELLK-------------------------------------------LGDLPTSGQIRLYDGRTGEQFE  1222
SAU             1001    WSTIE-------------------------------------------EAGMARDGKTVLYDGRTGEPFD  1027
BSU             1001    WETLE-------------------------------------------EAGMSRDAKTVLYDGRTGEPFD  1027
MSM              977    AGLLGSTL-----------------------------------PNRDGEVMVDADGKSTLFDGRSGEPFP  1011
MTB              985    QGLLSCTL-----------------------------------PNRDGDVLVDADGKAMLFDGRSGEPFP  1019
Consensus_aa:           b.hL............................................b..h..pGKhhL@DGRoGEPF.
Consensus_ss:           hhhhhhhhhh               hhhhhhhhhh          hhhhh       eeeee     eee

Conservation:            66 99 966 99 969699 999999699669999999999999999999699669999969999 99
TTH              985    GPIVVGQMFIMKLYHMVEDKMHARSTGPYSLITQQPLGGKAQFGGQRFGEMEVWALEAYGAAHTLQEMLT  1054
TAQ              985    GPIVVGQMFIMKLYHMVEDKMHARSTGPYSLITQQPLGGKAQFGGQRFGEMEVWALEAYGAAHTLQEMLT  1054
ECO             1223    RPVTVGYMYMLKLNHLVDDKMHARSTGSYSLVTQQPLGGKAQFGGQRFGEMEVWALEAYGAAYTLQEMLT  1292
SAU             1028    NRISVGVMYMLKLAHMVDDKLHARSTGPYSLVTQQPLGGKAQFGGQRFGEMEVWALEAYGAAYTLQEILT  1097
BSU             1028    NRVSVGIMYMIKLAHMVDDKLHARSTGPYSLVTQQPLGGKAQFGGQRFGEMEVWALEAYGAAYTLQEILT  1097
MSM             1012    YPVTVGYMYILKLHHLVDDKIHARSTGPYSMITQQPLGGKAQFGGQRFGEMECWAMQAYGAAYTLQELLT  1081
MTB             1020    YPVTVGYMYIMKLHHLVDDKIHARSTGPYSMITQQPLGGKAQFGGQRFGEMECWAMQAYGAAYTLQELLT  1089
Consensus_aa:           ..lsVGbM@hhKLhHhV-DKhHARSTGPYShlTQQPLGGKAQFGGQRFGEMEhWAhpAYGAA@TLQEhLT
Consensus_ss:           eeeee hhhhhhhhhhh   eeeee      ee                hhhhhhhhhh hhhhhhhhhh
```

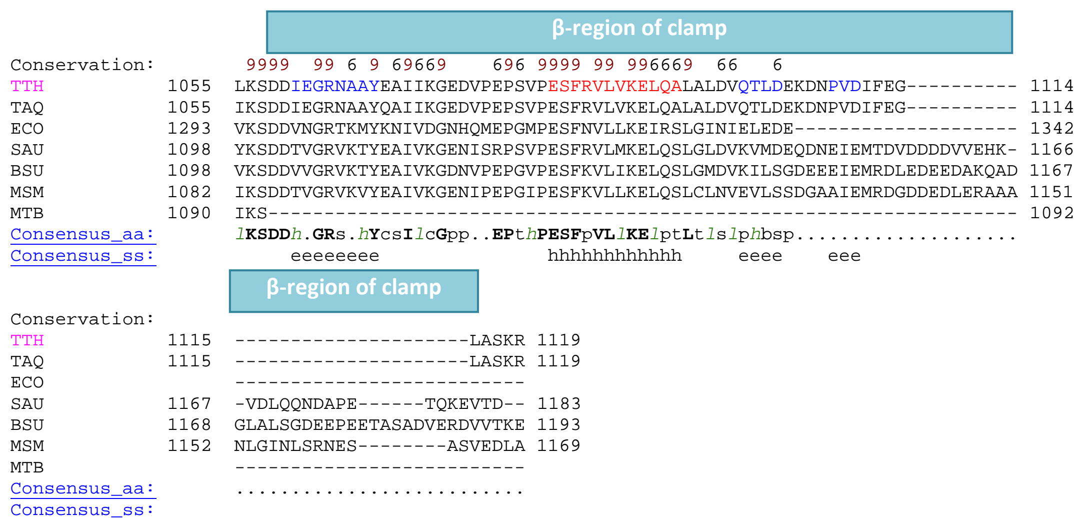


Figure 3.2 The structural domains of RpoB.

PROMALS3D (Pei *et al.*, 2008) was used to align the full *rpoB* protein between of MTB (AAA20242.2), MSM (YP_885753.1), *E.coli* (ECO) (NP_418414); *S. aureus* (SAU) (YP_003281434.1); *B. subtilis* (ZP_03589766.1); *T. aquaticus* (TAQ) (CAB65465) and *T. thermophilus* (TTH) (2A6E_M). In the PROMALS3D alignment, representative sequences are coloured in magenta, and residues forming alpha-helix are red and β-strands blue. The first and last residues of each block are numbered. Consensus predicted secondary structure symbols: alpha-helix: h; beta-strand: e. Structural domains of the β subunit of RNAP polymerase, determined using the *T. thermophilus* ternary elongation complex structure (Vassylyev *et al.*, 2007), were mapped as defined by Lane and Darst (2010b). Amino acid co-ordinates for each domain are as follows: β-lobe 1) Q22 – G125 and V336 – S392; β-lobe 2) R142 – A253; fork loop 2) P415 – E421 and D426 – R428; βa-10) V529 – D590; β-a11) F665 – K716; β-a12) A733 – K762; β-a13) D787 – V804; βa14) V823 – G894; β-flap) I705 – A828; flap tip) G763 – D784; flap tip helix) P769 – G779; β-clamp) D1058 – A1116 (continues in the β'); RRDR) F385 – G417. Residues that can interact with RIF are highlighted in **orange:** hydrogen bonding and **purple:** van der Waals interactions (Campbell *et al.*, 2001). Conservation indices above 4 are shown in the first line above each block.

3.4 Construction of suicide vector plasmids for allelic exchange

In this study, a cloning strategy was designed to introduce selected *rpoB* alleles (Table 3.1) into wild-type MSM using SDM to generate the specific point mutations, and allelic exchange for targeted mutagenesis. Although this strategy cannot eliminate the potential for second-site mutations, it does allow for the introduction of point mutations associated with RIF^{R} into the chromosome of MSM without the need to expose bacilli to RIF to select for mutation. The first step in this process involved the construction of suicide vectors containing the relevant mutant MSM *rpoB* allele. *E. coli* DH5α was used as the host organism for vector construction and manipulation, as described (see Materials and Methods 2.4.1).

3.4.1 Modification of p2NIL to remove *Acc*65I and *Bam*HI restriction sites

The core plasmid vector used in this study was p2NIL (Parish & Stoker, 2000). This vector carries an *aph*C gene which confers kanamycin resistance (KAN^R), thereby facilitating selection of *E. coli* and MSM transformants. Replication of p2NIL in *E. coli* begins at the o*ri*E site, and no mycobacterium-specific origin of replication is present. Therefore p2NIL is a suicide vector for mycobacteria. This vector also has a multiple cloning site, which includes a single *Pac*I site for convenient addition of the two-step selection cassette developed for use in mycobacteria by Parish and Stoker (Parish & Stoker, 2000). Naturally-occurring *Acc*65I and *Bam*HI restriction sites are located within the wild-type *rpoB* gene in MSM and were used for subsequent cloning (Figure 3.3). Therefore these sites were removed from the p2NIL backbone: that is, the vector was digested with *Acc*65I and the resulting 5’ overhangs filled in using the large (Klenow) fragment of *E. coli* DNA polymerase before re-circularization of the plasmid by ligation of the blunt ends. The *Bam*HI sites in p2NIL were removed in a similar process. Digestion and fill in of *Acc*65I and *Bam*HI sites, individually, resulted in a small increase (8 bp) in plasmid size. To confirm that both sites had been removed, an aliquot of the modified version of p2NIL – designated p2NILAB – was exposed to *Acc*65I and *Bam*HI. After overnight incubation, the plasmid remained undigested (Figure 3.3), verifying that these sites had been eliminated.

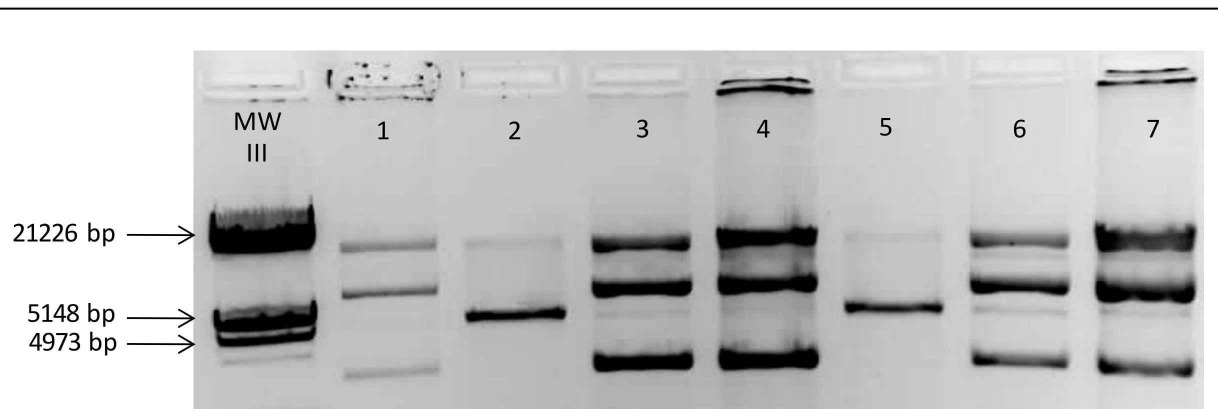


Figure 3.3 *Acc*65I and *Bam*HI sites were removed from p2NIL.

This figure represents two plasmids that were isolated from separate colonies. *lane 1*: standard p2NIL (undigested); *lane 2*: standard p2NIL (*Acc*65I digest); *lane 3*: p2NILAB from colony 2 (*Acc*65I digest); *lane 4*: p2NILAB from colony 17 (*Acc*65I digest); *lane 5:* standard p2NIL (B*am*HI digest); *lane 6:* p2NILAB from colony 2 (B*am*HI digest); p2NILAB from colony 17 (B*am*HI digest). The standard p2NIL vector is 4753bp; removal of *Acc*65I and *Bam*HI sites led to a size of 4761bp. In lanes 3 and 4, uncut DNA indicates that the putative p2NILAB vector was not cut with *Acc*65I. Similarly, in lanes 6 and 7, putative p2NILAB plasmid was not cut with *Bam*HI. This verified that the sites had been removed from p2NIL, to generate p2NILAB. MW III denotes molecular weight maker III (Roche, Germany). Fragments were visualised on a 1% agarose gel.

3.4.2 Construction of a template for SDM

A suitable template for SDM and suicide vector assembly was constructed by cloning wild-type MSM *rpoB* into p2NILAB. Primers were designed to amplify wild-type *rpoB* (Appendix 5.5) including 300 bp up-stream sequence thought to include the promoter (see Section 3.5.2). Artificial *Sca*I and *Hind*III restriction sites were included in the primer sequences to facilitate subsequent cloning into p2NILAB using the corresponding sites. The resulting 3893 bp PCR product was cloned into p2NILAB to generate p2*rpoB*. The integrity of the 8256 bp p2*rpoB* vector was established by restriction analysis and to confirm its suitability as a template for the subsequent construction of mutant *rpoB* alleles. Sequencing of the entire *rpoB* gene on the p2*rpoB* vector confirmed that no point mutations had been introduced into during cloning. Sequencing was of utmost importance, so that when SDM was applied, we were sure that the starting *rpoB* sequence was exactly the same as wild-type.

3.4.3 Introduction of RIF^R-associated point mutations: SDM

The p2*rpoB* construct was used as a template to facilitate introduction of relevant mutations into the *rpoB* gene. A PCR-based SDM strategy was employed to introduce five out of six mutations selected (Table 3.1), as described previously (Smith & Klugman, 1997). A schematic of SDM is presented in Figure 3.4. Briefly, this process involved two rounds of PCR. In the first round, a mutagenic primer was designed that contained the mutation of interest in the *rpoB* gene. This mutagenic "inner" primer was applied in a PCR reaction containing the p2*rpoB* template and a reverse "outer" primer to generate a PCR product that served as the "megaprimer" in the second PCR reaction, the remainder of the *rpoB* fragment was amplified (Figure 3.4), thereby generating a 1578 bp fragment (between *Acc*65I and *Bam*HI restriction sites). Given the length of the megaprimer used (530 bp – 600 bp, exact size depends on the mutation being introduced, see Materials and Methods 2.4.1.4).

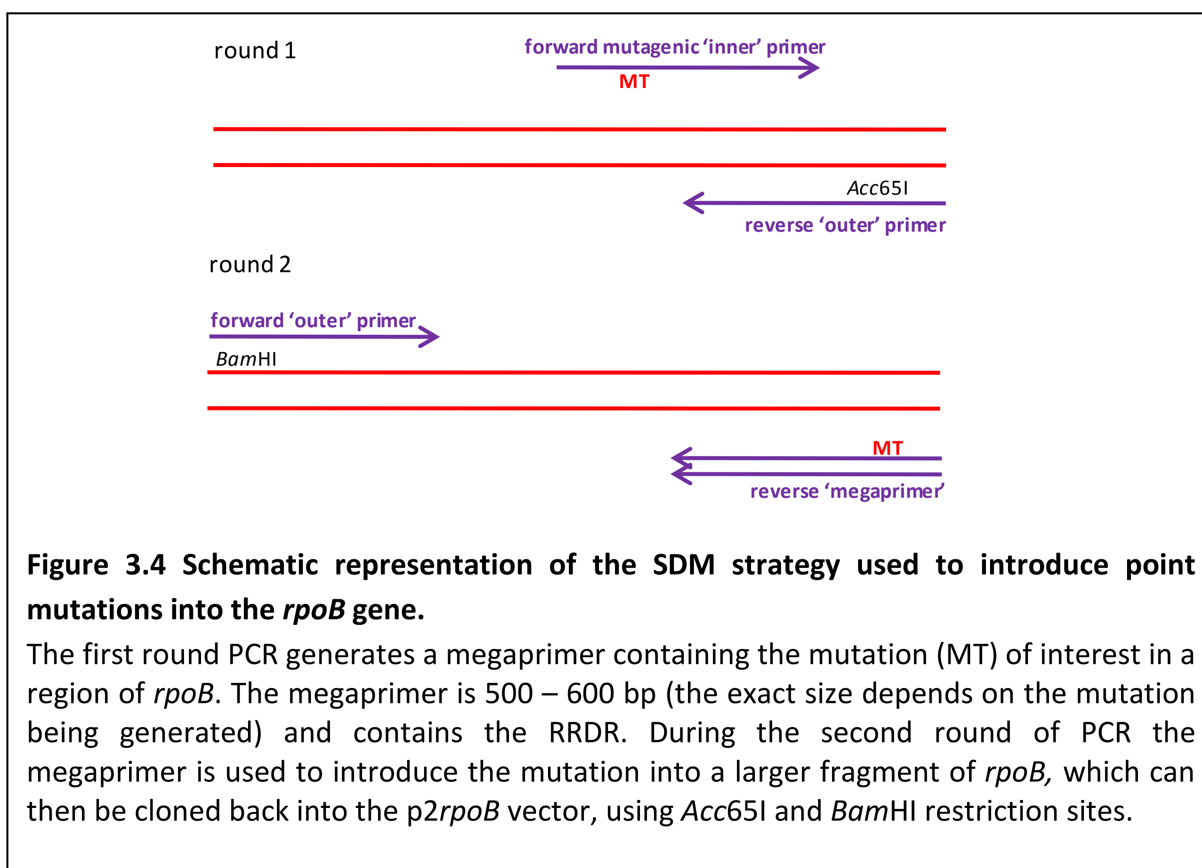


Figure 3.4 Schematic representation of the SDM strategy used to introduce point mutations into the *rpoB* gene.

The first round PCR generates a megaprimer containing the mutation (MT) of interest in a region of *rpoB*. The megaprimer is 500 – 600 bp (the exact size depends on the mutation being generated) and contains the RRDR. During the second round of PCR the megaprimer is used to introduce the mutation into a larger fragment of *rpoB,* which can then be cloned back into the p2*rpoB* vector, using *Acc*65I and *Bam*HI restriction sites.

3.4.4 Introduction of the 30 bp deletion: Xover-PCR

Five out of six constructs containing target RIF^R *rpoB* mutations were generated using the aforementioned SDM strategy. The outstanding construct containing a 30 bp deletion was generated via Xover-PCR (Table 3.1). MSM strains containing this mutation have been isolated at the MMRU (Digby Warner, unpublished results) and in independent studies (Malshetty *et al.*, 2010). MSM, via a recombination process, seems to be able to "loop out" this region of *rpoB* (Q513 – S522). This mutation is, therefore, is particularly interesting from a physiological point of view. The "megaprimer" method was not suited to production of this construct. Instead, Xover-PCR was employed, as previously described (Link *et al.*, 1997). This process involves two rounds of PCR, where products on either side of the targeted deletion are generated and used as a template in a subsequent round of PCR (Figure 3.5). In the first round of PCR, two separate reactions were set up. Reaction A was set up to amplify the region to the up-stream from the deletion (2237 bp). The reverse primer used in this reaction is unique in that it contains a 3' "tail" which is homologous to the 5' side of the region that is down-stream from the deletion. The use of this primer introduces a sequence, which is homologous to portions of the gene down-stream from the deletion. This facilitates homologous recombination during the second round of PCR. In reaction B, the region to the right of the deletion was generated using standard primers (1596 bp). Products from reactions A and B were used, in equimolar amounts, as templates in the final round of PCR (Figure 3.5). Two general primers were included in the reaction. Since the product of reaction A contains a region that is homologous to the matching end of the product of reaction B, the two products should recombine and the entire region should be amplified. It was important that in the final round of PCR, outer primers were excluded for the first 5 rounds of PCR. In this way products from reactions A and B were allowed to recombine, and were "filled in" by the DNA polymerase, and the full length product was augmented and, upon subsequent primer addition, amplified.

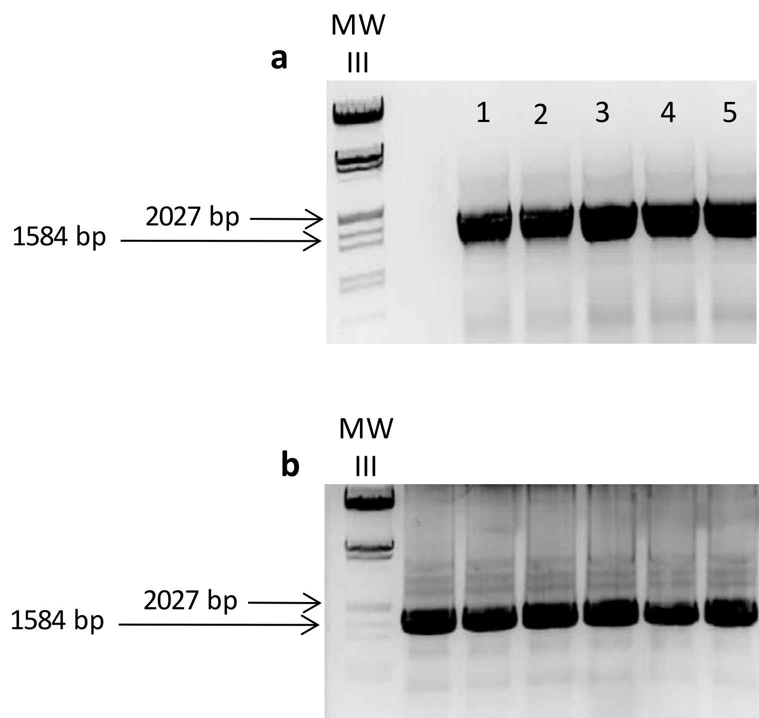


Figure 3.4(b) Gel images of final SDM and Xover PCR fragments generated cloning. The PCR products of the second round of SDM were separated on a 1% agarose gel. These fragments were purified and used for cloning into the modified p2NILAB vector. ***a)*** SDM fragments for single point mutations. Fragments containing specific mutations are in the following lanes: *lane 1*: S531L; *lane 2*: H526Y; *lane 3*: L521L; *lane 5*: Q513V *lane 4*: TH526TY ***b)*** Final Xover PCR fragments. All lanes represent this product. MW III denotes molecular weight marker III (Roche, Germany). Fragments were visualised on a 1% agarose gel.

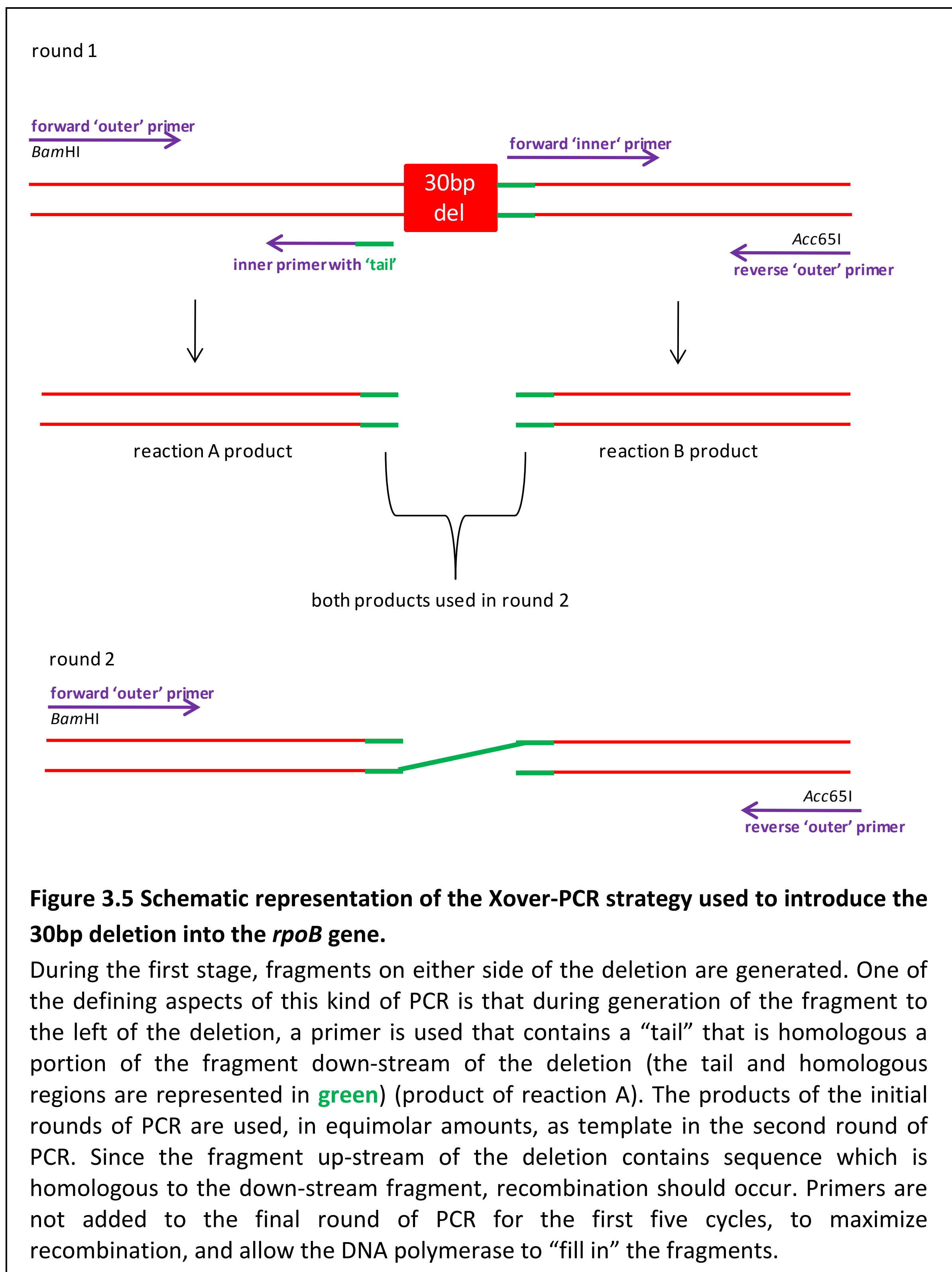


Figure 3.5 Schematic representation of the Xover-PCR strategy used to introduce the 30bp deletion into the *rpoB* gene.
During the first stage, fragments on either side of the deletion are generated. One of the defining aspects of this kind of PCR is that during generation of the fragment to the left of the deletion, a primer is used that contains a "tail" that is homologous a portion of the fragment down-stream of the deletion (the tail and homologous regions are represented in **green**) (product of reaction A). The products of the initial rounds of PCR are used, in equimolar amounts, as template in the second round of PCR. Since the fragment up-stream of the deletion contains sequence which is homologous to the down-stream fragment, recombination should occur. Primers are not added to the final round of PCR for the first five cycles, to maximize recombination, and allow the DNA polymerase to "fill in" the fragments.

After introduction of point mutations by SDM, and the 30 bp deletion by Xover-PCR, the resulting fragment was cloned into the p2*rpoB* template using *Acc*65I and *Bam*HI sites, in effect replacing the wild-type RRDR with a fragment containing the mutation of interest.

The successful introduction of the mutation was screened by AS-PCR and confirmed by sequencing the RRDR. The final allelic exchange construct was generated by cloning the *Pac*I cassette from (pGOAL19) into the corresponding site. Constructs therefore contained: the full length *rpoB* gene with a RIF^{R}-associated mutation (Table 3.1); a selection cassette containing *lac*Z, *sac*B, and the hygromycin resistance (HYG^{R}) gene; KAN resistance gene (KAN^{R}); and *E. coli* OriE (Figure 3.5). These constructs were designated p2*rpoB*^{MT}.

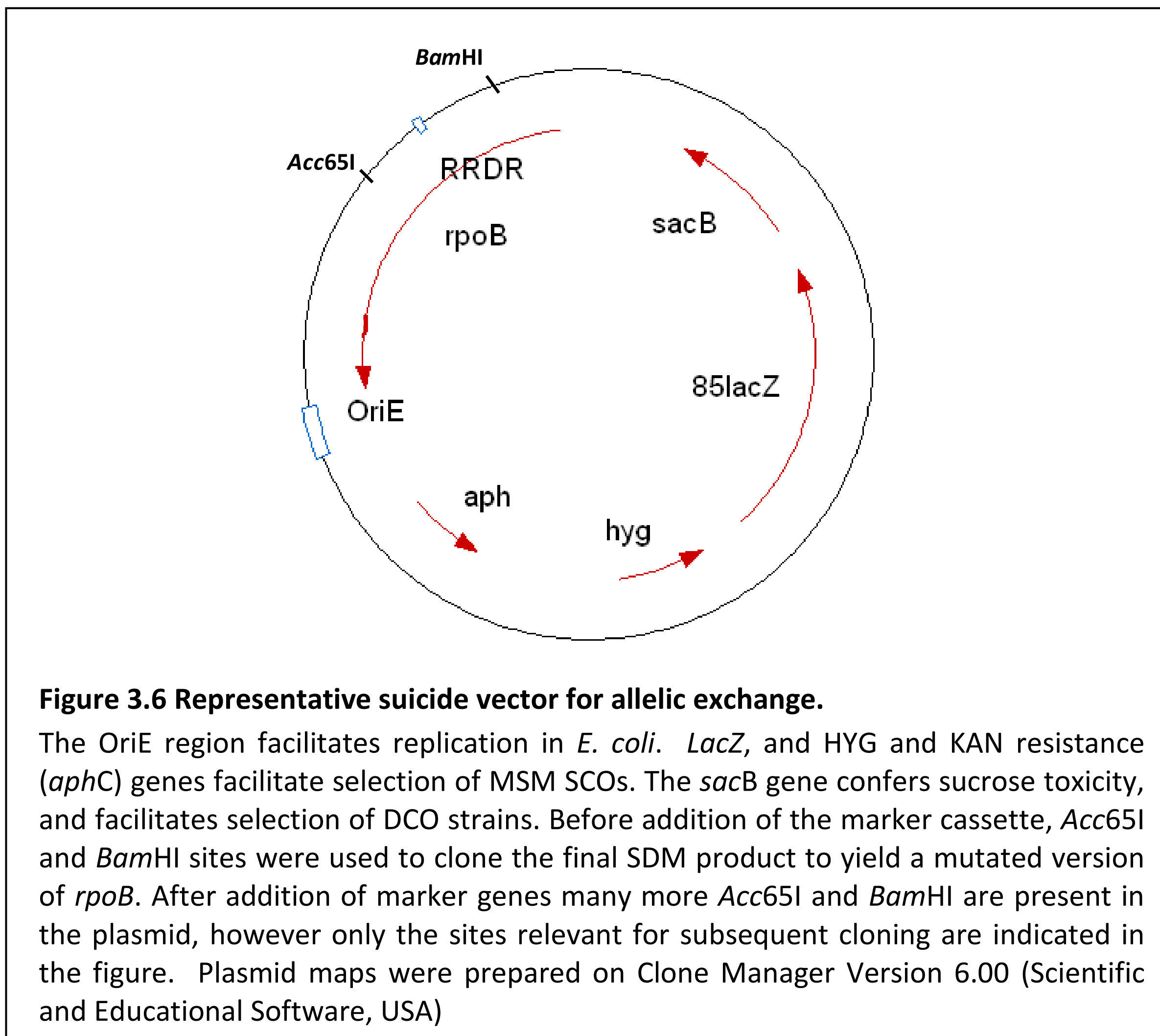


Figure 3.6 Representative suicide vector for allelic exchange.
The OriE region facilitates replication in *E. coli*. *LacZ*, and HYG and KAN resistance (*aph*C) genes facilitate selection of MSM SCOs. The *sac*B gene confers sucrose toxicity, and facilitates selection of DCO strains. Before addition of the marker cassette, *Acc*65I and *Bam*HI sites were used to clone the final SDM product to yield a mutated version of *rpoB*. After addition of marker genes many more *Acc*65I and *Bam*HI are present in the plasmid, however only the sites relevant for subsequent cloning are indicated in the figure. Plasmid maps were prepared on Clone Manager Version 6.00 (Scientific and Educational Software, USA)

3.4.5 AS-PCR to screen for mutations

Given that the desired outcome of the SDM process was the generation of a single point mutation that would differentiate the mutant allele from the wild-type, standard screening by restriction analysis could not be used to identify those constructs containing

the correct mutation. This limitation would apply similarly to any subsequent attempts to identify MSM transformants containing the mutant alleles, and so necessitated the optimization of an alternative screening tool. AS-PCR was also used to screen for the 30 bp deletion, as the sequence was different after removal of the nucleotides comprising the deletion.

AS-PCR has been utilized previously to detect RIF^R-associated mutations in clinical MTB isolates (Mokrousov *et al.*, 2003). Therefore, we investigated the applicability of this methodology as a screening tool, initially to identify desired mutant constructs, but subsequently to screen for successful MSM SCO and DCO mutants (see Section 3.5). To this end, primers were designed that would allow screening for mutants using AS-PCR (Appendix 5.5). In this modification of the standard PCR technique, three primers are used: two outer primers that bind to the sequence of interest allow for an internal control which establishes that the PCR reaction is working – in this case, a 1578 bp product should be generated which comprises the RRDR, and regions immediately up- and down-stream of the RRDR – and a third primer that is either specific for the wild-type sequence or the mutated allele at a specific site.

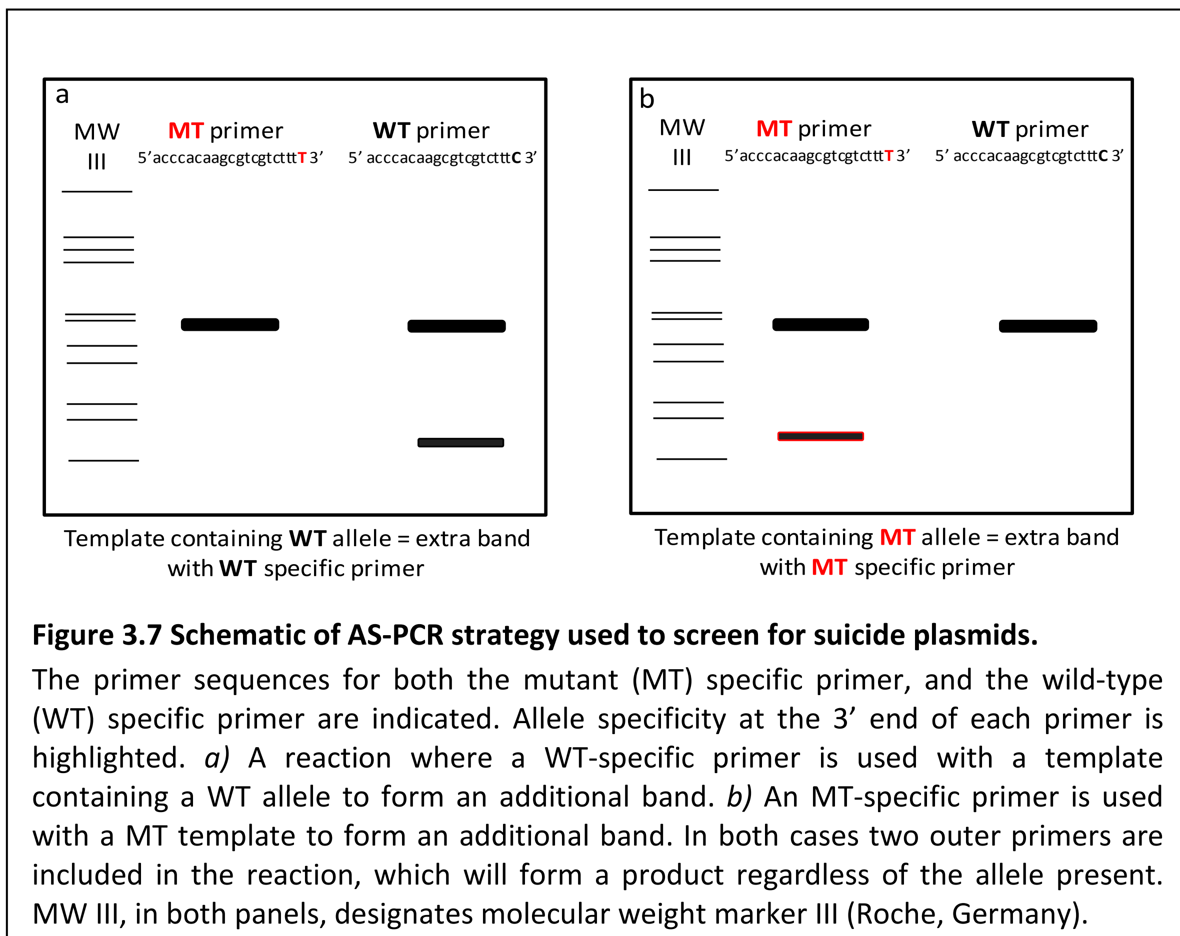


Figure 3.7 Schematic of AS-PCR strategy used to screen for suicide plasmids. The primer sequences for both the mutant (MT) specific primer, and the wild-type (WT) specific primer are indicated. Allele specificity at the 3' end of each primer is highlighted. *a)* A reaction where a WT-specific primer is used with a template containing a WT allele to form an additional band. *b)* An MT-specific primer is used with a MT template to form an additional band. In both cases two outer primers are included in the reaction, which will form a product regardless of the allele present. MW III, in both panels, designates molecular weight marker III (Roche, Germany).

The mutant-specific and wild-type-specific primers differ only by a single base pair at the 3' end (Figure 3.7). Therefore, the annealing temperature utilized for this PCR is critically important: if the annealing temperature is too low, then primers will bind non-specifically and will not differentiate between mutant and wild-type alleles; if too, high, no binding will occur and so no product will be generated. AS-PCR was successfully employed to screen for constructs containing desired point RIF^{R}-associated mutations. A gradient PCR was also performed to ensure that binding of AS-primers to the cognate fragments was not due to low annealing temperatures and therefore non-specific (Figure 3.8). The *rpoB* gene in plasmids that were confirmed via AS-PCR was sequenced to confirm that no unwanted mutations had been introduced during SDM.

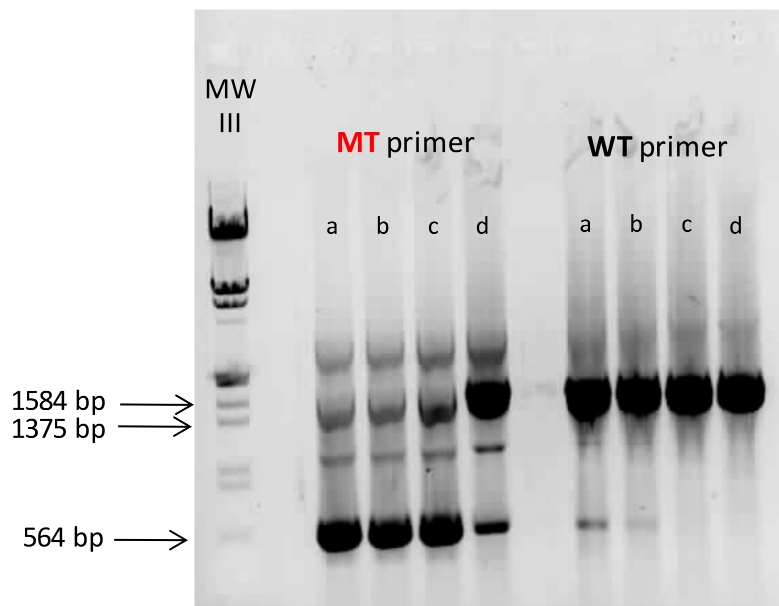


Figure 3.8 AS-PCR primers can detect the S531L mutation in a suicide construct.
Two outer primers, that bind regardless the presence of the mutation, produced a 1578 bp fragment. Reactions using the primer specific for the S531L mutant allele (MT primer) can be seen on the left. This primer yielded a smaller 546 bp band that was consistent with the presence of the mutant allele. On the right, a primer specific for the wild-type allele (WT primer) was used with the construct thought to contain the S531L mutation, and no extra bands were formed. Four different temperatures were used for the gradient PCR, and lanes wherein products for these temperatures are displayed as *a)* 55 °C *b)* 56.9 °C *c)* 61.1 °C *d)* 64.2 °C. Since the construct contained the mutation, the wild-type specific primer did not bind at higher temperatures and therefore no additional product was generated. Notably at the lowest temperature (55 °C), a band was formed with the wild-type primer – which illustrates the importance of annealing temperature for AS-PCR. MW III is molecular weight marker III (Roche, Germany). Fragments were visualised on a 1% agarose gel.

3.5 Allelic exchange to generate MSM strains containing RIF^R-associated mutations

Single point mutations in the RRDR are associated with high level RIF^R in clinical MTB isolates (Ramaswamy & Musser, 1998, Sandgren *et al.*, 2009). Therefore, it was important to develop a strategy that would allow introduction of these mutations into the chromosome of MSM, without any other point mutations, or the use of RIF to select for mutants. Such a strategy would allow for the evaluation of physiological implications of RIF^R-associated mutations – without any other alterations that may be induced by RIF selection. In addition the ability of these point mutations to confer high level RIF^R on their

own could be assessed. Allelic exchange methodology, which was pioneered in the MMRU in collaboration with a UK-based collaborator, Dr. Tanya Parish (Gordhan & Parish, 2001), has proved an invaluable technique for the genetic manipulation of mycobacteria including MSM and MTB and, importantly, allows the introduction of unmarked mutations into the mycobacterial chromosome (Gordhan & Parish, 2001, Parish & Stoker, 2000).

3.5.1 The first step of allelic exchange: Isolation of SCOs

The suicide vectors containing the specific *rpoB* mutations were delivered into MSM by electroporation. In a small proportion of bacilli, successful transformation is achieved which results in the site-specific incorporation of the suicide vector into the host chromosome by homologous recombination (Gordhan & Parish, 2001). The full process of allelic exchange, involves two selection steps, in which two distinct strains are generated, SCO and DCO mutants (described in Methods and Materials 2.4.2). The first of these are SCO mutants. As noted in Section 3.4.2, the suicide vector carries the *Pac*I selection cassette which comprises KAN^R and HYG^R genes and the *lac*Z gene. These enable the primary selection step, in which KAN- and HYG- resistant colonies that are blue on agar plates containing the chromogenic substrate, X-gal, are isolated. These SCOs contain two copies of the *rpoB* gene and are referred to as *rpoB*sco$^{WT/MT}$ (depending on the orientation, and mutation present in the strain). SCOs containing each of the relevant mutations (Table 3.1) were successfully generated.

3.5.2 Orientation of SCOs is important for *rpoB* expression

The orientation of SCO was important to determine if enough promoter sequence had been incorporated with the *rpoB* gene. During the first step of allelic exchange, a cross over occurs due to homologous recombination (Gordhan & Parish, 2001). If the cross over occurred anywhere up-stream of the site wherein the mutation was introduced, the mutated *rpoB* gene would be introduced up-stream of native *rpoB* (Figure 3.9a). This is referred to as an up-stream SCO. Had the cross over occurred down-stream of mutation site, the mutant *rpoB* gene would be incorporated down-stream of the native wild-type

rpoB gene, and a down-stream SCO would have been generated. Expression of the up-stream gene was driven by the native promoter present, regardless of whether it was the wild-type or mutated version of *rpoB*. Due to the cross over event and incorporation of vector sequence, the native promoter would not be present for expression of the down-stream gene. Promoter sequence would have had to be included on the suicide vector for the down-stream gene to have been expressed. As discussed in Section 3.6.1, a down-stream SCO containing an H526Y mutation (*rpoB*sco$^{WT/H526Y}$) was resistant to high levels of RIF, indicating that the down-stream gene is being expressed, and therefore that enough promoter sequence had been included in suicide vectors.

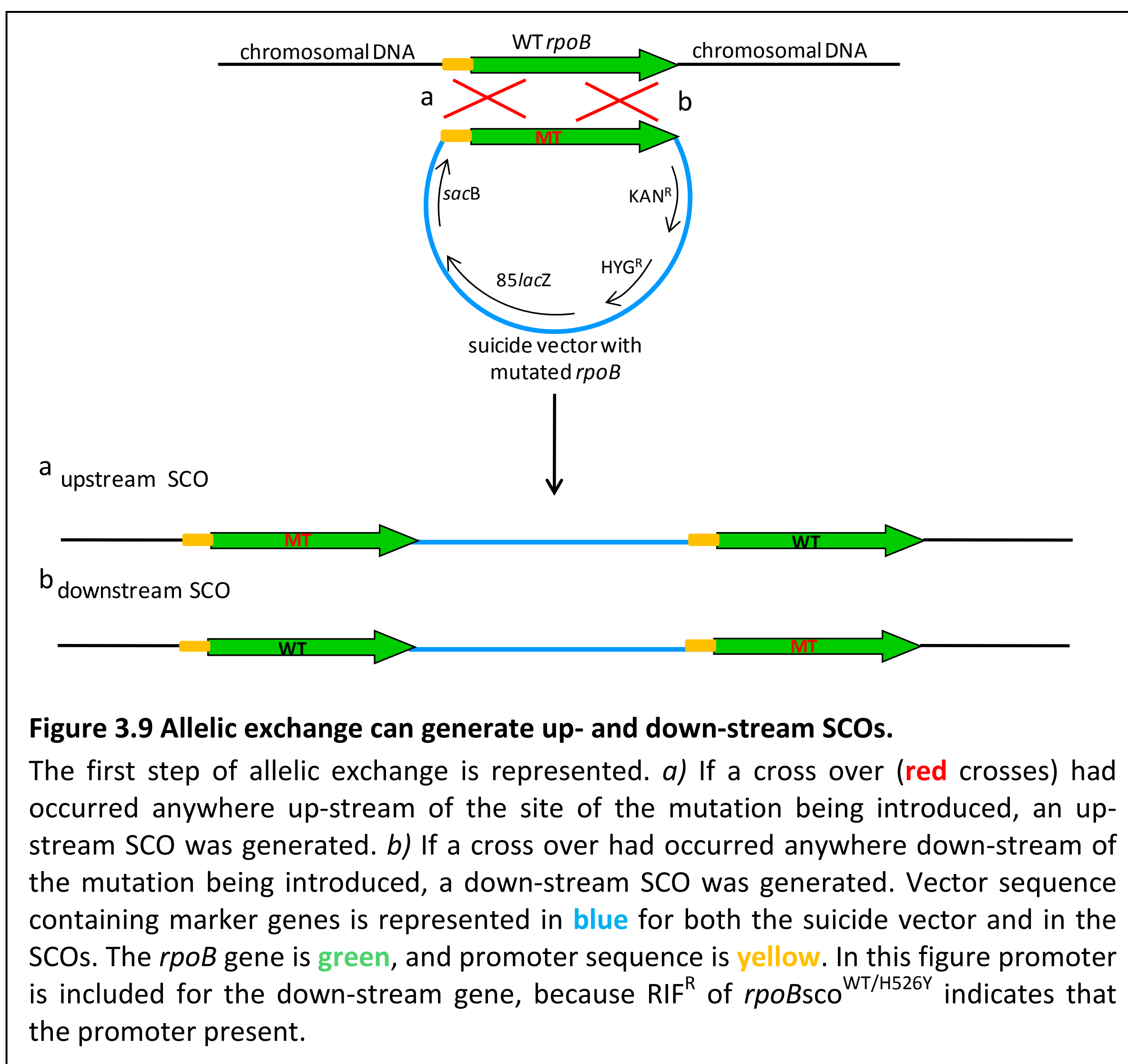


Figure 3.9 Allelic exchange can generate up- and down-stream SCOs.
The first step of allelic exchange is represented. *a)* If a cross over (**red** crosses) had occurred anywhere up-stream of the site of the mutation being introduced, an up-stream SCO was generated. *b)* If a cross over had occurred anywhere down-stream of the mutation being introduced, a down-stream SCO was generated. Vector sequence containing marker genes is represented in **blue** for both the suicide vector and in the SCOs. The *rpoB* gene is **green**, and promoter sequence is **yellow**. In this figure promoter is included for the down-stream gene, because RIF^R of *rpoB*sco$^{WT/H526Y}$ indicates that the promoter present.

To determine the orientation of the putative SCOs identified above (Section 3.5.1), a PCR-based strategy was designed. Primers were designed that were specific for either the up-stream or down-stream *rpoB* genes (Appendix 5.5). The forward primer that was specific for the up-stream gene was designed to bind chromosomal sequence that occurs up-stream of native *rpoB*; and the reverse primer was designed to bind chromosomal sequence down-stream of the native *rpoB* gene (Figure 3.10a). In this way primers [F1 + R1] should only amplify the up-stream *rpoB* gene, and primers [F2 + R2] should be specific for the down-stream gene, regardless of the origin of these alleles – that is, whether native wild-type or mutated (transformed) versions (Figure 3.10a). The product for primers [F1 + R1] was 3259 bp, while the primers [F2 + R2] yielded a product of 2770 bp (Figure 3.10b). Subsequently, AS-PCR was performed on these fragments to determine the nature of the amplified allele; that is, whether wild-type or mutated. The presence (or absence) of the mutation was confirmed by sequencing. Out of four S531L SCO isolates, two contained the mutated *rpoB* gene in up-stream orientation, and two in the down-stream orientation. Both H526Y SCO isolates contained the mutated *rpoB* gene in the down-stream orientation. Therefore, from our results, there appeared to be no significant bias towards either up-stream or down-stream insertion of the mutated *rpoB* gene for SCO strains examined.

At this point, it was decided that DCOs for only the S531L, H526Y and L521L mutations would be generated. The S531L and H526Y mutations are the most commonly observed mutations in clinical MTB isolates, and the L521L mutation is attractive as it has been reported to confer RIF^R (Yang *et al.*, 1998, Siddiqi *et al.*, 2002) but does not result in an amino acid change. Moreover, down-stream phenotypic characterisation was only conducted on these mutants, and in some cases just S531L mutants.

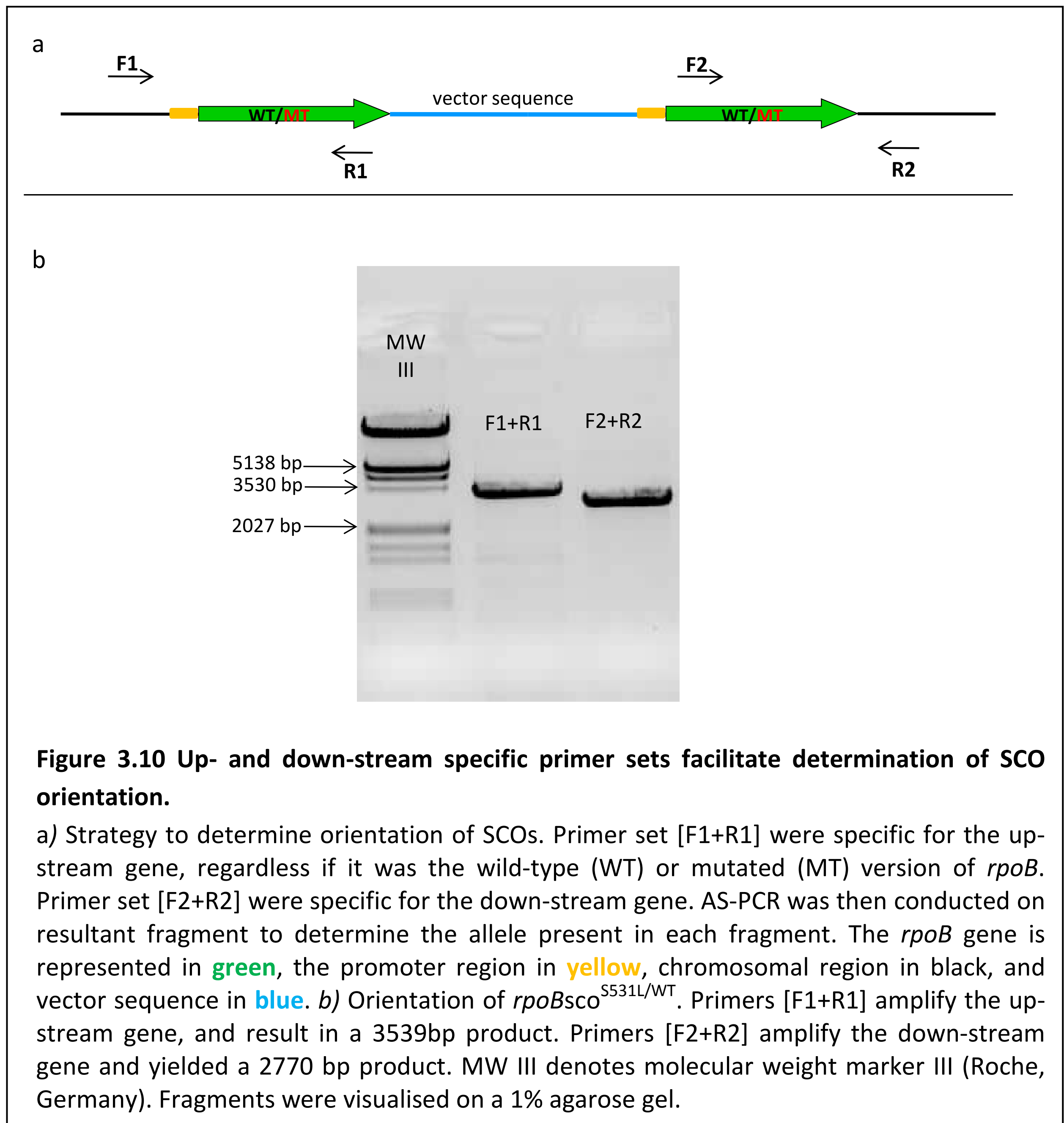


Figure 3.10 Up- and down-stream specific primer sets facilitate determination of SCO orientation.

a*)* Strategy to determine orientation of SCOs. Primer set [F1+R1] were specific for the up-stream gene, regardless if it was the wild-type (WT) or mutated (MT) version of *rpoB*. Primer set [F2+R2] were specific for the down-stream gene. AS-PCR was then conducted on resultant fragment to determine the allele present in each fragment. The *rpoB* gene is represented in **green**, the promoter region in **yellow**, chromosomal region in black, and vector sequence in **blue**. *b)* Orientation of *rpoB*sco$^{S531L/WT}$. Primers [F1+R1] amplify the up-stream gene, and result in a 3539bp product. Primers [F2+R2] amplify the down-stream gene and yielded a 2770 bp product. MW III denotes molecular weight marker III (Roche, Germany). Fragments were visualised on a 1% agarose gel.

3.5.3 The second step of allelic exchange: Generation of DCOs

SCO mutants contain two copies of the *rpoB* gene. Loss of one copy (preferably wild-type) requires a second step of allelic exchange, in which the SCOs are grown in liquid medium that does not contain any of the selective compounds. Allelic exchange depends on efficient homologous recombination and DNA replication therefore it is important that cells are given sufficient time and nutrients for cell division to occur (Gordhan & Parish, 2001). By removing the pressure of the selective antibiotic(s), a second cross over event occurs in which the marker cassette is effectively "looped out" of the chromosome

together with one allele. Successful DCO results in the retention of the mutated version of the *rpoB* gene in a strain referred to as a DCO mutant. However, cells may also exclude the marker cassette and the mutated *rpoB* gene, and revert to wild-type MSM. SCO cells grown in media without any selection compounds are likely to contain a mixed population of strains which include wild-type revertants, desired DCOs, and SCO strains which have not lost the marker cassette. The *Pac*I cassette also contains the s*ac*B gene, which confers sucrose toxicity. Therefore, to screen for DCOs, cultures of SCOs grown overnight in media without selective antibiotics were spread on solid media containing sucrose and X-gal. White colonies that were able to grow on these media were isolated, and DNA extracted using the colony boil protocol (see Materials and Methods 2.3.2.7).

Initially the AS-PCR strategy described in Section 3.4.4 was used to screen for mutations in DCO strains however it failed to detect the relevant mutation – even in control strains containing known *rpoB* mutations (control strains had been previously generated at the MMRU by UV exposure, Digby Warner, unpublished results). Optimisation of AS-PCR required several rounds of trial and error, and imposed a significant delay for DCO generation. The only difference between wild-type and mutant strains was a single point mutation. This would be unlikely to provide a unique restriction site and the mutation incorporated into the engineered rpoB gene would not change its size. Therefore Southern blotting would not be feasible to differentiate between wild-type and mutant strains.The final protocol for this version of AS-PCR required two rounds of PCR (see Materials and Methods: 2.4.2.5). During the first round, a small, 311 bp fragment, containing the RRDR, was isolated from the *rpoB* gene. The product was diluted (1/100) and used as the template in the second PCR reaction. Two separate reactions were then set up for each template. In each reaction a common reverse primer was used, however, the forward primer in each reaction was either specific for the wild-type allele or for the mutated version (Appendix 5.5*)*. The annealing temperature for both rounds of this PCR was increased from, 60 ^{o}C in the original AS-PCR, to 62 ^{o}C for this version, which greatly increased primer specificity. The smaller size of the template in the second reaction allowed for decreased annealing times (30 s for the original AS-PCR strategy vs. 15 s for this strategy), this minimised non-specific extension by FastStartTaq DNA Polymerase^{TM}

(Roche, Germany). Dilution to prevent excess of template also increased specificity of primers. Finally, for the second round of PCR, only 20 cycles were employed (whereas 30 cycles were employed for the original AS-PCR strategy), this was also found to minimise non-specific annealing.

If the correct primer was used with its cognate allele, a ± 170 bp product should be formed (the size of the band depends on the mutation being screened). Owing to the small size of these products, fragments were separated on a 2 % agarose gel, and the banding pattern indicated whether or not the mutation was present (Figure 3.11). DCOs containing the S531L mutation (Figure 3.11); H526Y; and L521L mutations were isolated in this way. In those cases where the AS-PCR indicated that there was a mutation present, the short fragment was sequenced to confirm the presence of the mutation, and DCO strains, designated *rpoB*^{MT}, were carried forward for phenotypic analysis.

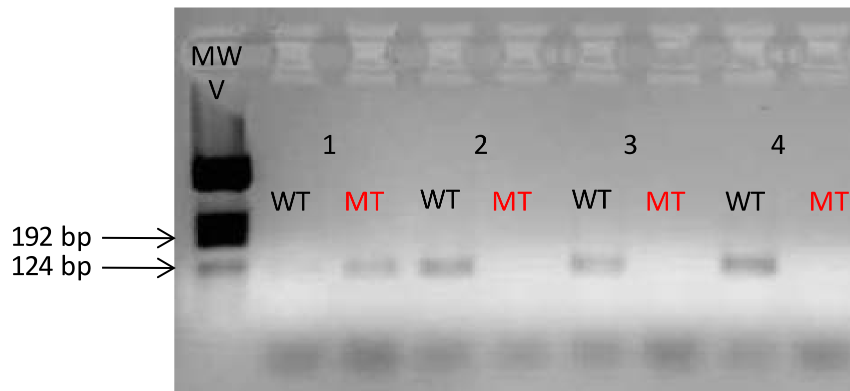


Figure 3.11 Colony one contains an S531L mutation in *rpoB* and is the *rpoB*^{S531L} DCO. A modified version of the AS-PCR was conducted to check if these colonies contained the S531L mutation. In contrast to the AS-PCR strategy to screen suicide plasmids, this version requires two rounds of PCR. Results from the second round of this PCR are shown, wherein an AS-forward primer and a general reverse primer are used. Products from reactions which contained wild-type specific primers are denoted by WT, while those which contain S531L-specific primers are indicated by MT. Results from four putative *rpoB*^{S531L} DCO colonies can be seen. Only colony one contains a mutant allele, revealed by the 109bp band formed by the S531L-specific primer (MT), and no band with the wild-type specific primer (WT). This was confirmed by sequencing. Colonies two – four contained a wild-type allele, and were wild-type revertants. Molecular weight marker V (Roche, Germany) is indicated by MW V. Fragments were visualised on a 1% agarose gel.

3.6 The impact of gene dosing for RIF^R

Approximately 95 % of RIF^R clinical MTB isolates contain mutations in a well-defined region of the *rpoB* gene (Ramaswamy & Musser, 1998). Structural analyses indicate that, of the 12 amino acids that surround the RIF binding pocket, 11 mutate in association with RIF^R (Campbell *et al.*, 2001). This provides significant support for necessity of these mutations for RIF^R. Molecular diagnostic assays have therefore been developed which utilize *rpoB* genotype as a proxy measure for RIF^R. However, to our knowledge, a RIF^R-associated *rpoB* mutation has not been introduced into the native site in chromosome of a susceptible mycobacterial strain – in the absence of RIF – and shown to confer high levels of resistance. Such a molecular investigation is essential to verify that RIF^R-associated mutations in *rpoB* are both necessary, and sufficient for RIF^R in mycobacteria. Therefore, while not the primary aim of this project, our ability to generate an *rpoB* point mutant in the absence of selection offered the opportunity of performing just such an investigation. That is, to compare RIF-susceptibility (RIF^S) of the parental, wild-type MSM with its derivative *rpoB* mutant strains.

3.6.1 RIF^S of merodiploid strains: SCOs containing the S531L or H526Y mutation are resistant to RIF

The key end-point of mutant generation was to isolate MSM mutants that were isogenic with the parental strain except for a single RIF^R-associated point mutation in the *rpoB* gene. While DCO generation was underway, the RIF^S of the SCO strains was evaluated. As noted above, these strains, contain both wild-type and mutated versions of the *rpoB* gene, and provided both genes are expressed, SCOs are essentially merodiploid. Some rare actinomycete species have two *rpoB* alleles and expression of the 'alternative' *rpoB* gene induces cryptic gene expression in *Streptomyces livdans* (Tala *et al.*, 2009). Furthermore, some of the nucleic acid differences between the two *S. lividans rpoB* alleles are analogous to those found in RIF^R isolates (Tala *et al.*, 2009).

If the mutated *rpoB* gene was in the up-stream orientation it would be expressed from the native promoter. However, if the mutated version of the *rpoB* gene was present in

the down-stream orientation, it would only be expressed if sufficient promoter sequence had been included in suicide vectors (see Section 3.5.2). Similar numbers of up- and down-stream SCO strains were generated. RIF^{S} of *rpoB*sco$^{S531L/WT}$ (an up-stream SCO); *rpoB*sco$^{WT/H526Y}$ (a down-stream SCO) and *rpoB*sco$^{L521L/WT}$ (up-stream SCO) were determined on solid media. Cells from these strains were grown to logarithmic phase, and a dilution series was prepared. Twenty microlitres from dilutions 10^{0} – 10^{-6} were plated on 7H10 media containing either 50 µg/ml; 100 µg/ml or 200 µg/ml (RIF_{50}; RIF_{100}; RIF_{200} respectively). To ensure that the strains had maintained SCO status, dilutions from each strain were also plated on 7H10 plates containing 20 µg/ml (KAN_{20}); 50 µg/ml (HYG_{50}) and X-gal. In all cases, blue colonies were produced on plates containing $KAN_{20}HYG_{50}$, confirming that these were still SCOs. The *rpoB*sco$^{S531L/WT}$ and *rpoB*sco$^{WT/H526Y}$ strains were able to grow on RIF_{200} containing solid media (Figure 3.12). In contrast, *rpoB*sco$^{L521L/WT}$ and wild-type MSM were not able to grow on any of the RIF concentrations tested. Decreased susceptibility to RIF of *rpoB*sco$^{S531L/WT}$ and *rpoB*sco$^{WT/H526Y}$ mutants but not *rpoB*sco$^{L521L/WT}$ provided the first indication that missense mutations on their own might confer RIF^{R}. Furthermore, decreased susceptibility of the *rpoB*sco$^{WT/H526Y}$ (a down-stream SCO) indicated that the mutated *rpoB* gene was being expressed when in the down-stream orientation, and provided evidence that the promoter region of *rpoB* is located in the 300 bp region up-stream of the *rpoB* gene.

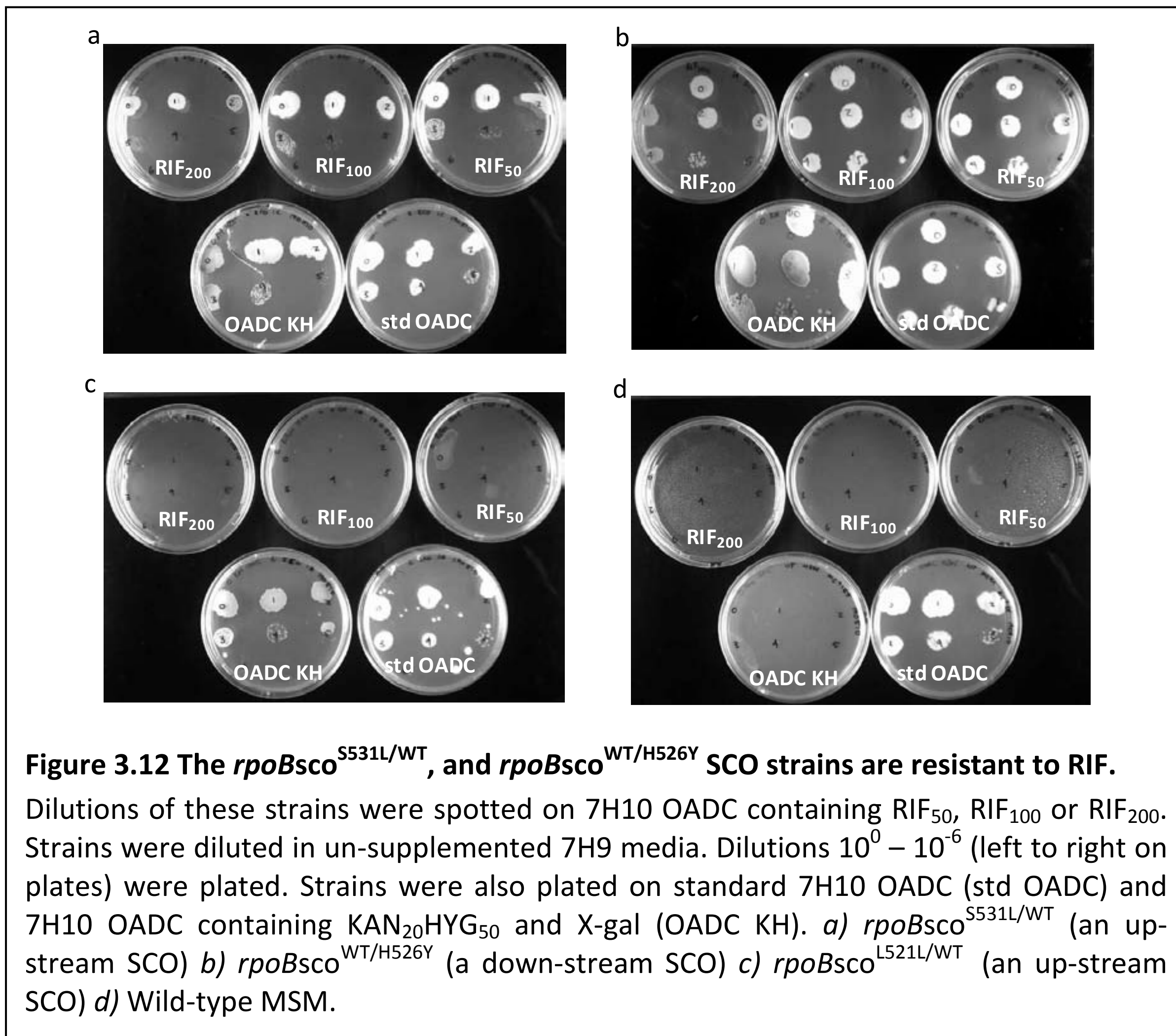


Figure 3.12 The *rpoB*sco$^{S531L/WT}$, and *rpoB*sco$^{WT/H526Y}$ SCO strains are resistant to RIF. Dilutions of these strains were spotted on 7H10 OADC containing RIF_{50}, RIF_{100} or RIF_{200}. Strains were diluted in un-supplemented 7H9 media. Dilutions 10^0 – 10^{-6} (left to right on plates) were plated. Strains were also plated on standard 7H10 OADC (std OADC) and 7H10 OADC containing $KAN_{20}HYG_{50}$ and X-gal (OADC KH). *a) rpoB*sco$^{S531L/WT}$ (an up-stream SCO) *b) rpoB*sco$^{WT/H526Y}$ (a down-stream SCO) *c) rpoB*sco$^{L521L/WT}$ (an up-stream SCO) *d)* Wild-type MSM.

3.6.2 The impact of gene dosing of *rpoB* on RIF^S: increased copy number of *rpoB* does not confer RIF^R

SCOs containing the S531L and H526Y mutations in one *rpoB* allele had significantly decreased susceptibility to RIF in comparison with the isogenic parental strain, an observation which suggested that these mutations might be sufficient for RIF^R. However, other studies have observed that for some antimicrobials over-expression of the target confers resistance (McMurry *et al.*, 1998, Kurth *et al.*, 2009). To determine whether increased expression of *rpoB* – or whether the presence of two alleles in the merodiploid strain – might result in increased RIF^R in the S531L and H526Y SCOs, *rpoB* was introduced into MSM in additional copies. An MSM strain containing two copies of the wild-type *rpoB* gene provided a specific control for determination of SCO susceptibility phenotypes described in Section 3.6.1. If increased expression of *rpoB* was contributing to RIF^R, these strains should have similar RIF^S to *rpoB*sco$^{S531L/WT}$ and *rpoB*sco$^{WT/H526Y}$ strains. The MSM

strain containing two copies of *rpoB* was generated via allelic exchange, using a suicide plasmid containing the wild-type *rpoB* gene. The implications of *rpoB* copy number were also evaluated by introducing both wild-type and mutated versions on a multicopy episomal (pOLYG) expression vector.

3.6.3 SCOs containing two copies of the wild-type *rpoB* gene are susceptible to RIF

First, to determine whether the presence of two copies of the *rpoB* gene (whether mutated or not) in the chromosome conferred RIF^R, a suicide vector containing wild-type *rpoB* was introduced into the MSM chromosome via electroporation as previously described (see Section 3.5.1). The resulting SCO mutant contained two copies of wild-type *rpoB* and was designated *rpoB*sco$^{WT/WT}$. As shown in Figure 3.13, the *rpoB*sco$^{WT/WT}$ mutant was as susceptible to RIF treatment as wild-type MSM. These data suggested that the possession of two copies of wild-type *rpoB* was not able to confer RIF^R but that, at the very least, one allele must contain a point mutation.

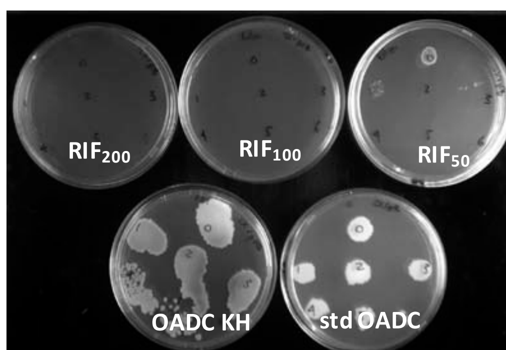


Figure 3.13 The *rpoB*sco$^{WT/WT}$ mutant is as susceptible to RIF as wild-type MSM. Dilutions of *rpoB*sco$^{WT/WT}$ were spotted on 7H10 OADC RIF_{50}, RIF_{100} or RIF_{200}. Strains were diluted in un-supplemented 7H9 media. Dilutions $10^0 – 10^{-6}$ (left to right on plates) were plated. The strains was also plated on standard 7H10 OADC (std OADC) and 7H10 OADC containing $KAN_{20}HYG_{50}$ and X-gal (OADC KH). Growth on $KAN_{20}HYG_{50}$ plates confirmed that *rpoB*sco$^{WT/WT}$ retained marker genes and remained SCO.

3.6.4 Expression of *rpoB* on pOLYG – a multicopy episomal plasmid

To examine whether the presence *rpoB* in multiple copies conferred resistance to RIF in MSM, the *rpoB* gene was cloned onto the multicopy episomal plasmid, pOLYG (O'Gaora

et al., 1997). Since SCO strains containing the S531L or H526Y mutations had exhibited reduced susceptibility to RIF, pOLYG constructs containing these alleles were also generated. These constructs – designated pG*rpoB*^{WT}, pG*rpoB*^{S531L}, and pG*rpoB*^{H526Y} – were introduced into MSM via electroporation. Surprisingly, very low transformation efficiencies were observed (Figure 3.14 and Table 3.2)

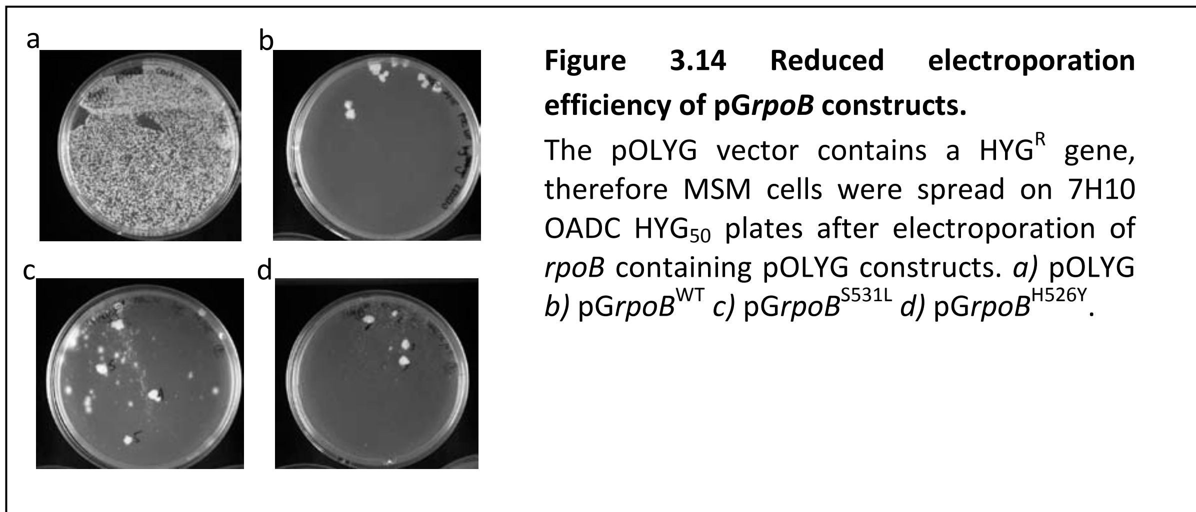


Figure 3.14 Reduced electroporation efficiency of pG*rpoB* constructs.
The pOLYG vector contains a HYG^{R} gene, therefore MSM cells were spread on 7H10 OADC HYG_{50} plates after electroporation of *rpoB* containing pOLYG constructs. *a)* pOLYG *b)* pG*rpoB*^{WT} *c)* pG*rpoB*^{S531L} *d)* pG*rpoB*^{H526Y}.

Table 3.2 Transformation efficiency of pG*rpoB* constructs[a]

Construct	**Electroporation efficiency (CFU/µg DNA)**
pOLYG	1 X 10^{4}
pG*rpoB*^{WT}	9
pG*rpoB*^{S531L}	4
pG*rpoB*^{H526Y}	5

[a]These experiments were conducted three times, however results from one experiment are represented in the table

The low transformation efficiency of all three *rpoB*-containing pOLYG constructs suggested that the increased expression of *rpoB* was toxic in MSM (Figure 3.14). Table 3.2 describes the relative electroporation efficiencies of pG*rpoB* constructs in comparison to pOLYG alone, in a representative electroporation. In this experiment, standard pOLYG was associated with an electroporation efficiency of 1 X 10^{4} CFU/µg DNA which is similar to published results (Warner *et al.*, 2006). In contrast, the pG*rpoB*^{WT} construct had a transformation efficiency of just 9 CFU/µg DNA. The pOLYG vector exists in the cell in multiple copies, and although this is likely to have increased expression of *rpoB*,

expression analysis such as RT-PCR is required to formally show over-expression. Given the low number of HYG^R transformants recovered from electroporations, and the resultant re-arrangement of plasmids within these strains (see Section 3.6.4), this kind of analysis was not possible. For the small number of transformed colonies that were obtained, RIF^S was evaluated by plating on solid media containing increasing RIF concentrations. Those MSM strains that were recovered containing either wild-type, S531L, or H526Y versions of *rpoB* on pOLYG did not grow on any RIF concentrations tested (Figure 3.15).

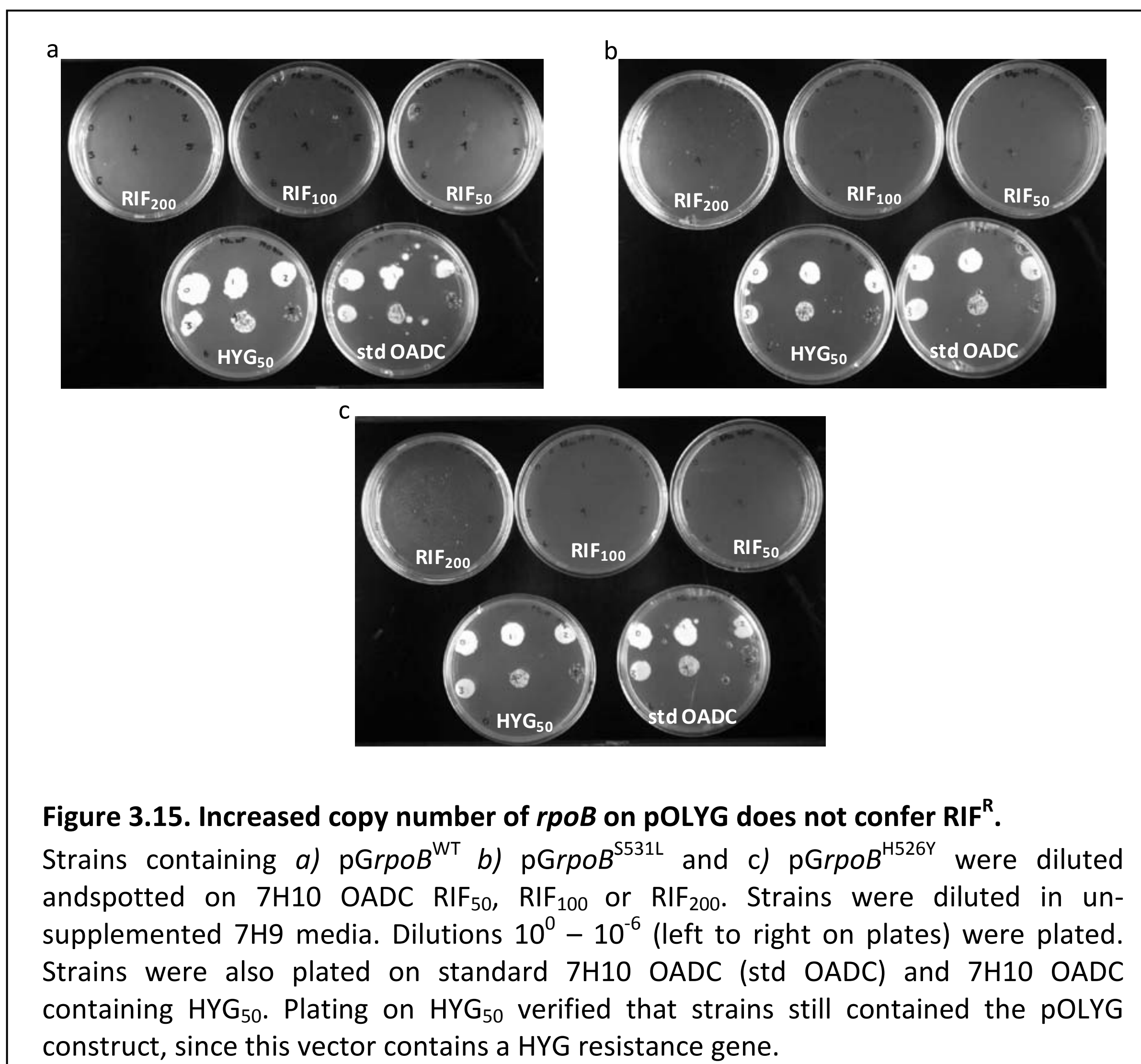
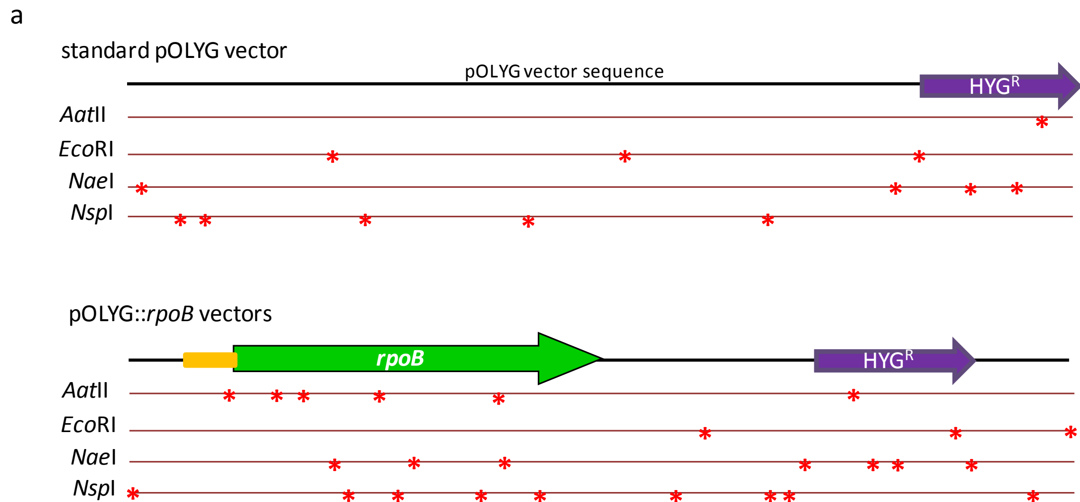


Figure 3.15. Increased copy number of *rpoB* on pOLYG does not confer RIF^R.
Strains containing *a)* pG*rpoB*^{WT} *b)* pG*rpoB*^{S531L} and c*)* pG*rpoB*^{H526Y} were diluted andspotted on 7H10 OADC RIF_{50}, RIF_{100} or RIF_{200}. Strains were diluted in un-supplemented 7H9 media. Dilutions 10^0 – 10^{-6} (left to right on plates) were plated. Strains were also plated on standard 7H10 OADC (std OADC) and 7H10 OADC containing HYG_{50}. Plating on HYG_{50} verified that strains still contained the pOLYG construct, since this vector contains a HYG resistance gene.

3.6.5 Expression of *rpoB* from pOLYG is toxic in mycobacteria

Previous studies have reported wholesale loss or re-arrangement of plasmids associated with toxicity, in mycobacteria (Stallings *et al.*, 2009). Therefore, to investigate the apparent toxicity of increased copy number of *rpoB*, we evaluated whether plasmid rearrangements had occurred in those MSM transformants that were recovered. Plasmids were recovered from MSM transformants by electroduction into *E.coli* (see Materials and Methods: 2.4.3.1). Plasmid DNA was isolated from *E. coli* and the integrity of the different *rpoB* containing pOLYG constructs (pG*rpoB*^{WT}; pG*rpoB*^{S531L}and pG*rpoB*^{H526Y}) evaluated by restriction analysis. Three colonies for each type of pOLYG construct was analysed for re-arrangement. Restriction enzymes were selected to ensure maximum coverage of the construct based on the sequence of the respective original vectors (Figure 3.16a). Results from this analysis indicated that significant rearrangement of plasmids had occurred upon introduction into MSM (Figure 3.16b). For example, digestion of the pG*rpoB*^{S531L} construct with *Eco*RI before electroporation yielded the expected 5561 bp, 2306 bp, and 1515 bp fragments. In contrast, the same construct yielded a significantly different restriction pattern (one fragment less than 3500 bp, one ± 2000 bp fragment, and one ± 1400 bp fragment) when recovered from a single MSM transformant (Figure 3.16). Similar evidence of large-scale rearrangement was observed with other restriction enzymes. Moreover, pG*rpoB*^{WT} and pG*rpoB*^{H526Y} constructs had also undergone significant rearrangement. This is consistent with a previous observation of significant instability of self-replicating plasmids expressing *rpoB* in MTB (Zaczek *et al.*, 2009).



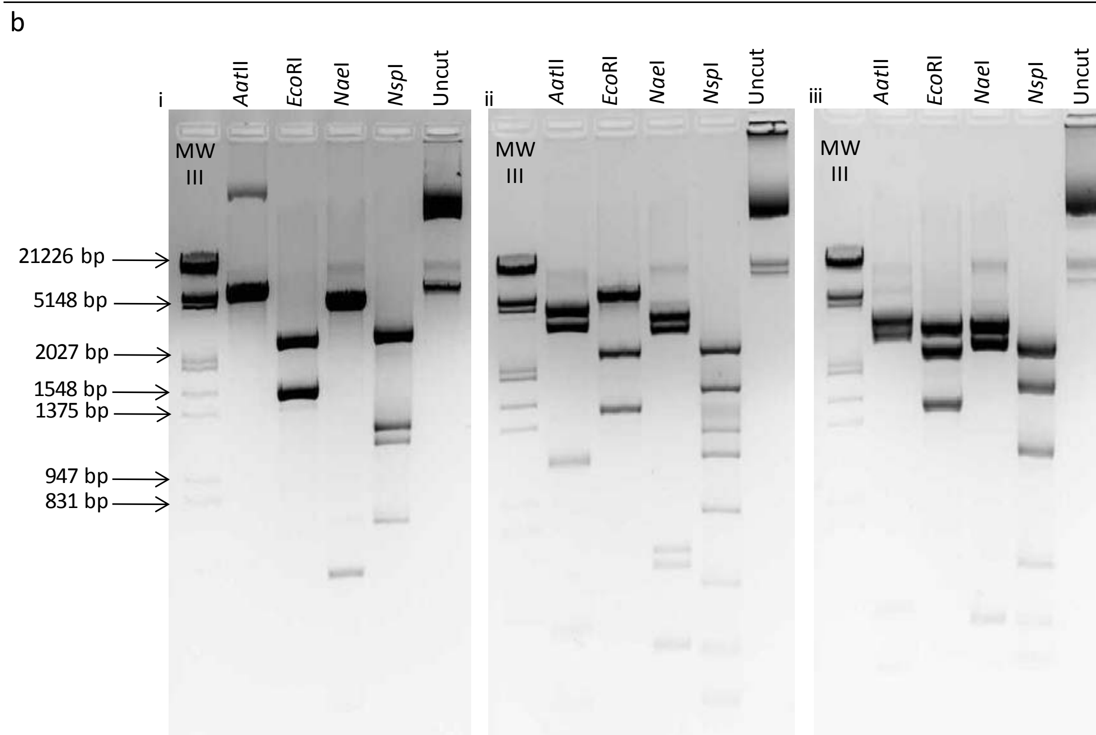


Figure 3.16 Over-expression of *rpoB* leads to plasmid re-arrangement.

Restriction sites for *Aat*II, *Eco*RI, *Nae*I and *Nsp*I on standard pOLYG and pG*rpoB* constructs. The HYG resistance gene is indicated in purple, the *rpoB* gene is indicated in green, and the promoter region in yellow. Cleavage sites are represented by *. This map is not drawn to scale, but rather intended to give some indication of regions cleaved by each enzyme. *b)* Cleavage pattern of pG*rpoB*^{S531L} constructs recovered after electroporation and digested with *Aat*II, *Eco*RI, *Nae*I and *Nsp*I. *i)* Standard pOLYG *ii)* pG*rpoB*^{S531L} before electroporation into MSM *iii)* pG*rpoB*^{S531L}. MW III is molecular weight marker III from Roche (Germany). Fragments were visualised on a 1 % agarose gel.

3.7 RIF^S of MSM mutants containing single RIF^R-associated mutations

The results of the RIF^S assay of the SCO strains confirmed the importance of *rpoB* mutations for RIF^R. Isolation of DCO mutants, which are isogenic to the parental strain except for a single RIF^R-associated *rpoB* mutation, would allow us to assess whether these mutations are sufficient for high-level RIF^R. In addition, DCO mutants would provide the means to determine the impact of *rpoB* mutations on MSM physiology.

3.7.1 Mutants containing S531L or H526Y mutations are able to grow on solid media containing relatively high concentrations of RIF

DCOs were generated which contained either the S531L, H526Y, or L521L mutation ($rpoB^{S531L}$; $rpoB^{H526Y}$; and $rpoB^{L521L}$ respectively). The RIF^S of resultant DCO strains was assayed on RIF containing solid media using the same methodology as described for RIF^S determination of SCO strains. Wild-type and SCO mutants were included in this assay (Figure 3.17). In addition, an S531L mutant generated previously by *in vitro* RIF selection following UV irradiation ($rpoB^{UV_S531L}$; Digby Warner, unpublished results) was included to provide an initial indication of whether the use of RIF selection to generate RIF^R mutants was associated with additional chromosomal or physiological changes that might impact RIF^S.

Cultures of each strain were grown to logarithmic phase and a dilution series was prepared in 7H10 OADC. Twenty microlitres from dilutions 10^0 – 10^{-6} were plated on 7H10 media containing either RIF_{50}; RIF_{100} or RIF_{200}. Dilutions from each strain were also plated on 7H10 plates containing KAN_{20}; $HYG_{50}$and X-gal to ensure that DCOs were not mistaken for SCOs and to confirm SCOs maintained their SCO status. Only SCO strains grew on $KAN_{20}HYG_{50}$ containing plates (Figure 3.17), confirming that AS-PCR to screen for DCOs had not mistakenly detected a mutation in an SCO. The L521L DCO strain did not grow on media containing any of the concentrations of RIF tested, but displayed significant growth on standard, antibiotic-free media. Similarly, wild-type MSM did not grow on plates containing RIF. In contrast DCO strains containing an S531L or and H526Y mutation grew on RIF_{50} OADC plates until a dilution of 10^{-4}, and moreover displayed

growth until a dilution of 10^{-3} on RIF_{200} plates. The *rpoB*^{UV_S531L} mutant displayed similar levels of growth on RIF containing plates, which was to be expected. However growth of *rpoB*^{S531L} and *rpoB*^{H526Y} strains on RIF concentrations of up to 200 µg/ml was a significant result: DCO mutants were generated in the absence of RIF, to minimise the possibility of introducing second site mutations, therefore the ability of MSM strains containing a single point mutation in the *rpoB* gene to withstand high concentrations of RIF provides significant evidence for the importance of these mutations for RIF^R. Initially growth up to RIF_{200} was evaluated. To obtain some indication of the maximum concentration of RIF that the mutants could withstand, growth was assayed on plates containing higher concentrations of RIF.

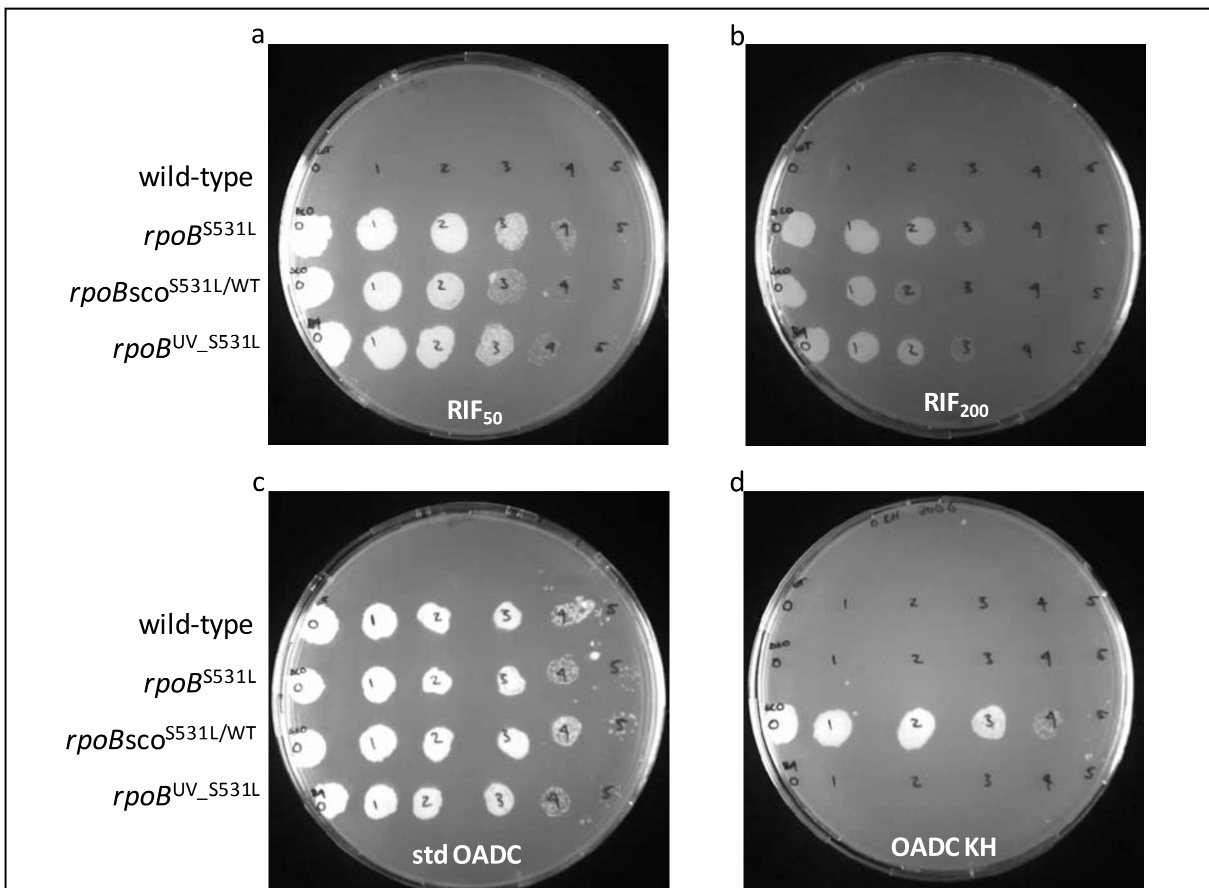


Figure 3.17 S531L mutants are resistant to high levels of RIF.
Dilutions of cells grown to logarithmic phase were plated on RIF containing media. Twenty microliters of dilutions 10^0 – 10^{-6} were plated (left to right on the plate).Dilutions were prepared in un-supplemented 7H9. Growth was evaluated on 7H10 OADC plates containing RIF_{50} and RIF_{200}; 7H10 OADC plates containing $KAN_{20}HYG_{50}$ (OADC KH) and standard 7H10 OADC plates (std OADC).

S531L and H526Y SCOs, and DCOs grew on RIF concentrations of up to 400 µg/ml (RIF_{400}) (Figure 3.18). The *rpoB*^{UV_S531L} mutant had comparable growth to DCO strains. Notably, at RIF_{200} or higher, growth of the *rpoB*sco$^{S531L/WT}$ and *rpoB*sco$^{WT/H526Y}$ only occurred up to dilutions of 10^{-2}; while *rpoB*^{S531L} and *rpoB*^{H526Y} DCO strains and *rpoB*^{UV_S531L} grew up to dilutions of 10^{-3} at RIF_{200}. This indicates that the wild-type version of the *rpoB* still present in SCO strains is being targeted by RIF. Conceivably RNAPs in the cell contain a mixture of mutated and wild-type β-subunits, and RIF is able to target the wild-type version. At high enough RIF concentrations, inhibition of sensitive RNAPs is probably sufficient to impede growth of SCO strains.

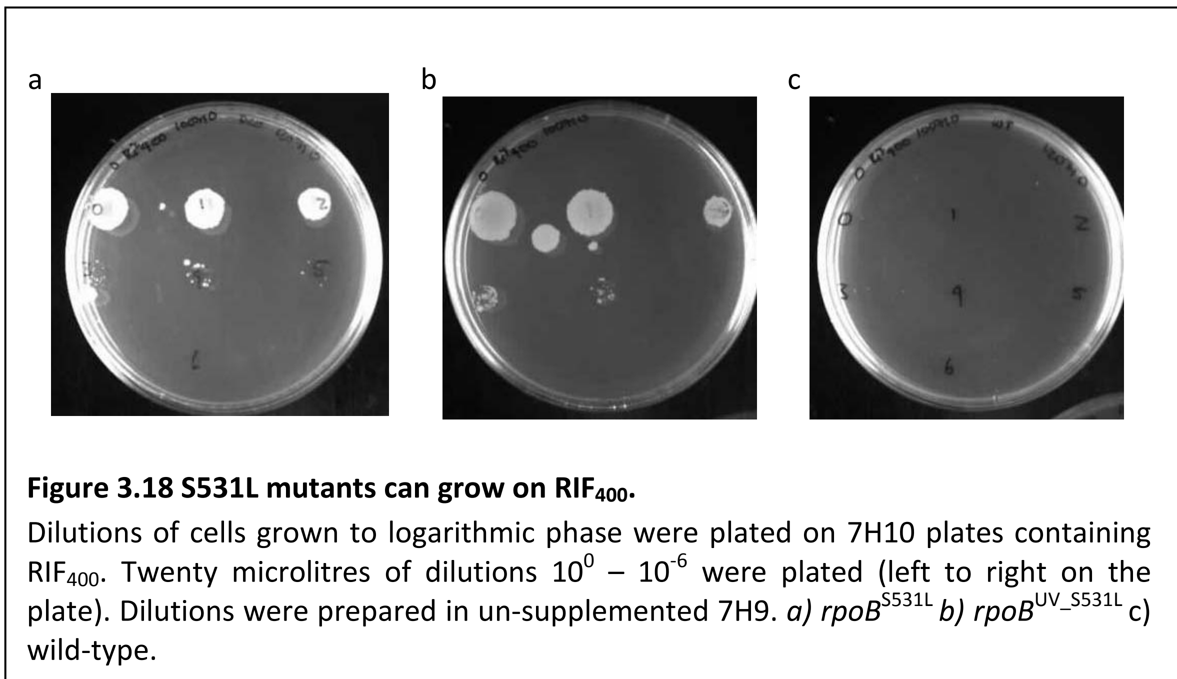


Figure 3.18 S531L mutants can grow on RIF_{400}.
Dilutions of cells grown to logarithmic phase were plated on 7H10 plates containing RIF_{400}. Twenty microlitres of dilutions 10^{0} – 10^{-6} were plated (left to right on the plate). Dilutions were prepared in un-supplemented 7H9. *a) rpoB*^{S531L} *b) rpoB*^{UV_S531L} c) wild-type.

3.7.2 MABA's were useful for MIC_{RIF} determination and for evaluation of physiological effects of RIF^{R}-associated mutations

The use of 96-well microtitre plates for MABA allows for evaluation of growth in the presence of serial two-fold dilutions of RIF and therefore allows for more accurate definition of the MIC_{RIF}. To set up the plates (Figure 3.19), liquid medium was added to all wells in the plate (except for the uppermost row, which served as a no growth control). As per the standard protocol, the outermost column served as a positive growth control, in which only growth medium and cells were added. Dimethyl suphoxide (DMSO) was used as solvent for RIF therefore the next column contained a dilution series of DMSO

only to evaluate its effect on strain growth. RIF was added to the uppermost wells, and diluted 'down' the columns. In this way a two-fold dilution series of RIF was established. MSM cells from a culture that had been grown to logarithmic phase were then diluted to an optical density (OD_{600}) of 0.05, and added to every well (except for the uppermost row which served as a contamination control). Plates were incubated at 37 ˚C for 48 hours, and Alamar blue added. Plates were incubated for a further 24 hours after the addition of Alamar blue and growth evaluated. Alamar blue is colorimetric growth indicator which contains resazurin. Resazurin is blue in its oxidised form (Mann & Markham, 1998). During bacterial growth the surrounding medium is reduced as bacteria use up oxygen. This drives the first step in resazurin reduction, and results in a colour change from blue to pink. Initial reduction of resazurin is irreversible, and produces a resofurin. Resofurin can undergo a reversible reduction to the form the colourless compound, dihydroresofurin (Mann & Markham, 1998). A change from blue to pink indicates growth, and wells that remain blue indicate that no growth has occurred. The MIC_{RIF} was defined as the concentration of RIF wherein growth was impeded. Therefore the transitional position on a plate where the blue to pink colour change ceased to occur was determined as the MIC_{RIF} (highlighted by yellow boxes in figures).

RIF^S determination is routinely performed using standard mycobacterial growth media, which contain glucose as principal carbon source. Mycobacteria possess an assortment of lipid metabolising enzymes (Cole *et al.*, 1998), and there is increasing evidence that lipids are the predominant carbon source during MTB infection (Savvi *et al.*, 2008, Brzostek *et al.*, 2009, Pandey & Sassetti, 2008, de Carvalho *et al.,* Marrero *et al.*, 2010, Garton *et al.*, 2008, Peyron *et al.*, 2008). During treatment with INH, a two kill step kill pattern is observed for MTB populations, which suggests that actively growing cells are killed rapidly by INH and more slowly growing cells take longer to kill (Jindani *et al.*, 2003). Moreover, Xie *et al.* (2005) show that of seventeen antibacterial compounds tested, only four were active against starved MTB populations (Xie *et al.*, 2005) and adjustment of carbon source can vary the types of molecules that are isolated from chemical library screens for activity against mycobacteria (Miller *et al.*, 2009, Pethe *et al.*, 2010). The ability to withstand antibiotic pressure is therefore intrinsically linked to physiological

pathways such as carbon metabolism. Therefore, the use of the MABA method to evaluate MIC_{RIF} facilitated the simultaneous investigation of two important questions raised by this project. That is, the ability of *rpoB* mutations to confer RIF^R, and the implication of these mutations for carbon source utilization.

3.7.3 MABA's in 7H9 OADC media reveal increased MIC_{RIF} for $rpoB^{S531L}$ and $rpoB^{H526Y}$

The MIC_{RIF} for $rpoB^{S531L}$, $rpoB^{H526Y}$ and $rpoB^{L521L}$ mutant strains, as well as the $rpoB^{UV_S531L}$ mutant and the wild-type strain, was initially determined in 7H9 OADC. The MIC_{RIF} for the $rpoB^{S531L}$ and $rpoB^{H526Y}$ mutant strains was 125 – 250 µg/ml (RIF_{125} – RIF_{250}) (Figure 3.19), while growth of $rpoB^{L521L}$ was inhibited at RIF concentrations between 0.98 µg/ml and 1.95 µg/ml. Wild-type growth was inhibited between 1.95 µg/ml and 3.9 µg/ml. These values are much lower than the published MIC_{RIF}, which range from 8 µg/ml (Piddock *et al.*, 2000) to 32 µg/ml (Alexander *et al.*, 2003), and probably reflects some of the difficulties in evaluating RIF^S in MSM (discussed in 4.4). In initial experiments, S531L and H526Y mutants were resistant to the highest concentration of RIF tested, namely 250 µg/ml. These strains grew at RIF_{125} – RIF_{250}, which were concentrations at the uppermost limit of the assay in this format. This was confirmed in subsequent experiments, which established an MIC_{RIF} for these mutants of 250–500 µg/ml (RIF_{250} – RIF_{500}).

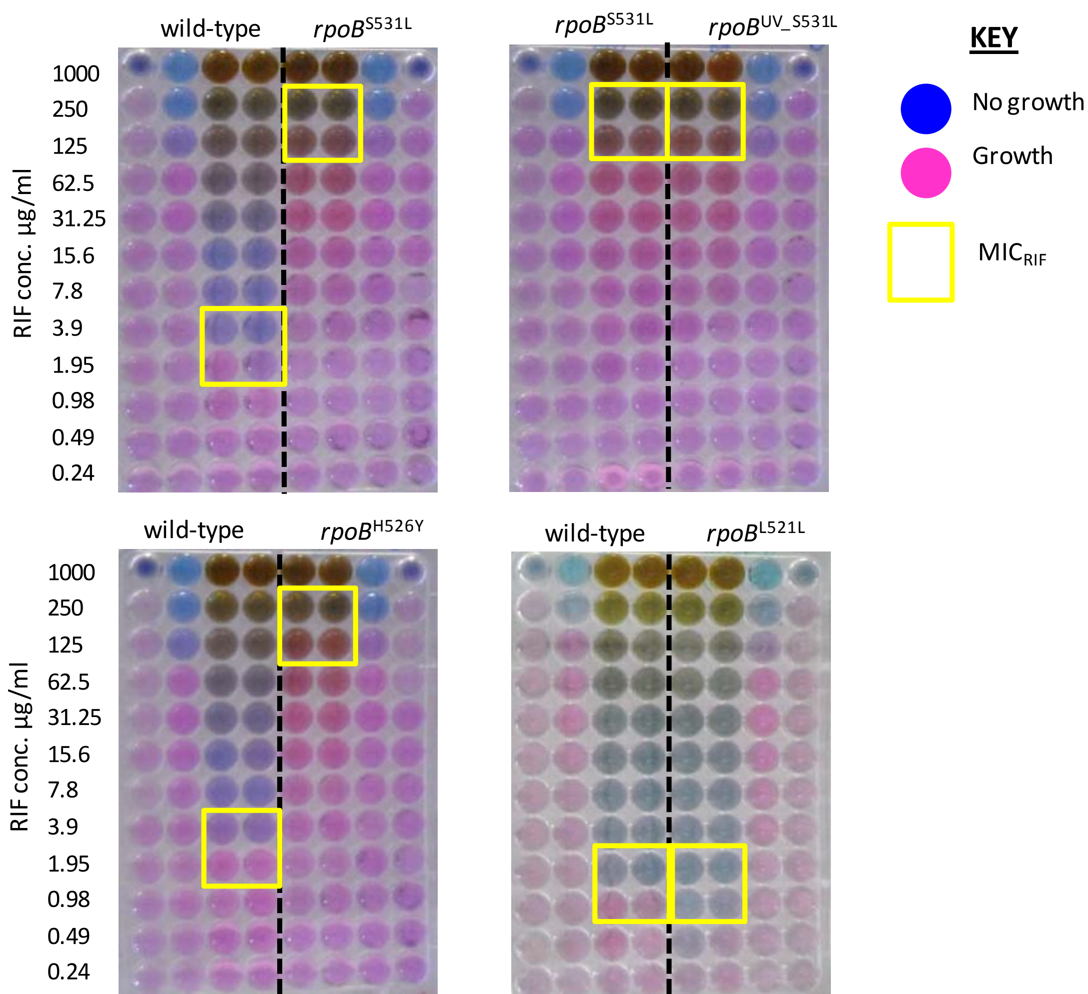


Figure 3.19 Decreased RIF-susceptibility of *rpoB*^{S531L} and *rpoB*^{H526Y} but not *rpoB*^{L521L} mutants.

The MIC_{RIF} is indicated in yellow boxes. No cells are added to the uppermost well, which serves as a contamination control – blue wells in this row confirm that there is no contamination. The first and second outermost columns serve as positive growth controls, and a control for the effect of the solvent respectively. 7H9 OADC medium was used in this assay. MIC's for different strains are as follows: wild-type: 1.95 – 3.9 µg/ml; *rpoB*sco^{S531L}: 125 – 250; *rpoB*^{S531L}: 125 – 250 µg/ml; *rpoB*^{UV_S531L} (mutant generated by UV exposure and selection on RIF): 125 – 250 µg/ml; *rpo*sco^{H526Y}: 125 – 250 µg/ml (not shown in figure); *rpoB*^{H526Y}: 125 – 250 µg/ml; *rpoB*sco^{L521L}: 1.95 – 3.95 µg/ml (not shown in figure); *rpoB*^{L521L}: 0.98 – 1.95 µg/ml (note, the MIC_{RIF} for wild-type in this plate is 0.98 – 1.95 µg/ml).

3.7.4 The EPI's, verapamil and reserpine, do not impact MIC_{RIF} of *rpoB*^{S531L}

The results presented above, and those from other studies, provide convincing evidence for the role of RRDR mutations in RIF^R. However, there are questions that surround the sufficiency of these mutations to confer the high levels of resistance seen in clinical isolates. Approximately 5 % of clinical isolates have no mutations in the RRDR (Telenti *et al.*, 1993, Ramaswamy & Musser, 1998), and the question of compensatory mutation has never been conclusively addressed. Among various possible mechanisms, there is increasing interest in the role of efflux in antibiotic resistance in bacterial pathogens, including MTB (Louw *et al.*, 2009). In *Salmonella* species, for example, fluoroquinolone resistance-associated *gyrA* mutants are more difficult to generate if efflux systems are inhibited (Ricci *et al.*, 2006). Moreover, a recent mathematical model postulates that activated efflux systems are required to fix and maintain resistance-associated point mutations in a population (Fange *et al.*, 2009). Mycobacterial genome sequences indicate the capacity to express efflux systems of every major class, however these systems are not well-characterised (Louw *et al.*, 2009). To evaluate the role of efflux in RIF^R in our system, efflux pump inhibitors (EPIs) were included in MABA assays of the S531L mutant strains. Two compounds were used in EPI assays to increase the possibility of inhibiting different types of efflux systems: reserpine, which specifically inhibits ATP-dependant pumps, and verapamil which impedes P-glycoprotein systems (Lechner *et al.*, 2008).

In agreement with previous observations (Piddock *et al.*, 2000), neither had significant impact on RIF metabolism. It is possible that MSM encodes efflux systems that are impervious to reserpine or verapamil, or that inhibition of one group of efflux pumps is offset by separate efflux systems; however, the investigation of these possibilities was beyond the scope of this study and, to some extent, obviated by subsequent results (see Section 3.7.5)

Table 3.3 MIC_{RIF} of S531L mutant strains in the presence of EPI's

EPI	MIC_{RIF} (µg/ml):							
	-EPI				+EPI			
	wild-type	*rpoB* S531L	*rpoB*sco S531L/WT	*rpoB* UV_S531L	wild-type	*rpoB* S531L	*rpoB*sco S531L/WT	*rpoB* UV_S531L
res[a]	1.95 – 3.9	125 – 250	125 – 250	125 – 250	3.9 – 7.8	125– 250	125 – 250	125 – 250
ver[b]	1.95 – 3.9	125 – 250	125 – 250	125 – 250	0.48– 0.98	125– 250	125 – 250	125 – 250

[a]Reserpine was included in assays at 20 µg/ml. [b]Verapamil was included in assays at 150 µg/ml

3.7.5 Restoration of the wild-type allele in *rpoB*^{S531L} restores RIF^S

Inhibition of selected efflux systems had insignificant impact on RIF^S in S531L mutants, so providing further support for the sufficiency of mutations in the RRDR for RIF^R. The wild-type allele was restored in the S531L mutant by a second round of allelic exchange. To this end, a suicide vector containing the wild-type *rpoB* allele was delivered into the S531L mutant strain by electroporation, and SCO strains isolated as described previously (see Section 3.5). The SCO mutants were then subjected to sucrose selection, and putative DCO strains screened by AS-PCR; in contrast to the generation of the *rpoB* mutants, in this case the opposite result, namely replacement of the mutant allele with the wild-type *rpoB*, was the desired outcome, and so the screen was conducted accordingly. Restoration of the wild-type *rpoB* allele in *rpoB*^{S531L} (strain designation *rpoB*$^{S531L \rightarrow WT}$) restored RIF^S to levels observed with wild-type MSM (Figure 3.20). In combination with the observed inability of the selected EPI's to impact RIF^S, this result provided very strong evidence that an S531L mutation, is sufficient on its own for high level RIF^R in MSM. However, these strains have not been subject to whole genome re-sequencing and so it remains a formal possibility that a second site mutation – in combination with an *rpoB* mutation - is required for high-level RIF^R in MSM.

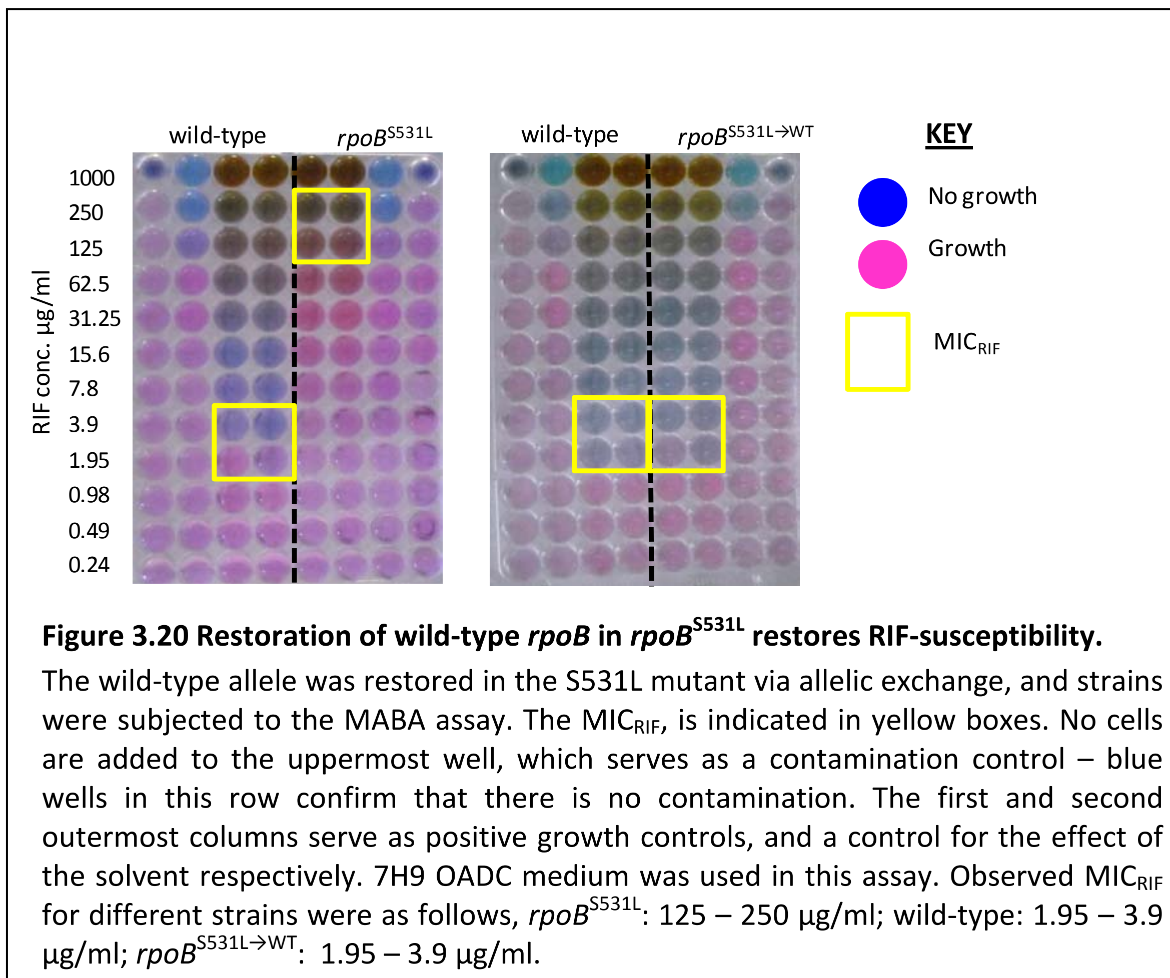


Figure 3.20 Restoration of wild-type *rpoB* in *rpoB*^{S531L} restores RIF-susceptibility. The wild-type allele was restored in the S531L mutant via allelic exchange, and strains were subjected to the MABA assay. The MIC_{RIF}, is indicated in yellow boxes. No cells are added to the uppermost well, which serves as a contamination control – blue wells in this row confirm that there is no contamination. The first and second outermost columns serve as positive growth controls, and a control for the effect of the solvent respectively. 7H9 OADC medium was used in this assay. Observed MIC_{RIF} for different strains were as follows, *rpoB*^{S531L}: 125 – 250 µg/ml; wild-type: 1.95 – 3.9 µg/ml; *rpoB*$^{S531L \rightarrow WT}$: 1.95 – 3.9 µg/ml.

3.7.6 Evaluation of potential carbon source specific effect on MIC_{RIF} of *rpoB*^{S531L} mutants

Conventional antibiotic targets are essential for function, and so are often associated with high levels of evolutionary conservation. It follows, therefore, that mutations in essential genes might have significant implications for bacterial function. There is evidence from an array of other bacterial species that mutations in *rpoB*, primarily those occurring in the RRDR, can have profound implications for physiology (Hosaka *et al.*, 2009, Maughan *et al.*, 2004, Nicholson & Maughan, 2002, Perkins & Nicholson, 2008, Jenkins *et al.*, 2009, Yu *et al.*, 2005). Notably, in *B. subtilis*, *rpoB* mutations analogous to those found in clinical MTB strains have been reported to confer the ability to subsist on carbon sources previously defined as non-utilizable (or non-nutritive) for the wild-type strain; for example, β-methyl-D-glucoside (Perkins & Nicholson, 2008). Moreover, as previously mentioned carbon source metabolism can significantly influence susceptibility to antibiotics (Xie *et al.*, 2005). We therefore investigated the potential influence of carbon source on MIC_{RIF}, as well as the possible impact of the S531L mutation on the

ability of MSM to utilize different carbon sources for growth *in vitro*. To this end, the MIC_{RIF} of selected *rpoB* mutants strains was determined by the MABA method (as previously described; Section 3.6.3.2) in a limited panel of media containing defined carbon sources: standard 7H9 medium supplemented with either C_4 (valerate) or C_5 (butyrate) lipids (Savvi *et al.*, 2008), as well as OADC (containing both oleic acid, a C_3 fatty acid; and dextrose, both of which could be used as carbon sources for growth) which is used routinely for mycobacterial growth assays (including MTB), or a GS mix which is used specifically for MSM in our laboratory. In some ways this study was designed to evaluate the feasibility of assaying the effects of single point mutations in *rpoB*, therefore a limited number of carbon sources were selected. There are several lines of evidence to indicate that fatty acids are the predominant carbon source during MTB infection (Brzostek *et al.*, 2009, Pandey & Sassetti, 2008, de Carvalho *et al.*, 2010, Marrero *et al.*, 2010, Garton *et al.*, 2008, Peyron *et al.*, 2008, McKinney *et al.*, 2000), therefore MABA's were focussed on fatty acid based carbon sources.

The MIC_{RIF} for the wild-type strain was significantly higher in lipid-based carbon sources (Table 3.4). In contrast, the MIC_{RIF} for S531L mutant strains did not differ in any of the carbon sources evaluated (Table 3.4). As described above (Section 3.6.), mutants were resistant to the highest concentration of RIF tested, therefore carbon source specific effects may not be discernable. Therefore, this assay was repeated using a higher starting concentration of RIF. In these experiments, plates were set up so that the highest concentration of RIF in which cells would be added was 1000 µg/ml. Owing to the deep red colour of RIF, colorimetric analysis of growth was not possible. Instead, growth was evaluated using pellet size. To ensure that pellets indicated growth and not cell debris, cell growth was evaluated by spotting on standard 7H10 OADC plates. The MIC_{RIF} for *rpoB*^{S531L} and *rpoB*^{H526Y} was 250 – 500 µg/ml. These results mirror those obtained on solid media, and provide further support for the importance of these mutations for RIF^R. In agreement with RIF^S determination on solid media (Section 3.7.1), SCO strains had slightly lower MIC_{RIF} values (125 – 250 µg/ml, for both *rpoB*sco$^{S531L/WT}$ and *rpoB*sco$^{WT/H526Y}$). This re-iterates that RIF^S in SCO strains is likely due to the presence of wild-type *rpoB*. Although growth was slower in butyrate- and valerate-containing media, there was no

difference in RIF^S for mutant strains. This is likely due to the very high MIC_{RIF} for mutant strains however, this result requires further investigation.

Table 3.4 MIC_{RIF} of S531L mutant strains in varied carbon sources

Growth medium	MIC$_{RIF}$ (µg/ml):				
	wild-type	*rpoB*^{S531L}	*rpoB*sco$^{S531L/WT}$	*rpoB*^{UV_S531L}	*rpoB*$^{S531L \rightarrow WT}$
7H9 GS	1.95 – 3.9	250 - 500	125 – 250	250 - 500	1.95 – 3.9
7H9 OADC	1.95 – 3.9	250 - 500	125 – 250	250 - 500	1.95 – 3.9
7H9 Valerate	15.6 – 31.25	250 - 500	125 – 250	250 - 500	7.8 – 15.6
7H9 Butyrate	15.6 – 31.25	250 - 500	125 – 250	250 - 500	7.8 – 15.6

3.8 Competition assays reveal a modest fitness defect for *rpoB*^{S531L}

Resistance-associated mutations in antibiotic targets are expected to impact physiology, and possibly to incur a fitness cost. Moreover, mathematical models predict that the spread of drug-resistant organisms is critically dependant on their relative fitness (Cohen and Murray, 2004). Previous studies have shown that RIF^R-associated *rpoB* mutations incur heterogeneous fitness costs in mycobacteria, which depend on strain background as well as the assay system employed (Gagneux *et al.*, 2006b, Mariam *et al.*, 2004, Billington *et al.*, 1999, Gillespie *et al.*, 2002). Therefore, we evaluated the relative fitness of the *rpoB* mutants *in vitro*. To this end, a competition assay was employed to determine if the *rpoB*^{S531L} mutant exhibited a growth phenotype (advantage or disadvantage) when grown in competition with the wild-type strain. Competition assays were set up as previously described (Gagneux *et al.*, 2006b). In all cases, wild-type MSM and the *rpoB*^{S531L} mutant were grown in axenic culture as controls. Colony counts on RIF_{200} plates would allow enumeration of *rpoB*^{S531L} strains, and subtraction of colony numbers on RIF_{200} plates, from total colonies on standard OADC plates would allow enumeration of the number of wild-type bacilli, thereby ensuring that the number of wild-type versus mutant baciili could be evaluated in the competition assay. Competition experiments were conducted with three separate sets of cultures and proceeded for 20 days (480 hours), until two sets became contaminated, and had to be discarded. At the outset, we were concerned that

plating on RIF may be problematic. Modelling and clinical data indicate that the likelihood of reversion to the wild-type allele in the absence of the antibiotic is low (Andersson & Levin, 1999). However, if this had occurred in our system, revertants would have disappeared into a population of wild-type strains, as the only way to differentiate between the wild-type strain and the mutant was by the RIF^R phenotype. Moreover, spontaneous mutation in the wild-type population may lead to RIF^R – especially at later stages of the experiment where nutrient limitation may impose a stress (Foster, 1998, Karunakaran & Davies, 2000, Cairns *et al.*, 1988, Bjedov *et al.*, 2003, Rosenberg & Hastings, 2004). This would bias results, and indicate falsely inflated number of *rpoB*^{S531L} mutant strains. However, this concern was allayed since no spontaneous RIF^R mutants were observed from axenic wild-type cultures for the duration of the experiment. This is in contrast to the idea that *rpoB* mutations increase in stationary phase (Foster, 1998, Karunakaran & Davies, 2000, Cairns *et al.*, 1988, Bjedov *et al.*, 2003, Rosenberg & Hastings, 2004) and differs from what was previously observed for mycobacteria (Karunakaran & Davies, 2000).

In axenic culture, wild-type and *rpoB*^{S531L} displayed very similar growth characteristics (Figure 3.21a) for the duration of the experiment. However, when strains are in direct competition, the *rpoB*^{S531L} mutant has a growth defect compared to the wild-type strain, and there was approximately one log difference in cells numbers for most of the experiment (Figure 3.21b). Slightly higher numbers of wild-type were seeded in the experiment, but this unlikely to account for differences in growth between the wild-type and mutant strain. The relative fitness of the mutant to the wild-type strain was calculated as previously described (Gagneux *et al.*, 2006b) and was found to be 0.87 (see Methods and Materials 2.4.6). This is a similar value to what was previously observed for the fitness of RIF-selected MTB mutants (Gagneux *et al.*, 2006b, Mariam *et al.*, 2004). The observed fitness cost is relatively low, and there were several observations which indicate the possibility that the number of *rpoB* mutants in the mixed culture may have been higher.

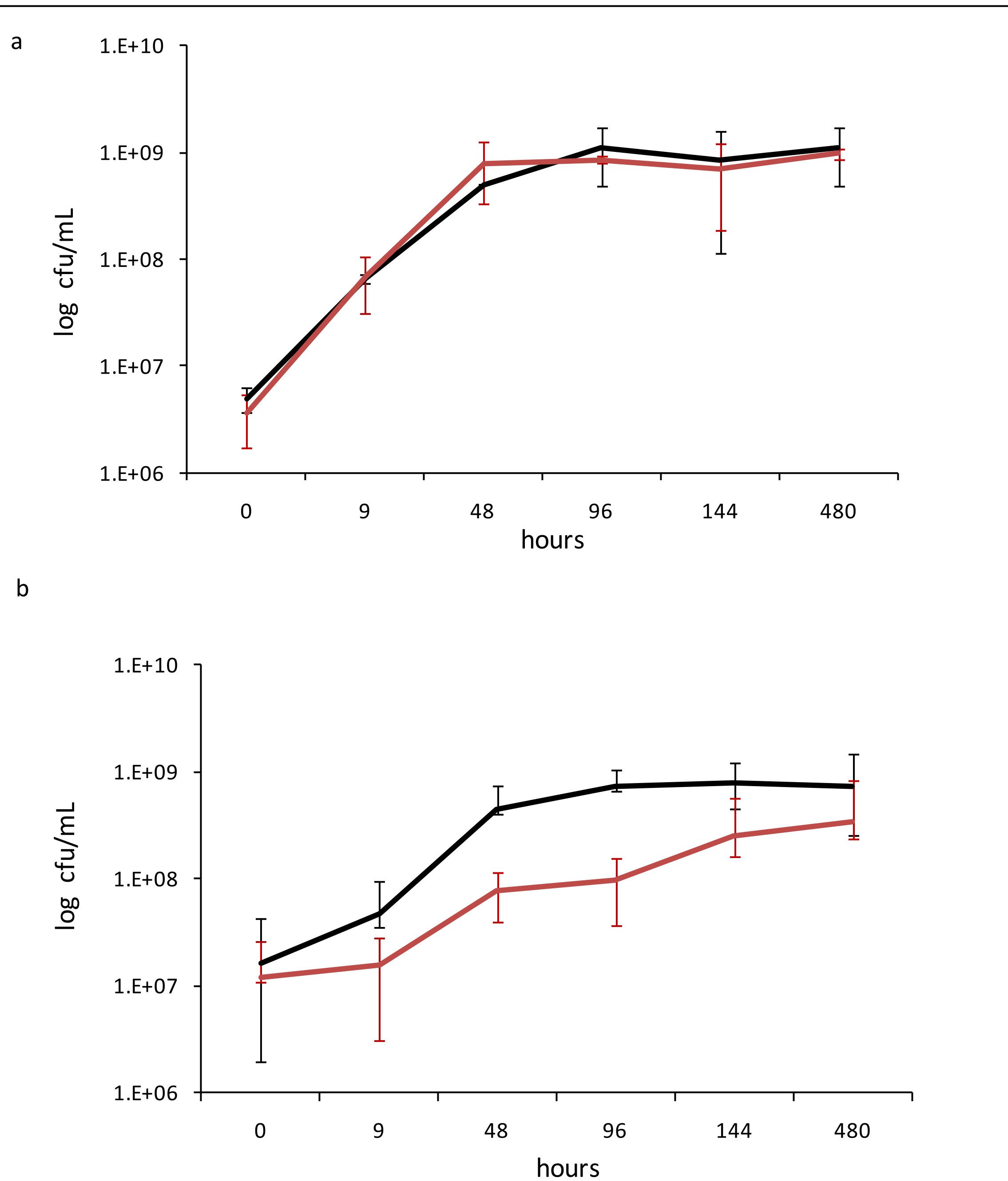


Figure 3.21 The *rpoB*^{S531L} mutant has a fitness defect when grown in competition with wild-type MSM in 7H9 OADC.

a) Growth of wild-type (——) and *rpoB*^{S531L} (——) in axenic culture. *b)* Growth of wild-type and *rpoB*^{S531L} in competition. Experiments were continued for 20 days (480 hours). Cells were enumerated by plating on standard 7H10 OADC and 7H10 OADC RIF_{200}. Experiments were conducted three times, and error bars represent standard deviation, which indicates how much variability there is between the calculated average, and individual measurements at each time point.

3.8.1 Concerns regarding the use of RIF^R to differentiate between wild-type and the *rpoB*^{S531L} mutant

As discussed above, there was some concern about using the RIF^R phenotype to differentiate between wild-type and *rpoB*^{S531L} strains. Observations during the competition experiment question the suitability of RIF^R for this purpose, and moreover raise important physiological questions.

The *rpoB*^{S531L} strain was grown in axenic culture, and growth was evaluated on standard 7H10 OADC plates, as well as 7H10 OADC RIF_{200} plates. Previous data indicated that the *rpoB*^{S531L} strain was resistant to levels of RIF exceeding 200 µg/ml on solid, and in liquid, media (Sections 3.7.1 and 3.7.3). Similar numbers of *rpoB*^{S531L} colonies should therefore be observed on standard OADC and RIF_{200} plates. We were surprised to see that the number of colonies on standard OADC plates was consistently higher than on RIF_{200} plates (average for the entire experiment: 5.1 X 10^8 CFU/ml (standard OADC) vs**.** 2.61 X 10^8 CFU/ml (RIF_{200})). RIF_{200} plates containing no colonies were observed on more than one occasion, and the decreased culturability of *rpoB*^{S531L} appeared to decrease as the experiment progressed (Figure 22). These observations raised questions regarding the ability of growth on RIF_{200} to accurately represent *rpoB*^{S531L} population size.

AS-PCR was performed on a selection of plates from different time points. In the competition culture, standard OADC plates should contain a mixture of wild-type and *rpoB*^{S531L} strains. To determine the proportion of strains in the mixed culture, the genotype of a selection of colonies from standard OADC plates was determined by AS-PCR. AS-PCR on DNA extracted from a selection of colonies would therefore give an idea of the proportion of strains that existed in the culture prior to plating. The genotype of twenty colonies from time points of 9 hours; 7 days (168 hours); 20 days (480 hours) and 48 days (1152 hours) were determined. Only 2 sets of cultures had previously become contaminated, and one set was maintained until 62 days, which enabled us to genotypically evaluate strains composition at 48 days. At 9 hours, AS-PCR indicated that 55 % of strains were wild-type, and 45 % were *rpoB*^{S531L} (Figure 3.22a). By day 20, the strain composition had changed to 70 % wild-type, and 30% mutant (Figure 3.22c). The

corresponding RIF_{200} plate had no colonies on it, and the number of *rpoB*^{S531L} mutants had been scored as zero (Figure 3.22c). AS-PCR and observations from axenic *rpoB*^{S531L} cultures indicate that colony numbers on RIF_{200} are not representative of the actual number of *rpoB*^{S531L} mutants present in the mixed culture. This evaluation was only conducted once and only on a selection of plates. It is therefore plausible that this is a rare phenomenon, and not representative of the whole experiment.

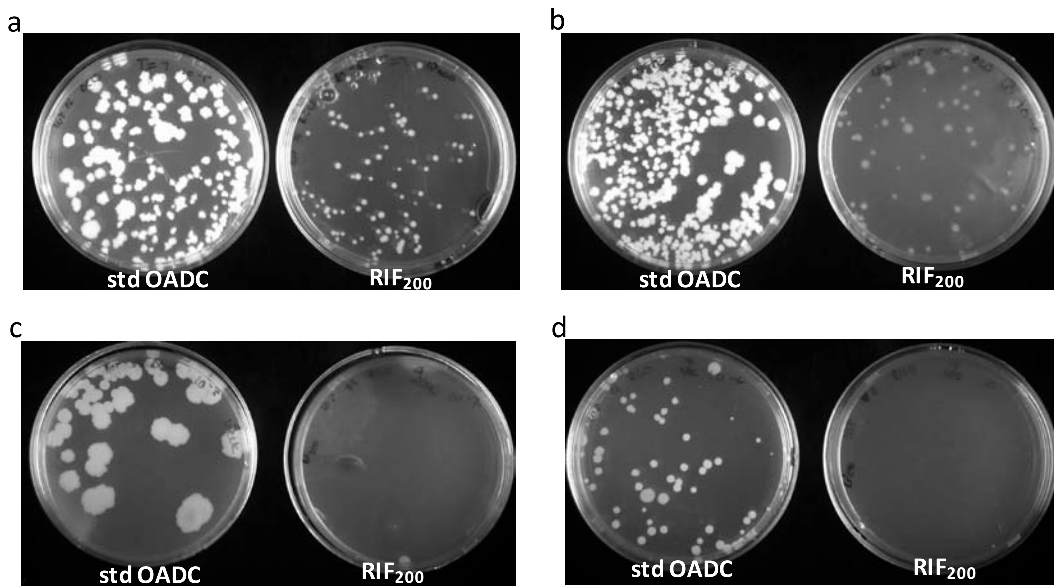


Figure 3.22 Growth on RIF_{200} plates does not represent the full extent of colonies present in the competition culture.

AS-PCR was conducted on a DNA from selection of colonies from various time points of the competition culture to determine the relative numbers of wild-type and *rpoB*^{S531L}. Colonies were selected from standard 7H10 OADC (std OADC plates), the corresponding 7H10 OADC RIF_{200} plates (RIF_{200}) are shown. Selected time points: *a)* 9 hours (55 % wild-type; 45 % *rpoB*^{S531L}) *b)* 7 days (52 % wild-type; 48 % *rpoB*^{S531L}) *c)* 20 days (70 % wild-type; 30 % *rpoB*^{S531L}) *d)* 48 days (85 % wild-type; 15 % *rpoB*^{S531L}). The numbers of *rpoB*^{S531L} colonies on RIF_{200} plates are consistently lower than indicated by AS-PCR. By day 20 and 48 several RIF_{200} plates containing no colonies were observed, even though AS-PCR reveals 30 % and 15 %, respectively, of strains are *rpoB*^{S531L}.

To evaluate the ability of *rpoB*^{S531L} strains coming out of the competition assay to grow in the presence of RIF, a 1 ml aliquot was removed from the mixed culture. Five hundred microliters was added to standard 7H9 OADC liquid media, and the other 500 µl to 7H9

OADC RIF_{200} media. The same was done for the wild-type axenic cultures, and OD_{600} was monitored daily, for 7 days. By the time these experiments were initiated, the assay was in its 9th week, and few colonies were observable on RIF_{200} plates. The OD_{600} of the axenic wild-type aliquot, in standard 7H9 OADC media, increased unexceptionally, by day 2 it was 0.77 and by day 7, it was 4.8. The OD_{600} of cells removed from the competition culture and added to in standard 7H9 OADC increased similarly to the wild-type (0.67 at day 2, and 5.4 at day 7). In contrast in 7H9 RIF_{200} broth, the aliquot from the competition culture had only reached an OD of 0.01 at day 2, and 0.44 at day 7. There was no increase in OD_{600} for the axenic wild-type cells in RIF_{200} media therefore it is unlikely that the delayed increase in OD was due to a late mutant emerging from the wild-type population. This indicates that *rpoB*^{S531L} mutants coming out of competition culture had difficulty growing in the presence from RIF, despite a MIC_{RIF} of 250 – 500 µg/ml.

The implication of these observations for this experiment is that the number of *rpoB*^{S531L} mutants represented on RIF_{200} plates, may have been lower than were actually present in the mixed culture, and results therefore may be skewed in favour of the wild-type strain. Perhaps more interesting though, is the question of why *rpoB*^{S531L} mutant strains, that are resistant to high levels of RIF, have reduced culturability on RIF as they come out of competition culture. This issue speaks to both key questions in this project – that is the sufficiency of single RIF^R-associated mutations to confer high level RIF^R, and the implications of these mutations for mycobacterial physiology.

4. Discussion

New anti-TB chemotherapeutics and improved control measures are direly needed to curb the burgeoning drug-resistant TB epidemic. An understanding of the how drug resistance develops, and how resistance impacts fitness, is crucial if new control strategies are to be effectively maintained. The majority (approximately 95 %) of all clinical RIF^R MTB isolates contain mutations in the *rpoB* gene (Ramaswamy & Musser, 1998, Telenti *et al.*, 1993). These mutations occur at over 25 sites in *rpoB*, and are associated with varying levels of RIF^R (Sandgren *et al.*, 2009). Notably, the same *rpoB*

allele can be associated with different levels of RIF^R in different strain backgrounds (Gagneux *et al.*, 2006b). This suggests that second-site mutations may be important in drug resistance-associated fitness, and highlights the need to introduce single, defined *rpoB* mutations into a strain in order to study the effects on drug resistance and physiology. For MTB, an obligate human pathogen, the ability to survive and be transmitted in a population is intricately dependent on physiological processes. Therefore, the impact of drug resistance-associated mutations on MTB physiology has fundamental consequences for fitness, and by association, the spread of drug-resistant strains. Moreover, outbreaks of multidrug-resistant (MDR)- and extensively drug-resistant (XDR)-TB raise significant biological questions surrounding the ability of MTB strains, which contain mutations in multiple essential genes, to maintain infection and transmission in a population.

This study was designed to address two broad, but interrelated, research questions: 1. The impact of specific drug resistance-associated mutations on mycobacterial physiology, in particular their consequences for perceived fitness; and 2. The contribution of single point mutations to resistance levels. To this end, a panel of MSM strains containing RIF^R-associated mutations was generated via SDM (Smith & Klugman, 1997) for point mutations and Xover-PCR (Link *et al.*, 1997) for the 30 bp deletion together with allelic exchange (Gordhan & Parish, 2001) (see Section 3.5). Observations of massively decreased RIF-susceptibility (RIF^S) of MSM strains containing S531L or H526Y mutations adds to the body of evidence indicating the importance of these mutations for RIF^R (Williams *et al.*, 1998, Miller *et al.*, 1994, Sandgren *et al.*, 2009, Ramaswamy & Musser, 1998, Telenti *et al.*, 1993). Moreover, an increase in susceptibility after restoration of the wild-type allele in the *rpoB*^{S531L} strain provides novel evidence to suggest an S531L mutation is sufficient for RIF^R (described in Section 3.7.5). The *rpoB*^{S531L} strain had a modest fitness defect when grown in competition with the wild-type strain (see Section 3.8). However, preliminary experiments indicate that more *rpoB*^{S531L} cells may have been present in the competition culture, than was represented by growth on RIF containing media. This raises the provocative possibility that the fitness of the mutant may actually

be higher than observed, and raises additional biological questions that require further investigation.

4.1 Advantages (and disadvantages) of the study design for this project

Since the study utilized MSM as research organism, an additional theme underlying this work was an assessment of the application of any techniques developed to MTB in future work. This is not a trivial point, and highlights our genuine concerns in setting up this project with our ability to generate *rpoB* point mutants in the absence of RIF selection and, more importantly, the type of assays that we might be apply to assess resultant fitness defects (or enhancements). These concerns are both physiological and experimental: Is it possible to generate a mutant that differs from its parent by only a single point mutation in a desired locus? How might an assay be designed so that it provides the best proxy measure of *in vivo* fitness? Is more than one assay required to approximate different stages of infection? How can the assay be set up to facilitate the differentiation of wild-type and mutant strains? *etc*. However, while the results presented here can therefore be considered "proof of concept" in many respects, there are certain key aspects of our results which we consider worth pursuing in their own right (in MSM) for their potential impact on our understanding of fundamental mycobacterial physiology.

Given the large number (approximately 25) of *rpoB* alleles associated with RIF^R in MTB, it was necessary to select a panel of mutants to replicate in MSM by SDM and allelic exchange. An analysis of the literature identified the two most common *rpoB* alleles observed in RIF^R MTB strains. This was subsequently confirmed when the TB Drug Resistance Database (TBDReaDB) was established (Sandgren *et al.*, 2009). The prevailing theory equates clinical frequency with fitness: that is, the most common mutations carry the least fitness cost (Gillespie *et al.*, 2002, Mariam *et al.*, 2004, Gagneux *et al.*, 2006b, Shcherbakov *et al.*, 2010). However, while intuitive, this impression overlooks fundamental uncertainties regarding the nature of the selective forces acting on clinical strains, both during infection (with or without antibiotic exposure) and, importantly during strain isolation and processing *in vitro*. A significant proportion of sputum bacilli,

for example, exist in a non-replicating state (Garton *et al.*, 2008), and similarly do not grow *in vitro* without stimulation by a Resuscitation Promoting Factor (RPF)-mediated growth factor (Mukamolova *et al.*, 2010).

In addition to elucidating the relative frequencies of the multiple RIF^R-associated mutations, our analysis of the literature also revealed previous reports indicating that a silent mutation, L521L, might be associated with RIF^R (Siddiqi *et al.*, 2002, Yang *et al.*, 1998). Structural analysis of the RNAP indicates that missense mutations in the RRDR of the *rpoB* gene decrease RIF binding efficiency (Campbell *et al.*, 2001). It was unlikely that this mutation was involved in RIF^R, however, if true, high level RIF^R resulting from a silent mutation would challenge this model. If should be noted that only the RRDR of the *rpoB* gene was analysed in these studies (Siddiqi *et al.*, 2002, Yang *et al.*, 1998). This and MIC_{RIF} data obtained in this study (see Section 3.6), suggests that other mutations in *rpoB* (or elsewhere) contributed to RIF^R in that strain. Nevertheless, the unusual nature of this mutation made it an attractive candidate for inclusion in our study together with the two most common mutations (Table 4.1), and three additional alleles that had been identified in unrelated work at the MMRU: an Q513V mutant that was isolated after UV-irradiation and selection on high level RIF, and found to exhibit differential growth on different growth media; secondly, a mutant containing a TH526TY mutation, which arises from a double transition mutation (**cc**ac → **tt**ac), and is commonly associated with error-prone repair of UV-induced DNA damage, by the SOS-inducible DNA polymerase III subunit, DnaE2 (Boshoff *et al.*, 2003, Warner *et al.*, 2010); and thirdly, a 30 bp deletion that has been isolated in MSM, in response to stress such as UV treatment and starvation, on multiple occasions at the MMRU (Digby Warner, unpublished results).

Table 4.1 Plasmid constructs and mycobacterial mutants generated in this study

Mutation[a]	**Reason for Study**	**Suicide vector**[b]	**SCO**[c]	**pOLYG vector**[d]	**DCO**[e]
WT *rpoB*	Template for SDM and control for complementation and over-expression analyses	✓	✓[f]	✓	N/A
S531L (t**c**g→t**t**g)	Mutation most commonly associated with RIF[R] in clinical MTB isolates (Sandgren *et al.*, 2009).	✓	✓	✓	✓
rpoB$^{S531L \rightarrow WT}$	Restoration of the wild-type allele in the S531L mutant background to determine if this restores RIF[S].	✓	N/A	N/A	✓[g]
H526Y (**c**ac→**t**ac)	Another common mutation in clinical MTB isolates (Sandgren *et al.*, 2009).	✓	✓	✓	✓
L521L (**c**tg→**t**tg)	Silent mutation associated with RIF[R] (Siddiqi *et al.*, 2002, Yang *et al.*, 1998).	✓	✓	X	✓
Q513V (**ca**g→**gt**g)	Preliminary research in the MMRU suggested that this mutation was associated with no growth phenotype in different media; used as a negative control.	✓	X	X	X
T-H526T-Y (ac**cc**ac→ac**tt**ac)	Double transition mutation commonly associated with DnaE2 activity in both MTB and MSM; implications for DNA damage-induced drug resistance (Boshoff *et al.*, 2003).	✓	✓	X	X

30 bp deletion (Q512 – S522) (cagttcatggaccag aacaacccgctgtcg)	During studies at the MMRU, this deletion was isolated in MSM on more than one occasion, in response to exposure to UV irradiation, starvation, and treatment with replication inhibitors (Digby Warner, unpublished). An analogous mutant was isolated in a recent study (Malshetty *et al.*, 2010).	✓	✓	X	X

[a]Mutations are numbered according to the system based on the *E. coli* genome, as per standard practice (Telenti *et al.*, 1993). [c]Suicide vectors, containing a mutated version of *rpoB*, which was generated via targeted mutagenesis (SDM or Xover-PCR). These constructs were used to introduce mutated *rpoB* alleles into the MSM chromosome during allelic exchange. [c]SCO MSM strains, generated during the first step of allelic exchange, contain two copies of the *rpoB* gene – a mutated version and the wild-type version. [d]pOLYG is an episomal, multicopy vector. The wild-type, S531L and H526Y versions of *rpoB* were added to this construct and introduced into MSM to examine implications of gene dosing for RIF^R. [e]DCO MSM strains generated from SCO strains, contain one copy of *rpoB* – the wild-type version has effectively been swapped for a mutated version. [f]This construct contains two copies of the wild-type *rpoB* gene, and was used as a control for evaluation of RIF^R in SCOs. [g]The wild-type *rpoB* allele was restored in the S531L mutant strain.

A number of studies have examined the implications of RIF^R-associated mutations for the fitness and physiology of mycobacteria (Gagneux *et al.*, 2006b, Mariam *et al.*, 2004, Billington *et al.*, 1999, Davies *et al.*, 2000, Gillespie *et al.*, 2002, Jenkins *et al.*, 2009, Malshetty *et al.*, 2010) and other organisms (O'Neill *et al.*, 2006, Yu *et al.*, 2005, Maughan *et al.*, 2004, Perkins & Nicholson, 2008, Hosaka *et al.*, 2009). However, these studies are united in their use of RIF to select for *rpoB* mutations, either in the laboratory, or for clinical isolates during patient treatment. A potential limitation of this approach is that second-site mutations may be introduced into the chromosome. Methodology employed in this study (SDM and allelic exchange) facilitated the generation RIF-naïve strains, that contained single point mutations commonly observed in RIF^R mycobacterial strains. Other techniques are available to introduce point mutations into a gene. Recombineering, which relies on recombination proteins from phage systems, presents an alternative to allelic exchange for replacement of the wild-type *rpoB* allele with a mutated version. This technique has been used to generate point mutations in mycobacterial genes, in some cases without the need for positive selection (van Kessel & Hatfull, 2008). However, many of preparative and screening procedures would be similar to those for allelic exchange. SDM would still have been required to generate a mutated version of *rpoB*, and an AS-

PCR based strategy would have been necessary to screen for resultant mutants. Moreover, generation of mutants by recombineering without positive selection results in similar transformant number to allelic exchange (van Kessel & Hatfull, 2008).

Three groups have previously introduced mutated versions of *rpoB,* containing common RIF^R-associated mutations, into MTB – both groups replaced a portion of a wild-type *rpoB* allele with the corresponding region from a resistant strain (Williams *et al.*, 1998, Zaczek *et al.*, 2009). In a separate study PCR-based SDM had been used to engineer mutations in the *rpoB* gene of MTB, which was then expressed in MSM (Miller *et al.*, 1994). In all three studies a significant decrease of RIF^S was observed upon introduction of the mutated *rpoB* allele. However, in all three cases mutants were generated which retained an intact wild-type *rpoB* gene. For an evaluation of the biology of mycobacterial strains containing mutated *rpoB* alleles, it was important that we generated strains wherein the wild-type allele had been removed.

4.2 Was MSM a good study organism for this project?

The fast growing, non pathogenic MSM mc^2155 was used as research organism in this study (Snapper *et al.*, 1990). MSM has been used extensively to investigate fundamental aspects of mycobacterial physiology, and often – though not in all cases – provides vital preliminary biological insights, as well as an indication of feasibility, prior to transferring the same research questions to MTB and other pathogenic mycobacterial species.

The use of MSM at the research organism in this study necessitated an evaluation of advantages and disadvantages of this system – keeping in mind the central aims of this study. Investigation of RIF^R in MSM is likely to yield interesting insights into mycobacterial biology, and is a good starting point to begin investigating drug resistance in MTB. However there are important physiological differences between these two species which must be considered. Comparative genomic studies have identified a high degree of evolutionary conservation among bacterial *rpoB* genes, which even show some relatedness to components of eukaryotic transcriptional machinery (Lane & Darst, 2010a, Lane & Darst, 2010b). At the amino acid level, the *rpoB* gene is 91% similar between MTB

and MSM, while the RRDR is fully homologous (Figure 3.2). RIF^R-associated amino acid substitutions in the RRDR are therefore, identical in both organisms. The DNA sequence of the RRDR is 91% identical between MTB and MSM, moreover specific bases that would be mutated are exactly the same. These similarities made MSM a suitable organism for study in this project.

Published MIC_{RIF} values for MTB (0.01 – 0.02 µg/mL) (Changsen *et al.*, 2003) are much lower than for MSM (8 – 32 µg/mL) (Piddock *et al.*, 2000, Alexander *et al.*, 2003). Increased tolerance to RIF for some mycobacteria results from the activity of various RIF-inactivating mechanisms. In MSM, for example, ribosylation of RIF by ADP ribosyltransferase leads to increased RIF tolerance (Quan *et al.*, 1997, Alexander *et al.*, 2003), while studies of the intracellular accumulation of radio-labeled RIF indicate that efflux systems are expressed at low levels in MSM, thereby reducing the intracellular concentration of active compound (Piddock *et al.*, 2000). Moreover, MSM contains an RNA binding protein (MsRbpA) that when over-expressed partially protects from the effects of RIF (Dey *et al.*, 2010). The expanded complement of RIF-inactivating mechanisms in MSM underscores the need to repeat this work in MTB, and poses some challenges in working with MSM to evaluate RIF^R and RIF^R-associated mutations. Initial MIC_{RIF} determination revealed that MSM mutants containing the S531L or H526Y mutation were resistant to concentrations of RIF in the 125 – 250 µg/mL range. This was the lower bound of MIC_{RIF} for these strains – MIC_{RIF} determination using higher concentrations of RIF revealed *rpoB*^{S531L} and *rpoB*^{H526Y} could grown in concentrations of RIF up to 250 – 500 µg/mL (Section 3.7). Repetition of these assays in analogous MTB strains is unlikely to yield such high absolute MIC_{RIF} values. However, it will be interesting to see if RIF^S decreases in MTB by the same proportion. As previously mentioned, important physiological and clinically relevant mycobacterial physiological features have been uncovered using MSM as a research organism. Even though, it will be important to repeat RIF^R determination using MTB, the primary aim of this project was to elucidate physiological effects of RIF^R, and MSM is a useful organism to do this in.

4.3 Merodiploidy and RIF^R

SCO recombinants are intermediates in the allelic exchange process and for most gene knock-out studies are merely a necessary step in the process of generating gene knock-out mutants. However in the case of *rpoB*, SCOs are potentially interesting given that some *Nonomuraea* and *Nocardia* sp. encode two paralogues of the *rpoB* gene (Ishikawa *et al.*, 2006, Vigliotta *et al.*, 2005). Moreover, nucleotide substitutions in the alternative *rpoB* gene in *Nonomuraea* sp. strain ATCC 39727 occur at the H526 and S531 loci – and at other sites within the RRDR (Tala *et al.*, 2009). The evaluation of RIF^S of SCO strains was therefore a potentially interesting avenue to pursue. Moreover, this kind of analysis provided the opportunity to investigate the impact of gene dosing on RIF^R.

During initial MABA assays, the highest concentration of RIF that cells were exposed to was 250 µg/mL. In these assays it appeared that S531L and H526Y SCOs were resistant to similar levels of RIF as their derivative DCO mutants. However, when higher concentrations were used in RIF^S assays, it became apparent that SCOs had a slightly lower MIC_{RIF} than DCO strains (Section 3.7.6). These results are consistent with previous reports that suggest that retention of the wild-type allele in a strain containing a mutated version of *rpoB* increases RIF^S (Williams *et al.*, 1998). Interestingly, Ilyina *et al.* (1971) report that, in merodipolid *E. coli* cells – containing both the wild-type *rpoB* gene and a RIF^R variant, the susceptible phenotype dominates (Ilyina *et al.*, 1971). The authors suggest that this is because the wild-type gene binds DNA more efficiently (Ilyina *et al.*, 1971). However, in our assays, the SCOs were associated with decreased susceptibility to RIF relative to wild-type MSM, thereby indicating that the mutated version of the *rpoB* was not dominated by wild-type. In addition, the observation that SCOs are more susceptible to RIF than DCOs suggests stochastic binding of wild-type and mutant RpoBs to DNA in the merodiploid strains. This conclusion is supported by the fact that the mutated version of *rpoB* is integrated at the native *rpoB* chromosomal site, so increasing the likelihood that native and mutated versions of *rpoB* are transcribed, and translated, at similar rates. The factors regulating the production and assembly of RNAP subunit components is not completely understood, however, β-subunit composition of assembled RNAP enzymes is probably a mixture – with some enzymes containing β-

subunit that is sensitive to RIF, and others containing a RIF-recalcitrant β-subunit. At relatively high concentrations of RIF, RNAP functionality is probably preserved by those enzymes that are made up of mutated versions of *rpoB*. As RIF concentrations increase, RNAP enzymes that contain a sensitive β-subunit are likely to be inhibited by RIF. However, it is not clear if this inhibition is adequate to prevent cell growth, and why mutated versions of *rpoB* can't take over function. Recently, Kohanski *et al.* (2009) suggested that bacterial death after antibiotic treatment is a result of oxidative stress mechanisms induced by target inhibition (Kohanski *et al.*, 2010, Kohanski *et al.*, 2008). The precise mechanisms may differ, but inhibition of RNAP may result in off-site toxic stress, which may partly explain inhibition of growth of SCOs at very high RIF concentrations.

4.4 Single S531L or H526Y point mutations in the *rpoB* gene drastically decrease RIF^S: does this infer sufficiency for RIF^R?

Several aspects of the directed mutagenesis and screening strategies employed in this study required several rounds of optimization to increase specificity. However, generation of these mutants, containing RIF^R-associated mutations, in the absence of RIF, allowed evaluation of the contribution of single point mutations in the *rpoB* gene – without second-site mutations in associated or complimentary genes – to RIF^R. RIF^S of *rpoB*^{S531L} and *rpoB*^{H526Y} mutants was significantly decreased in both solid and liquid media (wild-type: 1.95 – 3.9 µg/mL; *rpoB*^{S531L}/*rpoB*^{H526Y} mutant: 250 – 500 µg/mL). In contrast, the *rpoB*^{L521L} mutant displayed susceptibility levels similar to the wild-type. Only the RRDR region of the clinical isolates containing the L521L mutation was sequenced (Yang *et al.*, 1998, Siddiqi *et al.*, 2002), which indicates that either alternative *rpoB* mutations or other physiological mechanisms were involved in RIF^R.

Evaluation of the ability of mutant strains to grow on RIF containing solid media provided the first indication of RIF^S of strain generated in this study. MABA's were applied to determine a more precise MIC_{RIF} for mutant (Palomino *et al.*, 2002, Collins & Franzblau, 1997, Franzblau *et al.*, 1998). The utility of this method arises from the use of 96-well plates which facilitate assessment of growth across a large range of antibiotic

concentrations (the final concentration in a plate, represents 4000-fold dilution of the starting antibiotic concentration). Addition of Alamar Blue aids in evaluation of growth. Alamar Blue contains resazurin, and as growth occurs, oxygen usage by cells allows for reduction of resazurin, which then changes from blue to pink (Mann & Markham, 1998). Wild-type MSM has a relatively high MIC_{RIF} therefore elevated concentrations of RIF were required to evaluate the effect of the point mutation on RIF^R. The deep red colour of RIF made the use of Alamar Blue as a growth indicator problematic, especially in wells where the concentration of RIF was higher. Therefore pellet size was used as an additional growth measure and cell growth from selected wells was evaluated on 7H10 OADC (Appendix 5.2) plates to verify that growth was actually occurring.

The importance of S531L and H526Y mutations for RIF^R was reiterated by results of MABA's (see Section 3.7.3). These assays suggest that the main contributor to RIF^R was the single point mutation in *rpoB*. The fact that these strains were RIF-naïve until MIC determination supports this finding. To broaden the investigation of the sufficiency of RIF^R-associated mutations for resistance, the wild-type allele was restored in the *rpoB*^{S531L} mutant. When the leucine in the mutant RpoB was restored to the wild-type serine, RIF^S dropped to wild-type levels (Figure 3.20). This result provides important and novel evidence for the sufficiency of the S531L mutation for RIF^R. Experiments to restore the wild-type allele in the H526Y mutant are underway and it is likely that results will be similar. Restoration of the wild-type allele in the S531L mutant background provides good evidence that other mutations are not contributing to RIF^R, but does not rule out the possibility of second-site mutations that may compensate for fitness defects.

Structural evaluation of the RIF-RNAP complex shows that RIF binds to RNAP within the DNA/RNA channel (Campbell *et al.*, 2001). Mutations at eleven of the twelve amino acids that are directly involved in RIF binding have been observed in RIF^R organisms (Figure 3.2). These mutations often result in substitution of an amino acid with a small side chain, with an amino acid with a more bulky group, which probably interferes with RIF binding. Amino acids at positions 531 and 526 are involved in hydrogen bonding with RIF, and changes probably disrupt this bonding (Campbell *et al.*, 2001). Campbell *et al.* (2001) suggest that RIF impedes transcription by preventing bond formation between the 2nd

and 3rd trinucleotide in a growing RNA chain. Arstimovich *et al.* (2005) suggest an additional mechanism that may facilitate RIF inhibition of RNAP (Artsimovitch *et al.*, 2005): specifically that the tail of the RIF molecule provides an allosteric signal that disorders binding of the Mg^{2+} ion at the active site, and thereby slows catalysis. Both mechanisms probably play a role in RIF-mediated obstruction of transcription, and RIF^R-associated mutations at different positions in *rpoB* may be affecting different facets of RIF inhibition. Missense mutations that resulted in S531L and H526Y changes resulted in high level RIF^R, while the silent L521L mutation had no effect on RIF^R. Therefore, MIC_{RIF} determination during this study provide further support for the role of substituted amino acids – and decreased RIF binding – in RIF^R.

RNA transcription and DNA replication happen at the same time within a cell. Collision of transcriptional and replicative complexes can impact genetic stability due to replication stalling (Boubakri *et al.*, 2010). It is not clear what happens to RNAP after RIF inhibition. Does the enzyme disassociate and re-associate somewhere else on the chromosome? Is the RIF bound β-subunit degraded? If RIF inhibited RNAP increase DNA replication-stalling this could have severe consequences for genetic stability within the cell. Interestingly a recent study shows that *rpoB* mutations (within the RRDR) reduce the need for proteins that resolve stalled replication complexes (Baharoglu *et al.*, 2011). In the context of this study, it's particularly important to consider how *rpoB* mutations impact DNA fidelity. If introduction of *rpoB* mutations affects TCR or other DNA repair mechanisms, and results in increased mutagenesis, isogenicity of mutant strains may be compromised. RIF^R-associated mutations examined in previous studies for their role in TCR modulation are not commonly observed in clinical isolates (Trautinger *et al.*, 2002; Ganesan *et al.*, 2007; Baharoglu *et al.*, 2011). This does not, however, negate the need to consider the implication of RIF^R-associated mutations for isogenicity of mutant strains generated in this study. The implications of RIF^R-associated mutations on DNA repair provides an interesting area to explore in future studies. RIF^R-associated mutations in *rpoB* have pleiotropic effects for mechanistic and regulatory aspects of transcription – from transcription initiation (Campbell *et al.*, 2001) to pausing termination (Jin and Gross, 1998; Jin and Gross, 1991; Ingham & Furneaux, 2000), and regulation (Hu *et al.*, 2002, Xu *et al.*, 2002). In addition to impacting RIF binding and RIF^R, structural and mechanistic

implications of *rpoB* mutations are likely to be responsible for important physiological changes.

4.5 Are there other contributors to RIF^R?

Up to 95 % of RIF^R clinical isolates contain mutations in the RRDR (Telenti *et al.*, 1993, Ramaswamy & Musser, 1998). However, the genetic basis for resistance in the remaining strains remains unclear. In many studies, only the region of the *rpoB* gene containing the RRDR is sequenced, despite evidence that mutations in other regions of the *rpoB* gene can be associated with RIF^R. It is, therefore, possible that 5 % of RIF^R strains with no mutations in the RRDR may have mutations in other regions of *rpoB* (Heep *et al.*, 2001, Heep *et al.*, 2000b). The implications of these mutations for RNAP function, and bacterial physiology, are likely to be vastly different when compared to mutations found in the RRDR. EMB is an important first-line TB drug that is thought to interfere with cell wall biosynthesis (Belanger *et al.*, 1996). Introduction of an EMB^R-associated mutation into the MTB chromosome conferred some levels of resistance but not to the high levels seen in clinical isolates (Safi *et al.*, 2010). Such evidence indicates that there may be other mechanisms that play a role in resistance to other clinically relevant anti-TB drugs. Several factors are important for the ability of a drug to execute its effects on a target, including the rate of passive diffusion of the drug into the cell, the operation of active efflux pumps which serve to reduce the intracellular concentration of toxic compounds, enzymatic antibiotic inactivation mechanisms, and the intracellular concentration of the target. These are factors important in determining drug efficacy and are also therefore important for the development of drug resistance (Fange *et al.*, 2009) (Figure 4.1).

We were particularly interested in the role of efflux in RIF^R from a physiological perspective. There are several reports that indicate that active efflux systems are required for the development of resistance-associated mutations (Ricci *et al.*, 2006, Fange *et al.*, 2009, Lovmar *et al.*, 2009). Evaluation of mycobacterial genomes indicates that there is the potential for expression of several types of efflux systems (Louw *et al.*, 2009). It is likely that efflux systems in mycobacteria largely play a role in maintaining physiological homeostasis, and protecting the bacillus from the accumulation of toxic

compounds (Louw *et al.*, 2009) and extrusion of drugs may be an adaptive advantage conferred by some of these systems. In addition to extrusion of toxic substances – such as drugs – mathematical modelling (Fange *et al.*, 2009) and results from *Salmonella* experimental systems (Ricci *et al.*, 2006) suggest that activated efflux systems are absolutely required to 'prime' bacteria to generate and fix resistance-associated point mutations. In MTB, experiments indicate that inhibition of efflux decreases the mutation rate of EMB^R, and that efflux may facilitate concurrent development of resistance to EMB and INH (Srivastava *et al.,* 2010).

Inclusion of EPI's, verapamil and reserpine, in MABA plates had insignificant effects on RIF^S of wild-type or mutant strains. Reserpine inhibits ATP-dependant pumps, while verapamil impedes P-glycoprotein systems (Lechner *et al.*, 2008), and it is possible that mycobacterial efflux systems are recalcitrant to these compounds. Another possibility is that other extrusion systems compensated for inhibition of verapamil/reserpine sensitive efflux pumps. Reserpine and verapamil inhibit MSM efflux of ethidium bromide 2 – 4 fold (Lechner *et al.*, 2008). However, exposure to reserpine only slightly increases accumulation of radio-labelled RIF by MSM cells (Piddock *et al.*, 2000). The MABA method may not be sensitive enough to detect a change in efflux. However, if efflux systems were playing a significant role in resistance, inhibition should have detectable influence on MIC. MIC_{RIF} determination in the presence of EPI's indicates that efflux pumps (at least those inhibited by verapamil or reserpine) do not significantly contribute to RIF^R in MSM. However, this does not exclude a role for active efflux system in generating and fixing drug resistance-associated point mutations in target genes and remains to be explored.

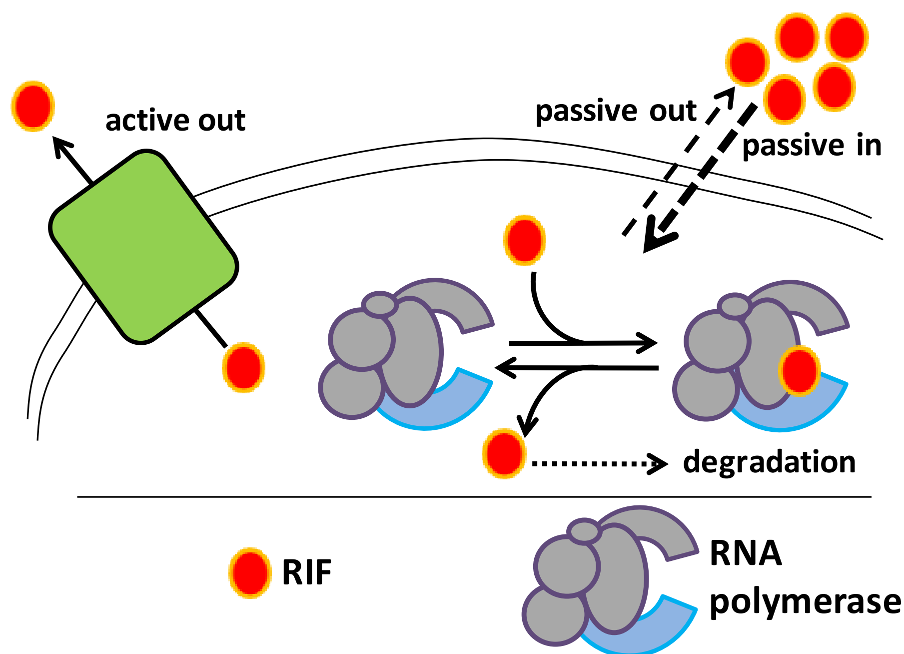


Figure 4.1 Factors important in drug efficacy and the development of drug resistance.

This figure was adapted from (Fange *et al.*, 2009) to reflect to the factors influencing the effective concentration of RIF required to inhibit bacillary growth by binding RNAP. Drug efficacy can be undermined at a number of levels: First, the drug needs to get into the cell, which often happens by passive diffusion (passive influx of the drug into the cell is indicated by thinker black arrows, while potential passive diffusion of the drug out of the cell is indicated by thinner black arrows). The efficacy of many drugs may be hindered at this step. The drug can be inactivated or degraded, a mechanism which decreases RIF^S in some mycobacteria, and related genera (Quan *et al.*, 1997). The drug needs to bind to the target, and the intracellular concentration of the drug required to inhibit all enzyme activity depends on the intracellular concentration of the target. Finally, intracellular drug concentration is likely to play a role in the development of target mutations (Fange *et al.*, 2009, Gagneux, 2009), and can be lowered by activate extrusion of drug by active efflux systems.

4.6 Expression of *rpoB* from an episomal plasmid is toxic in MSM: Does this impact RNAP subunit stoichiometry?

It has been suggested, and observed in some systems, that over-expression of an antibiotic target may contribute to resistance to the antibiotic (Kurth *et al.*, 2009, Huber *et al.*, 2009, McMurry *et al.*, 1998). To investigate the effect of gene dosing on RIF^R, wild-type, S531L and H526Y versions of the *rpoB* gene were cloned onto pOLYG – an episomal plasmid (O' Gaora *et al.*, 1997). Low electroporation efficiencies of pG*rpoB* constructs implied that expression of *rpoB* may from this vector may be toxic in MSM. Upon further investigation, significant re-arrangement was observed for pG*rpoB* constructs carrying both mutant and wild-type *rpoB* genes – an indication that plasmids were unstable upon introduction to MSM. The pOLYG vector exists in the cell in up to 23 copies (Huff *et al.*, 2010), and therefore expression from pOLYG is likely to confer increased expression of *rpoB.* Therefore toxicity of these plasmids may be related to over-expression of *rpoB*. In our system, electroporation efficiency was extremely low, and the plasmids recovered from the handful of HYG^R appeared to have undergone significant rearrangement, therefore *rpoB* expression levels could not be measured in these isolates. Similar findings were reported by Zaczek *et al.* (2009) who introduced mutated versions of *rpoB* into MTB on replicating vectors. However, they also reported constitutive expression of *rpoB* from an Hsp60 promoter that was integrated at the *att*B site. Little detail was provided for these strains and it is unclear whether *rpoB* was expressed from Hsp60 – RIF^R which may have indicated expression was not observed, and *rpoB* expression levels in these strains were not determined (Zaczek *et al.*, 2009).

A separate group did not observe similar toxicity when mutated versions of MTB *rpoB* gene were expressed from multi-copy episomal vectors (Williams *et al.*, 1998). It's possible therefore, that the particular episomal expression system used in ours and the Zaczek *et al.* (2009) study induces unfavourable levels of *rpoB* expression, while the system used in the Williams *et al.* (1998) study did not. Early reports indicate that expression of *rpoB* is regulated at the translational level. That is, if the *rpoB* gene were over-expressed without simultaneous increase of other RNAP components, translation of these subunits was downregulated (Passador & Linn, 1989). Subsequent studies have

shown regulation of *rpoB* expression at the transcriptional level (Dykxhoorn *et al.*, 1996). The inferred toxicity of *rpoB* in multi-copy seems to make intuitive sense: the product of *rpoB*, the β-subunit of RNAP, is part of an essential, multi-subunit enzyme whose components must be present in stoichiometric amounts.

4.7 Implications of RIF^R-associated mutations for mycobacterial physiology

At inception, the primary aim of this project was to generate strains containing RIF^R associated mutations, in the absence of RIF, so that the physiology of these strains could be evaluated without possible bias from second-site mutations. Studies in other organisms indicate that *rpoB* mutations, analogous to those found in clinical MTB isolates have broad implications for bacterial biology. In *B. subtilis*, for example, these mutations had broad implications for sporulation, germination and even carbon source metabolism (Maughan *et al.*, 2004, Perkins & Nicholson, 2008). In MTB, lowering the pH of a chemostat culture decreased the frequency with which mutants harbouring an S531L mutation were isolated (Jenkins *et al.*, 2009). This is significant, given that MTB is exists in acidic compartments with macrophages (Vandal *et al.*, 2009, Vandal *et al.*, 2008). Moreover, mutations in the *rpoB* gene can induce production of novel secondary metabolites in *Streptomyces* sp. (Hosaka *et al.*, 2009, Hu *et al.*, 2002, Xu *et al.*, 2002). Mutations in essential and highly conserved targets are expected to result in some cost to biological function of these proteins (Andersson & Levin, 1999, Andersson, 2006). Given that *rpoB* is highly conserved, and critical for bacterial function, it has long been suggested that RIF^R-associated changes in this gene would alter the functioning of RNAP, and therefore result in a fitness cost in the absence of RIF (Gagneux, 2009, Gagneux *et al.*, 2006b, Andersson, 2006, Andersson & Hughes). It is impossible to uncouple fitness from biological processes and therefore an understanding of the biological effects of drug resistance mutations is important for better TB management strategies, as well as well-informed drug discovery approaches.

4.7.1 Carbon source-specific effects for the MIC_{RIF} of wild-type but not mutant strains

The use of the MABA's allowed investigation of two central questions of this project: specifically, the potential physiological implications of RIF^R-associated mutations and the contribution of single *rpoB* point mutations to RIF^S. Antibiotic efficacy is dependent on the growth state of an organism (Xie *et al.*, 2005, Jindani *et al.*, 2003, Lipsitch & Levin, 1997). This has proved a significant challenge for TB control, since MTB can persist in the host for extended periods, where altered metabolism is one of the factors that lead to increased antibiotic tolerance (Barry *et al.*, 2009, Honer zu Bentrup & Russell, 2001). For an infectious pathogen, physiology will be affected by numerous factors including activity of host immune effectors, oxygen availability, and carbon source and co-factor availability (Warner & Mizrahi, 2007). All of these combine to form a network to increase or subvert antibiotic efficacy. We were particularly interested in the implications of carbon source metabolism for RIF^R because of evidence to suggest *rpoB* mutations drastically alter carbon source metabolism in *B. subtilis*. Mutations in *rpoB* – similar to those observed in clinical MTB strains – confer the metabolic ability to subsist on a variety of carbon sources not previously considered suitable for this growth of this organism (Perkins & Nicholson, 2008).

Assay of wide-ranging carbon sources was beyond the scope of this project therefore, MTB-relevant carbon sources were included in MABA assays. A number of recent studies indicate that, during infection, MTB probably subsists on lipid based carbon sources (Pandey & Sassetti, 2008, Savvi *et al.*, 2008, Brzostek *et al.*, 2009, Marrero *et al.*, 2010). Therefore MABA's were conducted in lipid based-media and the resultant MIC_{RIF} compared to those obtained in standard mycobacterial growth media. Interestingly, wild-type strains exhibited a 4-fold increase in MIC_{RIF} when grown in lipid containing media, while the MIC_{RIF} was lowest in GS containing media. From these assays, carbon source did not appear to have an effect on MIC_{RIF} for mutant strains. Further investigation is required, but it's possible that due to very high MIC_{RIF} observed for these strains, carbon source specific effects were not discernable. Nutrient usage of MSM is not well defined, and although MSM contains some cholesterol catabolising genes (Uhia *et al.*, 2011), it's not clear that MSM preferentially subsists on fatty-acids – as is suggested for MTB. It's possible that in MSM carbon source composition of the growth media results in altered

cell wall composition, which may decrease diffusion of RIF into the cell. These two possibilities are not mutually exclusive, and the physiological basis of this observation warrants further investigation. Inclusion of a broader range of compounds will form an interesting part of the future of this project. The Biolog Phenotype Microarray™ system (http://www.biolog.com/phenoMicro.shtml) allows evaluation of thousands of different phenotypes in a single run and was used by Perkins and Nicholson (2008) to uncover *rpoB*-associated metabolic capabilities in *B. subtilis* (Perkins & Nicholson, 2008). Although there are cost considerations, this system would be potentially useful to evaluate the effect of carbon source on MIC_{RIF}, and the physiological implications of RIF^R-associated mutations in mycobacteria.

Discordant drug-susceptibility results have been reported despite the use of standardized methodology for tests (Van Deun *et al.*, 2009). In this study, the MIC_{RIF} of wild-type MSM was notably higher in lipid-based carbon sources, which demonstrate how assay conditions can have significant implications for MIC values. This should be considered when evaluating resistance levels of clinical isolates. In these experiments, only one condition was altered, and it is likely that the adjustment of other conditions would also affect susceptibility levels.

4.7.2 A slight fitness defect for *rpoB*^{S531L} is revealed by competitive growth experiments

Mariam *et al.* (2004) observed that in three different assay systems (in axenic culture; *in vitro* competitive growth, and in macrophages), the inferred fitness of specific RIF^R mycobacterial strains varied depending on the assay employed. For example, in competition with wild-type, an H526Y mutant had the lowest fitness cost when compared to S531LY or S522L mutants, however in axenic culture S522L mutant displayed the lowest fitness cost. In a highly cited study, Gagneux *et al.* (2006b) show that the fitness of RIF^R strains is variable, and is dependent on the particular mutation, as well as the strain background. However, strains used in the aforementioned studies had been exposed to RIF prior to their assessment in these assays, and so it is possible that these strains had adapted to growth with a mutant RpoB, and perhaps even contained second-site mutations that could have contributed to any fitness phenotype observed. The

generation of RIF-naïve strains in this project, with a defined natural history (*in vitro* processing was kept to a minimum), facilitated physiological evaluation of RIF^R-associated mutations.

Fitness can be simply defined as the ability of an organism to survive and replicate (Andersson & Hughes, 2010), in a particular environment. Biological processes greatly impact the ability of the bacilli to do these things under a plethora of pressures that include host immune function, and antibiotic treatment (Warner & Mizrahi, 2007). Competition assay provided the opportunity to investigate the fitness of mutants, which would provide some indication of the degree of physiological effects. The use of RIF had been avoided at all costs during generation of *rpoB*^{MT} strains. However, the high MIC_{RIF} for *rpoB*^{S531L} (250 µg/mL – 500 µg/mL vs. 1.95 – 3.9 µg/mL for wild-type), permitted the use of RIF-containing plates (200 µg/mL RIF – RIF_{200}) to differentiate between wild-type and mutant strains. Therefore fitness was assayed as a function of growth and measured by colony number. A modest relative fitness defect was observed for the *rpoB*^{S531L} strain (Figure 21), which is similar to previous findings for MTB with the same mutation (Gagneux *et al.*, 2006b, Mariam *et al.*, 2004). However, several observations (discussed in the next Section) raise important questions regarding the suitability of RIF^R to distinguish between wild-type and mutant strains. Nutrient conditions faced by MTB during infection, or even MSM in the soil, are likely to differ fundamentally from lab conditions. Other conditions which aim to simulate factors prevalent *in vivo*, for MSM and MTB, could be incorporated into the competition assay. Competition experiments between strains containing other RIF^R-associated mutations, and the parental strain would also be informative. Further competition assays, using the strains generated in this project, are likely to yield interesting results, and have been included in design of the future of this project. Competition between strains containing different RIF^R-associated point mutations, or between strains containing the same mutations but selected in different ways would be especially interesting. For these assays RIF^R could not be used for strain differentiation – this underscores the need for a better strain differentiation system. It is anticipated that these experiments will be conducted in the near future.

Strain background can significantly impact observed fitness, for example an H526D mutation was associated with a lower fitness cost when introduced into the MTB T85 strain, than CDC1551 (Gagneux *et al.*, 2006b). Differences in observed fitness may be due to indirect biological effects of second-site mutation and these mutations may not be compensatory *per se*. Compensatory mutation, however, has important implications for physiology, and spread of drug-resistant strains. In *S. typhimurium*, direct compensation of functional disadvantages as result of STR^R can be ameliorated by secondary mutations (Bjorkman *et al.*, 1998, Bjorkman *et al.*, 1999). Furthermore, the type of compensatory mutation observed was dependent on the assay system used. Whole-genome re-sequencing of the strains generated was beyond the scope of the project, but would provide an interesting exploration of possible compensatory mutation after growth in stressful conditions. Epistatic interactions between RIF^R- and STR^R-associated mutations have been reported in *E. coli* (Chakrabarti and Gorini, 1977; Trindade *et al.*, 2009). Moreover it was recently shown that decreased requirement for RNAP, reduced fitness costs associated with RIF^R in *P. aeruginosa* (Hall *et al.*, 2011). Treatment of RIF^R strains with ribosome inhibitors, STR and chloramphenicol, lessened fitness costs as measured by growth. Treatment with ciprofloxacin, a DNA gyrase inhibitor did not have the same effect, which indicates that better growth of RIF^R mutants was not due to generalised antibiotic stress, or decreased growth (Hall *et al.*, 2011). STR and RIF are often used to treat MTB infected patients at the same time, and an investigation of the epistatic interactions between mutations associated with these drugs would be especially important for MDR-TB control.

4.8 Complexity of fitness assay and physiological characterisation

A core assumption of a competition assay is that strains maintain their fundamental identity. In other words, with regards to this study, that the *rpoB*^{S531L} retained the S531L mutation and the wild-type retained the wild-type allele. This, therefore, allows use of a selective agent to differentiate between the two. One possible deviation from this is spontaneous mutation of wild-type to mutant. In view of studies which suggest that mutation rate increases during stationary phase (Foster, 1998, Karunakaran & Davies, 2000, Cairns *et al.*, 1988, Bjedov *et al.*, 2003, Rosenberg & Hastings, 2004), mutation of

the wild-type strain was a concern for us. Many studies use *rpoB* mutations associated with RIF^R to study mutation rate, and since we used RIF to differentiate between wild-type and $rpoB^{S531L}$ strains, we were particularly concerned that colonies on RIF plates from may originate from mutated wild-type strains. To monitor this, wild-type cells from axenic culture were also plated on RIF_{200} plates. Our concerns were eliminated since no RIF^R mutants were observed on RIF_{200}, from axenic wild-type cultures, for the duration of the experiment. The "lack" of increased RIF^R mutation in stationary phase has been separately observed at the MMRU (Sam Barichevy, unpublished results; and Digby Warner, unpublished results). No RIF^R-specific mutation in the wild-type strain was a positive result – with regards to methodology for the competition assay, however, it contrasts to what was previously observed for MSM (Karunakaran & Davies, 2000). Karunakaran and Davies (1998) observed a steep increase of MSM colonies on RIF plates which they attributed to increased mutation rate during stationary phase. The differences between our study and that of Karunakaran and Davies could stem from the use of different MSM strains (in the Karunakaran and Davies study MSM mc^26 strain was used, while in this project MSM mc^2155 was used). Another notable difference is that our experiments were conducted using 7H9 OADC media, while the aforementioned group used Tryptic Soy Broth. Although an increase in RIF^R mutants during stationary phase growth is often attributed to elevated mutation rates; from an elegant set of molecular and microbiological experiments Wrande *et al.* (2008) suggest that *rpoB* mutations provide an advantage during stationary growth conditions and therefore outgrow parental strains. Carbon source specific differential growth has been observed for UV induced RIF^R MSM mutants (Digby Warner and Susan Puckett, unpublished results). An investigation of whether the differences between results generated in this study (and others at the MMRU), and those of Karanakarun and Davies (2008), are due to media-specific factors and resultant physiological impact for RIF^R mutants, presents an interesting experiment for the future.

Another assumption of a competition assay is that the strains remain receptive to the selective compounds used to differentiate between them. That is, that the wild-type strain remains RIF^S and that $rpoB^{S531L}$ strains remain resistant to RIF. AS-PCR revealed that by day 20 of the assay, 15 – 30 % of strains in the competition assay were of the $rpoB^{S531L}$

genotype, even when no colonies were present on RIF_{200} plates (after extended incubation) (Figure 3.22). Only 20 colonies were examined and genotyped by AS-PCR. Given the low numbers of colonies examined, it is possible that there was some sampling bias when selecting colonies genotype. This may explain why the numbers of mutant colonies actually go up by 3% from the 9 hour measurement to day 7 (Figure 22). However, there is a total lack of colonies on RIF_{200} plates at 20 days and at 48 days, while on standard OADC, AS-PCR indicates a considerably higher number of colonies are present in the culture. This observation has significant implications for the observed fitness of the S531L mutant and suggests that sub-population *rpoB*^{S531L} had lost the ability to grow on RIF. Additionally, an aliquot from the competition culture grew in RIF_{200} 7H9 OADC liquid media, albeit at a particularly slow rate (see Section 3.8.1). These observations require further investigation and raise important questions: If the *rpoB*^{S531L} was resistant to high levels of RIF going into the competition experiment, why upon isolation from competition, did a sub-population of exhibit decreased culturability on RIF? Why were some *rpoB*^{S531L}cells not able to grow on solid RIF_{200} media, but grew in equivalent liquid media?

Finding a better way to differentiate between strains in a competition assay will be an integral part of any future study. However, this will be challenging. Introducing resistance to a separate antibiotic as a marker is not ideal, as expression of the resistance gene may impact physiology in unknown ways, and introduce additional bias. Engineering of auxotrophy is also not perfect, and could indirectly impact fitness. A possible solution would be to introduce a silent point mutation, into a well characterized but non-related gene (“genome graffiti”). However differentiating colonies would still be cumbersome as sequencing or PCR characterization would be required. Fluorescently labeling one of the strains would make for easy and visually pleasing differentiation however colorimetric tags may impose some burden on the cell as they require extra resources to be produced and potential toxicity (Shaner *et al.*, 2005, Rang *et al.*, 2003). Unfortunately the issue of strain differentiation is a difficult one, and it is possible that all feasible options could influence physiology.

As evidenced by the preceding discussion, fitness of assay is complex and often fraught with possible bias. Development of an appropriate assay to measure the potential fitness and physiological implications of RIF^R-associated mutations remains a central aim of this work. Other model systems could be used to evaluate relative fitness, for example, growth in macrophage cells or in mice (strain differentiation would still be an issue in these models). Moreover, a targeted evaluation of physiological implications may also prove useful and an investigation of the consequences for the mechanistic aspects of transcription, such as promoter or sigma factor binding, is another potentially interesting avenue of investigation. For example, MSM RNAP containing a 30 bp deletion (Table 4.1) was refractory to transcriptional inhibition by RIF concentrations up to 1 mg/mL (Malshetty *et al.*, 2010). This mutation has been isolated on multiple occasions at the MMRU (Digby Warner, unpublished results) and SCOs containing this mutation have been generated (Table 4.1). From a physiological perspective, viability of a strain containing such a large deletion in RpoB raises interesting questions. MSM seems to have the ability to "loop out" this region of *rpoB*, the mechanism of mutagenesis for this kind of deletion is therefore potentially interesting from a DNA metabolism perspective.

4.9 A pathway to clinical relevance?

Multiple lines of evidence hint at the tantalizing possibility that drug resistance-associated mutations in target genes might influence fundamental physiological processes such as metabolism and regulation of the immune system – and, furthermore, might even confer beneficial capabilities in the absence of antibiotic pressure (Gagneux *et al.*, 2006, Maughan *et al.*, 2004, Perkins & Nicholson, 2008, Jenkins *et al.*, 2009, Wrande *et al.*, 2008). A primary objective of this project was to investigate this possibility through the application of genetic tools and microbiological techniques that constitute the core expertise of the MMRU. The construction by allelic exchange of a panel of mutants comprising *all* known drug resistance mutations and their evaluation in phenotypic assays is not feasible (Hazbon *et al*, 2008). Moreover, as noted throughout this thesis, *in vitro* assay systems of fitness are fraught with possible bias, and the results of fitness assessments likely to be influenced by numerous factors including the genetic background of a particular mutant strain. For this reason, we were determined to

complement our microbiological approach with other techniques. As whole-genome sequencing technologies become more sophisticated, and cheaper, an ever increasing amount of MTB sequence data is becoming available. As previously mentioned, there may be pressures which act during strain isolation and processing – this is of particular concern when working with model strains. The ever growing collection of sequence data for clinical isolates offers the prospect of studying clinical isolates that have undergone the least possible processing, and therefore minimal selection. This provides the closest approximation of conditions faced within the host. To get any closer, in-host sequencing would be required – which presents an exciting goal for MTB genomics.

Whereas classical molecular epidemiology for MTB is based primarily on DNA finger printing techniques, phylogenetic analyses using whole-genome sequence data are increasingly being utilized to supplement molecular epidemiological investigation (Motiwala *et al.*, 2010, Ioerger *et al.*, 2009, Ioerger *et al.*, 2010). For example, in a recent study, Ioerger *et al* (2010) used whole-genome sequence from MDR-TB clinical isolates in the Western Cape, to show that RIF^R-associated mutations appeared on multiple phylogenetic branches, indicating that these mutations were individually acquired. The information garnered from studies of this kind can be used to infer physiological characteristics of MTB. However, these studies are largely focused on investigating the transmission dynamics of MTB, the understanding of which is important for TB control. In contrast, our aim is to exploit sequence data to understand some of the implications of drug resistance mutations for mycobacterial physiology.

Evolutionary analyses can offer some insight into of how an organism (or functional units of an organism) has adapted to – and survived – the myriad of pressures it faces during its existence (Toft & Andersson, 2010). For an obligate intracellular pathogen, such as MTB, evolutionary changes (as inferred from sequence polymorphisms) can also elucidate the nature of the pressures imposed by its human host and, in so doing, illuminate aspects of host-pathogen interactions. One technique to investigate the types of constraints on a gene or region of a gene is to measure the ratio of non-synonymous to synonymous mutations (dN/dS) at a given site (Toft & Andersson, 2010, Koonin & Wolf, 2010). The dN/dS ratio is commonly used to assess adaptive dynamics in protein sites, and its

effectiveness is born out of the redundancy of the genetic code (Koonin & Wolf, 2010). A greater proportion of synonymous mutations at a given site, suggests the site is subject to purifying selection, whereas an increased preponderance of non-synonymous mutations indicates positive selection (Toft & Andersson, 2010, Koonin and Wolf, 2010). As might be expected, an analysis of this type does not come without its own caveats (Toft & Andersson, 2010). Due to small population sizes, and strong biases with regards to nucleotide composition, constraints on protein sequence may be relaxed in obligate intracellular pathogens. Resulting dN/dS values may therefore be exaggerated due to increased fixation of slightly deleterious mutations (Toft & Andersson, 2010). Moreover, increased fixation of slightly deleterious mutations in mycobacteria is a concern – in closely related organisms these mutations are less likely to be lost due to purifying selection (Hershberg *et al.*, 2008). Nevertheless, in full cognizance of the possible biases and confounding limitations of this approach, we have begun an evaluation of the dN/dS parameters at sites in *rpoB* that are associated with drug resistance, and those that are not. It is hoped that this kind of analysis will enhance and complement the conventional mycobacteriological techniques that comprise the bulk of this thesis.

Using genome comparisons, Hershberg *et al* (2008) uncovered previously under-appreciated levels of genomic diversity in MTB. Moreover, these authors concluded that much of the observed diversity had functional consequences and, further, that MTB was under reduced purifying selection, which is partially a result of the lack of horizontal gene transfer in MTB and extreme bottle necks during transmission (one bacilli is thought to be able to initiate an infection). In addition, they suggested that MTB lineages have some level of co-evolution with the populations they infect which, together with compartmentalisation within the host, leads to a highly segregated population structure. All of these factors function together to reduce effective population size and amplify the consequences of random genetic drift (Hershberg *et al*, 2008). A subsequent report by the same group suggested that regions of the MTB genome encoding known T-cell epitopes are more conserved than some essential genes (Comas *et al.*, 2010). This finding is counter-intuitive as one would expect antigenic variation to play a role in immune evasion. In both investigations, dN/dS analysis provided a useful tool, and the findings of this group motivated our interest in these questions, and in the potential application of

this methodology to *rpoB*. This is not without precedent: the dN/dS technique formed part of the methodology used by Gutierrez *et al.* (2005) to consider ancient ancestry of the MTB complex.

We intend to apply similar techniques to an evaluation of the emergence of drug resistance. Gutierezz *et al.* (2005) suggest that MTB emerged 35 000 years ago, antibiotics are therefore a relatively new selective pressure that the organism faces in its environment. How has MTB adapted to these new pressures, and if RIF^R-associated mutations have additional physiological consequences, might these mutants have existed before the application of RIF for TB treatment? If the emergence of RIF^R associated mutations is purely the result of antibiotic selection, one would expect regions of *rpoB* involved in RIF binding to have a higher dN/dS value than other sites. These values can be compared to other regions of *rpoB* that are functionally important and to other regions of the genome. In the longer term, values for *rpoB* can be compared to values for genes involved in resistance to other drugs, and to genes where selective dynamics are known.

To assist us with this analysis we approached Dr Pierre Durand, who recently established The Evolutionary Medicine Unit at the University of the Witwatersrand (Wits). Dr Durand has agreed to facilitate this analysis, and has already imparted valuable conceptual and technical knowledge. Computation of the dN/dS ratio requires high quality *rpoB* sequence data. Professor Rob Warren, who is a senior researcher at the sister CBTBR unit at Stellenbosch University, kindly provided whole-gene *rpoB* sequence from 22 clinical isolates from the Western Cape. Data from the TB Diversity Sequencing Project, which was set up by Sebastian Gagneux and Peter Small via the Broad Institute, is publically available on the TBDB database and is included in our analysis. The TB Diversity Sequencing Project sequence data includes strains from a variety of geographical regions. A total of 60 strains are available for analysis. Exploitation of sequence data provides an opportunity to examine the biology of drug-resistant clinical MTB strains and adds an additional dimension of clinical relevance to this project.

4.10 Concluding remarks and prospective outlook

The major motivation for this study was to gain some understanding of the physiological implications of RIF^R-associated mutations in mycobacteria. A panel of MSM strains containing clinically relevant RIF^R-associated mutations was generated in the absence of RIF. A combinatorial strategy of SDM and allelic exchange was selected so that likelihood of second-site mutation, which was a potential limitation in previous studies (Gagneux *et al.*, 2006b, Mariam *et al.*, 2004, Gillespie *et al.*, 2002), was minimised. Moreover, this methodology also allowed us to investigate a related issue, that is, the ability of single *rpoB* mutations to confer high level RIF^R, without second-site mutations. High level of RIF^R (MIC_{RIF} = 250 – 500 µg/mL) was observed for MSM strains containing S531L and H526Y mutations (Section 3.7.3). Moreover this is the first study, to our knowledge, to demonstrate that restoration of a wild-type allele in a RIF^R strain returns susceptibility to wild-type levels – as demonstrated for the $rpoB^{S531L}$ strain (see Section 3.7.5). Even though MSM is a useful study organism to investigate mycobacterial physiology, a necessary and natural follow up experiment is to introduce similar mutations in MTB and confirm the RIF^R phenotype in this clinically relevant organism. During this project, established genetic manipulation techniques were optimised to introduce drug resistance-associated mutations into the mycobacterial chromosome in the absence of selection. This will greatly assist during generation of analogous MTB mutants.

A moderate fitness cost was observed for the S531L mutant when compared with the wild-type strain. However, several observations suggest that more $rpoB^{S531L}$ cells were present in the competition assay, than was indicated by growth on RIF plates (see Section 3.8). These results exemplify just how challenging evaluation of fitness is. A fundamental component of this project, in future, will be to develop a defined assay to evaluate fitness of strains generated using methodology developed in this project. An improved method for differentiation of wild-type and mutant strains is crucial but, as discussed previously, development of such system is likely to be tricky. A carbon-source specific effect was observed for MIC_{RIF} of wild-type strains (described in Section 3.7.6). MIC_{RIF} evaluation was conducted in a limited panel of MTB-relevant carbon sources, and it would be interesting

to expand the panel of carbon sources evaluated – perhaps using The Biolog Phenotype Microarray™ system.

Every effort was undertaken to minimise second-site mutation during mutant generation. However, mutant strains were not re-sequenced and therefore the presence of second-site mutations can't be formally ruled out. In addition to re-sequencing of strains after mutant generation, strains could be sequenced at the start of, and at different time points, throughout a competition assay. Although technically demanding, this would allow elucidation of compensatory mutations, if present. Strain background can significantly impact the fitness costs of RIF^R mutants (Gagneux *et al.*, 2006b), and an exciting experiment, planned for the future, involves introduction of RIF^R-associated mutations in defined clinical MTB strains, and using aforementioned assay systems to evaluate potential physiological effects.

In this study we have taken the first step in characterising the physiological implications of RIF^R-associated mutations. During this process, additional biological questions have been exposed, that will make for interesting and important further study. Moreover, novel evidence for the sufficiency of *rpoB* mutations for RIF^R has been generated. Perhaps the most important outcome of this study, however, is that it generated a system whereby the physiological consequences of common drug resistance-associated mutations can be evaluated in a variety of strain backgrounds, including those that are clinically relevant.

5. Appendices

5.1: RIF^R-associated mutations included in this study

Table 5.1.1 Literature reviewed to determine the frequency of RIF^R-associated mutations in clinical and laboratory-derived MTB isolates

Paper	Region	Most common	2nd most common	3rd most common	Most rare
(Yuen *et al.*, 1999)	Australia	S531L tcg→ttg	H526Y cac→tac	D516Y gac→tac	S522Q tcg→cag
(Yang *et al.*, 1998)	Western Japan	S531L tcg→ttg	H526Y cac→tac	D516V gac→gtc	L521L ctg→ ttg
(Ramaswamy & Musser, 1998)	Various	S531L tcg→ttg	H526Y cac→tac	D516V gac→gtc	N518H aac→cac
(Mani *et al.*, 2001)	India	S531L tcg→ttg	H526D cac→gac	L511P ctg→ccg	L533L ctg→ctt
(Hwang *et al.*, 2003)	Taiwan	S531L tcg→ttg	H526Y cac→tac	D516V gac→gtc	L533P ctg→ccg
(Matsiota-Bernard *et al.*, 1998)	Greece	S531L tcg→ttg	H526D cac→gac	H526Y cac→tac	F508L ttc→ttg
(Williams *et al.*, 1998)	Various (mostly USA)	S531L tcg→ttg	H526Y cac→tac	H526Y cac→tac	
(Qian *et al.*, 2002)	East Asian countries	S531L tcg→ttg	D516V gac→gtc	H526Y cac→tac	Q513L caa→cta
(Pozzi *et al.*, 1999)	Italy	S531L tcg→ttg	H526D cac→gac	D516V gac→gtc	
(Gagneux *et al.*, 2006b)	Euro-American lineage	S531L tcg→ttg	H526Y cac→tac	S522L tcg→ttg	H526P cac→ccc
	East-Asian lineage	S531L tcg→ttg	H526D cac→gac		
(Ahmad *et al.*, 2002)	Kuwait and Dubai	S531L tcg→ttg	H526D cac→gac	H526D cac→gac	D516V gac→gtc
(Ahmad *et al.*, 2000)	Middle east	S531L tcg→ttg	D516V gac→gtc	H526P cac→ccc	
(Aristimuno *et al.*, 2007)	Venezeula	S531L tcg→ttg	D516V gac→gtc		

Table 5.1.2 Frequency of mutations associated with RIF^R in laboratory-derived RIF^R strains

Paper	Strain	Most common	2nd most common	3rd most common	Most rare
(Morlock *et al.*, 2002)	H37Rv	S531L tcg→ttg	H526Y cac→tac	S522L tcg→ttg	D516V gac→gtc
(Huitric *et al.*, 2006)	Beijing (B) or Non- Beijing (NB)	S531L tcg→ttg	H526Y cac→tac	H526P cac→ccc	N519L aac→aag
	Non- Beijing	S531L tcg→ttg	H526Y cac→tac	S522L tcg→ttg	L533P ctg→ccg
(Gagneux *et al.*, 2006b)	Euro-American lineage	H526Y cac→tac	S531L tcg→ttg	H526R cac→cgc	Q513L caa→cta
	East- Asian lineage	S531L tcg→ttg	H526Y cac→tac		
(Mariam, 2004)	Harlingen	H526Y cac→tac	S522 tcg→ttg	NO *rpoB* MUTATION BUT RESISTANT TO >32ug/ml RIF 37%	S531W tcg→tgg
MMRU (Warner *et al.*, unpublished data) UV-induced (2008)	South Africa (Various strains of MSM)	H526Y cac→tac	TH526TY acccac→acttac		S522L tcg→ttg
(Billington *et al.*, 1999)	A	H526Y cac→tac			
	B	S531L tcg→ttg			
	C	H526R cac→cga			

5.2 Bacterial growth media and supplements

The recipes described below are for 1 L of media; where smaller amounts were required, amounts were adjusted accordingly. Unless otherwise described, all media were prepared using distilled sterile H_2O (dsH_2O)

Table 5.2.1 *Escherichia coli* growth media

Growth media	Components
Luria Bertani Broth (LB) (liquid medium for *E. coli*)	10 g tryptone 10 g NaCl 5 g yeast extract 1L distilled water Autoclave for 15 mins at 121 ˚C[a]
Luria Bertani Agar (LA) (solid medium for *E. coli*)	10 g tryptone 10 g NaCl 5 g yeast extract 15 g bacteriological agar 1L distilled water Autoclave for 15 mins at 121 ˚C

[a]After autoclaving, media were allowed to cool to 55 ˚C before addition of supplements

Table 5.2.2 *Mycobacterium smegmatis* growth media

Growth media	Components	Rationale for use
Liquid media		
Middlebrook 7H9 Glucose-NaCl (GS)	4.7 g Middlebrook 7H9 (Difco™, USA) 0.2 % glycerol (Merck, Germany) Autoclave for 15 min at 121 ˚C[b] 0.05 % tween80 1 % GS	Middlebrook 7H9 is regularly used for mycobacterial growth analysis. Common supplements include GS and OADC. Both of these supplements contain a sugar as the primary carbon source.
Middlebrook 7H9 Oleic-acid-albumin-dextrose (OADC)	4.7 g Middlebrook 7H9 0.2 % glycerol Autoclave for 15 min at 121 ˚C 0.05 % tween80 10 % OADC	
Middlebrook 7H9 Valerate (Val)	4.7 g Middlebrook 7H9 0.085 % NaCl Autoclave for 15 min at 121 ˚C 0.05 % tween80 0.15 % Valerate 0.5 % Albumin	Work at the MMRU has recently shown that MTB and MSM are able survive on these compounds as the sole carbon source (Savvi *et al.*, 2008). Fatty acids are likely to be the prevalent carbon source during infection (Peyron *et al.*, 2008, McKinney *et al.*, 2000), and therefore the ability of *rpoB* mutants to utilize these compounds will be assessed. Valerate is an odd-chained fatty acid (C3), and butyrate is an even-chained fatty acid (C4). Upon β-oxidation, oddchained fatty acids yield propionyl CoA and acetyl CoA, while those with even chains only yield acetyl CoA (Savvi *et al.*, 2008). The ability of *rpoB* mutants to utilize two different metabolic pathways can thereforebe assessed.
Middlebrook 7H9 Butyrate (But)	4.7 g Middlebrook 7H9 0.085 % NaCl Autoclave for 15 min at 121 ˚C 0.05 % Tween80 0.1 % Butyrate 0.5 % Albumin	

Solid media		
Middlebrook 7H10 GS	19 g 7H10 agar (Difco™, USA) 0.5 % glycerol Autoclave for 15 min at 121˚C[c] 0.05 % tween80 1% GS	Middlebrook 7H9 is regularly used for mycobacterial growth analysis. Common supplements include GS and OADC. Both of these supplements contain a sugar as the primary carbon source[d].
Middlebrook 7H10 OADC	19 g 7H10 agar 0.5 % glycerol Autoclave for 15 min at 121˚C 0.05 % tween80 10 % OADC	

[c]After autoclaving, media were allowed to cool to 55˚C before addition of supplements
[d]Solid media supplemented with fatty acids s were not required for this study

Table 5.2.3 Medium supplements

Supplement	Details	Reason for use
Nutritional supplements		
OADC (Beckton Dickinson, USA)	0.05 % oleic acid-5 % albumin-2 % dextrose-0.004 % catalase Used as received from manufacturer Stored at 4 °C	Supplementation for MSM growth media. The predominant carbon source for mycobacterial growth was derived from these compounds.
Glucose-NaCl	Stock (100X) in distilled H_2O 20 % glucose 8.5 % NaCl distilled H_2O Filter sterilised (0.22 mm membrane) Stored at 4 °C	
Valerate (Sigma Aldrich, USA)	Used as received from manufacturer Stored at room temperature	
Butyrate (Sigma Aldrich, USA)	Stock (10 mg/ml) in dsH_2O Filter sterilised (0.22 mm membrane) Stored at 4 °C	
Polyoxyethylenesorbitan monooleate (Tween80) (Sigma-Aldrich, USA)	Stock (100X) in dH_2O 25 % Tween80 Filter sterilised (0.22 mm membrane) Stored at 4 °C	Supplementation for MSM growth media to prevent cell clumping

Antibiotics		
Kanamycin (KAN) (Roche Applied Sciences, Germany)	Stock (50 mg/ml) in dH_2O Filter sterilised (0.22 mm membrane) Stored at 4 ^oC	Selection of *E. coli* containing specific plasmids or for genetically manipulated MSM strains
Hygromycin (HYG) (Roche Applied Sciences, Germany)	Received at 50 mg/ml Stored at 4^oC in the dark	
Ampicillin (AMP) (Roche Applied Sciences, Germany)	Stock (100 mg/ml) prepared in 50% ethanol Filter sterilised (0.22 mm membrane) Stored at –20 ^oC	
Rifampicin (RIF) (Sigma Aldrich, USA)	Stock (100 mg/ml) in deionised dimethylformamide (DMF) (Sigma Aldrich, USA) Stored at 4 ^oC in the dark	To determine RIF resistance levels of MSM strains
Other selective compounds		
5-bromo-4-chloro-3-indolyl-β-galactoside (Xgal) (Bioline, USA)	Stock (20 mg/ml stock) in deionised DMF Stored at 4 ^oC in the dark	Selection of *E. coli* cells containing specific plasmids or for genetically manipulated MSM strains
Sucrose	Stock (25 %) in dH_2O Stored at room temperature	

5.3 DNA extraction solutions and DNA modifying enzymes

Table 5.3.1 DNA Extraction solutions

Solution	Components
Solutions for plasmid extraction from *E. coli*	
Solution I	50 mM glucose; 25 mM Tris-HCl (pH 8) 10 mM EDTA Autoclave for 10 min at 121 ^{o}C Stored at 4 ^{o}C
Solution II	1% SDS 0.2 M NaOH Autoclave for 10 min at 121 ^{o}C Stored at room temperature
Solution III	3 M potassium acetate 5 M acetic acid No autoclaving Stored at 4 ^{o}C
RNase A (Sigma Aldrich, USA)	Stock (10 mg/ml) made in distilled H_2O Boil for 10 min Stored at -20 ^{o}C.
Tris EDTA buffer (TE)	10 mM tris 1 mM EDTA (pH 8.0)
Solutions for chromosomal DNA extraction from MSM	
N-*cetyl-N,N,N*-trimethyl ammonium bromide (CTAB)/NaCl	4.1 % NaCl 10 % *N*-cetyl-*N*,*N*,*N*-trimethyl ammonium bromide Heat briefly at 65°C to dissolve
Lysozyme	10 mg/ml
Proteinase K	20 mg/ml
DNA purification solutions	
Phenol-chloroform	1:1 v/v solution
Chroloform-isoamyl alcohol	24:1 v/v solution

Table 5.3.2 Restriction endonucleases[a]

Enzyme[a]	Buffer (10X)	Recognition sequence and cleavage site
***Aat*II**	NEB buffer 4: 20 mM Tris-acetate 50 mM potassium acetate 10 mM Magnesium Acetate 1 mM Dithiothreitol pH 7.9 @ 25 °C	5’ GACGT|C 3’ 3’ C|TGCAG 5’
***Acl*I**	NEB buffer 4 100 µg/ml BSA	5’ AA|CGTT 3’ 3’ TT GCA|A
***Acc*65I**	NEB buffer 3: 50 mM Tris-HCl 100 mM NaCl 10 mM $MgCl_2$ 1 mM Dithiothreitol pH 7.9 @ 25 °C 100 µg/ml BSA	5’ G|GTACC 3’ 3’ CCATG|G 5’
***Bam*HI**	NEB buffer 3 100 µg/ml BSA	5’ G|GATCC 3’ 3’ CCTAG|G 5’
***Eco*RI**	Any NEB buffer	5’ G|AATTC 3’ 3’ CTTAA|G 5’
***Hind*III**	NEB buffer 2: 10 mM Tris-HCl 50 mM NaCl 10 mM $MgCl_2$ 1 mM Dithiothreitol pH 7.9 @ 25 °C 100 µg/ml BSA	5’ A|AGCTT 3’ 3’ TTCGAA 5’
***Nae*I**	NEB buffer 4	5’ GCC|GGC 3’ 3’ CGG|CCG 5’
***Nsp*I**	NEB buffer 2 100 µg/ml BSA	5’ RCATG|Y 3’ 3’ Y|GTACR 5’
***Sca*I**	NEB buffer 3	5’ AGT|ACT 3’ 3’ TCA|TGA 5’

[a]All restriction endonucleases were obtained from New England Biolabs, USA

Table 5.3.3 Other DNA modifying enzymes

Enzyme	Buffer (10X)	Functionality
Antarctic Phosphatase (New England Biolabs, USA)	Antarctic Phosphatase buffer: 50 mM Bis-Tris-Propane-HCl 1 mM $MgCl_2$ 0.1 mM $ZnCl_2$ pH 6.0 @ 25 °C	Removal of 5' phosphate groups during cloning
T4 DNA ligase (Roche Applied Sciences, Germany)	Ligation buffer: 660 mM Tris-HCl 50 mM $MgCl_2$ 10 mM dithioerythritol 10 mM ATP pH 7.5 @ 20 ° C	Catalyst for covalent bond formation between fragments during cloning
DNA large fragment I (Klenow)	NEB buffer 2	'Fill-in' of overhangs generated by restriction digestion

Table 5.3.4 DNA polymerases

PCR kit	Buffer	Reason for use
FastStart Taq DNA Polymerase™ (Roche Applied Sciences, Germany)	PCR reaction buffer™ (10X): 500 mM Tris/HCl 100 mM KC 50 mM $(NH_4)_2SO_4$ 20 mM $MgCl_2$ pH 8.3/ 25 °C GC solution™ is useful when amplifying high GC content DNA	Taq DNA polymerase does not have 3' → 5' proofreading ability, and therefore has an error rate of approximately 1.3 per 10^5 bases. This system was used for screening large numbers of colonies.
Phusion High Fidelity DNA Polymerase™ (New England Biolabs, USA)	Phusion High Fidelity PCR buffer™ contains 7.5 mM $MgCl_2$ GC solution™ is useful when amplifying high GC content DNA	This polymerase has a significantly lower error rate of 4.3 per 10^7 bases, and was therefore used when fidelity was paramount.

5.4 Agarose gel electrophoresis

Table 5.4.1 Agarose gel electrophoresis components

Component	Details
Agarose (Invitrogen, USA)	0.3 g agarose in 30 ml TAE buffer makes a 1 % agarose gel
Ethidium bromide (Invitrogen, USA)	Stock (10 mg/ml) made in dsH2O Stored at 4˚C in the dark
Loading dye	0.3 g Bromophenol blue 0.3 g Xylenol 93.6 ml 80 % glycerol 3 ml 0.5 M EDTA Made up to 250 ml with dsH2O Stored at 4 ˚C
tris-acetate-EDTA (TAE) buffer	Stock (50X) prepared as follows: 24.2 % Tris 5.71 % glacial acetic acid 0.05 M EDTA (pH 8.0) 50 X stock stored at room temperature 1 X electrophoresis running buffer stored at 4 ˚C

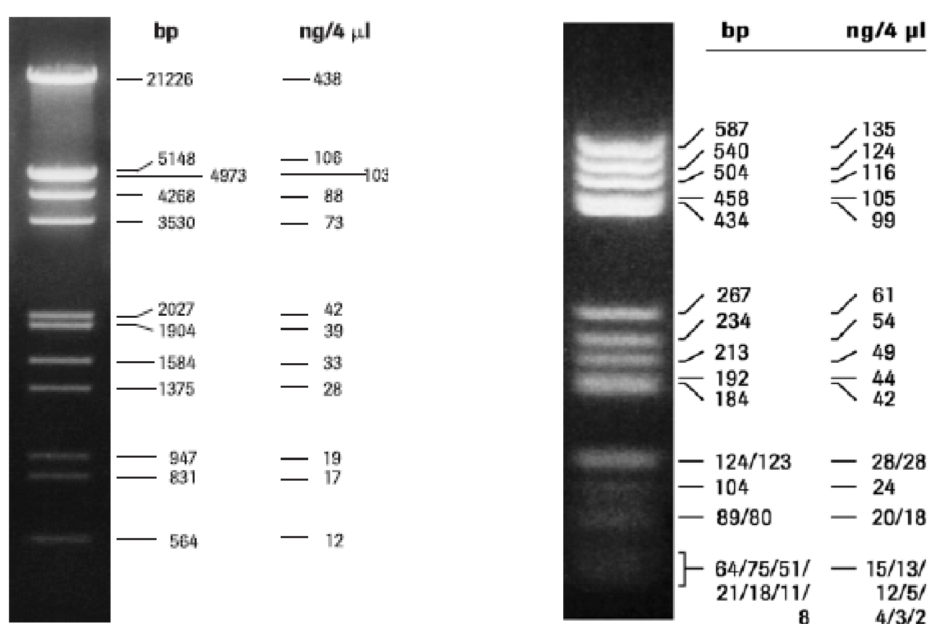


Figure 5.1 Molecular weight marker III and V (MW III and MW V).
MW III was used as standard marker. MW V was used to analyse small fragments generated during allele specific PCR. Both were supplied by Roche Applied Sciences (Germany).

5.5. Primers used in this study

Table 5.5.1 Primers used in this study

Name	Primer seq. (5'→3')	Description
	Isolation of MSM *rpoB* and SDM	
Whole*rpoB*F	cccccaccaaaaagcttgtt	Primers used to isolate the wild-type *rpoB* gene from the MSM chromosome. Underlined regions represent *Hind*III site in Whole*rpoB***F** and a *Sca*I site Whole*rpoB***R**.
Whole*rpoB*R	aacttgagtactgaggcggt	
SDM outer R	ttccataccggtacccacca	Outer general primers used in SDM
SDM outer F	Tccttctcggatccgcgct	
S531L mut F	cgtcgtcttttggcgctgggc	Mutagenic primers used to introduce specified mutations into the megaprimer during the first round of SDM
H526Y mut F	gggtctgacctacaagcgtcgt	
L521L mut F	gaacaacccgttgtcgggtct	
Q513V mut F	ccagctgtcggtgttcatggacca	
TH526TY mut F	cgggtctgacttacaagcgtcgt	
	Xover-PCR	
SDM outer R	ttccataccggtacccacca	Outer general primers used in Xover-PCR
SDM outer F	Tccttctcggatccgcgct	
Xover F inner	Ggtctgacccacaagcgt	Primers used in the first round of Xover-PCR. Xover F inner is used in reaction B, while Xover R inner is used in reaction A. The region that is homologous to the region down-stream from the deletion is underlined.
Xover R inner	acgcttgtgggtcagacc cgacagctggctggtgcc	
	Allele-specific PCR	
AS WILD S531L	acccacaagcgtcgtctttc	Allele specific primers for differentiation between specified mutations, and the wild-type allele at that site. These primers differed by a single base pair at the 3' end.
AS MT S531L	acccacaagcgtcgtctttt	
AS WILD H526Y	Cgctgtcgggtctgaccc	
AS MT H526Y	Cgctgtcgggtctgacct	
AS WILD L521L	catggaccagaacaacccgc	
AS MT L521L	catggaccagaacaacccgt	
AS WILD Q513V	Caccagccagctgtcgca	
AS MT Q513V	Caccagccagctgtcggt	
AS WILD TH526TY	Cgctgtcgggtctgaccc	
AS MT TH526TY	Cgctgtcgggtctgactt	
AS WT 30 bp del		
AS MT 30 bp del		
Seq F RRDR	tgagctgatccagaaccaga	Outer primers for standard AS-PCR (section 3.3.4); and primers used to isolate a smaller fragment for the modified version of AS-PCR (section 3.4.3). These primers were also used for sequencing.
Whole R3	Tcgggcacatgcggccg	

Sequencing primers		
Whole F	Gctgggtggcgggtgga	Primers used to sequence the p2rpoB template, and S531L and H526Y suicide vectors to check for any unwanted point mutations.
Whole F1	Tcgccgagctttccccca	
Whole F2	aacaagaagctgggcctgaa	
Whole F3	caccgaggaccgcgtcat	
Whole F4	Gcgcggcatcgtccgca	
Whole F5	Ccgtacccggtgacggtt	
Whole F6	agaagatcttcggaccgact	
Whole R	aagttgacgtctagcacgtaa	
Whole R1	Tggacgcccagtccggca	
Whole R2	atgatcgccacgagcaggt	
Seq R RRDR	cgatgaagccgaacgggtt	
Whole R4	Ttgtaacggcccacacgg	
Whole R5	Tctcctcgccgcggtcg	
Whole R6	Ttccggccggcaccgtc	
SCO orientation		
UpOri F	cccgcaacgattttcgctaa	Primers set 1 (F1 and R1) used to isolate the upstream *rpoB* gene in the SCO – regardless of whether the mutation is present
UpOri R	aagatcctgcccgtcgagga	
DownOri F	ctgcacgagggtcagactt	Primers set 2 (F2 and R2) used to isolate the downstream *rpoB* gene in the SCO – regardless of whether the mutation is present
DownOri F	aagtacaagcgtgtccgctt	

6. References

Adekambi, T., M. Drancourt & D. Raoult, (2009) The *rpoB* gene as a tool for clinical microbiologists. *Trends Microbiol* **17**: 37-45.

Adekambi, T., T. M. Shinnick, D. Raoult & M. Drancourt, (2008) Complete *rpoB* gene sequencing as a suitable supplement to DNA-DNA hybridization for bacterial species and genus delineation. *Int J Syst Evol Microbiol* **58**: 1807-1814.

Ahmad, S., G. F. Araj, P. K. Akbar, E. Fares, T. D. Chugh & A. S. Mustafa, (2000) Characterization of *rpoB* mutations in rifampin-resistant *Mycobacterium tuberculosis* isolates from the Middle East. *Diagn Microbiol Infect Dis* **38**: 227-232.

Ahmad, S., E. Mokaddas & E. Fares, (2002) Characterization of *rpoB* mutations in rifampin-resistant clinical *Mycobacterium tuberculosis* isolates from Kuwait and Dubai. *Diagn Microbiol Infect Dis* **44**: 245-252.

Alexander, D. C., J. R. Jones & J. Liu, (2003) A rifampin-hypersensitive mutant reveals differences between strains of *Mycobacterium smegmatis* and presence of a novel transposon, IS1623. *Antimicrob Agents Chemother* **47**: 3208-3213.

Altschul, S. F., W. Gish, W. Miller, E. W. Myers & D. J. Lipman, (1990) Basic local alignment search tool. *J Mol Biol* **215**: 403-410.

Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller & D. J. Lipman, (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**: 3389-3402.

Amaral, L., M. Martins, M. Viveiros, J. Molnar & J. E. Kristiansen, (2008) Promising therapy of XDR-TB/MDR-TB with thioridazine an inhibitor of bacterial efflux pumps. *Curr Drug Targets* **9**: 816-819.

Andersson, D. I., (2006) The biological cost of mutational antibiotic resistance: any practical conclusions? *Curr Opin Microbiol* **9**: 461-465.

Andersson, D. I. & D. Hughes, (2010) Antibiotic resistance and its cost: is it possible to reverse resistance? *Nat Rev Microbiol* **8**: 260-271.

Andersson, D. I. & B. R. Levin, (1999) The biological cost of antibiotic resistance. *Curr Opin Microbiol* **2**: 489-493.

Andries, K., P. Verhasselt, J. Guillemont, H. W. Gohlmann, J. M. Neefs, H. Winkler, J. Van Gestel, P. Timmerman, M. Zhu, E. Lee, P. Williams, D. de Chaffoy, E. Huitric, S. Hoffner, E. Cambau, C. Truffot-Pernot, N. Lounis & V. Jarlier, (2005) A diarylquinoline drug active on the ATP synthase of *Mycobacterium tuberculosis*. *Science* **307**: 223-227.

Aristimuno, L., M. Espana, A. Guilarte, C. Ramirez, J. Rojas, P. Gavin, A. I. Lopez-Calleja, M. A. Lezcano, M. J. Revillo, A. Cebollada, C. Martin & S. Samper, (2007) Multidrug-resistant *Mycobacterium tuberculosis* Beijing/W genotype in Venezuela. *J Med Microbiol* **56**: 1707-1708.

Artsimovitch, I., M. N. Vassylyeva, D. Svetlov, V. Svetlov, A. Perederina, N. Igarashi, N. Matsugaki, S. Wakatsuki, T. H. Tahirov & D. G. Vassylyev, (2005) Allosteric modulation of the RNA polymerase catalytic reaction is an essential component of transcription control by rifamycins. *Cell* **122**: 351-363.

Baharoglu, Z., R. Lestini, S. Duigou & B. Michel, (2011) RNA polymerase mutations that facilitate replication progression in the rep uvrD recF mutant lacking two accessory replicative helicases. *Mol Microbiol* **77**: 324-336.

Barrick, J. E., M. R. Kauth, C. C. Strelioff & R. E. Lenski, (2010) *Escherichia coli rpoB* mutants have increased evolvability in proportion to their fitness defects. *Mol Biol Evol*.

Barry, C. E., 3rd, (2001) *Mycobacterium smegmatis*: an absurd model for tuberculosis? *Trends in Microbiology* **9**: 473 - 474.

Barry, C. E., 3rd, H. I. Boshoff, V. Dartois, T. Dick, S. Ehrt, J. Flynn, D. Schnappinger, R. J. Wilkinson & D. Young, (2009) The spectrum of latent tuberculosis: rethinking the biology and intervention strategies. *Nat Rev Microbiol* **7**: 845-855.

Belanger, A. E., G. S. Besra, M. E. Ford, K. Mikusova, J. T. Belisle, P. J. Brennan & J. M. Inamine, (1996) The embAB genes of *Mycobacterium avium* encode an arabinosyl transferase involved in cell wall arabinan biosynthesis that is the target for the antimycobacterial drug ethambutol. *Proc Natl Acad Sci U S A* **93**: 11919-11924.

Billington, O. J., T. D. McHugh & S. H. Gillespie, (1999) Physiological cost of rifampin resistance induced *in vitro* in *Mycobacterium tuberculosis*. *Antimicrob Agents Chemother* **43**: 1866-1869.

Bjedov, I., O. Tenaillon, B. Gerard, V. Souza, E. Denamur, M. Radman, F. Taddei & I. Matic, (2003) Stress-induced mutagenesis in bacteria. *Science* **300**: 1404-1409.

Bjorkholm, B., M. Sjolund, P. G. Falk, O. G. Berg, L. Engstrand & D. I. Andersson, (2001) Mutation frequency and biological cost of antibiotic resistance in *Helicobacter pylori*. *Proc Natl Acad Sci U S A* **98**: 14607-14612.

Bjorkman, J., D. Hughes & D. I. Andersson, (1998) Virulence of antibiotic-resistant *Salmonella typhimurium*. *Proc Natl Acad Sci U S A* **95**: 3949-3953.

Bjorkman, J., P. Samuelsson, D. I. Andersson & D. Hughes, (1999) Novel ribosomal mutations affecting translational accuracy, antibiotic resistance and virulence of *Salmonella typhimurium*. *Mol Microbiol* **31**: 53-58.

Boor, K. J., M. L. Duncan & C. W. Price, (1995) Genetic and transcriptional organization of the region encoding the beta subunit of *Bacillus subtilis* RNA polymerase. *J Biol Chem* **270**: 20329-20336.

Borrell, S. & S. Gagneux, (2009) Infectiousness, reproductive fitness and evolution of drug-resistant *Mycobacterium tuberculosis*. *Int J Tuberc Lung Dis* **13**: 1456-1466.

Borukhov, S. & E. Nudler, (2003) RNA polymerase holoenzyme: structure, function and biological implications. *Curr Opin Microbiol* **6**: 93-100.

Borukhov, S., J. Lee & O. Laptenko, (2005) Bacterial transcription elongation factors: new insights into molecular mechanism of action. *Mol Microbiol* **55**: 1315-1324.

Borukhov, S. & E. Nudler, (2008) RNA polymerase: the vehicle of transcription. *Trends Microbiol* **16**: 126-134.

Boshoff, H. I., T. G. Myers, B. R. Copp, M. R. McNeil, M. A. Wilson & C. E. Barry, 3rd, (2004) The transcriptional responses of *Mycobacterium tuberculosis* to inhibitors of metabolism: novel insights into drug mechanisms of action. *J Biol Chem* **279**: 40174-40184.

Boshoff, H. I., R. Singh, and C. E. Barry, 3rd, (2008) Virulence and persistence mechanisms of *Mycobacterium tuberculosis*: Persistence mechanisms in latent disease. In: Handbook of Tuberculosis: Molecular Biology and Biochemistry*.* E. J. Rubin and S. H. E. Kaufmann (eds). Germany: Wiley-VCH, pp. 151 - 175.

Boshoff, H. I., M. B. Reed, C. E. Barry, 3rd & V. Mizrahi, (2003) DnaE2 polymerase contributes to *in vivo* survival and the emergence of drug resistance in *Mycobacterium tuberculosis*. *Cell* **113**: 183-193.

Boshoff, H. I. M., C. E. Barry, 3rd, & V. Mizrahi, (2004) Mutational Dynamics in *Mycobacterium tuberculosis*: implications for the evolution of drug resistance. *South African Journal of Science* **100**: 471 - 474

Boubakri, H., A. L. de Septenville, E. Viguera & B. Michel, (2010) The helicases dinG, rep and uvrD cooperate to promote replication across transcription units *in vivo*. *EMBO J* **29**: 145-157.

Burova, E. & M. E. Gottesman, (1995a) NusG overexpression inhibits Rho-dependent termination in Escherichia coli. *Mol Microbiol* **17**: 633-641.

Burova, E., S. C. Hung, V. Sagitov, B. L. Stitt & M. E. Gottesman, (1995b) Escherichia coli NusG protein stimulates transcription elongation rates in vivo and in vitro. *J Bacteriol* **177**: 1388-1392.

Brzostek, A., J. Pawelczyk, A. Rumijowska-Galewicz, B. Dziadek & J. Dziadek, (2009) *Mycobacterium tuberculosis* is able to accumulate and utilize cholesterol. *J Bacteriol* **191**: 6584-6591.

Cairns, J., J. Overbaugh & S. Miller, (1988) The origin of mutants. *Nature* **335**: 142-145.

Campbell, E. A., Korzheva N., Mustaev A., Murakami K., Nair S., Goldfarb A. & S. A. Darst, (2001) Structural mechanism for rifampicin inhibition of bacterial RNA Polymerase. *Cell* **104**: 901 - 912.

Campbell, E. A., O. Muzzin, M. Chlenov, J. L. Sun, C. A. Olson, O. Weinman, M. L. Trester-Zedlitz & S. A. Darst, (2002) Structure of the bacterial RNA polymerase promoter specificity sigma subunit. *Mol Cell* **9**: 527-539.

Cao, Z. W., L. Y. Han, C. J. Zheng, Z. L. Ji, X. Chen, H. H. Lin & Y. Z. Chen, (2005) Computer prediction of drug resistance mutations in proteins. *Drug Discov Today* **10**: 521-529.

Cars, O., L. D. Hogberg, M. Murray, O. Nordberg, S. Sivaraman, C. S. Lundborg, A. D. So & G. Tomson, (2008) Meeting the challenge of antibiotic resistance. *BMJ* **337**: a1438.

Chakrabarti, S. L. & L. Gorini, (1975) A link between streptomycin and rifampicin mutation. *Proc Natl Acad Sci U S A* **72**: 2084-2087.

Chakrabarti, S. L. & L. Gorini, (1977) Interaction between mutations of ribosomes and RNA polymerase: a pair of *strA* and rif mutants individually temperature-insensitive but temperature-sensitive in combination. *Proc Natl Acad Sci U S A* **74**: 1157-1161.

Changsen, C., S. G. Franzblau & P. Palittapongarnpim, (2003) Improved green fluorescent protein reporter gene-based microplate screening for antituberculosis compounds by utilizing an acetamidase promoter. *Antimicrob Agents Chemother* **47**: 3682-3687.

Chatterji, D. & A. K. Ojha, (2001) Revisiting the stringent response, ppGpp and starvation signaling. *Curr Opin Microbiol* **4**: 160-165.

Cohen, T., & M. Murray, (2004) Modeling epidemics of multidrug-resistant *M. tuberculosis* of heterogenous fitness. *Nature Medicine* **10**: 1117 - 1121.

Cohen, T., M. Murray, K. Wallengren, G. G. Alvarez, E. Y. Samuel & D. Wilson, (2010) The prevalence and drug sensitivity of tuberculosis among patients dying in hospital in KwaZulu-Natal, South Africa: a postmortem study. *PLoS Med* **7**: e1000296.

Cole, S. T., R. Brosch, J. Parkhill, T. Garnier, C. Churcher, D. Harris, S. V. Gordon, K. Eiglmeier, S. Gas, C. E. Barry, 3rd, F. Tekaia, K. Badcock, D. Basham, D. Brown, T. Chillingworth, R. Connor, R. Davies, K. Devlin, T. Feltwell, S. Gentles, N. Hamlin, S. Holroyd, T. Hornsby, K. Jagels, A. Krogh, J. McLean, S. Moule, L. Murphy, K. Oliver, J. Osborne, M. A. Quail, M. A. Rajandream, J. Rogers, S. Rutter, K. Seeger, J. Skelton, R. Squares, S. Squares, J. E. Sulston, K. Taylor, S. Whitehead & B. G. Barrell, (1998) Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* **393**: 537-544.

Collins, L. & S. G. Franzblau, (1997) Microplate alamar blue assay versus BACTEC 460 system for high-throughput screening of compounds against *Mycobacterium tuberculosis* and *Mycobacterium avium*. *Antimicrob Agents Chemother* **41**: 1004-1009.

Comas, I., J. Chakravartti, P. M. Small, J. Galagan, S. Niemann, K. Kremer, J. D. Ernst & S. Gagneux, (2010) Human T cell epitopes of *Mycobacterium tuberculosis* are evolutionarily hyperconserved. *Nat Genet* **42**: 498-503.

Connolly, L. E., P. H. Edelstein & L. Ramakrishnan, (2007) Why is long-term therapy required to cure tuberculosis? *PLoS Med* **4**: e120.

Conrad, T. M., M. Frazier, A. R. Joyce, B. K. Cho, E. M. Knight, N. E. Lewis, R. Landick & B. O. Palsson, (2010) RNA polymerase mutants found through adaptive evolution reprogram Escherichia coli for optimal growth in minimal media. *Proc Natl Acad Sci U S A*. doi:10.1073/pnas.0911253107

Daffe, M. & G. Etienne, (1999) The capsule of *Mycobacterium tuberculosis* and its implications for pathogenicity. *Tuber Lung Dis* **79**: 153-169.

Darst, S. A., (2001) Bacterial RNA polymerase. *Curr Opin Struct Biol* **11**: 155-162.

Davies, A. P., O. J. Billington, B. A. Bannister, W. R. Weir, T. D. McHugh & S. H. Gillespie, (2000) Comparison of fitness of two isolates of *Mycobacterium tuberculosis*, one of which had developed multi-drug resistance during the course of treatment. *J Infect* **41**: 184-187.

Davies, J., (2009) Everything depends on everything else. *Clin Microbiol Infect* **15 Suppl 1**: 1-4.

Davis, J. M. & L. Ramakrishnan, (2009) The role of the granuloma in expansion and dissemination of early tuberculous infection. *Cell* **136**: 37-49.

de Carvalho, L. P., S. M. Fischer, J. Marrero, C. Nathan, S. Ehrt & K. Y. Rhee, (2010) Metabolomics of *Mycobacterium tuberculosis* reveals compartmentalized co-catabolism of carbon substrates. *Chem Biol* **17**: 1122-1131.

Derbyshire, K. M. & S. Bardarov, (2000) DNA transfer in Mycobacteria: Conjugation and Transduction. In: Molecular Genetics of Mycobacteria. G. F. Hatfull & W. R. Jacobs, Jr. (eds). ASM Press, Washington D. C. pp 97-105.

Dey, A., A. K. Verma & D. Chatterji, (2010) Role of an RNA polymerase interacting protein, MsRbpA, from *Mycobacterium smegmatis* in phenotypic tolerance to rifampicin. *Microbiology* **156**: 873-883.

Donlan, R. M. & J. W. Costerton, (2002) Biofilms: survival mechanisms of clinically relevant microorganisms. *Clin Microbiol Rev* **15**: 167-193.

Dwyer, D. J., M. A. Kohanski & J. J. Collins, (2009) Role of reactive oxygen species in antibiotic action and resistance. *Curr Opin Microbiol* **12**: 482-489.

Dykxhoorn, D. M., R. Pierre & T. Linn, (1996) Synthesis of the beta and beta' subunits of *Escherichia coli* RNA polymerase is autogenously regulated *in vivo* by both transcriptional and translational mechanisms. *Mol Microbiol* **19**: 483-493.

Evans J., R. J. Wilkinson and H. Segal, (2011) The role of SigE in progression of rifampicin mono-resistant *Mycobacterium tuberculosis* to multidrug resistance. *J Antimicrob Chemother,* submitted.

Fange, D., K. Nilsson, T. Tenson & M. Ehrenberg, (2009) Drug efflux pump deficiency and drug target resistance masking in growing bacteria. *Proc Natl Acad Sci U S A* **106**: 8215-8220.

Foster, P. L., (1998) Adaptive mutation: has the unicorn landed? *Genetics* **148**: 1453-1459.

Franzblau, S. G., R. S. Witzig, J. C. McLaughlin, P. Torres, G. Madico, A. Hernandez, M. T. Degnan, M. B. Cook, V. K. Quenzer, R. M. Ferguson & R. H. Gilman, (1998) Rapid, low-technology MIC determination with clinical *Mycobacterium tuberculosis* isolates by using the microplate Alamar Blue assay. *J Clin Microbiol* **36**: 362-366.

Friedberg, E. C., P. L. Fischhaber & C. Kisker, (2001) Error-prone DNA polymerases: novel structures and the benefits of infidelity. *Cell* **107**: 9-12.

Gagneux, S., (2009) Fitness cost of drug resistance in *Mycobacterium tuberculosis*. *Clin Microbiol Infect* **15 Suppl 1**: 66-68.

Gagneux, S., M. V. Burgos, K. DeRiemer, A. Encisco, S. Munoz, P. C. Hopewell, P. M. Small & A. S. Pym, (2006a) Impact of bacterial genetics on the transmission of isoniazid-resistant *Mycobacterium tuberculosis*. *PLoS Pathog* **2**: e61.

Gagneux, S., C. Davis Long, P. M. Small, T. Van, G. K. Schoolnik, and B. J. M. Bohannan, (2006b) The competitive cost of antibiotic resistance in *Mycobacterium tuberculosis*. *Science* **312**: 1944 - 1946

Gagneux, S., K. DeRiemer, T. Van, M. Kato-Maeda, B. C. de Jong, S. Narayanan, M. Nicol, S. Niemann, K. Kremer, M. C. Gutierrez, M. Hilty, P. C. Hopewell & P. M. Small, (2006c) Variable host-pathogen compatibility in *Mycobacterium tuberculosis*. *Proc Natl Acad Sci U S A* **103**: 2869-2873.

Gagneux, S. & P. M. Small, (2007) Global phylogeography of *Mycobacterium tuberculosis* and implications for tuberculosis product development. *Lancet Infect Dis* **7**: 328-337.

Gandhi, N. R., A. Moll, A. W. Sturm, R. Pawinski, T. Govender, U. Lalloo, K. Zeller, J. Andrews & G. Friedland, (2006) Extensively drug-resistant tuberculosis as a cause of death in patients

co-infected with tuberculosis and HIV in a rural area of South Africa. *Lancet* **368**: 1575-1580.

Ganesan, A. K., A. J. Smith, N. J. Savery, P. Zamos & P. C. Hanawalt, (2007) Transcription coupled nucleotide excision repair in Escherichia coli can be affected by changing the arginine at position 529 of the beta subunit of RNA polymerase. *DNA Repair (Amst)* **6**: 1434-1440.

Garibyan, L., T. Huang, M. Kim, E. Wolff, A. Nguyen, T. Nguyen, A. Diep, K. Hu, A. Iverson, H. Yang & J. H. Miller, (2003) Use of the *rpoB* gene to determine the specificity of base substitution mutations on the *Escherichia coli* chromosome. *DNA Repair (Amst)* **2**: 593-608.

Garton, N. J., S. J. Waddell, A. L. Sherratt, S. M. Lee, R. J. Smith, C. Senner, J. Hinds, K. Rajakumar, R. A. Adegbola, G. S. Besra, P. D. Butcher & M. R. Barer, (2008) Cytological and transcript analyses reveal fat and lazy persister-like bacilli in tuberculous sputum. *PLoS Med* **5**: e75.

Gill, W. P., N. S. Harik, M. R. Whiddon, R. P. Liao, J. E. Mittler & D. R. Sherman, (2009) A replication clock for *Mycobacterium tuberculosis*. *Nat Med* **15**: 211-214.

Gillespie, S. H., (2002) Evolution of drug resistance in *Mycobacterium tuberculosis*: clinical and molecular perspective. *Antimicrob Agents Chemother* **46**: 267-274.

Gillespie, S. H., O. J. Billington, A. Breathnach & T. D. McHugh, (2002) Multiple drug-resistant *Mycobacterium tuberculosis*: evidence for changing fitness following passage through human hosts. *Microb Drug Resist* **8**: 273-279.

Glass, R. E., S. T. Jones & A. Ishihama, (1986a) Genetic studies on the beta subunit of *Escherichia coli* RNA polymerase. VII. RNA polymerase is a target for ppGpp. *Mol Gen Genet* **203**: 265-268.

Glass, R. E., S. T. Jones, V. Nene, T. Nomura, N. Fujita & A. Ishihama, (1986b) Genetic studies on the beta subunit of *Escherichia coli* RNA polymerase. VIII. Localisation of a region involved in promoter selectivity. *Mol Gen Genet* **203**: 487-491.

Goh, E. B., G. Yim, W. Tsui, J. McClure, M. G. Surette & J. Davies, (2002) Transcriptional modulation of bacterial gene expression by subinhibitory concentrations of antibiotics. *Proc Natl Acad Sci U S A* **99**: 17025-17030.

Goldman, S. R., R. H. Ebright & B. E. Nickels, (2009) Direct detection of abortive RNA transcripts *in vivo*. *Science* **324**: 927-928.

Gordhan, B. G., & T. Parish, (2001) Gene replacement using pretreated DNA. *Methods Mol Med*, **54**: 77-92.

Gumbo, T., (2010) New susceptibility breakpoints for first line antituberculosis drugs, based on antimicrobial PK/PD science and population pharmacokinetic variability. *Antimicrob Agents Chemother* **54**: 1484-1491.

Gutierrez, M. C., S. Brisse, R. Brosch, M. Fabre, B. Omais, M. Marmiesse, P. Supply & V. Vincent, (2005) Ancient origin and gene mosaicism of the progenitor of Mycobacterium tuberculosis. *PLoS Pathog* **1**: e5.

Guy, C. P., J. Atkinson, M. K. Gupta, A. A. Mahdi, E. J. Gwynn, C. J. Rudolph*, et al.*, (2009) Rep provides a second motor at the replisome to promote duplication of protein-bound DNA. *Mol Cell* **36**: 654-666.

Hall, A. R., I. C. James & C. Maclean, (2011) The fitness cost of rifampicin resistance in *Pseudomonas aeruginosa* depends on demand for RNA polymerase. *Genetics*. doi: 10.1534/genetics.110.124628

Hartkoorn, R. C., C. Sala, S. J. Magnet, J. M. Chen, F. Pojer & S. T. Cole, (2010) Sigma factor F does not prevent rifampin inhibition of RNA polymerase or cause rifampin tolerance in *Mycobacterium tuberculosis*. *J Bacteriol* **192**: 5472-5479.

Hazbon, M. H., A. S. Motiwala, M. Cavatore, M. Brimacombe, T. S. Whittam & D. Alland, (2008) Convergent evolutionary analysis identifies significant mutations in drug resistance targets of *Mycobacterium tuberculosis*. *Antimicrob Agents Chemother* **52**: 3369-3376.

Heep, M., B. Brandstatter, U. Rieger, N. Lehn, E. Richter, S. Rusch-Gerdes & S. Niemann, (2001) Frequency of *rpoB* mutations inside and outside the cluster I region in rifampin-resistant clinical *Mycobacterium tuberculosis* isolates. *J Clin Microbiol* **39**: 107-110.

Heep, M., S. Odenbreit, D. Beck, J. Decker, E. Prohaska, U. Rieger & N. Lehn, (2000a) Mutations at four distinct regions of the *rpoB* gene can reduce the susceptibility of *Helicobacter pylori* to rifamycins. *Antimicrob Agents Chemother* **44**: 1713-1715.

Heep, M., U. Rieger, D. Beck & N. Lehn, (2000b) Mutations in the beginning of the *rpoB* gene can induce resistance to rifamycins in both *Helicobacter pylori* and *Mycobacterium tuberculosis*. *Antimicrob Agents Chemother* **44**: 1075-1077.

Hershberg, R., M. Lipatov, P. M. Small, H. Sheffer, S. Niemann, S. Homolka, J. C. Roach, K. Kremer, D. A. Petrov, M. W. Feldman & S. Gagneux, (2008) High functional diversity in *Mycobacterium tuberculosis* driven by genetic drift and human demography. *PLoS Biol* **6**: e311.

Honer zu Bentrup, K. & D. G. Russell, (2001) Mycobacterial persistence: adaptation to a changing environment. *Trends Microbiol* **9**: 597-605.

Hosaka, T., M. Ohnishi-Kameyama, H. Muramatsu, K. Murakami, Y. Tsurumi, S. Kodani, M. Yoshida, A. Fujie & K. Ochi, (2009) Antibacterial discovery in actinomycetes strains with mutations in RNA polymerase or ribosomal protein S12. *Nat Biotechnol* **27**: 462-464.

Hu, H., Q. Zhang & K. Ochi, (2002) Activation of antibiotic biosynthesis by specified mutations in the *rpoB* gene (encoding the RNA polymerase beta subunit) of *Streptomyces lividans*. *J Bacteriol* **184**: 3984-3991.

Huber, J., R. G. Donald, S. H. Lee, L. W. Jarantow, M. J. Salvatore, X. Meng, R. Painter, R. H. Onishi, J. Occi, K. Dorso, K. Young, Y. W. Park, S. Skwish, M. J. Szymonifka, T. S. Waddell, L. Miesel, J. W. Phillips & T. Roemer, (2009) Chemical genetic identification of peptidoglycan inhibitors potentiating carbapenem activity against methicillin-resistant *Staphylococcus aureus*. *Chem Biol* **16**: 837-848.

Huff, J., A. Cysz, R. Landick, & M. Niederwies, (2010) Taking phage integration to the next level as a genetic tool for mycobacteria. Gene **468:** 8-19

Huitric, E., P. Verhasselt, A. Koul, K. Andries, S. Hoffner & D. I. Andersson, (2010) Rates and mechanisms of resistance development in *Mycobacterium tuberculosis* to a novel diarylquinoline ATP synthase inhibitor. *Antimicrob Agents Chemother*. **54:** 1022-1028

Huitric, E., J. Werngren, P. Jureen & S. Hoffner, (2006) Resistance levels and *rpoB* gene mutations among *in vitro*-selected rifampin-resistant *Mycobacterium tuberculosis* mutants. *Antimicrob Agents Chemother* **50**: 2860-2862.

Hwang, H. Y., C. Y. Chang, L. L. Change, S. F. Chang, Y. H. Chang, and Y. J. Chen, (2003) Characterization of rifampicin-resistant *Mycobacterium tuberculosis* in Taiwan. *Journal of Medical Microbiology* **52**: 239 - 245

Ilyina, T. S., M. I. Ovadis, S. Z. Mindlin, Z. M. Gorlenko & R. B. Khesin, (1971) Interaction of RNA polymerase mutations in haploid and merodiploid cells of *Escherichia coli* K-12. *Mol Gen Genet* **110**: 118-133.

Inaoka, T., K. Takahashi, H. Yada, M. Yoshida & K. Ochi, (2004) RNA polymerase mutation activates the production of a dormant antibiotic 3,3'-neotrehalosadiamine via an autoinduction mechanism in *Bacillus subtilis*. *J Biol Chem* **279**: 3885-3892.

Ingham, C. J. & P. A. Furneaux, (2000) Mutations in the β-subunit of the *Bacillus subtilis* RNA polymerase that confer both rifampicin resistance and hypersensitivity to NusG. *Microbiology* **146**: 3041-3049.

Ioerger, T. R., Y. Feng, X. Chen, K. M. Dobos, T. C. Victor, E. M. Streicher, R. M. Warren, N. C. Gey van Pittius, P. D. Van Helden & J. C. Sacchettini, (2010) The non-clonality of drug resistance in Beijing-genotype isolates of *Mycobacterium tuberculosis* from the Western Cape of South Africa. *BMC Genomics* **11**: 670.

Ioerger, T. R., S. Koo, E. G. No, X. Chen, M. H. Larsen, W. R. Jacobs, Jr., M. Pillay, A. W. Sturm & J. C. Sacchettini, (2009) Genome analysis of multi- and extensively-drug-resistant tuberculosis from KwaZulu-Natal, South Africa. *PLoS One* **4**: e7778.

Ishikawa, J., K. Chiba, H. Kurita & H. Satoh, (2006) Contribution of *rpoB2* RNA polymerase beta subunit gene to rifampin resistance in Nocardia species. *Antimicrob Agents Chemother* **50**: 1342-1346.

Jassal, M. & W. R. Bishai, (2009) Extensively drug-resistant tuberculosis. *Lancet Infect Dis* **9**: 19-30.

Jacobs W.R. Jr, Kalpan G.V., Cirillo J.D., Pascopella L., Snapper S.B., Udani R.A., Jones W., Barletta R.G. and B. R. Bloom, (1991). Genetic systems for mycobacteria. *Methods Enzymol.* **204**, 537–555.

Jenkins, C., J. Bacon, J. Allnutt, K. A. Hatch, A. Bose, D. M. O'Sullivan, C. Arnold, S. H. Gillespie & T. D. McHugh, (2009) Enhanced heterogeneity of *rpoB* in *Mycobacterium tuberculosis* found at low pH. *J Antimicrob Chemother*. **63:** 1118-1120.

Jiang, X., W. Zhang, Y. Zhang, F. Gao, C. Lu, X. Zhang & H. Wang, (2008) Assessment of efflux pump gene expression in a clinical isolate *Mycobacterium tuberculosis* by real-time reverse transcription PCR. *Microb Drug Resist* **14**: 7-11.

Jin, D. J. & C. A. Gross, (1988) Mapping and sequencing of mutations in the *Escherichia coli rpoB* gene that lead to rifampicin resistance. *J Mol Biol* **202**: 45-58.

Jin, D. J. & C. A. Gross, (1989) Characterization of the pleiotropic phenotypes of rifampin-resistant *rpoB* mutants of *Escherichia coli*. *J Bacteriol* **171**: 5229-5231.

Jin, D. J. & C. A. Gross, (1991) RpoB8, a rifampicin-resistant termination-proficient RNA polymerase, has an increased Km for purine nucleotides during transcription elongation. *J Biol Chem* **266**: 14478-14485.

Jin, D. J., W. A. Walter & C. A. Gross, (1988) Characterization of the termination phenotypes of rifampicin-resistant mutants. *J Mol Biol* **202**: 245-253.

Jindani, A., C. J. Dore & D. A. Mitchison, (2003) Bactericidal and sterilizing activities of antituberculosis drugs during the first 14 days. *Am J Respir Crit Care Med* **167**: 1348-1354.

Jones, D. T., (1999) Protein secondary structure prediction based on position-specific scoring matrices. *J Mol Biol* **292**: 195-202.

Karunakaran, P. & J. Davies, (2000) Genetic antagonism and hypermutability in *Mycobacterium smegmatis*. *J Bacteriol* **182**: 3331-3335.

Kaufmann, S. H. E. & B. D. Walker, (2009) AIDS and Tuberculosis: A deadly liasion. Germany: Wiley-VCH.

Reyrat J.M. and D. Khan, (2001) *Mycobacterium smegmatis*: an absurd model for tuberculosis? *Trends in Microbiology* **9**: 472 - 473.

Kim, M., E. Wolff, T. Huang, L. Garibyan, A. M. Earl, J. R. Battista & J. H. Miller, (2004) Developing a genetic system in *Deinococcus radiodurans* for analyzing mutations. *Genetics* **166**: 661-668.

Kohanski, M. A., M. A. Depristo & J. J. Collins, (2010) Sublethal antibiotic treatment leads to multidrug resistance via radical-induced mutagenesis. *Mol Cell* **37**: 311-320.

Kohanski, M. A., D. J. Dwyer, J. Wierzbowski, G. Cottarel & J. J. Collins, (2008) Mistranslation of membrane proteins and two-component system activation trigger antibiotic-mediated cell death. *Cell* **135**: 679-690.

Koonin, E. V. & Y. I. Wolf, (2010) Constraints and plasticity in genome and molecular-phenome evolution. *Nat Rev Genet* **11**: 487-498.

Kurth, D. G., G. M. Gago, A. de la Iglesia, B. Bazet Lyonnet, T. W. Lin, H. R. Morbidoni, S. C. Tsai & H. Gramajo, (2009) ACCase 6 is the essential acetyl-CoA carboxylase involved in fatty acid and mycolic acid biosynthesis in mycobacteria. *Microbiology* **155**: 2664-2675.

Kuznedelov, K. & K. Severinov, (2009) Recombinant bacterial RNA polymerase: preparation and applications. *Methods* **47**: 44-52.

Landick, R., J. Stewart & D. N. Lee, (1990) Amino acid changes in conserved regions of the beta-subunit of *Escherichia coli* RNA polymerase alter transcription pausing and termination. *Genes Dev* **4**: 1623-1636.

Lane, W. J. & S. A. Darst, (2010a) Molecular evolution of multisubunit RNA polymerases: sequence analysis. *J Mol Biol* **395**: 671-685.

Lane, W. J. & S. A. Darst, (2010b) Molecular evolution of multisubunit RNA polymerases: structural analysis. *J Mol Biol* **395**: 686-704.

Larsen M.H. (2000) Appendix 1. in Molecular genetics of mycobacteria. G.F. Hatfull & W.R. Jacobs, Jr,(eds) ASM Press. Washington, D.C. pp. 313-320.

Lechner, D., S. Gibbons & F. Bucar, (2008) Plant phenolic compounds as ethidium bromide efflux inhibitors in *Mycobacterium smegmatis*. *J Antimicrob Chemother* **62**: 345-348.

Levin, B. R., V. Perrot & N. Walker, (2000) Compensatory mutations, antibiotic resistance and the population genetics of adaptive evolution in bacteria. *Genetics* **154**: 985-997.

Link, A. J., D. Phillips & G. M. Church, (1997) Methods for generating precise deletions and insertions in the genome of wild-type *Escherichia coli*: application to open reading frame characterization. *J Bacteriol* **179**: 6228-6237.

Lipsitch, M. & B. R. Levin, (1997) The population dynamics of antimicrobial chemotherapy. *Antimicrob Agents Chemother* **41**: 363-373.

Livermore, D. M., (2003) Bacterial resistance: origins, epidemiology, and impact. *Clin Infect Dis* **36**: S11-23.

Louw, G. E., R. M. Warren, N. C. van Pittius, C. R. McEvoy, P. D. Van Helden & T. C. Victor, (2009) A balancing act: efflux/influx in mycobacterial drug resistance. *Antimicrob Agents Chemother* **53**: 3181-3189.

Lovmar, M., K. Nilsson, E. Lukk, V. Vimberg, T. Tenson & M. Ehrenberg, (2009) Erythromycin resistance by L4/L22 mutations and resistance masking by drug efflux pump deficiency. *EMBO J* **28**: 736-744.

Machowski, E. E., S. Dawes & V. Mizrahi, (2005) TB tools to tell the tale-molecular genetic methods for mycobacterial research. *Int J Biochem Cell Biol* **37**: 54-68.

Makinen, J., H. J. Marttila, M. Marjamaki, M. K. Viljanen & H. Soini, (2006) Comparison of two commercially available DNA line probe assays for detection of multidrug-resistant *Mycobacterium tuberculosis*. *J Clin Microbiol* **44**: 350-352.

Malhotra, S., V. J. Cook, J. N. Wolfe, P. Tang, K. Elwood & M. K. Sharma, (2010) A mutation in *Mycobacterium tuberculosis rpoB* gene confers rifampin resistance in three HIV-TB cases. *Tuberculosis (Edinb)* **90**: 152-157.

Malshetty, V., K. Kurthkoti, A. China, B. Mullick, S. Yamunadevi, P. B. Sang, N. Srinivasan, V. Nagaraja & U. Varshney, (2010) Novel insertion and deletion mutants of RpoB which render *Mycobacterium smegmatis* RNA polymerase recalcitrant to rifampicin mediated inhibition of transcription. *Microbiology* **156**: 1565-1573.

Mani, C., N. Selvakumar, S. Narayanan & P. R. Narayanan, (2001) Mutations in the *rpoB* gene of multidrug-resistant *Mycobacterium tuberculosis* clinical isolates from India. *J Clin Microbiol* **39**: 2987-2990.

Mann, C. M. & J. L. Markham, (1998) A new method for determining the minimum inhibitory concentration of essential oils. *J Appl Microbiol* **84**: 538-544.

Mariam, D. H., Y. Mengistu, S. E. Hoffner, and D. I. Anderson (2004) Effect of *rpoB* mutations conferring rifampin resistance on fitness of *Mycobacterium tuberculosis*. *Antimicrob Agents Chemother* **48**: 1289 - 1294.

Marrero, J., K. Y. Rhee, D. Schnappinger, K. Pethe & S. Ehrt, (2010) Gluconeogenic carbon flow of tricarboxylic acid cycle intermediates is critical for *Mycobacterium tuberculosis* to establish and maintain infection. *Proc Natl Acad Sci U S A* **107**: 9819-9824.

Martinez, J. L., A. Fajardo, L. Garmendia, A. Hernandez, J. F. Linares, L. Martinez-Solano & M. B. Sanchez, (2009) A global view of antibiotic resistance. *FEMS Microbiol Rev* **33**: 44-65.

Matsiota-Bernard, P., G. Vrioni & E. Marinis, (1998) Characterization of *rpoB* mutations in rifampin-resistant clinical *Mycobacterium tuberculosis* isolates from Greece. *J Clin Microbiol* **36**: 20-23.

Maughan, H., B. Galeano & W. L. Nicholson, (2004) Novel *rpoB* mutations conferring rifampin resistance on *Bacillus subtilis*: global effects on growth, competence, sporulation, and germination. *J Bacteriol* **186**: 2481-2486.

McKinney, J. D., K. Honer zu Bentrup, E. J. Munoz-Elias, A. Miczak, B. Chen, W. T. Chan, D. Swenson, J. C. Sacchettini, W. R. Jacobs, Jr. & D. G. Russell, (2000) Persistence of *Mycobacterium tuberculosis* in macrophages and mice requires the glyoxylate shunt enzyme isocitrate lyase. *Nature* **406**: 735-738.

McMurry, L. M., M. Oethinger & S. B. Levy, (1998) Triclosan targets lipid synthesis. *Nature* **394**: 531-532.

Miller, C. H., S. Nisa, S. Dempsey, C. Jack & R. O'Toole, (2009) Modifying culture conditions in chemical library screening identifies alternative inhibitors of mycobacteria. *Antimicrob Agents Chemother* **53**: 5279-5283.

Miller, L. P., J. T. Crawford & T. M. Shinnick, (1994) The *rpoB* gene of *Mycobacterium tuberculosis*. *Antimicrob Agents Chemother* **38**: 805-811.

Mitchison, D. A., (2000) Role of individual drugs in the chemotherapy of tuberculosis. *Int J Tuberc Lung Dis* **4**: 796-806.

Mokrousov, I., (2004) Multiple *rpoB* mutants of *Mycobacterium tuberculosis* and second-order selection. *Emerg Infect Dis* **10**: 1337-1338.

Mokrousov, I., T. Otten, B. Vyshnevskiy & O. Narvskaya, (2003) Allele-specific *rpoB* PCR assays for detection of rifampin-resistant *Mycobacterium tuberculosis* in sputum smears. *Antimicrob Agents Chemother* **47**: 2231-2235.

Moore, S. D. & R. T. Sauer, (2008) Revisiting the mechanism of macrolide-antibiotic resistance mediated by ribosomal protein L22. *Proc Natl Acad Sci U S A* **105**: 18261-18266.

Morlock, G. P., B. B. Plikaytis, and J. T. Crawford, (2002) Characterization of spontaneous, *in vitro* selected rifampin-resistant mutants of *Mycobacterium tuberculosis* strain H37Rv. *Antimicrob. Agents Chemother* **44**: 3298 - 3301.

Motiwala, A. S., Y. Dai, E. C. Jones-Lopez, S. H. Hwang, J. S. Lee, S. N. Cho, L. E. Via, C. E. Barry & D. Alland, (2010) Mutations in extensively drug-resistant *Mycobacterium tuberculosis* that do not code for known drug-resistance mechanisms. *J Infect Dis* **201**: 881-888.

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

Murakami, K. S., S. Masuda, E. A. Campbell, O. Muzzin & S. A. Darst, (2002a) Structural basis of transcription initiation: an RNA polymerase holoenzyme-DNA complex. *Science* **296**: 1285-1290.

Murakami, K. S., S. Masuda & S. A. Darst, (2002b) Structural basis of transcription initiation: RNA polymerase holoenzyme at 4 A resolution. *Science* **296**: 1280-1284.

Newell, K. V., D. P. Thomas, D. Brekasis & M. S. Paget, (2006) The RNA polymerase-binding protein RbpA confers basal levels of rifampicin resistance on *Streptomyces coelicolor*. *Mol Microbiol* **60**: 687-696.

Nicholson, W. L. & H. Maughan, (2002) The spectrum of spontaneous rifampin resistance mutations in the *rpoB* gene of *Bacillus subtilis* 168 spores differs from that of vegetative cells and resembles that of *Mycobacterium tuberculosis*. *J Bacteriol* **184**: 4936-4940.

Nicol, M. P. & R. J. Wilkinson, (2008) The clinical consequences of strain diversity in *Mycobacterium tuberculosis*. *Trans R Soc Trop Med Hyg* **102**: 955-965.

Niemann, S., C. U. Koser, S. Gagneux, C. Plinke, S. Homolka, H. Bignell, R. J. Carter, R. K. Cheetham, A. Cox, N. A. Gormley, P. Kokko-Gonzales, L. J. Murray, R. Rigatti, V. P. Smith, F. P. Arends, H. S. Cox, G. Smith & J. A. Archer, (2009) Genomic diversity among drug sensitive and

multidrug resistant isolates of *Mycobacterium tuberculosis* with identical DNA fingerprints. *PLoS One* **4**: e7407.

Nikaido, H., (1994) Prevention of drug access to bacterial targets: permeability barriers and active efflux. *Science* **264**: 382-388.

Nikaido, H., (2001) Preventing drug access to targets: cell surface permeability barriers and active efflux in bacteria. *Semin Cell Dev Biol* **12**: 215-223.

Nickels, B. E., S. J. Garrity, V. Mekler, L. Minakhin, K. Severinov, R. H. Ebright & A. Hochschild, (2005) The interaction between sigma70 and the beta-flap of Escherichia coli RNA polymerase inhibits extension of nascent RNA during early elongation. *Proc Natl Acad Sci U S A* **102**: 4488-4493.

Nomura, T., A. Ishihama, M. Kajitani, T. Takahashi, N. Nakada & K. Yoshinaga, (1984) Promoter selectivity of *Escherichia coli* RNA polymerase. II: Altered promoter selection by mutant holoenzymes. *Mol Gen Genet* **193**: 8-16.

O'Gaora, P., S. Barnini, C. Hayward, E. Filley, G. Rook, D. Young & J. Thole, (1997) Mycobacteria as immunogens: Development of expression vectors for use in multiple mycobacterial species. *Medical Principles and Practise* **6**: 91 - 96.

O'Neill, A., B. Oliva, C. Storey, A. Hoyle, C. Fishwick & I. Chopra, (2000) RNA polymerase inhibitors with activity against rifampin-resistant mutants of *Staphylococcus aureus*. *Antimicrob Agents Chemother* **44**: 3163-3166.

O'Neill, A. J., T. Huovinen, C. W. Fishwick & I. Chopra, (2006) Molecular genetic and structural modeling studies of *Staphylococcus aureus* RNA polymerase and the fitness of rifampin resistance genotypes in relation to clinical prevalence. *Antimicrob Agents Chemother* **50**: 298-309.

Ochi, K., (1987) Metabolic initiation of differentiation and secondary metabolism by *Streptomyces griseus*: significance of the stringent response (ppGpp) and GTP content in relation to A factor. *J Bacteriol* **169**: 3608-3616.

Oliver, A., R. Canton, P. Campo, F. Baquero & J. Blazquez, (2000) High frequency of hypermutable *Pseudomonas aeruginosa* in cystic fibrosis lung infection. *Science* **288**: 1251-1254.

Owens, J. T., R. Miyake, K. Murakami, A. J. Chmura, N. Fujita, A. Ishihama & C. F. Meares, (1998) Mapping the sigma70 subunit contact sites on Escherichia coli RNA polymerase with a sigma70-conjugated chemical protease. *Proc Natl Acad Sci U S A* **95**: 6021-6026.

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

Pandey, A. K. & C. M. Sassetti, (2008) Mycobacterial persistence requires the utilization of host cholesterol. *Proc Natl Acad Sci U S A* **105**: 4376-4380.

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

Park, J. S., M. T. Marr & J. W. Roberts, (2002) E. coli Transcription repair coupling factor (Mfd protein) rescues arrested complexes by promoting forward translocation. *Cell* **109**: 757-767.

Parwati, I., R. van Crevel & D. van Soolingen, (2010) Possible underlying mechanisms for successful emergence of the *Mycobacterium tuberculosis* Beijing genotype strains. *Lancet Infect Dis* **10**: 103-111.

Passador, L. & T. Linn, (1989) Autogenous regulation of the RNA polymerase beta subunit of *Escherichia coli* occurs at the translational level *in vivo*. *J Bacteriol* **171**: 6234-6242.

Pei, J. & N. V. Grishin, (2001) AL2CO: calculation of positional conservation in a protein sequence alignment. *Bioinformatics* **17**: 700-712.

Pei, J. & N. V. Grishin, (2007) PROMALS: towards accurate multiple sequence alignments of distantly related proteins. *Bioinformatics* **23**: 802-808.

Pei, J., B. H. Kim & N. V. Grishin, (2008) PROMALS3D: a tool for multiple protein sequence and structure alignments. *Nucleic Acids Res* **36**: 2295-2300.

Perkins, A. E. & W. L. Nicholson, (2008) Uncovering new metabolic capabilities of *Bacillus subtilis* using phenotype profiling of rifampin-resistant *rpoB* mutants. *J Bacteriol* **190**: 807-814.

Peterson, J. D., L. A. Umayam, T. Dickinson, E. K. Hickey & O. White, (2001) The Comprehensive Microbial Resource. *Nucleic Acids Res* **29**: 123-125.

Pethe, K., P. C. Sequeira, S. Agarwalla, K. Rhee, K. Kuhen, W. Y. Phong, V. Patel, D. Beer, J. R. Walker, J. Duraiswamy, J. Jiricek, T. H. Keller, A. Chatterjee, M. P. Tan, M. Ujjini, S. P. Rao, L. Camacho, P. Bifani, P. A. Mak, I. Ma, S. W. Barnes, Z. Chen, D. Plouffe, P. Thayalan, S. H. Ng, M. Au, B. H. Lee, B. H. Tan, S. Ravindran, M. Nanjundappa, X. Lin, A. Goh, S. B. Lakshminarayana, C. Shoen, M. Cynamon, B. Kreiswirth, V. Dartois, E. C. Peters, R. Glynne, S. Brenner & T. Dick, (2010) A chemical genetic screen in *Mycobacterium tuberculosis* identifies carbon-source-dependent growth inhibitors devoid of *in vivo* efficacy. *Nat Commun* **1**: 1-8.

Peyron, P., J. Vaubourgeix, Y. Poquet, F. Levillain, C. Botanch, F. Bardou, M. Daffe, J. F. Emile, B. Marchou, P. J. Cardona, C. de Chastellier & F. Altare, (2008) Foamy macrophages from tuberculous patients' granulomas constitute a nutrient-rich reservoir for *M. tuberculosis* persistence. *PLoS Pathog* **4**: e1000204.

Piddock, L. J., K. J. Williams & V. Ricci, (2000) Accumulation of rifampicin by *Mycobacterium aurum*, *Mycobacterium smegmatis* and *Mycobacterium tuberculosis*. *J Antimicrob Chemother* **45**: 159-165.

Potrykus, K. & M. Cashel, (2008) (p)ppGpp: still magical? *Annu Rev Microbiol* **62**: 35-51.

Pozzi, G., M. Meloni, E. Iona, G. Orru, O. F. Thoresen, M. L. Ricci, M. R. Oggioni, L. Fattorini and G. Orefici, (1999) *rpoB* mutations in multidrug-resistant strains of *Mycobacterium tuberculosis* isolated in Italy. *Antimicrob Agents Chemother* **37**: 1197 - 1199.

Qian, L., C. Abe, T. P. Lin, M. C. Yu, S. N. Cho, S. Wang & J. T. Douglas, (2002) *rpoB* genotypes of *Mycobacterium tuberculosis* Beijing family isolates from East Asian countries. *J Clin Microbiol* **40**: 1091-1094.

Quan, S., H. Venter & E. R. Dabbs, (1997) Ribosylative inactivation of rifampin by *Mycobacterium smegmatis* is a principal contributor to its low susceptibility to this antibiotic. *Antimicrob Agents Chemother* **41**: 2456-2460.

Ramaswamy, S. & J. M. Musser, (1998) Molecular genetic basis of antimicrobial agent resistance in *Mycobacterium tuberculosis*: 1998 update. *Tuber Lung Dis* **79**: 3-29.

Rang, C., J. E. Galen, J. B. Kaper & L. Chao, (2003) Fitness cost of the green fluorescent protein in gastrointestinal bacteria. *Can J Microbiol* **49**: 531-537.

Reynolds, M. G., (2000) Compensatory evolution in rifampin-resistant *Escherichia coli*. *Genetics* **156**: 1471-1481.

Ricci, V., P. Tzakas, A. Buckley & L. J. Piddock, (2006) Ciprofloxacin-resistant *Salmonella enterica serovar Typhimurium* strains are difficult to select in the absence of AcrB and TolC. *Antimicrob Agents Chemother* **50**: 38-42.

Roberts, J. & J. S. Park, (2004) Mfd, the bacterial transcription repair coupling factor: translocation, repair and termination. *Curr Opin Microbiol* **7**: 120-125.

Rosenberg, S. M. & P. J. Hastings, (2004) Adaptive point mutation and adaptive amplification pathways in the *Escherichia coli* Lac system: stress responses producing genetic change. *J Bacteriol* **186**: 4838-4843.

Roth, J., (2010) Genetic Adaptation: A New Piece for a Very Old Puzzle. *Current Biology* **20**: R15 - R17.

Rustad, T. R., A. M. Sherrid, & D. R. Sherman, (2008) Molecular mechanisms of dormancy and resuscitation: Role of latency in the global tuberculosis epidemic. In: Molecular mechanisms of dormancy and resuscitation: Role of latency in the global tuberculosis epidemic. E. J. Rubin & S. H. E. Kaufmann (eds). Germany: Wiley -VCH, pp. 287 - 299.

Safi, H., R. D. Fleischmann, S. N. Peterson, M. B. Jones, B. Jarrahi & D. Alland, (2010) Allelic exchange and mutant selection demonstrate that common clinical *embCAB* gene mutations only modestly increase resistance to ethambutol in *Mycobacterium tuberculosis*. *Antimicrob Agents Chemother* **54**: 103-108.

Sambrook, J, E.F. & Fritsch, T. Maniatis, (1989) Molecular cloning. A laboratory manual. Second Edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Sambrook, J., & D.W. Russell, (2001) Molecular cloning. A laboratory manual. Third Edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Sandgren, A., M. Strong, P. Muthukrishnan, B. K. Weiner, G. M. Church & M. B. Murray, (2009) Tuberculosis drug resistance mutation database. *PLoS Med* **6**: e2.

Savvi, S., D. F. Warner, B. D. Kana, J. D. McKinney, V. Mizrahi & S. S. Dawes, (2008) Functional characterization of a vitamin B12-dependent methylmalonyl pathway in *Mycobacterium tuberculosis*: implications for propionate metabolism during growth on fatty acids. *J Bacteriol* **190**: 3886-3895.

Scanga, C. A. & J. L. Flynn, (2010) Mycobacterial infections and the inflammatory seesaw. *Cell Host Microbe* **7**: 177-179.

Schnappinger, D., S. Ehrt, M. I. Voskuil, Y. Liu, J. A. Mangan, I. M. Monahan, G. Dolganov, B. Efron, P. D. Butcher, C. Nathan & G. K. Schoolnik, (2003) Transcriptional adaptation of *Mycobacterium tuberculosis* within macrophages: Insights into the phagosomal environment. *J Exp Med* **198**: 693-704.

Schulz Zur Wiesch, P., J. Engelstadter & S. Bonhoeffer, (2010) Compensation of fitness costs and reversibility of antibiotic resistance mutations. *Antimicrob Agents Chemother*. **54**: 2085-2095.

Scott, D. J., A. L. Ferguson, M. T. Gallegos, M. Pitt, M. Buck & J. G. Hoggett, (2000) Interaction of sigma factor sigmaN with *Escherichia coli* RNA polymerase core enzyme. *Biochem J* **352**: 539-547.

Shaner, N. C., P. A. Steinbach & R. Y. Tsien, (2005) A guide to choosing fluorescent proteins. *Nat Methods* **2**: 905-909.

Shcherbakov, D., R. Akbergenov, T. Matt, P. Sander, D. I. Andersson & E. C. Bottger, (2010) Directed mutagenesis of *Mycobacterium smegmatis* 16S rRNA to reconstruct the *in vivo* evolution of aminoglycoside resistance in *Mycobacterium tuberculosis*. *Mol Microbiol* **77**: 830-840.

Siddiqi, N., M. Shamim, S. Hussain, R. K. Choudhary, N. Ahmed, Prachee, S. Banerjee, G. R. Savithri, M. Alam, N. Pathak, A. Amin, M. Hanief, V. M. Katoch, S. K. Sharma & S. E. Hasnain, (2002) Molecular characterization of multidrug-resistant isolates of *Mycobacterium tuberculosis* from patients in North India. *Antimicrob Agents Chemother* **46**: 443-450.

Singer, M., D. J. Jin, W. A. Walter & C. A. Gross, (1993) Genetic evidence for the interaction between cluster I and cluster III rifampicin resistant mutations. *J Mol Biol* **231**: 1-5.

Smith, A. M. & K. P. Klugman, (1997) "Megaprimer" method of PCR-based mutagenesis: the concentration of megaprimer is a critical factor. *Biotechniques* **22**: 438, 442.

Smith, N. H., R. G. Hewinson, K. Kremer, R. Brosch & S. V. Gordon, (2009) Myths and misconceptions: the origin and evolution of *Mycobacterium tuberculosis*. *Nat Rev Microbiol* **7**: 537-544.

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

Srivastava, S., S. Musuka, C. Sherman, C. Meek, R. Leff & T. Gumbo, (2010) Efflux-pump-derived multiple drug resistance to ethambutol monotherapy in *Mycobacterium tuberculosis* and the pharmacokinetics and pharmacodynamics of ethambutol. *J Infect Dis* **201**: 1225-1231.

Stallings, C. L., N. C. Stephanou, L. Chu, A. Hochschild, B. E. Nickels & M. S. Glickman, (2009) CarD is an essential regulator of rRNA transcription required for *Mycobacterium tuberculosis* persistence. *Cell* **138**: 146-159.

Stewart, G. R., B. D. Robertson & D. B. Young, (2003) Tuberculosis: a problem with persistence. *Nat Rev Microbiol* **1**: 97-105.

Stinear, T. P., A. Mve-Obiang, P. L. Small, W. Frigui, M. J. Pryor, R. Brosch, G. A. Jenkin, P. D. Johnson, J. K. Davies, R. E. Lee, S. Adusumilli, T. Garnier, S. F. Haydock, P. F. Leadlay & S. T. Cole, (2004) Giant plasmid-encoded polyketide synthases produce the macrolide toxin of *Mycobacterium ulcerans*. *Proc Natl Acad Sci U S A* **101**: 1345-1349.

Strauss, O. J., R. M. Warren, A. Jordaan, E. M. Streicher, M. Hanekom, A. A. Falmer, H. Albert, A. Trollip, E. Hoosain, P. D. van Helden & T. C. Victor, (2008) Spread of a low-fitness drug-resistant *Mycobacterium tuberculosis* strain in a setting of high human immunodeficiency virus prevalence. *J Clin Microbiol* **46**: 1514-1516.

Tala, A., G. Wang, M. Zemanova, S. Okamoto, K. Ochi & P. Alifano, (2009) Activation of dormant bacterial genes by Nonomuraea sp. strain ATCC 39727 mutant-type RNA polymerase. *J Bacteriol* **191**: 805-814.

Telenti, A., P. Imboden, F. Marchesi, D. Lowrie, S. Cole, M. J. Colston, L. Matter, K. Schopfer & T. Bodmer, (1993) Detection of rifampicin-resistance mutations in *Mycobacterium tuberculosis*. *Lancet* **341**: 647-650.

Toft, C. & S. G. Andersson, (2010) Evolutionary microbial genomics: insights into bacterial host adaptation. *Nat Rev Genet* **11**: 465-475.

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

Traxler, M. F., D. E. Chang & T. Conway, (2006) Guanosine 3',5'-bispyrophosphate coordinates global gene expression during glucose-lactose diauxie in *Escherichia coli*. *Proc Natl Acad Sci U S A* **103**: 2374-2379.

Traxler, M. F., S. M. Summers, H. T. Nguyen, V. M. Zacharia, G. A. Hightower, J. T. Smith & T. Conway, (2008) The global, ppGpp-mediated stringent response to amino acid starvation in *Escherichia coli*. *Mol Microbiol* **68**: 1128-1148.

Trindade, S., A. Sousa, K. B. Xavier, F. Dionisio, M. G. Ferreira & I. Gordo, (2009) Positive epistasis drives the acquisition of multidrug resistance. *PLoS Genet* **5**: e1000578.

Trautinger, B. W. & R. G. Lloyd, (2002) Modulation of DNA repair by mutations flanking the DNA channel through RNA polymerase. *EMBO J* **21**: 6944-6953.

Tupin, A., M. Gualtieri, F. Roquet-Baneres, Z. Morichaud, K. Brodolin & J. P. Leonetti, (2010) Resistance to rifampicin: at the crossroads between ecological, genomic and medical concerns. *Int J Antimicrob Agents* **35**: 519-523.

Uhia, I., B. Galan, V. Morales & J. L. Garcia, (2011) Initial step in the catabolism of cholesterol by *Mycobacterium smegmatis* mc(2) 155. *Environ Microbiol*.

Uys, P. W., R. Warren, P. D. van Helden, M. Murray & T. C. Victor, (2009) Potential of rapid diagnosis for controlling drug-susceptible and drug-resistant tuberculosis in communities where *Mycobacterium tuberculosis* infections are highly prevalent. *J Clin Microbiol* **47**: 1484-1490.

van der Spuy, G. D., K. Kremer, S. L. Ndabambi, N. Beyers, R. Dunbar, B. J. Marais, P. D. van Helden & R. M. Warren, (2009) Changing *Mycobacterium tuberculosis* population highlights clade-specific pathogenic characteristics. *Tuberculosis (Edinb)* **89**: 120-125.

Van Deun, A., L. Barrera, I. Bastian, L. Fattorini, H. Hoffmann, K. M. Kam, L. Rigouts, S. Rusch-Gerdes & A. Wright, (2009) *Mycobacterium tuberculosis* strains with highly discordant rifampin susceptibility test results. *J Clin Microbiol* **47**: 3501-3506.

van Kessel, J. C. & G. F. Hatfull, (2007) Recombineering in *Mycobacterium tuberculosis*. *Nat Methods* **4**: 147-152.

van Kessel, J. C. & G. F. Hatfull, (2008) Efficient point mutagenesis in mycobacteria using single-stranded DNA recombineering: characterization of antimycobacterial drug targets. *Mol Microbiol* **67**: 1094-1107.

Vandal, O. H., C. F. Nathan & S. Ehrt, (2009) Acid resistance in *Mycobacterium tuberculosis*. *J Bacteriol* **191**: 4714-4721.

Vandal, O. H., L. M. Pierini, D. Schnappinger, C. F. Nathan & S. Ehrt, (2008) A membrane protein preserves intrabacterial pH in intraphagosomal *Mycobacterium tuberculosis*. *Nat Med* **14**: 849-854.

Vassylyev, D. G., S. Sekine, O. Laptenko, J. Lee, M. N. Vassylyeva, S. Borukhov & S. Yokoyama, (2002) Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 A resolution. *Nature* **417**: 712-719.

Vassylyev, D. G., M. N. Vassylyeva, A. Perederina, T. H. Tahirov & I. Artsimovitch, (2007) Structural basis for transcription elongation by bacterial RNA polymerase. *Nature* **448**: 157-162.

Vigliotta, G., S. M. Tredici, F. Damiano, M. R. Montinaro, R. Pulimeno, R. di Summa, D. R. Massardo, G. V. Gnoni & P. Alifano, (2005) Natural merodiploidy involving duplicated *rpoB* alleles affects secondary metabolism in a producer actinomycete. *Mol Microbiol* **55**: 396-412.

Viveiros, M., C. Leandro & L. Amaral, (2003) Mycobacterial efflux pumps and chemotherapeutic implications. *Int J Antimicrob Agents* **22**: 274-278.

Warner, D. F., (2010) The role of DNA repair in *M. tuberculosis* pathogenesis. *Drug Discov Today* **7**: e5-e11.

Warner, D. F., G. Etienne, X. M. Wang, L. G. Matsoso, S. S. Dawes, K. Soetaert, N. G. Stoker, J. Content & V. Mizrahi, (2006) A derivative of *Mycobacterium smegmatis* mc(2)155 that lacks the duplicated chromosomal region. *Tuberculosis (Edinb)* **86**: 438-444.

Warner, D. F. & V. Mizrahi, (2006) Tuberculosis chemotherapy: the influence of bacillary stress and damage response pathways on drug efficacy. *Clin Microbiol Rev* **19**: 558-570.

Warner, D. F. & V. Mizrahi, (2007) The survival kit of *Mycobacterium tuberculosis*. *Nat Med* **13**: 282-284.

Warner, D. F., D. E. Ndwandwe, G. L. Abrahams, B. D. Kana, E. E. Machowski, C. Venclovas & V. Mizrahi, (2010) Essential roles for imuA'- and imuB-encoded accessory factors in DnaE2-dependent mutagenesis in *Mycobacterium tuberculosis*. *Proc Natl Acad Sci U S A* **107**: 13093-13098.

Wegrzyn, A., A. Szalewska-Palasz, A. Blaszczak, K. Liberek & G. Wegrzyn, (1998) Differential inhibition of transcription from sigma70- and sigma32-dependent promoters by rifampicin. *FEBS Lett* **440**: 172-174.

Weinreich, D. M., R. A. Watson & L. Chao, (2005) Perspective: Sign epistasis and genetic constraint on evolutionary trajectories. *Evolution* **59**: 1165-1174.

Weinstein, E. A., T. Yano, L. S. Li, D. Avarbock, A. Avarbock, D. Helm, A. A. McColm, K. Duncan, J. T. Lonsdale & H. Rubin, (2005) Inhibitors of type II NADH: menaquinone oxidoreductase represent a class of antitubercular drugs. *Proc Natl Acad Sci U S A* **102**: 4548-4553.

Werner, F. & D. Grohmann, (2011) Evolution of multisubunit RNA polymerases in all three domains of life. *Nat Rev Microbiol* **9**: 85-98.

Westblade, L. F., E. A. Campbell, C. Pukhrambam, J. C. Padovan, B. E. Nickels, V. Lamour & S. A. Darst, (2010) Structural basis for the bacterial transcription-repair coupling factor/RNA polymerase interaction. *Nucleic Acids Res* **38**: 8357-8369.

WHO, (2009) Global Tuberculosis Control Report 2009.

Wichelhaus, T. A., B. Boddinghaus, S. Besier, V. Schafer, V. Brade & A. Ludwig, (2002) Biological cost of rifampin resistance from the perspective of *Staphylococcus aureus*. *Antimicrob Agents Chemother* **46**: 3381-3385.

Wilke, C. O., R. E. Lenski & C. Adami, (2003) Compensatory mutations cause excess of antagonistic epistasis in RNA secondary structure folding. *BMC Evol Biol* **3**: 3.

Williams, D. L., L. Spring, L. Collins, L. P. Miller, L. B. Heifets, P. R. Gangadharam & T. P. Gillis, (1998) Contribution of *rpoB* mutations to development of rifamycin cross-resistance in *Mycobacterium tuberculosis*. *Antimicrob Agents Chemother* **42**: 1853-1857.

Wrande, M., J. R. Roth & D. Hughes, (2008) Accumulation of mutants in "aging" bacterial colonies is due to growth under selection, not stress-induced mutagenesis. *Proc Natl Acad Sci U S A* **105**: 11863-11868.

Wu, S., S. T. Howard, D. L. Lakey, A. Kipnis, B. Samten, H. Safi*, et al.*, (2004) The principal sigma factor *sigA* mediates enhanced growth of *Mycobacterium tuberculosis in vivo*. *Mol Microbiol* **51**: 1551-1562.

Xie, Z., N. Siddiqi & E. J. Rubin, (2005) Differential antibiotic susceptibilities of starved *Mycobacterium tuberculosis* isolates. *Antimicrob Agents Chemother* **49**: 4778-4780.

Xu, J., Y. Tozawa, C. Lai, H. Hayashi & K. Ochi, (2002) A rifampicin resistance mutation in the *rpoB* gene confers ppGpp-independent antibiotic production in *Streptomyces coelicolor* A3(2). *Mol Genet Genomics* **268**: 179-189.

Yakhnin, A. V., H. Yakhnin & P. Babitzke, (2008) Function of the Bacillus subtilis transcription elongation factor NusG in hairpin-dependent RNA polymerase pausing in the trp leader. *Proc Natl Acad Sci U S A* **105**: 16131-16136.

Yam, K. C., I. D'Angelo, R. Kalscheuer, H. Zhu, J. X. Wang, V. Snieckus, L. H. Ly, P. J. Converse, W. R. Jacobs, Jr., N. Strynadka & L. D. Eltis, (2009) Studies of a ring-cleaving dioxygenase illuminate the role of cholesterol metabolism in the pathogenesis of *Mycobacterium tuberculosis*. *PLoS Pathog* **5**: e1000344.

Yang, B., H. Koga, H. Ohno, K. Ogawa, M. Fukuda, Y. Hirakata, S. Maesaki, K. Tomono, T. Tashiro & S. Kohno, (1998) Relationship between antimycobacterial activities of rifampicin, rifabutin and KRM-1648 and *rpoB* mutations of *Mycobacterium tuberculosis*. *J Antimicrob Chemother* **42**: 621-628.

Yang, X., N. M. Nesbitt, E. Dubnau, I. Smith & N. S. Sampson, (2009) Cholesterol metabolism increases the metabolic pool of propionate in *Mycobacterium tuberculosis*. *Biochemistry* **48**: 3819-3821.

Yano, T., L. S. Li, E. Weinstein, J. S. Teh & H. Rubin, (2006) Steady-state kinetics and inhibitory action of antitubercular phenothiazines on *Mycobacterium tuberculosis* type-II NADH-menaquinone oxidoreductase (NDH-2). *J Biol Chem* **281**: 11456-11463.

Yim, G., F. de la Cruz, G. B. Spiegelman & J. Davies, (2006) Transcription modulation of *Salmonella enterica* serovar Typhimurium promoters by sub-MIC levels of rifampin. *J Bacteriol* **188**: 7988-7991.

Yu, J., J. Wu, K. P. Francis, T. F. Purchio & J. L. Kadurugamuwa, (2005) Monitoring *in vivo* fitness of rifampicin-resistant *Staphylococcus aureus* mutants in a mouse biofilm infection model. *J Antimicrob Chemother* **55**: 528-534.

Yuen, L. K., D. Leslie & P. J. Coloe, (1999) Bacteriological and molecular analysis of rifampin-resistant *Mycobacterium tuberculosis* strains isolated in Australia. *J Clin Microbiol* **37**: 3844-3850.

Zaczek, A., A. Brzostek, E. Augustynowicz-Kopec, Z. Zwolska & J. Dziadek, (2009) Genetic evaluation of relationship between mutations in *rpoB* and resistance of *Mycobacterium tuberculosis* to rifampin. *BMC Microbiol* **9**: 10.

Zeibell, K., S. Aguila, V. Yan Shi, A. Chan, H. Yang & J. H. Miller, (2007) Mutagenesis and repair in *Bacillus anthracis*: the effect of mutators. *J Bacteriol* **189**: 2331-2338.

Zhang, G., E. A. Campbell, L. Minakhin, C. Richter, K. Severinov & S. A. Darst, (1999) Crystal structure of *Thermus aquaticus* core RNA polymerase at 3.3 A resolution. *Cell* **98**: 811-824.

Zhang, Y., (2005) The magic bullets and tuberculosis drug targets. *Annu Rev Pharmacol Toxicol* **45**: 529-564.

Zhou, Y. N. & D. J. Jin, (1998) The *rpoB* mutants destabilizing initiation complexes at stringently controlled promoters behave like "stringent" RNA polymerases in *Escherichia coli*. *Proc Natl Acad Sci U S A* **95**: 2908-2913.

Zumla, A., P. Mwaba, J. Huggett, N. Kapata, D. Chanda & J. Grange, (2009) Reflections on the white plague. *Lancet Infect Dis* **9**: 197-202.